Research article

Adaptive enhancement design of non-significant regions of a Wushu action 3D image based on the symmetric difference algorithm

Chao Zhao[1],*, Bing Li[2] and KaiYuan Guo[3]

[1] Wuhan Sport University, Wuhan 430079, China
[2] College of Physical Education, Chosun University, Gwangju 61452, South Korea
[3] Physical Education Institute, Yong In University, Yongin-si 17092, South Korea

* **Correspondence:** Email: zhaochao@whsu.edu.cn.

Abstract: The recognition of martial arts movements with the aid of computers has become crucial because of the vigorous promotion of martial arts education in schools in China to support the national essence and the inclusion of martial arts as a physical education test item in the secondary school examination in Shanghai. In this paper, the fundamentals of background difference algorithms are examined and a systematic analysis of the benefits and drawbacks of various background difference algorithms is presented. Background difference algorithm solutions are proposed for a number of common, challenging problems. The empty background is then automatically extracted using a symmetric disparity approach that is proposed for the initialization of background disparity in three-dimensional (3D) photos of martial arts action. It is possible to swiftly remove and manipulate the background, even in intricate martial arts action recognition scenarios. According to the experimental findings, the algorithm's optimized model significantly enhances the foreground segmentation effect of the backdrop disparity in 3D photos of martial arts action. The use of features such as texture probability is coupled to considerably enhance the shadow elimination effect for the shadow problem of background differences.

Keywords: symmetric difference algorithm; martial arts; action recognition; 3D image; adaptive enhancement

1. Introduction

Martial arts have been passed down throughout Chinese civilization's lengthy history as a special treasure of the Chinese people. Through constant evolution, martial arts have moved from their original purpose of neutralizing an enemy to developing moral character [1]. Chinese martial arts have a long history, and the public has always been interested in their transmission and spread. The Ministry of Education has incorporated martial arts into schools to support martial arts culture, spread martial arts spirit, improve students' physical condition, and support national heritage through educating the next generation of young people in martial arts. The first time martial arts were offered as a sport to middle school students was in Shanghai in 2016. Martial arts was also made a required subject in the middle school sports test in Shanghai starting in 2020.

The score for the martial arts portion of the secondary school entrance exam is calculated by averaging the results of many referees. Despite the fact that they all adhere to established grading guidelines, the subjective elements of the guidelines are too significant [2]. The referee must always maintain a high level of concentration, proper referee awareness, and outstanding physical fitness in the referee process, in addition to mastering the scoring regulations of the game skillfully. The evaluation of martial arts is excessively reliant on opinion. Therefore, it is critical to use computers as a more precise, efficient, and objective approach to aid in the evaluation of martial arts movements because there is a lack of a ruler for precise measurement [3]. The objective of the task is for the computer to automatically identify human behavior in the video to gather behavior data about the video object [4]. Separating video frames is the approach that researchers most frequently use while conducting an initial study of human motion recognition. Building a precise model of motion features allows researchers to recognize activities by first manually designing motion features that express human motions. However, manual feature acquisition requires a large amount of work and calculation. Human motion detection has advanced dramatically in recent years with the development of deep learning. Deep learning significantly lowers the cost of feature acquisition by automating the acquisition of motion features [5]. There have been positive study findings in the area of human motion detection up to this point, and the level of experimental accuracy is very high. However, there is still much work to be done before human motion recognition can be used in everyday life.

Numerous and significant technological advancements have occurred recently, such as human posture estimation technology for martial arts actions, which focuses on the task of identifying key human joints in images or videos and gathers the necessary skeleton data from joints. Based on this, computer vision-related information, particularly technology for estimating human posture, is used for the automatic comparison and analysis of videos, thereby achieving the automatic comparison and analysis of videos and completing the tasks of athlete guidance and evaluation [6]. The task of automatic comparison and analysis using video key frames encounters many challenges, just like other computer technology-related problems: for example, self-occlusion of the target, excessive amplitude of motion, change in shooting angle, and differences between moving individuals. These cause many issues in research on this task. Researchers continue to investigate and conduct research, despite the many challenges they encounter. They use a variety of techniques to apply comparative video analysis in sports teaching, event scoring, rehabilitation engineering, and other fields. They also continue to combine research with real-world projects to drive the rapid development of automatic video comparative analysis [7].

The symmetric difference algorithm-based adaptive enhancement design of non-significant sections of three-dimensional (3D) martial arts motion images is also crucial for MPEG-4 video coding. Thus, 3D martial arts action and moving object segmentation is a crucial study topic in the field of computer vision.

The contributions of this paper are as follows:

1) This paper includes a thorough study of the principle of a background difference algorithm, and a systematic summary and analysis of the advantages and disadvantages of various background difference methods. For several typical difficulties in the background difference algorithm, corresponding solutions are proposed. Then, the symmetric difference algorithm is proposed to automatically extract the empty background for the initialization problem of the background difference in 3D images of martial arts actions.

2) The optimized model of this algorithm greatly improves the foreground segmentation effect of the background difference in martial arts action 3D images. For the shadow problem of the background difference, the use of features such as texture probability is combined to greatly improve the shadow elimination effect.

3) After heat map prediction through the network, it is necessary to first complete the corresponding resolution recovery and then convert it to a coordinate representation to achieve a more realistic effect.

2. Related work

The use of martial arts action analysis video has expanded, and along with it, people's expectations of video analysis technology. In some ways, this increase in actual demand has also actively aided video analysis's advancement and development by forcing academics to gradually switch from studying low-level aspects to high-level features [8]. While this is being executed, higher semantic information in movies is mined and analyzed to complete the work of video analysis [9,10].

The primary subject of the study is the examination of baseball videos. The ability to recognize the throwing speed has been shown to have a substantial influence on event identification and video content retrieval. It may also be used to acquire movies of interest, which has a positive impact on the function of video event detection [11,12]. The main research content is the application of advanced semantics to football videos, which further restricts semantic matching. Simultaneously, the text event and video are synchronized through image processing, and video analysis of advanced semantic features is completed [13,14]. The main research content is on the visualization system. Motion analysis and trajectory analysis technology are used to analyze the events, regions, and personnel in the video, and improve the accuracy of video analysis [15,16]. The primary area of study is the automatic classification of basketball video clips, which may implement automatic player tracking and identify players based on the context of the video. The research simultaneously merges various elements to improve the video's tracking and analysis capabilities [17,18].

