Immersive Instruction of Dance Using 3D Capture Technology

Nangia, Shriya

Honors in Computer Science Thesis

Immersive Instruction of Dance Using 3D Capture Technology

Shriya Nangia, Computer Science B.S.

University of Iowa Holo Laboratory

Faculty Supervisor:

Tyler Bell

Abstract

Immersive environments and 3D capture technology are two powerful tools that have revolutionized the way we experience entertainment, education, architecture, engineering, and design. Virtual environments are computer-generated simulations with the aim to immerse users in a digital world. They can be designed for entertainment purposes, such as gaming, or for educational and training purposes, such as simulating real-world scenarios for military or medical training. 3D capture technology uses sensors or cameras to track the movements of an individual, capturing the motion and transforming it into a computer-generated animation or 3D model. It is widely used in the entertainment industry, particularly in the creation of animated movies and video games, as well as in scientific research and sports analysis. Immersive environments and 3D capture technology have great potential in the area of dance, both in terms of performance and training. Hence, in this paper, a project to create an immersive dance instruction environment using motion capture technology is developed and evaluated. Nonetheless, the quality of movement captured was influenced by the type of motions the instructor carried out.

Introduction

Immersive environments can be used in dance training, particularly for educational purposes. This can be useful for long-distance collaboration between dancers and choreographers, as they can work together in a shared virtual space. This allows students to learn from top dance instructors regardless of their location, helping to democratize access to dance education, and providing opportunities to people who may not have access to traditional dance schools.

I explored 2 main methods of 3D capture namely, using Microsoft's Kinect and the OptiTrack Motion Capture System. The Kinect uses a combination of infrared and RGB cameras to capture and track human movements in 3D space. It uses a process called "skeletal tracking" to build a physical body using a collection of points in 3D space. The system captures a depth image of the user and applies machine learning algorithms to identify the user's body parts and track their movements in real-time. The OptiTrack system builds a physical body using a collection of points in 3D space as well. The system uses multiple cameras to capture the movements of reflective markers placed on the subject's body. These markers are tracked in real-time, and their positions are used to create a 3D model of the subject's movements. Both technologies have been used in various applications, including gaming, healthcare, and education.

In this paper, I explore the construction of an immersive environment using the Unity Game Engine, placing an avatar animated using 3D capture technology and allowing users to interact with the avatar and the environment in a way that feels realistic.

Methods

Over a series of weeks, I captured 3D (RGBD) data of dance performers with the Microsoft Kinect to study dance movements in 3D space. The participants in this recording were adults dances from diverse backgrounds. A total of about 20 participants turned up for each session. The age range of the participants was 18 years and above. Participants did not necessarily have prior experience in various dance forms. Ethical considerations were taken into account, and informed consent was obtained from all participants prior to their involvement in the recording.

To capture the 3D (RGBD) data of the dance performers, two Microsoft Kinect v2 sensors were used. The Kinect v2 sensor combines an RGB camera and an infrared depth sensor, allowing for simultaneous capturing of color and depth information. This sensor was chosen due to its ability to provide detailed depth information and its suitability for capturing dynamic movements.

The data acquisition setup consisted of a large empty theatre, namely University of Iowa's Space Place Theatre in North Hall. There were adequate lighting conditions to ensure optimal performance of the Kinect sensor, and different stage lighting designs were used. The sensor was placed on a stable tripod at a height of approximately 1 meter. This height was chosen to capture the entire body of the participants during their dance movements. The participants were instructed to perform a dance routine led by the lead choreographer.

Before the data capture session, the participants were given a brief explanation of the study objectives and were provided with time to warm up and familiarize themselves with the stage area. They were also instructed to wear form-fitting clothing to ensure better visibility of body movements in the captured data. During the data capture session, each participant performed a predefined dance routine led by the lead choreographer, within the designated stage area. To ensure consistency, the same routine was performed by all participants. Each participant performed the routine two times to account for any variations or errors and to ensure optimal data had been captured.

