Research Article

Multidimensional Analysis of Physical Education Teaching Based on Big Data Action Scenarios

Feng Liu1 and Junjie Lv2

1Department of Physical Education Chengdu University of Information Technology, Sichuan Chengdu 610225, China
2Yulin Normal University Institute of Physical Education and Health, Yulin, 537000 Guangxi, China

Correspondence should be addressed to Junjie Lv; lvjunjie@ylu.edu.cn

Received 6 June 2022; Revised 7 August 2022; Accepted 10 August 2022; Published 6 September 2022

Copyright © 2022 Feng Liu and Junjie Lv. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Combined with the development of the current educational environment, physical education will replace English as the third subject of education in the future. The breakthrough in physical education teaching is also gradually changing from the general education form to today’s smart education. In the era of big data, physical education can also apply this technology. In action scenarios based on big data, operations such as correction of detailed actions or monitoring and identification of key actions are common. Through the computer vision system, the rational judgment of the computer can be used to give follow-up training points. Also able to store personal data during training. The presentation of algorithms cannot be avoided through computer vision. Based on Action Bank as the basic algorithm, this paper proposes a template research method based on multispectral clustering and has been applied in Action Bank. The tedious manual template selection is eliminated. This method replaces it to facilitate its dissemination in different databases. In this method, due to the slow speed of extracting features, a fast algorithm of quantitative Action Bank is extracted. The experimental part of the article compares whether the algorithm has been optimized in terms of performance before and after optimization. The resolution, time consumption, and detection errors of the Action Bank model are carried out. The experimental exploration and data collection and comparison are carried out. After the experimental optimization, the performance has been improved. By comparing the meanshift detection method and spatiotemporal action detection method with the Action Bank model mentioned in this article, the experimental data of resolution, time consumption, and detection error are compared, and the Action Bank model is obtained in terms of resolution, time consumption, and error detection. The time consumption is better than the other two algorithm models, but there is room for improvement in detecting errors, but the experimental results also meet the current detection requirements, and in the current physical education teaching, it also occupies the forefront of this field, at the forefront. Ministry of Education. Instead of English, physical education has become the third subject! On April 21, the Ministry of Education released the "Compulsory Education Physical Education and Health Curriculum Standard (2022 Edition)," and the new curriculum standard will be officially implemented in the fall semester of 2022. Among them, the proportion of the total class hours of “Sports and Health” is 10-11%, surpassing foreign language to become the third major subject in the primary and early stages. Physical education has already started its journey. Traditional physical education can no longer meet the development of the current environment. The arrival of the information society has also brought about the development and progress of physical education. Only by better combining the current information technology can it be satisfied. The needs of physical education is in the future.

1. Introduction

Sports analysis based on big data action scenes, through a large number of dynamic videos and example galleries, combined with the presentation of algorithms, a large number of algorithm experiments and detection experiments are carried out [1–5]. After many experiments, it is ensured that the algorithm can complete the given motion tracking task and motion data acquisition task under the condition of meeting the time limit requirement [6–8]. However, based
on the processing of factors such as lighting conditions, contrast, and jitter rate [9–11], the computational and data acquisition costs of the algorithm tend to be higher and higher. Therefore, in this paper, all the variables considered above are regarded as irrelevant variables, it is only a single case of discussing the motion state of the object, and the Action Bank algorithm will present the tracking effect in real time [12–13]. Before this, the pose estimation of a single variable and the giving of the probability function are the core points of the algorithm, and the extraction and recognition of feature points that show correlation with it and the construction of the kernel function are the first steps in the algorithm [14]. Therefore, to complete the computer action, all parties need to be contacted and cooperate closely. The recognition of the human body and the estimation of the posture are jointly completed, which has practical value for physical education teaching [15].

Combined with the development of the current educational environment, physical education will replace English as the third subject of education in the future. The breakthrough of physical education teaching is also gradually changing from the general education form to today’s smart education, etc., in the era of big data. This technology can also be used in physical education. In the action scene based on big data, the correction of detailed actions or the monitoring and identification of key actions are generally performed. Through the computer vision system, the rational judgment of the computer can give follow-up training. Points are also able to store personal data during training.

