ORIGINAL ARTICLE

HapMotion: motion-to-tactile framework with wearable haptic devices for immersive VR performance experience

Kyungeun Jung[1] · Sangpil Kim[2] · Seungjae Oh[3] · Sang Ho Yoon[1]

Abstract

We present a novel haptic rendering framework that translates the performer's motions into wearable vibrotactile feedback for an immersive virtual reality (VR) performance experience. Here, we employ a rendering pipeline that extracts meaningful vibrotactile parameters including intensity and location. We compute these parameters from the performer's upper-body movements which play a significant role in a dance performance. Therefore, we customize a haptic vest and sleeves to support vibrotactile feedback on the frontal and back parts of the torso and shoulders as well. To capture essential movements from the VR performance, we propose a method called motion salient triangle (MST). MST utilizes key skeleton joints' movements to compute the associated haptic parameters. Our method supports translating both choreographic and communicative motions into vibrotactile feedback. Through a series of user studies, we validate the user preference for our method compared to the conventional motion-to-tactile and audio-to-tactile methods.

Keywords Virtual reality · Haptics · Wearables · Performance

1 Introduction

The emergence of musical performances in VR has enabled audiences to actively participate in virtual concerts without physical presence. A 360-degree view of the virtual stage and real-time interaction with online participants Moritzen (2022) promoted the perception of "Being There" (Sense of Presence) in VR performances Charron (2017); Velt et al. (2015). As mentioned in Webb et al. (2016), resolving a few challenges (e.g., enabling performers to switch levels of social interactions or providing subtle physical cues of audience engagement) in social co-presence has the potential to provide distributed liveness for remote platforms like VR. To this end, previous works contributed to improving virtual avatars' motions in VR to improve co-presence Yakura and Goto (2020); Kaneko et al. (2018); Yan et al. (2020); Wang et al. (2020). Still, researchers focus primarily on visual and auditory cues which limit the possibility of further improving the sense of presence.

Recently, researchers start adopting multi-sensory stimuli which encompass visual and auditory cues as well as haptic stimuli Melo et al. (2020). For example, researchers translated the offline music performer's movement, physical interaction with the instrument, and audio output into vibrotactile feedback Turchet et al. (2021). Although previous work enriched the audience's experience by adopting haptic stimuli, it only triggered discrete haptic feedback based on specific gestures or signals rather than reflecting the continuous motions. Another work Abe et al. (2022) provided haptic feedback based on the audience's excitement level computed by biometric data (e.g., pulse). However, the biometric data could deliver haptic feedback that is not relevant to the performance context which lowers the sense of unity. Thus, we focus on forming a haptic rendering pipeline that

✉ Seungjae Oh
  oreo329@khu.ac.kr

✉ Sang Ho Yoon
  sangho@kaist.ac.kr

  Kyungeun Jung
  kyungeun.jung@kaist.ac.kr

  Sangpil Kim
  spk7@korea.ac.kr

[1] Graduate School of Culture Technology, KAIST, Daejeon 34141, Korea

[2] Department of Artificial Intelligence, Korea University, Seoul 02841, Korea

[3] Department of Software Convergence, Kyung Hee University, Suwon 17104, Korea
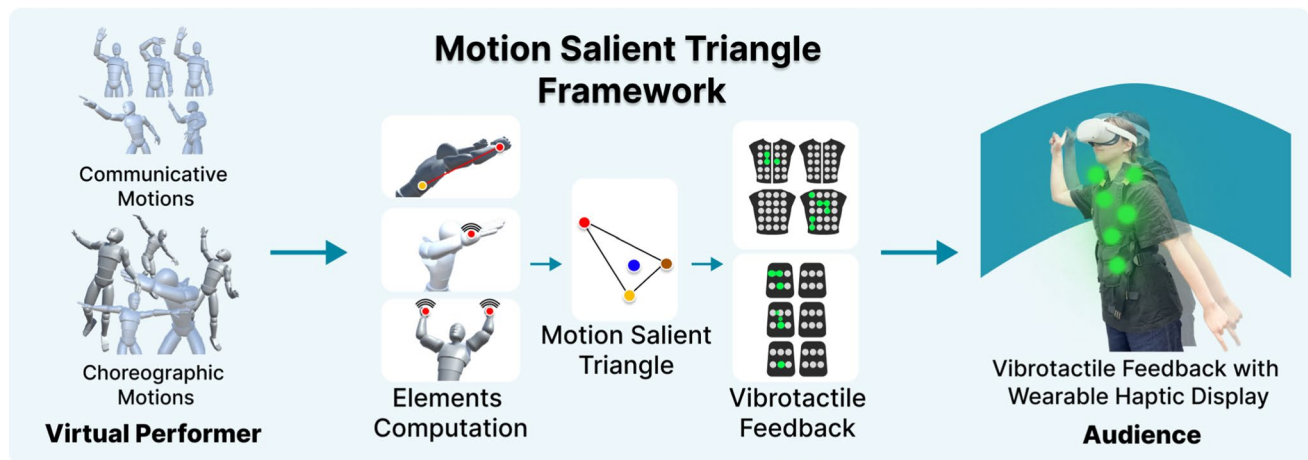
Fig. 1 HapMotion is a motion-to-tactile framework that translates the performer's motion in real time to enable an immersive VR performance experience

blended with the context of a VR performance such as the performer's choreographic and communicative motions.

Translation of visual and audio data into haptic feedback has been explored in many contexts. Previous work automatically generated haptic rendering parameters based on given input video and audio data Li et al. (2021); Kim et al. (2013); Rasool and Sourin (2014). For example, the saliency map of a movie frame integrated with psycho-acoustic features computes tactile intensity and rendering locations for vibrotactile feedback on a chair Li et al. (2021). However, extracting features from the two-dimensional (2D) video file was inefficient for getting precise motion information for immersive vibrotactile feedback. Getting detailed and specific motion data is left in the challenge, which entails context information as well. In our work, referencing Yun et al. (2021), which considers the wholesome context of interaction methods while enjoying first-person shooter games, we integrate our idea to detect the contextual information from the performer's movement.

Inspired by previous works, we further propose an automatic haptic rendering pipeline that translates the performer's full 3D motion data into meaningful vibrotactile feedback (Fig. 1). The higher immersive level would incur audiences to feel like being virtual performers, which would further enhance the sense of embodiment under a given VR performance. To transfer effective vibrotactile feedback to users, we employ an upper-body wearable configuration.

With the proposed motion-to-tactile framework, we aim to enrich the immersion of VR performances. To go beyond utilizing conventional visual and audio data, our approach translates tactile sensation based on 3D feature points directly acquired from the performer's motions.

By controlling the vibrotactile intensity and localization properties, we enhance audiences' embodiment and attention by generating a coherent tactile sensation with the VR performance.

In this work, we devise our translation pipeline to cover a variety of performer motion types including communicative (*e.g., waving hands, passing a mic toward audiences, hand clapping)* and choreographic motions. By enabling a new contextual experience in VR performances, we expect to further increase the level of participation and immersion compared to existing offline or VR concerts. Our contributions are as follows:

- A novel haptic rendering algorithm to translate performer's diverse motion context into haptic intensity and localization parameters;
- A novel motion-to-tactile framework that converts multimedia contents including performers' 3D motions and audio data into vibrotactile feedback;
- A wearable haptic interface to support a full 3D upper-body haptic feedback;
- Analysis of user studies demonstrating the user experiences with *HapMotion* for naturalness, immersion, satisfaction, consistency, and embodiment.

2 Related work

2.1 Vibrotactile translation of multimedia data

Researchers attempted to translate multimedia data into meaningful haptic feedback in order to enhance the given audiovisual experiences Danieau et al. (2012); MacLean et al. (2017). This translation has been applied in watching

videos Abdur Rahman et al. (2010), undergoing rehabilitation Alamri et al. (2010), playing games Rehman et al. (2008), and experiencing 4D films Seo et al. (2018). In our work, we aim to translate multimedia data from live VR performances into haptic feedback to promote immersive experiences.

Audio-to-Tactile Translation The audio-to-tactile translation was initially suggested since adding tactile experience induces easy communication of auditory information to users Chang and O'Sullivan (2005); Remache-Vinueza et al. (2021). Moreover, researchers demonstrated that the audio-to-tactile translation improved the user's perception of physical stimuli Imschloss and Kuehnl (2019) and media experience Mazzoni and Bryan-Kinns (2016). The earlier works often translated musical information to haptic sensation in a chair embedded with multiple vibrotactile actuators Karam et al. (2010); Nanayakkara et al. (2009); Hayes (2012); Yamazaki and Ohkura (2018); Fontana et al. (2016); Altinsoy and Merchel (2010). These works mainly focused on providing spatiotemporal vibration patterns to the body surface based on acoustic features from given audio. Recent works bring audio-to-tactile experience into wearable such as belts Yamazaki et al. (2017), gloves Enriquez et al. (2020), armbands Turchet and Barthet (2018), jackets Hashizume et al. (2018), and whole-body suits West et al. (2019). Still, tactile translation solely based on audio modality has limitations in conveying full media contexts to users.

Visual-to-Tactile Translation To go beyond acoustic feature-based tactile experience, visual-to-tactile translation has been explored to mediate immersive experiences Kim et al. (2013); Kruijff et al. (2017). In previous works, the addition of haptic feedback created from visual stimuli enabled rich multimodal sensations to improve immersion level Kim et al. (2010); Wilson et al. (2016); Lin et al. (2021). To be specific, researchers employed the visual saliency map representing contextual event locations to generate spatiotemporal vibrotactile effects Li et al. (2021). However, the context of visual stimuli remains in 2D RGB pixels which cannot express the full 3D information of state-of-art multimedia contents (e.g., 3D videos and motion-captured skeleton data).

Motion-to-Tactile Translation In terms of motion-to-tactile translation, there have been a few trials of translating the movement of one's character or camera's gait. Hapseat Danieau et al. (2012) translate the first-person point-of-view simulation into three force feedback devices, trying to mimic the sensation of a principal actor might have felt while recording with sensors attached to one's body. There was also research on considering the camera motion position as an input and translating it into six numbered ERM motors attached in moving chair Seo et al. (2018). As the availability of diverse choices to render in haptic technologies, there are no guidelines for rendering human movement that is explicitly associated with performance and performer

in a virtual environment. In our work, we further extend the capability of the vibrotactile translation framework by supporting 3D motion data. We minimized the redundancy of attaching sensors to the performer itself and supported audiences with wearable devices to enjoy the haptic experience anywhere they prefer. Here, we devise a full 3D upper-body wearable-based haptic interface that automatically renders the performer's motion into meaningful vibrotactile feedback.

2.2 Haptic feedback for immersive experience

The number of virtual performances surged over the past years with the advancement of commercially available VR headsets and users' preference for remote participation Charron (2017). To enrich the experience of participating in VR live performances, recent work Yakura and Goto (2020) focused on improving the motions of virtual audience avatars. Also, various commercial platforms the show must go beyond (The show must go beyond 2022; Emotionwave XR) for virtual performances have been launched to support live music performances of global artists in VR.

For haptic feedback in virtual performance, previous work improved the sense of unity by sharing responses through visual, auditory, and haptic stimuli based on audiences' biometric data Abe et al. (2022). In this study, the sense of unity and embodiment increased. Aligned with this direction, we propose a system that focuses on the translation of virtual performers' motions into vibrotactile feedback. To support more immersive haptic feedback, we cover the whole upper body including the shoulder by integrating haptic sleeves along with the haptic vest.

In order to adapt to the environment of the haptic feedback for virtual performance, referencing offline performance research was inevitable. The offline musical haptic wearable device for audiences Turchet et al. (2019) helped audiences' musical experiences by leveraging the sense of touch in terms of providing new capabilities for creative participation.

A broad range of tactile feedback approaches has been suggested for immersive musical experiences for musicians. For example, virtual drums, Cellomobo Berdahl et al. (2008), and reed Smyth et al. (2006), offer musicians a chance to interact with musical instruments integrated with a haptic system. Moreover, virtual violin supporting drawing a bow Nichols (2002) and AirPiano Hwang et al. (2017) based on a mid-air haptic system improved the players' experience. Still, the haptic feedback system that exclusively supports the performance context is rare. To this end, we propose a haptic system that directly applies to VR performances.

Haptic feedback supports to convey of meaningful information. Particularly it is used as an important part of storytelling in various fields. Feel Effects Israr et al. (2014)

devised the goal of haptic vocabulary and authoring methods for creating realistic haptic representations. Providing relevant and analogous haptic feedback considering the ongoing cues from the context of the environment is the most important feature for establishing a haptic sensation. In our work, we create novel haptic rendering algorithms to translate motion which often contains contextual information about the performance.

2.3 Design approach for rendering vibrotactile feedback

The eccentric rotating mass (ERM) motors are commonly used to implement vibrotactile feedback García-Valle et al. (2020). Although it is hard to control the frequency of motors Miklós and Szabó (2015), researchers still utilize ERMs because of their performance (e.g., strong vibrations) and scalability (e.g., low cost) Yun et al. (2019); Li et al. (2021); Tawa et al. (2021); Park and Choi (2018); Zhang et al. (2020). Previous researchers have developed various graphical editing tools that allow animators to render pipelines Schneider et al. (2015); Cuartielles et al. (2012). For example, perceptually optimizing interpolation algorithms for sparse vibrotactile grids Schneider et al. (2015) is one of the authoring approaches to feeling the haptic sensation from the media data. In order to build a comfortable and effective tool for real-time VR concert Lalioti et al. (2021), it is essential to translate a haptic authoring tool that provides automatically generated haptic feedback through the data-driven algorithm.

Recently, researchers have explored the method to generate haptic effects with automatic authoring pipeline Israr and Poupyrev (2011); Israr et al. (2014, 2016). For example, previous work defined the foundational library of usable haptic vocabulary to explicitly match linguistic phrases to the corresponding haptic patterns Israr et al. (2014). Recent works introduce the use of the designer's voice to design vibrotactile haptic feedback for iterative haptic design process Degraen et al. (2021) and the cross-modal information (diegetic audio and localization of sounding objects) to automatically generate tactile effects Zhang et al. (2020). These design approaches tell that context-based automatic haptic feedback generation would be essential to accommodate a complex set of multimedia data.

Inspired by previous works, we propose an automatic vibrotactile translation framework that utilizes 3D motion data from VR performances. For our framework, we focus on understanding the motions' representative features and associated context which is known to be crucial components in general performance Blumenfeld-Jones (2008). Unlike manual authoring tools, our work suggests a fully automated vibrotactile rendering framework that extracts the representative key points from the performer's motions and converts them into real-time vibrotactile feedback.