After the data capture session, the recorded RGBD data was transferred to a computer for further processing. The depth data, consisting of a depth map representing the distances of points from the sensor, was used to create a 3D point cloud of the participant's body. The RGB data, captured by the color camera, was used to overlay color information onto the 3D point cloud. I carried out quantitative analysis of the dance movements captured, which involved visual inspection and subjective interpretation of the dance movements based on the captured 3D data. Upon evaluating factors such as fluidity, and range of motion, from the aspect of using this capture data for avatar animation, we realised that using motion capture technology may deliver more favorable results.

To capture the dance movements for the animation of avatars, the OptiTrack motion capture system was then utilized. The decision to use the OptiTrack system was based on its high precision and accuracy in capturing the intricate movements of the dance performer. Unlike the Microsoft Kinect, which was used for 3D (RGBD) data capture in a previous phase of the study, the OptiTrack system provided more detailed and reliable motion data suitable for animation purposes.

The OptiTrack system consists of multiple high-resolution cameras strategically placed around the capture area. These cameras track the position and orientation of reflective

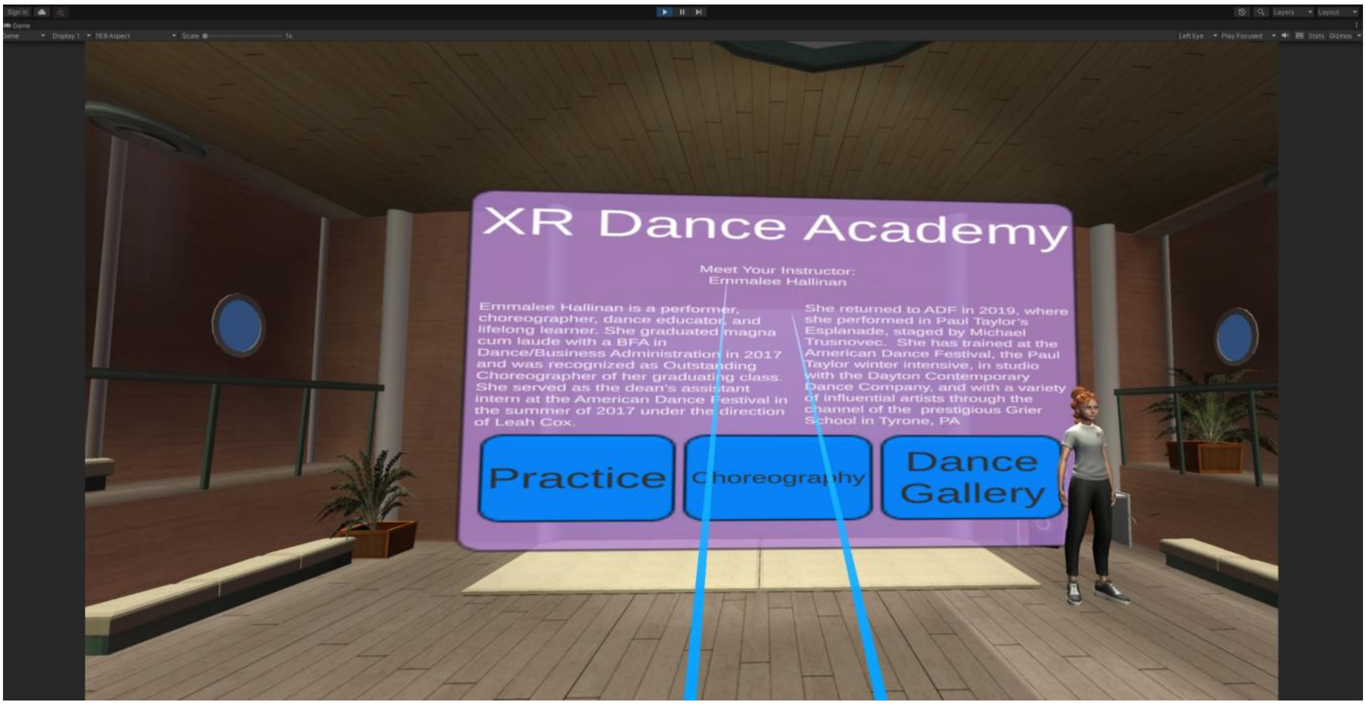
markers placed on the dancer’s body, allowing for precise capture of their movements in three-dimensional space.

