EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2024)
M. Corsini, D. Ferdani, A. Kuijper, and H. Kutlu (Editors)

Short Paper

Documenting Dunhuang Dance Using Motion Capture Technology

Zeyu Wang[1], Chengan He[3], Zhe Yan[1], Jiashun Wang[4], Yingke Wang[5], Junhua Liu[6], Angela Shen[7], Mengying Zeng[2], Holly Rushmeier[3], Huazhe Xu[8], Borou Yu[2], Chenchen Lu[2], Eugene Y. Wang[2]

[1]The Hong Kong University of Science and Technology (Guangzhou), China, [2]Harvard University, Cambridge, USA, [3]Yale University, New Haven, USA, [4]Carnegie Mellon University, Pittsburgh, USA, [5]Stanford University, Stanford, USA, [6]The Chinese University of Hong Kong, Shenzhen, China, [7]Northeastern University, Boston, USA, [8]Tsinghua University, Beijing, China

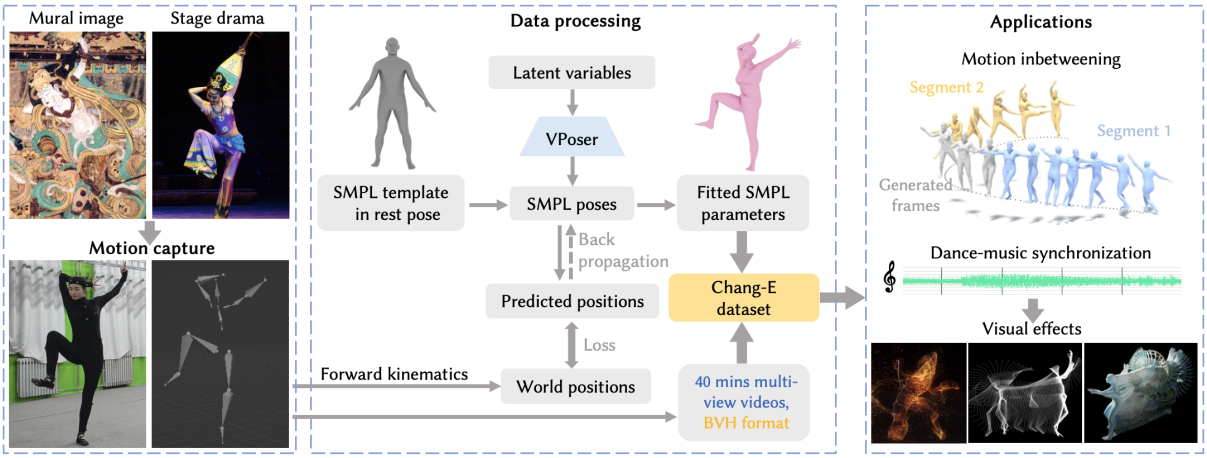


Figure 1: *The generation process of the Chang-E dataset and its applications. We captured dance artists performing Dunhuang dance using Motion Capture technology, fitted the 3D motion data into SMPL format, and further created reinterpretations and visual effects based on the dataset. The mural image is from Mogao Cave 112 in Dunhuang. The stage drama image is from Silk Road, Flower, and Rain.*

Abstract

This paper introduces a pipeline for documenting Chinese Classical Dunhuang Dance using motion capture technology. Our captured dataset includes full-body movements documented across eight categories, totaling 40 minutes of professional dance (preview available at https://cislab.hkust-gz.edu.cn/projects/chang-e/). The dataset supports creative applications for Dunhuang dance culture, showcased in a new media immersive exhibition. We used motion inbetweening to concatenate dance sequences and synchronized them with music through retiming techniques, enhancing rhythm and harmony. Visual effects were applied to digital dancers, achieving visually appealing results that echo Buddhist meditation and bodily cognition. The Chang-E dataset enables digital preservation and creative reimagination of Dunhuang dance, offering high-quality data and an interdisciplinary collaboration framework for future graphics and cultural heritage research.

