# Building a Parallel Corpus for English Translation Teaching Based on Computer-aided Translation Software

Xiaolin Wang

Sichuan university of culture and arts, Sichuan Mianyang 621000, China, sanmao120627@126.com

Corresponding author: Xiaolin Wang, sanmao120627@126.com

**Abstract.** With the increasing level of modern information technology application, computer data processing technology has gradually become the main and classical auxiliary technology application means in the daily working life of the majority of social members. In the practical workflow of dance art design, in order to continuously meet the new demands put forward by the members of the community to dance art appreciation, and at the same time, better improve the quality and rationality of dance movement design, the staff gradually choose to use the computer as an auxiliary tool, applied to the practical design process. In this paper, while the research on artificial intelligence is going deeper and deeper, the artistic value hidden in the dance movement design process is gradually discovered. In this paper, the application of computer technology aids to the design of dance movements is taken as the main research background. On the basis of a simple analysis of the practical content of dance movement design, the recent development and innovative approaches of various dance movement designs with the assistance of computer software technology are explained and introduced in detail.

**Keywords:** computer assistance; English translation software; corpus construction.
**DOI:** https://doi.org/10.14733/cadaps.2021.S4.175-185

## 1 INTRODUCTION

The implementation and execution of dance movement design, driven and influenced by computer-aided action, actually refers to the designer's actual body language, and through the integration of different forms of emotions, the goal of integrating dance movement design and emotional content is achieved. Bothma et al. [1] have pointed out that the presentation of various dance movements of the dancers themselves can even be regarded as a materialized mode of expression of dance design emotions in the body language. For this reason, if designers want to design more excellent dance movement works, they must first fully understand the complexity of each dancer's body parts and the various movement designs in the process of constructing different dance formations [2]. In the process of constructing different dance forms, the complexity of various movement designs [3]. In terms of the basis of dynamic design of body composition patterns, the basic composition patterns of dance movement people's own body movements are not much different from those of other ordinary members of society. However, the proportional demands on the body structure of dance performers are often more stringent than those of ordinary people. In general, the basic skeletal structure and production of dancers is characterized by a strict adherence to the basic growth characteristics of three long and one short. Nooh et al. [4] proposed that only those dancers who are born with this production, through continuous efforts and training, can maximize the performance effect of the various dance movement designs while maintaining the advantages of basic musculoskeletal growth. Among the relevant academic articles on dance movement dynamic design essentials, famous dance movement designers have made an in-depth analysis of rigid human movement. Caro-Codón et al. [5] proposed that dance designers have pointed out that if a dance performer is placed in a virtualized design environment through the projection of skeletal perspective, it is possible to maximize the performance effect of the finished dance design.

The body structure of a dance performer can be seen as a form of rigid union placed in a virtual skeletal construction environment. In the practical work of rigid dance movement design, technicians often optimize the design of a given dance movement by referring to real dancers and the changing state of their own body muscles [6]. In Western capitalist countries with a higher level of economic development, the implementation of rigid dance movement design also selectively combines the use of virtualized media in the creation of dance art, so as to improve the creativity of dance movement design and better enhance the artistic appreciation value of the designated dance art creation products [7]. The theoretical content of applying the functional sections of computer-virtualized movement design to the process of opening dance movement design work was first proposed in the mid-1950s in the international context. Western capitalist countries with a high level of capital and economic development took the lead in proposing the application of computer animation in the design of modern dance movements, combined with the technical means of design compilation. On the basis of the dance preparation system, which was produced by Júnior et al. [8], technicians have conducted repeated experiments and analyses on the application of new dance movement design methods in combination with the application of computerized dance movement design model plates [9].

With the help of the new computer function processing technology, the time required for dance movement design has gradually shortened from the previous working cycle of ten to twenty-four hours to the current time fluctuation range of five to ten minutes. Exner et al. [10] proposed that the degree of damage to the dancers' own skeletal development and muscle molecular structure caused by various dance movements has been further enhanced from the original ten-digit fuzzy estimation to a more accurate measurement range of two decimal places. Furthermore, with the continuous evolution of dance types and forms and the proliferation of dance videos, how to navigate dance videos quickly and efficiently is now the main problem faced. The problem of complex and repetitive movements in music and dance videos poses a problem for the analysis and recognition of dance movements [11]. Therefore, in order to increase the rate of dance video analysis and reduce the computational complexity, extracting keyframes of music and dance videos will be an effective solution. Keyframes are defined as some image frames in the video that have little redundancy and are representative of the video. For dance videos, keyframe extraction is the process of removing redundancy, and at a certain moment when the music is expressive and the difference between the current frame and the previous keyframe is large, the image frame is identified as a new keyframe. Through the keyframe collection, the user can get a quick understanding of the content of the dance video, which shortens the time the user spends browsing the dance video and saves the storage space of the video.

