# Automatic Generation of Haptic Motion Effects Expressing Human Dance

Jaehyeok Ahn* POSTECH
Seungmoon Choi† POSTECH

2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR) | 979-8-3315-3645-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/VR59515.2025.00058

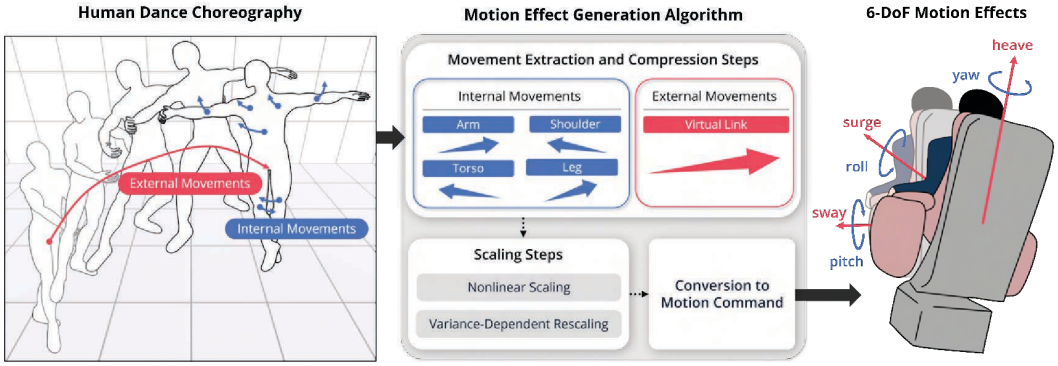



Figure 1: Generating haptic motion effects expressing human dance to enhance the multisensory experiences.

## Abstract

Haptic motion effects refer to physical stimuli that stimulate the human vestibular system for various purposes, such as flight simulation and entertainment. Motion effects are intensively utilized in 4D films, which provide viewers with multisensory special effects with audiovisual stimuli. This paper focuses on the automatic creation of motion effects that express the physical movements in human dance. Such effects can significantly enhance the user's multisensory experiences of watching and feeling a human dance. Taking the human motion capture data as input, our algorithms compute full six-degree-of-freedom motion effects featuring: (1) Separate processing of the external and internal movements of a dancer's body; (2) Decomposing the internal movements into a few segments to obtain optimal motion effects between clarity and expressiveness; and (3) Merging of all external and internal movements with appropriate scaling and weights. Our novel method can elicit better user experiences than two previous motion effect generation algorithms, which respectively represent the movements in general articulated bodies and all dynamic components in the scene.

**Index Terms:** Motion Effects, Haptics, Automatic Generation, Dance, User Experience

## 1 Introduction

In extended reality (XR), multisensory effects are crucial to eliciting users' unified and enhanced perception and bridging the gap between the virtual and physical realms. Motion effects are one type of multisensory effect that presents vestibular stimuli to the user, typically using a motion platform that stimulates the user's entire body. This motion-enabled setup is widely used in many XR applications, e.g., virtual reality (VR) games, 4D rides, and 4D films.

This paper addresses the feasibility of motion effects expressing human motion, especially *dance*. Dance is the artistic sequence of human body movements based on the designed choreography, communicating the dancer's intent and emotions [47]. Dance has no linguistic barriers, and this universality has accelerated technological advancements that afford more immersive and interactive dance-viewing experiences. For example, 3D VR180 cameras enable the reproduction of pop idol performances and dances in VR, where users can adjust their viewpoints with head movements. However, research on other multisensory effects is still immature.

We propose a computational algorithm that automatically generates a command to a six-degree-of-freedom (DoF) motion platform from the motion capture data of human dance (Section 3). Our work follows a general framework of Han et al. [9] for motion effects that express the movements of multiple articulated bodies, e.g., two fighting humans. As a specialization, we aim to enhance the user's multisensory experience of watching and feeling dance. Unlike [9], our algorithm has different treatments for external and internal body movements, and the internal movements of a few body segments contribute to the motion effect individually, all to maximize the expressiveness of motion effects while maintaining their clarity (Section 4). We evaluate the user experiences elicited by our algorithm through two user studies, one assessing different weighting policies for designing our algorithm (Section 5) and the other comparing ours with previous algorithms with similar purposes (Section 6). It is followed by conclusions, limitations, and a plan for future work (Section 8).

To our knowledge, this work presents the first effective algorithm dedicated to expressing human dance actions using a 6-DoF motion chair, which can be regarded as the main contribution of this paper.

## 2 Related Work

### 2.1 Representing Human Dance

Dance is a complex sequence of body movements. Extensive research has been conducted to enable articulated bodies, such as robots and virtual characters, to depict human dance movements.

A robot that dances is technically challenging, requiring advanced motion planning, balance maintenance, coordination, and control. The research has focused on applying a dancer's movements to humanoids, ensuring that robots can perform natural and similar dance motions while considering their physical limitations. Researchers used motion capture data to simulate specific parts of the human body [16, 31] or to control the robot's movements based on the overall motion of the dancer [19, 32, 50]. Other modalities

*e-mail: ajh131@postech.ac.kr
†e-mail: choism@postech.ac.kr

2642-5254/25/$31.00 ©2025 IEEE
DOI 10.1109/VR59515.2025.00058

related to dance were also harnessed, sound [2, 39], audiovisual information [33], and the emotions conveyed by music [37, 49].

A robot can also dance with a human partner. The information on the human dancer is commonly passed to the robot through interaction force [20, 21, 28, 43], along with other modalities such as motion capture [38] and laser range sensing [45]. Recently, Ladenheim et al. [22] developed a motion-activated wearable robot that combines a human and a robot to create a cyborg character on stage. Robots can also serve as tutors to teach dance to humans [34].

Unlike robots, virtual characters have fewer physical constraints, which is advantageous for a richer expression of dance. Kim et al. [17] proposed an algorithm that analyzes music in real time and synchronizes motion clips with the analyzed data. Mousas [30] used a Hidden Markov Model to control the dance movements of a virtual partner based on the user's performance. Gonen et al. [7] presented a motion generation method for both robots and virtual characters from the data collected from real people.

Translating dance movements into multisensory effects using devices that do not bear kinematic similarity to human limbs is a particularly challenging task. To our knowledge, a recent work by Jung et al. [14] is unique in this research space. This method uses full-body vibrotactile cues to represent dance movements in a VR performance, using the concept of a motion-salient triangle. Our work attempts to achieve similar goals using a 6 DoF-motion platform.

### 2.2 Automatic Generation of Motion Effects

Originally developed for training pilots and astronauts, haptic motion simulation has since expanded into various entertainments, including theme park rides, 4D cinemas, and virtual reality games. Lee et al. [25] categorized such motion effects into four groups: camera-, object-, sound, and context-based effects. Camera-based motion effects, frequently employed in 4D attractions for point-of-view (POV) scenes, mimic the camera's movements and enhance the sense of immersion. Object-based motion effects operate the motion chair to emphasize the movements of objects of interest in the scene. Sound-based motion effects produce short and intense movements to simulate sound-triggered events, such as explosions and gunshots. Lastly, context-based motion effects use the scene's narrative or thematic contexts, such as stepping on the accelerator, to augment indirect cues for their application.

After a conceptual framework by Shin et al. [40], Lee et al. [25] presented a near real-time authoring algorithm of camera-based motion effects for point-of-view (POV) shots in videos. It estimates the 6-DoF camera extrinsic information from optical flow and feeds it to a washout filter to make a 3-DoF motion command. This method was extended by Lim et al. [27], which added high-frequency motion representing the ground texture, such as roller coaster rails and cobblestones. Luan et al. [29] demonstrated a method of capturing motion footage using a motion camera and converting its trajectory to a 6-DoF motion effect.

As for object-based motion effects, Lee et al. [24] represented the movements of an object projected on the screen as motion effects while tracking the object's motion using computer vision techniques. Han et al. [10] introduced the concept of a *motion proxy* in the camera space to make a 3-DoF motion effect compressed from the 6-DoF translation and rotation of the object. This approach was further developed to express the movements of multiple articulated bodies as motion effects [9], which merges all link movements collected from motion capture data. Recently, Han et al. [8] proposed a motion generation algorithm that considers the 3D movements of all pixels in the 2D scene using scene flow estimation and saliency detection models powered by deep learning. This method is comprehensive as it responds to object motions and all other motions in the scene, including the camera's. Our work presented in this paper is concerned with object-based motion effects and specializes in expressing human dance motion.