CCS Concepts

• ***Computing methodologies*** → ***Animation;*** • ***Applied computing*** → ***Arts and humanities;***

1. Introduction

The Mogao Caves in Dunhuang, Gansu province, China, are a world-renowned UNESCO Cultural Heritage Site, featuring murals that depict Chinese classical Dunhuang dance. Digitally preserving and reinterpreting Dunhuang dance has been challenging due to the need for expertise in both art history and computer graphics. We introduce the first open comprehensive motion capture dataset of current Dunhuang dance, *Chang-E*, which includes skeleton data, body mesh, and multi-view videos, enabling diverse applications.

Constructed between the 4th and 14th centuries, the Mogao Caves are adorned with murals and sculptures representing a unique Buddhist artistic achievement. These murals depict dancers

DOI: 10.2312/gch.20241250

with fluid, angular, and curved movements and distinctive gestures, influenced by the folk dances of Eurasia along the Silk Road. Integral elements such as long sashes and musical instruments, including the Pipa, Drum, and Flute, are seamlessly woven into the dance motions. Since the 1979 dance drama *Silk Road, Flower, and Rain*, scholars have endeavored to reconstruct and reinterpret Dunhuang dance, establishing the well-known Dunhuang dance style.

Documenting and preserving dance is a long-studied endeavor. Dance research, creation, and performance are taught and communicated orally and bodily from one generation to the next. While many motion capture datasets exist in the computer graphics community, dance-related content is scarce, especially for Chinese classical dance. In this paper, we present our interdisciplinary effort on building the first Dunhuang dance motion capture dataset in the world, *Chang-E*. Our contributions include:

- A Dunhuang dance motion capture dataset with full-body performances by professional dancers in 8 categories, totaling 40 minutes of high-quality data. This dataset is compatible with other motion capture datasets and can be used for various applications, including training neural networks for new dance genres.
- A technical pipeline using motion inbetweening and retiming methods to concatenate dance segments and synchronize them with background music, enabling the visualization of digital dancers and special effects.
- An interdisciplinary framework between dance and computing communities, leading to a new media exhibition that allows visitors to experience Dunhuang dance in immersive environments, promoting the preservation of intangible cultural heritage.

2. Related Work

For intangible cultural heritage like dance, digital preservation becomes increasingly important. Dance digitization began in the early 20th century with pictures and videos. Zhanna [Zha20] visually reconstructed Kazakh folk dance to preserve cultural heritage. Dhanapalan [Dha18] used a 16-camera MoCap system to digitize Kathakali dance from India. Aristidou et al. [ASC*15] combined MoCap and Laban Movement Analysis (LMA) to identify dance style qualities, aiding motion comparison and evaluation.

As one of the most popular technologies for recording human movements, motion capture has been utilized by a growing number of researchers to collect dance data. Human3.6M [IPOS13] contains 3.6 million 3D articulated poses from professional actors. The Human Motion Database [GFB12] includes data from 10 subjects performing 17 activities. The KIT Whole-Body Human Motion Database [Chr16] includes motions, anthropometric measurements, and scene setups. Recent datasets like DNA-Rendering [CCF*23]offer extensive annotations and large-scale captures. Besides, there are specific activity datasets including CMU-MMAC [DlTHB*08] for cooking tasks and a karate dataset by Szczęsna et al. [SBP21] for analyzing sports techniques.

There have been efforts to create dance databases [CTL*21, LZZ*23, ANBH23]. The AIST Dance Video Database [TFHG19] offers 13,939 street dance videos. AMASS [MGT*19] consolidates 15 MoCap datasets into a unified framework, including over 300 subjects and 11,000 motions.

However, despite the contribution of these datasets, there are very few datasets containing Chinese dance. The National University of Singapore's database [YYK*23] includes minimal Chinese dance data. There is still a notable gap in datasets focusing on traditional Chinese dance, particularly Dunhuang dance.