In the practical workflow of dance art design, in order to better improve the quality and rationality of dance movement design while continuously satisfying the new demands put forward by the members of the society for the appreciation of dance art, the staff gradually chooses to apply computers as auxiliary tools in the practical design process. In this paper, while the research on artificial intelligence is going deeper and deeper, the artistic value hidden in the dance movement design process is gradually discovered. In this paper, the application of computer technology aids to the design of dance movements is taken as the main research background. On the basis of a simple analysis of the practical content of dance movement design, the recent development and innovation of various dance movement designs with the assistance of computer software technology are explained and introduced in detail.

# 2 COMPUTER-AIDED EXTRACTION OF DANCE MOVES

## 2.1 Keyframe Extraction for Music and Dance Videos

With the continuous evolution of dance genres and forms and the increasing number of dance videos, how to navigate through dance videos quickly and efficiently is the main problem faced today. The problem of complex and repetitive movements of music and dance videos poses a problem for the analysis and identification of dance movements. Therefore, in order to increase the rate of dance video analysis and reduce the computational complexity, extracting keyframes of music and dance videos will be an effective solution. Keyframes are defined as some image frames in the video that have little redundancy and are representative of the video. For dance videos, keyframe extraction is the process of removing redundancy, and at a certain moment when the music is expressive and the difference between the current frame and the previous keyframe is large, the image frame is identified as a new keyframe. Through the keyframe collection, the user can get a quick understanding of the content of the dance video, which shortens the time the user spends browsing the dance video and saves the storage space of the video.

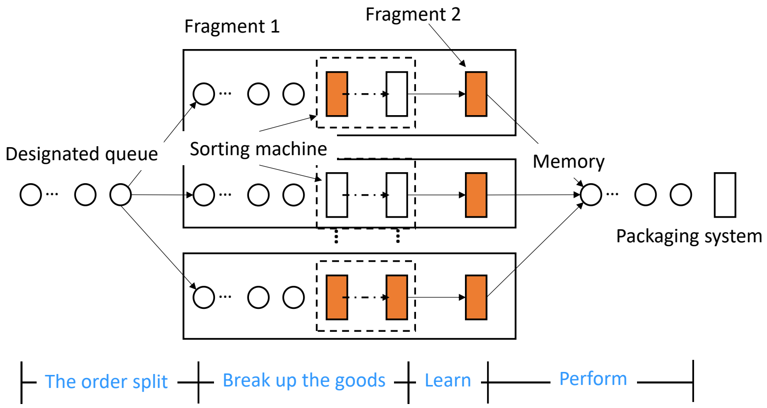


**Figure 1**: Keyframe extraction idea.

The variability of dance movements and the problem of too many redundant movements pose a great challenge for keyframe extraction. In order to extract a set of keyframes that has few redundancies and can summarize the content of the video, this paper will perform optical flow computation on a sequence of images of a dance movement video after framing, which is able to match large displacements and also allows optical flow estimation for smaller objects. The optical flow map reflects the movement direction and speed of the dance movements, while the entropy calculation counts the amount of information in each optical flow image. However, for music and dance videos, the musical features are also crucial. Therefore, this paper will extract the envelope line features and music energy features of music and fuse the music features with the entropy sequence to obtain a music-related entropy sequence. Finally, this paper identifies the next frame that is greater than the threshold as a keyframe, and the set of keyframes is the video summary of the dance video. The main idea of this method is shown in Figure 1.

