Embodied Dance in Folk Cultural Space: Immersive Folk Dance Teaching Application to Enhance Cultural Experience

Mingyang Su
Tsinghua Shenzhen International Graduate School, Tsinghua University
Shenzhen, China
sumy22@mails.tsinghua.edu.cn

Yun Xie
Tsinghua Shenzhen International Graduate School, Tsinghua University
Shenzhen, China
894436717@qq.com

Binlin Feng
Shenzhen University
Shenzhen, China
fengbinlin2021@email.szu.edu.cn

Wen Ku
Tsinghua Shenzhen International Graduate School, Tsinghua University
Shenzhen, China
kuwen1816@163.com

Chao Liu
The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
cliu009@connect.hkust-gz.edu.cn

Xiaomei Nie
Tsinghua Shenzhen International Graduate School, Tsinghua University
Shenzhen, China
nie.xiaomei@sz.tsinghua.edu.cn

Xiu Li*
Tsinghua Shenzhen International Graduate School, Tsinghua University
Shenzhen, China
li.xiu@sz.tsinghua.edu.cn

Figure 1: The user controls his/her virtual avatar through body movements (A); the user controls hi/her virtual avatar to talk to the virtual non-player character in the game in a virtual cultural space(B); the user controls his/her virtual avatar to dance with the virtual non-player character in the game in a virtual cultural space (C); the virtual non-player character shows dance in the virtual cultural space (D); the user talks to the virtual non-player character in the game (E).

ABSTRACT

Current research on folk dance digitization focuses mainly on visualizing dance movements, ignoring the user's cultural experience. Using Hmong dance as a case study, we developed an application for folk dance digitization that allows users to interactively embodied in a folk cultural space, enhancing their cultural experience. We emphasize embodied participation in the folk cultural space through physical interaction for a more profound cultural experience. Through comparative experiments and interviews, we verified the effectiveness of this method in enhancing cultural experience. The study results provide new ideas and methods for the digitalization of folk dance research, making the cultural heritage of folk dance more attractive and engaging.

*Corresponding author

CHCHI 2024, November 22–25, 2024, Shenzhen, China

ACM ISBN 979-8-4007-1389-7/24/11.
https://doi.org/10.1145/3758871.3758917

ACM Reference Format:
Mingyang Su, Yun Xie, Binlin Feng, Wen Ku, Chao Liu, Xiaomei Nie, and Xiu Li. 2024. Embodied Dance in Folk Cultural Space: Immersive Folk Dance Teaching Application to Enhance Cultural Experience. In *Chinese CHI 2024*

(CHCHI 2024), November 22–25, 2024, Shenzhen, China. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3758871.3758917

1 INTRODUCTION

Folk dance is an important part of intangible cultural heritage, serving as a representation of specific folk traditions and cultures [49, 64]. Digitizing folk dance can make it more widely shared, and allow people to experience and learn folk dance in a novel way [22, 34, 52]. However, current research on the digitalization of folk dance still has some limitations regarding providing in-depth cultural experiences to users.

Early research focused on constructing ethnographies or databases by recording and storing dance motions [5, 7, 18, 19, 48, 62], to explore how to build connections between digitized folk dances and their cultural context [12, 14, 16, 52]. While these efforts contribute to preserving cultural heritage, they often ignore that the dance is dynamic and its interaction with participants. Static ethnographies and databases lack embodied participation, where users simply search and view dance animations, making it challenging to comprehend the cultural significance behind the dance motions.

It has been shown that embodied participation is an effective way to enhance user cognition. Many studies constructed ethnographies or databases by recording and storing dance motions, allowing users to engage in learning folk dance [2, 23, 34, 38, 39]. However, these studies often only focus on motions without considering the need to relate folk dances to cultural context [33, 51, 52], lacking an explanation of the cultural connotations behind the dance motions in embodied participation.

To address that, we take Hmong dance as an example and develop an immersive dance teaching application that enhances cultural experience through embodied dance teaching. We designed an immersive virtual cultural space with folk features, where users control their avatars to learn folk dance through body motions. Additionally, we allow the virtual humans to wear folk costumes and perform folk dances in the virtual environment. As shown in Fig. 1, users can interact with virtual humans through their virtual avatars to grasp cultural knowledge and instruction about folk dances. Our contributions include:

(1) We designed an embodied dance teaching application based on virtual avatars and motion recognition technology.
(2) We placed the folk dance content in its related virtual cultural space, comprehensively recording and preserving the dancers' motions, props, and environment.
(3) We further conducted a user study to evaluate the effectiveness of our application. The results showed that this method significantly increased users' understanding and interest in folk dance culture.

2 RELATED WORK

2.1 Content Representation in Digital Dance

2.1.1 Traditional Content Representation. Early studies stored digital dance using video or images, recording the motion, forms, and performances of dances. These media became common forms of preservation for documenting dance culture [21, 29, 35, 53, 63]. For instance, the SFDC project collected videos of traditional dances from around the world[11], while the Folk Dance Federation of California project also preserved a significant number of related folk dance videos[44].

With advancements of motion capture technology, the digital reproduction of dance in virtual space has become feasible[13, 17, 37, 53]. This technology records dancers' motions as digital animations, allowing for the permanent preservation of dance and the build of an extensive database[30]. The Wholodance project developed a dance animation tool that records the dancer's motion data through a motion capture system and records it as skeletal animation[50].

To enhance the understanding of folk dances, these animations are interpreted by textual information. For example, the WebDANCE project has developed an e-learning platform for teaching traditional dances, which uses a motion capture system to generate 3D animations. And digital humans and explainable text are presented for teaching [30]. However, these digital humans are usually placed in abstract virtual spaces, wearing ordinary clothes, and are disconnected from the folk cultural connotations of the dance.

Although these studies have succeeded in achieving digital preservation of folk dance, they generally lack integration with the cultural context of folk dance. Winnicott believes that cultural experience is latent in the potential space between cultural individuals and the environment[59]. Folk dance has its specific cultural connotation, and its cultural expression should not be dropped in the digitizing folk dance[26, 56]. The folk costume and environment play an important role in folk dance. Still, at present, the most research ignores the cultural context brought by the costume and environment and only records the dance animation.

2.1.2 Folk Culture Space. Pierre Bourdieu's field theory defines a field as "a network or structure of objective relationships between positions defined by objective criteria" [24]. From this perspective, folk dance is a specific activity carried out by people with specific identities in a specific cultural context, covering psychological, spatial, and other elements. In the study of "folk dance," the field can be regarded as the embodiment of a folk cultural space, which includes not only folk dance animation but also necessary elements, such as folk costumes, local scenes, and the participants themselves.

Some current studies attempt to provide a more culturally contextualized visual representation of dance. In the study of the digitization of Cypriot dance[52], the consistency of the dance with the cultural context was considered, with virtual characters dressed in local Cypriot costumes and performing in real spaces. Iris et al. [33] provide a generally rural environment in a VR space, which contributes to the user's understanding of the traditional folk dances. The digitization of dance in Pauliteiros provides the user with motion data and allows for the creation of virtual dancers in typical Pauliteiros costumes for animation[51].

