Advances in Science and Technology Research Journal

Advances in Science and Technology Research Journal, 18(8), 217–227
https://doi.org/10.12913/22998624/193180
ISSN 2299-8624, License CC-BY 4.0

Received: 2024.08.25
Accepted: 2024.10.04
Published: 2024.11.01

Identifying and Animating Movement of Zeibekiko Sequences by Spatial Temporal Graph Convolutional Network with Multi Attention Modules

Maria Skublewska-Paszkowska[1], Pawel Powroznik[1], Marcin Barszcz[1*], Krzysztof Dziedzic[1], Andreas Aristodou[2,3]

[1] Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland
[2] University of Cyprus, 75 Kallipoleos Str., Nicosia 1678, Cyprus
[3] CYENS Centre of Excellence, 23 Dimarchias Square, Nicosia 1016, Cyprus
* Corresponding author's e-mail: m.barszcz@pollub.pl

ABSTRACT

Folk dances, integral components of intangible cultural heritage (ICH), are both fleeting and fragile. However, with the rapid advancement of computer vision techniques, there arises an opportunity to document and safeguard these cultural expressions for future generations. This study aims to identify the distinctive dance sequences and characteristics of Zeibekiko, a popular Greek folk solo dance found in variations across Greece, Cyprus, and the Aegean region of Minor Asia, and translate them into a virtual 3D environment. Utilizing a state-of-the-art optical motion capture system featuring active markers (the PhaseSpace X2E system), precise recordings of the Zeibekiko dance are achieved. The three-dimensional spatial data derived from the dancer's movements serves as the foundation for classification, accomplished through a Spatial Temporal Graph Convolutional Network with Multi Attention Modules (ST-GCN-MAM). This innovative architecture strategically employs attention modules to extract key features of the dance from primary areas of the upper and lower parts of human body. With high level accuracy, the proposed tool accurately detected and recognized Zeibekiko sequences. Ensuring the precise alignment of captured points with corresponding bones or anatomical features in the 3D dancer model is essential for seamless and authentic animations. Advanced visualization and animation techniques are then employed to translate these points into smooth, realistic character movements, preserving their inherent dynamics and expressions. As a result, a faithful virtual rendition of the dance is achieved, capturing its authenticity and beauty. Such a solution holds potential applications in gaming, video production, or virtual museum exhibits dedicated to showcasing folk dances.

Keywords: graph convolutional network, attention model, motion capture, intangible cultural heritage, 3D model, animation, Zeibekiko.

INTRODUCTION

Animation involves creating moving images and has become increasingly significant over the years, impacting various fields such as entertainment, science, education, medicine, and industry. It plays a crucial role in the entertainment sector, from classic animated films to modern 3D animations and video games, captivating audiences of all ages with remarkable visual experiences that entertain, engage, and sometimes move them emotionally. Iconic films like "The Lion King", "Avatar," and "Shrek" have transformed perceptions of animated films, proving that animated characters can evoke strong emotions like those portrayed by live actors. In education, animation helps visualize and simulate complex concepts that are difficult to explain with static images or text. In fields like medicine, the military, and aviation, animation enables the creation of realistic simulations, invaluable for training and real-world preparation. In marketing and advertising, dynamic animated ads are more effective in capturing customer attention compared to static ads. Animation also serves as a unique form of artistic expression, offering creators limitless opportunities to unleash their creativity. Like painters, animators use their skills and technology to produce works that express their visions and emotions. Over the last decade, the rapid development of 3D technology has significantly advanced animation, enabling more realistic effects. In character animation, techniques like Motion Capture (MoCap) and 3D animation are revolutionizing fields from films to virtual, augmented, and mixed reality, transforming how we interact with digital worlds.

MoCap technology has become essential in modern film production, gaming, and scientific research. It precisely captures actors' movements, transferring them to digital characters, playing a key role in preserving and showcasing intangible cultural heritage (ICH), such as traditional dances. This technology accurately records complex choreographies that are integral to many cultures, allowing these dances to be visualized in detail, invaluable for future generations and cultural researchers. MoCap facilitates the creation of virtual dance libraries for educational and cultural promotion worldwide. Traditionally, performances like dance and rituals, lacking physical documentation, were preserved mainly through films or photos, making them vulnerable to being forgotten. With the advancement of 3D animation, these arts can now be faithfully transferred to a virtual 3D space, serving as documentation and protection against the loss of cultural heritage. By using the MoCap technology, we can offer engaging presentations of this heritage, immersing younger generations in the virtual world.

This main contribution of this work is to identify the unique and distinctive sequences and characteristics of Zeibekiko, a popular Greek folk dance from Greece, Cyprus, and the Aegean region of Asia Minor, and translate them into a virtual 3D environment. The Zeibekiko dance performances were captured using the PhaseSpace X2E motion capture system with active markers. From this data, ten sequences were extracted to define the dance. A Spatial Temporal Graph Convolutional Network with Multi Attention Modules (ST-GCN-MAM) was used to recognize these sequences, focusing on key features from the upper and lower body. The recognized sequences were visualized through animations, showcasing the distinct movements and characteristics of the Zeibekiko dance.

RELATED WORKS

Digitalization of traditional dance is crucial for preserving and protecting cultural heritage. This aspect, developed over generations, is often forgotten and endangered, making it essential to safeguard traditional folk dances for future generations [1]. MoCap techniques, beyond videos and images, play a major role in accurately recording three-dimensional movements [2]. Research indicates MoCap as the most promising method for preserving intangible cultural heritage, allowing precise capture of dancers' movements for realistic animations used in education and entertainment [3]. Indeed, MoCap allows one to create three-dimensional digital models of movements that can be analyzed or reconstructed, e.g., in the form of realistic animations for educational or entertainment purposes.