Setting joint martial arts movements effectively uses the forward or backward kinematics method [19]. It is possible to determine each linked limb's position by keyframing the joint rotation angle. The forward kinematics method is the overall name for this technique [20]. Researchers interested in joint martial arts frequently used it because it was the first to provide a matrix description approach to describe the position of each joint through the relative coordinate system. However, it can be exceedingly challenging for a novice martial artist to create realistic motion by

setting key frames for each joint [21]. A practical solution is to record the spatial motion data of human joints through real-time input devices, that is, a motion capture method. To overcome the lack of flexibility of this method, researchers edit the captured data by mixing martial arts action curves, which makes it possible to establish a reusable motion database [22,23]. A novel idea for motion retargeting is proposed. This technology is ideal for processing motion capture martial arts actions because it can transfer a character's martial arts movements to another character with the same joint structure but differing joint lengths while maintaining the character's original quality [24,25].

An intuitive approach that uses a hierarchical workspace for each joint segment is provided in terms of algorithm optimization for the martial arts action model, which minimizes the mobility of the joint position as much as feasible [26]. The issue with this approach is that the user has no control over the outcome. The solution found for complicated joint structures cannot be the solution that produces natural motion. It has been suggested that students should construct joints via inverse motion [27]. In their method, users specify the world coordinate system position of the foot and then use the pseudo inverse Jacobian matrix to solve the rotation angle from the foot to the hip joint. An advantage of the kinematics solution is that constraints can be set on some key positions of joints [28]. For example, when people bend their knees, they can restrain their feet on the floor while leaning down. Similarly, when people walk, they first rotate their bodies around one foot, then around two feet, and then around the other foot [29]. Inverse kinematics is frequently used to resolve joint martial arts moves that have restrictions, which is comparable with choosing one solution out of several to satisfy the constraints. The joint tree can be rebuilt if only one point is limited [30].

In martial arts action model, computer, and dynamics. Explicitly express the kinematic and dynamic constraints, and then solve these equations [31]. Unfortunately, this method is computationally expensive. A method for combining forward kinematics and inverse kinematics for joint motion editing is proposed. Martial artists can make interactive goal-based modifications to existing joint movements. The key idea of the method is to insert the required joint space motion into the inverse kinematics control mechanism. An interactive method is proposed to control the motion of biped joint animals through kinematic constraints. These constraint models can comprehend the properties of motion and regulate a person's stability and balance. In this study, a technique is suggested for fusing dance with martial arts motions. The forward kinematics and inverse kinematics martial arts action setting methods are all provided by the martial arts action software Maya, suffrage, lamias, and Bove front. Softimage's Actor module allows users to set joint martial arts movements, whereas Maya is a program for character martial arts movements.

To summarize, the adaptive enhancement design of non-significant regions of martial arts action 3D images based on the symmetric difference algorithm is an important reference for the close combination of martial arts and computer science.

3. Symmetric difference algorithm

Typically, an image sequence serves as the research subject for the segmentation of martial arts movement. In addition to being a function of spatial position, an image sequence is also a function that evolves through time. A frame is a collection of related images. A still image is a result of the spatial position, which is unaffected by changes in time. A single still image cannot adequately convey an object's motion. Generally, the image sequence can be expressed as

$$\left\{ f\left(x,y,t_0\right), f\left(x,y,t_1\right),\ldots, f\left(x,y,t_{n-1}\right)\right\} to_0 < t_1 < \ldots < t_{n-1}. \tag{1}$$

Then, the acquisition time interval of two adjacent frames is defined as

$$t_k = t_k - t_{k-1}, k = 1, 2, \ldots n - 1, \tag{2}$$

where it is generally believed that $\Delta t_k (k = 1, 2, \ldots n - 1) = \Delta t(k = 1, 2, \ldots, n - 1)$; that is, it is generally believed that the acquisition time interval of all images is equal.

The application of computer vision technology spans a wide range of industries. As a result, motion target segmentation technology is also used for many research objects in various martial arts movement analysis fields, and must meet various application objectives and use environment criteria. The objects of concern are moving cars and people on foot. Its signal source is a TV video signal (or carrier frequency) signal captured by a conventional standard (PAL or N system) color or black and white TV signal camera, or a computer common compressed format digital signal that is compressed, transmitted, and saved with this signal as the source. It should be noted that because of the technical characteristics of interlaced scanning, one frame of an image is acquired in two interlaced fields; hence, a "sawtooth" is generated at the edge of the moving object, which affects target segmentation; however, this problem does not exist in an N system signal.

Figure 1. Three low space filters for 3D images of martial arts actions.

Figure 2. Laplace edge enhancement of a blurred 3D image of martial arts actions.

Image blurring is the opposing notion to image enhancement in the adaptive enhancement design of non-significant parts of martial arts action 3D images, and beginners may also believe that this processing is not very effective. In fact, it can be a very helpful approach in some circumstances. For instance, high-frequency noise can be reduced to make the image more aesthetically acceptable by the application of a particular amount of blurring. When the image's uninteresting backdrop is blurred, the unblurred portion of the image is highlighted; that is, this processing is a helpful processing technique. The convolution kernel technique can also be used to blur the image. Figure 1 shows that the convolution kernel is 5 × 5 and each convolution coefficient is l. In fact, the fuzzy

convolution operation averages all pixel values in the neighborhood. The Laplace edge enhancement of the blurred 3D image of martial arts actions is shown in Figure 2.