3. Data Collection

Due to the high professionalism of Dunhuang dance, we invited a team from China Dance Academy, which is one of the highest institutions for Chinese dance education and assessment. The team included a Dunhuang dance professor/choreographer with over 40 years of experience, and four professional dancers (two females, two males) with over 10 years of experience.

3.1. Motion Capture System Deployment

We defined a capture volume of $7m \times 6m \times 2.8m$ to accommodate rapid circular spins and dynamic movements with props like long silk. Our motion capture setup consisted of 12 high-speed cameras equipped with $6mm$ lenses, designed to capture full-body movements. These cameras were mounted on a ceiling-mounted steel frame around the capture volume (Fig. 2a).

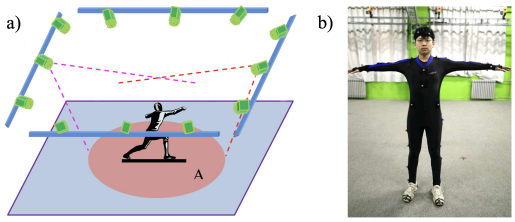


Figure 2: *Motion capture environment. a) Diagram of the camera installation layout in our MoCap system. The motion capture cameras (green) are evenly distributed on the support frame (blue). The circular area A on the ground represents the effective activity space for motion capture. b) T-pose for system initialization.*

During the dance performance, the dancers wore specialized motion capture suits with 38 reflective markers attached. The placement of these markers included four around the head, one on the neck, one on each shoulder, two on the chest, one on the back, four around the pelvis, and three on each arm, hand, leg, and foot.

As depicted in the murals, performers usually wear long skirts, drape silk ribbons, and use props, but motion capture systems can only track body markers, not props. Flowing silk can occlude markers, resulting in incomplete data. Therefore, for motion capture, dancers only wore close-fitting garments, though their movements mirrored those in traditional costumes. We also recorded some segments with props for reference.

3.2. Motion Capture Process

We first performed camera calibration and marker point calibration for the MoCap system. We used a T-shaped calibration pole to track three points on the pole's movement trajectories, with each of the 12 cameras capturing around 3,000 evenly distributed points within

the motion capture area. These points helped minimize 3D point reprojection error and optimize the intrinsic and extrinsic camera parameters concurrently. Thus we established a fixed coordinate system within the capture volume.

During motion capture, dancers started with a T-Pose (Fig. 2b), followed by automatic capture and real-time 3D reconstruction of feature points. In the 3D view, the initial frame was matched with the marker point model, and subsequent frames were tracked to maintain marker consistency over time, enhanced by machine learning algorithms.

Captured data had artifacts, such as missing or inaccurate feature points, requiring post-processing. We assessed data quality and manually adjusted marker mapping where needed. To address shaky points, we applied a smoothing algorithm, greatly enhancing the accuracy of the captured motion.

4. Data Processing

Our motion capture produced human skeleton data in BVH format, which we then used to generate mesh geometry. Some previous methods estimate the 3D body shape and deformation using existing MoCap marker sets, such as MoSh++ [MGT*19]. In this section, we present a novel method to convert BVH motion data to the Skinned Multi-Person Linear Model (SMPL) [LMR*15]. In the context of Dunhuang dance, the SMPL format offers several advantages including continuous mesh deformation and ease of integration with deep learning frameworks. This establishes a solid foundation for facilitating future research and creative applications based on the captured data.

4.1. BVH to SMPL

The procedure of converting Dunhuang dance motion capture data from BVH to SMPL parameters involves unifying the coordinates of each joint in both formats into the global coordinate system and using machine learning methods to obtain the poses and translation parameters in SMPL.

We first employ forward kinematics (FK) to convert local joint transformations into world coordinates for a hierarchy of joints. Each joint's local rotation is converted from Euler angles to quaternions for stability and efficiency. The global rotation of a joint is computed by multiplying its parent's global rotation with its local rotation. Using this global rotation and a positional offset, the global position of each joint is determined. This recursive method starts from the root joint and moves to the leaf joints, computing world coordinates from local definitions.