A study on the Eyo masquerade dance demonstrated the power of MoCap systems—optical, inertial, and volumetric—in capturing movements with and without costumes [4]. This method effectively preserves folk performances. Another study focused on learning the Czech folk dance, Hornemcanska, using VR and feedback, showing improved results in dance performance compared to professional dancers [5]. The Tang Dynasty dance from China was captured using Vicon MoCap and presented in VR, alongside tomb murals to visualize dance, music, and scenes, further proving the importance of MoCap in cultural preservation [6]. This conclusion was supported by another study, which emphasized MoCap's advantages in creating animations that preserve and teach traditional dances, capturing complex three-dimensional details that traditional methods (e.g., text, images, videos) cannot [7].

MoCap technology not only preserves dance movements but also aids in creating comprehensive databases for educational purposes, enriching the teaching process and allowing for a detailed understanding of traditional dances. Kaiqiang Sun's work focuses on MoCap technology in recording, visualizing, and teaching dances, emphasizing its transformative potential in traditional dance education and its versatile applications in multimedia, sports training, and rehabilitation [8].

More recently, machine learning and deep learning techniques analyze dance movements to identify patterns, classify forms, and assess technique [9]. Digital dance repositories not only preserve dance but also serve as educational tools, enabling students and researchers to explore various techniques. MoCap, combined with VR technology, creates an immersive mechanism for presenting recorded dance movements and mapping their trajectories in 360 degrees [10]. This experience helps users understand the technique and performance of traditional dances, enhancing interest in intangible cultural heritage. Studies also explore dance pose and gesture recognition, with techniques like fuzzy L membership functions and Hidden Markov Models (HMM) applied to various dances, achieving high accuracy in gesture classification [11, 12, 13]; for instance, six gestures of Likok Pulo Dance from Aceh, captured with a Kinect sensor, have achieved 94.87% accuracy [13]. Research on unsupervised dance figure analysis through video recording highlights the use of HMM for segmenting video recordings to identify repeating patterns and generate animated avatars, beneficial for dance education and interactive entertainment [14].

The successful classification of folk arts, such as Dai, Tibetan, Wei, Mongolian, and Miao dances, shows the effectiveness of VR applications in enhancing user awareness and learning through entertainment [15]. The integration of motion capture data into Labanotation for dance movement annotation demonstrates advancements in preserving folk dances, while other studies employ classifiers like k-NN and TreeBagger for movement annotation in Malaysian and Indian dances [16, 17, 18, 19]. Motion capture data also facilitates the classification of Chinese classical dance postures and recognition of cultural dances, utilizing neural networks with high accuracy [20, 21]. The discussed research underscores the significance of using computer technologies in recording and animating dance movements, enhancing both preservation and education of traditional dances.

MATERIALS AND METHODS

Traditional part-based frameworks do not widely use deeply trained networks to enhance both part localization and feature representation learning. In this paper, we briefly describe the Spatial Temporal Graph Convolutional Network with Multi Attention Modules (ST-GCN-MAM). Its unique architecture focuses on the attention modules to extract the characteristic features of dance in four parts of the human body. Our approach allows for nearly cost-free computation of part attentions, and the network can be trained end-to-end. The architecture of the network, as depicted in Figure 1, includes convolution, channel grouping, and part classification subnetworks. Initially, the network processes a full-size input represented as a skeleton graph through convolutional layers to extract region-based feature representations. Next, it generates multiple part attention maps via channel grouping and weighting layers, using a sigmoid function to produce probabilities. These part representations are then derived by pooling from region-based feature representations using a spatial attention mechanism. Finally, fully-connected and Softmax predict a set of probability scores for each category. The ST-GCN-MAM is optimized to convergence by iteratively learning a Softmax classification loss for each part.

Dataset

In this study, we utilized data from the Dance Motion Capture Database (DMCD) [22], focusing on the culturally significant Zeibekiko dance. The PhaseSpace Impulse X2E motion capture system [9], employing active optical motion capture via modulated LEDs, was utilized to record and preserve these intangible cultural expressions. This system provides high-frequency optical tracking at up to 960Hz, ensuring precise three-dimensional motion data capture while preserving natural human proportions and movements.

The DMCD comprises performances by skilled dancers from diverse cultural organizations and dance schools, forming a comprehensive dataset of over 250 dance performances featuring more than 30 folk dances, some of which are included in EUROPEANA, the European Digital Library of Cultural Heritage. We carefully selected dance data from expert male and female performers with extensive experience in cultural organizations and dance schools. Specifically, we included five variations of the Zeibekiko dance performed by four different dancers, alongside performances of 14 other Greek and Cypriot folk dances (such as Tatsia, Pentozali, Antikristos, Haniotikos, Hasapiko, Tsamikos, among others) by five different dancers. This diverse dataset enabled us to capture a wide spectrum of dance styles and expressions.

For efficient storage and computational processing, we stored the motion capture data in both C3D and BVH formats, initially sampled at varying rates from 120 to 480 frames per second. However, to optimize efficiency without sacrificing temporal information, we downsampled the data to 24 frames per second. This approach preserved essential characteristics of the dance movements while reducing data size and computational demands. The data was segmented and labeled by a professional dancer associated with the Nicosia Folk Association. It's noteworthy that the raw C3D data was used without any post-processing or cleaning. Fortunately, the data was largely free of artifacts and occlusions, ensuring its reliability and suitability for analysis.