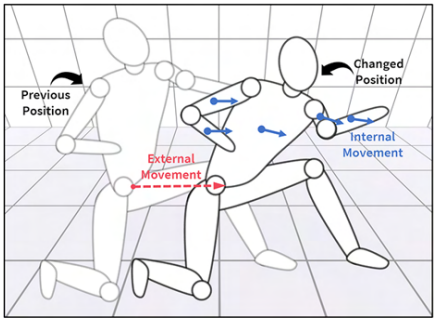



Figure 2: Concepts of external and internal movements.

Other notable results include algorithms that make motion effects by processing sound [25, 55] and a method that combines camera- and object-based motion effects into one motion effect while preserving their perceptual effects [35].

## 3 Research Problem

For our purpose, the human body can be modeled as an articulated structure composed of many links and joints. We assume that the motion data of the human body expressing a dance are provided, e.g., by motion capture or from an existing human motion database. Our goal is to process the motion data and make the 6-DoF command to a motion platform to convey the core information in the dance motion. This research problem is formulated as follows.

We represent a human body using a scene graph. The homogeneous transformation matrix from the world coordinate frame (index 0) to the base node (index 1) of the graph is denoted by ${}^0\boldsymbol{T}_1$. The transformation matrix for a node $i$ in the graph is represented by

$$ {}^0\boldsymbol{T}_i = {}^0\boldsymbol{T}_1 \, {}^1\boldsymbol{T}_i, \tag{1} $$

where ${}^1\boldsymbol{T}_i$ forms a continuous transformation path from the root node to node $i$. Then, a dance data is expressed by

$$ \boldsymbol{D} = \bigcup_{i=0}^{K} {}^{i-1}\boldsymbol{T}_i[n], \tag{2} $$

where $K$ is the number of nodes (links) in the graph, and $n$ is the discrete time index with the sampling period of $T_s$ and the dance duration of $T$ ($0 \le n \le T/T_s$).

From this input data of $\boldsymbol{D}$, we aim to compute the motion command $\boldsymbol{m}[n]$ to the motion chair to use, where

$$ \boldsymbol{m} = (\boldsymbol{m}^{rtn}, \boldsymbol{m}^{tln}), \tag{3} $$

$$ \boldsymbol{m}^{rtn} = (roll, pitch, yaw), \tag{4} $$

$$ \boldsymbol{m}^{tln} = (surge, sway, heave). \tag{5} $$

$\boldsymbol{m}^{rtn}$ and $\boldsymbol{m}^{tln}$ are the rotation and translation components. This command $\boldsymbol{m}$ must convey the essential information in the complex human dance movements to the user while satisfying the limited performance specifications of the motion chair. This problem corresponds to information compression; the motion chair has only six free motion directions, while the human dance movements have a considerably greater DoF.

## 4 Motion Effect Generation Algorithm

This section describes our algorithm that generates motion effects that adequately express a human dance. The algorithm consists of seven steps, each explained in one subsection below.

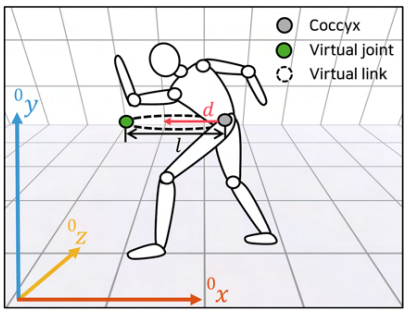



Figure 3: Virtual link representing the external movement.

### 4.1 Step 1: Movement Extraction

We categorize dance movements into external and internal movements. External movements refer to the translation and rotation of the dancing character's entire body in the world space and are represented by ${}^0\boldsymbol{T}_1$. Internal movements mean the movements of all body links except the base and are contained in $\boldsymbol{D}^- = \boldsymbol{D} - \{{}^0\boldsymbol{T}_1\}$. Dancers deliver intentions and emotions through external and internal movements of the body following the choreography.

We consider the external and internal movements separately because external movements can mask the effects of internal movements on motion effects. Figure 2 shows an example where the internal movements involving leaning the upper body forward and curling the left arm inward look insignificant compared to the body's large external movement. Representing internal movements properly as motion effects requires independent processing from external movements. All external and internal movements must be incorporated to express complex body motions for dance.

Our algorithm receives dance motion data as input. The first step is extracting the motion data's external and internal movements. We describe the external movements by the homogeneous transformation ${}^0\boldsymbol{T}_1$ from the world frame to the base link's frame of the human body. The base link must represent the transformation of the entire body, and we choose the coccyx (tailbone joint) as the base link. Its transformation information is included in most motion capture data. The head and limbs mainly cause the internal movements, and their description should remain independent of the external movements ${}^0\boldsymbol{T}_1$. Using coccyx as the base link for external proxy computation allows the least interference by internal movements. Thus, we extract the rest of the transformation matrices, $\boldsymbol{D}^- = \{{}^1\boldsymbol{T}_2, \cdots, {}^{N-1}\boldsymbol{T}_N\}$, to represent the internal movements.

### 4.2 Step 2: Movement Compression

The next step is to compress the large amount of information in the extracted movements into a few variables adequate for generating motion effects. To this end, we extend the concept of *motion proxy*, a 3D point that represents the collective information about the motion of an object of interest. The motion proxy's effectiveness has been demonstrated in a few previous studies [10, 9]. We define two types of motion proxies: external and internal proxies for external and internal movements.

The external motion proxy is defined by adding a virtual link to the body's base link (coccyx), as shown in Figure 3. The virtual joint is positioned at a distance $l$ from the coccyx's local coordinate frame along the outward normal of the coccyx's plane. Its end point is denoted by ${}^0\boldsymbol{p}^{vl}$ in the world coordinate frame. Then, the external motion proxy $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)$ is set as

$$\boldsymbol{\mu} = {}^0\boldsymbol{p}^{vl}. \tag{6}$$

This motion proxy merges the translation and rotation of the virtual link into the translation of one point in the world space, as in [10]. The parameter $l$ adjusts the extent of the body rotation's contribution to the motion proxy $\boldsymbol{\mu}$ relative to the body translation.

For internal movements, we consider that the fixation time of observers watching a dancer depends on the body part [36, 48] and decompose the body into four segments:

$$S = \{Torso, Arms, Legs, Shoulders\}. \tag{7}$$

This arrangement is determined as the best trade-off between using a lower number (high clarity but low expressiveness) and a higher number (low clarity but high expressiveness) after numerous tests of various dance data. Finer segmentation, e.g., separating the left and right arms or including the hands, feet, or head, degrades the clarity, whereas coarse segmentation compromises the expressiveness.

For body segment $s \in S$, we define its internal motion proxy as

$$\boldsymbol{v}_s = \frac{1}{\sum_{j=1}^{N_s} a_{s,j}} \sum_{i=1}^{K_s} a_{s,i}\, {}^1\boldsymbol{r}_{s,i}. \tag{8}$$

Here, $K_s$ is the number of links in segment $s$. ${}^1\boldsymbol{r}_{s,i}$ is the rotation angle vector (roll, pitch, yaw) of link $i$ in $s$ in reference to the base body link 1 (representing an internal motion). $a_{s,i}$ is the normalized visual area of link $i$, calculated as the area on the display screen to which its mesh is projected.

This definition is based on two observations. First, all human joints are rotational, so only rotation variables are included in the formulation. Second, adjacent joints are connected by a link, and we use the link's volume projected to the viewer's eyes, $a_{s,i}$, as the weight. This strategy was demonstrated to be effective in [9].

The four internal motion proxies $\boldsymbol{v}_s (s \in S)$ are combined into a single internal motion proxy in later steps.

### 4.3 Step 3: Differentiation and Filtering

The external motion proxy $\boldsymbol{\mu}$ and the internal motion proxy $\boldsymbol{v}_s$ of each body segment $s$ are numerically differentiated to the velocity terms. This step is necessary to ensure that motion effects reflect only relative motions. The final step, which generates the commands for the motion chair, also requires the velocity input. We filter out the noise amplified by the numerical differentiation using a Wavelet denoising technique, which decomposes the signal into multiple frequency bands and removes noise from each band [6, 5]. This filtered velocity, denoted by $\dot{\boldsymbol{\mu}}$, is passed to the next step without further processing.