SMPL has 24 joints while BVH has 19 joints and 5 end-sites. We manually picked 21 matching joint pairs between BVH and SMPL, ensuring a one-to-one semantic correlation for fitting the SMPL parameters. We employ an optimization process to convert captured BVH keypoints to SMPL parameters (Fig. 1). The parameters include root orientation $\vec{\varphi} \in \mathbb{R}^6$, root translation $\mathbf{t} \in \mathbb{R}^3$, body pose $\vec{\omega}$, and a scaling factor s. The body poses $\vec{\omega} \in \mathbb{R}^{23\times 3}$ denotes the axis angles for 23 joints corresponding to their parent joints. $\vec{\omega} = \mathcal{V}(\vec{z})$ is decoded using a fixed Variational Human Body Pose Prior (VPoser) $\mathcal{V}$ [PCG*19]. $\vec{z} \in \mathbb{R}^{32}$ is the latent space of VAE initialized as the input of VPoser. The SMPL model takes the root orientation and body poses as a whole of full body poses $\vec{\theta} = [\mathcal{M}(\vec{\varphi}), \mathcal{V}(\vec{z})]^T \in \mathbb{R}^{24\times 3}$, where $\mathcal{M} : \mathbb{R}^6 \mapsto \mathbb{R}^3$ is a function that maps 6D rotations to axis angles. Hence, we take the parameters $\Phi = \{\vec{\theta}, \mathbf{t}, s\}$ as input. By performing the forward pass of the SMPL model with a fixed $\vec{\beta} = \vec{0}$, we get the estimated joints' global positions $\hat{\mathbf{J}}$:

$$\hat{\mathbf{J}} = \mathrm{SMPL}(\vec{\theta}, \mathbf{t}, s). \tag{1}$$

We build an index permutation function P mapping the estimated joint to the corresponding captured keypoint, and perform the optimization by minimizing an MSE loss evaluated every N frames:

$$\arg\min_{\{\Phi_i\}_{i=1}^N} \frac{1}{N\times K} \sum_{t=1}^{N} \|P(\hat{\mathbf{J}}_t) - \mathbf{J}_t\|_2^2. \tag{2}$$

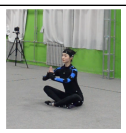
We compute the gradients with PyTorch and minimize the loss with the SGD optimizer. The optimization process consists of 3 stages, each with a specific learning rate schedule for different parameters. For the parameter $\vec{\varphi}$, its learning rate of 3 stages is set to 0, 0.1, and 0.1, respectively. For the parameter $\vec{z}$, its learning rate of 3 stages is set to 0, 0, and 0.1, respectively. For the parameter $\mathbf{t}$, its learning rate is set to 100 throughout the 3 stages. For the parameter s, its learning rate is set to 1 throughout the 3 stages. An optimization loop is executed for the specified number of iterations for each stage (100, 300, and 10,000 times respectively). After the optimization loop, parameter values are extracted and stored as a `.npz` file in the format aligning with AMASS, thus can be read directly using the SMPL add-on in Blender.

4.2. Post-processing

To correct inaccuracies in the perceived ground height and foot positions in 3D character animation using SMPL, we first calculated toe velocities as the finite difference between consecutive frames and extracted the y-coordinates of the left toe, right toe, and root joint. Frames with toe velocities exceeding a predefined threshold were excluded to filter out airborne foot positions. Using DBSCAN clustering, we grouped similar static foot height values and calculated the median height within each cluster. The smallest median foot height was adjusted by an offset to estimate the ground height, which was then subtracted from all joints' y-coordinates to align the SMPL model with the ground, ensuring accurate ground contact and correcting foot positioning errors.

We also applied other post-processing steps to correct errors that may have been introduced during the conversion process. We first identify problematic frames that typically contain erroneous or noisy data, unusual subject movement, or conversion issues. Once identified, these frames are corrected using spherical linear interpolation (Slerp) [Sho85]. Simultaneously, root translation is linearly interpolated over the same frames. We also transformed all the data to a z-up coordinate system, thus aligning the data with the convention in AMASS. Finally, we eliminated idle frames where no significant action occurs from the final data.