Data preparation

Utilizing the C3D Python library and based on specified frame numbers, the following sequences were extracted from the C3D files. The professional dancers identified ten sequences from a recorded Zeibekiko dance, as outlined in Table 1.

SPATIAL TEMPORAL GRAPH ATTENTION NETWORK WITH MULTI ATTENTION MODULES FOR BODY PART LOCATION IN 3D SPACE

The architecture of the ST-GCN-MAM is illustrated in Figure 1. The ST-GCN-MAM, proposed in this article, comprises four independent components with identical structures, each designed to model the dependencies of the upper left body part, upper right body part, lower left body part, and lower right body part of the data.

Each component consists of a convolution layer, few Spatial-Temporal-Graph-Attention-Module (ST-GAM) blocks, and an output block. The initial convolution layer captures correlations between input features and generates multiple feature maps. The ST-GAM blocks model spatial-temporal dependencies, with each ST-GAM block incorporating skip connections to prevent over smoothing. The outputs of the four components, denoted as *Y1*, *Y2*, *Y3*, and *Y4*, are combined into the final output *Y* by the multicomponent fusion module.

The ST-GAM includes two Graph Attention Convolution (GAC) blocks and a two Graph Convolution Network (GCN) layers, as well as, two Gated Temporal Convolutional Layer (GTCL) and two Batch Normalization (BN) blocks, as illustrated in Figure 2. Spatial-Temporal-Graph-Attention-Module architecture. Additionally, the two GAC layers within the same ST-GAC block have shared weights (WS) and there are skip connection (SC) between corresponding GTCL. The GTCL block is designed to capture temporal dependencies, while the GAC layer is used to identify correlations between nodes in spatially sparse data.

Attention models have been widely applied in the field of computer vision [23, 24, 25]. These models gather information from a limited visual area. The chosen region, from which data is extracted, is identified by specific activity patterns, textures, colors, or shapes. Typically, an attention model is structured as an encoder-decoder. It is extensively utilized in various applications such as traffic forecasting [26], sequences labelling [27, 28], recommendation systems [29], and document classification [30]. According to [31],

Table 1. Zeibekiko sequences

No. sequence	Sequence name
1	Left Single Turn
2	Right foot strike
3	Seat on the right
4	Seat on the left
5	Leg kick with counterclockwise turn
6	Step right
7	Hand clap shot floor
8	Double hand clap
9	Double pretense right
10	Back heel strike

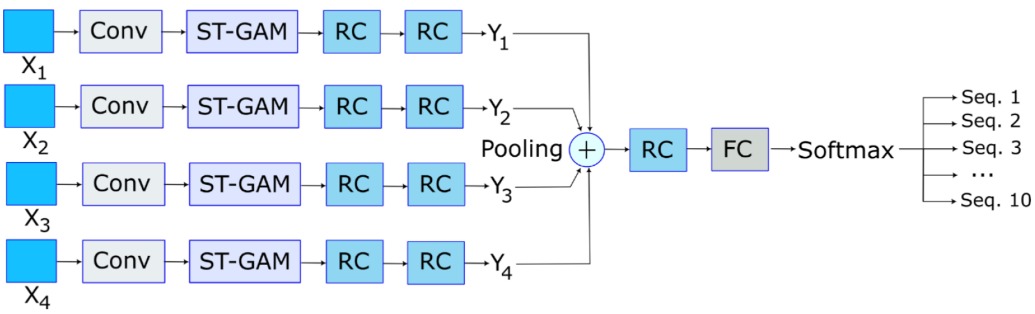


Figure 1. ST-GCN-MAM architecture

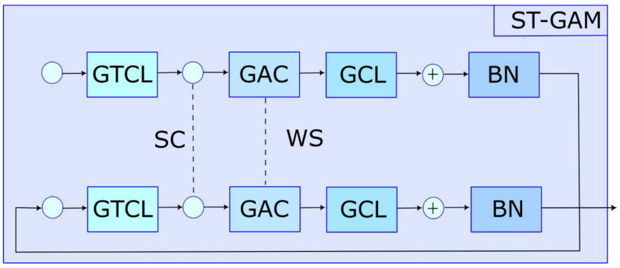


Figure 2. Spatial-Temporal-Graph-Attention-Module architecture

attention models can be broadly categorized into hard and soft attention. In this study, the soft attention model was employed. The primary function of the attention model is to retain information about the selected sections of the model, while the secondary function is to designate the context vector, which plays a crucial role in predicting dancer positions.

Given that capturing spatial correlations in spatially sparse data is challenging, we utilize multiple stacked GAM blocks to better exploit relationships and improve predictions. Each block consists of one GCN layer and one GAC layer. The use of GAC with shared parameters within these blocks may also help mitigate the over-smoothing issue associated with GCN.

Gated temporal convolution layer

Drawing inspiration from Graph WaveNet [32], we casual dilated convolution to capture temporal dependencies. Unlike traditional 1D convolution, dilated causal convolution skips a fixed number of steps to perform the convolution operation. By stacking multiple layers of dilated causal convolutions, the receptive field can increase exponentially. Additionally, when processing input sequences of the same length, dilated convolution requires fewer parameters and achieves faster training speeds compared to Recurrent Neural Networks (RNNs). Given the scattered and distant spatial distribution of use points, we believe that the relationship between these points has a delay effect, making dilated convolution more scalable. To better learn temporal information, we employ a gated mechanism based on dilated causal convolution (GTCL) to control the flow of information [33]:

$$\text{output} = \tanh(\omega_1 \times \mathrm{X} + \mathrm{b}) \circ \sigma(\omega_2 \times \mathrm{X} + \mathrm{c}) \quad (1)$$

where: b, c, ω indicate learning parameters, $\circ$ is the elementwise product, * denotes convolution operation and σ is an activation function.