2. Multidimensional Analysis of Physical Education Teaching Based on Big Data Action Scenarios

This paper will complete the rapid and accurate identification, detection, and storage of sports-related activity data based on big data action scenarios [16–19]. The basic big data and physical education analysis is shown in Figure 1.

Based on the theoretical support of behaviorism and cognitivism learning theory, wisdom education, connectivism, the technical support of 3D energy characteristics, directional energy characteristics, and other technologies, this paper will jointly analyze physical education in big data action scenarios. This paper will design a feasible and scientific action collection model through the comprehensive application of the key technologies of motion data collection and human body model system plus theoretical support and can be effectively applied to basic physical education, such as the collection and collection of basic motion actions, demonstration of basic sports actions, and correction of sports actions, based on big data technology, it is very convenient to store and reference actions in the future [20–23].

2.1. Introduction of Action Bank Model

2.1.1. 3D Direction Energy Feature. The dynamic activities of people can be seen as the decomposition of energy in multiple directions. If effectively decomposed, then, the action of only one point can be considered, the energy of which can converge in different directions of the space-time cube. As an expression for the minimum activity level, this simplest decomposition is the most basic method for determining activity location [24, 25].

The decomposition of the energy in the space-time domain is achieved by a generalized filter of the third-order Gaussian derivative, which is labelled \( G_{\mathbf{\tilde{\theta}}} (x) \), where \( \mathbf{\tilde{\theta}} \) is used to represent the direction of the filter at this point, and \( x \) represents the position of the energy in the space-time three-dimensional space at this point. The operation for a certain point energy set is as follows:

\[
\tilde{E}_{\mathbf{\tilde{\theta}}}(x) = \sum_{\mathbf{\theta} \in \Omega} \left( G_{\mathbf{\tilde{\theta}}} \ast I \right)^2.
\]

(1)

can be obtained by conventional methods using variable filters:

\[
\hat{\mathbf{\theta}} = \cos \left( \frac{\pi i}{4} \right) \hat{\theta}_a(\mathbf{n}) + \sin \left( \frac{\pi i}{4} \right) \hat{\theta}_b(\mathbf{n}),
\]

(2)

where \( \hat{\theta}_a = \mathbf{n} \times \mathbf{\tilde{z}} || \mathbf{n} \times \mathbf{\tilde{z}} \|, \hat{\theta}_b = \mathbf{n} \times \hat{\theta}_a(\mathbf{n}), \mathbf{\tilde{z}} \) is a vector, and the range is in \( 0 \leq i \leq 3 \), and the group can decompose the energy into any direction in the frequency domain, namely,

\[
E_n(x) = \sum_{i=0}^{N} E_{\hat{\theta}}(x).
\]

(3)

\( \hat{\theta}(x) \) is one of several different directions in equation (2), each \( E_{\hat{\theta}}(x) \) computed with a directional energy filter.

In order to eliminate the influence of the direction energy itself and the corresponding relationship, it is necessary to unify the directions generated by equation (3), and it is necessary to divide the energy in each direction by the total energy:

\[
\tilde{E}_{\mathbf{\tilde{\theta}}}(x) = \frac{\tilde{E}_{\mathbf{\tilde{\theta}}}(x)}{\sum_{j=0}^{M} \tilde{E}_{\mathbf{\tilde{\theta}}}(x) + \varepsilon}.
\]

(4)

Expression \( M \) representing the number of energy directions is selected, and \( \varepsilon \) enters the value required to prevent some total energy from being too small to cause instability. Typically, \( \varepsilon \) a seventh energy measure is entered to indicate no structural action:

\[
\tilde{E}_{\mathbf{\tilde{\theta}}}(x) = \frac{\varepsilon}{\sum_{j=0}^{M} \tilde{E}_{\mathbf{\tilde{\theta}}}(x) + \varepsilon}.
\]

(5)

From the concept that using the formulas from (1) to (5), the energy spectrum of different planes in the space domain and the energy \( x \) \( y \) \( t \) expressed in different directions of space can be extracted from the character activity sequence. The energy in different directions is decomposed into a point, the active sample template is combined with
the video waiting to be detected, and the positioning of the active sample can be carried out.

