Analysis of Sanda action based on CCNN model of IBS distance measurement

Shan Li
Shandong Management University, Shandong Province, China

Corresponding author and e-mail: Shan Li, 14438120180132@sdmu.edu.cn

Abstract. IBS is a similarity measurement method of symbol sequences based on order statistics, and CCNN is a method based on CNN feature extraction and IBS distance measurement, which can construct crossover probability graph through image crossover comparison. IBS distance and Noribs distance are used to define the similarity between crossover probability graphs. For Sanda movement, it is difficult to carry out big data analysis with a small number of samples and the movement style of athletes will change greatly. In this paper, a CCNN model based on IBS distance measurement is proposed to analyze the action of Sanda. Firstly, the motion pictures obtained were de-noised and then simplified into a sequence diagram of human skeleton nodes by Delaunay triangulation. Then, 80% of the samples were input into the CCNN network for training and 20% for algorithm verification. The experiment found that in the case of a small number of samples, the accuracy of the simplified image analysis was 10% higher than that of the unsimplified image. In addition, compared with the accuracy of K-distance classification, Bayesian network classification and decision tree classification, it is found that the model can improve the accuracy by 2%-11.75% due to the small number of samples, and has higher accuracy and applicability.

1. Introduction
Sanda, also called sanshou, is an unarmed combat fighter project that two people use the traditional martial arts of kicking, hitting, tumbling and other offensive and defensive techniques to subdue the other side according to certain rules. It is an important form of competition. The Chinese martial arts can be divided into the ancient sanshou and the modern sanda. Its highly abstract, attacking skills roughly divided into two kinds of movement forms: One is a linear method. The other is a type curve method. The traditional motion analysis is based on manual analysis. With the continuous maturity of image processing technology and the continuous improvement of information level, the establishment of moving image information management system has become a research hotspot. Image processing informationization has a broad development prospect [1-2] because of its high precision, high energy, accurate processing, image recognition, large volume data processing and strong flexibility, which can combine the characteristics of image information and freely choose the freedom and clarity of image [2]. When analyzing the motion image, images must be taken first. However, due to the light, weather, image pixel and other reasons, the collected image will have noise. In order to reduce the noise, increase the clarity of the picture and the accuracy of image processing, Tian Bin adopts the corresponding wavelet coefficients by threshold to remove noise, which can effectively restrain movements in the image noise and improve the image SNR movements. For blurry or dim images,
Retinex can enhance the image to enhance the robustness of the image's light change. By doing so, Retinex can improve the contrast of the moving image. Through experimental comparison, it is found that the average misclassification rate of the motion classification method based on image processing technology is 5%. And the average rejection rate is 3%. The average misscore rate of traditional motion classification methods is about 10%. And the average misscore rate is about 6%. Therefore, the classification error rate and rejection rate of the motion classification method based on image processing technology are far lower than the comparison algorithm, which can effectively reduce the motion classification error, overcome the defect of the current motion classification speed is slow and improve the motion classification efficiency [3]. In the process of image processing, it is necessary to carry out feature processing on the acquired image. The traditional feature extraction is mainly based on manual extraction, that is, it is necessary to set the target feature which is beneficial to classification to carry out feature extraction. At present, the commonly used target features are divided into visual features, transformation coefficient features, statistical features and algebraic features, etc. There are two defects in manual extraction. The first is that the algorithm is cumbersome and inefficient. The second is that the artificial features to be extracted are easily disturbed by the complexity of the target. In order to improve these two defects and improve the efficiency of image processing, the current improved algorithms mainly include the self-learning feature extraction which uses the server loaded with the algorithm to train the self-extracting features from a large number of samples. Through various region selection algorithms, the global feature extraction and other feature extraction algorithms based on the K-Fan model of the target image are set. For image classification, the current mainstream methods include k-distance classifier, Naive Bayes network classifier, Extended Naive Bayes network classifier, decision tree classifier, SVM classifier, Rocchio classifier, support vector machine, artificial neural network classifier and so on. At present, image feature extraction and classification algorithms are mainly improved in the direction of simplicity, recognition reliability, efficiency and effectiveness [4]. Sports action pictures contain many elements, which need to be simplified. In the process of image simplification, Guo Jianping, Li Xi et al, established a new image convolution model for the automatic intelligent analysis of karate athletes' technical action recognition, movement frequency statistics and track tracking. The action pictures with many elements were simplified into Delaunay triangulation maps based on skeleton node composition and local optimization processing, which improved the identification of behavior patterns. Through comparative experiments, it is found that the topology information capacity constructed by this method is larger and more conducive to the accuracy of behavior recognition effect. And the behavior recognition algorithm improves the intelligence and precision of technical and tactical analysis. But the lack of moving sequence information in the link of trajectory analysis reduces the practicability of trajectory analysis [5]. However, the image processing at present is mostly based on the analysis and comparison of a large number of data. Sanda action analysis is a project with strong timeliness. The movement style of the same player in different periods will change. And the collection of Sanda action pictures is small. So it is difficult to carry out a large number of comparison and classification. In the research to solve similar problems, Xu Ziyi from Nanjing University adopted CCNN to classify the artistic styles of artists' works and achieved good results. CCNN can process and classify pictures with a small number of samples, but no one has applied CCNN to the research on the analysis of free fighting movements. In this paper, the CCNN model based on IBS distance measurement is adopted to identify and classify the Sanda movements according to the characteristics of less samples of Sanda action pictures. Selecting four famous sanda athletes and its motion picture will collect pictures of 80% for neural network training, 20% for algorithm validation. After preprocessing the training data, input the pre trained convolution neural network for feature extraction, measure the similarity with the existing data based on IBS idea, and finally output the similarity output result of Sanda action.
2. Principles and algorithms
The human skeleton node sequence diagram model was obtained by obtaining nodes from images and generating Delaunay triangulation maps. In the process of Delaunay triangulation, each node needs to be locally optimized. The steps of local optimization are as follow. The first step is that two adjacent triangles form a quadrilateral. The second step is that make a circle around the two triangles and check whether the other points are inside the circle. Figure 1 shows the generation process of human skeleton node sequence diagram model.