Graph attention convolution block

Since GTCL facilitates the creation of connections between nodes by using hierarchy in their levels of importance, and GCN is a computationally advantageous model of convolutional networks capable of operating in non-Euclidean spatial structures, we have devised a graph attention convolution block (GAC). This block is directly connected with GCN layer to register spatial dependencies. In this study, we incorporate the GCN layer proposed in [32] for further imaging of hidden spatial connections based on GAC. The formulation of the whole block is as follows [33]:

$$\mathrm{Z} = \sum_{k=0}^{K} \mathrm{P_f}^k \mathrm{X} \mathrm{W}_{k1} + \mathrm{P_b}^k \mathrm{X} \mathrm{W}_{k2} + \widetilde{\mathrm{A}^k_{apt}} \mathrm{X} \mathrm{W}_3 \quad (2)$$

where: $\mathrm{A}^k_{apt} \in \mathbb{R}^{N' x N'}$ denotes an adjacency matrix (normalized with self-loops), $\mathrm{X} \in \mathbb{R}^{N' x\, D}$ represents an input data, $\mathbf{N}'$ is the number of available nodes, D signifies the feature dimension, $\mathrm{W} \in \mathbb{R}^{D x\, M}$ are the learning parameters, and $\mathrm{Z} \in \mathbb{R}^{N' x\, M}$ denotes the output. In case of a directed graph, the diffusion process occurs in two directions: P^k_f represent the forward direction, P^k_b denotes the backward one and *k* is the order of diffusion.

Multicomponent fusion

In the multicomponent fusion block, *Y1*, *Y2*, *Y3*, and *Y4* are concatenated along the feature axis and treated as feature vectors representing different spatial-temporal dependencies. Subsequently, two convolution layers with ReLU activation functions are employed to learn the correlations among four components and the attributes of each prediction time step. The outputs of the four components are fused as follows [33]:

$$\tilde{Y} = W_2 \times ReLU(W_1 \times Y_1||Y_2||Y_3||Y_4||) \quad (3)$$

where || denotes the concatenation operation and × represents the convolution operation.

CLASSIFIER EVALUATION

In the field of machine learning, several typical measures are used for evaluating classifiers such as: accuracy, precision, recall and F1 score [34, 35]. *Accuracy* is determined by dividing the number of correctly predicted instances by the total number of instances. It provides a general measure of how well the classifier is performing across all classes. *Precision* measures the accuracy of the positive predictions. It is defined as the ratio of true positive instances to the sum of true positive and false positive instances. High precision indicates that the classifier has a low false positive rate. Recall measures the ability of the classifier to identify all relevant instances. It is computed as the ratio of true positive instances to the sum of true positive and false negative instances. F1 score represents the harmonic mean of precision and recall. It delivers a unified metric that balances both precision and recall, making it particularly valuable when the class distribution is imbalanced.

The confusion matrix results, as shown in Figure 3, offer a comprehensive view of the classification accuracy and errors for the model. This visual representation aids in understanding where the model excels and where it may fall short. Furthermore, Figure 4 illustrates the training performance over 10 independent runs, showcasing the model's learning curve and stability across different iterations. This figure helps to assess the robustness and reliability of the training process. To provide a deeper insight into the model's performance, the resulting measures, including precision, recall, F1 score, and other relevant metrics, are meticulously documented in Tables 2 to 5. These tables offer a granular analysis of the model's capabilities, enabling a thorough evaluation of its strengths and areas for improvement.

DANCE SEQUENCES ANIMATION

One of the leading tools for processing data from Motion Capture systems is Autodesk MotionBuilder. offering robust capabilities in importation, mapping, retargeting, and animation editing. Its versatility allows users to craft lifelike character animations applicable across diverse fields. However, leveraging these tools demands advanced proficiency, presenting both a challenge and an opportunity to achieve highly realistic animation effects. This section outlines the process of generating Zeibekiko dance animations through advanced MoCap data processing techniques using Autodesk MotionBuilder. The process encompassed several pivotal stages: motion data registration, cleaning, importation, mapping, calibration, retargeting, and final animation adjustments. Similar methodologies have been

