Dance Action Capture and CAD Design Based on Big Data Algorithm Optimization

Jing Zhao[1] and Jing Li[2]

[1]School of Music, Henan Polytechnic University, Jiaozuo, Henan 454000, China, zhaojing@hpu.edu.cn
[2]School of Music, Henan Polytechnic University, Jiaozuo, Henan 454000, China, Lijinghappygirl@126.com

Corresponding author: Jing Zhao, zhaojing@hpu.edu.cn

Abstract. The primary objective of this investigation is to refine dance motion capture and computer-aided design (CAD) processes through the utilization of big data algorithms, thereby addressing concerns related to precision and productivity in current technologies. During the data acquisition phase, suitable dance motion capture devices and CAD design tools are selected to gather authentic dance movement information and CAD specifications from dancers. Following this, the raw data undergoes preprocessing, which involves cleaning, organization, and format conversion. In the subsequent model development stage, leveraging big data algorithms alongside graph convolution networks (GCN) and other advanced techniques, an enhanced model for dance motion capture and CAD design is formulated. To validate the model's efficacy, a series of comparative experiments are devised. The experimental findings reveal that the model excels in accurately capturing dance movements, achieving an approximately 96% precision rate. Moreover, it demonstrates a capacity to rapidly generate high-calibre 3D models and exhibits remarkable adaptability to diverse dance styles. These advantages make this method have broad application prospects and potential in practical application.

Keywords: Big Data; Dance Action Capture; CAD Design; Graph Convolution Network
DOI: https://doi.org/10.14733/cadaps.2024.S21.68-83

1 INTRODUCTION

With the advancement of science and technology, dance motion capture has emerged as a novel technique, garnering considerable attention in dance artistry. With the rapid development of multimedia data and artificial intelligence technology, human motion recognition has become a research hotspot in the field of computer vision. Among them, human dance action recognition based on big data algorithm optimization has received widespread attention due to its potential application value in entertainment, education, health, and other fields. Ahmad et al. [1] conducted a comprehensive investigation on human dance action recognition based on Graph Convolutional

Neural Network (GCN). Traditional CNN and RNN face challenges in handling dance movements with complex spatiotemporal relationships and dynamic changes. To address these issues, Graph Convolutional Neural Networks (GCN) have been introduced into the field of dance action recognition. GCN can capture the spatial relationships and temporal dependencies between skeleton joints, thereby more accurately identifying dance movements. The data preprocessing stage mainly includes operations such as cleaning, annotating, and skeletonizing the original data; The model construction phase involves designing an appropriate GCN architecture to extract effective features. This technology offers precise recording of a dancer's movements, unlocking new avenues for dance choreography, instruction, and staging. Nevertheless, current motion capture systems face limitations like insufficient accuracy and subpar real-time functionality, hindering their widespread practical use. Meanwhile, the dance world has witnessed the growing application of CAD technology.

Facial expression recognition has a wide range of applications in fields such as human-computer interaction, robotics, security, and healthcare. However, in the natural environment of the wild, the difficulty of facial expression recognition greatly increases due to factors such as changes in lighting, diversity in facial orientation, and occlusion. For this purpose, Bodapati et al. [2] proposed a deep convolutional neural network (CNN) architecture for field facial expression recognition. The recognition of facial expressions not only relies on local features but also needs to consider the contextual information of the image. Therefore, we combined the fully convolutional network (FCN) with CNN to achieve the fusion of contextual information. To achieve facial expression classification, we used a softmax classifier. This classifier can generate probability distributions for each category to determine the expression category of the input image. The experimental results show that compared with traditional facial expression recognition methods, the proposed deep CNN architecture has higher accuracy and robustness in field facial expression recognition. At the same time, we also found that the architecture performs well in dealing with lighting changes, facial orientation diversity, occlusion, and other issues. CAD enables the vivid representation of dance moves via 3D models, equipping dance creators with more streamlined and powerful design aids. Dance action facial recognition is a technology that combines computer vision and machine learning, aimed at identifying different dance actions by analyzing facial image sequences. This technology can help dance coaches better understand students' dance skills, provide more accurate feedback, and also provide new research tools for dance research. Chaabane et al. [3] proposed a dance action face recognition method based on statistical features and an SVM classifier. In order to convert continuous image sequences into formats that can be input into machine learning models, we use Dynamic Time Warming (DTW) method to encode feature sequences. DTW can adapt to sequences of different lengths and speeds, allowing for the comparison of feature sequences of different dance movements. It used a facial image dataset containing different dance movements for experiments, including various dances such as jazz, samba, street dance, etc. The experimental results show that the dance action face recognition method based on statistical features and SVM classifier has high accuracy and robustness. For some complex and fast dance movements, this method can still accurately identify them. However, traditional CAD processes are often unwieldy and time-consuming, failing to cater to the demands of rapid dance creation and iteration.

Big data encompasses vast datasets that surpass the processing capabilities of conventional software tools, thereby impeding timely capture, management, analysis, and organization. These datasets are crucial for guiding strategic business decisions. The synthesis of dance action images is an important research direction in the fields of computer vision and machine learning, aiming to generate realistic dance action images through image processing and machine learning techniques. Chandra et al. [4] proposed a dance action synthesis image method using hybrid neural networks and particle swarm optimization techniques. A hybrid neural network is a neural network structure that combines a Convolutional Neural Network (CNN) and a Recurrent Neural Network (RNN). CNN is used to extract local features of images, while RNN is used to process sequence data, such as time series data. In dance action synthesis images, hybrid neural networks can effectively combine the spatiotemporal information of dance actions to generate realistic dance images. To train and optimize hybrid neural networks, a dataset containing a large number of dance action images is required. The dataset should include dance action images with different styles, angles, and lighting conditions so that the model can learn feature representations in various situations. In practical terms, dance motion capture and CAD design, when bolstered by big data algorithm optimization, hold immense promise. This integrated approach not only elevates the precision and efficiency of motion capture and modelling for dance-related endeavours but also enriches virtual reality and game production with more lifelike dance animations. The posture and movements in dance training are crucial for improving dance skills and expressiveness. However, traditional dance training mainly relies on the observation and guidance of coaches, lacking objective and quantitative evaluation methods. To address this issue, Chen et al. [5] proposed a dance training posture motion capture research method based on mathematical similarity matching statistical analysis. It preprocesses the collected raw data, including denoising, normalization, and other operations, to improve the accuracy and reliability of the data. Perform statistical analysis on preprocessed data using mathematical similarity matching algorithms. This algorithm can classify and identify similar postures or movements based on the similarity of dance movements. Based on the results of similar matching, capture and evaluate the motion of dance training postures. By comparing standard movements or postures, the effectiveness of dance training can be objectively evaluated, and targeted feedback and suggestions can be provided.