Random noise in the image is efficiently removed in the 3D image model of martial arts motion. This is because, after the pixels are sorted, the pixels in the neighborhood that have random, abrupt changes in brightness values are either ranked at the front of the queue or at the bottom. The Euclidean distance, Makowski distance, and Chebyshev distance are three widely used similarity measuring algorithms based on the distance function between samples. The precise computation is

$$d(x,y) = \sqrt{\left(x_0 - y_0\right)^2 + \left(x_1 - y_1\right)^2 + \cdots + \left(x_n - y_n\right)^2} = \sqrt{\sum_{i=0}^{n}\left(x_i - y_i\right)^2}. \tag{3}$$

To complete the position detection of human joint locations in unrestricted photos or videos, the estimation of martial arts movements and postures is primarily used. The heat map positioning method is typically used in current models for research. This method creates a heat map for each joint point and uses the probability value of the joint point position as the heat map's response value. The most likely joint point coordinates are then represented by the size of the response value. Figure 3 depicts the common martial arts movements and entire attitude estimation process. Any large resolution boundary box image should first be reduced to the requisite small resolution size before being sent to the attitude estimate model to complete the prediction of the heat map. To obtain the appropriate joint point position coordinates in the original image, it is necessary to restore the anticipated heat map to a specific resolution, which allows it to be converted to the original coordinate system. The maximal active position is known as the anticipated position.

Figure 3. Common martial arts movements.

In this section, a number of mild treatments are applied to the model structure based on the HRNet model of the martial arts body posture estimate. In parallel, the Small HRNet model is proposed, and built using DARK's data encoding and decoding technology, which can essentially maintain the same detection accuracy. The lightweight model proposed in this study was tested on two open datasets, COCO and MPII, to confirm its applicability and accuracy. The model's parameters and size reduced considerably.

The Small HRNet model's network structure primarily consists of three stages, that is, stages 1, 2,

and 3, and consists of parallel connected subnets. Each subnet is resolved from top to bottom. Stage 1 consists of a subnet with the highest resolution, which is made up of a bottleneck module (Figure 4 shows a schematic diagram of the structure of a bottleneck module); Stage 2 consists of two parallel subnets, each of which consists of a Convblock module; and Stage 3 consists of three subnets, each of which consists of a Smallblock module. The three steps mentioned above are connected to each other and combined, which gives the overall network a parallel structure of three subnets. This structure enables the network to enhance multi-scale information fusion while maintaining high resolution. Figure 5 shows a schematic diagram of the fusion method used between various resolution features.
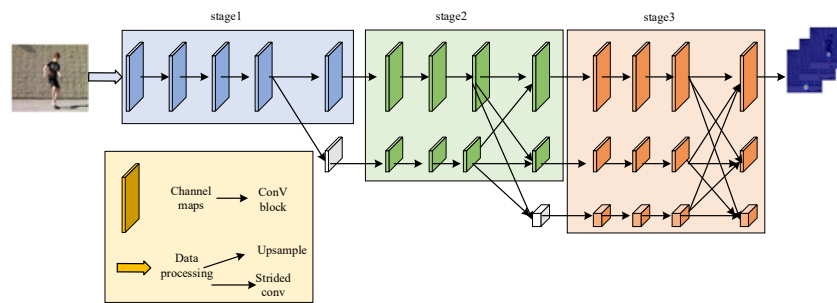

Figure 4. Schematic diagram of the bottleneck module structure of 3D images of martial arts actions.
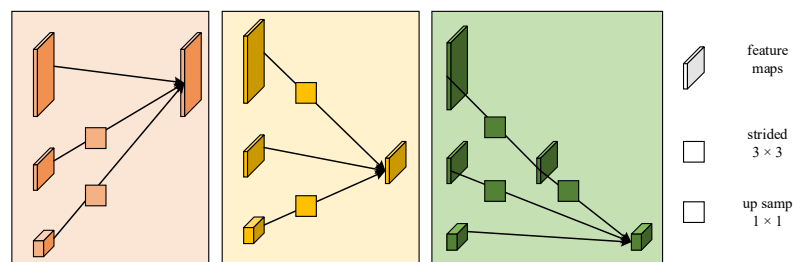

Figure 5. Schematic diagram of the fusion method used between various resolution features.

Deepwise convolution and pointwise convolution are the two main components of the depth separable convolution process used in martial arts action recognition. Both processes have the same effect of completing feature acquisition; however, depth separable convolution is more efficient than conventional convolution because it uses fewer computing resources and has fewer parameters. Figure 6 shows an action convolution diagram. If the input is a 5×5-pixel three-channel color image, and simultaneously, if the number of output channels is 4 and the dimension of the convolution kernel is 3×3, then the convolution kernel size is $3 \times 3 \times 3 \times 4$. After the convolution layer is passed, the result is four feature maps. Currently, the pixel size of the result also depends on whether there is padding during convolution. If there is padding, the output size is the same as the input size, that is, 5×5. If there is no padding, the output size is three smaller than the input size $\times 3$. At this time, the convolution layer has four filters, each filter has three convolution kernels, and each convolution kernel has a size of 3×3. Therefore, the parameter quantity of the conventional convolution layer can be calculated as $4 \times 3 \times 3 \times 3 = 108$.
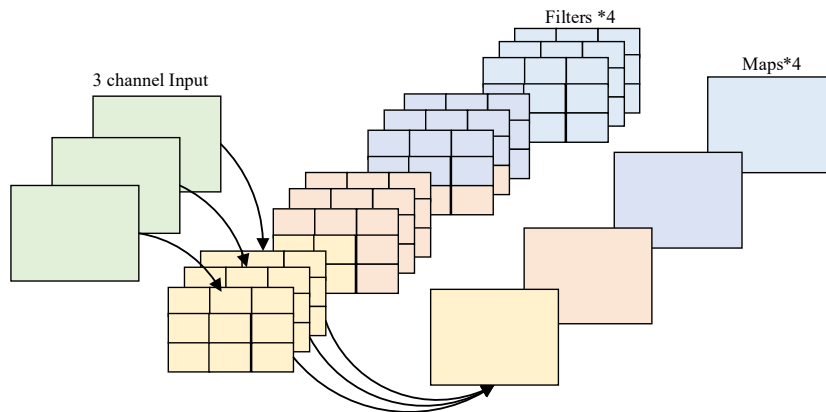
Figure 6. 3D diagram of martial arts action convolution.