The traditional optical flow algorithm is limited at certain moments, and although it is able to track the target in the image, the optical flow calculation is not very effective for motion discontinuities, occlusion phenomena and large displacements. Therefore, the optical flow method will be relevant by continuously analyzing the problem, process and exploring the corresponding solutions. In the keyframe extraction algorithm based on motion analysis, they were the first to introduce the optical flow method to perform optical flow computation on image sequences under each lens and select local minima as keyframes. However, the lens segmentation is not accurate enough and the traditional optical flow method is poorly targeted. Therefore, in this paper, the LDOF (Large Displace Optical Flow) method proposed will be used to obtain the motion characteristics of dance videos. This can be used for dance videos to accurately calculate every change in dance movements without missing those small body parts (hands, feet, etc.).

$$E(w) = E_c(w) + \alpha E(smooth) + \beta Match \tag{1}$$

$$E_c(w) = \int_{color} \psi(I_3(x) - I(y))dx \tag{2}$$

$$E(smooth) = \int_{color} \nabla \psi(q(x) - movement)dx \tag{3}$$

$$Match = \int_{color} (\nabla \psi(q(x))^2)dx \tag{4}$$

where $\alpha$, $\beta$ and $\gamma$ are adjustable weighting parameters and E is an assumption of luminance invariance that applies to both color and grayscale images. The effect of illumination is unavoidable, so a gradient constraint w is added to reduce the effect of illumination. w is then smoothed by H(smooth). The last two items are used to construct descriptor matches and their minima through a multivariate model and optimization.

## 2.2 Image Motion Integration

Image entropy has been widely used in extracting key frames. The work done by image entropy focuses on the statistics of the features in the image, in the reported, the entropy value of each image in the image sequence is calculated and then the entropy value corresponding to the current frame is compared with the previous frame, if it is greater than the threshold value it is considered as a key frame. The entropy calculation for the image frames takes into account only the features of the image, but for the dance video, it is the variation of the movement that is critical. Therefore, the method used in this paper is to calculate the entropy value of each optical flow diagram derived from the above stage, as the input of that stage, and this paper calculates the entropy value on the grayscale image, although there is information loss compared to the color image, but the information required for the entropy calculation does not change, and the speed of the calculation has been improved, and it is also easy to store. By calculating the entropy value of the current optical flow diagram in chronological order.

$$img\_i = \sum_{i}^{q} \int (\nabla \psi(q(x))^2)dx \tag{5}$$

Dance is a kind of art with human body movements as the main form of expression. Among the different subjects of dance works, dance and music have the most intimate relationship, and the accompanying music in a dance work plays an extremely crucial role in the quality of a dance work. Without good music, it is very difficult to produce beautiful dance works. Dance needs music to render the atmosphere and support the theme. The accompanying music in a dance video fully reflects the characteristics of the dance movements. At the choreographic stage, the dance needs to be choreographed in different forms based on the characteristics of the music. Likewise, the soundtrack needs to be matched with appropriate accompanying music according to the rhythm and style of the dance movements. Scheirer has experimentally demonstrated that the envelope line features reflect the original information of the music. Therefore, in this paper, the envelope line feature curve of the music is fused with the entropy value sequence to obtain an entropy value sequence that is related to the music. In this paper, the accompanying music will be extracted from the video, the WAV format will be read, and then the envelope line features and energy features of the music will be extracted to prepare for the subsequent feature fusion, which is useful for event detection of the music. Figure 2 is a graph of the result of envelope line feature extraction of the accompanying music of the dance video Theodora_Afraid_Wide. It shows the entire profile of the audio signal.