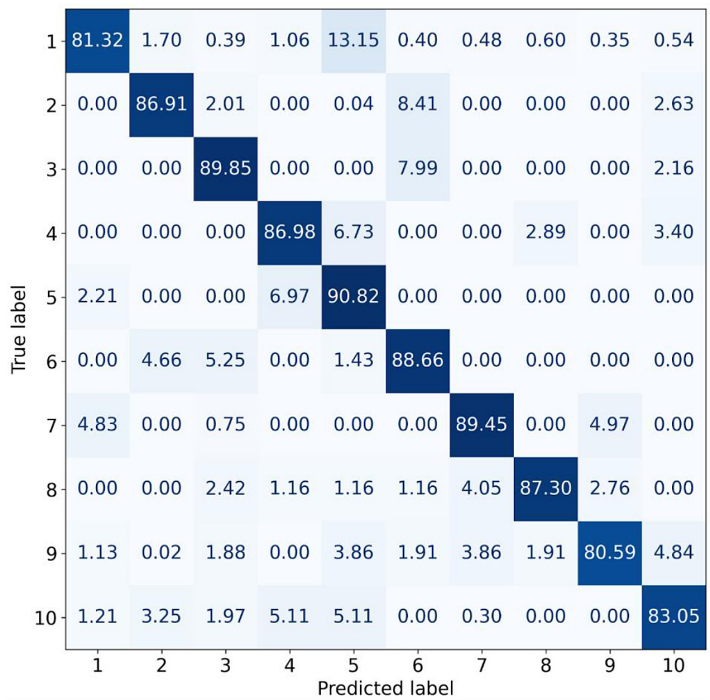


Figure 3. Confusion matrix for Zeibekiko sequences classification

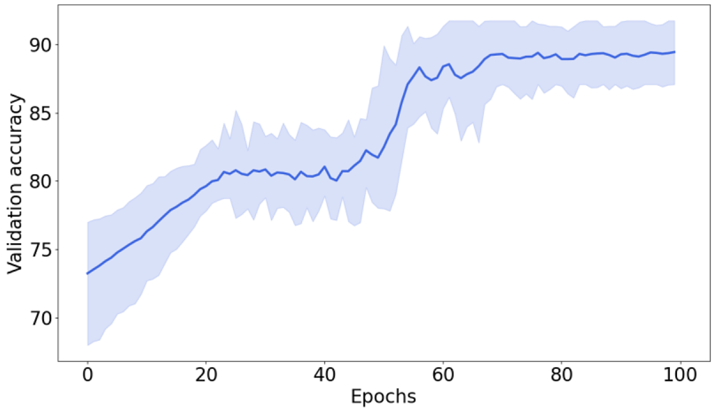


Figure 4. Validation accuracy for ten independent trails. The bold line indicates the average accuracy, while shadows the minimums and maximums values

Table 2. Obtained accuracy results (%)

Parameter	1	2	3	4	5	6	7	8	9	10
Mean	81.32	86.91	89.85	86.98	90.82	88.66	89.45	87.30	80.59	83.05
Max	86.85	89.89	92.65	88.94	94.34	91.72	93.12	90.23	86.90	86.27
Min	79.53	79.95	80.54	79.98	83.71	83.62	86.95	86.44	79.71	79.51
±SD	5.87	9.61	6.72	4.59	2.88	4.63	5.51	2.21	7.36	8.14

Table 3. Obtained precision results (%)

Parameter	1	2	3	4	5	6	7	8	9	10
Mean	84.41	90.15	93.07	90.26	93.87	91.18	92.83	90.52	83.65	86.09
Max	88.50	92.03	95.59	93.21	96.18	94.17	94.93	92.74	87.87	89.46
Min	76.26	84.14	84.74	84.69	89.13	84.93	85.07	84.33	75.38	78.71
±SD	5.74	4.17	3.96	5.37	3.05	3.24	5.52	5.57	5.92	6.38

Table 4. Obtained recall results (%)

Parameter	1	2	3	4	5	6	7	8	9	10
Mean	94.68	89.39	71.72	78.69	75.47	71.94	78.99	89.11	96.82	90.47
Max	96.47	91.39	76.92	87.98	83.89	76.54	82.60	92.10	97.72	92.74
Min	92.82	87.34	64.93	74.77	74.91	68.88	70.37	85.36	95.14	86.14
±SD	3.21	3.17	8.83	6.02	5.76	4.19	5.82	5.98	4.75	4.86

Table 5. Obtained F1-score results (%)

Parameter	1	2	3	4	5	6	7	8	9	10
Mean	87.52	88.25	79.84	82.75	73.87	79.23	84.05	88.30	87.98	86.62
Max	90.37	90.19	84.03	86.45	88.25	83.22	87.02	90.90	90.37	89.49
Min	84.30	85.71	73.52	79.42	68.84	76.07	77.02	84.84	84.12	82.74
±SD	4.46	4.54	8.15	8.82	8.57	5.54	5.05	7.61	4.27	4.39

applied in other studies, such as the creation of traditional Chinese dance animations [36] and the development of ergonomic models [37] using MotionBuilder for real-time assessment of work-related musculoskeletal risks based on MoCap-recorded movement histories.

Initially, the dancer's movement data during the Zeibekiko dance was recorded using the

