SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(1), 808.
https://doi.org/10.62617/mcb808

Article

Badminton striking motion recognition based on dense trajectory algorithm

Youfeng Yang[1], Yin Zhang[2], Chang Li[3,*]

[1] College of Physical Education, Heze University, Heze 274000, China
[2] Department of Culture and Science and Technology, Shaanxi Provincial Party School, Xi'an 710061, Shaanxi, China
[3] Department of Physical Education, XingTai Medical College, Xingtai 054000, China
*** Corresponding author:** Chang Li, lichang19861107@126.com

CITATION

Yang Y, Zhang Y, Li C. Badminton striking motion recognition based on dense trajectory algorithm. Molecular & Cellular Biomechanics. 2025; 22(1): 808. https://doi.org/10.62617/mcb808

ARTICLE INFO

Received: 15 November 2024
Accepted: 27 November 2024
Available online: 9 January 2025

Abstract: With people paying more attention to health and quality of life, more and more people are keen to exercise through various sports, such as badminton, which is suitable for all ages. With the development of the new era, artificial intelligence and robots have gradually entered people's lives, among which badminton robot is an intelligent robot that can compete with people in real time. The robot can not only capture and track badminton moving at high speed, but also has a high-speed motion control system to accurately complete the striking action. Therefore, the main research content of this paper is based on badminton robot vision system to capture, identify and analyze badminton players' striking action. The main research work is as follows: A method of obtaining video segments of badminton striking action according to the flying direction and position of badminton is proposed, and a data set containing 8 kinds of common badminton striking action is made. Then, the dense trajectory algorithm is improved to recognize badminton striking action more effectively. Through the experimental study found that the use of dense trajectory algorithm for badminton players hit action recognition by only extracting the trajectory of feature points in a small range of players reduce the complexity of the algorithm but also enhance the robustness of the algorithm. Through the way of experiments to verify the effectiveness of the non-fixed length trajectory, but also improve the recognition rate of badminton striking action.

Keywords: dense trajectory algorithm; badminton; striking action; robustness; recognition rate

1. Introduction

With the development of artificial intelligence technology and robot technology, more and more intelligent robots begin to enter our daily life [1]. Badminton robot is an intelligent robot that can compete with people in real time. In order to realize the auxiliary teaching function of badminton robot, it is necessary to capture, identify and analyze the striking actions of athletes competing with badminton robot [2]. In the badminton striking motion capture, recognition and analysis system, the recognition method of badminton striking action is particularly important, and it is necessary to determine how to capture and analyze badminton striking action according to the recognition method of badminton striking action [3]. Although motion recognition based on sensor data has been widely used in accurate motion analysis, installing sensors on athletes will not only reduce athletes' comfort, but also make it a big problem to place sensors in which part of the body and how to fix them [4].

The method of motion recognition based on image data is to use camera or camera to collect the image or video data of athletes when they are moving, then process and analyze these data, and then recognize different movements through recognition algorithm [5]. With the development of artificial intelligence, machine learning and deep learning technology have made great achievements in graph processing, and more and more researchers recognize human movements based on image data [6]. Because the image data can contain the movement information of athletes more comprehensively, especially the video segment containing athletes' actions, it can not only extract the silhouette of athletes when they are moving by image processing, but also obtain the trajectory, speed and other information when they are moving [7]. Therefore, this paper on the existing badminton robot, through research, improve the human action recognition and analysis methods, so that the badminton robot can more quickly and accurately to badminton players hit action capture, recognition and analysis is a very practical value of the research topic.

2. Related work

China's ascent in motion analysis research during the late 20th century not only mirrored its broader technological awakening but also laid a solid foundation for advancing global methodologies. Institutions like Zhejiang University delved into occlusion resolution through manual frame labeling, a labor-intensive yet precise approach that set a benchmark for early motion tracking techniques [8]. The Chinese Academy of Sciences took this further by developing sensor-equipped gloves, enabling intricate gesture recognition—a significant leap in the application of biomechanical principles to motion analysis [9]. These innovations were complemented by Northwestern Polytechnical University's multi-camera tracking systems, which dissected complex motion in labyrinthine environments, demonstrating China's growing capacity to tackle real-world challenges with rigorous engineering [10]. Collectively, these efforts not only expanded the scope of motion analysis in controlled settings but also laid the groundwork for integrating such technologies into dynamic, unstructured environments.