2.1.2. Template Matching of Directional Energy Features

(1) Space-Time Template Matching. To detect movement in large query campaigns, the 3D model needs to slide across all positions in the space-time cube. At each location, the similarity between the raw energy signature of the model and the directional energy within the coverage of the space-time cube is calculated at the same time as the distance and then added together:

$$M(x) = \sum_{m} m[S(u), T(u - x)].$$

Here, moving along the three axes of space-time space \(m[S(u), T(u - x)]\) refers to the similarity between the features \(S\) of the query task activity and the template features \(T\).

In fact, many methods for calculating the relativity of histograms can be used to calculate the similarity between energy characteristics, and the Babbitt coefficient is used here. For histograms \(p\) and histograms \(Q\), if each contains a channel, the Bap coefficients are defined as follows:

$$m(P, Q) = \sqrt{\sum_{b=1}^{B} P_b Q_b^*}. \tag{7}$$

The calculation of the correlation value is limited between 0 and 1, with closer to 0 indicating a complete mismatch, and closer to 1 the higher the similarity.

The final step of detection selects \(M\) the most likely matching point in the space-time cube, which is the most likely template movement position. The local maxima of the space-time cube are achieved under low braking.

(2) Template Matching of Assigned Values. In tasks that match a specific template, it may be necessary to set different values for different areas of the template. This can be done by changing the global negotiation function:

$$M(x) = \sum_{m} w(u)m[S(u), T(u - x)].$$

(3) Height Matching. In the previous field of pattern recognition, high matching usually consists of several methods: (1) transition from fuzzy to clear search using spatiotemporal pyramids, (2) preliminary estimation of the requester’s location from particle samples, (3) evaluation using template subclasses, and (4) the matching calculation is terminated early. The results of these calculations can lead to target loss. The method used to determine the motion position is described here. In equation (7), the mode distance is equal to calculating the sum of the associations between each channel:

$$M(x) = \sum_{u} \sum_{b} \sqrt{S_b(u)} \sqrt{T_b(u - x)} = \sum_{b} \sqrt{S_b^*} \sqrt{T_b^*}. \tag{9}$$

This * represents interrelated operations, and \(b\) represents the index number of bins.

Therefore, the correlation distance can be efficiently calculated using the convolution theorem from Fourier transform to frequency space, so that the time-consuming correlation operation in complex space becomes the multiplication of space points in the frequency range:

$$M(x) = F^{-1}\left\{ \sum_{b} F\{S_b^*\} F\{T_b^*\} \right\}. \tag{10}$$

Expressions \(F\{\cdot\}\) and \(F^{-1}\{\cdot\}\) represent the Fourier transform and the inverse Fourier transform, respectively, and \(T_b^*\) represent the corresponding templates. In the implementation process, the Fourier transform can be effectively realized by the fast discrete Fourier transform.
2.1.3. Complexity Analysis. Let \( W_{(T,S)}, H_{(T,S)}, \) and \( D_{(T,S)} \) specify the width, height, and frame number of the template video \( T \) and the video to be detected, respectively. The S Action Bank algorithm can be regarded as two parts, one part is the generation of 3D directional energy features, and the other part is the combination based on this feature. The time at which the 3D directional energy feature \( O(W \times H \times D \times k) \) appears as, where \( k \) represents the length of the filter. In the matching process, equation (6) points out the complexity of the space-time field matching based on the Buckley coefficient:

\[
O \left( \sum_{i \in (T,S)} W_i H_i D_i \right).
\]

Formulas (2)–(10) convert the correlation operation into the frequency domain product and divide the 3D FFT into 1D FFT for calculation, and the complexity is

\[
O[BW_5 H_5 S_5 (\log_2 D_5 + \log_2 W_5 + \log_2 H_5)].
\]

It can be obtained that an efficient template matching method is adopted in the frequency range, which greatly reduces the computational complexity in the matching process and enables the directional energy characteristics to be processed quickly.