Additionally, multimodal databases are increasingly being applied to the recording and presentation of dance culture. The i-Treasures project has established the first multimodal dataset for intangible cultural heritage, integrating images, videos, text, and motion-capture animations[15]. This approach provides a more comprehensive solution for documenting dance culture, making the process more refined and complete.

In summary, the current digitization of folk dance primarily relies on videos, text, or motion-capture data, with attempts to incorporate corresponding cultural contexts. These studies highlight

that the digitization of dance requires not only advancements in technology but also an emphasis on restoring and integrating its cultural background. Placing dance within an environment that is connected to its cultural context enables a more thorough preservation and transmission of its cultural connotations.

2.2 User Participation in Dance Digitization

Barbara pointed out that the digitization of tangible culture is "object-oriented," focusing on static records [36], while Christoph emphasized that intangible culture should be "human-oriented," highlighting the key role of its carriers and dynamic expressions [60]. As a dynamic intangible cultural heritage, the digitization of ethnic dance particularly requires user participation.

2.2.1 Digital Dance Database. Folk dance can be recorded in different ways, and early digital studies relied more on the recording and retrieval of motions and the presentation of forms such as text, video, and graphic symbols [21, 29, 35, 53, 63] . Even though these studies successfully record dance content and forms, objects have limited cultural dissemination of such intangible things as dance.

Both video and motion-capture animation databases provide users with convenient ways to search and watch dances [30, 50]. For example, Matus and colleagues developed a 3D visualization application for demonstrating and teaching ethnic dance from different regions of Slovakia, allowing users to access and learn these dances anytime. Similarly, the Terpsichore project focuses on researching, archiving, and presenting folk dance as an intangible cultural heritage, making dance professionals, educators, the creative industry, and the general public accessible to it.[6].

In addition, multimedia content has further expanded the interactive methods for folk dance. The AniAge project[1] aims to preserve Southeast Asian performing arts, particularly local dances at risk of vanishing. Sarah Kenderdine's team recreates Confucian rituals through theatre, venues, and object replication based on "The Book of Rituals," offering cultural re-enactment via an app and interactive page [31]. However, it still revolves around user retrieval and viewing.

However, these databases focused on visualizing cultural elements and restoring existing dance culture, resulting in a relatively monotonous and dull user experience. Purely visual viewing makes it difficult for users to grasp the cultural essence of the dance. As an intangible cultural heritage, ethnic dance fundamentally emphasizes embodied user participation.

2.2.2 Embodied Dance. In dance teaching, embodied interaction is widely used to enhance users' perception of dance [47]. The basic idea of embodied interaction is that human cognitive processes are not only based on the brain's functioning but are generated through bodily perception and interaction with the surrounding environment. Embodied interaction emphasizes the important role of perception and the body in cognitive activity [3, 27]. Still, it does not limit itself to emphasizing the physicality of cognition, but rather the dynamic unity of cognition, body, and environment.

Teaching dance with embodied interaction inspires us that motions can be used as the primary methods of interaction. Users can immerse themselves in the folk space and interact with virtual digital entities through body motions and dialogues, thus gaining access to the cultural connotations behind dances. Zhou et al. [64] provided an overview of the development of dance in human-computer interaction over the past two decades, highlighting body motion as a key medium for communication. Drawing on theories of bodily interaction [25] and empirical studies, previous research has explored innovative uses of bodily signals and the concept of kinesthetic creativity [9, 28, 58].

Grammatikopoulou et al. [47] developed a Kinect-based framework for the gamification of ICH dance, focusing on teaching motion in ICH dance, from loading motion capture data to customizing the design of dance games. [23] created a virtual 3D game environment that allowed users to observe dance motions from any angle. In [2], dance learning was implemented through a game interface, where the agent represented by the user was located in the center of the scene and was able to imitate the teacher agent located in the same space. However, these studies focused more on learning dance motions in the process of digitization while ignoring the cultural experience of dance.

Overall, current research on the digitizing folk dance mainly includes two forms of user participation: first, searching and watching dances through 3D visualization platforms; and second, engaging in embodied participation for dance learning. However, few studies integrate users' embodied participation with culturally contextualized spaces, enabling a comprehensive understanding of folk culture.

3 PROTOTYPE DESIGN AND IMPLEMENTATION

Therefore, this study aims to place users' embodied participation within specific folk cultural spaces. Through immersive interaction, it aims to help users fully experience the richness of folk dance.

To achieve this goal, we developed a multi-modal interactive application that digitizes and supports the learning of Hmong folk dance. Drawing on embodied cognition theory [24, 57], we designed interaction mechanisms that encourage users to engage with dance movements physically. In parallel, informed by field theory, we curated the application's visual and spatial content. The interface and features of the application are illustrated in Fig. 2.

3.1 Emobodied Interaction

Embodied interaction allows users to learn folk dances by engaging with virtual entities in a culturally situated virtual space. This approach not only helps users master dance gestures with real-time feedback but also immerses them in the rhythm, aesthetics, and spatial context of ethnic traditions, deepening their cultural understanding beyond what traditional methods offer.

Therefore, we adopt monocular motion capture as the primary interaction method, enabling users to perform full-body dance movements freely within a culturally immersive virtual environment. Compared to traditional VR-based systems that rely on wearable devices, this approach offers greater convenience and accessibility by eliminating hardware constraints [40, 54]. Specifically, we employ MediaPipe-based posture estimation to extract users' postures from monocular video input, allowing for real-time embodied interaction with virtual elements. This technique is particularly suitable for dance learning applications, where users benefit from

both natural movement and immediate visual feedback, thereby enhancing engagement and cultural immersion.

Figure 2: Embodied dance teaching in the folk culture space.

To maintain continuity and minimize disruption, we introduce voice interaction as an auxiliary input method. Through voice commands, users can communicate with the virtual dance agent, retrieval dance materials, and engage in the experience without leaving the immersive cultural space. Our voice interaction is powered by Azure's speech recognition technology.

3.2 Folk Cultural Space

According to Bresnahan's framework [10], the field of folk dance can be understood through three dimensions: performance environment, performance content, and performance culture. The performance environment includes the dancers, props, timing, and spatial settings involved in the presentation of folk dances. The performance content specifically refers to the physical dance movements of folk dance, including dance postures, dance rhythm. The performance culture encompasses the cultural narratives behind the dance, expressed through background stories, character roles, and symbolic meanings embedded in the performance.

For the performance environment, we expressed cultural semantics by designing virtual scenes and characters rooted in Hmong traditions. Using MAYA, we modeled 3D characters wearing distinctive Hmong costumes, such as bird-inspired garments, and constructed Hmong-style architectural environments in Unreal Engine to recreate authentic performance settings.

For the performance content, we employed the VICON full-body optical motion capture system to record 3D dance animations. By capturing dancers' skeletal motion data, we built a folk dance dataset that objectively documents movement patterns and temporal rhythm.