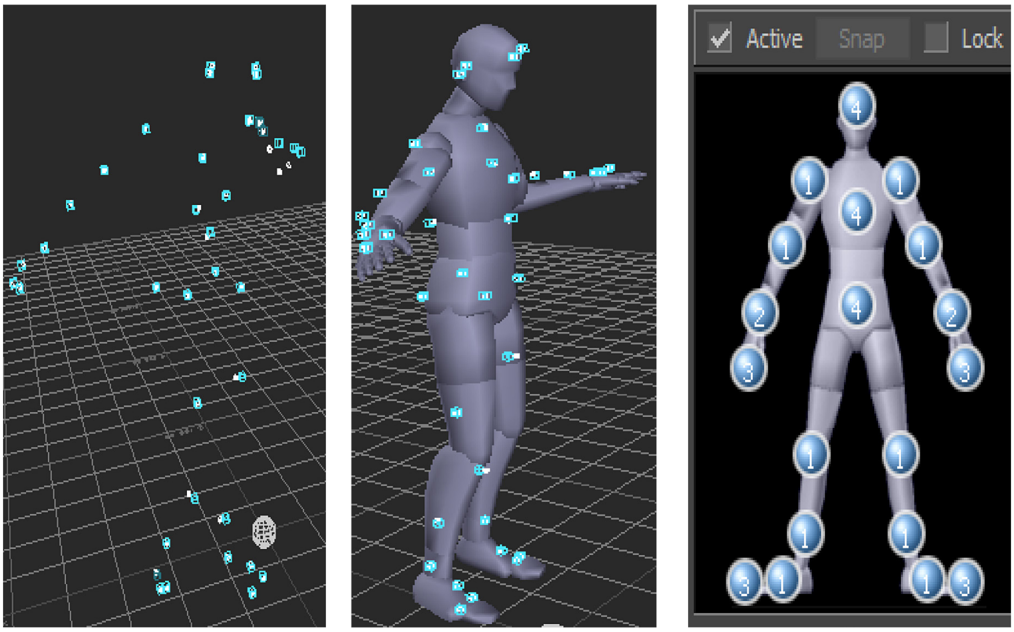
MoCap system, ensuring precise capture of every gesture via multiple body markers. This data, stored in C3D format, detailed marker positions in three-dimensional space. To address noise and errors like missing or inaccurately placed markers, a meticulous cleaning process was undertaken. Techniques such as interpolation to fill data gaps and noise filtering were employed to enhance data fidelity. Then, calibration and normalization procedures followed to ensure consistency and readiness of the motion data for subsequent processing stages. Upon importing the cleaned data into Autodesk MotionBuilder, it was visualized as an animated point cloud (Figure 5a). A critical aspect involved mapping this point cloud onto a virtual actor model. This process entailed precise fitting of the anatomical model to the point data (Figure 5b) and subsequently linking these points to the skeletal model's bones (Figure 5c). This methodology enabled the creation of a dynamic 3D model controlled by the point cloud, laying the foundation for further animation refinement.

In the next step of animation creation, the movements were retargeted from one skeleton (the actor's skeleton) to another (the skeletal structure of a realistic 3D avatar model). This critical step facilitates the adaptation of recorded movements to a 3D model with a different skeletal configuration. Autodesk MotionBuilder provides advanced tools that automate the transfer of movements between skeletons, ensuring adjustments for optimal realism. Additional calibration and refinement were essential to rectify any discrepancies in movement representation on the wireframe, ensuring fidelity to the dancer's original motions.

The final stage involved further refining the animation to maximize realism and align with project specifications. Advanced techniques such as layering were employed, allowing for the addition of supplementary movements or corrections without altering the base animation. Animation blending facilitated seamless transitions between different movement sequences, crucial for maintaining naturalness and continuity. Additionally, control functions were defined to automate specific movements such as walking or hand gestures, enhancing overall naturalness and realism. Through the application of advanced techniques and tools within Autodesk MotionBuilder, the process of animating the Zeibekiko dance yielded a realistic and dynamic portrayal that authentically captures the intricacy and allure of this traditional dance (Figure 6).

DISCUSSION

MoCap systems are increasingly popular due to their advantages such as precision, fluid



(a) (b) (c)

Figure 5. (a) animated point cloud, (b) matching the actor model to point data, (c) assigning animated points to the model bones

Figure 6. Zeibekiko dance animation sequences

motion capture, and the ability to create detailed three-dimensional models. These capabilities allow for the creation of realistic animations that are challenging to achieve using traditional methods. Motion capture can be executed through various techniques including optical systems [4, 6], inertial systems [38], and those utilizing wearable sensors [39, 40].

The classification of three-dimensional data remains a complex task. Machine learning methods have been employed to recognize various types of dances [15] and to assess dancers' techniques by identifying characteristic patterns [3, 16]. Sequencing dance movements is a common focus in scientific studies, with approaches in the literature including RNNs, CNNs, Generative Adversarial Networks (GANs), and their variants [41, 42]. Models such as k-NN classification [16, 18], TreeBagger [16], SVM classifiers [19], and CNN approaches [21] have demonstrated high accuracy in identifying sequences from MoCap data. Additionally, novel approaches like the Multi-Scale Graph Attention Network with Gated Recurrent Score have been proposed for tasks such as Labanotation generation [17]. Similarly, in [43], input sequences are segmented into individual action instances, each instance is recognized, and the execution of each action is evaluated solely based on the positions of 3D skeleton joints.

This work introduces a new approach called ST-GCN-MAM, consisting of four independent soft attention modules designed to model dependencies across the upper left body, upper right body, lower left body, and lower right body segments. Applied to Zeibekiko dance sequences, this architecture achieved satisfactory accuracy, up to 94.34%, indicating its suitability for animating classified Zeibekiko sequences.

CONCLUSIONS

The paper explores a classifier employing four attention modules for classifying dance sequences, demonstrating effective albeit not top-tier accuracy in comparison to existing literature. It innovatively uses 3D point data instead of image-based inputs, offering insights into the benefits of this approach for analyzing and classifying dance. Despite not achieving peak performance, the study underscores the adequacy of the classifier within this novel framework and suggests avenues for further refinement in dance sequence analysis.

The process of animating Zeibekiko dance using Motion Capture and Autodesk MotionBuilder techniques enabled detailed visualization of this important cultural heritage element. Through precise data processing and animation, realistic representations capturing the specificity and dynamics of Zeibekiko were achieved. These animations serve not only to preserve and promote traditional dance but also as valuable educational and research tools, enhancing understanding and appreciation of cultural richness. MoCap and advanced animation software are pivotal in preserving cultural heritage by accurately recording and reproducing intricate dance forms that might otherwise be lost. This technological process not only documents dance traditions but also facilitates their analysis and reinterpretation, crucial for educational and research contexts.