Clearly, there is no requirement that these photographs be placed next to each other. In accordance with particular specifications, researchers can choose which photographs to analyze at specific intervals, for example, choosing one pair of images for three successive frames. Figure 7 illustrates this algorithm's optimization concept.
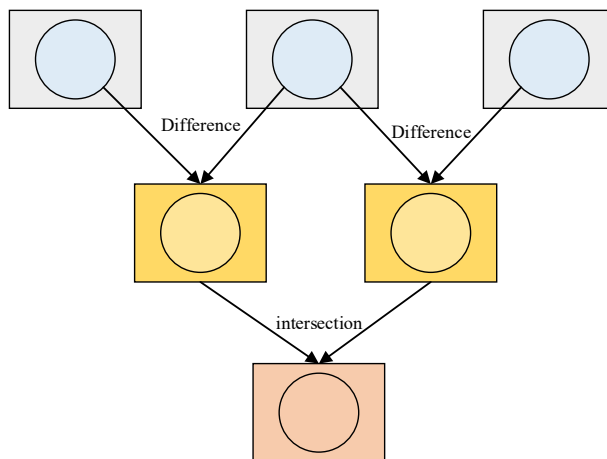

Figure 7. Optimization concept of the symmetric difference algorithm for martial arts movements.

In the model, if the image has L gray levels, the inter class variance is

$$\sigma_B^2 = \omega_0 \left(g_0 - k \right)^2 + \omega_1 \left(g_1 - k \right)^2, \tag{4}$$

where

$$\omega_0 = \sum_{i=1}^{k} p_i, \quad \omega_1 = \sum_{i=k+1}^{L} p_i = 1 - \omega_0 a_j(x,y) = \begin{cases} 1, & \left| f_j(x,y) - f_{j-1}(x,y) \right| > \xi \\ 0, & \left| f_j(x,y) - f_{j-1}(x,y) \right| \leq \xi \end{cases}, \quad j = 1,2,\ldots N. \tag{5}$$

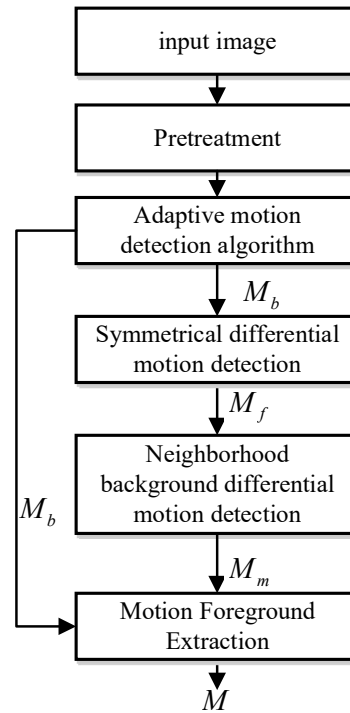The final algorithm flow chart of the martial arts 3D action recognition model is shown in Figure 8.

Figure 8. Wushu action recognition.

4 Methods

4.1. Selection of the research objects

In this study, 11 top athletes (first-class and above) from the Wuhan Institute of Physical Education's martial arts team who effectively used the whip leg method were the subjects. Prior to the test, the researchers determined that none of the participants had engaged in rigorous training during the previous 24 hours, they had no sports-related injuries within the previous three months, and their physical condition and athletic prowess were both normal, as indicated in Table 1.

Table 1. Selection of the research objects for Wushu action recognition.

Number of people (n)	Age (yr)	Height (cm)	Body weight (kg)	Years of practicing martial arts (yr)
11	19.8 ± 2.9	175.3 ± 4.4	66.7 ± 9.4	4.8 ± 2.9

After the 3D image recognition algorithm processes the martial arts activity, the subjective examination of the image quality has certain advantages over objective analysis in that it allows for the selection of an important target or locally relevant area for observation and evaluation. However, this approach is not predicated on a certain mathematical model. There are scoring discrepancies between different observers, even when they evaluate the quality of the same image, if there is no precise quantitative measurement of image quality. This is illustrated in Table 2 by the fact that observers' subjective thinking, outside interference, and other factors affect their evaluations of image quality.

Table 2. Fixed evaluation criteria for the 3D image recognition of martial arts actions.

Grade	Fixed evaluation scale	Relative evaluation scale
1	Excellent	comparative images
2	good	comparative images
3	ordinary	Average level of a group of comparison images
4	Poor	Below the average level of a group of comparative images

As the two most commonly used reference datasets for estimating human posture, in this experiment, the MPII and COCO datasets were used. The MPII dataset, which includes more than 40,000 distinct gesture annotations, was compiled from YouTube and consists of 25,000 photos with annotation data. The complete dataset contains 410 human activities with activity tags. The material is extremely extensive and many human targets have missing or obscured data. The MPII dataset contains 16 bone joint sites for the human body. The COCO dataset originated from the Microsoft Corporation and has a wide range of applications. It includes image segmentation, object detection, and the detection of human bone joints. The dataset contains 200,000 images with annotation information, including 250,000 individual annotations. The human body is marked as 17 joint points, and the coordinates and visibility of each joint point are recorded.

4.2. Model building

In this study, the various infrared image detail enhancement algorithms mentioned above are analyzed. The adaptive algorithm model of martial arts action recognition is shown in Figure 9.
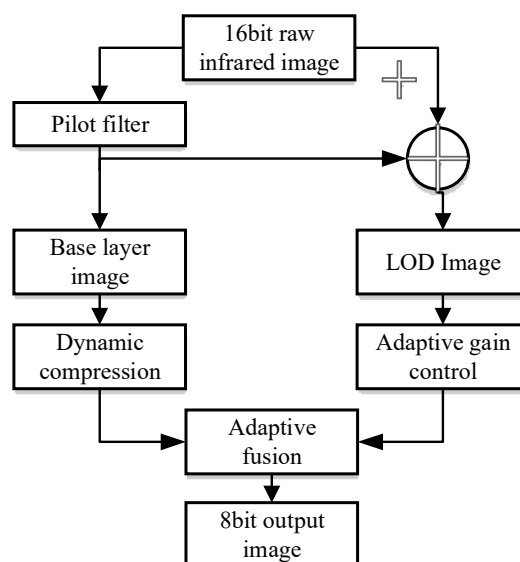
Figure 9. Adaptive algorithm model of martial arts action recognition.