3 Design space

In this section, we share survey results on how performers' motions occur during offline and VR performances. We analyzed the recorded videos of both offline and VR performances to categorize and quantify the types of motions and their occurrence frequencies.

3.1 Offline and VR performance survey

To better understand which motions could provide effective haptic feedback to promote an immersive experience, observing the types and characteristics of motions that occurred during offline and virtual performances is essential. Before getting into the deep analysis, we define the motion-type terminology: choreographic motion as the dance movements following the music during the concert or a technique of combining movements and performing them in dance Bisig (2022); Pehkonen (2017), and communicative motion as the actions that have functions to share their emotions and nonverbal communication between audiences Kaneko et al. (2018).

3.1.1 Survey method

We conducted a survey and analysis on the videos demonstrating the audiences' responses after the performer's actions. For our survey, we categorize the motion as communicative motion if performers aim to bond with the audience and gain social interaction (*e.g., Handing a microphone and waving to the audience*). We validate our categorization by checking the audiences' responses right after the communicative motion occurred. For example, we checked whether audiences shout, raise and clap, or sing along during offline performances and raising visualization feedback or adding text comments (e.g., Justin Biebers' Virtual Experience: flying hearts Bieber (2021)) during VR performances.

3.1.2 Live offline performances

We first listed up the sound source by selecting the two highest-ranked songs per year (between 2010 and 2021) from the most trustworthy music chart like *BILLBOARD HOT 100* Billboard (2022). We mainly picked songs that have 95~140 beats per minute (BPM) and support both choreographic and communicative motion. We categorize the tempo of music as *slow* (95-115 BPM), *medium* (115-135 BPM), and *fast* (135-140 BPM) Karageorghis et al. (2011).

Fig. 2 The footage of offline and VR performances. **A** and **C** represent communicative motions, while **B** and **D** indicate choreographic motions
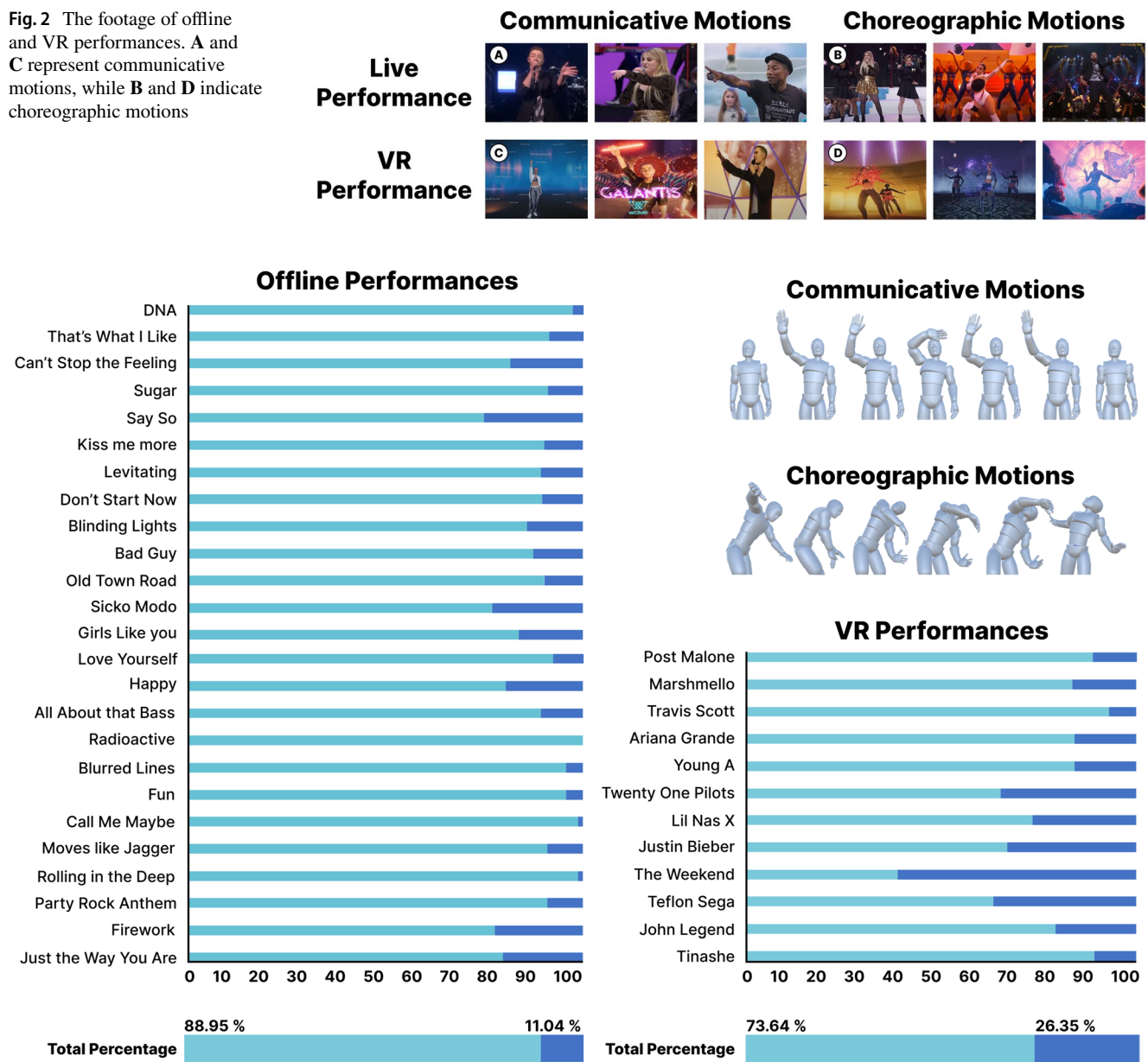

Fig. 3 The ratio of communicative motion and choreographic motion from 25 live offline performances (Left) and 12 virtual performances video clips (Right)

We found out that audiences' responses were apparent after the performer's communicative motion for songs with over 100 BPM. In this work, we focused on translating a solo artist rather than a group since a solo artist's motions clearly show representative choreographic motion and communicative motion (Fig. 2).

3.1.3 VR performances

All VR performances support converting performers into avatars where real performers go through a motion capture system to convert their motion data into live avatars.

We selected 15 pre-recorded VR live performances with 250 min of playing time from YouTube. Like offline concert selection, we choose the most viewed VR concert videos.

3.2 Survey takeaway and design considerations

Both offline and virtual performances are all about the music, choreography, and nonverbal communication between the audience and the players Kaneko et al. (2018).

Figure 3 shows the ratio of communicative and choreographic motions from the performance. According to our analysis, we found that live offline performances generally

consisted of 88.95% of choreography and 11.04% of communicative motions. We observed that the ratio of choreographic motions decreases to 73.64% and the communicative motions increase to 26.35% for VR performances.

Both offline and VR performances show higher occupation rates of choreographic motions compared to communicative motions. However, VR performances show a 10% higher rate than offline performances for communicative motions. This result indicates that we need to consider how to translate communicative motions in a more adaptive way for futuristic performance. As shown in Fig. 4a, communicative motions utilize more hands in order to convey nonverbal interaction such as pointing at audiences to induce sing-along or clapping.

Choreographic motions also consisted of movements that mainly use the upper body, which is regarded as the placement of the hands, and wrists in meaningful spaces. Therefore, we should focus on translating both the detailed and macro-level flow of upper-body motions. Inspired by previous works Tsai et al. (2022); Li et al. (2021); Fang (2021); Gonzalez-Franco and Peck (2018), we focused on five key aspects including naturalness, immersion, satisfaction, consistency, and embodiment when designing our system. To achieve this goal, we apply the following design considerations.

Naturalness To transfer both the sophisticated and overall flow of the performer's movements, we apply vibrotactile sensation to the whole upper body. We customize the haptic sleeves on both shoulders to fulfill the missing region from the existing haptic vest.

Immersion To render immersive vibrotactile feedback, we control spatiotemporal parameters with various warping approaches. We adjust vibrotactile amplitudes based on physics-based elements (e.g., acceleration, and distance from the audience's body) rather than applying constant intensity.

Satisfaction Musical interaction with choreographic motions is crucial in supporting accessibility and satisfaction for audiences Veronesi (2014). To this end, we devise our system to support audio-to-haptic feedback along with the vibrotactile feedback rendered from the performer's motions. We assume that the multimodal (music and motion) vibrotactile rendering approach could improve the haptic experience while enjoying VR performance.

Consistency To maintain consistency in the haptic experience, it is essential to provide integrated haptic feedback based on both performance context as well as performer's direct motion flow. To support this, we propose a novel rendering algorithm that operates a set of vibrotactile devices to convey the intended contexts from the performer's motions in real time with zero latency.

Embodiment Previous study showed that the embodiment of an outgroup can enhance empathy Thériault et al. (2021). Along this line, we believe that feeling the third person's motions would improve the embodiment of the virtual performer and enhance the virtual performance's immersion level. We aim to provide vibrotactile feedback that translates haptic location in a mirrored way. Here, *Body Swapping* is a key approach where synchronizing movements from both users result in an illusion of body ownership analogous to other bodily illusions Botvinick and Cohen (1998). Syncing movements in a mirrored way as if a user is glancing at the mirror shows enhancement in the relationship between the two participants. Thus, we horizontally flipped the locations of tactile feedback.
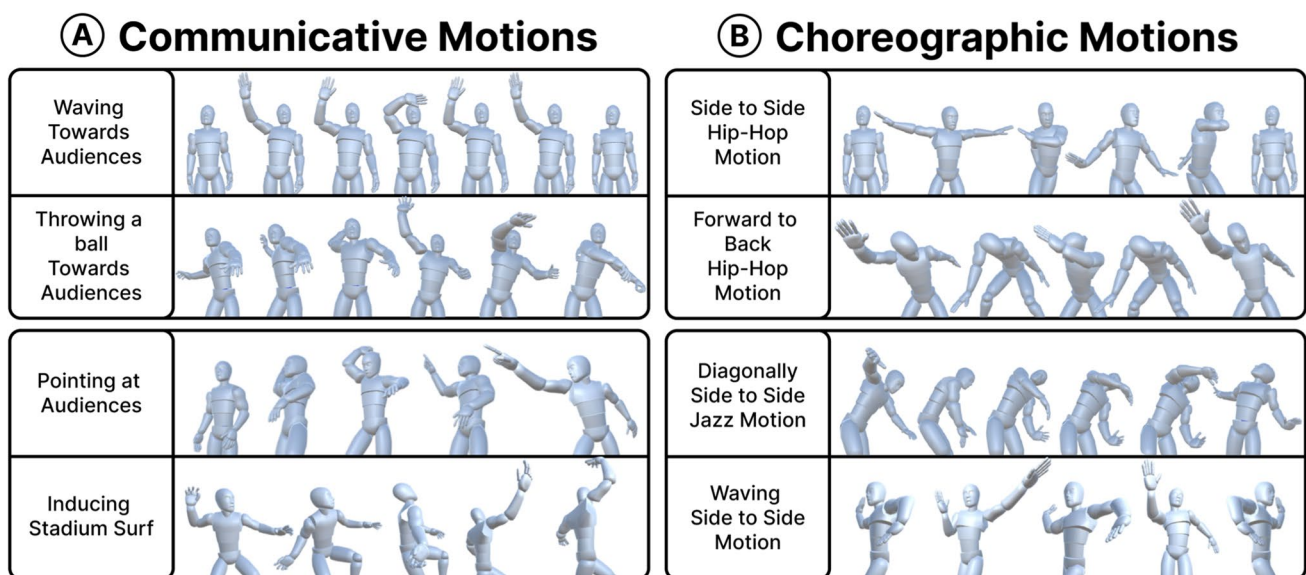

Fig. 4 **A** Exemplary communicative motions and **B** chosen choreographic motions excerpted from the survey

4 Motion-to-tactile translation approach

Previous researchers introduced a haptic effect using an RGB-image-based visual saliency map that works with audio data Li et al. (2021). In our case, we add 3D motion data to accommodate every movement of the performer into a meaningful haptic effect. We propose a motion salient triangle (MST) that aims to effectively translate characteristics of movements into vibrotactile haptic feedback. In this section, we describe our novel rendering design approach using the proposed MST. Our rendering approach using MST processes spatiotemporal parameters extracted from three-dimensional (3D) joint coordinates. Furthermore, one-dimensional (1D) haptic phantom sensation Park and Choi (2018) is adopted in order to express the detailed flow of the performer's motions during consecutive frames. We support robust real-time data processing without much data loss. Therefore, our method achieves a high-level correlation between vibrotactile effects and virtual performers' movement to improve the audience's experience in virtual performance.

4.1 Computing motion salient triangle from key element vertices

MST is a key motion event localization method for translating one's motion into vibrotactile feedback. As mentioned in Sect. 3.2, a large portion of choreographic and communicative motions includes upper-body movements. Moreover, we analyze that hands joint coordinates play a crucial role in upper-body movements such as *Handing the microphone to the audience* and *Inducing the audience to do Mexican surf*. For this reason, we assign hand joint coordinates as active

joint coordinates J_A that represents rich information from motion. In this work, we formulate 3D joint coordinates as $J = (x, y, z)$.

We further define root joint coordinates (J_R) and the center of mass of torso coordinates (J_T). As shown in Fig. 5, J_R represents a stable point on the shoulder located opposite to J_A side which reflects the balanced position while carrying out diverse motions. Since the shoulders' translation displacement is low compared to other joints during the performer's motion, we pick shoulders for J_R Golomer et al. (2009). J_T provides a stable point inside of a torso, which mostly sticks to its initial position. Using these two stationary points, our proposed algorithm considers not only micro-level motion flow but also the macro-level stream of movement in continuous frames. We name J_A, J_T and J_R as key element vertices which required to form MST.

By concatenating these key element vertices, we generate a 3D polygon. MST-based algorithms employ real-time human body tracking consisting of 32 joints from *Azure Kinect DK* Microsoft (2020). We designate J_A as either Hand$_{Right}$ or Hand$_{Left}$ joint given from the Azure Kinect. We place J_R in either Shoulder$_{Left}$ or Shoulder$_{Right}$ which is symmetrically opposite side of the J_A.