Before the motion capture session, the instructor was given a brief explanation of the objectives and was provided with time to warm up and familiarize themselves with the capture area. Reflective markers were then attached to specific anatomical landmarks on the dancer's body, including joints, limbs, and torso, to facilitate accurate tracking.

During the motion capture session, the instructor performed a predetermined dance routine within the capture area. The routine included a combination of basic dance movements and choreographed sequences specific to the dance style, Hip Hop.

After the motion capture session, the recorded motion data from the OptiTrack system was transferred to a computer for further processing. The OptiTrack software, Motive, provided was used to calibrate and synchronize the captured data from multiple cameras. This involved refining the captured motion data by adjusting keyframes, smoothing motion curves, and fine-tuning joint rotations to enhance the quality and aesthetic appeal of the animation. The motion data, after processing, was imported into the Unity Game Engine to animate virtual avatars that replicated the dance movements of the instructor. The motion data was applied to a rigged character model to animate it in a realistic manner. The animated avatars were evaluated based on their resemblance to the original dance routine performed by the instructor.

After capturing the motion data and applying the post-processing techniques to refine the accuracy and precision of the captured movements, the processed motion capture data was exported as an “.fbx” file format compatible with Unity Game Engine.

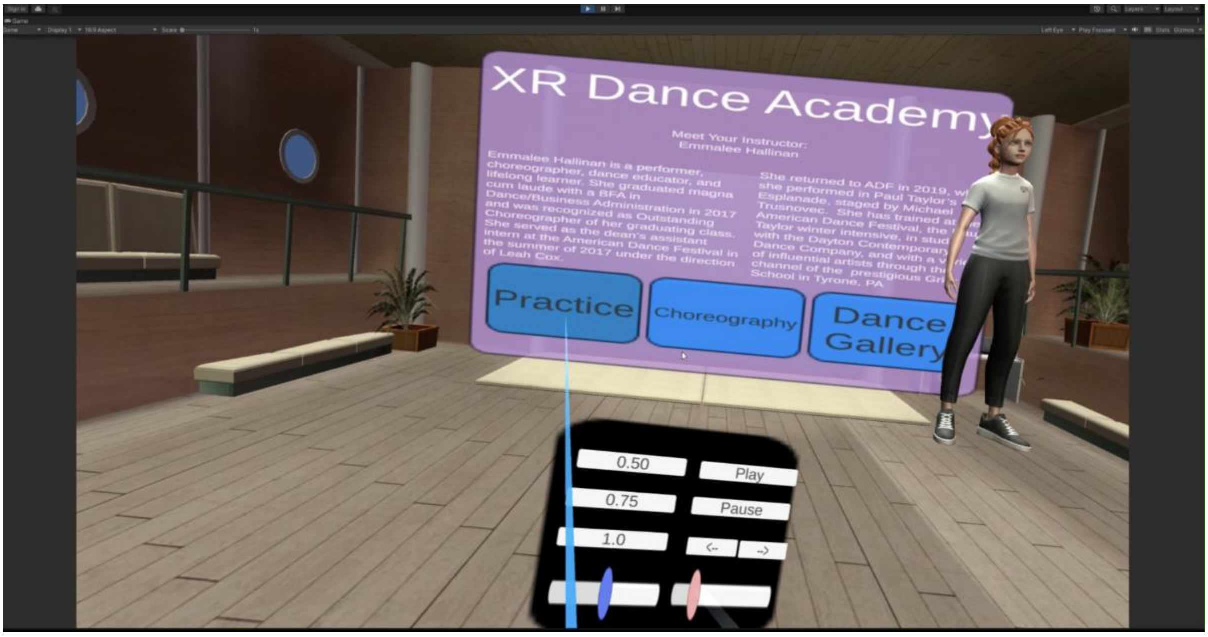


The motion capture data was mapped onto the participant's chosen avatar in Unity’s Animator which enabled the avatar to accurately replicate the recorded hip hop dance movements. To represent the dance instructor in a virtual reality environment, an avatar was generated using ReadyPlayerMe, a platform that allows users to create customized virtual avatars. The instructor had the option to create an avatar that slightly resembled their physical appearance and personal style.

The virtual reality dance studio was designed as the primary environment for participants to learn and practice hip hop dance. The studio was created using Unity and featured a visually immersive setting with appropriate lighting, a dance floor, and a mirrored wall to allow participants to observe their movements.

To enhance the learning experience, several customization options were provided to users. Users could choose from a selection of 12 basic hip hop dance moves, each of which could be learned individually.

Additionally, they had the option to select a full choreography that combined the 12 basic steps into a 1-minute routine. To accommodate individual preferences and learning styles, users are given control over various parameters within the virtual reality dance studio. They could adjust the speed of each dance movement, position and placement of their avatar dance instructor, and even alter the size and rotation of the instructor's avatar allowing them to observe the steps both from the front and back. These controls were accessible through a virtual watch interface on the user's wrist, utilizing finger gestures for selection, without the need for external controllers due to the implemented hand tracking functionality.