Table 1: *Sample pose and associated information of eight dance categories in the Chang-E dataset.*

Dance	 *Flying Apsaras*	 *Lotus Steps*	 *Thirty-Six Postures*	 *Thirty-Six Postures*	 *Revelation Meditation*	 *Sogdian Whirl (2 sequences)*	 *Playing the Pipa behind the Back (2 sequences)*	 *Lei Gong Drum*
Duration	6 min 32 s	3 min 51 s	5 min 29s	4 min 31s	3 min 24 s	1 min 18 s & 1 min 20 s	2 min 8 s & 1 min 55 s	4 min 51 s
Dancer(s)	1 female	1 female	2 females	2 males	1 female & 1 male	1 male	1 male	2 males

4.3. Dataset Content

We recorded video clips at 30 frames per second (FPS) with accompanying background music and simultaneously captured raw motion data in the BVH format. And we obtained the SMPL format data through data processing. The duration of these clips ranged from 1 to 6 minutes. Eight distinct Dunhuang dance themes (Table 1) were selected for our study including *Flying Apsaras*, *Ji Yue Tian's Thirty-Six Postures* (including male and female versions), *Lei Gong Drum*, *Revelation Meditation* (including male and female versions), *Lotus Steps*, *Sogdian Whirl* (2 different sequences) and *Playing the Pipa Behind the Back* (2 sequences). To capture a complete view of dance movements, we positioned 4 video cameras around the dancer and any body parts obscured from one camera's perspective could still be captured from another viewpoint. All the recorded videos and the captured BVH data were synchronized for consistency. Following the process of data alignment and removal of irrelevant segments, the total duration of the dance sequences reached approximately 40 minutes. Our final dataset contains the dance sequences with 4 camera views, BVH format, converted SMPL format, and rendered SMPL video (Fig. 3). The captured movements cover all the basic classic postures of contemporary Dunhuang dance, including standing, floor work, jumping, spinning, and continuous movements.

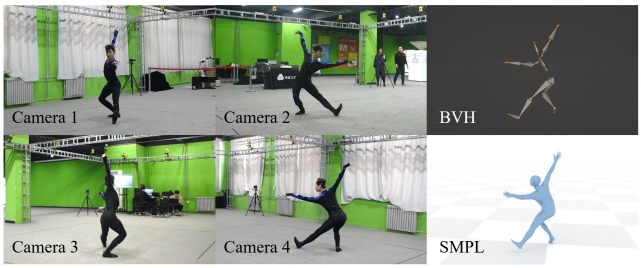


Figure 3: *Sogdian Whirl in 4 camera views, recorded skeleton data in BVH format and generated body mesh in SMPL format.*

5. Applications

Dunhuang dance, inspired by the flying apsaras in ancient murals, is an evolving art form that blends historical and sociocultural elements. Our motion capture dataset *Chang-E* enables artists to reimagine Dunhuang dance with new visual elements, including synchronized music and particle effects.

5.1. Dancing in Heaven

Dunhuang's imagery of flying apsaras captures the delicate balance of grace and strength, transcending physical limitations. In the drama *Flying Apsaras*, dancers use lifting platforms and "iron boots" to shift their center of gravity, allowing for extraordinary leaning and flying poses.

We aim to create complex movements and flying poses that are challenging for human dancers. Our approach extracts two segments in different dance themes from the *Chang-E* dataset. Then we use Slerp [Sho85] to seamlessly connect two sequences by generating intermediate poses, ensuring smooth transitions between distinct motion sequences.

We manipulated the digital avatar to follow a user-specified trajectory through the integrated sequences. This trajectory governs both the position and orientation of interpolated poses, allowing the digital character to perform dance moves naturally along a predefined path. To maintain the authenticity of the dance, we fixed the root orientations of specific frames, ensuring that key parts of the performance, such as spinning, retain their original postures.