2.1.4. Extracting Features for Motion Localization. Dynamic motion can be regarded as energy moving in different directions, then when only one point is considered as energy, if it can be effectively decomposed, it can be regarded as the combination of energy in different directions for this point on the space-time cube. This decomposition is the lowest and most basic activity expression in the positioning algorithm.

The energy decomposition in the space-time cube is achieved by a 3rd-order Gaussian derivative and a three-bit filter, which can be written as \( G_{\theta}(x) \), where \( \theta \) represents the direction of the filter at that point, and the unknown \( x \) represents the point position in the spacetime cube. Among them, the value of the point corresponding to the video to be detected around the point \( \Omega \) of the filter is

\[
E_{\theta}(x) = \sum_{x=\Omega} (G_{\theta} * I)^2.
\]

Third-order filter banks can be obtained with conventional tunable filters:

\[
\hat{\theta}_i = \cos \left( \frac{\pi i}{4} \right) \hat{\theta}_a(\hat{n}).
\]

The correct way to write this formula should be

\[
\hat{\theta}_i = \cos \left( \frac{\pi i}{4} \right) \hat{\theta}_a(\hat{n}) + \sin \left( \frac{\pi i}{4} \right) \hat{\theta}_b(\hat{n}),
\]

where \( \hat{\theta}_a = \hat{n} \times \hat{e}_x/||\hat{n} \times \hat{e}_x||, \hat{\theta}_b = \hat{n} \times \hat{\theta}_a(\hat{n}), \hat{e} \) is a vector, the range is in \( 0 \leq i \leq 3 \), and the group can decompose the energy into any direction in the frequency domain, namely,

\[
E_{\theta}(x) = \sum_{i=0}^{3} E_{\theta i}(x).
\]

\( \hat{\theta}(x) \) is one of several different directions in equation (2), each \( E_{\theta}(x) \) is computed with a directional energy filter.

Finally, it is found that rest energy \( E_r \) and unstructured energy have \( E_s \) nothing to do with motion, they can be used as separation salient energy, and other five kinds of energy can be obtained. These five energies can be set uniformly and finally form five channels of eigenvalues.

For the Action Bank detector, it defines 7 original space-time energies, they are static function \( E_s \), rightward energy \( E_r \), upward energy \( E_u \), downward energy \( E_d \), flickering energy \( E_f \), and unstructured energy \( E_o \).

The added formula is

\[
E_i = E_r - E_s - E_o.
\]

3. Improvement and Optimization of the Algorithm

If a hyperplane is found by a linear classifier in a dimensional vector, then, there is the following equation:

\[
w^T x + b = 0.
\]

There is one, the two-dimensional plane hyperplane is a straight line, \( w \) is the normal vector, and \( b \) refers to a segment on the intercept. After obtaining the hyperplane, the formula of the classification function will become:

\[
f(x) = w^T x + b.
\]

As can be seen from the formula, \( f(x) = 0 \) the points \( x \) are points at the hyperplane, \( f(x) < 0 \) or \( f(x) > 0 \), the samples to be classified can be divided into two categories. So \( f(x) < 0 \), the categories of \( y \) sum can be +1 sum and \( f(x) > 0 \), respectively, -1.

From the point to the hyperplane is expressed by the certainty or validity of the classification result \( ||w^T x + b|| \). Absolute values can be removed by using categorical markers and then entering the \( y \) formula for the interval between functions:

\[
\tilde{y} = y(\tilde{w}^T x + b) = yf(x).
\]

Since the size of the range of the function is affected by the size of the set parameters, for example, the parameters \( ||w|| \) and \( ||b|| \) are modified proportionally, and the range of the function will also be affected and will be changed proportionally, so the geometric range is introduced to adjust the range of the function:

\[
\tilde{y} = \frac{\tilde{y}}{||w||} = y \left( \frac{w^T x + b}{||w||} \right).
\]
As the interval of the geometric interval becomes larger and larger, the confidence of the classification becomes higher and higher. When the training example \( T \) contains \( n \) points, we will make these \( n \) points have a reasonable distance from the reference vector in the hyperplane.