The filtered velocity of $\boldsymbol{v}_s$ goes through another high-pass filter (5th-order Butterworth, cutoff frequency 0.45 Hz). This double-filtered velocity is denoted by $\dot{\boldsymbol{v}}_s$. It is because the low-frequency components in the input variables tend to push the motion platform to its workspace boundary, and it is customary to apply a high-pass filter to prevent this problem [25]. External movements are substantially slower than internal movements. Thus, the low-frequency energy of $\dot{\boldsymbol{\mu}}$ is quite small. It does not cause the motion chair saturation problem, and we use $\dot{\boldsymbol{\mu}}$ without high-pass filtering.

### 4.4 Step 4: Normalization and Nonlinear Scaling

In usual human motion, slow movements do not convey meanings or belong to the preparatory phase for larger movements. In contrast, slow movements in dances can serve as a means of expressing emotions and intentions. Therefore, a method that controls the extent to which slow (or fast) movements are reflected is useful for creating effective motion effects for a specific choreography.

To this end, we define a nonlinear scaling rule as follows: for a finite sequence of 3D vector $\boldsymbol{h}[n] = (h_1[n], h_2[n], h_3[n])$,

$$\hat{h}_k[n] = \left( \frac{h_k[n]}{\max_i \| \boldsymbol{h}[i] \|} \right)^{\alpha}, \quad k \in \{1,2,3\} \tag{9}$$

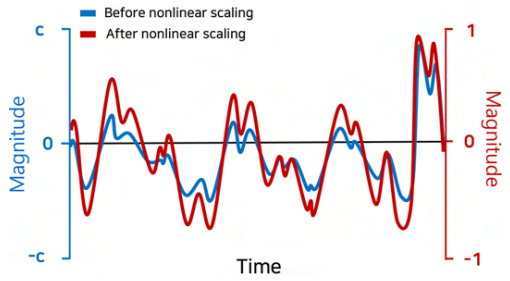



Figure 4: Comparison of signals before (blue) and after the nonlinear scaling (red). $\alpha = 0.8$. $c$ is an arbitrary constant.

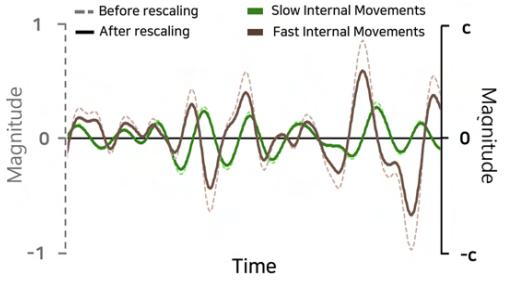



Figure 5: Comparison between before (dotted line) and after variance-dependent rescaling (solid line). $c$ is an arbitrary constant.

where $\alpha > 0$ is an exponent. This mapping normalizes each component in $\boldsymbol{h}$ using the maximum magnitude of $\boldsymbol{h}$ and scales it using a power function. As $\alpha$ decreases from 1 to 0, small values in $\boldsymbol{h}$ are more emphasized in $\hat{\boldsymbol{h}}$ (Figure 4). Using this rule, $\dot{\boldsymbol{\mu}}$ is scaled to $\hat{\dot{\boldsymbol{\mu}}}$, and each $\dot{\boldsymbol{v}}_s$ is converted to $\hat{\dot{\boldsymbol{v}}}_s$. This step helps the generated motion effects clearly express even detailed and slow movements.

### 4.5 Step 5: Merging Internal Motion Proxies

The previous step passes the normalized and scaled velocity variables of the internal motion proxies, $\hat{\dot{\boldsymbol{v}}}_s$, for the four body segments (*Torso*, *Arms*, *Legs*, and *Shoulders*). This step combines the four variables into one internal motion proxy.

The passed internal motion proxies are represented in the body's base coordinate frame as in (8). Merging them makes a single motion proxy that preserves all internal movement information independent of the body's orientation to the viewer. For example, suppose a dancer moves arms while spinning in one position. In our algorithm, the spinning is represented by the external motion proxy using the world frame. In contrast, the arm motion is expressed by the internal motion proxy using the body frame. Thus, the dance choreography is expressed as a whole. This arrangement is generally effective according to our tests using various dance videos.

To design an effective merging scheme, we consider a fact in human vestibular perception that the human sensitivity is lower for yaw motion than for roll and pitch motion [26]. In our rules, the yaw of the merged motion proxy is determined solely by the motion proxy for shoulders. Shoulder motion encompasses comprehensive upper body movements, and humans tend to focus more on the upper body than the lower body when watching dances [36]. We avoid making the yaw of the merged motion proxy more complex by injecting additional information. Then, the roll and pitch of the merged motion proxy are computed from the rest of the internal proxies with weights. In summary, the merged internal motion proxy, $\hat{\dot{\boldsymbol{v}}} = (\hat{\dot{v}}_{roll}, \hat{\dot{v}}_{pitch}, \hat{\dot{v}}_{yaw})$, is computed by:

$$\hat{\dot{v}}_{roll} = \sum_{s \in S^-} w_s \hat{\dot{v}}_{s,roll}, \tag{10}$$

$$\hat{\dot{v}}_{pitch} = \sum_{s \in S^-} w_s \hat{\dot{v}}_{s,pitch}, \tag{11}$$

$$\hat{\dot{v}}_{yaw} = \hat{\dot{v}}_{Shoulders,yaw}, \tag{12}$$

where $S^- = S - \{Shoulders\} = \{Torso, Arms, Legs\}$, and $w_s$ is the merging weight of segment $s$. $w_s$ has crucial effects on the resulting motion commands, and we delve into this issue in Section 5.

### 4.6 Step 6: Variance-Dependent Rescaling

The external and internal motion proxies, $\hat{\dot{\boldsymbol{\mu}}}$ and $\hat{\dot{\boldsymbol{v}}}$, passed from the previous step are normalized. This step restores their original scales so that users can experience the motion effects adequately.

For the external motion proxy, we use the following simple rule:

$$\tilde{\dot{\boldsymbol{\mu}}} = \max\left(\|\dot{\boldsymbol{\mu}}\|\right) \hat{\dot{\boldsymbol{\mu}}}. \tag{13}$$

For the internal motion proxy $\hat{\dot{\boldsymbol{v}}}$, we use a deliberate scheme because the scales of internal movements vary significantly between choreographies. Some dance genres, such as ballet, feature many slow internal movements, while others, such as hip-hop, exhibit much faster movements. The range of internal movement speeds must be considered for the resulting motion effects to express the dance characteristics appropriately. Contrarily, the external movements demonstrate a relatively consistent range of movement speed across dance genres, and the simple restoration in (13) is sufficient.

Motion effect designers can rescale the internal motion proxies to tune the motion effects to their best intention. However, it may require many iterations, and we provide a heuristic rule that works well for most dance genres according to our tests:

$$\tilde{\dot{v}}_{roll} = \frac{\max_{s \in S^-} \max_n \|\dot{\boldsymbol{v}}_s[n]\|}{\log_{10} \sigma_{roll}} \hat{\dot{v}}_{roll} \tag{14}$$

$$\tilde{\dot{v}}_{pitch} = \frac{\max_{s \in S^-} \max_n \|\dot{\boldsymbol{v}}_s[n]\|}{\log_{10} \sigma_{pitch}} \hat{\dot{v}}_{pitch} \tag{15}$$

$$\tilde{\dot{v}}_{yaw} = \frac{\max_n \|\dot{\boldsymbol{v}}_{Shoulders}[n]\|}{\log_{10} \sigma_{yaw}} \hat{\dot{v}}_{yaw}, \tag{16}$$

where $\sigma$ represents the standard deviation of $\|\dot{\boldsymbol{v}}_{Shoulders}\|$ for yaw, and the average of the standard deviations of $\|\dot{\boldsymbol{v}}_s\|$ over *Torso*, *Arms*, and *Legs* for roll and pitch. Here, the magnitude rescaling term using the maximum is penalized using the standard deviation. It has the effect that the motion proxy velocity that contains large fluctuations is less amplified while one that includes small variations is more emphasized. This effect is smoothed by adding the logarithm function for the standard deviation, as shown in Figure 5.

### 4.7 Step 7: Motion Predictive Control

Finally, we generate a command to the motion platform using a standard Model Predictive Control (MPC) method. MPC is a general framework in control engineering that makes the system output converge to the reference over time by solving an optimization problem with a given process model. When applied to motion command generation, the performance limitations of the motion platform can be formulated into the constraints of the optimization problem. This advantage makes MPC a preferred choice for motion generation over other methods, such as washout filters [3].