Based on the above, this research will be devoted to building an efficient and accurate dance action capture model and applying it to CAD design, so as to realize fast capture, accurate reproduction and efficient editing of dance actions. Its main findings and innovations are embodied in the following aspects:

⊖ By introducing a big data algorithm, the efficient processing and accurate analysis of dance movement data are realized, and valuable feature information and motion rules are extracted, which provides strong support for subsequent model construction and optimization.

⊜ Using GCN's powerful spatial feature extraction ability, we successfully captured the complex spatial relationship and motion features between different parts of the dancer's body, which effectively improved the precision and stability of dance action capture.

⊛ Regarding the optimization of CAD modelling for dance movements, the deep learning-based approach presented in this research notably enhances modelling efficiency and quality. Additionally, it bolsters the model's adaptability to varying dance styles and complexities, broadening its potential for digital dance expression and visualization.

The structure of this article can be summarized as follows:

Section 1: Introduction

This opening segment outlines the research's background, objectives, and significance. It also surveys the current state of related research and existing outcomes, providing a foundational understanding of the study's importance and direction.

Section 2: Theoretical Foundations and Literature Review

Here, the theoretical underpinnings and critical technologies involved in the study are introduced, encompassing big data algorithms, dance motion capture, and CAD modelling. This section lays the groundwork for subsequent research and also summarizes pertinent literature.

Section 3: Constructing a Dance Motion Capture Model with GCN

This part delves into the utilization of big data algorithms for capturing and processing dance movements. It covers data collection, preprocessing, feature extraction, and the construction of the algorithm and model.

Section 4: CAD Modeling Optimization and Dance Movement Simulation Experiments

In this section, the focus is on using the captured dance movement data from the previous section for CAD modelling. The precision and efficiency of modelling are enhanced through optimization algorithms. Additionally, the design and execution of simulation experiments are detailed, along with a comprehensive analysis and discussion of the experimental results. This section validates the effectiveness and advantages of the big data algorithm-based approach to dance motion capture and CAD modelling optimization through comparative experiments.

Section 5: Conclusion and Outlook

The closing section offers a comprehensive summary of the entire study, highlighting key findings and innovations. It also identifies research limitations and potential future directions. Furthermore, it provides a forward-looking perspective on the application prospects and development trends of dance motion capture and CAD modelling optimization technology leveraging big data algorithms.

2 THEORETICAL BASIS AND LITERATURE REVIEW

Aerobics is a dynamic and rhythmic sport that requires high levels of posture and movement. In order to improve training effectiveness and accuracy, an effective method is needed to capture and analyze postures and movements in aerobics training. Chen and Alsemineari [6] proposed a method based on mathematical similarity-matching statistical analysis for capturing posture movements in aerobics training. In order to verify the effectiveness of the mathematical similarity-matching statistical analysis method in capturing posture movements in aerobics training, we conducted a series of experiments. The experimental results show that this method can accurately capture and analyze postures and movements in aerobics training, and has good stability and reliability. In addition, through comparative experiments, we also found that this method has high accuracy and robustness in handling complex and dynamic aerobics movements. Traditional 3D modelling and animation production methods often require a lot of manpower, time, and resources, and the recognition and modelling effects of single-view images are not ideal. To address these issues, Dvorožňák et al. [7] proposed a single view method for dance 3D modelling and animation based on big data algorithm optimization. It uses the extracted features and labels to train a deep-learning model. The goal of this model is to predict the 3D pose of the human body based on a single-view image. Using Generative Adversarial Networks (GANs) as the infrastructure, the model is trained through optimization algorithms to improve its accuracy and robustness. It uses a publicly available dance image dataset for experimentation and compares the proposed method with other methods. The experimental results show that this method can generate more accurate and realistic 3D human models and dance animations under single-view conditions. Meanwhile, by optimizing the application of algorithms, we further improved the performance and robustness of the model.

With the rapid development of artificial intelligence and computer vision technology, human motion recognition has been widely applied in many fields. Especially in the field of dance, accurately identifying and understanding dance movements is crucial for applications such as dance analysis, dance teaching, virtual reality, and augmented reality. In recent years, deep learning, especially graph neural networks (GNNs), has demonstrated strong capabilities in handling dance movements with complex spatiotemporal relationships and dynamic changes. Feng and Meunier [8] explored human dance action recognition technology based on deep learning graph neural networks. Deep learning is a branch of machine learning that simulates the hierarchical structure of the human brain by constructing multi-layer neural networks, enabling the processing and analysis of complex data. Graph neural networks are a special type of deep learning model specifically designed for processing graph-structured data. In dance action recognition, graph neural networks can effectively capture spatial and temporal dependencies in dance skeleton sequences. With the advancement of technology and the advent of the big data era, music and dance, two forms of art, are also undergoing unprecedented changes. In contemporary times, the optimization of big data algorithms has brought new perspectives and possibilities for the creation, performance, and dissemination of music and dance. Especially with the introduction of nonlinear processes and flowing roles, music and dance art become more dynamic and vibrant. In traditional music and dance creation, the process is often linear, following a fixed process from conception to implementation, and then to completion. However, in the era of big data, algorithm optimization has made non-linear creativity possible. This non-linear process allows artists to constantly explore, trial and error, and discover new ideas and expressions during the creative process. By analyzing a large amount of music and dance data, algorithms can discover potential connections between various styles, rhythms, and movements, thereby providing inspiration for artists. Meanwhile, non-linear processes also mean that artists can continuously adjust and optimize their works based on feedback, making their creations more efficient and precise. However, in the era of big data, these roles have become more fluid and diverse. Through big data analysis, everyone can become a creator and performer of music, making the expression of music and dance more diverse [9].