Referencing from computing the center of mass of human body segments Adolphe et al. (2017), we first consider the spine naval point as the center of mass in the human body. We carry out $r = \frac{R \cdot l}{Q}$ using Unity 3D engine. Here, R is the value of reactive force, which accounts for value 1, l refers to the length of the lever, which computed the height of the virtual character, and lastly, Q refers to a mass of the human body which calculated automatically. We then finally calibrate the center of mass of the torso through the mentioned equation.
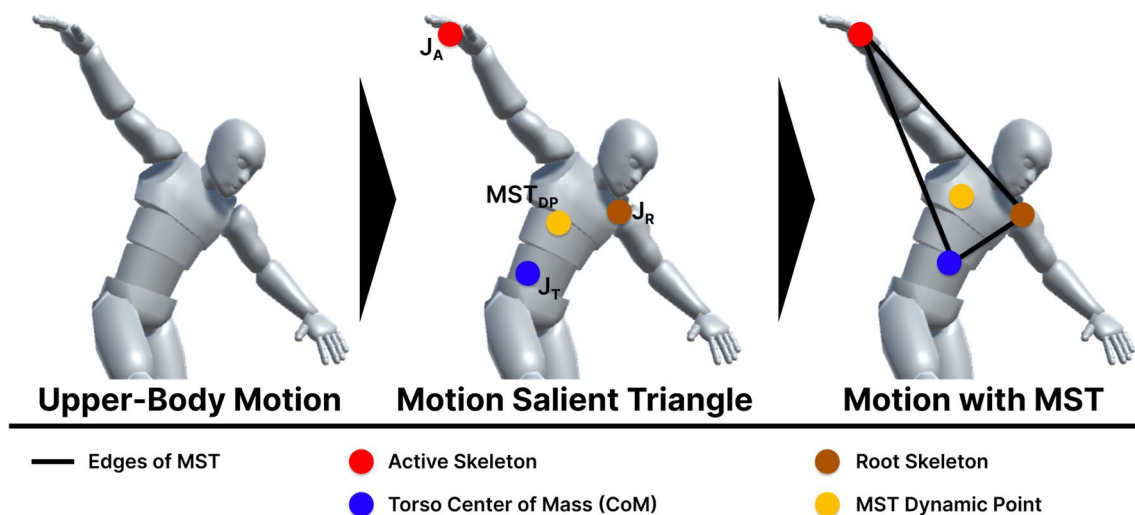

Fig. 5 Overall concept of MST. From the original upper-body motion (Left), we extract key elements vertices for MST (Middle). Then, we concatenate the vertices with edges (Right) and create a 3D triangle called motion salient triangle (MST)

MST Dynamic Point After creating the 3D triangle, we compute MST dynamic point (MST_{DP}) as shown in Eq. 1. Here, J_C refers to a centroid in MST. ω_{Torso}, ω_{Active}, and ω_{Root} indicate the weighting coefficients for each key element vertex. We set MST_{DP} with weighted distance from each key element vertices which translates the direct flow of the movement. For the initial frame, we set the ω values as 1 and adjust them afterward according to the movement of the performer.

$$\text{MSP}_{DP} = J_C + \frac{(J_A - J_C) \cdot \omega_{Active} + (J_R - J_C) \cdot \omega_{Root} + (J_T - J_C) \cdot \omega_{Torso}}{\omega_{Active} + \omega_{Root} + \omega_{Torso}}, \tag{1}$$

Figure 6 illustrates the overall system flow of our proposed haptic translation method described as follows:

1. Collect an offline performer's 3D joint data (Azure Kinect) in real-time and transfer joint data to a virtual avatar in the Unity plugin.
2. Compute the J_T point if the current frame is not the initial frame.
3. Set Joint_{Right} and Joint_{Left} as potential active joints and keep track of both distances to J_T. By comparing computed distances, we determine the number of active joints, J_A.
4. Compute acceleration of $J_A(s)$. If two J_As are above the threshold, we set two J_As for computing MST_{DP}. If one J_A is above the threshold, we set one J_A and J_R.
5. Distribute localization weights to each key element vertex (See 4.2.2) and compute MST_{DP} using Eq. 1.
6. Process mapping and warping of MST_{DP} (Fig. 7). If the MST_{DP} is inside of the bounded area, tactile location is assigned through the 3D warping method. If not, we consider it as surface direct mapping.
7. Set intensity level based on the distance value from MST_{DP} to J_T.
8. Increase the haptic intensity level if MST_{DP}'s acceleration goes above the threshold, mainly mentioned in Sect. 4.3. If 3D warping vertices to the exception nodes are illustrated in Fig. 11a, we employ 1D phantom sensation when adjusting the intensity level (See Fig. 11b).

4.2 Translating tactile location

4.2.1 Rendering MST dynamic point

The proposed algorithm maintains the controlled proportion of the distance between MST_{DP} and the surface of the torso. In our system, we stream direct *vector3*-type raycasting to the target point (TP). TP is the centroid among four representative joint coordinates, including the front and back of the torso and left/right shoulders (Fig. 7a).

Figure 8a shows the top and side view of the warping range and how raycasting stimulates each haptic node. The range of the warping boundary is set based on the range of motion (ROM) data from our previous surveys (See Sect. 3.2). Therefore, the performer's maximum and minimum X and Z ROM become the range of the X-axis and Z-axis of the warping boundary. We measure the maximum and minimum length between J_A and the local coordinate of the performer's J_T to define the range of warping availability. In general, the maximum and minimum ranges come out as 185 cm and 13 cm. If MST_{DP} is out of these ranges, we adjust MST_{DP} to the closest boundary.

As shown in Fig. 8a, there are two exemplary cases for assigning haptic feedback by 3D warping. By applying a homogeneous matrix, we convert MST_{DP} from 3D coordinates to the 2D haptic display nodes. Figure 8b indicates the surface mapping, which occurs when MST_{DP} hits the haptic display node directly. We further explain in more detail about 3D warping and direct surface mapping.

When it comes to the raycasting in the middle of consecutive actuator nodes (Node_A and Node_B, Node_C, Node_D), we deploy modified 1D phantom Park and Choi (2018). Figure 8c shows the haptic output after raycasting. As the ray hits among random actuator nodes, it will eventually be actuated at the same time with different intensity levels due to the 1D phantom sensation and 2D Grid-based sensation. We will give a description in Sect. 4.3.4 about two main cases of additive sensation.

3D Warping Fig. 8a illustrates how we process raycasting to warp the MST_{DP} to the haptic display node. The raycasting starts from MST_{DP} to the target point. If the ray hits the node and locates within the boundary, we set it as a haptic proxy.

In order to convey a natural and embodied user experience, our system aims for mirrored haptic feedback from the virtual performer. This denotes that audiences feel the flipped feedback of the performer's motions as if they are watching the mirror of their performer. We consider that mirrored rendering design enhances the level of embodiment while experiencing the vibrotactile feedback as mentioned in Sect. 3.2.

Direct Surface Mapping When MST_{DP} is smaller than the minimum warping range shown in Fig. 7b, MST_{DP} generally locates directly on the surface nodes of the performer's torso. In this case, these surface nodes become haptic proxies to transfer tactile feedback as shown in Fig. 8b.

Out-of-Range Projection Since our system supports real-time rendering, strong treatment for unexpected cases is inevitable. If MST_{DP} has been positioned out of the pre-calculated maximum range, we redefine the excluded MST_{DP} to the closest coordinate in the maximum range (Fig. 8c).
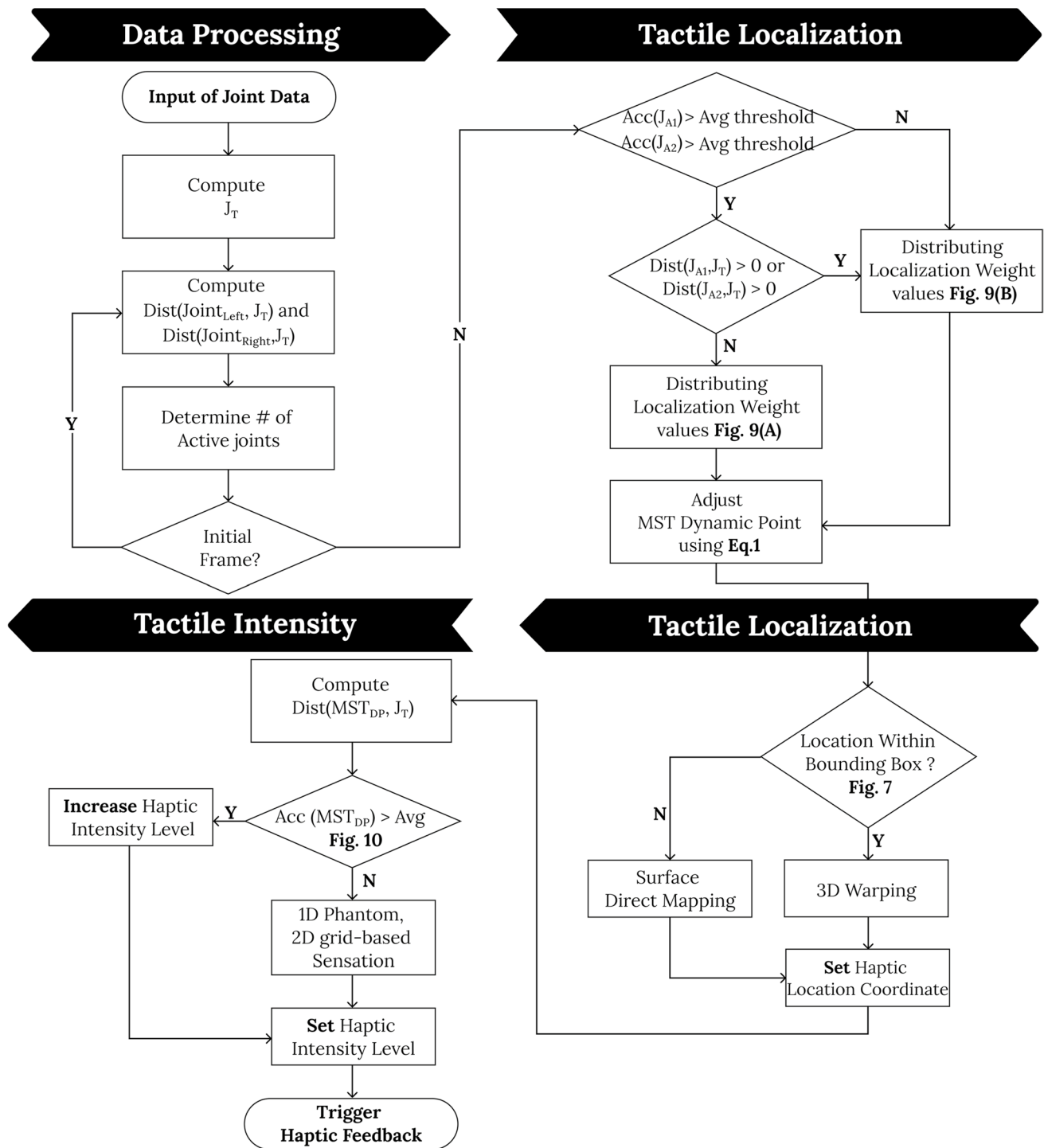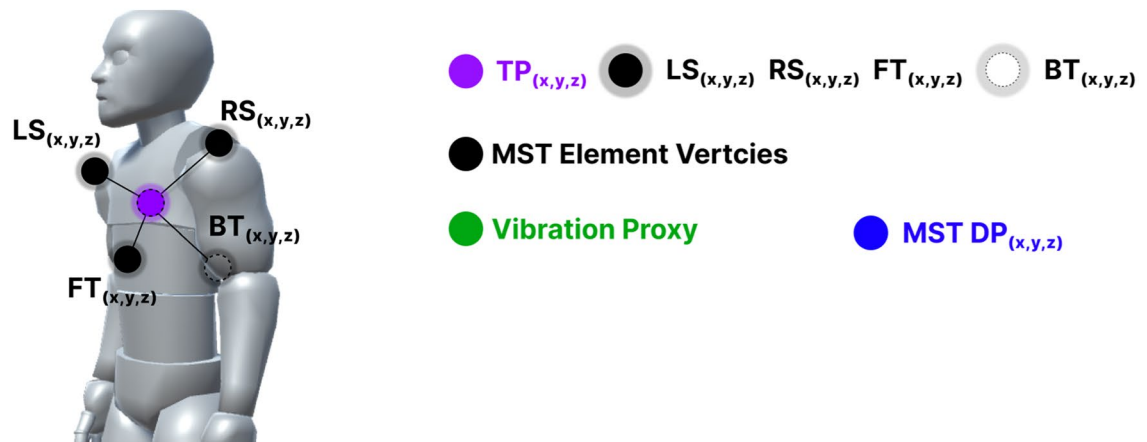
Fig. 6 Overall system flow of MST-based algorithms including data processing, tactile localization, and tactile intensity adjustment

4.2.2 Integrating motion context to MST dynamic point

To cover the various motion types from the virtual performance, we adjust weight distribution when computing MST_{DP}. We consider single active joint and dual active joints conditions.

For weight distribution, we compare each appointed J_A's acceleration data in every frame. In terms of the acceleration threshold, we compute the average acceleration for the previous three frames. If the acceleration value of the current frame is higher than the real-time threshold, we append the value of weight $(A_{t-2} - A_{t-1})$ on the J_A. Therefore, the
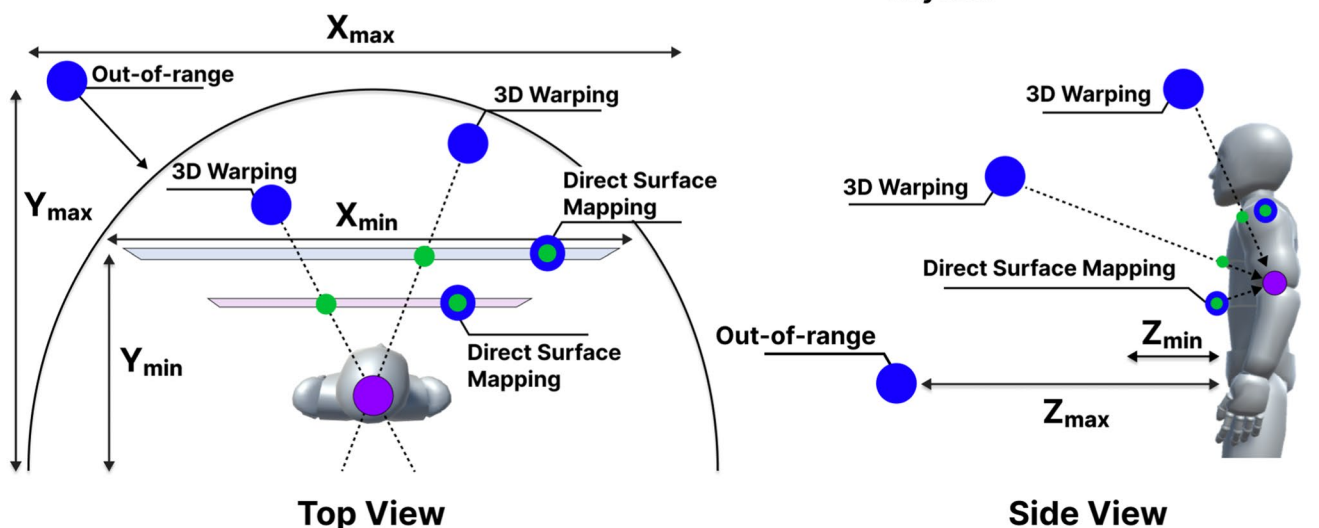
Fig. 7 Overall process of warping and mapping of MST$_{DP}$. **A** First, we compute the target point (TP). **B** Then, we process 3D warping or direct mapping within the defined ranges. There could be several mapping cases including 3D warping, direct surface mapping, and out-of-range projection. (See Fig. 8 as well.)

calculated weight value will be applied in real time with Eq. 1.