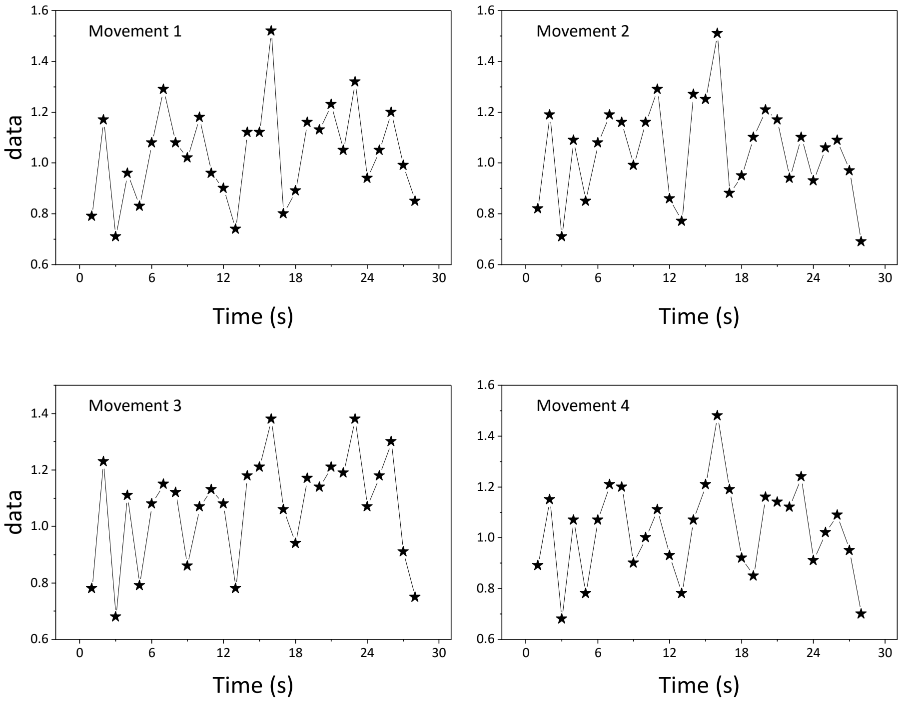


**Figure 2:** Envelope characteristics of the accompanying music.

The extraction of the energy signature of the audio is firstly performed by windowing the audio x( j) to obtain the kth frame. The audio signal exists in y, the length of y is N, the sampling rate is fs, the length of each take is wlen, the displacement of the front and back frames is dis, and the overlap between the two frames is olap =wlen -dis. Therefore, for an audio signal of length N.

$$fs = (N - drop) / drop_i \tag{6}$$

The average amplitude of the audio is then calculated, i.e., the energy signature of the audio. The audio feature sequence and the entropy value sequence can then be aligned. Finally, the feature fusion is performed by multiplying the audio feature sequence with the entropy value sequence to obtain a music-related entropy value sequence.

## 3 DANCE MOVEMENT DESIGN

### 3.1 Dance Video Clip Search

Video clip retrieval is performed by finding video clips in the video that are similar to the content of the query clip and then locating the video clip in some way. Before similarity matching, the structure of the current video should be considered and which level of the video should be selected to perform the similarity matching process. The introduction of keyframes opens up a new way of thinking. Keyframes can summarize a video well and improve the efficiency of browsing the video. The feature selection focuses on the retrieval rate and the comprehensiveness of the description. After the clips are matched for similarity, some video clips with high similarity will be extracted. However, how to measure the similarity of a video clip is not only based on some basic feature matches, but also on the relationship between the video levels of these matches. For example, chronological order, granularity size and interference factors.

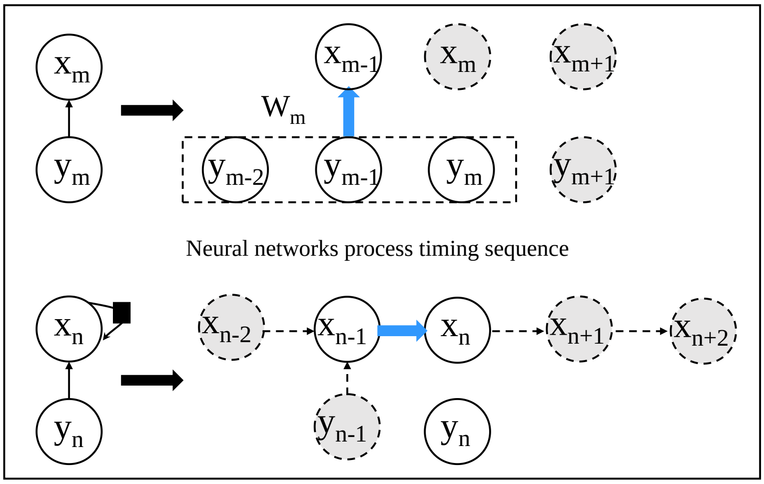


**Figure 3:** Video clip similarity matching.

The feature factor represents the similarity between the query fragment and the similar fragment obtained by feature extraction, feature description and other operations, which occupies a larger weight in the similarity result of two video clips. If the feature value of both is larger, it means that the similarity between two video clips is higher. The video clip is composed of a number of shots. There are more similar frames within shots and little difference between shots, which leads to high self-similarity of the video clips. When the query clip is compared to it, multiple patterns emerge. As shown in Figure 3, since the number 5 in the similar clip p3V is similar to the numbers 6 and 7 in the query clip, resulting in a many-to-one situation, the similarity of similar clip p1V is significantly better than that of similar clip p3V, and the granularity value of similar clip p1V is greater. In practical applications, the above points are comprehensively weighed according to the content and structural characteristics of the video and weight values are assigned to them, based on which the video clips with higher similarity to the query clips can be found.