Human motion recognition and prediction are important research directions in the field of computer vision. It has broad application prospects in fields such as human-computer interaction, virtual reality, and motion analysis. In recent years, with the development of deep learning technology, significant progress has been made in human motion recognition and prediction based on three-dimensional skeletons. The method proposed by Li et al. [10] is based on a co-occurrence graph neural network, aiming to extract effective feature representations from 3D skeleton data. A symbiosis graph neural network is a type of graph neural network that can capture spatial and temporal relationships between skeletons. It first constructs a skeleton graph using the position and orientation information of skeleton joints. Then, a co-occurrence graph neural network is used to model the skeleton graph to capture the spatial and temporal dependencies between skeletons. By training and optimizing the co-occurrence graph neural network, we can extract effective feature representations from skeleton data. Human dance motion recognition is an important research direction in the field of computer vision, which has wide applications in dance analysis, motion capture, virtual reality, and other fields. With the development of deep learning technology, skeleton-based dance action recognition methods have received widespread attention. Liang et al. [11] proposed a non-uniform motion aggregation method based on graph convolutional networks for human dance motion recognition. Traditional dance motion recognition methods mainly rely on manually designed features, such as joint angles, speed, etc. These methods have limited recognition effectiveness for complex dance movements and dynamic changes. In recent years, deep learning techniques, especially convolutional neural networks and recurrent neural networks have achieved significant results in the fields of image recognition and video processing. On this basis, skeleton-based deep learning methods have been widely applied in dance action recognition. With the advent of the digital age, 3D dance modelling and animation have become a hot research field. Traditional 3D dance modelling methods typically require a large amount of annotated data, which not only increases the cost of data collection but also limits the model's generalization ability. To address this issue, Protopapadakis et al. [12] proposed an unsupervised 3D dance method based on big data algorithm optimization using stacked autoencoders. Collect a large amount of dance video data from publicly available dance video datasets. These data are not annotated and are used to train unsupervised 3D dance models. It utilizes deep learning techniques to extract key features from video frames, such as human posture, action trajectories, etc. These features will be used to train stacked autoencoders. Using the extracted features as input, the dance video data is encoded and decoded using a stacked autoencoder. This process is unsupervised and aims to learn the intrinsic structure and patterns of video data.

As an art form, the learning process of dance requires a lot of practice and feedback. In order to improve learning efficiency, Raheb et al. [13] proposed a dance interactive learning system optimized based on big data algorithms, aiming to provide learners with a real-time and personalized learning environment. It combines dance practice with game elements to make the learning process more interesting and attractive. By setting up challenge and reward mechanisms, stimulate learners' interest and motivation in learning. Using virtual reality (VR) technology to create a realistic dance scene for learners. By simulating the actual performance environment, it helps learners better adapt to the stage and improve their performance. Develop personalized teaching plans based on the learner's learning progress and ability level. The system automatically adjusts learning resources and difficulty to meet the needs of different learners. In addition to visual and kinesthetic stimuli, the system can also utilize various sensory stimuli such as audio and touch to assist teaching. For example, using sound feedback to help learners better grasp the sense of rhythm; and Providing guidance on force and posture through tactile devices, enhancing the learning experience and perception ability. With the advent of the big data era, dance motion recognition, as an important branch of computer vision, has also ushered in new development opportunities. Traditional dance action recognition methods often rely on manually designed features and simple models, and cannot fully explore and utilize the rich information contained in large-scale data. To address this issue, Wu et al. [14] proposed a perceptual-rich graph learning method based on big data algorithm optimization, aiming to improve the accuracy and robustness of dance action recognition. In early research, dance action recognition mainly relied on traditional computer vision techniques and manually designed features. However, these methods often fail to effectively handle complex dance movements and dynamic changes and are highly dependent on data size and quality. In recent years, the rapid development of deep learning technology has provided new solutions for dance action recognition. Deep learning technology can automatically learn the feature representation of data, with strong representation ability and robustness. Dance training is a process that requires long-term persistence and precise guidance. With the development of artificial intelligence and machine learning technology, reinforcement learning has been applied in dance training to provide personalized training plans and improve training effectiveness. Xin [15] explored the influencing factors of dance training effectiveness based on reinforcement learning and evaluated these factors. Accurately capturing the dance movements of learners is a prerequisite for reinforcement learning. Timely feedback can help learners understand their shortcomings and adjust training strategies accordingly. A good learning environment and abundant learning resources can help improve learning efficiency and provide learners with more opportunities to practice and explore. Personalized customization based on individual differences among learners and adapting to the ability needs of different learners are important factors in achieving effective training. Reinforcement learning-based dance training provides learners with a more personalized and efficient learning approach. By evaluating the influencing factors, we can further optimize the training system and improve training effectiveness.

3 CONSTRUCTION OF A DANCE MOTION CAPTURE MODEL BASED ON GCN

Dance motion capture technology has emerged in recent years alongside the swift advancement of computer vision and sensing technologies. This technology involves the capture of a dancer's bodily movements, which are then transformed into digital data for precise recording and analysis of dance moves. Currently, this technology relies heavily on a range of sensors and camera equipment, including optical motion capture and inertial sensor motion capture systems. The former system employs multiple high-speed cameras to detect marker points on the dancer's body, utilizing computer vision algorithms to reconstruct the dancer's three-dimensional motion path. This system has the advantages of high capture precision and strong real-time performance, but it is greatly affected by environmental factors such as light and occlusion. The inertial sensor motion capture system measures the motion information of the dancer, such as acceleration and angular velocity, by wearing multiple inertial sensors on the dancer, so as to calculate the dancer's motion trajectory. This system is portable and easy to use, but the capture precision is relatively low. Dance motion capture technology has a wide application prospect in dance creation, teaching and performance. In dance teaching, teachers can use this technology to evaluate and correct students' movements in real time to improve the teaching effect; In dance performance, this technology can be used to realize real-time rendering and interaction of dance movements and enhance the audience's viewing experience.