Regarding dual active joints, key element vertices consist of two J_A and sole J_T. In this case, we compute both the acceleration value and distance value. If the acceleration values in the current frame t for both active joints are higher than the real-time computing average threshold, we once again define there are two active joints to render. Then, we compute the distance value from left J_A to MST$_{DP}$ and right J_A to MST$_{DP}$ at the same time. By comparing the distance value of each active joint, we distribute the weight value $(A_{t-2} - A_{t-1})$ to the joint which records a higher distance

from MST$_{DP}$. Figure 9b adapts the condition for two active joints.

4.3 Translating tactile intensity

4.3.1 Hardware intensity calibration

To translate tactile intensity with a set of hardware, we first define the hardware calibration coefficient (C) to provide precise tactile stimuli. This calibration is to identify the relationship for our input–output actuators, which certifies the liable following results. We measured the output acceleration from each eccentric rotating mass (ERM)
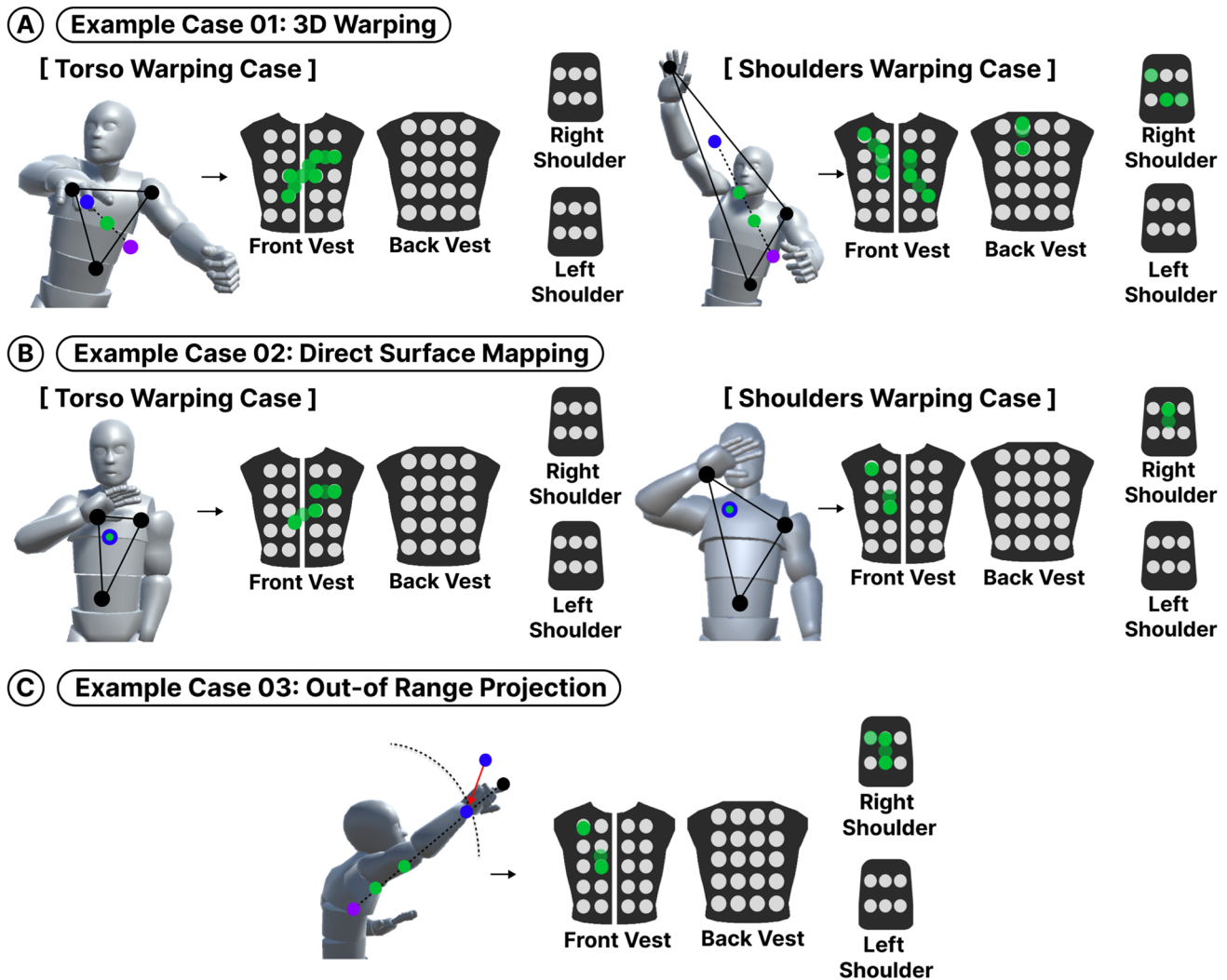
Fig. 8 We show examples of **A** 3D warping, **B** direct surface mapping, and **C** out-of-range projection while employing and MST-based localizing algorithm

using a high-precision 9DoF IMU (SparkFun, ICM-20948) while changing the input amplitude. The measured acceleration in each condition was fit to linear interpolation. For the output amplitude, corresponding vibration frequency from the bHaptics vest and sleeve recorded the vibrotactile actuators (range 1.00~4.37 G). Here, (G) refers to gravitational acceleration. The most effective vibrotactile frequency for human perception lies between 130 and 230 Hz Sun et al. (2022). To satisfy both vibrotactile intensity and frequency, we set the value C as 6 which corresponds to level 6 of bHaptics's intensity parameter (3.16 G with 142 Hz).

4.3.2 Intensity control strategy

To accurately simulate the sensation of the upper-body movement, we adjust the intensity level according to the distance of MST_{DP} to J_T. Therefore, controlling a fine level of intensity is necessary Li et al. (2021). We control the ERMs' intensity parameter value to effectively convey the performer's motions. Depending on the distance value of MST_{DP} to J_T, the level of tactile intensity is linearly combined. The larger ROM gets, the higher the tactile amplitude is. By adjusting tactile intensity based on the distance which represents the quantity of the motion from the performer, users would easily notice the flow of movements from the performer. The proposed intensity control strategy would benefit motions that contain precise and dynamic contexts like choreographic and communicative motions.
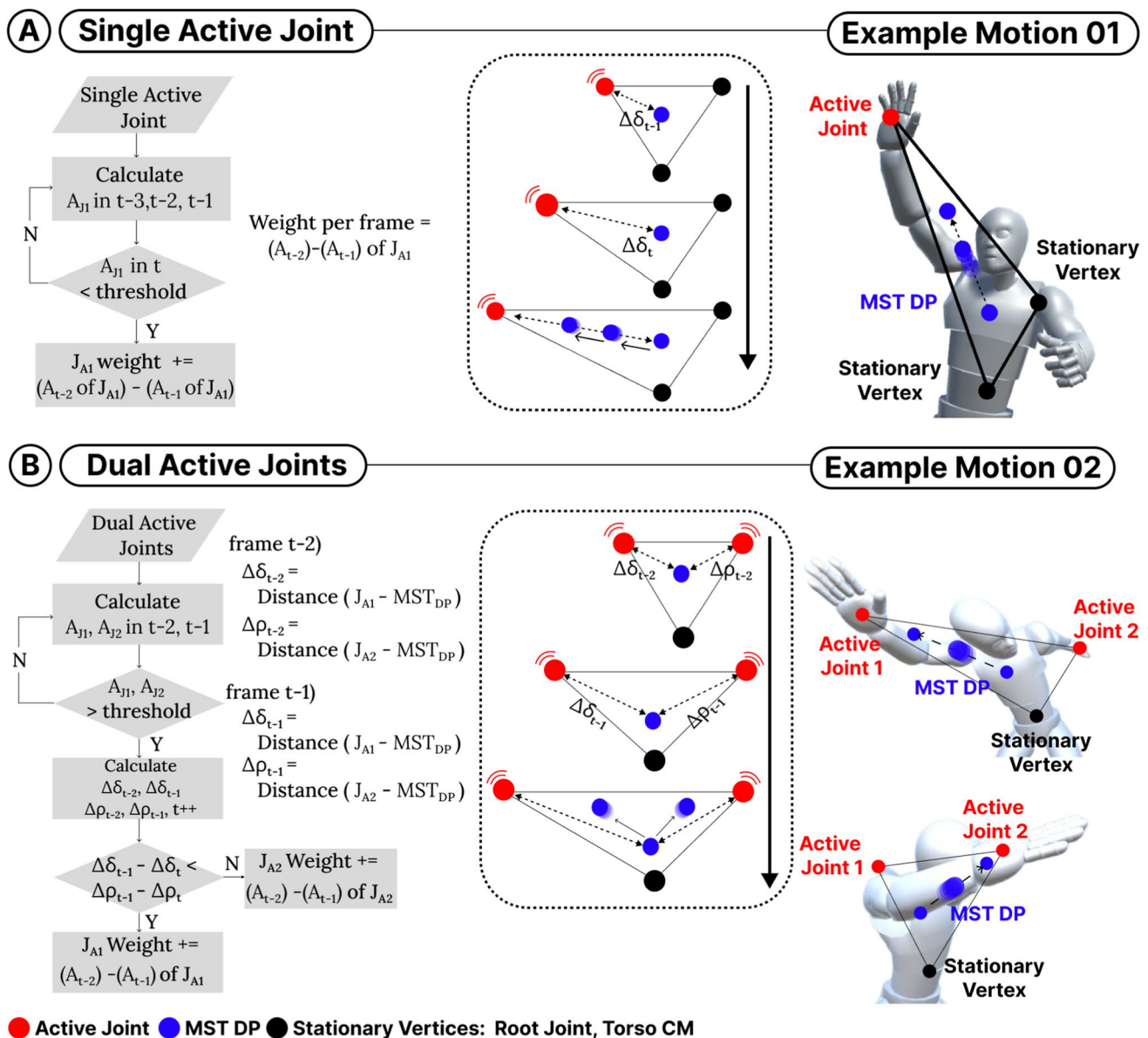
Fig. 9 Overall workflow and visualization of weight distributions for tactile translation. We consider both **A** single active joint and **B** dual active joint cases

$$I_t = \left(\alpha \cdot D_t \cdot C + (1 - \alpha) \cdot I_{t-1}\right) \qquad (2)$$

Equation 2 is based on an exponential filter that uses exponentially weighted averaging to produce an output value. Here, I_t, α, D_t refer to the total intensity value, smoothing factor, and distance measurement of two vertices accordingly. In our work, we set α as 0.5 where we distribute the same importance weight to the current frame (t) and the previous frame (t-1).

As mentioned previously, we adjust C to transfer the intended tactile intensity to the bHaptics vest and sleeve. Thoroughly stated in Sect. 4.3.1, we confirm bHaptics's level

6 intensity parameter is the most comfortable value Maereg et al. (2017). Thus, we set the value C as level 6.

4.3.3 Intensity distribution based on motion dynamics

As previously mentioned, $Distance(J_T$ - $\text{MST}_{DP})$ indicates the distance between MST_{DP} and J_T as shown in Fig. 10. We increase the intensity as the distance gets larger which intuitively conveys the performer's movement into a tactile experience. We also accommodate dynamic motions by controlling intensity level based on the active joint's (J_A) acceleration. The wholesome
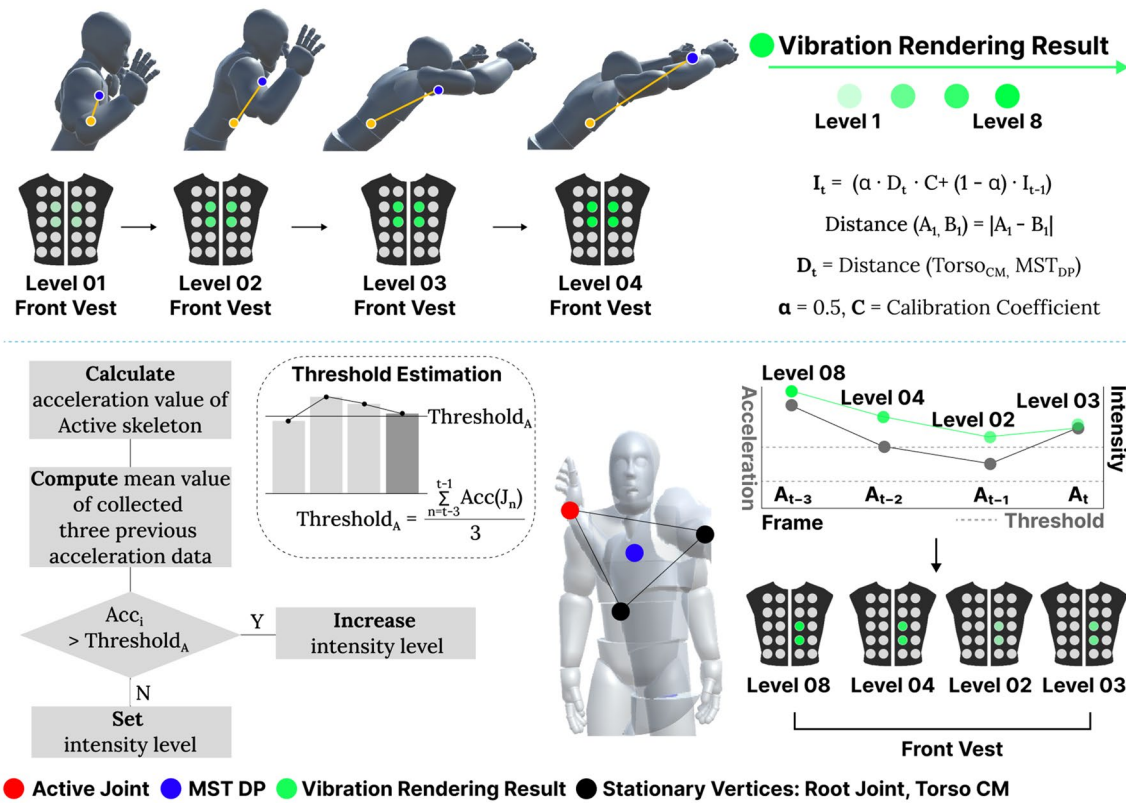
Fig. 10 Overall methods for intensity control. (Top) The intensity increases as the distance gets larger from the performer's body. (Bottom) We employ an acceleration threshold and use average accelera-tion to further control the tactile intensity. When the average accelera-tion is lower than the threshold, the intensity level decreases and vice versa

intensity is modified if the acceleration exceeds the set threshold which is the mean acceleration value from the recent three frames. If the acceleration of the current frame exceeds, maximum level 2 (1.23 G with 103 Hz) intensity is added regarding the minimum value of human noticeable intensity value Verrillo (1966). Therefore, the maximum intensity will be accounted for in intensity level 8.