5.2. Dancing to the Music

Music and dance are intrinsically linked through rhythm, involving the temporal distribution of events like musical sounds or bodily movements. The Dunhuang murals depict dancers with instruments such as the pipa, guqin, sheng, and drums, synchronizing movements with Chinese classical music to evoke beauty and meaning. We used the dance retargeting algorithm by Davis and Agrawala [DA18] to synchronize edited Dunhuang dance with background music. The steps include: extracting visual beats from the source video and musical beats from the target audio, using a dynamic programming algorithm to align these beats, applying a time-warping function to synchronize the video's timeline with the audio's, and computing the duration and repetition for each frame in the warped timeline based on the target audio. Finally, we rearranged poses in the motion sequence according to the synchronized

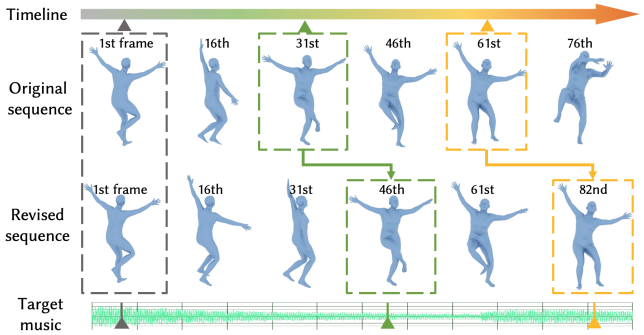


Figure 4: *Dance-music synchronization. The original sequence has 76 frames and we showcase every 15th frame (i.e., frames 1, 16, 31, 46, etc.). The original video is warped to a longer duration video to match the beat of the target music.*

frames, creating a new array of synchronized poses aligning with the audio rhythm.

5.3. Visual Effects

We aimed to present the beauty of Dunhuang dance vividly by integrating visual effects with our dataset. By utilizing various platforms, we combined digitized Dunhuang dance with modern VFX technology to enhance its aesthetic appeal.

On Unity, we imported BVH data as stick figures and added particle effects, creating dynamic visuals where flames danced (Fig. 5a) and water flowed (Fig. 5b) with the dancer's movements. On the Embergen platform, we overlaid a Bodhisattva-shaped skin onto the digital character (Fig. 5c) and added smoke effects, giving the appearance of smoke dispersing as the dancer moved (Fig. 5d). Using OpenFrameworks, we connected BVH joints with lines, creating a stick-figure motion overlay that enhanced the dance's dynamism (Fig. 5e). On Houdini, the digital figure in SMPL format was rendered with realistic motion and colors, closely mirroring Dunhuang mural dancers (Fig. 5f).

These visualizations not only enrich the public's experience of Dunhuang dance but also showcase the potential of merging traditional art with modern digital technology.

5.4. Cave Dance Exhibition

We launched the Cave Dance exhibition in immersive environments at Harvard University in the US (Fig. 6) and in Shanghai, China. The Harvard exhibition lasted over a year, attracting thousands of visitors. This exhibition offers an unconventional new media experience, inviting audiences into a dynamic universe of Dunhuang dance and Buddhist culture. It allows the exploration of themes like corporeality, life, and spiritual transcendence, creating a vibrant environment where tradition and technology coalesce.

6. Conclusion

This paper introduces a pipeline for documenting Chinese Classical Dunhuang Dance using motion capture technology. The resulting dataset *Chang-E* is the first open comprehensive motion capture dataset of Dunhuang dance. The dataset includes motion data performed by professional dancers, available in multiview video, BVH, and SMPL body mesh formats. We employed motion inbetweening, dance-music synchronization, and visual effects to create digital reenactments of ancient Dunhuang dance murals.

Our work has the following limitations. First, the passive marker-based MoCap system necessitates dancers to wear a customized suit, preventing the capture of traditional costumes and props. Second, we used Slerp for motion inbetweening, but advanced methods such as deep learning-based motion synthesis could generate more realistic sequences. Additionally, our dataset lacks hand gestures, which are critical in Dunhuang dance. Future work will explore custom hand-tracking systems and parametric models.