Specifically, we use the following formulation similar to Han et

al. [10]: For a timed function of 3D vector, $\boldsymbol{q}(t)$,

$$\begin{aligned} \boldsymbol{m} = \arg\min_{\boldsymbol{m}} \quad & \|\beta \boldsymbol{q} - \dot{\boldsymbol{m}}\|^2 + \varepsilon \|\boldsymbol{m}\|^2 \\ \text{subject to} \quad & \dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\dot{\boldsymbol{m}},\ \dot{\boldsymbol{m}}(t) = C\boldsymbol{x}(t) + D\dot{\boldsymbol{m}}, \\ & \dot{\boldsymbol{m}}(t) = \frac{d}{dt}\boldsymbol{m}(t),\ |\boldsymbol{m}(t)| \le \boldsymbol{m}_{max}, \end{aligned} \tag{17}$$

where $\boldsymbol{m}$ is a 3-DoF motion command, and $\boldsymbol{m}_{max}$ contains the maximum motion ranges that the motion platform allows on the three axes. $\beta$ is a manually determined scale factor from the virtual world to the physical world, considering the workspace of the motion platform and user experience. $\varepsilon$ determines how fast $\boldsymbol{m}$ converges to zero (the neutral position), an important requirement for motion commands. $\boldsymbol{x}$ is a ($10 \times 3$) state vector that contains the past and present information of the system and reflects the future output. $A$, $B$, $C$, $D$ are the state matrices of dimensions ($10 \times 10$), ($10 \times 3$), ($3 \times 10$), and ($3 \times 3$), respectively. In our MPC implementation, we configure $A$, $B$ and $C$ as a delayed differentiator that the system accounts for the past state to produce a gradual response to input changes. $D$ is set to zero. This configuration ensures that the output is driven by the dynamic evolution of $\boldsymbol{x}$ and not directly by $\boldsymbol{m}$, leading to a more gradual and controlled response [4].

The above MPC problem is solved twice, once with $\boldsymbol{q} = \tilde{\boldsymbol{\mu}}$ and the other with $\boldsymbol{q} = \tilde{\boldsymbol{v}}$, with possibly different values of $\beta$ and $\varepsilon$. The former generates a 3-DoF motion command expressing the external movements, and we use that command, denoted by $\boldsymbol{m}^{tln}$, to determine the translation of the motion platform. Similarly, the latter produces a 3-DoF motion command, $\boldsymbol{m}^{rtn}$, for the internal movements mapped to the motion platform's rotation. This mapping strategy, which fully utilizes the six DoFs, aligns the dancer's position in the virtual space with the chair position on the platform and represents the dancer's link rotations through the chair rotation.

Finally, we combine the two 3-DoF motion commands to make a full 6-DoF motion command:

$$\boldsymbol{m} = \left(\boldsymbol{m}^{tln}, \boldsymbol{m}^{rtn}\right). \tag{18}$$

This command is used to control the movement of the 6-DoF platform in accordance with the dance video.

### 4.8 Implementation

Our implementation is tuned to a commercial 6-DoF motion platform (MotionSystems, PS-6TM-150 [42]), which has a wide motion rage: surge (-10 cm, 12 cm), sway (-10 cm, 10 cm), heave (-11 cm, 12 cm), roll (-26.0°, 26.0°), pitch (-25.0°, 25.6°), and yaw (-22.5°, 22.5°). Excessive platform movements can cause motion sickness, so we limit the maximum motion range to $\pm 12°$ for all rotation axes and $\pm 6$ cm for translation axes. This platform is sufficiently powerful to provide consistent motion stimuli regardless of the weight differences between adults.

For this motion chair, we set $l = 0.5$ to define the virtual link (Figure 3). We use $\alpha = 0.8$ for the motion proxy scaling in (9). For MPC, $\beta^{tln} = 3.85$ and $\beta^{rtn} = 0.13$, while $\varepsilon^{tln} = \varepsilon^{rtn} = 0.1$. Unity is used to provide graphics to users.

## 5 User Study on Weighting Policies

The internal movements of different body segments are merged into one motion proxy using weights (Section 4.5). How to determine the weights is expected to have critical effects on the motion effects' perceptual quality. This section introduces a few policies designed for merging weights. Since no objective criteria exist as to the transformation quality from dance to motion effects, we evaluate the weighting policies by conducting a user study for their perceptual, cognitive, and subjective quality. All user studies reported in this paper were approved by the Institutional Review Board at the authors' institution (PIRB-2024-E005).

### 5.1 Weighting Policies

Three weighting policies are considered in this user study.

#### 5.1.1 Uniform (UNI)

The uniform weighting method combines the three internal motion proxies in equal proportions: for $s \in S^- = \{Torso, Arms, Legs\}$.

$$w_s[n] = \frac{1}{3}. \tag{19}$$

#### 5.1.2 Salient (SAL)

Rapidly moving objects have high visual saliency, and this policy emphasizes the internal movements of faster segments:

$$w_s[n] = \frac{\|\hat{\boldsymbol{v}}_s[n]\|}{\sum_{s \in S^-} \|\hat{\boldsymbol{v}}_s[n]\|}. \tag{20}$$

#### 5.1.3 Selective (SEL)

This method selects only one motion proxy that is the most salient at the moment. However, excessive switching between different motion proxies is harmful, and we use a selection method for merging two motion proxies in four steps, as proposed by Park et al. [35]: (1) Compute the weights of SAL at index $n$ for each motion proxy; (2) Set the weight of the motion proxy with the highest SAL value to 1 and the weights of the others to 0; (3) Remove short intervals where the weight changes frequently; and (4) Apply sinusoidal smoothing to the transition intervals where the weight changes occur. Further details can be found in [35].

We apply this method twice to merge the three internal motion proxies. First, $\hat{\boldsymbol{v}}_{Torso}$ and $\hat{\boldsymbol{v}}_{Arms}$ are merged to form an internal proxy representing the upper body movements. Second, this merged proxy is further combined with $\hat{\boldsymbol{v}}_{Legs}$ to make a motion proxy expressing the whole internal movements. This policy concentrates the motion effect more on the upper body, as the audience attends more to the dancer's upper body [36].

### 5.2 Methods

#### 5.2.1 Participants

We recruited 20 participants (14 males and 6 females, 21–38 years old with an average age of 25.1). None of them reported known sensorimotor disorders. Before the experiment, participants were provided with a written document detailing the objective and procedure of the study. Then, they signed a consent form to confirm their agreement to participate. After the experiment, participants received approximately USD 15 as a token of appreciation.

#### 5.2.2 Devices

We used the 6-DoF motion platform in Section 4.8. Participants sat in the motion chair and watched a 75-inch screen (Innos Corp., S7501KU); see Figure 6. They wore noise-canceling headphones (Bose, QC 45) that played white noise to block auditory cues.

#### 5.2.3 Experimental Conditions

We selected four animations of a single dancer from AwesomeDog's motion capture dataset [1]. The animations were played in a virtual environment using Unity. This animation playback was recorded in a video, and the video was played to participants so that they could see all the dancers' movements. Table 1 explains the characteristics of each dance choreography. For each of the four videos, we generated three sets of motion effects from the motion capture data using the three weighting methods, UNI, SAL, and SEL, as described in Section 5.1. Therefore, the user study had two independent variables (*Motion Effect* and *Video*). The videos and motion effects are available in the supplemental video.

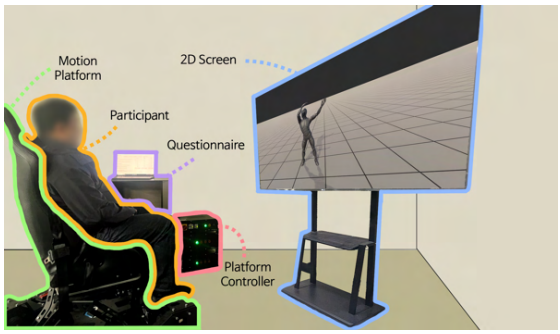



Figure 6: User study configuration.