Usually, the expense of martial arts movement instruction is taken into account. The image is downsampled to lower the image resolution before it is used to train the pose estimation network. This process is referred to as going from coordinate to heat map coordinate coding because it enables the network to use the heat map as a label to complete training and then use Gaussian blur to adopt the shape of the heat map. By contrast, it is important to complete matching resolution recovery before the heat map predictions made through the network are converted into coordinate representations. Coordinate decoding is the process of converting the coordinates from the heat map.

In Experiment 1, which included 22,246 samples in total, the validation set in the MPII dataset was selected as the test sample, which included 2958 samples. Information about the bone joint points marked in the MPII dataset is shown in Figure 10.

Figure 10. Collection information diagram of Wushu action data points based on the MPII dataset.

In Experiment 2, which included 149,813 samples in total, the validation set in the COCO dataset was selected as the test sample, which included 6352 samples. Information about bone joint points marked in the COCO dataset is shown in Figure 11.
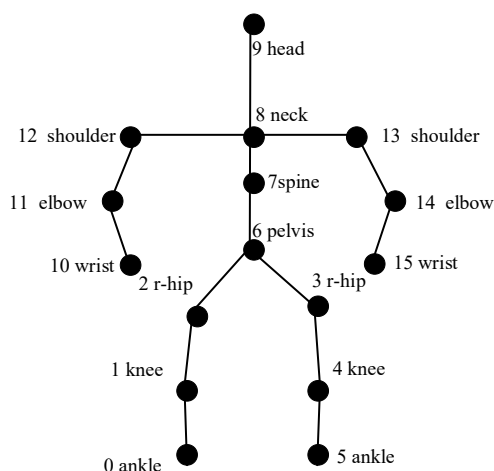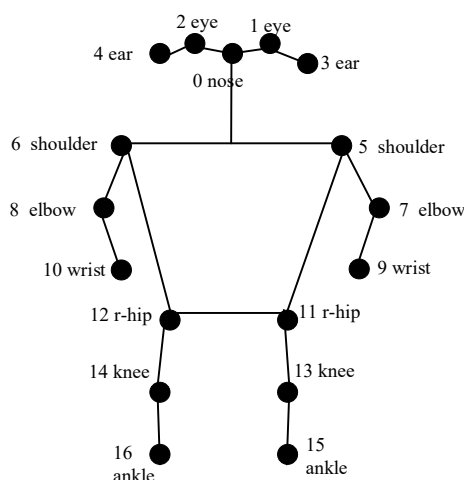

Figure 11. Collection information diagram of Wushu action data points based on the COCO dataset.

5. Case study

5.1. Effect evaluation

To determine whether the output image produced by the algorithm was better on the original basis and more suitable for human eyes to observe, in addition to determining whether the algorithm used to process the image was qualified or had flaws, the quality of the infrared image that was processed by the algorithm needed to be evaluated using some method. However, because no single assessment method exists that can reliably assess the quality of the image processed by the algorithm and generate consistent findings, such an evaluation has never been standardized. The current research's findings primarily compare the subjective and objective qualities of the processed photographs. When different observers use a subjective awareness discrimination algorithm to process images, the results vary greatly because of the different preferences and experiences of the observers. Figure 12 illustrates how to use data from the objective evaluation to correct the divergence caused by the subjective evaluation as closely as possible to improve the evaluation result for an image.
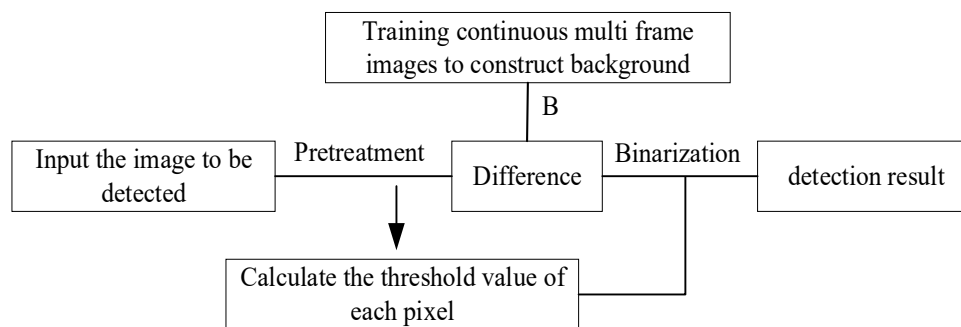
Figure 12. Flow chart of the motion detection algorithm for a complex background.

Figure 12 illustrates how martial arts techniques based on symmetric difference algorithms can distinguish moving objects from complex backgrounds, such as the undulating water surface of a swimming pool, and are suited for settings with somewhat steady lighting conditions. However, the system is unable to accurately detect moving objects in conditions with frequent light transition, vigorous background movement, and poor weather. Figure 13 and the model factor formula demonstrate this.