CAD design is a technical means to assist designers in product or engineering design with computers as the main tool. In the dance realm, CAD design finds its primary use in dance costumes, stage scenery, and the visual representation of dance moves. For dance costume design, CAD technology streamlines the process by enabling designers to swiftly sketch clothing styles, efficiently choose and blend fabrics, virtually preview costume effects, and thus elevate both design speed and quality. In terms of stage scenery design, CAD offers a three-dimensional modelling and rendering tool that facilitates quick design iterations and previews. By converting dance movements into 3D models or animations, they've empowered choreographers and dancers with a more intuitive understanding and revision process for dance routines. In the past few years, numerous scholars have dedicated themselves to extensive research in dance motion capture and CAD design. Regarding dance motion capture, the emphasis has been on refining the accuracy and real-time responsiveness of capture techniques, as well as exploring innovative methods and algorithms. For instance. In the domain of CAD design, research efforts center on enhancing design workflow efficiency and enhancing the user's design experience. The foundation of constructing an accurate action capture model lies in the comprehensive collection of dance movement data. To guarantee both diversity and precision in our dataset, we employed multiple data acquisition strategies. Firstly, we engaged professional dancers to perform a range of dance routines, capturing their motion trajectories using a state-of-the-art optical motion capture system. Secondly, we augmented our dataset with a wealth of dance movement data sourced from open dance databases. Throughout the data collection process, we maintained a controlled environment, minimizing the impact of variables such as lighting and occlusions on data quality. Additionally, dancers received proper guidance to ensure natural and fluid performances.

During the preprocessing phase, we underwent a rigorous data cleaning and organization process. This involved eliminating any abnormal or erroneous data resulting from equipment malfunctions or operator errors. Subsequently, we synchronized and aligned the data temporally and spatially, ensuring seamless integration between datasets from different sources. Furthermore, we normalized the data to mitigate the influence of individual body differences among dancers on motion capture accuracy.

GCN, a deep learning architecture tailored for processing graph-structured data, plays a pivotal role in our analysis. By defining convolution operations on the graph, GCN effectively captures the spatial relationships between nodes and extracts valuable feature information. For a detailed visual representation of the GCN model structure, please refer to Figure 1.

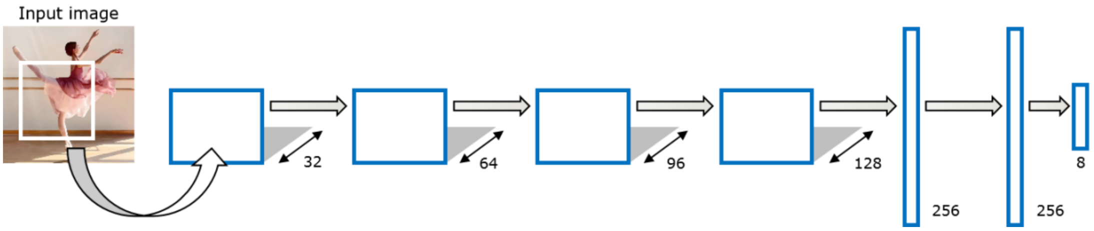

Figure 1: Model structure diagram.

To determine the motion posture data for bones and joints in the human motion model, converting sensor-derived posture data is essential. Establishing an apt coordinate system facilitates swift posture transformations. Computing the mutual transformation of a common point $P_A\ X,Y,Z$ across three coordinate systems involves translational T and rotational R adjustments among the axes.

$$P_A^{'} = RP_A + T \tag{1}$$

The rotational measure often assumes a baseline angle of 0 or 180 degrees for every axis, whereas the translational measure signifies the 3D length parameters of the model's skeletal structure.

In the context of human joint recognition, Kinect connects individuals through their joint points, enabling it to determine the coordinate locations of humans by assessing both frontal and lateral joint points. The acquisition of pixel-specific information is facilitated through the recognition and inference of body components. Subsequently, the density estimator for these body components is delineated as follows:

$$f_c\ \hat{x} \propto \sum_{i=1}^{N} w_{ic} \exp\left(-\left\|\frac{\hat{x}-\hat{x}_i}{b_c}\right\|^2\right) \tag{2}$$

Where $\hat{x}$ is the 3D space coordinate, N is the number of pixels, w_{ic} is the pixel weight, $\hat{x}_i$ represents the projection of pixel x_i into the world space, and b_c represents the width of each component.

The precise driving equation can be ascertained through forward dynamics:

$$\vec{S}_i = \vec{R}_{parent\ i} * l_i \vec{O}_i + \vec{S}_{parent\ i} \tag{3}$$

Where $\vec{S}_i, \vec{R}_{parent\ i}$ represents the location of both the i key point and its parent node, while $\vec{O}_i$ denotes the relative direction vector (a unit vector indicating direction) extending from the parent node to the present i key point, the rotational and extensional relationships can be conveyed by utilizing the bone length l_i alongside the rotation matrix $\vec{R}_i$ encompassed within the shape parameters.

In dance action capture, the dancer's body can be represented as a graph structure composed of joint points. Each joint point corresponds to a node in the graph, and the connection relationship between joint points corresponds to the edge in the graph. By using GCN to process such graph structure data, we can effectively capture the spatial relationship and motion characteristics of dancers' body parts. In the dance motion capture model, GCN plays a pivotal role in extracting distinct features from the dancer's joint data. These extracted features then serve as inputs for the downstream neural network model, facilitating deeper processing and analysis. This integrated approach enables precise capture and streamlined modelling of dance movements.

To mitigate challenges posed by motion variations during feature extraction, we compute the centroid points of human contours independently and designate them as the reference points for image coordinates. This ensures a consistent and reliable basis for feature extraction, enhancing the overall accuracy and efficiency of the motion capture process. Assuming x_c, y_c represents the centroid coordinate, N_b denotes the cumulative count of pixels along the boundary, and x_i, y_i designates the pixel points lining the perimeter, we can formulate as follows:

$$\begin{cases} x_c = \dfrac{1}{N_b} \displaystyle\sum_{i=1}^{N_b} x_i \\ y_c = \dfrac{1}{N_b} \displaystyle\sum_{i=1}^{N_b} y_i \end{cases} \tag{4}$$

To reduce the effects of image scale and signal duration on training and recognition, we use minimum-maximum normalization and equal-interval resampling to adjust signal amplitude and length.