To address the importance of learning the history and origins of a dance form, particularly in the context of hip hop, I developed an Immersive Virtual Reality (VR) "History Gallery" to incorporate this educational element alongside the practical aspects of learning the dance moves. It allows individuals to explore the roots of hip hop in an engaging and interactive manner.

The VR History Gallery was designed as an immersive environment that simulates a museum-like setting. It was created within the Unity Game Engine, ensuring a realistic and visually appealing experience for users. The gallery consists of various panels dedicated to significant artists and the origins of specific dance moves within hip hop. To facilitate user exploration and learning, users navigate through the VR environment using joystick controls, allowing them to move freely and explore. Users can interact with the content by clicking on various user interface elements (UI). By clicking on specific UI elements, users can access detailed information about artists, historical events, and the origins of specific hip hop dance moves.

Results and Discussion

The motion capture of the dancer using the chosen motion capture system yielded favorable results, capturing the overall dance movements accurately. However, there is a scope for improvement in capturing finer details, such as finger movements. By increasing the number of markers and improving their placement, it would be possible to enhance the precision of capturing these intricate details, providing a more comprehensive representation of the dancer's performance.

Hip Hop is a dance style known for its intricate footwork, which presented challenges for the motion capture system. Breakdancing, a significant component of hip hop, involves performers executing complex movements that require the body to twist into various angles. Unfortunately, the motion capture system struggled to accurately capture these intricate body movements. Specifically, it faced difficulty in distinguishing when the legs were crossed over one another, potentially due to markers being blocked by each other or not captured correctly. This limitation hampered the system's ability to accurately represent and analyze the complex footwork and twisting movements in hip hop and breakdancing.

To mitigate the limitations of the motion capture system, the dance steps had to be recorded at a significantly slower speed and later adjusted to match the original pace during post-processing. This approach aimed to enhance the accuracy of the motion capture system by allowing it to capture the movements more effectively. Slowing down the dance performance provided the system with more time to accurately track the markers. However, this workaround added an additional step to the data processing pipeline and may not fully compensate for the system's inherent limitations.

Certain limitations in detecting ankle, knee, and hip rotation in all three planes resulted in the need to correct the recorded animation in Motive Software. The motion capture system was unable to capture these rotations accurately, leading to discrepancies in the recorded motion. Through the use of Motive, adjustments and corrections were made to the recorded motion captures to improve the fidelity and alignment of the animations. The software's skeleton solver proved particularly useful in resolving occlusion issues caused by blocked markers.