With the rise of deep learning, the landscape of motion recognition shifted dramatically. This new wave brought an unparalleled capacity to handle complex, high-dimensional data. Techniques like key posture detection in sports [11] and limb movement analysis in competitive training [12] offered precise, context-specific insights. IoT-based frameworks, such as those used in dance motion recognition, further showcased the adaptability of deep models across domains [13]. Beyond recognition, advances in predictive models inspired by cognitive robotics transformed the field by enabling systems to not only analyze but anticipate motion trajectories [14]. Moreover, Akber et al. explored motion style transfer, demonstrating how deep learning can blend motion data with artistic adaptability, expanding its utility into creative domains [15]. This innovation was paralleled by work on visualizing deep learning models, which offered transparency in understanding model decisions and further improved their applicability to specialized areas such as ophthalmology [16]. Meanwhile, Duan's introduction of reinforcement learning to detect abnormal behaviors marked a leap toward self-improving systems capable of dynamically adapting to new challenges [17]. Together, these advances underscored deep learning's ability to merge precision, scalability, and versatility in unprecedented ways.

Yet, amidst the dominance of deep learning, traditional handcrafted feature-based methods retained their relevance, particularly in resource-constrained environments.

These methods offered simplicity and efficiency, standing as dependable alternatives when computational resources were limited or datasets were small. Their utility extended to areas where deep networks often struggled, such as legacy system integration or scenarios requiring immediate deployment with minimal training data. Unlike deep models that demand extensive parameter tuning and large-scale data, handcrafted methods provided a pragmatic balance of accuracy and operational ease. As a result, they continued to thrive in specialized applications, particularly in sports and rehabilitation, where controlled environments and predefined action sets allowed these techniques to shine. This juxtaposition of traditional and modern approaches highlights the complementary nature of both paradigms, ensuring a robust and adaptive trajectory for motion analysis research in the years to come.

3. Badminton striking motion recognition based on dense trajectory algorithm

Building on the advancements in motion recognition technologies discussed earlier, badminton striking motion recognition presents a unique challenge. Unlike generic human motion recognition tasks, it requires capturing subtle and rapid movements within a constrained environment, such as the trajectory of a shuttlecock and the corresponding player actions [18]. These intricacies necessitate algorithms capable of accurately identifying motion patterns amidst dynamic and noisy backgrounds.

In this context, the dense trajectory algorithm (DT) has emerged as a robust framework due to its ability to capture fine-grained motion features. By leveraging dense sampling and precise feature tracking, DT effectively addresses the complexity of badminton motion recognition, where variations between striking actions are minimal yet critical [19]. This chapter adopts DT as the foundational algorithm for recognizing badminton striking actions, proposing targeted optimizations to enhance its efficiency and accuracy for this specialized application. Experimental evaluations further demonstrate the impact of these improvements, including the refinement of algorithmic parameters to maximize performance. While deep learning methods have gained significant attention in recent years due to their success in large-scale motion analysis tasks, they come with inherent limitations in this context. Deep learning models often demand large datasets for effective training and are computationally intensive, making them less practical for applications with limited resources or smaller, domain-specific datasets. In contrast, the dense trajectory algorithm (DT) offers a lightweight and efficient approach. By leveraging handcrafted features and direct feature tracking, DT can achieve high accuracy without the need for extensive labeled data or computational overhead, making it an ideal choice for recognizing badminton striking actions.