In a particular example, regarding some communica-tive and choreographic motions, whose active joint(s) has the same speed but translates forward or backward in continuous frames, it is inevitable to render the different levels of intensity on a haptic display. As our demonstrat-ing vibrotactile intensity translation reflects the amount of displacement between two main element vertices, we guarantee to improve the different types of motions either that account for big translation but also small translation.

4.3.4 1D phantom and 2D grid-based tactile sensation

In order to convey subtly ruled intensity, we provide the phantom sensation inspired by Park and Choi (2018). When the ray cast hits the computed bounded area, which is settled in the width of length between adjacent nodes divided by ten units (Unity 3D), we add the supplement vibration intensity to the primarily computed intensity. The intensities for each node will be adjusted along with the divided units, gained by multiplication of value K, the distance between nodes, and the normalized portion of distance a_n.

Figure 11a implies the 1D phantom sensation in between two consecutive nodes. By tracking the destination of the ray based on the normalized distance portion, the closest node to the ray will be designated as the main node and regarded as the starting point with the coordinate of (0,0). If the ray hits near $Node_A$, the intensity level will increase with the computed value, $K \cdot (1 - \alpha_n)$ $Node_B$ gains $K \cdot \alpha_n$, while $Node_A$ gains $K \cdot (1 - \alpha_n)$.

In the case of 2D grid-based tactile sensation, Fig. 11b indicates cases when a ray hits among four adjoined
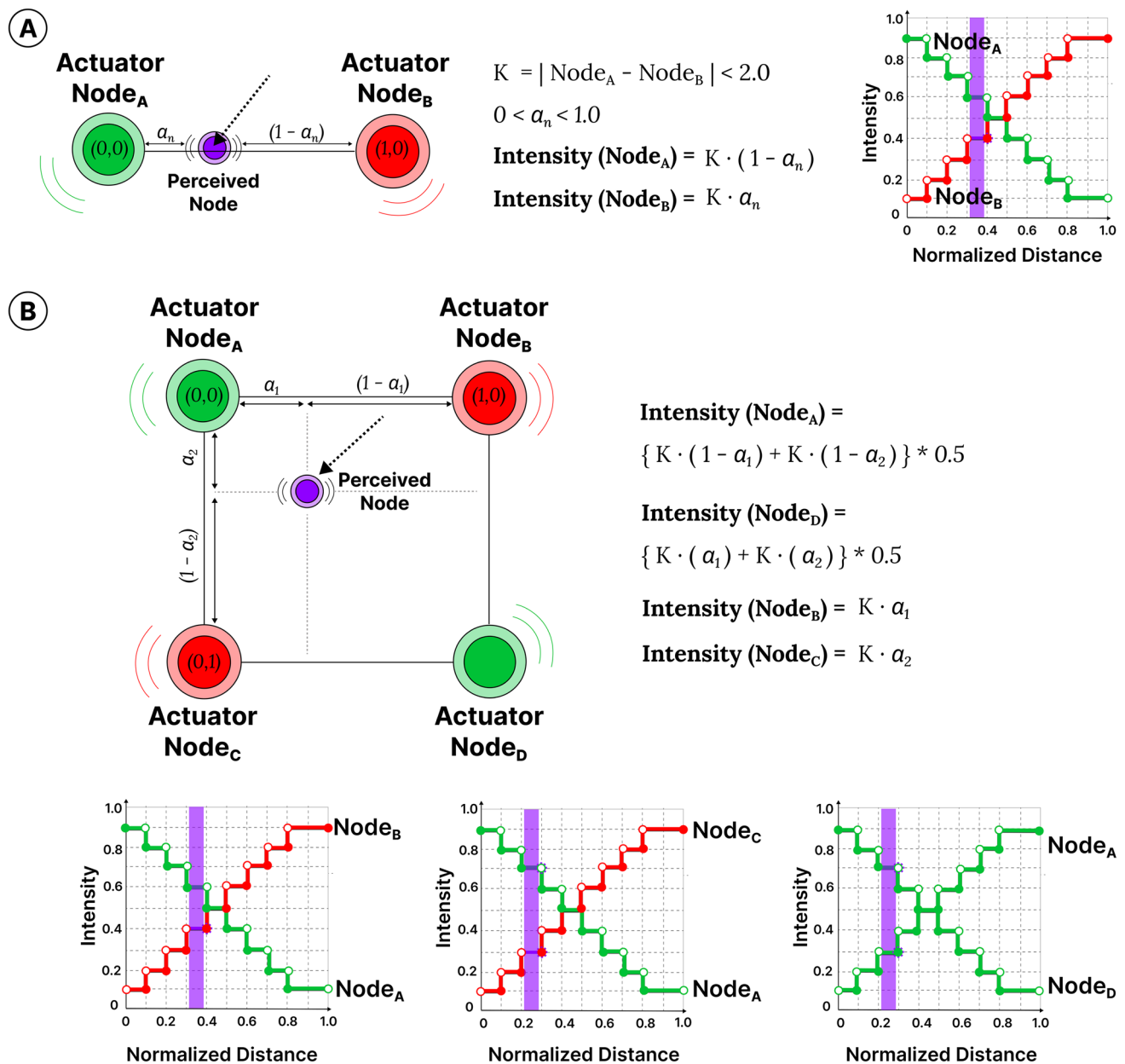
Fig. 11 Overall cases of node processing approach. **A** Case of node processing approach for 1D phantom tactile sensations and intensity plot for $Node_A$ and $Node_B$. **B** Case of node processing approach for 2D grid-based method and intensity plot for $Node_A$ and rest of the nodes. Here, the X- and Y-axis refer to normalized distance and intensity, respectively. The highlighted vertical line in plots refers to the illusory locations where the perceived node locates

nodes. This rendering rule is extended from the previously mentioned 1D phantom sensations. We distribute the computed intensity separately to four nodes located near the perceived node by following functions. In 2D grid-based sensation, the closest node to the destination of the ray is regarded as the main node ($Node_A$). We examine three correlation sets between the main node and supplementary nodes ($Node_B$, $Node_C$, $Node_D$). Following two nodes $Node_B$, $Node_C$ comply with the rule of 1D

phantom sensation. With respect to the intensity of $Node_D$, we averaged the distributed value of $Node_B$ and $Node_C$, which allows experiencing the continuity of transition in consideration of connected nodes. Regarding the intensity of $Node_A$, we arranged the multiplying coefficient value to a properly measured value, 0.5 in order to prevent the node from high-intensity saturation.

5 Hardware and software configuration

5.1 System overview

Fig. 12 demonstrates four main steps in overall hardware workflow. First, we collect 3D joint coordinates from the offline performer with Azure Kinect. Once a set of point cloud data is captured from offline performers, we convert these data into a virtual performer in Unity 3D. Then, we extract joint-based spatiotemporal parameters from the converted avatar, which are used as input to our MST-based algorithm to carry out the tactile translation. Lastly, we apply our real-time tactile translation to a proposed full upper-body wearable haptic interface covering the torso and shoulders.

5.2 System configuration

To convey the performer's upper-body motion to users, we configure our prototype with two different types of wearable haptic devices. We employ Tactsuit X40 and pairs of Tactosy from bHaptics (2019) where we place the Tactosy sleeves to each of shoulders as shown in Fig. 12. Every vibrotactile display module is wireless and battery-powered. The suit consisted of 40 individually controllable ERMs (20 ERMs on front and back) with a weight of 1.7 kg. For the sleeve, we alter its intended equipped location to the shoulder instead of the forearm. Each of haptic sleeves consist of 6 individually controllable ERMs with a weight of 0.32 kg. We used Oculus Quest 2 Inc (2020) as our VR platform. We utilize the bHaptics Unity plugin to assign vibration location and intensity from our algorithm.

6 User experience study

We carry out multiple user studies to validate our proposed MST-based algorithms with the hardware and software configuration shown in Sect. 5. We collect participants' subjective ratings through two different user studies. In the first study, we compare our pipeline with the baseline approach to confirm how our algorithm performs in translating discrete choreographic and communicative motion sequences. For the second study, we set up several virtual concert scenes and collected subjective ratings to compare various media-to-vibrotactile translation approaches and their combination (Fig. 13).

6.1 Study setup

We recruited 24 participants (11 male and 13 female) for the experiment with a mean age of 25 (ranging from 20 to 37). No participants reported any sensory disorders that could affect their auditory, vision, or haptic perception. All participants had experience using a head-mounted device (HMD), and 12 participants had experience with a haptic suit. All participants mentioned they had no basic knowledge about VR performance, so we thoroughly explained the concept of a virtual performance before carrying out the study. Researchers went through with participants on user study procedures and equipped them with HMD and customized wearable haptic hardware as shown in Fig. 12.

The training session was given to the participants before conducting an actual session. During the main session, the participants verbally answer in-VR questionnaires (Table 1) after experiencing each method. We collect subjective ratings on the given haptic rendering methods using a 7-point Likert scale (1=Strongly Disagree, 7=Strongly Agree). We give 10–15 s break after experiencing each method to
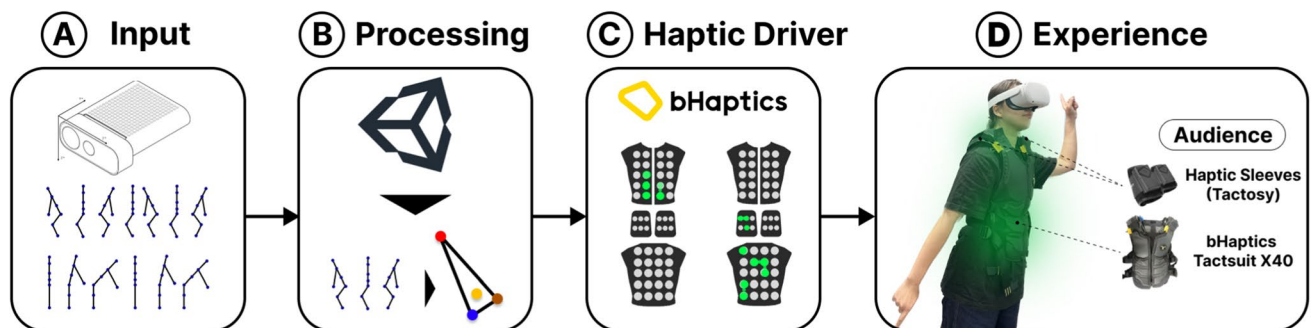

Fig. 12 System overview. **A** We collect 60 Hz 3D motion data from the virtual performer, **B** compute the joint-based spatiotemporal measurements, and **C** transfer the associated driving parameters to the haptic driver based on our proposed algorithm. **D** Users experience real-time tactile feedback translated from the performer's motions wearing a customized haptic vest along with haptic sleeves on both shoulders

prevent user adaptation and fatigue. The entire study took about 1.5–2 h.

6.2 Questionnaires

We devise questionnaires to investigate whether the proposed MST-based rendering pipeline enhanced the VR experience regarding naturalness, immersion, consistency, satisfaction, and embodiment. The higher ratings from these aspects would ensure the validity of our system. Since translating motion data into vibrotactile feedback in real time requires a direct response, which affects users' experience and satisfaction Lin et al. (2021), we added a questionnaire

for the latency. We ask users whether tactile feedback was translated on time along with the performer's motion (visual aid). Table 1 shows questionnaires used in both studies, and we slightly modify words to better represent each study context.

6.3 Study 1: motion-to-tactile framework performance

In this study, we compare two different haptic rendering approaches shown in Fig. 14. We select six motion sequences consisting of four choreographic motions (*Side to side hip-hop dance, Forward to Back hip-hop dance,*

Fig. 13 User study setup. **A** Participants experience the virtual concert using the given HMD. **B** Wearing the customized upper-body haptic device. **C** We carry out in-VR questionnaires to collect subjective ratings
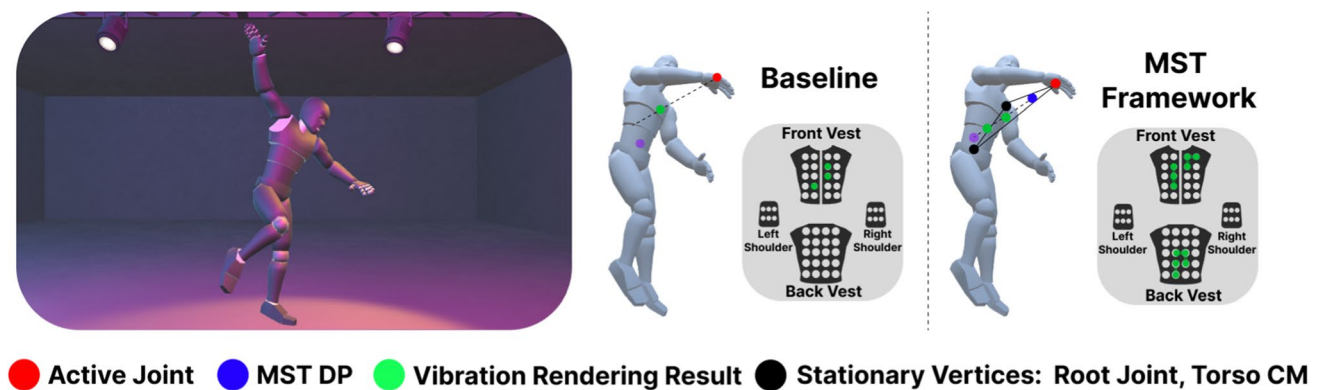

Fig. 14 (Right) Scene of user HMD. (Left) We show two rendering results of baseline and MST-based pipeline

Table 1 Questionnaires for user study 1 and 2

Naturalness	I was able to relate the wearable vibrations to the actions of the performer (context of the performance)
Immersion	The haptic effects make me more immersed in the movement of the performer (into the short scene of the concert)
Satisfaction	I enjoyed the experience of vibration sensations while watching the performer's movement (this clip of the concert)
Consistency	I feel haptic feedback is well matching to the performer's motion
Latency	To what extent did you feel the latency between the performer's movement and the haptic feedback?
Embodiment	The haptic effects increase feeling as is my body was located where the virtual performer is

Diagonally side to side Jazz motion, and Waving side to side motion) and two communicative motions (*Waving toward audiences*(motion 1), *Throwing a ball/mic toward audiences*(motion 4) from Fig. 4. Each motion lasts for about 20~25 s. A total of 12 combinations (2 haptic rendering methods × 6 motions) were tested. The presentation order of the motion conditions was randomized for each participant, and that of the rendering approaches was randomized within all motion conditions.