For the performance culture, we designed digital Hmong characters who appear in culturally contextualized virtual scenes. These agents wear traditional costumes, reproduce ethnic dance movements, and give explanations. Through light interaction, users can both observe the dances and deepen their understanding of Hmong heritage.

4 USER STUDY

We aimed to evaluate the effectiveness of the cultural experience of the application. We designed two independent variables, **Representation** and **Participation**, based on folk cultural space and embodied interaction. We used a customized questionnaire as the metrics to explore the effectiveness of folk cultural space and embodied interaction in enhancing cultural experience. In addition, we conducted semi-structured interviews at the end of the study to collect qualitative feedback.

4.1 Study Design

To assess the efficacy of our application design, we will investigate if embodied interaction in folk cultural space, as opposed to conventional digital storage and static user engagement, favors the ethnic dance culture experience. On the basis of this, we put out the subsequent experimental hypotheses:

H1: Users obtain a better cultural experience from folk cultural space compared to digital videos.

H2: Users obtain a better cultural experience from embodied interaction participation compared to passive search and viewing.

We examine users' cultural engagement experiences in various settings at the application to further support our hypotheses. We introduced two independent variables, **Representation** and **Participation**. We manipulated their values to investigate how different combinations affect users' cultural experience.

While traditional dance digitization often relies on text, video, and symbolic representations, we introduced a more immersive folk cultural space condition for comparison, as shown in Fig. 3 right:

VIDEO: folk dances were presented in video format.

VERB: During the performance, voice and text were used to explain cultural elements corresponding to the movements.

ART: The performance included culturally appropriate costumes and background scenes.

MULTIMODAL: A virtual dancer performed within a folk environment, accompanied by real-time explanations of dance-related cultural content.

In order to compare the embodied interaction with the static engagement of traditional digital dance application, we investigated

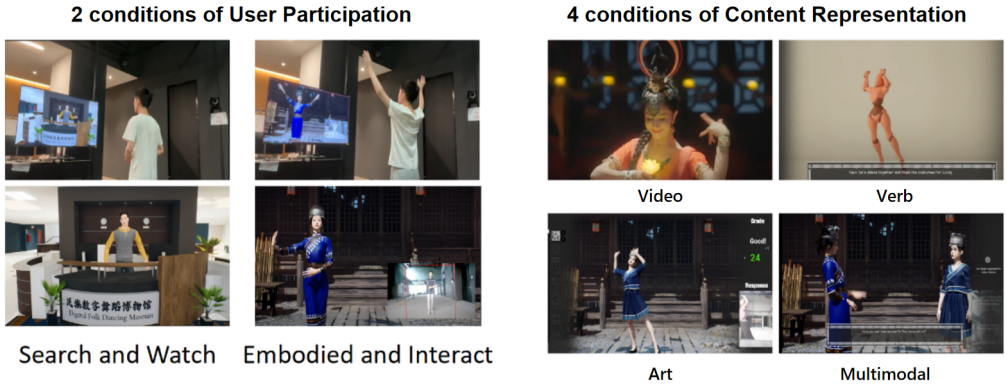

Figure 3: Independent variable. Two conditions of Participation(left). Four conditions of Representation(right).

the cultural experiences of participants in two different conditions, as shown in Fig. 3 left:

SEARCH: Participants engaged by searching and viewing.

EMBODIED: Participants actively participated in folk dance teaching.

The FRACH framework [4] extends the ISO/IEC-25010 quality model by introducing immersiveness and collaboration as key dimensions for evaluating cultural heritage game experiences. Based on this model, we selected three dimensions relevant to our study: knowledge acquisition, satisfaction, and usability, as well as the immersiveness dimension. In addition, we incorporated dimensions from the MEEGA+ model [46], including fun, learnability, accessibility, and control, to complement our evaluation of the learning experience. Using these eight dimensions, we designed a 7-point Likert-scale questionnaire to assess both cultural and user experiences. Participants were asked to rate each dimension after completing the task:

Q1: I can immerse myself in the experience.

Q2: I find the experience process enjoyable.

Q3: During the experience, I can control my own pace.

Q4: I feel the application's interactions are user-friendly.

Q5: I can learn about useful Hmong culture from the application.

Q6: I can understand the content expressed by the application.

Q7: I can easily gain an understanding of Hmong culture from the application.

Q8: I am satisfied with the entire application.

We adopted a within-subject design and followed a Latin Square Design for counterbalance. For the questionnaire, We first tested the data for normality using Shapiro-Wilk tests. If normality was satisfied, we conducted an Analysis of Variance; if not, we used the Mann-Whitney U test to analyze and report the results.

4.2 Participants,Apparatus and Procedure

We recruited 24 participants, 12 males and 12 females, to maintain gender balance, with a mean age of 21.5 and a variance of 1.49. Our framework focused on a broader, extra-cultural audience; none of the participants knew enough about Hmong dance culture. Nine had some dance experience, and fifteen were dance novices.

To ensure a smooth experimental procedure, we prepared a comprehensive setup including the testing environment, interactive system, and evaluation tools. The experimental space measured approximately 2 × 2 square meters. Participants engaged with multimodal content through a high-definition camera and a 72-inch display equipped with voice recognition, allowing for natural body and voice-based interaction. All necessary hardware and software components were carefully calibrated to provide a consistent and high-quality interactive experience.

At the beginning of the experiment, participants completed a demographic questionnaire and signed an informed consent form. We then introduced the research background and objectives, followed by a brief assessment of their familiarity with Hmong culture to establish a baseline understanding. Next, we provided a detailed explanation of the experimental procedure, including the three task steps and relevant technical information. Participants received a walkthrough of the application interface and interaction methods to ensure they were comfortable with the system. During this familiarization phase, we answered any questions and clarified key features as needed. After this adaptation period, each participant interacted with the Hmong Dance application under different content representation and user participation conditions.

As Shown in Fig 3, participants stood in the experimental space. In the **SEARCH**, they could interact using voice commands or mouse and keyboard inputs to search for or browse dances they wished to view and learn. The search database was pre-constructed by our researchers based on motion capture animation data and implemented in the Unreal Engine. In the **EMBODIED**, users learned dances by following digital human demonstrations, with the system providing feedback on the accuracy of their movements through motion recognition technology. Various modal representations were also pre-stored in the database: **VIDEO** from online, while **VERB** and **ART** were motion-captured animations presented by different digital humans in the Unreal Engine.

After each condition, evaluation forms were given out to participants to get their subjective opinions. We conducted interviews after respondents completed the questionnaires to get more in-depth input.

Figure 4: Stacked divergent bar chart for custom questionnaire.The upper left is Q1, the upper right is Q2, the lower left is Q3, and the lower right is Q4.

5 RESULT

5.1 Reliability and validity

The standardized Cronbach's alpha value was discovered to be 0.850 after reliability analysis, indicating that the questionnaire has a generally good level of reliability. The reliability coefficient did not significantly increase when any particular item was left off the questionnaire, indicating that no revisions are required.