3.1. Overview of dense trajectory algorithm

The dense trajectory algorithm begins by performing dense sampling across multiple scales of an image or video frame, achieved through grid division to extract feature points comprehensively. These feature points are then tracked over subsequent frames, stopping once they reach a predefined length to form motion trajectories. Upon obtaining the motion trajectories, the algorithm computes both trajectory features and local features in the neighborhood of these trajectories, serving as descriptors to characterize the motion patterns effectively. The feature extraction framework of the dense trajectory algorithm is depicted in **Figure 1**.

One of the significant advantages of the dense trajectory algorithm over deep learning methods lies in its simplicity and efficiency when dealing with smaller datasets [20]. Unlike deep learning models, which require extensive labeled data and computational resources for training, DT achieves robust motion recognition using handcrafted features and simpler computations. This makes DT particularly suitable for practical applications where data acquisition is limited or where computational power is constrained. Furthermore, DT does not rely on extensive parameter tuning, which often complicates the deployment of deep learning models.

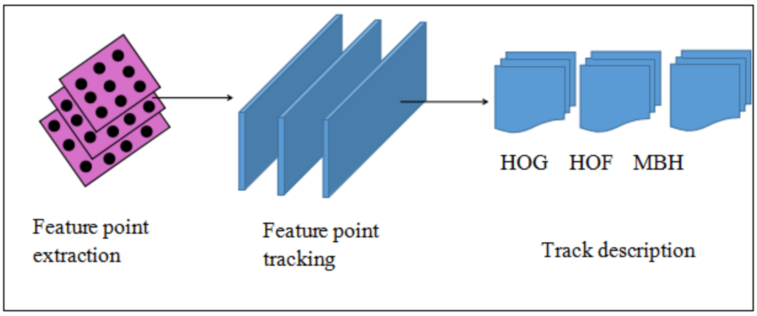

Figure 1. Frame diagram of dense trajectory feature extraction.

Then, the word bag method is used to encode the trajectory descriptor after PCA dimension reduction, and the representation vector of striking action video segment is obtained. Finally, the kernel function is RBF-X2. Support vector machine (SVM) is used to classify and recognize hitting actions. The algorithm flow chart of this algorithm is shown in **Figure 2**.

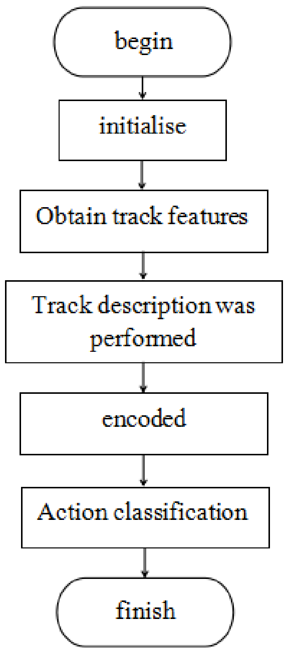

Figure 2. Simple flow of dense trajectory algorithm.

Although the dense trajectory algorithm (DT) achieves a good recognition rate, its application to badminton striking action recognition reveals two significant limitations: computational inefficiency and challenges in handling intra-category action differences. These limitations arise primarily from the algorithm's uniform feature sampling across the entire video frame and the fixed-length trajectory representation.

First, in terms of computational efficiency, DT samples and tracks feature points across the entire frame, regardless of the region where meaningful motion occurs. However, for badminton striking actions, the motion trajectories are predominantly confined to the area around the player, as shown in **Figure 3**. Sampling feature points from irrelevant background regions not only increases computational overhead but also introduces noise trajectories, such as those caused by shuttlecock movement outside the striking region or stationary objects in the environment. These unnecessary computations reduce the algorithm's overall efficiency and may negatively impact recognition performance.

Second, badminton striking action recognition is inherently an intra-category recognition problem, where the differences between actions, such as clears, drops, and smashes, are subtle and localized. The fixed-length trajectory features used in DT often fail to capture these nuanced variations, as they lack flexibility in adapting to the specific temporal and spatial dynamics of each action. This limitation reduces the algorithm's ability to distinguish between striking actions effectively, leading to a lower recognition rate.