The biggest problem in the optimization of \( w \) and \( b \) parameters in the hyperplane range is the problem of geometric intervals:

\[
\max \hat{y}_i s.t. y_i (w^T x_i + b) = \hat{y}_i \geq \hat{y}.
\]  

(22)

The interval of the value will change with the change of the parameters. In order to facilitate the optimization and export of the function, the interval of the function is substituted by \( \hat{y} = 1 \), and the following formula is obtained:

\[
\max \frac{1}{||w||} s.t. y_i (w^T x_i + b) = \hat{y}_i \geq \hat{y}.
\]  

(23)

The equation can also be written in the following form:

\[
\min \frac{1}{2} ||w||^2, s.t. y_i (w^T x_i + b) \geq 1.
\]  

(24)

### Table 1: Experimental simulation data table.

<table>
<thead>
<tr>
<th>Example</th>
<th>Action data storage(s)</th>
<th>Action error detection times</th>
<th>Action correction times</th>
<th>Error correction rate (%)</th>
<th>Satisfaction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running action</td>
<td>162</td>
<td>36</td>
<td>Twenty three</td>
<td>63.8</td>
<td>80.6</td>
</tr>
<tr>
<td>Shooting action</td>
<td>195</td>
<td>45</td>
<td>30</td>
<td>66.7</td>
<td>85.3</td>
</tr>
<tr>
<td>Table tennis swing</td>
<td>145</td>
<td>33</td>
<td>Twenty one</td>
<td>63.6</td>
<td>96.3</td>
</tr>
<tr>
<td>Badminton swing</td>
<td>207</td>
<td>56</td>
<td>31</td>
<td>55.3</td>
<td>84.2</td>
</tr>
</tbody>
</table>

### Table 2: Experimental data comparison table.

<table>
<thead>
<tr>
<th>Example</th>
<th>Error detection times</th>
<th>Error correction times</th>
<th>Effective error correction</th>
<th>Error correction rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional physical education</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td>Action Bank model</td>
<td>16</td>
<td>12</td>
<td>10</td>
<td>75</td>
</tr>
</tbody>
</table>

### Table 3: Comparison of detection amount and resolution (%).

<table>
<thead>
<tr>
<th>Example</th>
<th>Primary model</th>
<th>Optimization model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running action</td>
<td>69.2</td>
<td>70.56</td>
</tr>
<tr>
<td>Shooting action</td>
<td>79.3</td>
<td>82.01</td>
</tr>
<tr>
<td>Table tennis swing</td>
<td>85.2</td>
<td>87.21</td>
</tr>
<tr>
<td>Badminton swing</td>
<td>85.6</td>
<td>86.24</td>
</tr>
</tbody>
</table>

The quadratic programming problem with linear constraints introduces Lagrange multipliers \( a \), and the objective function is as follows:

\[
L(w, b, a) = \frac{1}{2} ||w||^2 - \sum_{i=1}^{n} a_i (y_i (w^T x_i + b) - 1). 
\]  

(25)

Converted to the dual problem, fixed \( a \), minimize the \( w \) sum with \( b \) respect to the objective function \( L \):

\[
\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^{n} a_i y_i x_i.
\]  

(26)
Formula (26) in the article is solved together based on the formula (25), but only one formula cannot be obtained, and another formula should be added to solve it together.

\[
\frac{\partial L}{\partial w} = 0 \Rightarrow \sum_{i=1}^{n} a_i y_i = 0. \tag{27}
\]

From equation (27) into equation (26), we can get

\[
\max \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i=1}^{n} a_i a_j y_i y_j x_i^T x_j, s.t., a_i \geq 0, i = 1 \cdots n. \tag{28}
\]

Equation (28) can be quickly solved by the sequence minimum optimization algorithm, which ensures the efficiency of the SVM algorithm.

4. Experimental Simulation

In Action Bank, hand-picked templates were chosen with an average pixel of 50, 120, and a time length of 40-50 frames.