![Model generation process diagram of human skeleton node sequence diagram.](image)

Based on IBS idea, Noribs distance is defined firstly.

\[
NorIBS = \frac{\sum |R_i - R_j| SE_i}{D \sum SE_i}
\]

(1)

\[
SE_j = E_i(P) + E_j(P)
\]

(2)

\[
E(P) = -p \log p
\]

(3)

\[
Re \text{vIBS} = NorIBS(random(p_i))
\]

(4)

Where \(R\) is the probability and order of the ith filter of the x image, and \(SE\) is the sum of entropy of the probability distribution of the two images. When two images are similar, Noribs tends to 0, and the greater the distance, the smaller the similarity. In this paper, a distance ModIBS is defined again, where \(\beta\) is a coefficient, generally 2-4. The value of ModIBS is 0-1. When this value tends to 0, it means that the two pictures are similar, and when this value tends to 1, it means that the styles are different.

\[
ModIBS = e^{-(Re \text{vIBS} - NorIBS)\beta}
\]

(5)

CCNN(Cross Contrast Neural Network) model is a method based on CNN feature extraction and IBS distance measurement, which aims to measure the similarity of different images and realize multiple image classification on this basis. The input of CCNN model is a group of images of a specific size, and a deep convolutional neural network is used to extract the features of this group of images, and a group of crossover probability graphs is constructed by cross-comparing the images. On the basis of IBS, a Modibs distance is defined to characterize the similarity between vectors of the crossover probability graphs. The CCNN model consists of two parts. The first part is feature extraction. In this part, different features are extracted by using the implicit information generated by comparison between input images, and a set of cross probability graphs of feature vectors are obtained. The second part is the distance measurement calculation based on crossover probability graph, which
aims to measure the similarity between the images. These two parts are not independent, but connected together to form a complete network, which can be used for gradient propagation and training. The overall structure of CCNN is shown in Figure 2.

After feature extraction by CCNN, cross-comparison and similarity measurement are required for the extracted features. The feature vector of any image is expressed as.

\[ P_n = [p_1, p_2, \ldots, p_i, \ldots, p_n], n \in [1, N] \]

\[ p_i = \sum \sum \text{softsign}(\max(v_{ij}, 0)) / \sum \sum \text{softsign}(\max(v_{ij}, 0)), i \in [1, D], j \in [1, k] \]

Vij represents the JTH pixel value of the ith convolution filter of the image; N represents the number of input images. And the dimension of the last convolution layer is D; K is the size of the output of the last convolution layer. Combine images in random pairs, setting the object label to 0 if both images belong to the same category.