Table 1: Characteristics of the dance data used in the user study.

| Dance tag | Description | Duration | Sources |
|---|---|---|---|
| *Ballet* | **Ballet:** Basic slow limbs' movements and turns. | 47s | 3'49"-4'20" and 4'56"-5'10 in [53] |
| *Hiphop* | **Hip-Hop Dance:** Dynamic limbs' movements and turns. | 31s | 5'42"-6'11" in [52] |
| *Nitty* | **1960s Dance Style:** Wriggling dances to the music's rhythm. | 29s | 8'34"-9'03" in [52] |
| *Charleston* | **Charleston Dance:** Dance with fast kicks and arm swings. | 33s | 2'06"-2'38" in [54] |

### 5.2.4 Procedure

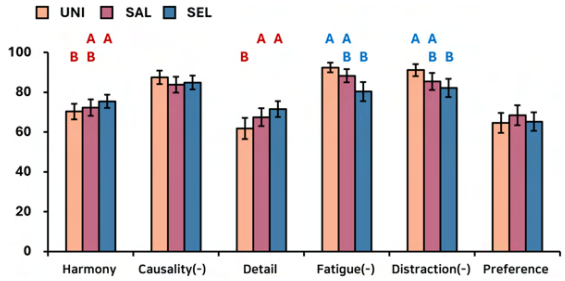
Participants had a practice session to familiarize themselves with motion effects. During this session, they experienced all three motion effects for the video they would evaluate in the first block of trials of the main session. The main session comprised four blocks of trials, one for each video. Each block had three trials in which the three motion effects were presented with the video. In each trial, participants experienced the pair of a video and a motion effect twice and answered a questionnaire that included the eight questions in Table 2. The clarity of motion effect expression was evaluated by Q1 (Harmony) and Q2 (Causality), and the expressiveness was by Q3 (Detail). The order of the video was randomized across participants, and the order of the motion effects per video was balanced across participants using the balanced Latin square. Participants took a 2-minute break after the practice session and a 3-minute break after each block of trials in the main session. They could also rest whenever needed.

Table 2: Questions used in the user study.

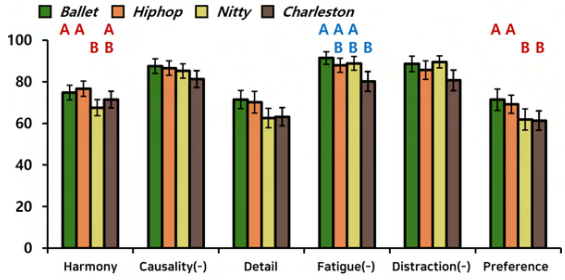
| No. | Measure | Question |
|---|---|---|
| Q1 | Harmony | The motion effect matched the dancer's movements. |
| Q2 | Causality (-) | I experienced motion effects that I did not understand why they had been provided. |
| Q3 | Detail | The motion effect described the dancer's detailed movements. |
| Q4 | Fatigue (-) | I felt tired after experiencing the motion effect. |
| Q5 | Distraction (-) | I was distracted from watching the video by the motion effect. |
| Q6 | Preference | I liked the motion effect. |
| Q7 | Motion Sickness | I felt motion sickness or dizziness. |
| Q8 | Free comments | Please leave a comment regarding the motion effect. |

* A '-' sign indicates a question of negative meaning.

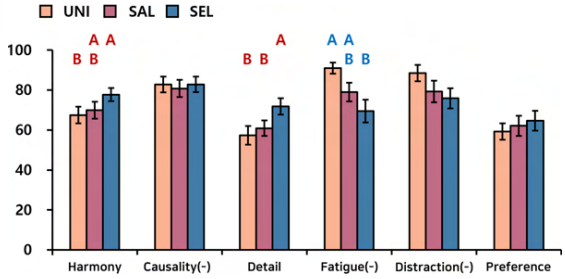
For each question from Q1 to Q6, participants expressed their extent of agreement on a scale from 0 to 100 by selecting a position on a horizontal line. The binary question Q7 was responded to by using a radio button. For Q8, participants entered their comments using a keyboard. In particular, we explicitly instructed participants to leave a comment when they experienced motion sickness. We also had a free interview with participants after the experiment.




(a) Average scores for motion effects




(b) Averages scores for videos




(c) Individual scores for video *Charleston*

Figure 7: Results of User Study 1. The scales of negative questions (represented by a '-' sign in the labels) are inverted so that a higher score indicates a better performance. Error bars represent standard errors. Motion effect sets paired with different letters showed significant differences by the SNK (red) or DSCF (blue) tests.

## 5.3 Results

To check the normality of the collected data, we applied the Yeo-Johnson power transform [46] to the data and then conducted the Shapiro-Wilk test. The responses to the positive questions (Q1, Q3, and Q6) mostly followed normal distributions, but this was not the case for the negative questions (Q2, Q4, and Q5). The responses to the negative questions were densely clustered in the low score ranges. Participants appeared to have a particular response bias of hesitating to give high scores to the negative questions.

The scores for the positive questions did not satisfy the sphericity assumption of ANOVA (Mauchly's test). Therefore, we performed two-way repeated-measures ANOVA with Greenhouse-Geisser corrections using *Motion Effect* and *Video* as the independent variables. For the negative questions, we performed one-way Kruskal-Wallis tests using *Motion Effect* or *Video* as the independent variable, respectively. We then applied the Student-Newman-Keuls (SNK) and Dwass-Steel-Critchlow-Fligner (DSCF) tests on the significant

cases of positive and negative questions, respectively, for post-hoc multiple comparisons. In what follows, we denote the results of ANOVA by AN and the Kruskal-Wallis by KW.

Figure 7 plots the experimental results. *Motion Effect* (Figure 7a) was significant in four questions: Q1 (Harmony; $F(1.98, 37.65) = 3.47$, $p = .0417$) and Q3 (Detail; $F(1.50, 28.57) = 7.59$, $p = .0045$) by AN, and Q4 (Fatigue; $\chi^2(2) = 17.0708$, $p = .0002$) and Q5 (Distraction; $\chi^2(2) = 10.1208$, $p = .0063$) by KW. *Video* (Figure 7b) was significant for Q1 (Harmony; $F(2.60, 49.43) = 5.12$, $p = .0033$), Q3 (Detail; $F(2.85, 54.20) = 3.77$, $p = .0172$), and Q6 (Preference; $F(2.52, 47.87) = 6.16$, $p = .0022$) by AN and Q4 (Fatigue; $\chi^2(3) = 11.0224$, $p = .0116$) by KW. For Q1, Q3, and Q6, we also conducted one-way ANOVA as simple-effect tests for each video, followed by the SNK tests for the significant cases. Significant results were observed only for *Charleston* (Figure 7c). The complete statistics results are provided in the supplemental material.

### 5.4 Discussion

#### 5.4.1 Effects of Motion Effects

On average (Figure 7), SAL and SEL scored significantly higher in Q1 (Harmony) and Q3 (Detail) than UNI. SAL and SEL can express the dancer's detailed internal movements by emphasizing rapid segment movements. The participants found such motion effects to be well aligned with the dancer's choreographies. For example, 12 participants (Ps) who experienced SEL for *Charleston* (featuring many fast kicks and arm swings) commented, "The chair well represented the dynamic movement of the limbs." This merit was supported by the significant differences observed only in *Charleston* (Figure 7c). However, SEL exhibited significantly worse performance in Q4 (Fatigue) and Q5 (Distraction) than UNI. It may be because SEL (also SAL) makes relatively large motion effects for the movements of small-area links; 11 Ps said, "SEL was excessive in expressing either the entire or partial movements, leading to increased fatigue and distraction."

UNI expresses all segment movements with equal weights and may miss some detailed internal movements. Fourteen Ps mentioned, "UNI failed to represent some dancer movements, or its intensity was weak." However, UNI was preferred by those who considered the detailed motion effects uncomfortable. Fourteen Ps said, "UNI was smooth, making it comfortable to watch."

Regarding Q7 (Motion Sickness), only one P experienced motion sickness with UNI, two with SAL, and three with SEL. They reported that experiencing motion sickness affected only that trial, with no or negligible effects on the other trials.

Our algorithm controls the chair's yaw motion using the internal movements of *Shoulder*. This method makes the motion effects, especially those representing the torso's movements, more intuitive, allowing users to easily understand what the effects express. Two Ps commented, "The motion effect in expressing the movements was proper. The clear representation of the torso's turning actions was especially appealing." This advantage seems to have contributed to the high Q2 (Causality) scores of over 80 in all three conditions.

#### 5.4.2 Effects of Videos

As in Figure 7b, *Video* had significant effects in Q1 (Harmony), Q4 (Fatigue), and Q6 (Preference). Notably, the participants preferred *Ballet* and *Hiphop* to *Nitty* and *Charleston* with significant differences. A distinct pattern was that *Ballet* and *Hiphop* included many body turns, while *Nitty* and *Charleston* did not. Our system represents body turns as external movements by stimulating the user in the surge and sway axes. This method can be regarded as effective, as *Ballet* and *Hiphop* showed significantly higher harmony scores than *Nitty* and *Charleston*. In the raw data, when turns were present, all scores positively increased compared to when turns were absent.