Traditional physical education can no longer be satisfied. In the current educational environment, based on the technology of Action Bank, a multidimensional analysis of physical education is carried out. Traditional sports can only provide simple action guidance for human actions under the intuitive vision of the human body. However, in the big data motion scene, it is possible to store the data of the human body dynamic motion accurate to the frame, to score and correct the accuracy of the dynamic motion, and to match the corresponding data based on the stored dynamic motion data. Intuitive analysis of sports. Now, with the technology of Action Bank, the experimental simulation of the existing sports actions is carried out to compare the improvement of the Action Bank model compared with the traditional physical education. Now conduct model experiments based on a certain action data, and get Table 1, as shown in the table.

The above is a demonstration of the experimental data after the model simulation experiment. It can be seen that the error correction rate of the experimental dynamic data is still maintained at more than half. Now, the same example is artificially tested, and then, the obtained data are compared. Table 2 is as shown in Figure 2.

As shown in the table, it can be intuitively found that the traditional physical education is manual inspection, so the error detection, error correction, and effective error correction are all the same, but the effective error detection of the model is more than the manual error detection by 2. From this, it can be seen that manual error detection will still be more or less affected by factors such as environment, man-made, and force majeure, while the model will not be affected by so many factors. Unless there is an error in the algorithm itself, the shortcomings of the model are reflected in the correction. The error rate may be due to the lack of control data. In the future experiments, the basic behavior will be stored, and the database will be improved to achieve the optimal operation of the model.

4.1. Performance Comparison before and after Model Optimization. Now collect and collect an example action in a big data action scene. By comparing the resolution of the experiment, the consumption of recognition time, and the number of recognition errors, it is concluded whether the optimization of the model has a substantial effect on the model. Promote.

4.2. Fiction Resolution Comparison. Comparison of unit detection volume and resolution is as shown in Table 3 and as shown in Figure 3.

It can be seen from the figure that the blue one is the primary model, and the orange one is the optimized model. It can be seen intuitively that the optimized model has a significantly improved resolution than the primary model. According to
the table, it can be seen that in the resolution data, there is also a significant improvement in accuracy.

4.3. Time Consumption Comparison of Unit Detection Amount. In the action scene of big data, the time requirements for detection are relatively strict, and the general detection must ensure the timeliness, so that the data obtained from the detection can be effectively used.

Comparison of time consumption per unit detection amount is as shown in Table 4 and as shown in Figure 4.

The figure shows that the blue one is the primary model, and the orange one is the optimized model. It can be seen intuitively that the time consumption of the optimized model is shorter, and it can better meet the requirements of the experiment for timeliness.

4.4. Comparison of the Number of Recognition Errors per Unit Detection Amount. The detection of model performance can be intuitively analyzed by the number of errors in the number of experiments. By comparing the number of errors before and after the optimization of the experimental model, the performance of the model before and after optimization can be compared, as shown in Table 5 and as shown in Figure 5.

In the figure, the blue one is the primary model, and the orange one is the optimized model. It can be seen intuitively that after the model is optimized, the number of detection errors has dropped significantly, meeting the performance requirements of the experiment.

To sum up, the performance of the optimized model is more in line with the inspection requirements, the resolution of the experiment is optimized, and the motion data obtained by the inspection can be stored more accurately. The reduction in the number of recognition errors is an essential improvement in the performance of the model.

Tables 1 to 3 in the article are mainly for the comparison before and after the optimization of the algorithm used in the article, based on the simulation and data collection of the experiment based on the resolution of the algorithm data, time consumption, and the number of detected errors. Finally, the experimental results are obtained, and the optimized algorithm is much better than the initial algorithm in terms of resolution, time consumption, or the number of detection errors.

4.5. Performance Comparison before and after Model Optimization. Based on the experimental data in Section 4.2 of the article, this paragraph compares the algorithm used in this article by introducing the meanshift algorithm and spatiotemporal detection method and compares the resolution, time consumption, and number of detection errors. The Action Bank model used in this paper is better than the other two algorithms in terms of resolution and time consumption, but it is slightly inferior to the spatiotemporal detection method in the number of errors in checking. Therefore, in the future algorithm optimization and improvement, it will be more inclined to modify the detection of the algorithm.