At the same time, the KMO value of the questionnaire validity is 0.841, there is a correlation if it is greater than 0.8, and Bartlett's sphericity test ($p < .001$) is significant, which meets the requirements of factor analysis.

5.2 Quantitative Results

We collected the results of the questionnaires from the participants, conducted Shapiro-Wilk tests on them, and satisfied the normal distribution. These results were then further analysed using two-way ANOVA and post hoc tests. For this we obtained the following results for the data:,Q1 to Q4 as shown in Fig. 4, Q5 to Q8 as shown in Fig. 5

Q1: I can immerse myself in the experience. We found that the **Participation** showed significance ($F = 61.873, p < .001$) indicating the presence of a main effect and that the mode of engagement would have a differential relationship on immersion. Meanwhile, **Presentation** style showed a significant main effect ($F = 104.978, p < .001$). Different **Representation** also had a differential relationship on immersion. Post hoc multiple comparisons showed that **EMBODIED** scores were significantly higher than **SEARCH** ($p < .001$), and both **ART** and **MULTIMODAL** were higher than **VERB** ($p < .001, p < .001$) and **VIDEO** ($p < .001, p < .001$). The interaction effect between **Participation** and **Representation** was not significant ($F = 1.491, p = 0.219 > .05$).

Q2: I find the experience enjoyable. Participation showed a significant main effect on the enjoyment of the experience ($F = 46.982, p < .001$), as did **Representation** ($F = 92.353, p < .001$). Post hoc multiple comparisons showed that the **EMBODIED** had significantly higher scores than the **SEARCH** ($p < .001$), and **MULTIMODAL** received significantly higher scores than **ART** ($p < .01$) and much higher scores than **VERB** and **VIDEO** ($p < .001$ for all). Additionally, **ART** received significantly higher scores than **VERB** and **VIDEO** ($p < .001$ for both). There was no significant interaction effect between **Participation** and **Representation** ($F = 0.842, p = 0.473 > .05$).

Q3: During the experience, I can control my own pace. Participation showed a significant main effect ($F = 140.258, p < .001$), indicating that there is a main effect of **Participation** on perceived control. **Representation** did not show a significant main effect ($F = 1.271, p > .05$). There was also no significant interaction effect between **Participation** and **Representation** ($F = 2.634, p > .05$), suggesting that there is no interaction between the two factors. Post hoc multiple comparisons revealed that the **SEARCH** had significantly higher scores than the **EMBODIED** ($p < .001$).

Q4: I feel that the application's interaction is user-friendly. Participation exhibited a significant main effect ($F = 9.274, p < .01$), indicating that there is a main effect of **Participation** on user-friendliness. **Representation** also showed a significant main effect ($F = 60.528, p < .001$), indicating that **Representation** has a significant impact on user-friendliness. There was no significant interaction between the two independent variables ($F = 0.227, p > 0.05$). Post hoc multiple comparisons revealed that **EMBODIED** scores were higher than **SEARCH** ($p < .05$), **MULTIMODAL** and **VERB** were both higher than **ART** ($p < .001, p < .001$) and **VIDEO** ($p < .001$), and **ART** was significantly higher than **VIDEO** ($p < .001$).

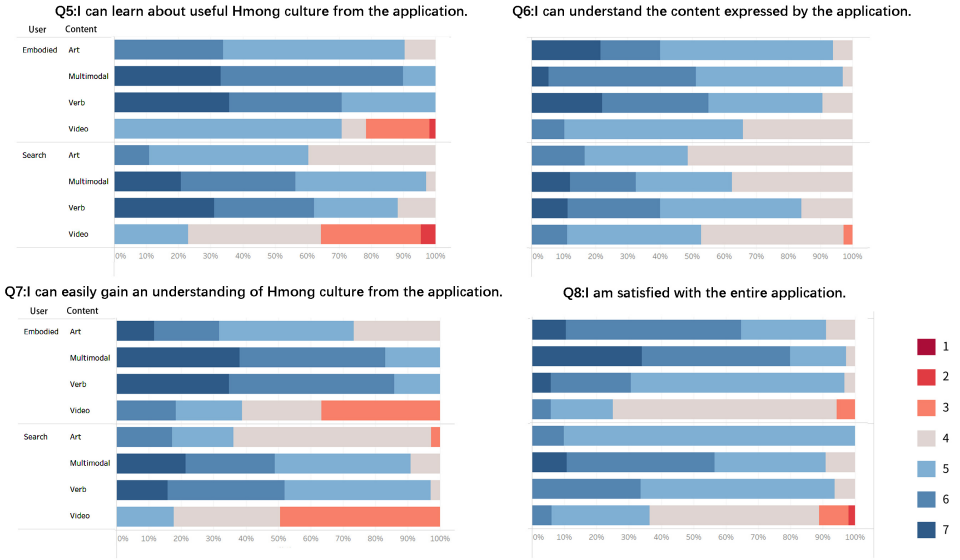

Figure 5: Stacked divergent bar chart for custom questionnaire.The upper left is Q5, the upper right is Q6, the lower left is Q7, and the lower right is Q8.

Q5: I can learn about useful Hmong culture from the application. We found a significant main effect for **Participation** ($F = 17.946, p < .01$) and a significant main effect for **Representation** ($F = 58.828, p < .001$). There was no significant interaction effect between **Participation** and **Representation** ($F = 0.258, p > .05$). Post hoc multiple comparisons showed that **EMBODIED** was rated higher than **SEARCH** ($p < .01$), and **MULTIMODAL** and **VERB** were both significantly higher than **ART**($p < .001, p < .01$) and **VIDEO** ($p < .001$ for all comparisons), with **ART** also being significantly higher than **VIDEO** ($p < .01$).

Q6: I can understand the content presented by the application. Participation presented significance in the comprehensibility dimension ($F = 17.297, p < 0.001$), while **Representation** presented significance ($F = 7.956, p < .001$), suggesting that the main effect was present for both. No significance was presented between **Participation** and **Representation** ($F = 1.729, p > 0.05$), indicating that there was no interaction effect between the two. We further conducted post hoc multiple comparisons and found that **EMBODIED** was higher than **SEARCH** ($p < .001$), **MULTIMODAL** with **VERB** scores were significantly higher than **VIDEO** ($p < .01, p < .001$).

Q7: I can easily understand the Hmong culture from the application. We analyzed the data and found a significant main effect of **Participation** ($F = 21.559, p < .01$), and there was also a significant main effect of **Representation**($F = 68.816, p < .001$). Post hoc analysis revealed that **EMBODIED** was higher than **SEARCH** ($p < .001$), **MULTIMODAL** and **VERB**'s ratings were both higher than **ART** ($p < .001, p < .001$), and all three were higher than **VIDEO** ($p < .001, p < .001, p < .001$).

Q8: I am satisfied with the entire application. Participation showed a significant main effect ($F = 7.083, p < .01$), and **Representation** also displayed a significant main effect ($F = 44.554, p < .001$), indicating main effects for both factors that influence satisfaction. However, there was no significant interaction effect between **Participation** and **Representation** ($F = 1.693, p > .05$), suggesting no interaction between the two factors. Post hoc analysis showed that **EMBODIED** was rated higher than **SEARCH** ($p < .001$), and **MULTIMODAL**, **VERB**, and **ART** were all rated higher than **VIDEO** ($p < .001$ for all).