#### 5.4.3 Summary

Each of the three conditions of the weights for combining the internal motion proxies has advantages and disadvantages. UNI effectively conveys overall internal movements but fails to capture detailed dance actions. In contrast, SEL accurately represents the detailed link movements but carries a higher risk of inducing motion sickness. SAL serves as an intermediate solution. Consequently, no preference differences were found among the three conditions. However, the three motion effects showed a high degree of causality, enabling the motion effects to occur at the user's desired timing.

## 6 Final Evaluation

This section reports another user study that compares our method with two similar existing algorithms that rely on the concept of motion proxy in terms of user experiences.

### 6.1 Motion Effect Generation Algorithms

#### 6.1.1 Ours: Selective SEL

We selected SEL for our algorithm among the three weighting policies tested. It demonstrated better expressiveness (Harmony and Detail) than the others.

#### 6.1.2 Multiple Articulated Bodies: MAB

In Han et al. [9], an algorithm is presented to automatically generate motion effects that express the movements of multiple articulated bodies from motion capture data, e.g., two fighting humans. A motion proxy is defined for each link in the articulated bodies, and the link motion proxies are combined to delineate the entire scene motion. In particular, all steps are focused on the viewer's perception, projecting all data and parameters into the camera space. Our work can be regarded as a successor of this framework in [9] as a specialization to human dances.

Han et al. used a commercial 3-DoF motion chair for four persons [9]. We implemented and turned their algorithm to work well with our 6-DoF motion chair for one person. The *SIZE* method in [9] was used to determine the weights of multiple links, which we also use to merge the internal movements. This version is called MAB (Multiple Articulated Bodies) in the rest of this paper.

#### 6.1.3 Pixel Motion Estimation: PME

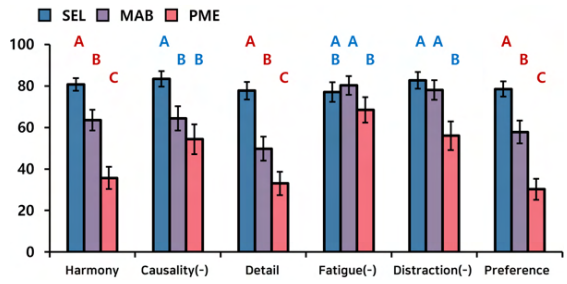
A more recent work of Han et al. [8] uses a scene flow model [51] to extract a 3D displacement vector for each pixel in the 2D scene. Each pixel's visual saliency is also estimated using a saliency detection model [41]. Then, the 3D motions of all pixels in the scene are combined using their saliency values as weights to compose a motion proxy that comprehensively expresses all motion components. This framework is very general and capable of generating different classes of motion effects, such as camera- and object-based motion effects, using the same method.

We implemented the pixel-level motion generation algorithm of Han et al. [8] to work well with our 6-DoF motion chair. We chose the SH condition (subject-relative visual mode and MPC with a high-pass filter), which demonstrated the best overall performance. This algorithm will be called PME (Pixel Motion Estimation).

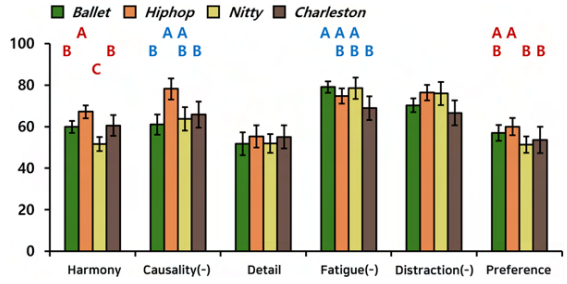
Both MAB and PME use only roll, pitch, and heave, as the original algorithms were designed for 3-DoF motion chairs. We acknowledge that our algorithm SEL, which utilizes the full 6-DoF motion, has inherent advantages over MAB and PME in this regard. We are unaware of other 6-DoF motion effect generation algorithms suitable for comparison with our algorithm.

### 6.2 Methods

The experimental methods were very similar to those of the previous experiment. The same details are not repeated for brevity.




(a) Arranged by *Motion Effect*




(b) Arranged by *Video*

Figure 8: Average scores obtained from User Study 2. The notation methods are the same as those of Figure 7.

### 6.2.1 Participants

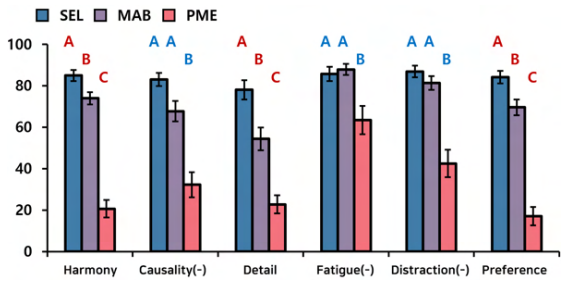
We recruited twenty volunteers (13 males and 7 females, 19–31 years old, average age 23.2, not participated in the previous experiment) for this user study. They were paid approximately USD 18 after the experiment.

### 6.2.2 Experimental Conditions

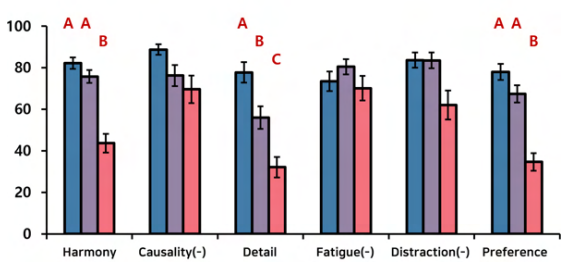
This study also used the choreography sequences in Table 1. For each of the four videos, we generated three sets of motion effects using the three algorithms: SEL, MAB, and PME. The perceived intensity of motion effects significantly impacts user experiences [12, 13]. We adjusted the scales of the three motion effect sets to equalize their perceived intensity levels. To this end, an adequate motion scale was set for SEL based on tests with five participants. Then, the perceptual strengths of MAB and PME were made the same as SEL by a method of adjustment with six participants.

## 6.3 Results

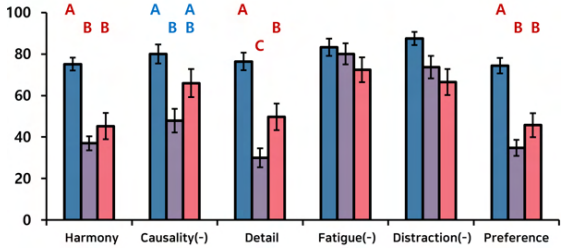
The data analysis methods were the same as the first user study. Figure 8 shows the average scores obtained in this study. *Motion Effect* was significant in all questions: Q1 (Harmony; $F(1.72, 32.72) = 73.43$, $p < .0001$), Q3 (Detail; $F(1.58, 30.07) = 42.63$, $p < .0001$), and Q6 (Preference; $F(1.80, 34.26) = 78.21$, $p < .0001$) by AN, and Q2 (Causality; $\chi^2(2) = 39.1418$, $p < .0001$), and Q4 (Fatigue; $\chi^2(2) = 7.1395$, $p = .0282$), and Q5 (Distraction; $\chi^2(2) = 37.7049$, $p < .0001$) by KW. *Video* was significant in Q1 (Harmony; $F(2.20, 41.74) = 14.39$, $p < .0001$) and Q6 (Preference; $F(2.35, 44.57) = 4.76$, $p = .0099$) by AN, and Q2 (Causality; $\chi^2(3) = 13.0843$, $p = .0045$) and Q4 (Fatigue; $\chi^2(3) = 8.2507$, $p = .0411$) by KW. As simple effect tests, we ran one-way repeated-measures ANOVA on the data of each *Video* with an independent variable of *Motion Effect*. The results are marked in Figure 9. Complete statistics results are provided in the supplemental material.



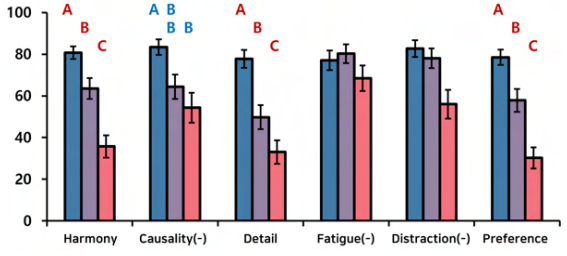

(a) *Ballet*




(b) *Hiphop*




(c) *Nitty*




(d) *Charleston*

Figure 9: Individual scores collected from User Study 2. The notation methods are the same as those of Figure 7.