Zeibekiko dance animation contributes significantly to cultural research by enabling detailed studies of its steps, rhythm, and style with precision. Realistic animations can serve as effective educational tools in schools and universities, fostering interactive learning about dance culture and history. Despite challenges such as data cleaning and model calibration, the process of creating Zeibekiko dance animations presents opportunities for advancing research and technology in digital preservation of cultural heritage. Further improvements in data processing and animation techniques could enhance realism and accuracy, potentially addressing concerns about the authenticity of digitally rendered dance.

Looking ahead, advancements in MoCap and animation technologies hold promise for furthering the preservation and promotion of cultural heritage. Integration with augmented reality and virtual reality stands to offer immersive experiences that deepen cultural education and public engagement in safeguarding cultural treasures. In conclusion, utilizing Autodesk MotionBuilder and Motion Capture for Zeibekiko dance animation represents a significant stride in modernizing the preservation of intangible cultural heritage. These advanced tools not only document and analyze traditional dances but also elevate their visibility and societal value.

REFERENCES

1. A. Aristidou, A. Chalmers, Y. Chrysanthou, C. Loscos, F. Multon, J. E. Parkins, B. Sarupuri, E. Stavrakis. Safeguarding our Dance Cultural Heritage. Eurographics Tutorials, April 26, 2022.
2. Guo Y., Wang X. A Spatiotemporal data acquisition and processing method for Ansai waist drum based on motion capture. In: E3S Web of Conferences. 2020, 179, 01024.
3. Reshma M.R., Kannan B., Jagathy Raj, V.P., Shailesh, S., Cultural heritage preservation through dance digitization: A review. Digit. Appl. Archaeol. Cult. Herit. 2023, 28, e00257. https://doi.org/10.1016/j.daach.2023.e00257.
4. Ami-Williams T., Serghides C.-G., Aristidou A. Digitizing Traditional Dances Under Extreme Clothing: The Case Study of Eyo. J. Cult. Herit. 2024, 67, 145-157.
5. Kico I., Zelníček D., Liarokapis F. Assessing the learning of folk dance movements using immersive virtual reality. In: 2020 24th International Conference Information Visualisation (IV). IEEE. 2020, 587-59.
6. Jie G., Chaetnalao A. Application of motion capture technology in the digital recreation of Tang Dynasty mural art: a case study of Han Xiu's Tomb. J. Arts Technol. 2023, 14(4), 298-319.
7. Peng Y. Research on dance teaching based on motion capture system. Math. Probl. Eng., 2022, 1-8. https://doi.org/10.1155/2022/1455849.
8. Sun K. Research on Dance Motion Capture Technology for Visualization Requirements. Sci. Program. 2022, 1–8. https://doi.org/10.1155/2022/2062791.
9. Aristidou A., Shamir A., Chrysanthou Y. Digital dance ethnography: Organizing large dance collections. J. Cult. Herit. 2019, 12, 1-27.
10. A. Aristidou, N. Andreou, L. Charalambous, A. Yiannakidis, Y. Chrysanthou. Virtual Dance Museums: the case of Greek/Cypriot folk dancing. In EUROGRAPHICS Workshop on Graphics and Cultural Heritage, GCH'21, Bournemouth, United Kingdom, November 2021.
11. Hariharan D., Acharya T., Mitra S. Recognizing hand gestures of a dancer. In: Pattern Recognition and Machine Intelligence: 4th International Conference, PReMI 2011, Moscow, Russia, June 27-July 1, 2011. Proceedings 4, 186-192. Springer Berlin Heidelberg.
12. Saha S., Ghosh L., Konar A., Janarthanan R. Fuzzy L membership function based hand gesture recognition for Bharatanatyam dance. In: 2013 5th International Conference and Computational Intelligence and Communication Networks. 2013, 331-335. IEEE.
13. Anbarsanti N., Prihatmanto A. S. Dance modelling, learning and recognition system of aceh traditional dance based on hidden Markov model. In: 2014 International Conference on Information Technology Systems and Innovation (ICITSI). 2014, 86-92. IEEE.
14. Ofli F., Erzin E., Yemez Y., Tekalp A.M., Erdem C.E., Erdem A.T., Abaci T., Ozkan M.K. Unsupervised dance figure analysis from video for dancing Avatar animation. In: 2008 15th IEEE International Conference on Image Processing. Presented at the 2008 15th IEEE International Conference on Image Processing - ICIP 2008, IEEE, San Diego, CA, pp. 1484–1487. https://doi.org/10.1109/ICIP.2008.4712047.
15. Cai W., Liu W. Innovative Strategies of Virtual Reality Technology in Ethnic Dance Inheritance. Appl. math. nonlinear sci. 2024, 9(1).
16. Chaudhry H., Tabia K., Rahim S. A., BenFerhat S. Automatic annotation of traditional dance data using motion features. In: 2017 International Conference on Digital Arts, Media and Technology (ICDAMT). 2017, 254-258. IEEE.
17. Li M., Miao Z., Lu Y. LabanFormer: Multi-scale graph attention network and transformer with gated recurrent positional encoding for labanotation generation. Neurocomputing. 2023, 539, 126203.
18. Ma-Thi C., Tabia K., Lagrue S., Le-Thanh H.,

Bui-The D., Nguyen-Thanh T. Annotating movement phrases in Vietnamese folk dance videos. In: Advances in Artificial Intelligence: From Theory to Practice: 30th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2017, Arras, France, June 27-30, 2017, Proceedings, Part II 30. 2017, 3-11. Springer International Publishing.