The baseline method directly maps the active joint to the haptic feedback. This approach is a widely used and conventional approach Schneider et al. (2015) where the haptic feedback is rendered based on a key factor like active joint information. To solely compare the performance difference, we use the same warping algorithm (Fig. 7) and hardware configuration for both methods. The only difference is that "Baseline" method uses the active joint while the "MST-based pipeline" employs MST_{DP} to deliver haptic experiences.

Results and Discussion We conduct a two-way within-subjects analysis of variance (ANOVA) first and carry out one-way ANOVA for the cases presenting meaningful interaction effects. Then, we run Tuckey's HSD test for each of the 6 subjective ratings to confirm the effects of the haptic rendering methods and their significance.

We look at the effects of haptic rendering methods with the 6 discrete motions by analyzing two-way ANOVA results (Table 2). Excluding "latency," different haptic rendering methods show statistically significant main effects with $p < 0.001$. The results indicate that the haptic rendering methods largely affect the user experience (effect size η^2 near 0.05). For naturalness (5.74 vs. 3.55), immersion (5.77 vs. 3.38), satisfaction (5.57 vs. 3.52), consistency (5.8 vs. 3.24), and embodiment (4.84 vs. 3.68), MST rendering pipeline shows a much higher score than the baseline. This result indicates our algorithm successfully translates choreographic and communicative motions into vibrotactile feedback. For the motion with a large range of motions (e.g., Motion 5), we notice higher consistency, satisfaction, naturalness, and immersion level.

In terms of "latency," a lower score indicates better performance meaning there was no latency for the vibrotactile feedback. Both haptic rendering methods present fast and responsive tactile stimuli to users, and we found no significant differences in latency for main and interaction effects. Therefore, we continued our statistical analysis excluding this subjective area.

Figure 15 shows the average Likert score of the 6 subjective areas from participants. In all motions, we observe that the average rating for our algorithm is superb to the baseline approach in general. To statistically confirm the validity of these observations, we perform a one-way ANOVA to assess the performance and effectiveness of the

Table 2 Two-way ANOVA results for six subjective ratings

Measure	Factor	Statistics	Effect size
Naturalness	Motions	$F(5,115) = 1.2205, p = 0.304$	0.009
	Methods	$F(1,23) = 14.195, p < 0.001$***	0.469
	Interaction	$F(5,115) = 4.4299, p < 0.001$***	0.041
Immersion	Motions	$F(5,115) = 2.4616, p = 0.037$*	0.019
	Methods	$F(1,23) = 14.195, p < 0.001$***	0.434
	Interaction	$F(5,115) = 11.82, p < 0.001$***	0.089
Satisfaction	Motions	$F(5,115) = 2.1516, p = 0.065$	0.023
	Methods	$F(1,23) = 14.195, p < 0.001$***	0.521
	Interaction	$F(5,115) = 13.64, p < 0.001$***	0.095
Consistency	Motions	$F(5,115) = 1.5363, p = 0.184$	0.011
	Methods	$F(1,23) = 14.195, p < 0.001$***	0.458
	Interaction	$F(5,115) = 7.56, p < 0.001$***	0.064
Latency	Motions	$F(5,115) = 2.19, p = 0.060$	0.029
	Methods	$F(1,23) = 3.675, p = 0.068$	0.006
	Interaction	$F(5,115) = 1.844, p = 0.111$	0.015
Embodiment	Motions	$F(5,115) = 3.65, p = 0.004$***	0.045
	Methods	$F(1,23) = 3.675, p = 0.068$	0.423
	Interaction	$F(5,115) = 2.2294, p < 0.056$[a]	0.015

Each row shows the effect size and *p* values according to each factor, including motions, methods, and interaction between motions and methods. Both significant ratings and marginally significant ratings are in bold

*,**,*** = *p* value <0.05, 0.01 and 0.001. [a] = Marginally significant rating

proposed MST-based algorithm compared to the "Baseline" method.

According to Table 3, for "Motion 2–6," overall user experience with the MST-based algorithm comes out to be superior to the baseline, which results in statistically significant main effects ($p < 0.05$) on most of the subjective ratings in Table 3. For "Motion 1," we only see a statistically significant effect on naturalness. We notice that the absolute magnitude of the Likert score for "Baseline" is particularly high compared to other motions since *Waving* or *Inducing shouting* included in this communicative motion is hard to produce effective haptic feedback experiences with an active joint or MST_{DP}. Still, our algorithm shows better ratings in general.
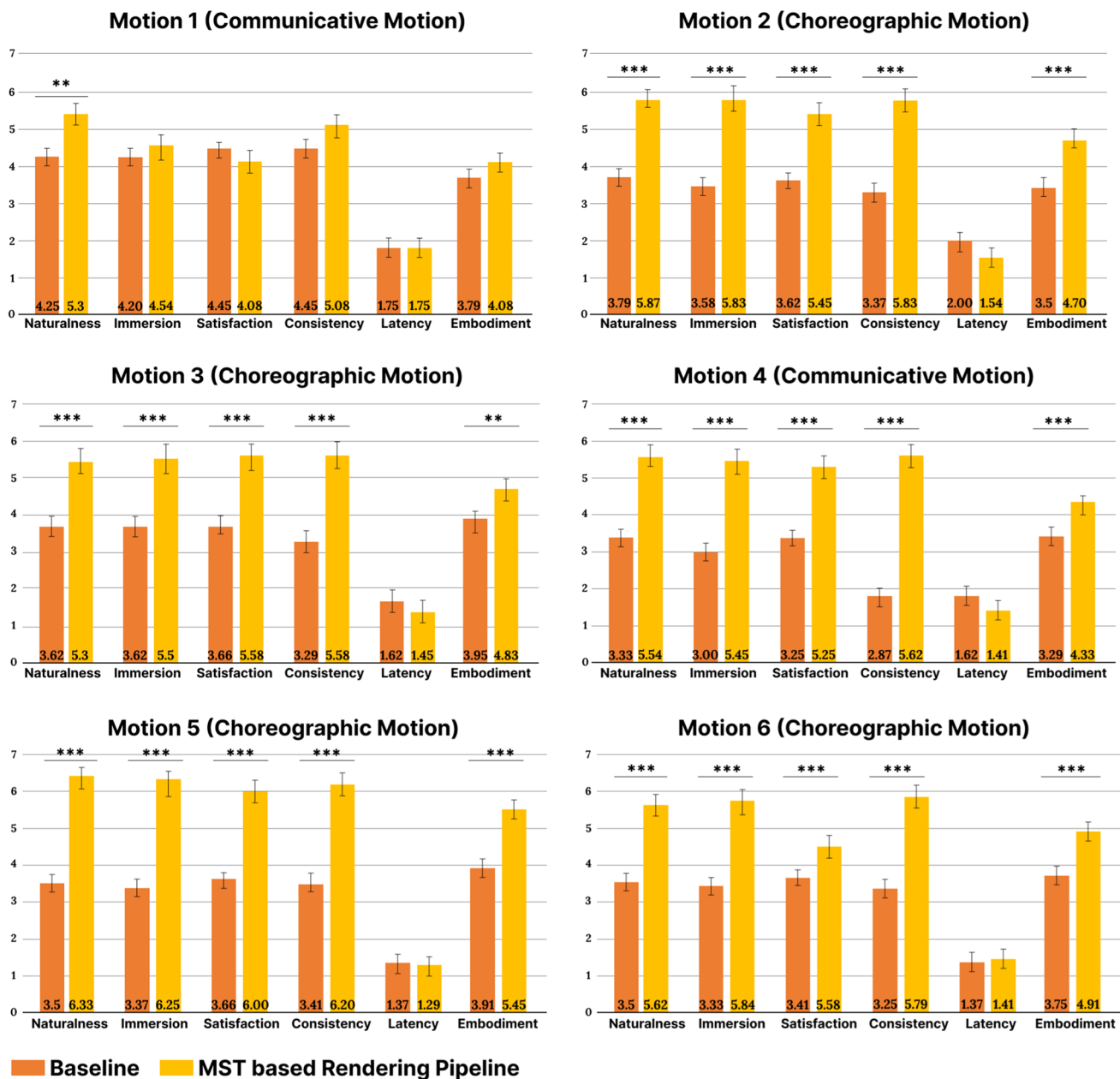
Fig. 15 The user experience results for given haptic feedback in terms of naturalness, immersion, satisfaction, consistency, latency, and embodiment. The error bars represent standard errors. **,*** = p value < 0.01, and 0.001

6.4 Study 2: tactile translation preference

Our proposed motion-to-tactile framework showed promising results for translating a virtual performer's motion into effective vibrotactile feedback. We further investigate the holistic user experience using our framework along with or without a conventional audio-to-tactile approach. We would like to find out subjective ratings and user preference under three different conditions: (1) *motion-to-tactile*, (2) *audio-to-tactile*, and (3) *audio and motion (multimodal)-to-tactile* (Fig. 16).

For *motion-to-tactile* condition, we employ an MST-based algorithm from Sect. 4. For *audio-to-tactile* translation, we utilize the audio-to-haptics feature from bHaptics (2019) which provides several audio-to-tactile themes with varying frequencies. We choose *POP* theme which supports 80–90 Hz, as this frequency range effectively conveys a variety of audio cues, such as sudden changes in pitch and rhythm. Lastly, we test a combination of *audio-to-tactile* and *motion-to-tactile* translation methods. Here, we set the intensity ratio of audio-to-motion as 2 to 3 since it provides adequate and comfortable feedback. We notice that increasing

the audio intensity ratio generally overwhelms the whole sensation which is not desirable for a balanced mixture.

We select motions for each study scene from Sect. 3.2. To appropriately distribute choreographic and communicative motions that align with the audio context, we refer to the actual performance of each selected song (Table 4). As

a result, user study scenes consist of approximately 80% choreographic motions and 20% communicative motions.

In summary, we empirically test 9 combinations (3 haptic translation methods × 3 scenes) for multimedia rendering conditions in virtual concerts (Table 4). The presentation order of the three scenes was randomized for each

Table 3 One-way ANOVA results for each motion

	Naturalness	Immersion	Satisfaction	Consistency	Embodiment
Motion 1	7.6496	0.91013	0.32354	2.3491	1.5138
	$p = 0.011*$	$p = 0.350$	$p = 0.575$	$p = 0.139$	$p = 0.231$
	$\eta_2 = 0.133$	$\eta_2 = 0.015$	$\eta_2 = 0.004$	$\eta_2 = 0.040$	$\eta_2 = 0.010$
Motion 2	14.195	14.195	14.195	14.195	14.195
	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$
	$\eta_2 = 0.646$	$\eta_2 = 0.707$	$\eta_2 = 0.577$	$\eta_2 = 0.744$	$\eta_2 = 0.168$
Motion 3	14.195	14.195	14.195	14.195	10.226
	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$	$p = 0.004**$
	$\eta_2 = 0.591$	$\eta_2 = 0.503$	$\eta_2 = 0.407$	$\eta_2 = 0.585$	$\eta_2 = 0.098$
Motion 4	14.195	14.195	14.195	14.195	14.195
	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$
	$\eta_2 = 0.637$	$\eta_2 = 0.708$	$\eta_2 = 0.565$	$\eta_2 = 0.146$	$\eta_2 = 0.591$
Motion 5	14.195	14.195	14.195	14.195	14.195
	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$
	$\eta_2 = 0.779$	$\eta_2 = 0.792$	$\eta_2 = 0.676$	$\eta_2 = 0.735$	$\eta_2 = 0.338$
Motion 6	14.195	14.195	14.195	14.195	14.195
	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$	$p < .001***$
	$\eta_2 = 0.647$	$\eta_2 = 0.723$	$\eta_2 = 0.587$	$\eta_2 = 0.718$	$\eta_2 = 0.150$

The first row of each motion represents $F(1,23)$ value, and the second row with $p < 0.001$ or $p < 0.05$ indicates there is a significant effect

Both significant ratings and marginally significant ratings are in bold. *,**,*** = p value <0.05, 0.01 and 0.001
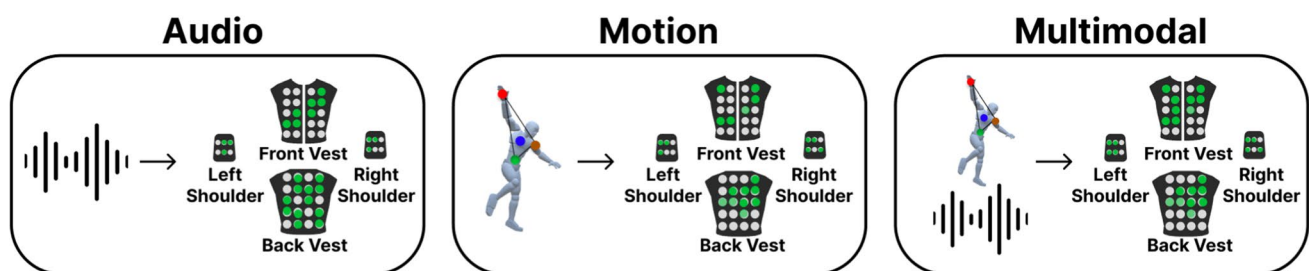

Fig. 16 We confirm user experience in VR performance with (Left) audio-to-tactile, (Middle) motion-to-tactile, and (Right) audio and motion-to-tactile translations

Table 4 Detail VR performance scene information for Study 2

Scene	Title—musician (Year)	Duration (s)	BPM	Motion components
Scene 1	Happy-Pharrell Williams (2015)	41 s	160	Choreographic Motion 82% (33.62s) Communicative Motion 18% (7.38s)
Scene 2	Levitating-Dua Lipa (2022)	38 s	103	Choreographic Motion 88% (33.44s) Communicative Motion 18% (4.56s)
Scene 3	Blurred Lines = Robin Thicke (2013)	42 s	120	Choreographic Motion 90% (37.8s) Communicative Motion 10% (4.2s)

Table 5 Two-way ANOVA results for 3 different media-to-tactile translation methods for each scene's dependent variables: naturalness, immersion, satisfaction, consistency, latency, and embodiment

Measure	Factor	Statistics	Effect size
Naturalness	Scene	$F(2,46) = 1.07$, $p = 0.352$	0.005
	Methods	$F(2,46) = 23.12$, $p < 0.001$***	0.155
	Interaction	$F(4,92) = 1.51$, $p = 0.204$	0.018
Immersion	Scene	$F(2,46) = 1.64$, $p = 0.205$	0.010
	Methods	$F(2,46) = 19.26$, $p < 0.001$***	0.150
	Interaction	$F(4,92) = 1.10$, $p = 0.363$	0.013
Satisfaction	Scene	$F(2,46) = 1.69$, $p = 0.196$	0.008
	Methods	$F(2,46) = 19.82$, $p < 0.001$***	0.164
	Interaction	$F(4,92) = 1.01$, $p = 0.408$	0.012
Consistency	Scene	$F(2,46) = 4.204$, $p = 0.021$*	0.020
	Methods	$F(2,46) = 18.50$, $p < 0.001$***	0.163
	Interaction	$F(4,92) = 0.820$, $p = 0.516$	0.010
Latency	Scene	$F(2,46) = 2.56$, $p = 0.088$	0.019
	Methods	$F(2,46) = 5.18$, $p = 0.009$**	0.020
	Interaction	$F(4,92) = 2.98$, $p = 0.023$*	0.022
Embodiment	Scene	$F(2,46) = 1.07$, $p = 0.353$	0.005
	Methods	$F(2,46) = 6.62$, $p = 0.003$**	0.055
	Interaction	$F(4,92) = 1.63$, $p = 0.173$	0.015

Significant ratings are in bold. *,**,*** = p value $p < 0.05$, 0.01 and 0.001

participant, and that of the rending modality methods was randomized within all scenes.