## 6.4 Discussion

### 6.4.1 Effects of Main Factors

SEL outperformed MAB or PME with significant differences in almost all measures (Figure 8a). Eighteen Ps noted on SEL, "It represented limb movements more vividly than the others." in relation to Q3 (Detail). It can contribute to the higher scores of SEL in Q1 (Harmony) and Q2 (Causality). Although a detailed representation of motion effects can induce higher risk of motion sickness as mentioned in Section 5.4.1, SEL maintained competitive scores in Q4

(Fatigue) and Q5 (Distraction) to MAB and PME. Motion cues that match visual cues can reduce motion sickness [18], and it appears that SEL adequately matches the motion chair movements with the dancer's visual cues. Consequently, SEL obtained the best score of Q6 (Preference) among the three motion effects.

Both MAB and PME used a single motion proxy, which seems insufficient to adequately capture the complex dance movements. Sixteen Ps who experienced PME and 18 Ps who experienced MAB reported, "The motion effects failed to represent certain dance movements or expressed them with weaker intensity." In particular, PME received the significantly lowest scores in Q1 (Harmony), Q3 (Detail), and Q6 (Preference). It may be because that the videos' background used in the experiment was static, monochromatic, and repetitive, degrading the performance of the scene flow estimation model [51] and the overall use experiences.

Figure 8b shows that the effect of *Video* followed a similar trend to the previous experiment. That is, the two dance choreographies that included body turns, *Ballet* and *Hiphop*, received generally more favored responses. In the next section, we examine the data in the individual videos more carefully.

#### 6.4.2 Interaction Effects

The key contribution of our algorithm lies in its ability to separate the external and internal movements and represent them using the independent DoFs of the motion chair. It proved fruitful in *Hiphop* (Figure 9b), which emphasizes leg actions with large horizontal external displacements. Since MAB and PME treat all link motions identically, they provided only simple motion chair movements, mainly reflecting the dancer's horizontal translation. It seems to be the primary cause for the significantly lower scores in Q3 (Detail). It is encouraging that *Hiphop* received higher scores in most measures than the other choreographies, highlighting the effectiveness of our key ideas (Figure 8b).

Similar results were observed in *Charleston*, where the choreography had simultaneous external and internal movements (Figure 9d). Most Ps reported, "SEL well captured the dynamic link movements," while seven Ps commented for MAB and PME, "The motion effects commonly failed to express the limb movements."

Our algorithm uses nonlinear scaling to emphasize small dance movements. Its benefit was evident in *Ballet*, which involved slow internal movements. Here, SEL received very favorable responses, including a high preference score of 78, significantly greater than those of MAB and PME (Figure 9a). One P praised SEL for *Ballet*, "The motion effect captured even the subtle movements of the limbs and enhanced the immersion. I felt like the chair motion perfectly matched the dancer's turn and movements." Conversely, MAB and PME use merged link movements without scaling, which can obscure low-speed dance actions. Seven Ps left the same responses on MAB and PME for *Ballet*, "The chair movements either failed to represent specific dance actions or lacked sufficient details."

Our algorithm merges the internal movements using a few internal motion proxies. It was particularly advantageous in *Nitty*, which included many actions where the dancer's upper and lower body moved in opposite directions. In such cases, MAB and PME, which relied on a single proxy, had the effects of the upper and lower body movements cancel out each other, resulting in abnormally weak effects. In addition, Ps often experienced sudden and sharp motion effects when the movements were not canceled out. Two Ps commented, "When the character rotated, I felt the chair rotated excessively." It explains that MAB and PME received significantly lower Q2 (Causality) and Q4 (Fatigue) scores in *Nitty* than SEL, unlike the other dance choreographies (Figure 9c).

#### 6.4.3 Motion Sickness

Motion sickness was reported by four Ps with SEL, two Ps with MAB, and two Ps with PME out of the 20 Ps. Prior to the experiment, we tuned the intensity and duration of all motion effects to be similar, as they influence motion sickness [11, 15]. Hence, the overall level of motion sickness was low, but SEL tends to cause more than MAB and PME.

Our algorithm is to express dance, which consists of many vigorous actions. However, the lengths of dance scenes in our target applications are relatively short, e.g., a few minutes. These two facts have contradictory effects on motion sickness. Keeping it in mind, we need to control the motion scaling parameters, $\beta^{tln}$ and $\beta^{rtn}$, for the best trade-off between achieving good expressiveness and avoiding motion sickness.

#### 6.4.4 Summary

Our algorithm SEL demonstrated substantially higher scores in all measures about user experiences than the other algorithms MAB and PME implemented from the previous work. This result indicates that the features we devised to express human dance are effective for motion effects. The overall level of motion sickness was also low. However, it is emphasized that our algorithm utilizes the full six DoFs of the motion platform, while the other two methods use only three DoFs because of their original designs.

## 7 Limitations and Future Work

Our algorithm has a few limitations and the associated topics for future work: (1) Our algorithm assumes that high-quality 3D motion data for human dance is available and cannot be applied to general 2D scenes. Methods for human 3D pose estimation in 2D videos are rapidly evolving in computer vision, and integrating them with our algorithm can improve the applicability to a great extent. (2) Our algorithm considers only the dancer's visual movements. Dance generally involves music that modulates perceptual properties [23] and emotions [44]. Adding such components as an input to motion effect generation can further improve user experiences, although our understanding of the motion effects' influence on emotion is immature yet [12, 13]. (3) This work targeted regular users, but collaboration with dance experts is likely to unveil valuable insights to further improve the system. (4) Our method may be effective for other similar genres of human motion, and this direction of extension may bring fruitful outcomes.

## 8 Conclusions

This paper has presented an automatic generation algorithm designed to express human dance using a six-DoF motion simulator for improved multisensory experiences. Our algorithm independently processes a dancer's external and internal movements and decomposes the internal movements into those of a few body segments. These movements are represented by multiple motion proxies, and they are scaled and merged into two motion proxies, external and internal, using nonlinear scaling and weights. These two motion proxies are converted to motion commands using the standard MPC. The perceptual performance of our algorithm is validated through two user studies. Additionally, the limitations of our current system and plans for improvements are also discussed.


## Acknowledgments

This research was supported in part by research grants (2022R1A2C2091161 and RS-2024-00451947) from NRF and an ITRC (RS-2024-00437866) from IITP of Korea.


## References


[1] AwesomeDog. Awesomedog mocap dance packs, 2023. 5

[2] T. Bi, P. Fankhauser, D. Bellicoso, and M. Hutter. Real-time dance generation to music for a legged robot. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1038–1044. IEEE, 2018. 2

[3] E. Camacho and C. Alba. *Model Predictive Control*. Advanced Textbooks in Control and Signal Processing. Springer London, 2013. 4

[4] S. Casas, R. Olanda, and N. Dey. Motion cueing algorithms: a review: algorithms, evaluation and tuning. *International Journal of Virtual and Augmented Reality (IJVAR)*, 1(1):90–106, 2017. 5

[5] D. L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, 1995. 3

[6] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994. 3

[7] E. C. Gonen, Y. J. Chae, and C. Kim. Imitation of human upper-body motions by humanoid robots. In *16th International Conference on Ubiquitous Robots*, pp. 334–341. IEEE, 2019. 2

[8] S. Han, J. Ahn, and S. Choi. Generating Haptic Motion Effects for General Scenes to Improve 4D Experiences. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*, pp. 51–60, 2024. 2, 7

[9] S. Han, J. Park, and S. Choi. Generating haptic motion effects for multiple articulated bodies for improved 4d experiences: A camera space approach. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–17. ACM, 2023. 1, 2, 3, 7

[10] S. Han, G. Yun, and S. Choi. Camera space synthesis of motion effects emphasizing a moving object in 4d films. In *2021 IEEE Virtual Reality and 3D User Interfaces*, pp. 670–678, 2021. 2, 3, 5

[11] T. Irmak, V. Kotian, R. Happee, K. N. de Winkel, and D. M. Pool. Amplitude and temporal dynamics of motion sickness. *Frontiers in Systems Neuroscience*, 16:866503, 2022. 9

[12] D. Jeong, S. H. Han, D. Y. Jeong, K. Kwon, and S. Choi. Investigating 4d movie audiences' emotional responses to motion effects and empathy. *Computers in Human Behavior*, 121:106797, 2021. 8, 9