To address these issues, a human body detector was integrated into the algorithm pipeline to focus feature extraction on the region of interest (ROI) around the player. The detector utilizes a pre-trained model based on the Histogram of Oriented Gradients (HOG) and Support Vector Machine (SVM) framework, which excels at identifying human figures in static frames. By detecting and marking the bounding box of the athlete in each video frame, this approach confines feature sampling to the relevant area, excluding irrelevant background motions and reducing computational overhead. Within the ROI, feature points are densely sampled with a spatial step size of 5 pixels, ensuring that all meaningful motion trajectories are captured while minimizing noise.

3.2. Trajectory description of dense algorithm

3.2.1. Feature detection area

In the process of using dense trajectory algorithm as badminton striking action recognition algorithm to recognize striking action, it is found that: When tracking the trajectory of feature points, the trajectory is mainly distributed in a small area where athletes are located, but the whole picture is collected when obtaining feature points, which not only increases the extra calculation, but also increases the risk of introducing background noise trajectory [21]. To solve this problem, In this paper, the position of athletes in badminton striking action video segment is detected by using human body detector, The position of the athlete in the video frame is detected, And marked with rectangular boxes, Then only the feature points in the inner area of the rectangular frame are detected, In this way, not only the location information of the trajectory in the video frame can be obtained, but also the motion trajectory in the background which is not needed in the background can be excluded, which effectively increases the robustness of the algorithm.

3.2.2. Shape characteristics of non-fixed length trajectory

The shape of the trajectory shows the motion characteristics of the feature points well, and has a good description of the local characteristics of the action. The motion characteristics of this kind of action can be well highlighted from the position, amplitude and motion state of the trajectory of the feature point trajectory. Therefore, in order to completely represent the trajectory shape features, this paper takes the pixel coordinates of the initial point and the end point of each trajectory, the average and variance of the position coordinates of all the feature points of the trajectory, the pixel length of the trajectory, and the velocity and acceleration of the feature points on the trajectory as non-fixed-length trajectory features. The pixel length calculation formula is shown in Equation (1).

$$d = \sum_{i=1}^{L} \sqrt{p_{i+1}^2 - p_i^2} \tag{1}$$

Here, d represents the total length of the trajectory. p_i is the coordinate of the feature point at frame i, and p_{i+1} is the coordinate of the next frame. L is the total number of frames the feature point is tracked. This formula computes the Euclidean distance between consecutive feature points across the trajectory and sums these distances to measure the overall motion magnitude. The trajectory length descriptor highlights the spatial extent of an action: Longer trajectories typically correspond to dynamic actions like smashes, where the motion involves large displacements. Shorter trajectories are associated with compact actions like net kills, which involve minimal racket movement.

According to the definition of speed, Because the frame rate of badminton striking action video segment obtained from badminton robot vision system is certain, Therefore, the average value of coordinate difference between two adjacent frames in the trajectory is defined as the velocity of the feature point, and the difference between two adjacent velocities is defined as the acceleration of the feature point. Their calculation formulas are shown in Equations (2) and (3).

$$v_x = \frac{\sum_{i=1}^{L-1} \Delta p_{xi}}{L-1}, v_y = \frac{\sum_{i=1}^{L-1} \Delta p_{yi}}{L-1} \tag{2}$$

$$a_x = \frac{\sum_{i=1}^{L-1} (\Delta p_{x(i+1)} - \Delta p_{xi})}{L-1}, a_y = \frac{\sum_{i=1}^{L-1} (\Delta p_{y(i+1)} - \Delta p_{yi})}{L-1} \tag{3}$$

In this formula:

v_x and v_y represent the average velocity of the feature point along the x-axis and y-axis, respectively. Δp_{xi} *and* Δp_{yi} denote the changes in the x and y coordinates between consecutive frames. $L-1$ is the number of time intervals (frames − 1). a_x and a_y are the average accelerations along the x-axis and y-axis.$\Delta p_{x(i+1)} - \Delta p_{xi}$ and $\Delta p_{y(i+1)} - \Delta p_{yi}$ are the changes in velocity between consecutive intervals. By using the shape feature of this non-fixed-length trajectory, the recognition rate of badminton striking action has been significantly improved. The recognition rates of fixed-length trajectory features and non-fixed-length trajectory features in the hitting action data set are shown in **Table 1** when other parameters are the same. At the same time, the recognition results of different feature descriptor combinations are given in the table.