In fact, connected regions are marked using connected operators, and the area of each marked moving region is calculated. An area smaller than a preset threshold can be deleted [32], that is, small area noise can be removed using this operation to make the model fitting effect more accurate:

$$R(x,y) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M}\left[S_{x,y}(i,j) \times T(i,j)\right]^2}{\sum_{i=1}^{M}\sum_{j=1}^{M}\left[S_{x,y}(i,j)\right]^2}. \tag{6}$$

```
                          ┌──────────┐
                          │  start   │
                          └──────────┘
                               │
                          ┌─────────────────┐
                          │ Get Initial     │
                          │ Background      │
                          └─────────────────┘
                               │
                          ┌─────────────────┐
                          │ Three frame image│
                          │ sequence f₁·f₂·f₃│
                          └─────────────────┘
```

Three frame image sequence $f_1 \cdot f_2 \cdot f_3$

Interframe difference between f_1 and f_2	Interframe difference between f_1 and f_2
Interframe difference filtering operation	Interframe difference filtering operation
Binarization of inter frame difference and edge detection	Binarization of inter frame difference and edge detection

Image "AND" operation to get motion contour of intermediate frame

Morphological operation: fill the internal cavity: clear small noise: smooth the edge of the object

Connected annotation operation: remove the remaining large area of noise; Dimensioning multiple motion regions

Each motion area of connected annotation is made into an external rectangle

Background model update: update the background in the area outside the external rectangles

Segmentation of moving objects by background subtraction information+inter frame difference information reading

The second morphological operation: fill the internal cavity: remove small noise: smooth the edge of the object
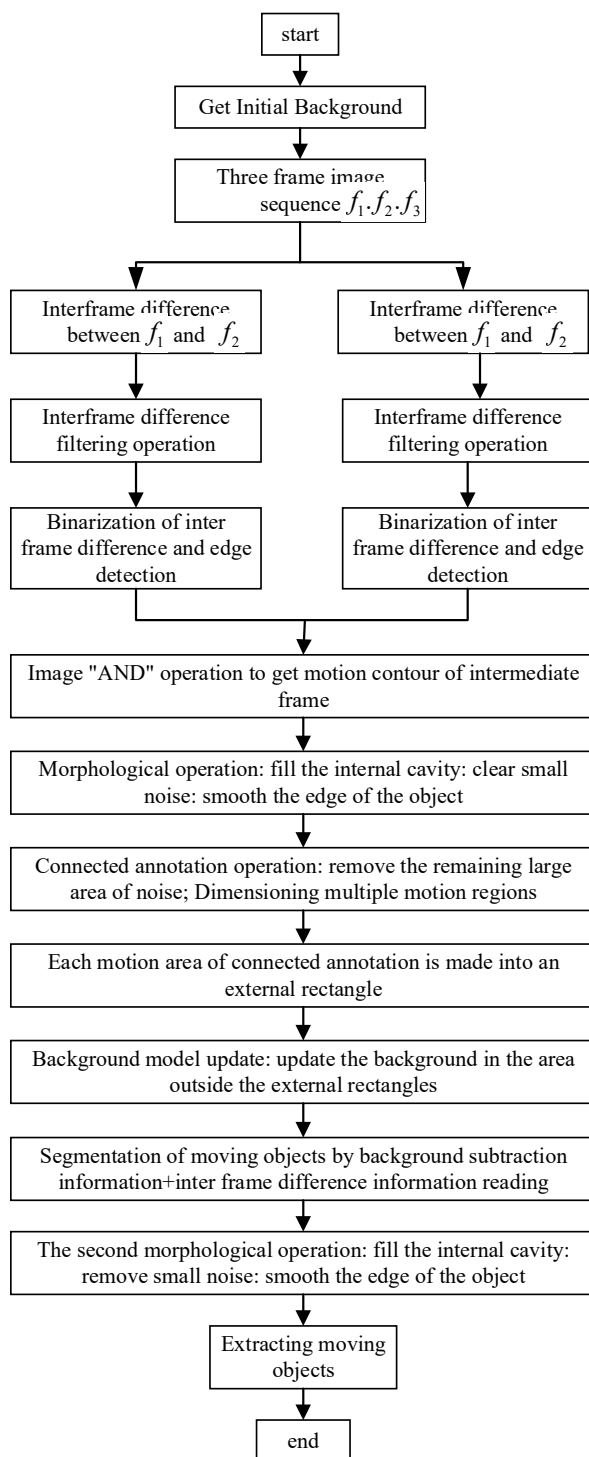
Extracting moving objects

end

Figure 13. Final optimized recognition path.

5.2. Objective analysis of action recognition

In the scientific study of martial arts, a mathematical statistical approach to evaluate the total quality is 3D action recognition. It is difficult to assess the size of a particular target or the level of detail information in a particular area of the image using this approach of mathematical statistical analysis; they can only be calculated using the total infrared image. Furthermore, after the algorithm

is processed, the data evaluated by the objective algorithm cannot be used directly as the identification standard of infrared image quality; it can only reflect some performance indicators of the infrared image and also has a reference value when the infrared image quality is evaluated [33,34]. The average gradient expression can be written as indicated in Table 3 to make the calculation easier.

Table 3. Gradient experiment results for the Wushu action image.

algorithm	HE algorithm	BF&DDE algorithm	AGF&DDE algorithm	Algorithm in this paper
Figure 4-1	6.6297	9.7483	22.2439	10.6818
Figure 4-2	3.6958	4.9213	19.1207	7.3939
Figure 4-3	2.0242	4.1454	10.5497	4.7786
Figure 4-4	2.2679	4.4704	11.4852	5.9031
Figure 4-5	4.2481	6.7935	18.9236	8.2017
Figure 4-6	2.3143	6.4146	15.0491	6.6783
Figure 4-7	2.5783	2.4298	15.4495	5.4668
Figure 4-8	2.1266	4.3669	13.3653	5.7347
Figure 4-9	4.2108	5.5429	16.6565	8.6547

The evaluation of an algorithm's quality for a high dynamic range infrared video image with the non-significant area adaptive enhancement of a martial arts action 3D image is heavily influenced by both the algorithm's execution duration and the processing outcome of the infrared image. Because the study's research object is a dynamic infrared video image, if the algorithm takes too long to process each frame, the dynamic infrared video image will not play smoothly, which gives the viewer the impression that something is "stuck." In the comparison experiment for the algorithm processing time, 10 frames 384 × Based on the 16-bit infrared image of 288, the algorithm in this study was compared with the HE algorithm, BF&DDE algorithm, and AGF&DDE algorithm in terms of processing time. The comparison results for algorithm processing are shown in Table 4.