In conclusion, *Chang-E* revitalizes the ancient dance forms depicted in the Dunhuang Caves through modern technology, fostering interdisciplinary collaboration in cultural heritage research. This open dataset aims to contribute to art history, new media art, and fields in computer graphics like AI generative methods, dynamic 3D reconstruction, digital humans, and character animation.

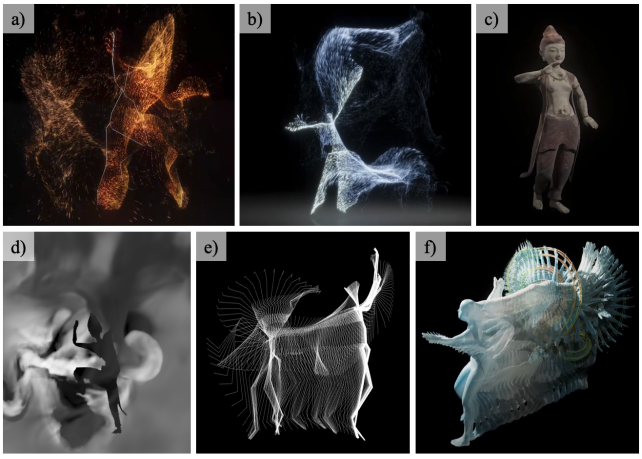


Figure 5: *Visual effects showcase. a) and b) Particle effects produced by Unity such as dancing flame and flowing water. c) and d) Bodhisattva-skinned digital character produced by Embergen. e) Dance sequences produced by Open Frameworks. f) Dance sequences produced by Houdini.*

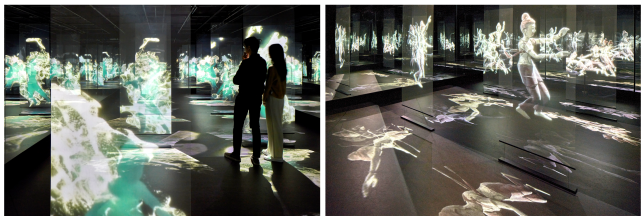


Figure 6: *Audience immersed in the Cave Dance exhibition. The colors, light and shadow, and spatial atmosphere in the Dunhuang caves are presented through the graceful dance movements of digital dancers and the use of massive visual effects.*

References

[ANBH23] ALEXANDERSON S., NAGY R., BESKOW J., HENTER G. E.: Listen, denoise, action! audio-driven motion synthesis with diffusion models. *ACM Trans. Graph. 42*, 4 (Jul 2023). doi:10.1145/3592458. 2

[ASC*15] ARISTIDOU A., STAVRAKIS E., CHARALAMBOUS P., CHRYSANTHOU Y., HIMONA S. L.: Folk Dance Evaluation Using Laban Movement Analysis. *J. Comput. Cult. Herit. 8*, 4 (Aug 2015). URL: https://doi.org/10.1145/2755566, doi:10.1145/2755566. 2

[CCF*23] CHENG W., CHEN R., FAN S., YIN W., CHEN K., CAI Z., WANG J., GAO Y., YU Z., LIN Z., ET AL.: Dna-rendering: A diverse neural actor repository for high-fidelity human-centric rendering. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)* (Los Alamitos, CA, USA, Oct 2023), IEEE Computer Society, pp. 19925–19936. doi:10.1109/ICCV51070.2023.01829. 2

[Chr16] CHRISTIAN MANDERY AND ÖMER TERLEMEZ AND MARTIN DO AND NIKOLAUS VAHRENKAMP AND TAMIM ASFOUR: Unifying representations and large-scale whole-body motion databases for studying human motion. *IEEE Transactions on Robotics 32*, 4 (2016), 796–809. 2

[CTL*21] CHEN K., TAN Z., LEI J., ZHANG S.-H., GUO Y.-C., ZHANG W., HU S.-M.: Choreomaster: Choreography-oriented music-driven dance synthesis. *ACM Trans. Graph. 40*, 4 (Jul 2021). doi:10.1145/3450626.3459932. 2