5.3 Qualitative Results

A semi-structured interview was conducted after each participant had finished the experiment. During this process, one researcher would fully interact with the participant, guiding the participant to express his/her views. Meanwhile, the other researcher would record the interviews using audio recording software. To analyze the data further, after all the experimental sessions, the two researchers converted the conversations into text and open-coded all the data to find potential codes and themes for thematic analysis. For this purpose, the following themes were obtained:

5.3.1 Cultural Understanding. Cultural experience, the fundamental task of our experimental study, was to explore the impact of the differences between different forms of content presentation and different ways of user engagement on users' cultural experience. Thus, we focused on exploring participants' cultural experiences and interviewed them about their perceptions of folk cultures in dance museums.

On **VERB**, participants generally agreed that text was their primary way of learning about dance culture. Several participants stated, *"Textual explanations are very important in helping us understand the cultural meanings behind the motions during the dance experience."*(P13). For participants who were not culturally familiar, it was often difficult to understand the cultural connotations

directly from the dance motions; therefore, proper textual explanations helped them understand more easily. *"This text explanation was clear and helped me understand the Hmong culture."* (P8) Another participant said, *"I can now understand the connotations of the motions through explanations; for example, the motion 'Duck Step' looks very much like a duck walking, but before, I didn't think that it originated from ducks."* (P3) In the motion disassembly and participants' embodied interactions, the textual explanations effectively aided participants' understanding of the motions, *"After the explanations, it was easier for me to understand this complex three-beat motion, which I had thought was a complicated motion, but now I can learn the motion better."* (P6)

Understanding Hmong culture came from participants' experiences gained through **EMBODIED**. On the one hand, the embodied participation brought fun to the participants. It made it easier for them to understand the culture behind the dance: *"I still didn't quite understand it just by listening to the explanation, but dancing along made it clearer."* (P5). Another participant said, *"Watching more is not as impressive as dancing once."* (P2) Participants also felt that the embodied approach left a deeper impression: *"I remembered the motion."* (P2) *"I have a very strong memory of this monkey-like motion."* (P3)

5.3.2 Embodied Interaction. Embodied interaction was another focus of our research, and we explored the impact of this form of embodied interaction on participants' cultural experiences. Most of the participants reported that **EMBODIED** could give them a deeper understanding of the cultural meaning behind the dance. When they mimicked the digitized dancers' motions, they were learning dance techniques and experiencing a cultural story. *"Through the actual dance motions, I feel closer to the culture of this people"* (P6), and *"When imitating the dance steps, I seem to feel the story and emotion behind the motion"* (P9). The embodied interaction provided them with an immersive way of learning about the culture that was not limited to visual and auditory senses but also bodily perception.

We also found that embodied interaction allowed users to be more active and engaged in the learning process than traditional viewing learning. *"When I was involved in the dance, I felt that I was no longer an outsider but became part of the culture"* (P12), and *"Learning the dance physically was more interesting than just watching the video and made me remember the cultural features"* (P15). This interactive approach increased their engagement and reinforced learning. The museum mobilized participants' voice and body interactions, freeing them from the limitations of the keypad. *"I could dance along"* (P6). Overall, participants rated the **EMBODIED** as enhancing their immersion so that they were not just sitting in front of a computer experiencing the digital museum. Still, they were actually *"moving their whole body"* (P4). On the contrary, participants were less positive about the immersiveness of **SEARCH** (5 times), saying, *"I know exactly what I'm doing now; I'm not in the virtual world anyway"* (P7), *"I can only click with the mouse, it lacks the sense of reality"* (P10).

Additionally, participants reflected on some of the challenges in **EMBODIED**, finding the form of embodied participation too difficult. *"Sometimes it felt a bit difficult to follow the pace of the digital dancers, and it took some time to get used to it"* (P4), and *"I would like to have more options for personalized adjustments so that I can learn at my own pace and ability"* (P10).

5.3.3 Immersive Culture Space. Participants emphasized enhancing the cultural experience through the folk cultural space. In **ART**, they felt that the art design that matched the folk dances made them more able to feel the cultural elements in the dances. *"The art presentation of the Hmong style made me appreciate the culture in it more."* (P17). Several participants said that the folk costumes and scenes were an inseparable part of the dance, and *"The combination of Hmong costumes and scenes made this dance more complete."*(P23). Compared to the scenes that lacked artistic expression, participants felt that *"The dance presentation without artistic elements lacked a sense of entry, and there was a cognitive gap."* (P22). Especially when the unique costumes of the Hmong people appeared, *"A unique feeling of the exotic flavor came to my mind."*

Meanwhile, *"cultural immersion"* was the most mentioned feeling by the participants during the whole experience, and an important source of this immersion is the virtual cultural space constructed by folk scenes. *"The virtual scene of the Hmong village greatly enhanced my sense of immersion."* (P1), *"The appearance of the Hmong digital characters attracted my attention."* (P7), *"The folk costumes on the digital characters attracted me."* (P3), and *"the Hmong architecture gave me stronger cultural resonance than the blank walls."* (P8). Participants agreed that the art scene design and digital characters related to Hmong culture constituted a strong immersion experience.

However, a few participants indicated that sudden textual interpretations (**VERB**) interrupted their continuous emotional experience while immersed in the dance experience. *"I don't like to be interrupted by sudden explanations while following the dance; it spoils my immersion."*(P4) and they preferred purely ethnocultural spaces and lilting music, *"just music and fine arts, no need for voice explanations to disturb me."* (P4)

6 DISCCUSION

Through quantitative analyses of customized questionnaires and in-depth analyses of user interviews, we have revealed the impact of different digital representations of folk dance and different ways of user engagement on the cultural experience. However, the specific differences, strengths, and weaknesses of each condition will be discussed in more detail in this section. We will also outline the limitations of these findings and suggest possible directions for future research.

6.1 Enhancing Cultural Experiences through Embodied Interaction

Embodied interaction enables users to participate directly in folk dance through bodily movement, allowing them to act as both observers and active experiencers. Compared to traditional VR-based applications, monocular motion capture offers a greater sense of freedom and immersion, as it removes the need for wearable equipment [40, 54]. This form of deep engagement facilitates not only the learning of dance techniques, but also an embodied experience of the cultural narratives and emotional expressions embedded in the dance [41, 57]. In contrast, traditional approaches such as video-based folk dance corpora often render users passive viewers, limiting their cultural engagement. Through motion imitation,

users internalize both physical skills and symbolic meaning, transmitting cultural emotions through bodily sensation to cognitive memory, a process shown to enhance cultural retention and affective resonance [55].