Predicting 3D posture, which relies heavily on determining an individual's intersection angles, is a complex task that requires considerable effort and research. The challenge lies in the fact that 2D posture, projected onto the image plane, changes significantly with viewpoint, lacking invariance. Our proposed approach aims to reconstruct 3D posture from its 2D representation.

After obtaining the 3D posture, normalization becomes essential. Using 14 3D joint points, we normalize by setting the distance from the neck to the hip center as the unit of length. The hip position is defined as the midpoint between the left and right hips, with the neck serving as the origin of the coordinate system. The normalization process for the body joints is as follows:

$$J_i^{'} = \frac{J_i - J_{neck}}{H_{nh}} \tag{5}$$

Where $J_i \in R^2$ represents the location of the body joint i, $J_i^{'}$ denotes the normalized position of J_i, J_{neck} indicates the neck position, and H_{nh} measures the distance from the neck to the hip.

The construction process of the dance action capture model based on GCN is as follows:

Definition of graph structure: according to the dancer's body structure, a graph structure is defined, in which nodes represent the dancer's joint points and edges represent the connection relationship between the joint points. This graph structure will be used as the input of GCN.

Feature extraction: Employing GCN, the features inherent in the dancer's joint data are extracted. This process enables GCN to discern the spatial relationships and motion patterns among joint points, encapsulating these insights within vectors.

Model training: The extracted features are fed into a downstream neural network for rigorous training. A comprehensive dataset of dance movements guides this supervision, facilitating the refinement of model parameters via the backpropagation algorithm.

Motion capture and reconstruction: Once trained, the model proves adept at capturing and reconstructing novel dance movements. By simply introducing new dance action data, the model seamlessly generates corresponding capture outcomes. These results manifest as either a 3D skeletal animation or a precise sequence of motion trajectories.

In our investigation, the displacement field, arising from the amalgamation of optical flows across consecutive frame pairs, emerges as the foundational attribute of optical flows.

The optical flow utilizes a transformative method to ascertain the displacement vector field d_t for pixel x,y across two successive time frames, denoted as t and $t+1$. Furthermore, it pinpoints the function u,v that minimizes the energy function. To compute the optical flow, a blend of data terms and smoothing elements is employed to enhance the overall energy function. This mathematical concept can be formulated as follows:

$$E_{Global} = E_{Data} + \lambda E_{Smooth} \tag{9}$$

In this context, E_{Data} refers to a data item that gauges the coherence between the optical flow and the input image. E_{Smooth} represents a smoothing term, while λ signifies the flow field, which tends to vary in a gradual manner. Lastly, E_{Global} denotes the optimization of the overall energy.

Let pixel $p\ x,y$ possess a gray value of $i\ x,y,t$ at time t. Following the elapse of time Δt, the 2 pixel shifts to position $x+\Delta x, y+\Delta y, t+\Delta t$, and its gray value transforms to $I\ x+\Delta x, y+\Delta y, t+\Delta t$. Since both points correspond to the identical pixel, their gray values remain consistent, leading to the derivation of the subsequent equation:

$$I\ x+\Delta x, y+\Delta y, t+\Delta t\ = I\ x,y,t \tag{7}$$

Optical flow demonstrates vulnerability to noise, scale variations, and motion direction shifts. In contrast, the optical flow direction histogram excels at encapsulating motion data while maintaining robustness against scale and directional changes, making it an indispensable asset in numerous motion recognition research frameworks. By amalgamating the optical flow histograms of all cells contained within a predefined block, we can concatenate them to craft the block's distinctive optical flow histogram feature vector.

To standardize the histogram, we utilize the $2L$ - $norm$ normalization technique:

$$2L\text{ - }norm, v \leftarrow \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}} \tag{8}$$

Finally, the feature vectors, which are derived from the optical flow histograms of each block, are combined to form the comprehensive HOF (Histogram of Optical Flow) feature for the whole image. In order to ascertain the exact dimensions, one may employ the following formula:

$$V = binNum \times cellNum \times blockNum \tag{9}$$

In this setting, $binNum$ represents the number of directional columns, $cellNum$ indicates the tally of cell grids contained within every block, and $blockNum$ embodies the overall count of blocks existent within the image.

In this way, this article successfully constructs a dance action capture model based on GCN. The model can not only accurately capture the movements of dancers, but also effectively deal with complex dance movements and changeable body postures. This provides strong support for the subsequent application of dance creation, teaching and performance.

4 OPTIMIZATION OF CAD MODELING AND SIMULATION EXPERIMENT OF DANCE MOVEMENTS

4.1 Optimization of CAD Modeling of Dance Movements

In the field of dance, CAD modelling mainly involves the digital expression and visual display of dance movements. By capturing the movements of dancers, these movements are transformed into 3D models by using CAD software, so as to realize the accurate recording and reproduction of dance movements. The basic principles of CAD modelling include geometric modelling, motion modelling and physical modelling [16]. Geometric modelling pays attention to the mathematical expression of object shape and structure; Motion modelling involves the position and attitude changes of objects in space; Physical modelling considers the deformation and motion law of objects when they are acted by external forces. In the CAD modelling of dance movements, these principles are applied to capture and express the dancer's body posture and motion trajectory.

Aiming at the problems existing in the process of CAD modelling of dance movements, this study proposes to use a big data algorithm to optimize. First of all, by collecting a large number of dance movement data, a rich data set is constructed. Then, the algorithm is used to analyze and process these data, and useful feature information and motion law are extracted. In the process of optimization, this article pays special attention to the following aspects: first, improve the precision and efficiency of modelling, and ensure that the captured dance movements can be accurately transformed into 3D models; The second is to enhance the expansibility and reusability of the model so that it can adapt to different styles of dance movements and scene requirements; The third is to optimize the user experience and operation process, and reduce the difficulty of use and learning cost. In this paper, the HOG feature is obtained from the image subsequent to edge feature manipulation. This feature serves as a crucial aspect in target detection for both computer vision and digital image processing. The procedure for extracting HOG features, at a glance, is depicted in Figure 2.