By introducing the meanshift detection method and the spatiotemporal action detection method, the action scenes of several cases are detected and analyzed, and the experimental models or algorithms are compared by comparing the resolution, time consumption, and number of experimental errors of the experiments. Finally, the Action Bank model is obtained. Whether it meets the requirements for detection in terms of performance.

The meanshift algorithm refers to the use of the gradient of the probability density to solve and find the optimal solution.

4.5.1. Spatiotemporal Detection Method. The early spatiotemporal action detection is to process frame by frame to obtain the bounding box and action category of the characters in each frame and then connect these boxes along the time dimension to form the spatiotemporal action detection result.
Now for the detection of the experimental resolution, based on the data collected and sorted out, the results are shown in Table 6, as shown in Figure 6.

In the figure, the blue is the Action Bank model, the orange is the meanshift detection method, and the gray is the spatiotemporal action detection method. It can be seen intuitively from the figure that the blue line in the figure is the resolution data of the Action Bank model. Both are at the top of the datasheet, and it follows that the Action Bank model is superior to the other two models in terms of resolution.

The time consumption of the referenced model or algorithm is now tested to determine whether the timeliness of the model or algorithm meets the testing requirements. The testing data is shown in Table 7, as shown in Figure 7.

In the figure, the blue one is the Action Bank model, the orange one is the meanshift detection method, and the gray one is the spatiotemporal action detection method. You can intuitively see the blue line, that is, the Action Bank model time consumption data is at the lowest end of the table. It can be concluded that the model is better than other models in terms of time consumption, but based on the consideration of timeliness, the time consumption of all models meets the requirements for timeliness.

The number of inspection errors is the most intuitive data comparison for the performance of the model or algorithm. The results of the inspection data are shown in Table 8, as shown in Figure 8.

In the figure, the blue one is the Action Bank model, the orange one is the meanshift detection method, and the gray
one is the spatiotemporal action detection method. It can be seen intuitively that in the table, the data of the gray part of the spatiotemporal action detection method is relatively lower than the rest. There are two models and algorithms, but the number of errors of the Action Bank model is also within the scope of the inspection requirements, so it basically meets the requirements in terms of performance.

Tables 4 to 6 of the article are experimentally compared with the algorithm used in this article by introducing the meanshift algorithm and spatiotemporal detection method, and by comparing the resolution, time consumption, and number of detection errors, and finally get the results used in this article. The Action Bank model is better than the other two algorithms in terms of resolution and time...
consumption, but it is slightly inferior to the spatiotemporal detection method in the number of errors in checking. However, the requirements for detection are also met.

To sum up, the Action Bank model is better than the other two models and algorithms in terms of resolution and time consumption. Although it does not perform relatively well in terms of the number of detected errors, but based on performance considerations, the Action Bank model fully meets the requirements and provides directions for the future development and optimization of the model.

This paper studies the human activity recognition and retrieval of big data videos, improves the application of the Action Bank method in the context of big data, and has achieved satisfactory results, but there are still deficiencies in some aspects and room for improvement. In the future work, the author will continue to seek improvements in at least the following two points: (1) the method of template learning and (2) the setting of the parameters of the quantization Action Bank algorithm.

5. Conclusion

Combined with the above, the article compares the resolution, detection time consumption, and number of detection errors before and after the optimization of the model through experimental simulation and obtains the experimental data. It is concluded that the performance of the optimized model is significantly improved than that of the primary model. Then, the meanshift detection method and the spatiotemporal action detection method are introduced. By comparing the resolution, detection time consumption, and detection error times of these two algorithms with the Action Bank model, the resolution and detection time of the Action Bank model are obtained. The consumption is better than the other two algorithms, but it is weaker than the spatiotemporal detection method in terms of detecting errors. Meet the trends of the current environment.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

Acknowledgments

This work was supported by 2022 Guangxi University Young and Middle-aged Teachers’ Basic Research Ability Improvement Project; Guangxi Minority Sports to Help Rural Revitalization Path Research; Project no.: (2022KY0559).

References