Results and Discussion We conduct a two-way ANOVA on each subjective rating.

As shown in Table 5, different media-to-translation methods show a statistically significant effect in all subjective ratings ($p < 0.01$ or $p < 0.001$). We observe a main effect of scene on consistency, but the participants reported high consistency scores across all scenes (4.90, 4.48, 4.91 for Scenes 1-3). The study did not find significant interaction effects between media-to-tactile methods and scene types except for latency. The rating for latency was 2.01 in the audio-to-tactile approach, which was the largest value among other approaches. We conduct post hoc analysis with Tukey's Honest Significant Difference (HSD) test to further investigate the effects of various media-to-tactile approaches.

Figure 17 illustrates the post hoc test results with the grouping labels where the different letter indicates significantly different groups. We observe that the proposed *motion-to-tactile* ($\mu=5.38$, SD=0.12) method receives more positive feedback from users than *audio-to-tactile* ($\mu=4.09$, SD=0.17) or *multimodal-to-tactile* ($\mu=4.76$, SD=0.10). The results clearly show that appending tactile feedback created from performer motions induces an immersive VR performance experience.

Moreover, we directly ask users to rank their preferred media-to-tactile translation methods based on their experiences from the study. We also obtain verbal feedback to fully understand users' opinions. Our results show that users prefer *motion-to-tactile* (58.4%) technique compared
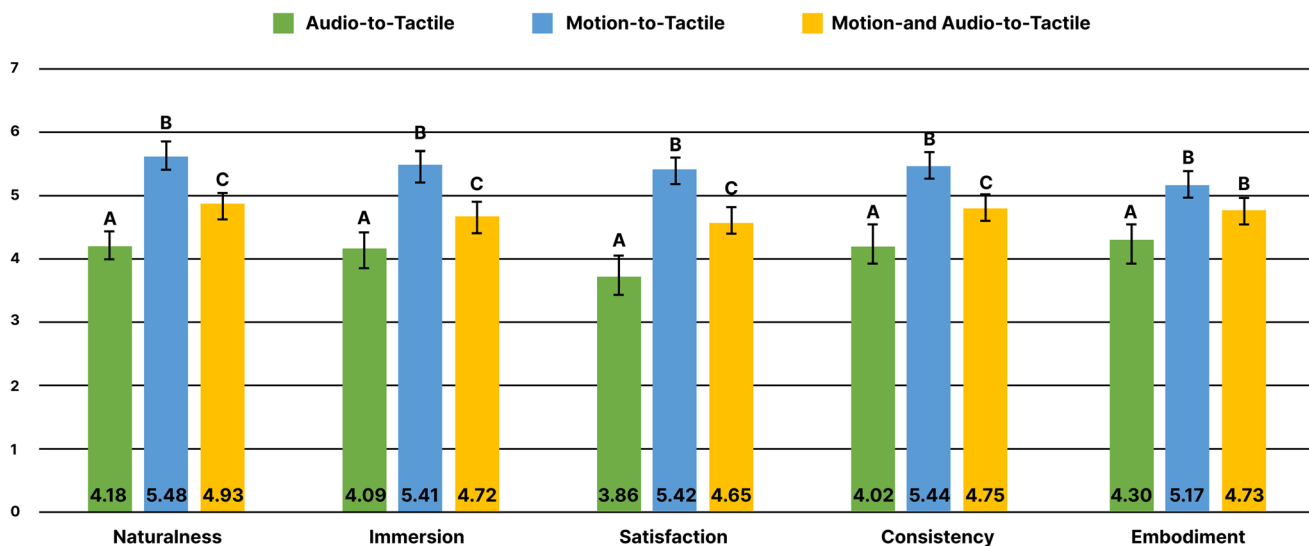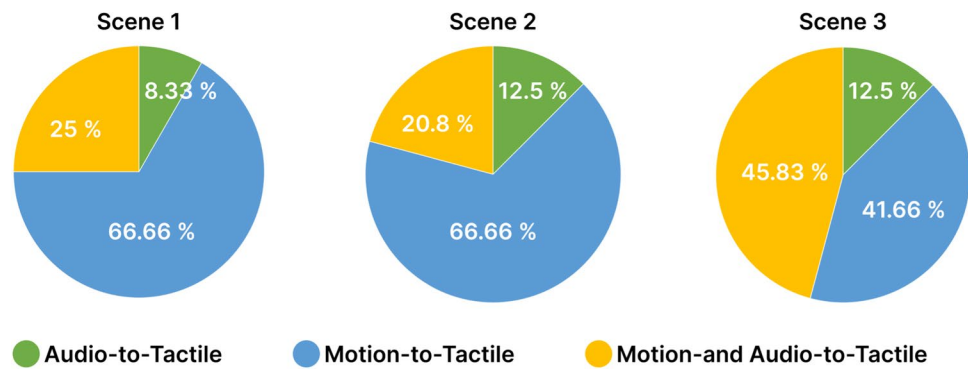

Fig. 17 Result of Tukey's HSD test for subjective ratings (naturalness, immersion, satisfaction, consistency, and embodiment) with various media-to-tactile translation methods. Error bars represent standard errors. The conditions grouped with the same alphabet did not show statistically significant differences

Fig. 18 Preference from users on 3 different tactile translation methods for user experience

to *audio-to-tactile* (11.11%), and *multimodal-to-tactile* (30.54%) approaches.

As shown in Fig. 18, we observe users prefer *motion-to-tactile* and *multimodal-to-tactile* over *audio-to-tactile*. Surprisingly, the users choose haptic feedback created from performer motions computed with MST algorithms as the most preferred experience. This tells that users favor connecting haptic feedback toward visual cues (virtual performer's motion). However, users prefer both *multimodal-to-tactile* and *motion-to-tactile* in Scene 3. Unlike other scenes where communicative motions are given regardless of audio context (e.g., maintain same audio for hand waving), we design Scene 3 with the balanced allocation of communicative motions for the given audio context (e.g., slow tempo/low volume for hand waving). The results tell that reflecting motions from performers, whether it is aligned with the audio context or not, still improves the user experience for VR performance over the conventional audio-to-tactile method.

We notice that the users least preferred the sole *audio-to-tactile* approach. One participant (P6) mentioned that it is always better to have a haptic feedback feature while watching the virtual performance. However, P6 denoted that it was hard to extract any context from a given haptic experience. Another participant (P2) commented that the MST-based algorithm felt more immersive since choreographic motions and given audio fully express the performance context through vibrotactile feedback. This implies that users consider "Motion" as a key factor for tactile translation. From Fig. 18, the results show that "Motion" integrated tactile translation methods (*motion-to-tactile* and *multimodal-to-tactile*) were preferred by more than 90% of users. Thus, we believe that the proposed tactile translation method shows great potential in facilitating immersive experiences for virtual performance.

7 General discussion and conclusion

We propose MST-based algorithms that provide contextual motion-to-tactile translation and enable sophisticated real-time haptic experiences. Throughout the user studies, participants report that the haptic experience with VR performance is consistent and well designed to support VR concerts. Discussing our design considerations, we identify several design guidelines and challenges along with future works.

Multimodal-to-tactile Framework From the study result, the *motion-to-tactile* approach is preferred the most by users. However, we observe that users prefer *multimodal-to-tactile* for the scene with carefully designed motion allocation with a given audio context. Since most VR performance scenes would be designed with careful motion allocation along with audio context, we encourage utilizing the *multimodal-to-tactile* approach.

Embodiment for VR performance The main objective of our work is to provide an immersive VR performance experience by reflecting the performer's motion translated to the users. In our studies, we observe that users feel like dancing even in a seated position when applying *motion-to-tactile* translation. Moreover, our approach further enhances the bond between the VR performer and the user. Aligned with previous research Thériault et al. (2021), we observe that users perceive the same motion as the virtual character using our method. This tells that our method has the potential in providing a sense of presence and realism as well as increasing the bond with the remote performer in VR performance.

Future Work In this work, we mainly focus on a virtual solo performer. The current MST-based algorithm is hard to reflect multiple performers' motions into representative haptic feedback. Therefore, for future work, it is in our interest to find out an effective method to exert meaningful haptic feedback from multiple performers. Potential solutions would be tracking user's attentions (e.g., eye-tracking and head-tracking) to efficiently reflect the overall haptic experience.

We foresee that the MST algorithm's flexibility lends itself to the expansion of other types of promising performance platforms such as augmented (AR) and mixed reality (MR). For instance, AR/MR can enhance the live concert experience by adding digital overlays to the real-world environment. It is possible to correlate our tactile translation with the digital contents overlaid live performance. Overall, the use of tactile feedback in live concerts opens up a new opportunity to support engagement and creativity and enhance the cultural experience for both performers and fans.

We present a novel media-to-tactile translation method based on an MST-based framework. This translates the performer's choreographic as well as communicative motions into meaningful vibrotactile feedback. We also customize an upper-body wearable haptic interface to provide full 3D haptic feedback to reflect the performer's various motions with appropriate haptic feedback. Through user studies, we confirm the proposed algorithm's performance over the conventional approach in terms of subjective ratings and user preference. Our work will enable an immersive VR performance experience by proposing a novel motion-to-tactile framework reflecting contextual information about the performance.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval The authors declare that they received consent from all the participants before participating in the study. This work was approved by the Institutional Review Board of KAIST (Korea Advanced Institute of Science and Technology) (IRB number: KH2022-105)

References

Abdur Rahman M, Alkhaldi A, Cha J, El Saddik A (2010) Adding haptic feature to youtube. In: Proceedings of the 18th ACM international conference on multimedia, pp 1643–1646. https://doi.org/10.1145/1873951.1874310

Abe M, Akiyoshi T, Butaslac I, Hangyu Z, Sawabe T (2022) Hype live: biometric-based sensory feedback for improving the sense of unity in vr live performance. In: 2022 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW), pp 836–837. https://doi.org/10.1109/VRW55335.2022.00269. IEEE

Adolphe M, Clerval J, Kirchof Z, Lacombe-Delpech R, Zagrodny B (2017) Center of mass of human's body segments. Mech Mech Eng 21(3):485–497

Alamri A, Kim H-N, Cha J, El Saddik A (2010) Serious games for rehabilitation of stroke patients with vibrotactile feedback. Int J Comput Sci Sport (Int Assoc Comput Sci Sport) 9(2):52–64

Altinsoy ME, Merchel S (2010) Cross-modal frequency matching: Sound and whole-body vibration. In: International workshop on haptic and audio interaction design, pp 37–45. https://doi.org/10.1007/978-3-642-15841-4_5. Springer

Berdahl E, Steiner H-C, Oldham C (2008) Practical hardware and algorithms for creating haptic musical instruments. In: NIME, vol 8, pp 61–66

Bieber J Wave Presents: Justin Bieber - an interactive virtual experience. https://www.youtube.com/watch?v=UAhGvhvcoyY (2021)

Billboard: Billboard Hot 100. https://www.billboard.com/charts/hot-100/ (2022)

Bisig D (2022) Generative dance-a taxonomy and survey. In: Proceedings of the 8th international conference on movement and computing, pp. 1–10. https://doi.org/10.1145/3537972.3537978

Blumenfeld-Jones D (2008) Dance, choreography, and social science research. Handbook of the arts in qualitative research: perspectives, methodologies, examples, and issues, pp 175–184

Botvinick M, Cohen J (1998) Rubber hands 'feel' touch that eyes see. Nature 391(6669):756–756. https://doi.org/10.1038/35784

bHaptics: Tactsuit, full body haptic suit for VR - Tactsuit. https://www.bhaptics.com/tactsuit/ (2019)

Chang A, O'Sullivan C (2005) Audio-haptic feedback in mobile phones. In: CHI'05 extended abstracts on human factors in computing systems, pp 1264–1267. https://doi.org/10.1145/1056808.1056892

Charron J-P (2017) Music audiences 30: Concert-goers' psychological motivations at the dawn of virtual reality. Front Psychol. https://doi.org/10.3389/fpsyg.2017.00800

Cuartielles D, Göransson A, Olsson T, Stenslie S (2012) Developing visual editors for high-resolution haptic patterns. In: The Seventh international workshop on haptic and audio interaction design, Lund, Sweden, pp 42–45

Danieau F, Lécuyer A, Guillotel P, Fleureau J, Mollet N, Christie M (2012) Enhancing audiovisual experience with haptic feedback: a survey on HAV. IEEE Trans Haptics 6(2):193–205. https://doi.org/10.1109/TOH.2012.70

Danieau F, Fleureau J, Guillotel P, Mollet N, Lécuyer A, Christie M (2012) Hapseat: producing motion sensation with multiple force-feedback devices embedded in a seat. In: Proceedings of the 18th ACM symposium on virtual reality software and technology, pp 69–76. https://doi.org/10.1145/2407336.2407350

Degraen D, Fruchard B, Smolders F, Potetsianakis E, Güngör S, Krüger A, Steimle J (2021) Weirding haptics: In-situ prototyping of vibrotactile feedback in virtual reality through vocalization. In: The 34th Annual ACM symposium on user interface software and technology, pp 936–953. https://doi.org/10.1145/3472749.3474797

Emotionwave XR: We are creating new musical experiences in an extended reality space. https://emotionwave.com/xr/