ORB, an acronym for Oriented FAST and Rotated BRIEF (o FAST and r BRIEF), is a fast and efficient algorithm for feature point extraction and description. The algorithm was proposed in 2011 by Ethan Rublee et al. ORB features are divided into two parts, one is the key point and the other is the descriptor. The extraction of ORB features is generally divided into two steps: o FAST feature point detection: finding the feature points in the image, which differs from FAST feature point detection in that ORB improves it by adding the main direction of the feature point, giving it rotation invariance. r BRIEF feature descriptor, which is a binary string of points in the region of the detected feature point, is a ORB feature extraction is an improvement on FAST feature extraction, which is fast feature detection with the idea that a pixel is more likely to be a feature if there is a large difference between that pixel and the surrounding pixels. radius to get 16 pixels on the same circle. A pixel P is treated as a feature point if the 16 pixels on the circle have a value of N consecutive pixels that is greater than the brightness of p plus the threshold X, or less than the brightness of p minus the threshold X. The above operation is performed iteratively to detect feature points for each pixel in the image as shown in Figure 4.

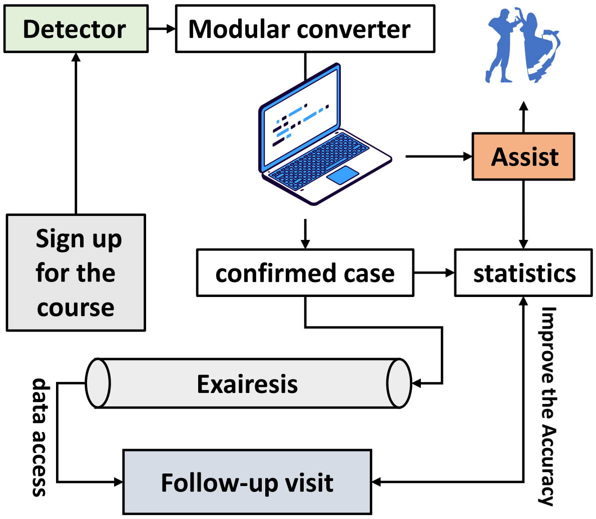


**Figure 4:** Pixel dot feature detection.

After extracting keyframes from a dance video, a sequence of image frames that summarize the video content is obtained. The dance video retrieval library consists of a number of keyframes and the video clips they summarize. It is formed by first extracting keyframes from dance videos, and then segmenting each dance video by an audio-based segmentation processing method. In the dance video clip retrieval, the query clip to be retrieved is matched with the video clip in the video library, and it can be seen from the previous video clip retrieval methods that the retrieval time can be greatly shortened by extracting keyframes for the video and comparing them with the keyframes in the footage. In this paper, improvements are made from two aspects: on the one hand, at the keyframe extraction stage, after analyzing the two datasets used in this paper, the keyframe extraction work is carried out by using the movement of each dance movement in the dance video, which improves the accuracy of keyframe extraction and thus the accuracy of dance video retrieval. On the other hand, multiple factors are used in the retrieval, which makes the similar videos obtained have high quality characteristics.

First, keyframes are extracted for the dance video. The dance video is composed of a series of dance movements, with more or less movement embodied between the coherent dance movements, and the movement information in the dance video is represented by the optical flow, and then entropy is used to count the information in each optical flow graph. The entropy sequence is fused with the musical features to obtain a music-related entropy sequence. Subsequently, the keyframes are selected by thresholding, and the best threshold for the video is selected by comparing it with the set of keyframes selected by multiple users when the threshold is set. This results in a set of keyframes of a dance video for subsequent video clip retrieval. After that, in the dance video retrieval phase, ORB features are used to extract feature points between the query clip and the video keyframes in the dance video database by comparing every (3, 5 or 7) frames of the query clip with all the keyframes to find matching feature points. The maximum number of feature points corresponding to the image frame in the query clip is found in the result for the video clip in which the keyframes are located. This step is based on the video abstract video clip retrieval method, with the idea that the keyframes can summarize the video content, and the comparison retrieval is performed in the form of video abstract. Finally, the similarity model is constructed through the consideration of multiple factors. The similarity model is constructed based on feature factors, temporal order factors and interference factors. Each factor is given a different weight depending on the video. Experimentally, a pair of weights suitable for the video is selected. The most similar video clip is obtained in this way as shown in Figure 5.