The embodied interaction further strengthened the cultural experience brought by the cultural field. Participants agreed that by dancing or engaging in interactions with the digital human guides, it was easier for them to integrate themselves into the Hmong cultural scene. This sense of actual participation enhanced their cultural immersion, which increased the learning attractiveness and strengthened the users' understanding of and emotional commitment to the culture[43, 61]. At the same time, the embodied interaction eliminates the limitations of keyboard and mouse, and the dialogue with the digital person also seems to place the user into the same space for cultural exchange across time and space.

6.2 Building Cultural Context through Folk Cultural Space

Folk cultural space can effectively construct immersive cultural fields by engaging multiple sensory channels, including visual, auditory, and embodied modalities. Compared to traditional video representations, the **multimodal** and **art** received higher ratings for cultural engagement. Participants appreciated the detailed presentation of Hmong architecture and costumes, frequently citing these visual and spatial elements as key contributors to cultural immersion. This supports prior research suggesting that the digitization of folk dance should incorporate not only bodily motion but also culturally relevant spatial and aesthetic contexts [33, 52].

In addition to visual and spatial elements, the **multimodal** and **verb** provided cultural explanations alongside the dance performance. This was especially helpful for participants unfamiliar with Hmong culture, who reported that textual or verbal descriptions enhanced their understanding of the dance's symbolic meaning. In the **verb**, participants followed cultural annotations while engaging in body movements, allowing them to link abstract narratives with physical experience. However, when text was presented alone without accompanying visuals or embodied elements, participants described the experience as "less intuitive" and "cognitively detached." This highlights the need for an integrated multimodal approach in which visual design (**art**), bodily interaction, and narrative explanation (**verb**) form a cohesive cultural space.

Nevertheless, some participants expressed concerns that the textual or voice-over content in **verb** disrupted the immersive experience. They noted that hearing a narrator's voice, rather than the voice of the digital dancer, created a sense of dissonance that broke the illusion of a coherent cultural space [32]. Future implementations may improve this by synchronizing the narrative voice with the digital dancer and embedding cultural expression more seamlessly within the embodied interaction.

6.3 Lack of Freedom and Tediousness in Text

The results of our study have highlighted some challenges and areas for improvement. Firstly, participants perceived that embodied interaction seemed to offer less freedom and control than the **search**, a viewpoint validated by the actual ratings. They felt their autonomy was restricted when following the avatar in embodied interaction, preventing them from exploring the content freely. While traditional **search** lacked professional digital guides and opportunities for cultural dance experiences, participants could freely search and explore information in the dance corpus. These findings require us to pay attention to balancing the guidance of the avatar and the user's freedom [8, 20].

Concerns were also raised regarding the textual explanations presented in the **verb** condition. Some participants reported that the written content was overly lengthy and occasionally monotonous, which diminished their enjoyment and reduced engagement. These observations are consistent with the limitations in cultural experience identified in the quantitative results. To enhance user engagement and maintain continuity in digital cultural heritage experiences, future designs should consider more vivid and interactive ways of presenting cultural explanations. One effective strategy could be the integration of text with visual elements [42], enabling richer multimodal storytelling.

6.4 Limitations and Future work

While our system demonstrates the potential of embodied interaction in folk cultural spaces, several limitations remain. First, monocular motion capture technology still faces challenges in accurately and consistently recognizing users' body movements [45]. Delayed or missing feedback sometimes caused users to question whether their movements were properly captured, which disrupted immersion and affected the learning experience.

Second, the complexity and cultural specificity of folk dance movements result in noticeable differences in performance and comprehension between beginners and experienced users. Many novice participants found it difficult to follow the digital instructor, leading to frustration. This highlights the need for more adaptive systems that can provide instruction based on users' skill levels and offer supportive guidance during learning.

Future research could explore how embodied interaction within culturally immersive environments promotes user engagement and cross-cultural understanding. Customizing content delivery based on users' cultural interests, preferences, and interaction patterns may improve both accessibility and learning outcomes. Researchers may also consider applying embodied interfaces and multimodal content to a broader range of cultural contexts and user demographics to improve inclusivity and relevance.

7 CONCLUSION

This study investigated how to enhance cultural experiences in the digitization of folk dance through embodied interaction in a culturally grounded virtual environment. We developed an interactive folk dance teaching application that integrates motion capture, visual design, and cultural narration to support immersive cultural learning.

Through comparative experiments, we confirmed the effectiveness of embodied interaction within folk cultural spaces in promoting cultural understanding. Elements such as culturally inspired visual design, narrative explanations, and dance animations significantly contributed to constructing a rich cultural context and enhancing users' sense of immersion.

Furthermore, physical participation enabled users to engage more deeply with the emotional and symbolic dimensions of folk dance. This embodied engagement not only improved users' comprehension of cultural content but also fostered stronger affective connections, making the learning experience more meaningful and memorable.

ACKNOWLEDGMENTS

This work was supported by a research grant from Shenzhen Key Laboratory of Next Generation Interactive Media Innovative Technology(Funding No: ZDSYS20210623092001004), and the Center for Social Governance and Innovation at Tsinghua University.

REFERENCES

[1] Telli Abdelmoutia, Thi Chau Ma, Bourahla Mustapha, Tabia Karim, and Benferhat Salem. 2018. AniAge Ontology for Movement Classification in Vietnamese Dance. (2018).

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

[3] Michael L Anderson. 2003. Embodied cognition: A field guide. *Artificial intelligence* 149, 1 (2003), 91–130.

[4] Roberto Andreoli, Angela Corolla, Armando Faggiano, Delfina Malandrino, Donato Pirozzi, Mirta Ranaldi, Gianluca Santangelo, and Vittorio Scarano. 2017. A Framework to Design, Develop, and Evaluate Immersive and Collaborative Serious Games in Cultural Heritage. *J. Comput. Cult. Herit.* 11, 1, Article 4 (dec 2017), 22 pages. https://doi.org/10.1145/3064644

[5] Andreas Aristidou, Panayiotis Charalambous, and Yiorgos Chrysanthou. 2015. Emotion analysis and classification: understanding the performers' emotions using the LMA entities. In *Computer Graphics Forum*, Vol. 34. Wiley Online Library, 262–276.

[6] Andreas Aristidou, Ariel Shamir, and Yiorgos Chrysanthou. 2019. Digital dance ethnography: Organizing large dance collections. *Journal on Computing and Cultural Heritage (JOCCH)* 12, 4 (2019), 1–27.

[7] Andreas Aristidou, Efstathios Stavrakis, and Yiorgos Chrysanthou. 2014. LMA-based motion retrieval for folk dance cultural heritage. In *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: 5th International Conference, EuroMed 2014, Limassol, Cyprus, November 3-8, 2014. Proceedings 5*. Springer, 207–216.

[8] Matthew Ball and Vic Callaghan. 2012. Explorations of autonomy: an investigation of adjustable autonomy in intelligent environments. In *2012 Eighth International Conference on Intelligent Environments*. IEEE, 114–121.

[9] Bertha Bermudez and Chris Ziegler. 2014. Pre-choreographic movement kit. In *Proceedings of the 2014 International Workshop on Movement and Computing*. 7–12.

[10] Aili Bresnahan. 2014. Improvisational artistry in live dance performance as embodied and extended agency. *Dance Research Journal* 46, 1 (2014), 85–94.