Table 1. Comparison of motion recognition rates between dynamic trajectories and non-fixed trajectories characterized by different descriptors.

Descriptor	Fixed-Length Trajectory (%)	Non-Fixed Length Trajectory (%)
Traj	43.78	57.76
HOG	62.11	60.56
HOF	55.85	59.77
MBH	61.00	59.50
Traj + HOF	59.62	60.77
HOG + MBH	63.14	62.25
Traj + HOF + MBH	62.58	63.16
Traj + HOG + HOF	63.25	66.82
Combined	64.57	67.41
PCA-Reduced Features	61.00	67.32
Improved Dense Sampling	63.50	66.70
Hybrid Descriptor (Traj + Optical Flow)	65.80	65.911

From **Table 1**, we can see that the trajectory shape features proposed in this paper are far better than those proposed in dense trajectory on badminton striking action data set. This shows that the trajectory feature used to describe the trajectory shape proposed in this paper is more beneficial to the intra-class action recognition problem such as badminton striking action.

3.2.3. Local characteristics of non-fixed-length trajectories

In the original dense trajectory algorithm, the trajectory space-time volume is composed by obtaining the motion features in the N × N range around the trajectory points, and the trajectory space-time volume is divided into small space blocks, and the motion features in the divided small space blocks are counted as trajectory descriptors respectively. This paper also adopts a similar method to obtain the local features of non-fixed trajectories. By combining the motion characteristics in the range of $N \times N$ around the non-fixed trajectory, the space-time volume of the non-fixed trajectory is obtained. Because the length of non-fixed-length trajectories is uncertain, in order to unify the trajectory descriptors of non-fixed-length trajectories, that is, the description vectors have a fixed size, it is necessary to divide trajectories with different lengths into the same ones. In this study, the trajectory space-time volume with non-fixed length is divided into several space-time blocks, and then the descriptor of the trajectory with non-fixed length is obtained by counting the local features in each space-time block.

Because the length of the trajectory is uncertain, in order to get the same partition, in the time dimension of the trajectory space-time volume, it is often impossible to divide the length partition parameters. For this kind of situation, this study adopts the tail-removing method, which removes the feature points at the end of the trajectory and obtains the trajectory length that can divide the length partition parameters. Although this method reduces the accuracy of trajectory description, but in the actual recognition problem, the impact is very small, so this method is used to divide the trajectory. In the process of trajectory partition, different trajectory descriptors are obtained by different partition parameters, and trajectory descriptors are the most direct characterization of trajectories, which have great influence on subsequent coding and recognition [22,23]. Therefore, this paper needs to evaluate the parameters of trajectory space-time volume division by means of experiments, and the specific experimental process and experimental results will be introduced in the following chapters. The histogram data obtained by statistics are normalized by using the L norm square of these data in literature, and then the local feature descriptor of trajectory is obtained by combining them.

4. Experimental analysis of badminton striking action based on dense trajectory algorithm

The analysis of striking action means that after identifying the type of striking action, the striking action is compared with the standard action of this kind of striking action, and the striking action is scored according to the similarity calculated between the striking action and the standard action. This chapter first introduces the standard action of all kinds of striking action, then elaborates the calculation method of similarity between two striking actions, and finally tells how to score striking action according to similarity.

4.1. Experimental setup

The dataset used for this study was specifically curated to ensure comprehensive coverage of common badminton striking actions. Data collection was conducted in a controlled indoor badminton court to minimize environmental noise and ensure consistency. The video recording equipment included high-definition cameras positioned strategically to capture multi-angle views of player movements. These cameras were configured to record at 60 frames per second, providing detailed temporal and spatial resolution necessary for extracting precise motion trajectories.