Motive was instrumental in successfully editing and trimming the recorded motion captures to the desired length. The software's capabilities allowed for precise control over the captured data, enabling the removal of unwanted segments and refining the animations. This editing process helped streamline the data and ensure that only the desired portions were used for further analysis and animation.

The motion capture results demonstrated the effectiveness of the system in capturing the overall dance movements of the performer. However, challenges were encountered in capturing finer details like finger movements and accurately representing complex footwork and twisting movements in hip hop and breakdancing. These limitations may be addressed by improving marker placement, density, and tracking capabilities.

Recording the dance steps at a slower speed and correcting the motion captures in post-processing enhance accuracy only to a certain extent, albeit with additional processing steps. The use of Motive played a crucial role in editing and refining the recorded motion captures, resolving occlusion issues, and trimming the animations to the desired length.

Overall, while the motion capture results were generally favorable, it revealed limitations in capturing fine details, complex footwork, and twisting movements. Corrections and adjustments were made using Motive to enhance the accuracy and alignment of the recorded motion captures. Further improvements in motion capture technology and techniques are necessary to fully capture the intricacies of dance movements and support more precise analysis and animation

Conclusion and Future Implementation

To enhance the virtual reality dance experience, future implementations would be focused on incorporating networking concepts to enable multiple students to join classes simultaneously. This would transform the learning environment into a more social and collaborative setting, where students can interact with each other, share their progress, and learn together. Additionally, expanding the classes to include other dance forms such as jazz, ballet, and cultural dances like Kathak, Chinese dance, or Japanese dance would allow students to explore a wider range of dance styles and increase their cultural and historical awareness.

To provide a more comprehensive learning experience, in future implementations I would develop dedicated VR galleries for each dance form. These galleries would serve as interactive spaces where users can learn about the significant people, histories, and cultural contexts associated with each dance style. Incorporating multimedia elements, such as videos, audio narrations, and interactive displays, would further engage users and deepen their understanding of the diverse dance forms from around the world. Additional exhibits and sections can be added to cover a broader range of artists, dance moves, and historical events. Furthermore, adding interactive elements such as mini-games, quizzes, and collaborative features can enhance user engagement and provide a more dynamic and comprehensive learning experience. These interactive elements would provide opportunities for users to test their knowledge, reinforce learning concepts, and gamify the overall experience. By incorporating challenges, rewards, and progress tracking, users would be motivated to actively participate in the learning process and further deepen their understanding of the dance forms.

To address the limitations of occlusion and marker visibility, using a more advanced motion capture system is crucial. This would involve utilizing better marker tracking technologies that reduce occlusion and improve marker recognition. By capturing movements with higher precision and accuracy, the animated avatar in the virtual reality dance studio would provide a seamless and less confusing representation of the instructor's movements, allowing users to follow along more effectively.

In addition to the future implementations mentioned earlier, another important aspect would be to conduct a user test study to evaluate the overall success and effectiveness of the XR Dance Academy application. This study would provide valuable insights into the user experience, learning outcomes, and areas for further improvement. The user test study would involve recruiting a diverse group of participants with varying levels of dance experience and familiarity with virtual reality technology. These participants would be randomly assigned to either an experimental group, which would use the XR Dance Academy Application, or a control group, which would follow a traditional online method of learning dance. The study would assess several key factors, including user satisfaction, learning outcomes, engagement, and usability of the application. The data collected could include survey responses, performance evaluations, and qualitative feedback gathered through interviews or group discussions.

In conclusion, while there is ample room for future developments and expansions, the implementation of the virtual reality dance experience could be a promising tool for learning and practicing dance. The immersive environment, customizable features, and historical context provided by the VR History Gallery would enhance participants' engagement and appreciation for learning dance forms, increasing cultural awareness, and fostering a sense of community among dancers worldwide.