19. Samanta S., Purkait P., Chanda B. Indian classical dance classification by learning dance pose bases. In: 2012 IEEE Workshop on the Applications of Computer Vision (WACV). 2012, 265-270. IEEE.
20. Zhang N. Identification Model of Writhing Posture of Classical Dance Based on Motion Capture Technology and Few-Shot Learning. Comput Intel Neurosc. 2022, 1, 8239905.
21. Mindoro J. N., Festijo E. D., de Guzman M. T. G. A comparative study of deep transfer learning techniques for cultural (aeta) dance classification utilizing skeleton-based choreographic motion capture data. In: 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE). 2021, 74-79. IEEE.
22. Stavrakis E., Aristidou A., Savva M., Himona S. L., Chrysanthou, Y. Digitization of cypriot folk dances. In: Progress in Cultural Heritage Preservation: 4th International Conference, EuroMed 2012, Limassol, Cyprus, October 29–November 3, 2012. 404-413. Springer Berlin Heidelberg..
23. Nowomiejska K., Powroznik P., Skublewska-Paszkowska M., Adamczyk K., Concilio M., Sereikaite L., Zemaitiene R., Toro M.D., Rejdak R. Residual Attention Network for distinction between visible optic disc drusen and healthy optic discs. Opt. Lasers Eng. 2024, 176, 108056.
24. Skublewska-Paszkowska M., Powroznik P., Lukasik E. Attention Temporal Graph Convolutional Network for Tennis Groundstrokes Phases Classification. In: 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), July 2022, July, 1-8. IEEE.
25. Liu S., Dai S., Sun J., Mao T., Zhao J., Zhang H. Multicomponent Spatial-Temporal Graph Attention Convolution Networks for Traffic Prediction with Spatially Sparse Data. Comput. Intel. Neirosc. 2021, 1, 9134942.
26. Bai J., Zhu Y., Song Y., Zhao L., Hou Z., Du R., Li H. A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting. ISPRS Int. J.—Geo-Inf. 2021, 10, 485.
27. Skublewska-Paszkowska M., Powroznik P. Temporal pattern attention for multivariate time series of tennis strokes classification. Sensors. 2023, 23(5), 2422.
28. Skublewska-Paszkowska, M., Powroznik, P., Lukasik, E. Attention Temporal Graph Convolutional Network for Tennis Groundstrokes Phases Classification. In: 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). 2022, 1-8. IEEE.
29. Xiao, J.L.; Ye, H.; He, X.; Zhang, H.; Wu, F.; Chua, T. Attentional Factorization Machines: Learning the Weight of Feature, Interactions via Attention Networks. arXiv 2017, arXiv:1708.04617.
30. Pappas N., Popescu-Belis A. Multilingual Hierarchical Attention Networks for Document Classification. arXiv. 2017, 1707.00896.
31. Bahdanau D., Cho K., Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate. arXiv 2014, 1409.0473.
32. Wu Z., Pan S., Long G., Jiang J., Zhang C. Graph wavenet for deep spatial-temporal graph modeling. arXiv preprint. 2019, 1906.00121.
33. Fang S., Zhang Q., Meng G., Xiang S., Pan C. GST-Net: Global spatial-temporal network for traffic flow prediction. In: IJCAI, August 2019, 2286-2293.
34. Powroznik P., Czerwinski D. Spectral methods in Polish emotional speech recognition. Adv. Sci. Technol. Res. J. 2016, 10(32), 73-81
35. Skublewska-Paszkowska M., Powroznik P., Barszcz M., Dziedzic K. Dual Attention Graph Convolutional Neural Network to Support Mocap Data Animation. Adv. Sci. Technol. Res. J. 2023, 17(5) 313-325.
36. Miao Y. Dance pose capture and recognition based on heterogeneous sensors. Procedia Comput. Sci. 2023, 228, 171–184. https://doi.org/10.1016/j.procs.2023.11.021.
37. Qianwen, L. Application of motion capture technology based on wearable motion sensor devices in dance body motion recognition. Meas.: Sens. 2024, 32, 101055.
38. Al-Faris M., Chiverton J., Ndzi D., Ahmed A.I. A Review on Computer Vision-Based Methods for Human Action Recognition. J. Imaging .2020, 6, 46. https://doi.org/10.3390/jimaging6060046.
39. Gupta, S., Singh, S. Indian dance classification using machine learning techniques: A survey. Entertain. Comput. 2024, 50, 100639. https://doi.org/10.1016/j.entcom.2024.100639
40. Muhamada A.W., Mohammed A.A., Review on recent Computer Vision Methods for Human Action Recognition. ADCAIJ. 2022, 10, 361–379. https://doi.org/10.14201/ADCAIJ2021104361379.
41. Patrona F., Chatzitofis A., Zarpalas D., Daras P. Motion analysis: Action detection, recognition and evaluation based on motion capture data. Pattern Recognit. 2018, 76, 612–622. https://doi.org/10.1016/j.patcog.2017.12.007.
42. Wang M., Yu R. Digital production and realization for traditional dance movements based on Motion Capture Technology. The Frontiers of Society, Science and Technology. 2022, 4(11).
43. Eldar R., Fisher-Gewirtzman D. Ergonomic design visualization mapping- developing an assistive model for design activities. Int. J. Ind. Ergon. 2019, 74, 102859.