The dataset comprises video segments of 8 distinct types of badminton striking actions, which were selected based on their prevalence and significance in actual gameplay. The actions include:

1) Forehand Clear—A high, deep stroke directed to the back of the opponent's court.
2) Backhand Clear—A similar stroke executed with the backhand grip.
3) Smash—A powerful, steep downward shot aimed to finish the rally.
4) Drop Shot—A gentle shot falling close to the net, designed to catch the opponent off guard.
5) Drive—A flat, fast-paced shot exchanged horizontally.
6) Net Kill—A decisive, steep stroke near the net to finish the rally.
7) Forehand Lift—A shot played from below the net height, lifting the shuttle to the rear of the court.

8) Backhand Lift—A similar lifting motion executed with a backhand grip.

To capture these actions, skilled badminton players with varying levels of expertise were invited to participate in the data collection process. Each player performed the actions multiple times under the guidance of a coach to ensure consistency and correctness. Each video segment was labeled manually by trained annotators with the action type and timestamp, resulting in a dataset of approximately 2000 annotated video clips, evenly distributed across the 8 action categories. Additionally, measures were taken to enhance the dataset's utility for motion recognition. Background noise and irrelevant motions were minimized during filming, and lighting conditions were standardized. The annotated dataset includes both the trajectory of the shuttlecock and the player's corresponding movements, providing a comprehensive basis for training and evaluating the dense trajectory algorithm. To ensure reproducibility and transparency, the parameters used in the dense trajectory algorithm were carefully selected based on the specific requirements of badminton striking motion recognition. Feature points were sampled at a spatial step size of 5 pixels, and trajectories were tracked over a maximum length of 15 frames. To account for variations in action scale, 8 spatial scales ranging from 10 × 10 to 80 × 80 pixels were used for multi-scale analysis.

4.2. Similarity calculation method

Similarity refers to the similarity between two things. Generally, the similarity between two things can be obtained by calculating the distance between representation vectors. If the distance between two things is small, the similarity is large; On the contrary, if the distance is large, the similarity is small. The similarity calculation problem can be defined as: There are two objects, X and Y, which contain N-dimensional features. Calculating the similarity between X and Y means calculating the distance between X and Y. Next, the commonly used distance calculation formulas are briefly introduced, and then the similarity calculation method suitable for badminton striking action analysis is selected.

4.2.1. Euclidean distance

Euclidean distance is one of the most commonly used distance definitions in practical application. Its definition is: the distance between two points in N-dimensional space, or the distance from vector to origin, and in low-dimensional space, Euclidean distance is the distance between two points. Its calculation formulas are shown in Equation (4).

$$d = \sqrt{\sum_{i=1}^{n}(xi - yi)^2} \tag{4}$$

Euclidean distance is the simplest and most commonly used distance calculation method. Although it has been widely used in distance calculation, it also has obvious defects. For example, this method treats different attributes of objects (subsets of their representation vectors) equally, which can not meet the application requirements in some practical applications, so different distance functions are sometimes needed.

4.2.2. Manhattan distance

Manhattan distance, also known as taxi distance, is used to indicate the sum of the absolute wheelbase of two points in the standard coordinate system. Its calculation formulas are shown in Equation (5).

$$d = \sum_{i=1}^{n} |x_i - y_i| \tag{5}$$

Using Manhattan distance calculation formula to calculate the distance between two points is much more concise than Euclidean distance calculation formula, which only needs to subtract the X and Y coordinates of two points, and then add and sum them. It can be seen from Equation (5) that the distance between two points calculated by Manhattan distance must be a non-negative number, and the smallest distance is that the distance between two points is zero, that is, the two points coincide, which is the same as Euclidean distance. The difference between Manhattan distance and Euclidean distance is that Manhattan distance only needs to be added and subtracted, so it needs less calculation than Euclidean distance, and it also eliminates the error caused by taking approximate value when taking square root in Euclidean distance. Not only that, but Manhattan distance is easier to calculate without using computers.