[11] Stockton Folk Dance Camp. 2024. Dance Videos. https://stocktonfolkdancecamp.org/dance-videos/

[12] Giacomo Cozzani, Francesca Pozzi, Francesca Maria Dagnino, AV Katos, and EF Katsouli. 2017. Innovative technologies for intangible cultural heritage education and preservation: the case of i-Treasures. *Personal and Ubiquitous Computing* 21 (2017), 253–265.

[13] M Delbridge. 2015. Motion Capture in Performance: An Introduction. In *MOTION CAPTURE IN PERFORMANCE: AN INTRODUCTION*. 1–94. https://doi.org/10.1057/9781137505811

[14] Kosmas Dimitropoulos, Sotiris Manitsaris, Filareti Tsalakanidou, Spiros Nikolopoulos, Bruce Denby, Samer Al Kork, Lise Crevier-Buchman, Claire Pillot-Loiseau, Martine Adda-Decker, Stephane Dupont, et al. 2014. Capturing the intangible an introduction to the i-Treasures project. In *2014 international conference on computer vision theory and applications (VISAPP)*, Vol. 2. IEEE, 773–781.

[15] Kosmas Dimitropoulos, Sotiris Manitsaris, Filareti Tsalakanidou, Spiros Nikolopoulos, Bruce Denby, Samer Al Kork, Lise Crevier-Buchman, Claire Pillot-Loiseau, Martine Adda-Decker, Stephane Dupont, et al. 2014. Capturing the intangible an introduction to the i-Treasures project. In *2014 international conference on computer vision theory and applications (VISAPP)*, Vol. 2. IEEE, 773–781.

[16] Kosmas Dimitropoulos, Filareti Tsalakanidou, Spiros Nikolopoulos, Ioannis Kompatsiaris, Nikos Grammalidis, Sotiris Manitsaris, Bruce Denby, Lise Crevier-Buchman, Stephane Dupont, Vasileios Charisis, et al. 2018. A multimodal approach for the safeguarding and transmission of intangible cultural heritage: The case of i-Treasures. *IEEE intelligent systems* 33, 6 (2018), 3–16.

[17] DMCD. 2019. Dance Motion Capture Database. http://dancedb.cs.ucy.ac.cy/.

[18] Anastasios D Doulamis, Athanasios Voulodimos, Nikolaos D Doulamis, Sofia Soile, and Anastasios Lampropoulos. 2017. Transforming Intangible Folkloric Performing Arts into Tangible Choreographic Digital Objects: The Terpsichore Approach.. In *VISIGRAPP (5: VISAPP)*. 451–460.

[19] Katerina El Raheb and Yannis Ioannidis. 2013. Dance in the world of data and objects. In *International Conference on Information Technologies for Performing Arts, Media Access, and Entertainment*. Springer, 192–204.

[20] Marco Gillies and Daniel Ballin. 2004. Integrating autonomous behavior and user control for believable agents. (2004).

[21] Nikolaos Grammalidis, Kosmas Dimitropoulos, Filareti Tsalakanidou, Alexandros Kitsikidis, Pierre Roussel, Bruce Denby, Patrick Chawah, Lise Buchman, Stèphane Dupont, Sohaib Laraba, et al. 2016. The i-treasures intangible cultural heritage dataset. In *Proceedings of the 3rd International Symposium on Movement and Computing*. 1–8.

[22] Athina Grammatikopoulou, Sohaib Laraba, Ozan Sahbenderoglu, Kosmas Dimitropoulos, Stella Douka, and N Grammalidis. 2019. An adaptive framework for the creation of exergames for intangible cultural heritage (ICH) education. *Journal of Computers in Education* 6, 3 (2019), 417–450.

[23] Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014. Does gamification work?–a literature review of empirical studies on gamification. In *2014 47th Hawaii international conference on system sciences*. Ieee, 3025–3034.

[24] Mathieu Hilgers and Éric Mangez. 2014. Bourdieu's theory of social fields: Concepts and applications. (2014).

[25] Kristina Höök, Martin P Jonsson, Anna Ståhl, and Johanna Mercurio. 2016. Somaesthetic appreciation design. In *Proceedings of the 2016 chi conference on human factors in computing systems*. 3131–3142.

[26] Laurel Horton and Paul Jordan-Smith. 2004. Deciphering folk costume: Dress codes among contra dancers. *Journal of american folklore* 117, 466 (2004), 415–440.

[27] Scott Hotton and Jeff Yoshimi. 2010. The dynamics of embodied cognition. *International Journal of Bifurcation and Chaos* 20, 04 (2010), 943–972.

[28] Stacy Hsueh, Sarah Fdili Alaoui, and Wendy E Mackay. 2019. Understanding kinaesthetic creativity in dance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[29] Rajkumar Kannan, Frederic Andres, and Christian Guetl. 2010. DanVideo: an MPEG-7 authoring and retrieval system for dance videos. *Multimedia Tools and Applications* 46 (2010), 545–572.

[30] Vicky Karkou, Sophia Bakogianni, and Evangelia Kavakli. 2008. Traditional dance, pedagogy and technology: an overview of the WebDANCE project. *Research in Dance Education* 9, 2 (2008), 163–186.

[31] Sarah Kenderdine, Lily Hibberd, Jeffrey Shaw, Tsong-Zung Chang, and Yumeng Hou. 2023. *Archery Rites: Remaking Confucian Rites*. Springer Nature Singapore, Singapore, 249–278. https://doi.org/10.1007/978-981-16-8321-3_13

[32] Peter Khooshabeh, Morteza Dehghani, Angela Nazarian, and Jonathan Gratch. 2017. The cultural influence model: When accented natural language spoken by virtual characters matters. *AI & society* 32 (2017), 9–16.

[33] Iris Kico, Milan Dolezal, Nikos Grammalidis, and Fotis Liarokapis. 2020. Visualization of folk-dances in virtual reality environments. In *Strategic Innovative Marketing and Tourism: 8th ICSIMAT, Northern Aegean, Greece, 2019*. Springer, 51–59.

[34] Iris Kico, Nikos Grammalidis, Yiannis Christidis, and Fotis Liarokapis. 2018. Digitization and visualization of folk dances in cultural heritage: a review. *Inventions* 3, 4 (2018), 72.

[35] Eugenia S Kim. 2011. ChoreoSave: A digital dance preservation system prototype. *Proceedings of the American Society for Information Science and Technology* 48, 1 (2011), 1–10.

[36] Barbara Kirshenblatt-Gimblett. 2014. Intangible heritage as metacultural production. *Museum International* 66, 1-4 (2014), 163–174.

[37] Alexandros Kitsikidis, Kosmas Dimitropoulos, Stella Douka, and Nikos Grammalidis. 2014. Dance analysis using multiple Kinect sensors. In *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, Vol. 2. 789–795.