[DA18] DAVIS A., AGRAWALA M.: Visual Rhythm and Beat. *ACM Trans. Graph. 37*, 4 (Jul 2018). doi:10.1145/3197517.3201371. 4

[Dha18] DHANAPALAN B.: Kathakali and motion capture: An experimental dialogue between indian classical dance and technology. *The International Journal of New Media, Technology and the Arts 13* (Jan 2018), 15–24. doi:10.18848/2326-9987/CGP/v13i01/15-24. 2

[DlTHB*08] DE LA TORRE F., HODGINS J., BARGTEIL A., MARTIN X., MACEY J., COLLADO A., BELTRAN P.: *Guide to the Carnegie Mellon University Multimodal Activity (CMU-MMAC) Database*. Tech. Rep. CMU-RI-TR-08-22, Pittsburgh, PA, Apr 2008. 2

[GFB12] GUERRA-FILHO G., BISWAS A.: The Human Motion Database: A Cognitive and Parametric Sampling of Human Motion. *Image Vision Comput. 30*, 3 (Mar 2012), 251–261. doi:10.1016/j.imavis.2011.12.002. 2

[IPOS13] IONESCU C., PAPAVA D., OLARU V., SMINCHISESCU C.: Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence 36* (Dec 2013). doi:10.1109/TPAMI.2013.248. 2

[LMR*15] LOPER M., MAHMOOD N., ROMERO J., PONS-MOLL G., BLACK M. J.: SMPL: A Skinned Multi-Person Linear Model. *ACM Trans. Graph. 34*, 6 (Oct 2015). doi:10.1145/2816795.2818013. 3

[LZZ*23] LI R., ZHAO J., ZHANG Y., SU M., REN Z., ZHANG H., TANG Y., LI X.: Finedance: A fine-grained choreography dataset for 3d full body dance generation. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)* (Los Alamitos, CA, USA, Oct 2023), IEEE Computer Society, pp. 10200–10209. doi:10.1109/ICCV51070.2023.00939. 2

[MGT*19] MAHMOOD N., GHORBANI N., TROJE N. F., PONS-MOLL G., BLACK M.: Amass: Archive of motion capture as surface shapes. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (2019), pp. 5441–5450. doi:10.1109/ICCV.2019.00554. 2, 3

[PCG*19] PAVLAKOS G., CHOUTAS V., GHORBANI N., BOLKART T., OSMAN A. A. A., TZIONAS D., BLACK M. J.: Expressive Body Capture: 3D Hands, Face, and Body from a Single Image. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 10975–10985. 3

[SBP21] SZCZĘSNA A., BŁASZCZYSZYN M., PAWLYTA M.: Optical Motion Capture Dataset of Selected Techniques in Beginner and Advanced Kyokushin Karate Athletes. *Scientific Data 8* (Jan 2021). doi:10.1038/s41597-021-00801-5. 2

[Sho85] SHOEMAKE K.: Animating Rotation with Quaternion Curves. In *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1985), SIGGRAPH '85, Association for Computing Machinery, p. 245–254. doi:10.1145/325334.325242. 3, 4

[TFHG19] TSUCHIDA S., FUKAYAMA S., HAMASAKI M., GOTO M.: AIST Dance Video Database: Multi-Genre, Multi-Dancer, and Multi-Camera Database for Dance Information Processing. In *Proceedings of the 20th International Society for Music Information Retrieval Conference* (Delft, Netherlands, Nov 2019). 2

[YYK*23] YING G. J., YIN K., KUMAR K. D., GENG H., MAHADEVAN S. C.: NUS Motion Capture Database, Accessed: 2023. URL: https://mocap.cs.sfu.ca/nusmocap.html. 2

[Zha20] ZHANNA R.: Digitization of intangible cultural heritage as a method to save and actualize it on the example of kazakh folk dance. In *Proceedings of the 6th International Conference on Engineering & MIS 2020* (New York, NY, USA, 2020), ICEMIS '20, Association for Computing Machinery. doi:10.1145/3410352.3410748. 2