[13] D. Y. Jeong, S. H. Han, S. Choi, D. Jeong, and K. Kwon. Investigating perceived emotions and affects of a scene, and the user satisfaction with motion effects in 4d movies. *International Journal of Industrial Ergonomics*, 85:103173, 2021. 8, 9

[14] K. Jung, S. Kim, S. Oh, and S. H. Yoon. Hapmotion: motion-to-tactile framework with wearable haptic devices for immersive vr performance experience. *Virtual Reality*, 28(1):13, 2024. 2

[15] R. S. Kennedy, K. M. Stanney, and W. P. Dunlap. Duration and exposure to virtual environments: sickness curves during and across sessions. *Presence: Teleoperators & Virtual Environments*, 9(5):463–472, 2000. 9

[16] C. Kim, D. Kim, and Y. Oh. Adaptation of human motion capture data to humanoid robots for motion imitation using optimization. *Integrated computer-aided engineering*, 13(4):377–389, 2006. 1

[17] G. Kim, Y. Wang, and H. Seo. Motion control of a dancing character with music. In *Proceedings of 6th IEEE/ACIS International Conference on Computer and Information Science*, pp. 930–936. IEEE, 2007. 2

[18] J. Kim, J. Hwang, and T. Park. Effect of motion cues on simulator sickness in a flight simulator. In *Virtual, Augmented and Mixed Reality. Design and Interaction: 12th International Conference, VAMR 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I 22*, pp. 493–506. Springer, 2020. 9

[19] S. Kim, C. Kim, B. You, and S. Oh. Stable whole-body motion generation for humanoid robots to imitate human motions. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2518–2524. IEEE, 2009. 1

[20] T. Kobayashi, E. Dean-Leon, J. R. Guadarrama-Olvera, F. Bergner, and G. Cheng. Whole-body multicontact haptic human–humanoid interaction based on leader–follower switching: A robot dance of the "box step". *Advanced Intelligent Systems*, 4(2):2100038, 2022. 2

[21] K. Kosuge, T. Hayashi, Y. Hirata, and R. Tobiyama. Dance partner robot-ms dancer. In *Proceedings of 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 4, pp. 3459–3464. IEEE, 2003. 2

[22] K. Ladenheim, R. McNish, W. Rizvi, and A. LaViers. 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*, pp. 243–251, 2020. 2

[23] J. Lee and S. Choi. Real-time perception-level translation from audio signals to vibrotactile effects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2567–2576, 2013. 9

[24] J. Lee, B. Han, and S. Choi. Interactive motion effects design for a moving object in 4d films. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pp. 219–228, 2016. 2

[25] J. Lee, B. Han, and S. Choi. Motion effects synthesis for 4d films. *IEEE Transactions on Visualization and Computer Graphics*, 22(10):2300–2314, 2016. 2, 3

[26] J. Lee, J. Park, and S. Choi. Absolute and differential thresholds of motion effects in cardinal directions. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. Association for Computing Machinery, 2021. 4

[27] B. Lim, S. Han, and S. Choi. Image-based texture styling for motion effect rendering. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10, 2021. 2

[28] Z. Liu, Y. Koike, T. Takeda, Y. Hirata, K. Chen, and K. Kosuge. Development of a passive type dance partner robot. In *Proceedings of 2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 1070–1075. IEEE, 2008. 2

[29] H. Luan, Y. Wang, L. Huang, L. Wang, G. Lv, W. Gai, X. Luan, and C. Yang. Multi-sensory consistency experience: A 6-dof simulation system based on video automatically generated motion effects. In *Computer Graphics International Conference*, pp. 462–473. Springer, 2023. 2

[30] C. Mousas. Performance-driven dance motion control of a virtual partner character. In *Proceedings of IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 57–64. IEEE, 2018. 2

[31] S. Nakaoka, A. Nakazawa, F. Kanehiro, K. Kaneko, M. Morisawa, and K. Ikeuchi. Task model of lower body motion for a biped humanoid robot to imitate human dances. In *Proceedings of 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3157–3162. IEEE, 2005. 1

[32] A. Nakazawa, S. Nakaoka, K. Ikeuchi, and K. Yokoi. Imitating human dance motions through motion structure analysis. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, pp. 2539–2544. IEEE, 2002. 1

[33] M. Ohkita, Y. Bando, Y. Ikemiya, K. Itoyama, and K. Yoshii. Audio-visual beat tracking based on a state-space model for a music robot dancing with humans. In *Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5555–5560. IEEE, 2015. 2

[34] D. F. Paez Granados, B. A. Yamamoto, H. Kamide, J. Kinugawa, and K. Kosuge. Dance teaching by a robot: Combining cognitive and physical human–robot interaction for supporting the skill learning process. *IEEE Robotics and Automation Letters*, 2(3):1452–1459, 2017. 2

[35] J. Park, S. Han, and S. Choi. Merging camera and object haptic motion effects for improved 4d experiences. In *2023 IEEE International Symposium on Mixed and Augmented Reality*, pp. 1036–1044. IEEE, 2023. 2, 5

[36] R. Ponmanadiyil and M. H. Woolhouse. Eye movements, attention, and expert knowledge in the observation of bharatanatyam dance. *Journal of eye movement research*, 11(2), 2018. 3, 4, 5

[37] R. Qin, C. Zhou, H. Zhu, M. Shi, F. Chao, and N. Li. A music-driven dance system of humanoid robots. *International Journal of Humanoid Robotics*, 15(05):1850023, 2018. 2

[38] O. E. Ramos, N. Mansard, O. Stasse, C. Benazeth, S. Hak, and L. Saab. Dancing humanoid robots: Systematic use of osid to compute dynamically consistent movements following a motion capture pattern. *IEEE Robotics & Automation Magazine*, 22(4):16–26, 2015. 2

[39] J.-H. Seo, J.-Y. Yang, J. Kim, and D.-S. Kwon. Autonomous humanoid robot dance generation system based on real-time music input. In *2013 IEEE International Workshop on Robot and Human Communication*, pp. 204–209. IEEE, 2013. 2

[40] S. Shin, B. Yoo, and S. Han. A framework for automatic creation of motion effects from theatrical motion pictures. *Multimedia Systems*,

20:327–346, 2014. 2

[41] Y. Su, J. Deng, R. Sun, G. Lin, H. Su, and Q. Wu. A unified transformer framework for group-based segmentation: Co-segmentation, co-saliency detection and video salient object detection. *IEEE Transactions on Multimedia*, 2023. 7

[42] M. Systems. Motion platforms: Ps-6tm-150. https://motionsystems.eu/product/motion-platforms/ps-6tm-150/. Accessed: 2024-05-11. 5

[43] T. Takeda, Y. Hirata, and K. Kosuge. Dance step estimation method based on hmm for dance partner robot. *IEEE Transactions on Industrial Electronics*, 54(2):699–706, 2007. 2

[44] D. Vastfjall. A review of the musical mood induction procedure. *Musicae scientiae*, 6(1; SPI):173–212, 2002. 9

[45] H. Wang and K. Kosuge. Control of a robot dancer for enhancing haptic human-robot interaction in waltz. *IEEE Transactions on Haptics*, 5(3):264–273, 2012. 2

[46] S. Weisberg. Yeo-johnson power transformations. *Department of Applied Statistics, University of Minnesota. Retrieved June*, 1:2003, 2001. 6

[47] Wikipedia contributors. Choreography. https://en.wikipedia.org/wiki/Choreography, 2024. Accessed: 2024-05-17. 1

[48] M. H. Woolhouse and R. Lai. Traces across the body: influence of music-dance synchrony on the observation of dance. *Frontiers in human neuroscience*, 8:965, 2014. 3

[49] G. Xia, J. Tay, R. Dannenberg, and M. Veloso. Autonomous robot dancing driven by beats and emotions of music. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pp. 205–212, 2012. 2

[50] K. Yamane and J. Hodgins. Simultaneous tracking and balancing of humanoid robots for imitating human motion capture data. In *Proceedings of 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2510–2517. IEEE, 2009. 1

[51] G. Yang and D. Ramanan. Upgrading optical flow to 3d scene flow through optical expansion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1334–1343, 2020. 7, 9

[52] YouTube. Awesome dog mocap dance set 3 - iclone, 2020. 6

[53] YouTube. Motion capture ballet dance set 3 demo - awesome dog mocap, 2020. 6

[54] YouTube. Solo charleston fbx mocap dance pack demo, 2023. 6

[55] G. Yun, H. Lee, S. Han, and S. Choi. Improving viewing experiences of first-person shooter gameplays with automatically-generated motion effects. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2021. 2