Table 4. Comparative experimental results of 3D image recognition algorithms for martial arts actions based on the symmetric difference algorithm.

Algorithm Time	HE algorithm processing time	BF&DDE algorithm processing time	AGF&DDE algorithm processing time	Processing time of this algorithm
Urban scene	1.638	15.767	60.103	1.883

The algorithm proposed in this study took 1.882 seconds longer than the HE algorithm to process 10 frames of images for martial arts action 3D image recognition based on the symmetric difference algorithm, which required the least amount of time for infrared image processing. Both the HE algorithm and proposed algorithm processed dynamic video images, and both algorithms played

them without getting stuck. Both the BF&DDE and AGF&DDE algorithms had a lengthy processing time. The dynamic compression algorithm used by the AGF&DDE algorithm caused it to run too slowly. However, the proposed algorithm used the method of creating compressed arrays. Both the BF&DDE algorithm and AGF&DDE algorithm processed dynamic video images in a state of poor playback, which displayed a poor playback effect. Additionally, the algorithm removed some minor background movements, such as swaying leaves and shaking grass, because of the use of the neighborhood difference; hence, they did not affect the detection results. The experiment demonstrated the effectiveness of the moving object detection method presented in this study, which extensively uses spatiotemporal data. The detection times of various algorithms are listed in Table 5 in the appropriate order.

Table 5. Detection time of martial arts action 3D image recognition based on the symmetric difference algorithm.

Background shape	Mixed Gauss model method	Classification method based on pixel gray level	CSTMODA algorithm in this paper
Time consumption (s/frame)	0.05	0.066	0.09

6. Conclusions

Computer-based human behavior analysis primarily concentrates on the identification, detection, tracking, and description of human objects, in addition to the comprehension of human behavior. A computer was used in this study to achieve the intellectualization of sports and to assist in the training of athletes. The basis of studying human behavior is the human posture estimate. To complete the extraction of the athletes' motion features, it can be used to gather information on the joints of athletes. This fundamental data can then be investigated for human behavior analysis, and it offers tremendous convenience for subsequent human behavior analysis. An in-depth design methodology for adaptive Wushu action recognition enhancement based on human posture assessment was presented in this study. The importance of this research was examined first. Second, the present state of motion recognition research was explained, its drawbacks were discussed, such as its huge network model and low efficiency, and a motion recognition method based on key point data classification performed by humans was proposed.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Conflict of interest

The authors declare that they have no conflicts of interest regarding this work.

References

1. A. Tulendiyeva, T. Saliev, Z. Andassova, A. Issabayev, I. Fakhradiyev, Historical overview of injury prevention in traditional martial arts, *Sport Sci. Health*, **17** (2021), 837–848. https://doi.org/10.1007/s11332-021-00785-0

2. H. Liang, Evaluation of fitness state of sports training based on self-organizing neural network, *Neural Comput. Appl.*, **33** (2021), 3953–3965. https://doi.org/10.1007/s00521-020-05551-w

3. S. Starke, Y. Zhao, F. Zinno, T. Komura, Neural animation layering for synthesizing martial arts movements, *ACM Trans. Graphics*, **40** (2021), 1–16. https://doi.org/10.1145/3450626.3459881

4. M. Toshpulatov, W. Lee, S. Lee, A. H. Roudsari, Human pose, hand and mesh estimation using deep learning: a survey, *J. Supercomput.*, **78** (2022), 7616–7654. https://doi.org/10.1007/s11227-021-04184-7

5. Z. J. Zha, J. Liu, T. Yang, Y. Zhang, Spatiotemporal-textual co-attention network for video question answering, *ACM Trans. Multimedia Comput. Commun. Appl.*, **15** (2019), 1–18. https://doi.org/10.1145/3320061

6. H. Kwon, C. Tong, H. Haresamudram, Y. Gao, G. D. Abowd, N. D. Lane, et al., IMUTube: Automatic extraction of virtual on-body accelerometry from video for human activity recognition, in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **4** (2020), 1–29. https://doi.org/10.1145/3411841

7. D. A. Kumar, A. S. C. S. Sastry, P. V. V. Kishore, E. K. Kumar, Indian sign language recognition using graph matching on 3D motion captured signs, *Multimedia Tools Appl.*, **77** (2018), 32063–32091. https://doi.org/10.1007/s11042-018-6199-7

8. L. H. Long, Role of artificial intelligence algorithm for taekwondo teaching effect evaluation model, *J. Intell. Fuzzy Syst.*, **40** (2021), 3239–3250. https://doi.org/10.3233/JIFS-189364

9. B. M. Craig, A. J. Lee, Stereotypes and structure in the interaction between facial emotional expression and sex characteristics, *Adapt. Hum. Behav. Physiol.*, **6** (2020), 212–235. https://doi.org/10.1007/s40750-020-00141-5

10. Z. Wu, S. Shen, X. Lian, X. Su, E. Chen, A dummy-based user privacy protection approach for text information retrieval, *Knowledge-Based Syst.*, **195** (2020), 105679. https://doi.org/10.1016/j.knosys.2020.105679

11. R. Zhang, F. Torabi, G. Warnell, P. Stone, Recent advances in leveraging human guidance for sequential decision-making tasks, *Auton. Agents Multi-Agent Syst.*, **35** (2021), 1–39. https://doi.org/10.1007/s10458-021-09514-w

12. Z. Wu, S. Xuan, J. Xie, C. Lin, C. Lu, How to ensure the confidentiality of electronic medical records on the cloud: A technical perspective, *Comput. Biol. Med.*, **147** (2022), 105726. https://doi.org/10.1016/j.compbiomed.2022.105726

13. A. K. Mackenzie, M. L. Vernon, P. R. Cox, D. Crundall, R. C. Daly, D. Guest, et al., The multiple object avoidance (MOA) task measures attention for action: Evidence from driving and sport, *Behav. Res. Methods*, **54** (2022), 1508–1529. https://doi.org/10.3758/s13428-021-01679-2