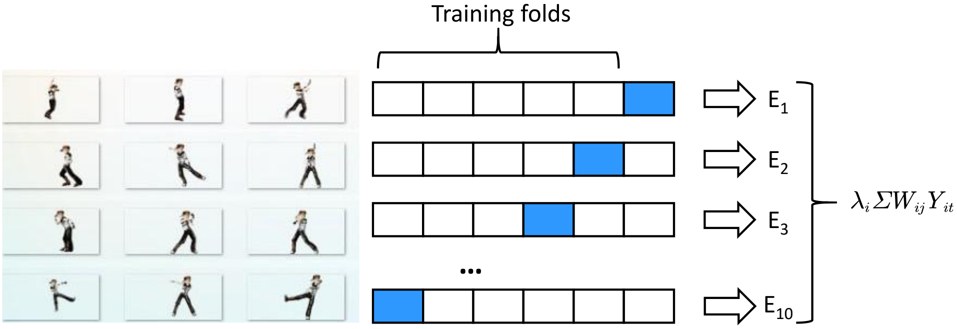


**Figure 5:** Video Motion Weighting Analysis.

### 3.2 Design of Dance Data Sets based on Device Recording

Prior to recording the dance video, the following preparations are made: first, the motion acquisition area is framed with a reflector ball, then the data is pre-captured by holding a calibration stand and swinging it around the acquisition area, and then the stand is placed in the area to set the origin. The height, weight, and width of each bone are measured for each performer, and the information is recorded in the software application. Lastly, the person's joints were placed at the point of origin. In the dance video recording phase, five dancers performed the eight dances in Table 5.1 and recorded the videos for subsequent processing. In the video recording, the five dancers are represented by the capital letters A, B, C, D, and E. The eight dance movements are distinguished by the following labels. For example, A2, C7, E4 and so on. The dancers perform in the pre-framed collection area. For different dance combinations, they are divided into "total" and "score". For the different dance combinations, they are divided into "total" and "sub", which are always one combination, and the performer has to perform all the dance moves in each combination. For different dance datasets, the first step is to perform keyframe extraction, and the keyframes of the dances can be used to represent the content of the video. According to the results of keyframe extraction, in the subsequent dance video retrieval, the most similar dance videos are obtained by matching the keyframes and video frames for the corresponding retrieval operation. The list of dance video recordings is shown in Figure 6.

A clip search of dance videos reveals that the dance movements in dance videos are closely related to music. And in different categories of dance videos, there are more or less video clips, which are similar to a video under different categories, but they express different themes or emotions. In other words, the same dance moves in different combinations will produce a new dance video. In the experimental process, the weights set in the similarity matching model will have some influence on the experimental accuracy. The three factors related to accuracy, feature factor, temporal order factor and interference factor, all have different weights. However, the main factor that determines the similarity between two clips is the feature factor, which is the value of the feature obtained from matching the information of the image. Therefore, the feature factor should have a higher weight in comparison.