3. Experimental design
Firstly, the 400 pictures of the four Sanda players were de-noised and then the pictures were processed into the sequence map of human skeleton. The images obtained after processing were grouped. And 80 of them were selected as algorithm verification. And 320 were used for convolutional network training. After processing and grouping, 80 images were classified by CCNN, K-distance, Bayesian network and decision tree, respectively. Finally, the accuracy of each method was compared. The process is shown in Figure 3 and Figure 4.
4. Results and discussion
Check data, sanda attack the main way for the elbow and kick, will be collected by the four athletes (A - D) 400 sanda movement images of the same players in different periods and different grouping, and then respectively according to the characters of the elbow and legs similarity matching, to select one of the two people to display the results as shown in figure 5.
Figure 5. Comparison of movement similarity of the same athlete in different periods.

As can be seen from Figure 5, the movement styles of the same player are different in different periods, and some players even have great changes. However, the changes are few in number, and the changes will last for a long time after the change of tactics, and the fluctuation range is limited. The results of similarity comparison between different players are shown in Figure 6.

As can be seen from Figure 6, there is no obvious rule of movement correlation between different players, and there is a phenomenon of great difference in movement style. By comparing Figure 5 and Figure 6, it can be seen that:

- The movement style of the same Sanda player may change significantly in different periods.
- There will be great differences in movement styles between different players.

At the same time, due to the small number of Sanda action pictures, it is difficult to carry out big data analysis. 50 pieces of the same data were selected for CCNN classification, K-distance classification, Bayesian network classification and decision tree classification respectively. The sample size was gradually increased to 400 and the gradient was increased to 50. Comparison of accuracy of each classification method is shown in Figure 7.
As can be seen from Figure 7, when the number of samples is large, the accuracy gap between the methods is small due to the support of big data. However, when the number of samples is small, CCNN has its unique advantages and its accuracy is 2%-11.75% higher than other methods. It is difficult to make big data analysis on the pictures of Sanda movements due to the difficulty in obtaining materials, and the same athlete's movement style may change greatly in different periods, which is nonlinear to a certain extent. Therefore, compared with the traditional classification algorithm supported by big data, the CCNN network based on IBS has the characteristics of accurate classification in a small number of samples, and has great advantages in the field of scattered image classification.

The accuracy of the original image without Delaunay triangulation simplification was compared with that after simplification, and the comparison results were shown in Figure 8.

As can be seen from Figure 8, the analysis accuracy of the simplified Delaunay triangulation image in the system is up to 10% higher than that of the original image. Under the condition of a large number of samples supporting the original image, the system accuracy does not improve much after
the simplification of the original image. However, when the number of samples is relatively small, the simplification of the original image can effectively improve the image analysis accuracy. Due to the small number of samples in the pictures of Sanda and other sports, it is difficult to carry out big data analysis, and the movement style of athletes will change greatly. In the analysis of the movement of Sanja and other sports, the extracted action pictures can be first simplified into the sequence diagram of human skeleton nodes by Delaunay triangulation, and then input into the trained CCNN network. It can achieve a relatively accurate analysis of images in the case of a small number of samples.

5. Conclusions
In this paper, CCNN based on IBS distance measurement is proposed to analyze the action pictures of Sanda according to the characteristics of less samples. Firstly, the obtained original image was mute and then simplified into a sequence diagram of human skeleton nodes by Delaunay triangulation, and then input into the trained CCNN network. It was found that under the condition of the small sample size after simplified image analysis accuracy than not simplified picture accuracy is increased by 10%, and K-distance classification, bayesian network classification and the accuracy of the decision tree classification comparison found that method supported by a large number of samples in the accuracy of difference is not big, but the sample size is small, CCNN high accuracy 2% 11.75% than the other three methods, shows that this method analysis of sanda, and other sports pictures has the advantage.

References
[1] Tian Qinqin, Chen Lang, Zhang Xiangqing. Intelligent processing method for image information recognition [J]. China New Communications, 2020, 22(05):128.
[2] LIANG Feng, ZHANG Zhili, LI Xiangyang, et al. Human Lower Limb Motion Recognition in Human Motion Capture Process [J]. Journal of Computer-Aided Design & Computer Graphics, 2015, 27 (12): 2419-2427.
[3] TIAN Bin. Research on Motion Classification Based on Image Processing Technology [J]. Modern Electronic Technology, 2021, 44(03):49-53
[4] Wang Weinan, Yang Chaohong. Review of target recognition methods based on image processing technology [J]. Computer and Information Technology, 2019, 27(06):9-15.
[5] Wang Weinan, Yang Chaohong. Review of target recognition methods based on image processing technology [J]. Computer and Information Technology, 2019, 27(06):9-15.