By constructing a depth neural network model, the dancer's motion trajectory is taken as input, and the corresponding 3D model parameters are output. During the training process, the parameters and structure of the model are constantly adjusted, so that it can better fit the actual data and improve the precision accuracy.

4.2 Design and Implementation of Simulation Experiment

In order to verify the effect of the optimized dance motion capture model in CAD modelling, a series of simulation experiments were designed and implemented. This section selects a representative dance action as the experimental object and uses the optimized model to capture and model it. Then, the modelling results are compared and analyzed with traditional CAD modelling methods. In the simulation experiment, this article pays special attention to the following aspects: first, the accuracy and integrity of the model capturing dance movements; Second, the efficiency and quality of generating 3D models from models; third is the adaptability of the model to dance movements with different styles and difficulties. Figure 3 is a continuous dance modelling action frame.

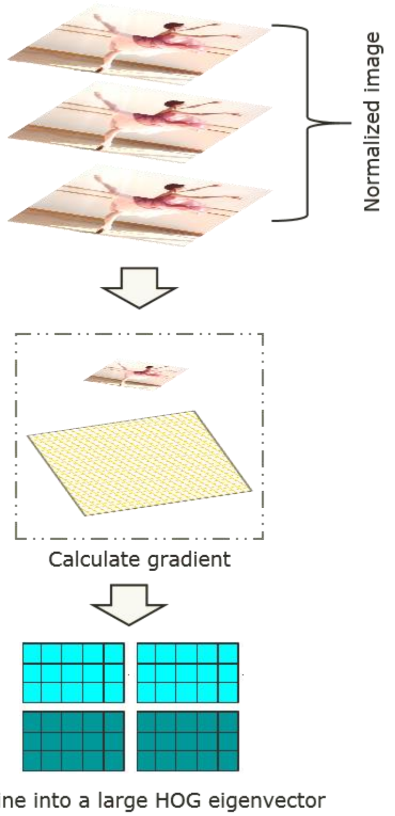

Figure 2: Extraction process of hog features.

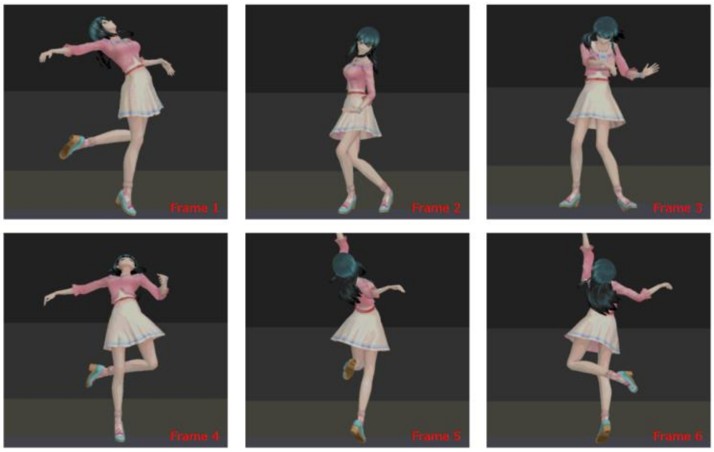

Figure 3: Dance modelling effect.

Accuracy means that the model can accurately capture the details of dancers' movements, including joint angles and motion trajectories. The accuracy of the model capturing dance movements is shown in Figure 4.

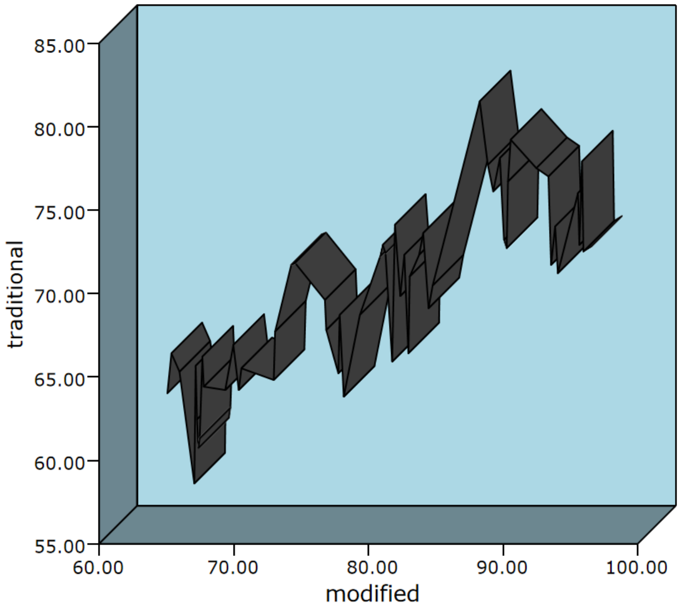

Figure 4: Accuracy of a model capturing dance movements.

Figure 4 illustrates the model's proficiency in capturing dance movements, achieving an impressive accuracy of approximately 96%. This high proportion underscores the model's exceptional ability to accurately detect and recognize dance movements. More precisely, it signifies that 96% of the dance movements evaluated by the model are precisely captured and identified, while only 4% remain either inaccurate or unrecognized. Despite the model's remarkable accuracy, it is imperative to acknowledge and address the 4% margin of error. Future endeavours could focus on refining the algorithm, expanding the training dataset, and enhancing the model's architecture to further minimize the error rate and elevate the model's precision. Additionally, Figure 5 exhibits the efficiency of generating a 3D model, whereas Figure 6 showcases the quality of the generated model. The results in Figure 5 show that the time consumption of the model in generating the 3D model is significantly reduced and the efficiency is significantly improved. This is mainly due to the application of optimization algorithms and the reasonable allocation of computing resources. The optimization algorithm can reduce unnecessary calculations while ensuring the precision of the model, thus improving the calculation efficiency. Ensuring the judicious allocation of computing resources allows the model to fully leverage these resources during calculations, thereby preventing any wastage. As evident from Figure 6, the quality of the 3D model produced by our model has undergone significant enhancement. The generated 3D model not only boasts high fidelity but also accurately replicates the movements and forms of dancers, exhibiting richer and more intricate details. This substantial improvement is primarily attributed to the incorporation of big data algorithms and advancements in model optimization technology. Big data algorithms enable the extraction of valuable feature information from extensive dance movement datasets, providing robust data support for model development. Furthermore, model optimization techniques refine the precision and stability of the model, ensuring that the generated 3D model meets the demands of practical applications. Figure 7 illustrates the model's recognition accuracy across different styles of dance movements.