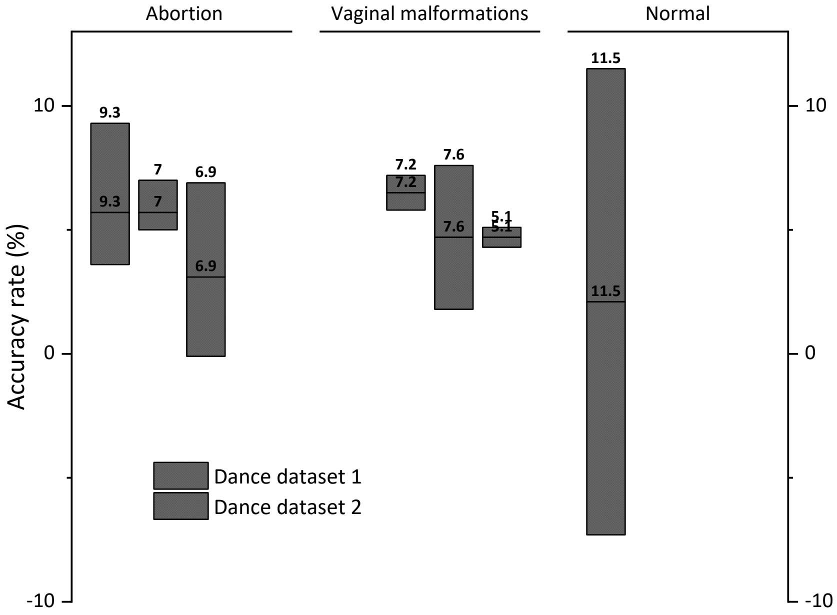


**Figure 6:** Accuracy of video clips in the two dance datasets under different weights.

By setting the parameters in the experiment, the corresponding comparison experiments are performed in two dance datasets. In this paper, dance video retrieval is performed by extracting keyframes from music and dance videos. In video retrieval, keyframes can be used to represent the main content of a video, which reduces the computational effort. The keyframe extraction method for dance videos according to Chapter 3 is compared with the method in the literature, and in the experiments, since the dance movements of dance videos mirror their accompanying music, the corresponding entropy values change when the accompanying music keeps changing over time, so that the heavy beats can be effectively distinguished from the beats with more movement. However, the foregoing focuses on selecting candidate keyframes by clustering them and then ranking them to select more representative keyframes. The method selects a less redundant set of keyframes, but the selected keyframes are not accurate enough and differ significantly from the user's assessment, mainly because the method ignores audio features. In the experiment, the thresholds are selected according to the evaluation criteria through iterations, and a set of keyframes with high evaluation coefficients is selected.

The wide range and speed of video distribution makes the amount of video data available on the web huge. When performing content-based video retrieval, the diversity of videos and their repetitiveness become a problem in the field of video retrieval. In recent years, research in the retrieval field has been continuously favored by scholars at home and abroad. In this paper, the research related to video retrieval is aimed at dance videos with varied morphologies. However, there are few publicly available dance datasets, and this paper uses the Dance Motion Capture

Database dance dataset and the dance dataset recorded by professional dancers performing dances through the Vicon motion capture system. The main focus of this paper is on video clip retrieval based on keyframe extraction from music and dance videos. The keyframes are extracted from the music and dance videos, and a feature sequence with audio variation features is obtained by fusing the musical features with the movement features of the dance movements by frame rate, and thresholds are selected by measuring the values of the evaluation system. The experimental results show that the method proposed in this paper is effective in extracting the set of keyframes that summarize the video content as shown in Figure 7.