14. Z. Wu, S. Shen, H. Li, H. Zhou, C. Lu, A basic framework for privacy protection in personalized information retrieval: An effective framework for user privacy protection, *J. Organ. End User Comput.*, **33** (2021), 1–26. https://doi.org/10.4018/JOEUC.292526

15. M. Rana, V. Mittal, Wearable sensors for real-time kinematics analysis in sports: a review, *IEEE Sens. J.*, **21** (2020), 1187–1207. https://doi.org/10.1109/JSEN.2020.3019016

16. Z. Wu, G. Li, S. Shen, X. Lian, E. Chen, G. Xu, Constructing dummy query sequences to protect location privacy and query privacy in location-based services, *World Wide Web*, **24** (2021), 25–49. https://doi.org/10.1007/s11280-020-00830-x

17. F. Malawski, Depth versus inertial sensors in real-time sports analysis: a case study on fencing, *IEEE Sens. J.*, **21** (2020), 5133–5142. https://doi.org/10.1109/JSEN.2020.3036436

18. Z. Wu, S. Shen, H. Zhou, H. Li, C. Lu, D. Zou, An effective approach for the protection of user commodity viewing privacy in e-commerce website, *Knowledge-Based Syst.*, **220** (2021), 106952. https://doi.org/10.1016/j.knosys.2021.106952

19. E. Kon, B. D. Matteo, P. Verdonk, M. Drobnic, O. Dulic, G. Gavrilovic, et al., Aragonite-based Scaffold for the treatment of joint surface lesions in mild to moderate osteoarthritic knees: results of a 2-year multicenter prospective study, *Am. J. Sports Med.*, **49** (2021), 588–598. https://doi.org/10.1177/0363546520981750

20. Z. Liu, L. Li, S. Liu, Y. Sun, S. Li, M. Yi, et al., Reduced feelings of regret and enhanced fronto-striatal connectivity in elders with long-term Tai Chi experience, *Social Cognit. Affective Neurosci.*, **15** (2020), 861–873. https://doi.org/10.1093/scan/nsaa111

21. K. Petri, P. Emmermacher, M. Danneberg, S. Masik, F. Eckardt, S. Weichelt, et al., Training using virtual reality improves response behavior in karate kumite, *Sports Eng.*, **22** (2019), 1–12. https://doi.org/10.1007/s12283-019-0299-0

22. R. Lozada-Yánez, N. La-Serna-Palomino, F. Molina-Granj, Augmented reality and MS-kinect in the learning of basic mathematics: KARMLS case, *Int. Educ. Stud.*, **12** (2019), 54–69. https://doi.org/10.5539/ies.v12n9p54

23. J. C. Zhou, J. M. Sun, W. S. Zhang, Z. F. Lin, Multi-view underwater image enhancement method via embedded fusion mechanism, *Eng. Appl. Artif. Intell.*, **121** (2023), 105946. https://doi.org/10.1016/j.engappai.2023.105946

24. P. Parrend, P. Collet, A review on complex system engineering, *J. Syst. Sci. Complexity*, **33** (2020), 1755–1784. https://doi.org/10.1007/s11424-020-8275-0

25. X. Huang, R. Ball, W. Wang, Comparative study of industrial design undergraduate education in China and USA, *Int. J. Technol. Des. Educ.*, **31** (2021), 565–586. https://doi.org/10.1007/s10798-020-09563-4

26. H. W. Wu, E. Fajiculay, J. F. Wu, C. C. S. Yan, C. P. Hsu, S. H. Wu, Noise reduction by upstream open reading frames, *Nat. Plants*, **8** (2020), 474–480. https://doi.org/10.1038/s41477-022-01136-8

27. E. O. Abiodun, A. Alabdulatif, O. I. Abiodun, M. Alawida, A. Alabdulatif, R. S. Alkhawaldeh, A systematic review of emerging feature selection optimization methods for optimal text classification: the present state and prospective opportunities, *Neural Comput. Appl.*, **33** (2021), 15091–15118. https://doi.org/10.1007/s00521-021-06406-8

28. H. Dong, L. Zhao, Y. Shu, N. N. Xiong, X-ray image denoising based on wavelet transform and median filter, *Appl. Math. Nonlinear Sci.*, **5** (2020), 435–442. https://doi.org/10.2478/amns.2020.2.00062

29. L. Jiang, T. Zhang, Y. Feng, Identifying the critical factors of sustainable manufacturing using the fuzzy DEMATEL method, *Appl. Math. Nonlinear Sci.*, **5** (2020), 391–404. https://doi.org/10.2478/amns.2020.2.00045

30. J. Feng, M. Meng, S. Liu, X. Zhang, J. Yuan, Z. Zhang, Prediction of Chinese automobile growing trend considering vehicle adaptability based on Cui–Lawson model, *Appl. Math. Nonlinear Sci.*, **5** (2020), 367–376. https://doi.org/10.2478/amns.2020.2.00054

31. Z. Lao, D. Pan, H. Yuan, J. Ni, S. Ji, W. Zhu, et al., Mechanical-tunable capillary-force-driven self-assembled hierarchical structures on soft substrate, *ACS Nano*, **12** (2018), 10142–10150. https://doi.org/10.1021/acsnano.8b05024

32. X. Luo, C. Zhang, L. Bai, A fixed clustering protocol based on random relay strategy for EHWSN, *Digital Commun. Networks*, **9** (2023), 90–100. https://doi.org/10.1016/j.dcan.2022.09.005

33. J. C. Zhou, L. Pang, W. S. Zhang, Underwater image enhancement method by multi-interval histogram equalization, *IEEE J. Oceanic Eng.*, **48** (2023), 474–488. https://doi.org/10.1109/JOE.2022.3223733

34. Z. K. Wang, H. L. Zhen, J. D. Deng, Q. F Zhang, X. J. Li, M. X. Yuan, et al., Multiobjective optimization-aided decision-making system for large-scale manufacturing planning, *IEEE Trans. Cybern.*, **52** (2022), 8326–8339. https://doi.org/10.1109/TCYB.2021.3049712