Enriquez K, Palacios M, Pallo D, Guerrero G (2020) Sense: sensory component VR application for hearing impaired people to enhance the music experience. In: 2020 15th Iberian conference on information systems and technologies (CISTI), pp 1–6. https://doi.org/10.23919/CISTI49556.2020.9140447. IEEE

Fang CM, Harrison C (2021) Retargeted self-haptics for increased immersion in vr without instrumentation. In: The 34th Annual

ACM symposium on user interface software and technology, pp 1109–1121. https://doi.org/10.1145/3472749.3474810

Fontana F, Camponogara I, Cesari P, Vallicella M, Ruzzenente M An exploration on whole-body and foot-based vibrotactile sensitivity to melodic consonance (2016). 11390/1095380

García-Valle G, Arranz-Paraíso S, Serrano-Pedraza I, Ferre M (2020) Estimation of torso vibrotactile thresholds using eccentric rotating mass motors. IEEE Trans Haptics 14(3):538–550. https://doi.org/10.1109/TOH.2020.3048290

Golomer E, Toussaint Y, Bouillette A, Keller J (2009) Spontaneous whole body rotations and classical dance expertise: how shoulder-hip coordination influences supporting leg displacements. J Electromyogr Kinesiol 19(2):314–321. https://doi.org/10.1016/j.jelekin.2007.08.004

Gonzalez-Franco M, Peck TC (2018) Avatar embodiment towards a standardized questionnaire. Front Robot AI. https://doi.org/10.3389/frobt.2018.00074

Hashizume S, Sakamoto S, Suzuki K, Ochiai Y (2018) Livejacket Wearable music experience device with multiple speakers. In: International conference on distributed, ambient, and pervasive interactions, pp 359–371. https://doi.org/10.1007/978-3-319-91125-0_30. Springer

Hayes L (2015) Skin music (2012) an audio-haptic composition for ears and body. In: Proceedings of the 2015 ACM SIGCHI conference on creativity and cognition, pp 359–360. https://doi.org/10.1145/2757226.2757370

Hwang, I., Son, H., Kim, J.R (2017) Airpiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback. In: 2017 IEEE world haptics conference (WHC), pp 213–218. https://doi.org/10.1109/WHC.2017.7989903. IEEE

Imschloss M, Kuehnl C (2019) Feel the music! exploring the crossmodal correspondence between music and haptic perceptions of softness. J Retail 95(4):158–169. https://doi.org/10.1016/j.jretai.2019.10.004

Inc, F.: Oculus Quest 2 (2020). http://www.oculus.com/quest2

Israr A, Zhao S, Schwalje K, Klatzky R, Lehman J (2014) Feel effects: enriching storytelling with haptic feedback. ACM Trans Appl Percep (TAP) 11(3):1–17. https://doi.org/10.1145/2641570

Israr A, Poupyrev I (2011) Control space of apparent haptic motion. In: 2011 IEEE world haptics conference, pp 457–462. https://doi.org/10.1109/WHC.2011.5945529. IEEE

Israr A, Zhao S, McIntosh K, Schwemler Z, Fritz A, Mars J, Bedford J, Frisson C, Huerta I, Kosek M et al. (2016) Stereohaptics: a haptic interaction toolkit for tangible virtual experiences. In: ACM SIGGRAPH 2016 studio, pp 1–57. https://doi.org/10.1145/2929484.2970273

Kaneko T, Tarumi H, Kataoka K, Kubochi Y, Yamashita D, Nakai T, Yamaguchi R (2018) Supporting the sense of unity between remote audiences in vr-based remote live music support system ksa2. In: 2018 IEEE international conference on artificial intelligence and virtual reality (AIVR), pp 124–127. https://doi.org/10.1109/AIVR.2018.00025. IEEE

Karageorghis CI, Jones L, Priest D-L, Akers RI, Clarke A, Perry JM, Reddick BT, Bishop DT, Lim HB (2011) Revisiting the relationship between exercise heart rate and music tempo preference. Res Q Exerc Sport 82(2):274–284. https://doi.org/10.1080/02701367.2011.10599755

Karam M, Branje C, Nespoli G, Thompson N, Russo FA, Fels DI (2010) The emoti-chair: an interactive tactile music exhibit. In: CHI'10 extended abstracts on human factors in computing systems, pp 3069–3074. https://doi.org/10.1145/1753846.1753919

Kim Y, Cha J, Ryu J, Oakley I (2010) A tactile glove design and authoring system for immersive multimedia. IEEE MultiMed 17(3):34–45. https://doi.org/10.1109/MMUL.2010.5692181

Kim M, Lee S, Choi S (2013) Saliency-driven real-time video-to-tactile translation. IEEE Trans Haptics 7(3):394–404. https://doi.org/10.1109/TOH.2013.58

Kruijff E, Marquardt A, Trepkowski C, Schild J, Hinkenjann A (2017) Designed emotions: challenges and potential methodologies for improving multisensory cues to enhance user engagement in immersive systems. Vis Comput 33(4):471–488. https://doi.org/10.1007/s00371-016-1294-0

Lalioti V, Ppali S, Thomas AJ, Hrafnkelsson R, Grierson M, Ang CS, Wohl BS, Covaci A (2021) Vr rehearse & perform-a platform for rehearsing in virtual reality. In: Proceedings of the 27th ACM symposium on virtual reality software and technology, pp 1–3. https://doi.org/10.1145/3489849.3489896

Li Y, Yoo Y, Weill-Duflos A, Cooperstock J (2021) Towards context-aware automatic haptic effect generation for home theatre environments. In: Proceedings of the 27th ACM symposium on virtual reality software and technology, pp 1–11. https://doi.org/10.1145/3489849.3489887

Lin Y-H, Wang Y-W, Ku P-S, Cheng Y-T, Hsu Y-C, Tsai C-Y, Chen MY (2021) Hapticseer: a multi-channel, black-box, platform-agnostic approach to detecting video game events for real-time haptic feedback. In: Proceedings of the 2021 CHI conference on human factors in computing systems, pp 1–14. https://doi.org/10.1145/3411764.3445254

MacLean KE, Schneider OS, Seifi H (2017) Multisensory haptic interactions: understanding the sense and designing for it. The handbook of multimodal-multisensor interfaces: foundations, user modeling, and common modality combinations 1:97–142. https://doi.org/10.1145/3015783.3015788

Maereg AT, Nagar A, Reid D, Secco EL (2017) Wearable vibrotactile haptic device for stiffness discrimination during virtual interactions. Front Robot AI 4:42. https://doi.org/10.3389/frobt.2017.00042

Mazzoni A, Bryan-Kinns N (2016) Mood glove: a haptic wearable prototype system to enhance mood music in film. Entertain Comput 17:9–17. https://doi.org/10.1016/j.entcom.2016.06.002

Melo M, Gonçalves G, Monteiro P, Coelho H, Vasconcelos-Raposo J, Bessa M (2020) Do multisensory stimuli benefit the virtual reality experience? a systematic review. IEEE Trans Vis Comput Graph. https://doi.org/10.1109/TVCG.2020.3010088

Microsoft: Azure Kinect DK. https://learn.microsoft.com/ko-kr/azure/kinect-dk/ (2020)

Miklós Á, Szabó Z (2015) Simulation and experimental validation of the dynamical model of a dual-rotor vibrotactor. J Sound Vib 334:98–107. https://doi.org/10.1016/j.jsv.2014.06.011

Moritzen K (2022) Opening up virtual mosh pits: music scenes and in-game concerts in fortnite and minecraft. J Sound Music Games 3(2–3):115–140. https://doi.org/10.1525/jsmg.2022.3.2-3.115

Nanayakkara S, Taylor E, Wyse L, Ong SH (2009) An enhanced musical experience for the deaf: design and evaluation of a music display and a haptic chair. In: Proceedings of the Sigchi conference on human factors in computing systems, pp 337–346. https://doi.org/10.1145/1518701.1518756

Nichols C (2002) The vbow: development of a virtual violin bow haptic human-computer interface. In: Proceedings of the 2002 conference on new interfaces for musical expression, pp 1–4. https://doi.org/10.1017/S135577180200211X

Park G, Choi S (2018) Tactile information transmission by 2d stationary phantom sensations. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–12. https://doi.org/10.1109/TOH.2020.3002696

Pehkonen S (2017) Choreographing the performer-audience interaction. J Contemp Ethnogr 46(6):699–722. https://doi.org/10.1177/0891241616636663

Rasool S, Sourin A (2014) Image-driven haptic rendering. Trans Comput Sci XXIII Special Issue Cyberworlds 42:58–77. https://doi.org/10.1007/978-3-662-43790-2_4

Remache-Vinueza B, Trujillo-León A, Zapata M, Sarmiento-Ortiz F, Vidal-Verdú F (2021) Audio-tactile rendering: a review on technology and methods to convey musical information through the sense of touch. Sensors 21(19):6575. https://doi.org/10.3390/s21196575

Schneider OS, Israr A, MacLean KE (2015) Tactile animation by direct manipulation of grid displays. In: Proceedings of the 28th annual ACM symposium on user interface software & technology, pp 21–30. https://doi.org/10.1145/2807442.2807470

Seo J, Mun S, Lee J, Choi S (2018) Substituting motion effects with vibrotactile effects for 4d experiences. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–6. https://doi.org/10.1145/3173574.3174002

Smyth T, Smyth TN, Kirkpatrick AE (2006) Exploring the virtual reed parameter space using haptic feedback. In: 2006 IEEE Workshop on multimedia signal processing, pp 45–49. https://doi.org/10.1109/MMSP.2006.285266. IEEE

Sun Z, Zhu M, Shan X, Lee C (2022) Augmented tactile-perception and haptic-feedback rings as human-machine interfaces aiming for immersive interactions. Nat Commun 13(1):5224. https://doi.org/10.1038/s41467-022-32745-8

Tawa S, Nagano H, Tazaki Y, Yokokohji Y (2021) Extended phantom sensation: vibrotactile-based movement sensation in the area outside the inter-stimulus. Adv Robot 35(5):268–280. https://doi.org/10.1080/01691864.2020.1854114

The show must go beyond. (2022). https://wavexr.com/

Thériault R, Olson JA, Krol SA, Raz A (2021) Body swapping with a black person boosts empathy: using virtual reality to embody another. Q J Exp Psychol 74(12):2057–2074. https://doi.org/10.1177/17470218211024826

Tsai H-R, Liao Y-S, Tsai C (2022) Impactvest: rendering spatio-temporal multilevel impact force feedback on body in vr. In: Proceedings of the 2022 CHI conference on human factors in computing systems, pp 1–11. https://doi.org/10.1145/3491102.3501971

Turchet L, Barthet M (2018) Co-design of musical haptic wearables for electronic music performer's communication. IEEE Trans Human-Mach Syst 49(2):183–193. https://doi.org/10.1109/THMS.2018.2885408

Turchet L, West T, Wanderley MM (2021) Touching the audience: musical haptic wearables for augmented and participatory live music performances. Personal Ubiquitous Comput 25(4):749–769. https://doi.org/10.1007/s00779-020-01395-2

Turchet L, Barthet M (2019) Haptification of performer's control gestures in live electronic music performance. In: Proceedings of the 14th international audio mostly conference: a journey in sound, pp 244–247. https://doi.org/10.1145/3356590.3356629

ur Rahman S, Sun J, Liu L, Li H (2008) Turn your mobile into the ball: rendering live football game using vibration. IEEE Trans Multimed 10(6):1022–1033. https://doi.org/10.1109/TMM.2008.2001352

Velt R, Benford S, Reeves S, Evans M, Glancy M, Stenton P (2017) Towards an extended festival viewing experience. In: Proceedings of the ACM international conference on interactive experiences for TV and online video, pp 53–62. https://doi.org/10.1145/3077548.2745206

Veronesi D (2014) Correction sequences and semiotic resources in ensemble music workshops: the case of conduction®. Soc Semiot 24(4):468–494. https://doi.org/10.1080/10350330.2014.929393

Verrillo RT (1966) Vibrotactile sensitivity and the frequency response of the Pacinian corpuscle. Psychon Sci 4(1):135–136

Wang Y, Lu L, Zeng Q, Li J (2020) An audience positioning method for high-fidelity sound system. In: 2020 IEEE 6th international conference on computer and communications (ICCC), pp 1151–1155. https://doi.org/10.1109/ICCC51575.2020.9344903. IEEE

Webb AM, Wang C, Kerne A, Cesar P (2016) Distributed liveness: understanding how new technologies transform performance experiences. In: Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing, pp 432–437. https://doi.org/10.1145/2818048.2819974

West TJ, Bachmayer A, Bhagwati S, Berzowska J, Wanderley MM (2019) The design of the body: suit: score, a full-body vibrotactile musical score. In: International conference on human-computer interaction, pp 70–89. https://doi.org/10.1007/978-3-030-22649-7_7. Springer

Wilson G, Freeman E, Brewster S (2016) Multimodal affective feedback: Combining thermal, vibrotactile, audio and visual signals. In: Proceedings of the 18th ACM international conference on multimodal interaction, pp 400–401. https://doi.org/10.1145/2993148.2998522

Yakura H, Goto M (2020) Enhancing participation experience in vr live concerts by improving motions of virtual audience avatars. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR), pp 555–565. https://doi.org/10.1109/ISMAR50242.2020.00083. IEEE

Yamazaki R, Ohkura M (2018) Affective evaluation while listening to music with vibrations to the body. In: International conference on applied human factors and ergonomics, pp 379–385, Springer. https://doi.org/10.1007/978-3-319-94944-4_41

Yamazaki Y, Mitake H, Oda R, Wu H-H, Hasegawa S, Takekoshi M, Tsukamoto Y, Baba T (2017) Hapbeat: single dof wide range wearable haptic display. In: ACM SIGGRAPH 2017 emerging technologies, pp 1–2. https://doi.org/10.1145/3084822.3084843

Yan S, Yan X, Shen X (2020) Exploring social interactions for live performance in virtual reality. In: SIGGRAPH Asia 2020 Posters, pp 1–2. https://doi.org/10.1145/3415264.3425466

Yun G, Lee H, Han S, Choi S (2021) 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. https://doi.org/10.1145/3411764.3445358

Yun G, Oh S, Choi S (2019) Seamless phantom sensation moving across a wide range of body. In: 2019 IEEE world haptics conference (WHC), pp 616–621. https://doi.org/10.1109/WHC.2019.8816104. IEEE

Zhang K, Kim LH, Guo Y, Follmer S (2020) Automatic generation of spatial tactile effects by analyzing cross-modality features of a video. In: Symposium on spatial user interaction, pp 1–10. https://doi.org/10.1145/3385959.3418459