[38] Alexandros Kitsikidis, Kosmas Dimitropoulos, Deniz Uğurca, Can Bayçay, Erdal Yilmaz, Filareti Tsalakanidou, Stella Douka, and Nikos Grammalidis. 2015. A game-like application for dance learning using a natural human computer interface. In *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being: 9th International Conference, UAHCI 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part III 9*. Springer, 472–482.

[39] Sohaib Laraba and Joëlle Tilmanne. 2016. Dance performance evaluation using hidden Markov models. *Computer Animation and Virtual Worlds* 27, 3-4 (2016), 321–329.

[40] Chen Liu, Anna Wang, Chunguang Bu, Wenhui Wang, and Haijing Sun. 2021. Human Motion Tracking with Less Constraint of Initial Posture from a Single RGB-D Sensor. *Sensors (Basel, Switzerland)* 21 (2021). https://doi.org/10.3390/s21093029

[41] Mariana Lozada and Natalia Carro. 2016. Embodied action improves cognition in children: Evidence from a study based on Piagetian conservation tasks. *Frontiers in Psychology* 7 (2016), 393.

[42] Qiduo Lu, Chenhao Zhu, and Xia Ye. 2022. Research on Multimodal Summarization by Integrating Visual and Text Modal Information. In *2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*. IEEE, 882–889.

[43] Elise J Matthews, Laurie Clune, Florence Luhanga, and Reny Loewen. 2021. The impact of cultural immersion international learning experiences on cultural competence of nursing students: A critical integrative review. *Journal of Professional Nursing* 37, 5 (2021), 875–884.

[44] Folk Dance Federation of California. 2024. Welcome to the Folk Dance Federation of California. https://www.folkdance.com/

[45] Shaohua Pan, Qi Ma, Xinyu Yi, Weifeng Hu, Xiong Wang, Xingkang Zhou, Jijunnan Li, and Feng Xu. 2023. Fusing monocular images and sparse imu signals for real-time human motion capture. In *SIGGRAPH Asia 2023 Conference Papers*. 1–11.

[46] Giani Petri, Christiane Gresse von Wangenheim, and Adriano Borgatto. 2016. MEEGA+: An Evolution of a Model for the Evaluation of Educational Games. (07 2016).

[47] Eftychios Protopapadakis, Athina Grammatikopoulou, Anastasios Doulamis, and Nikos Grammalidis. 2017. Folk dance pattern recognition over depth images acquired via kinect sensor. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42 (2017), 587–593.

[48] Ioannis Rallis, Nikolaos Doulamis, Anastasios Doulamis, Athanasios Voulodimos, and Vassilios Vescoukis. 2018. Spatio-temporal summarization of dance choreographies. *Computers & Graphics* 73 (2018), 88–101.

[49] MR Reshma, B Kannan, VP Jagathy Raj, and S Shailesh. 2023. Cultural heritage preservation through dance digitization: A review. *Digital Applications in Archaeology and Cultural Heritage* (2023), e00257.

[50] Anna Rizzo, Katerina El Raheb, Sarah Whatley, Rosemary E Kostic Cisneros, Massimiliano Zanoni, Antonio Camurri, Vladimir Viro, Jean-Marc Matos, Stefano Piana, Michele Buccoli, et al. 2018. WhoLoDancE: whole-body interaction learning for dance education. In *CEUR Workshop Proceedings*. CEUR Workshop Proceedings, 41–50.

[51] João Paulo Sousa, Patricia Cordeiro, Rogério Tavares, and João Victor Gomide. 2023. Dancing Into the Digital Age: Experimenting the Digitization of the Pauliteiros Folk Dances. In *2023 18th Iberian Conference on Information Systems and Technologies (CISTI)*. 1–6. https://doi.org/10.23919/CISTI58278.2023.10211933

[52] Efstathios Stavrakis, Andreas Aristidou, Maria Savva, Stephania Loizidou Himona, and Yiorgos Chrysanthou. 2012. Digitization of Cypriot Folk Dances. In *Progress in Cultural Heritage Preservation*, Marinos Ioannides, Dieter Fritsch, Johanna Leissner, Rob Davies, Fabio Remondino, and Rossella Caffo (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 404–413.

[53] Efstathios Stavrakis, Andreas Aristidou, Maria Savva, Stephania Loizidou Himona, and Yiorgos Chrysanthou. 2012. Digitization of Cypriot Folk Dances. In *Proceedings of the 4th International Conference on Progress in Cultural Heritage Preservation* (Limassol, Cyprus) *(EuroMed'12)*. Springer-Verlag, Berlin, Heidelberg, 404–413. https://doi.org/10.1007/978-3-642-34234-9_41

[54] Z. Su, Lan Xu, Dawei Zhong, Zhong Li, Fan Deng, Shuxue Quan, and Lu Fang. 2021. RobustFusion: Robust Volumetric Performance Reconstruction Under Human-Object Interactions From Monocular RGBD Stream. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45 (2021), 6196–6213. https://doi.org/10.1109/TPAMI.2022.3215746

[55] Tsz Wing Tsang and Hui Jing Lu. 2022. Hand movement improves word memory of Grade 1 students. *International Journal of School & Educational Psychology* 10, 3 (2022), 408–417.

[56] June A Vail. 2003. Staging" Sweden": a typology for folk dance in performance. *Scandinavian Studies* 75, 1 (2003), 89–102.

[57] Rémi A van Compernolle and Lawrence Williams. 2013. Group dynamics in the language classroom: embodied participation as active reception in the collective Zone of Proximal Development. *Classroom Discourse* 4, 1 (2013), 42–62.

[58] Doug Van Nort. 2015. [radical] signals from life: from muscle sensing to embodied machine listening/learning within a large-scale performance piece. In *Proceedings of the 2nd International Workshop on Movement and Computing*. 124–127.

[59] Donald W. Winnicott. 2016. 429The Location of Cultural Experience. In *The Collected Works of D. W. Winnicott: Volume 7, 1964 - 1966*. Oxford University Press. https://doi.org/10.1093/med:psych/9780190271398.003.0075

[60] Christoph Wulf. 2020. Performativity and dynamics of intangible cultural heritage. In *Ritual, Heritage and Identity*. Routledge India, 76–94.

[61] Ping Yang. 2018. Experiential learning opportunities: Removing language barriers and maximizing cultural immersion. In *Study abroad contexts for enhanced foreign language learning*. IGI Global, 189–217.

[62] Yang Yang, Howard Leung, Lihua Yue, and Liqun Deng. 2013. Generating a two-phase lesson for guiding beginners to learn basic dance movements. *Computers & Education* 61 (2013), 1–20.

[63] Peng Ye and Yaolin Zhou. 2013. The framework of chinese intangible cultural heritages metadata standards. In *2013 International Workshop on Computer Science in Sports*. Atlantis Press, 221–225.

[64] Qiushi Zhou, Cheng Cheng Chua, Jarrod Knibbe, Jorge Goncalves, and Eduardo Velloso. 2021. Dance and Choreography in HCI: A Two-Decade Retrospective. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Yokohama</city>, <country>Japan</country>, </conf-loc>) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 262, 14 pages. https://doi.org/10.1145/3411764.3445804