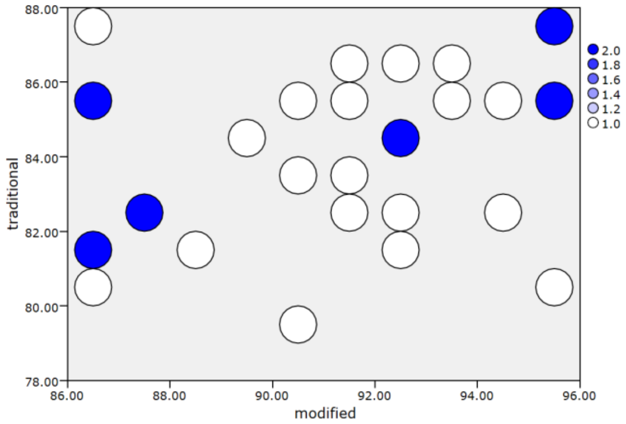

Figure 5: Efficiency of 3D model generation.

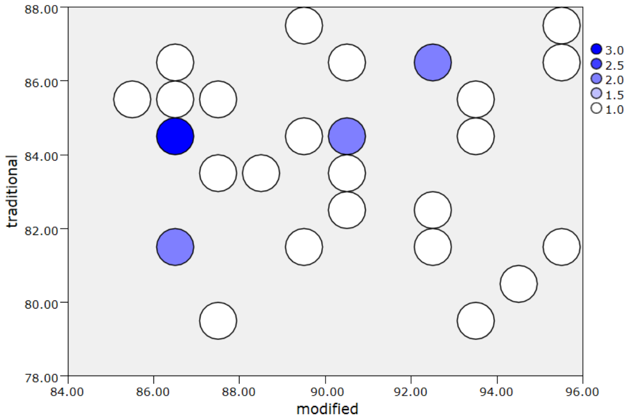

Figure 6: Quality of 3D model generated.

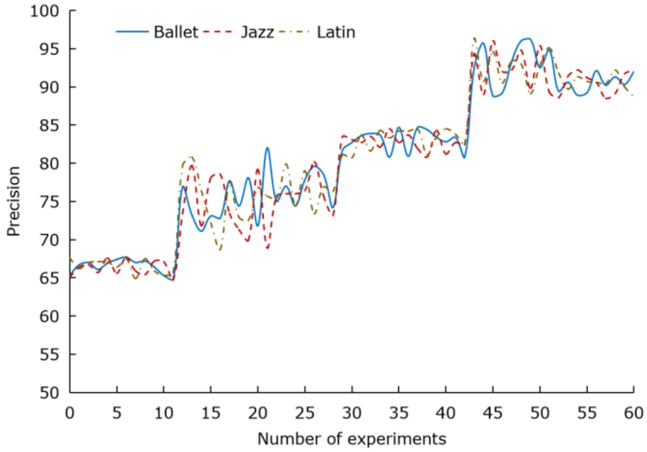

Figure 7: Model identification precision.

From the overall trend, the model shows a relatively high recognition accuracy for different styles of dance movements. This shows that the model has strong generalization ability and can adapt to various styles of dance movements. This is very important in practical application, because different dance styles correspond to different movement characteristics and laws, and the model needs to be flexible and adaptable enough to accurately capture and identify these movements. For some common dance styles with obvious movement characteristics, such as modern dance and ballet, the recognition precision of the model is very high, almost reaching more than 95%. This shows that the model is very accurate in capturing and identifying the movements of these dance styles, and can effectively extract the key features and movement rules of the movements. However, for some dance styles with complex action characteristics and high requirements for details, such as street dance and folk dance, the recognition precision of the model is relatively low, but it is also within an acceptable range. This is because the movements of these dance styles change rapidly and the consistency between movements is strong, which brings certain challenges to the capture and recognition of the model. But even so, the model can still capture and identify the main features and movement trends of these dance movements.

4.3 Analysis of Simulation Experiment Results

The simulation results in the last section show that the optimized dance motion capture model has achieved remarkable improvement in CAD modelling. Specifically, the following aspects have been significantly improved:

First, the accuracy and integrity of the model to capture dance movements have been significantly improved, with an accuracy of about 96%. Through the training and learning of deep learning algorithms, the model can capture the dancer's movement trajectory and posture change more accurately, and transform it into a 3D model completely.

Secondly, the efficiency and quality of 3D model generation have been improved obviously. The optimized model adopts more efficient algorithms and data structures and can generate high-quality 3D models in a short time. Furthermore, the model has good expansibility and reusability and can adapt to different styles and difficulties of dance movements.

Thirdly, the adaptability of the model to dance movements with different styles and difficulties has been enhanced. Whether it is a common dance style or a dance style with complex motion characteristics, the model can capture and identify its motion characteristics and motion rules well. Through the optimization and learning of big data algorithms, the model can better understand and handle dance movements with different styles and difficulties, which provides a wider application space for dance creation and teaching.

In conclusion, the dance action CAD modelling optimization approach presented in this article, which is rooted in big data algorithms, has yielded outstanding outcomes. This method serves as a robust technical foundation for the digital representation and visual exhibition of dance artistry.

5 CONCLUSIONS

The aim of our study centers on refining dance motion capture and CAD design processes through the utilization of big data algorithms. This approach targets enhanced precision in capturing dance movements and greater efficiency in CAD design. Our findings reveal that the model excels in accurately capturing intricate dance movements, achieving an impressive accuracy rate of approximately 96%. This signifies that the model adeptly grasps and identifies minute details of dancers' movements, encompassing joint angles, motion trajectories, and other nuances.