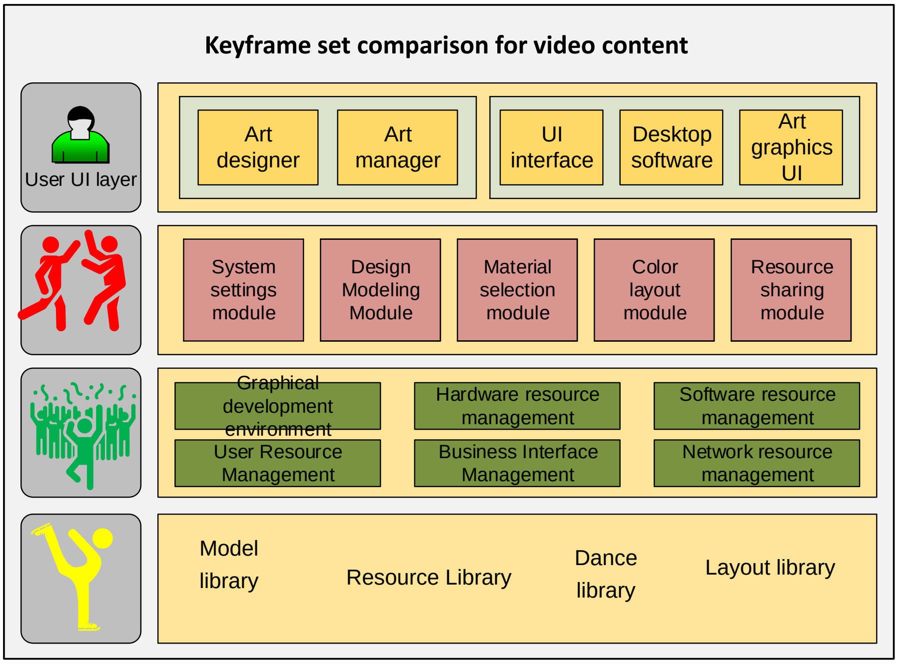


**Figure 7:** Keyframe set comparison for video content.

## 4 CONCLUSION

With the help of new computer function processing technology, the time required to design dance movements has gradually been reduced from the previous ten to twenty-four-hour work cycle to the current five-to-ten minute time fluctuation range. The degree of damage to the dancers' own skeletal development and muscle molecular structure caused by various dance movements has been further enhanced from the original ten-digit fuzzy estimation to a more accurate measurement range of two decimal places. Furthermore, with the continuous evolution of dance types and forms and the proliferation of dance videos, the main problem now is how to navigate dance videos quickly and efficiently. The problem of complex and repetitive movements in music and dance videos poses a problem for the analysis and recognition of dance movements. Therefore, in order to increase the rate of dance video analysis and reduce the computational complexity, extracting keyframes of music and dance videos will be an effective solution. Keyframes are defined as some image frames in the video that have little redundancy and are representative of the video. For dance videos, keyframe extraction is the process of removing redundancy, and at a certain moment when the music is expressive and the difference between the current frame and the previous keyframe is large, the image frame is identified as a new keyframe. Through the keyframe collection, the user can get a quick understanding of the content of the dance video, which shortens the time the user spends browsing the dance video and saves the storage space of the video.

*Xiaolin Wang*, https://orcid.org/0000-0002-5543-3300

**REFERENCES**

[1] Bothma, T.-J.; Prinsloo, D.-J.; Heid, U.: A taxonomy of user guidance devices for e-lexicography, Lexicographica, 33(1), 2018, 391-422. https://doi.org/10.1515/lexi-2017-0018

[2] Coley, C.-W.; Barzilay, R.; Jaakkola, T.-S.: Prediction of Organic Reaction Outcomes Using Machine Learning, ACS Central Science, 3(5), 2017, 434-443. https://doi.org/10.1021/ACSCENTSCI.7B00064

[3] Hoffman, C.; Johnson, T.; Junge, M.: Recurrence and Transience for the Frog Model on Trees, Annals of Probability, 45(5), 2017, 2826-2854. https://doi.org/10.1214/16-AOP1125.

[4] Nooh, A.; Aoude, A.; Fortin, M.: Use of Computer Assistance in Lumbar Fusion Surgery: Analysis of 15 222 Patients in the Acs-Nsqip Database, Global Spine Journal, 7(7), 2017, 617-623. https://doi.org/10.1177/2192568217699193

[5] Caro-Codón, J.; Rey, J.-R.; Cruz, J.-D.: Precision and Reproducibility of Non-Automatic Measurement of the Qrs Complex in Potential Candidates for Cardiac Resynchronization Therapy, Journal of Electrocardiology, 57, 2019, 90-94. https://doi.org/10.1016/J.JELECTROCARD.2019.08.011

[6] Kohli, M.; Alkasab, T.; Wang, K.: Bending the Artificial Intelligence Curve for Radiology: Informatics Tools from Acr and Rsna, Journal of the American College of Radiology, 16(10), 2019, 1464-1470. https://doi.org/10.1016/J.JACR.2019.06.009

[7] Burns, D.-C.; Mazzola, E.-P.; Reynolds, W.-F.: The Role of Computer-Assisted Structure Elucidation (Case) Programs in the Structure Elucidation of Complex Natural Products, Natural Product Reports, 36(6), 2019, 919-933. https://doi.org/10.1039/C9NP00007K.

[8] Júnior, F.-V.-C.: The Optimal Multilinear Bohnenblust–Hille Constants: A Computational Solution for the Real Case, Numerical Functional Analysis and Optimization, 39(15), 2018, 1656-1668. https://doi.org/10.1080/01630563.2018.1489414

[9] Hahn, E.-A.; Downing, N. -R.; Stout, J. -C.; Miner, J. -A.; Carlozzi, N. -E.: Understanding the Need for Assistance with Survey Completion in People with Huntington Disease, Quality of Life Research, 27(3), 2018, 801-810. https://doi.org/10.1007/S11136-017-1747-6

[10] Exner, P.; Lotoreichik, V.; Pérez-Obiol, A.: On the Bound States of Magnetic Laplacians on Wedges, Reports on Mathematical Physics, 82(2), 2018, 161-185. https://doi.org/10.1016/S0034-4877(18)30084-3

[11] Borwein, J.-M.; Lindstrom, S.-B.; Sims, B.: Dynamics of the Douglas-Rachford Method for Ellipses and P -Spheres, Set Valued and Variational Analysis, 26(2), 2018, 385-403. https://doi.org/10.1007/S11228-017-0457-0