Moreover, notable advancements have been made in the efficiency and quality of 3D model generation. The model is now capable of producing high-calibre 3D models in a fraction of the time previously required. Additionally, the model demonstrates remarkable adaptability to a diverse range of dance styles. Whether confronted with a common dance form or one characterized by complex motion patterns, the model adeptly captures and interprets its unique movement signatures and rules. This underscores the model's strong generalization capabilities, enabling it to seamlessly adapt to a wide array of dance styles.

The research presented in this article carries significant theoretical and practical implications. From a theoretical perspective, this study contributes to the exploration of big data algorithms in the context of dance motion capture and CAD design, thereby enriching and advancing theoretical research in related domains. By constructing a dance motion capture model rooted in big data algorithms, we anticipate elevating the precision of motion capture and providing robust support for the digital evolution of dance artistry.

On a practical level, the outcomes of this research hold direct applicability in dance creation, education, and performance. By refining dance movement capture technology, dancers and choreographers can more efficiently capture and document dance movements, ultimately enhancing the quality and expediting the process of dance creation. Furthermore, improvements in the CAD design process empower dance creators with greater convenience in designing and modifying dance movements, leading to shortened creative cycles and reduced costs.

Beyond these immediate applications, this study serves as a catalyst for the deeper integration of dance art and modern technology. By introducing contemporary scientific and technological advancements such as big data algorithms, dance art is poised to transcend traditional boundaries and achieve cross-disciplinary integration with fields like digital media and virtual reality. This fusion promises to usher in a new era of immersive and captivating visual experiences for audiences.

Looking ahead, we recognize the importance of enhancing the user experience. Future efforts will focus on streamlining the learning curve, minimizing costs, and prioritizing the user-friendliness and accessibility of the interface. This approach aims to make dance motion capture and CAD design technology based on big data algorithm optimization more accessible and enjoyable for a broader audience.

Jing Zhao, https://orcid.org/0000-0002-3823-4879
Jing Li, https://orcid.org/0009-0006-0922-3317

REFERENCES

[1] Ahmad, T.; Jin, L.; Zhang, X.; Lai, S.; Tang, G.; Lin, L.: Graph convolutional neural network for human action recognition: A comprehensive survey, IEEE Transactions on Artificial Intelligence, 2(2), 2021, 128-145. https://doi.org/10.1109/TAI.2021.3076974

[2] Bodapati, J.-D.; Srilakshmi, U.; Veeranjaneyulu, N.: FERNet: a deep CNN architecture for facial expression recognition in the wild, Journal of the Institution of Engineers (India): Series B, 103(2), 2022, 439-448. https://doi.org/10.1007/s40031-021-00681-8

[3] Chaabane, S.-B.; Hijji, M.; Harrabi, R.; Seddik, H.: Face recognition based on statistical features and SVM classifier, Multimedia Tools and Applications, 81(6), 2022, 8767-8784. https://doi.org/10.1007/s11042-021-11816-w

[4] Chandra, D.; Hegadi, R.; Karmakar, S.: Facial expression synthesis images using a hybrid neural network with particle swarm optimisation techniques, International Journal of Biomedical Engineering and Technology, 31(1), 2019, 64-83. https://doi.org/10.1504/IJBET.2019.101052

[5] Chen, Q.; Albarakati, A.; Gui, L.: Research on motion capture of dance training pose based on statistical analysis of mathematical similarity matching, Applied Mathematics and Nonlinear Sciences, 7(2), 2021, 127-138. https://doi.org/10.46719/dsa202029335

[6] Chen, Q.; Alsemmeari, R.-A.: Research on aerobics training posture motion capture based on mathematical similarity matching statistical analysis, Applied Mathematics and Nonlinear Sciences, 7(2), 2021, 203-216. https://doi.org/10.46719/dsa202029340

[7] Dvorožňák, M.; Sýkora, D.; Curtis, C.; Curless, B.; Sorkine, H.-O.; Salesin, D.: Monster mash: a single-view approach to casual 3D modeling and animation, ACM Transactions on Graphics (ToG), 39(6), 2020, 1-12. https://doi.org/10.1145/3414685.3417805

[8] Feng, M.; Meunier, J.: Skeleton graph-neural-network-based human action recognition: A survey, Sensors, 22(6), 2022, 2091. https://doi.org/10.3390/s22062091
[9] Hsueh, S.; Alaoui, S.-F.; Mackay, W.-E.: Deconstructing creativity: non-linear processes and fluid roles in contemporary music and dance, Proceedings of the ACM on Human-Computer Interaction, 3(10), 2019, 1-21. https://doi.org/10.1145/3359305
[10] Li, M.; Chen, S.; Chen, X.; Zhang, Y.; Wang, Y.; Tian, Q.: Symbiotic graph neural networks for 3D skeleton-based human action recognition and motion prediction, IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(6), 2021, 3316-3333. https://doi.org/10.1109/TPAMI.2021.3053765
[11] Liang, C.; Yang, J.; Du, R.; Hu, W.; Tie, Y.: Non-uniform motion aggregation with graph convolutional networks for skeleton-based human action recognition, Electronics, 12(21), 2023, 4466. https://doi.org/10.3390/electronics12214466
[12] Protopapadakis, E.; Rallis, I.; Doulamis, A.; Doulamis, N.; Voulodimos, A.: Unsupervised 3D motion summarization using stacked auto-encoders, Applied Sciences, 10(22), 2020, 8226. https://doi.org/10.3390/app10228226
[13] Raheb, K.-E.; Stergiou, M.; Katifori, A.; Ioannidis, Y.: Dance interactive learning systems: A study on interaction workflow and teaching approaches, ACM Computing Surveys (CSUR), 52(3), 2019, 1-37. https://doi.org/10.1145/3323335
[14] Wu, C.; Wu, X.-J.; Kittler, J.: Graph2Net: Perceptually-enriched graph learning for skeleton-based action recognition, IEEE Transactions on Circuits and Systems for Video Technology, 32(4), 2021, 2120-2132. https://doi.org/10.1109/TCSVT.2021.3085959
[15] Xin, L.: Evaluation of factors affecting dance training effects based on reinforcement learning, Neural Computing and Applications, 34(9), 2022, 6773-6785. https://doi.org/10.1007/s00521-021-06032-4