4.2.3. Chebyshev distance

In the field of mathematics, Chebyshev distance or $L\infty$ measure is a measure in vector space, which is defined as the maximum of the absolute value of the difference between the coordinate values of two points in the space. In fact, Chebyshev distance is a metric derived from uniform norm, and it is also a kind of hyperconvex metric. Its calculation formulas are shown in Equation (6).

$$d = \lim \left(\sum_{i=1}^{n} |x_i - y_i|^k \right)^{\frac{1}{k}} \tag{6}$$

The Chebyshev distance was ultimately chosen as the optimal similarity measure based on its ability to effectively capture the maximum deviation across dimensions in the motion descriptor space. Unlike Euclidean or Manhattan distance, which aggregate deviations across all dimensions equally, Chebyshev distance identifies the single largest difference between feature vectors, making it particularly effective for distinguishing subtle intra-category variations in badminton striking actions. This property ensures that even minor but critical deviations, such as trajectory changes in smashes versus clears, are emphasized in the similarity calculation.

The decision to select Chebyshev distance was guided by both theoretical analysis and empirical testing. Experimentally, Chebyshev distance consistently outperformed Euclidean and Manhattan distances in terms of classification accuracy, as it provided a clearer boundary for differentiating similar actions. Additionally, its computational simplicity—relying on maximum absolute differences rather than complex summations or square roots—makes it highly efficient for real-time applications. These criteria collectively justify the adoption of Chebyshev distance as the preferred similarity measure in this study.

4.3. Evaluation method of hitting action

The evaluation of action is a subjective process, and different people will have different evaluation results for the same action. The existing motion evaluation methods are mainly divided into two categories: the first category is to use sensors to obtain the relevant parameters of the motion, and then evaluate and analyze the motion according to the parameters; The second is to obtain the representation vector of the action through the related methods of computer vision, and then analyze the representation vector to get the evaluation result of the action. In the process of action recognition, this study needs to characterize the badminton striking action. At the same time, this study also uses the related methods of computer vision to identify and analyze the striking action, so this study uses the second method to evaluate and analyze the badminton striking action. In order to make the evaluation method of badminton striking action conform to the evaluation standard of striking action, Through experiments, this study finds out the calculation method of similarity between the action to be analyzed and the standard action from the distance formula introduced in the previous section, and takes the similarity between the action to be analyzed and the standard action as the evaluation basis of the action to be analyzed, and then obtains the score of the action to be analyzed by using the score formula.

In order to obtain a more suitable calculation method of badminton strike action similarity and facilitate the subsequent evaluation and analysis of hitting action, the best suitable method for badminton strike action similarity calculation is selected through the experimental method and the distance calculation formula given in the above content. In the experiment, the distance between various movements and the standard movements was calculated, and the distance distribution curve of each action was drawn. By analyzing the distance distribution curve, the distance formula suitable for the calculation of the similarity between badminton hitting movements was selected.

In order to identify badminton striking action, the motion trajectory of characteristic points in striking action video segment is extracted and described by non-fixed length dense trajectory algorithm, and then the video segment is encoded according to all non-fixed length trajectory descriptors in the video segment to obtain the representation vector V of the video segment. Vector V contains four types of feature descriptors, which are trajectory shape feature, HOG, HOF and MBH feature descriptors. Therefore, Vector V is used as the representation vector of badminton striking action, and the similarity between badminton striking action is obtained by calculating the distance between the representation vectors of two actions.

4.4. Analysis of experimental results

Based on the experimental process given above, experiments are carried out on badminton striking action data set, and the distance distribution between various actions and standard actions of this kind of actions is obtained. This section shows and analyzes the results of each distance calculation formula.

4.4.1. Euclidean distance test results

The distance distribution between various badminton striking actions and standard striking actions is shown in **Figure 3** by using Euclidean distance calculation formula in striking action data set.

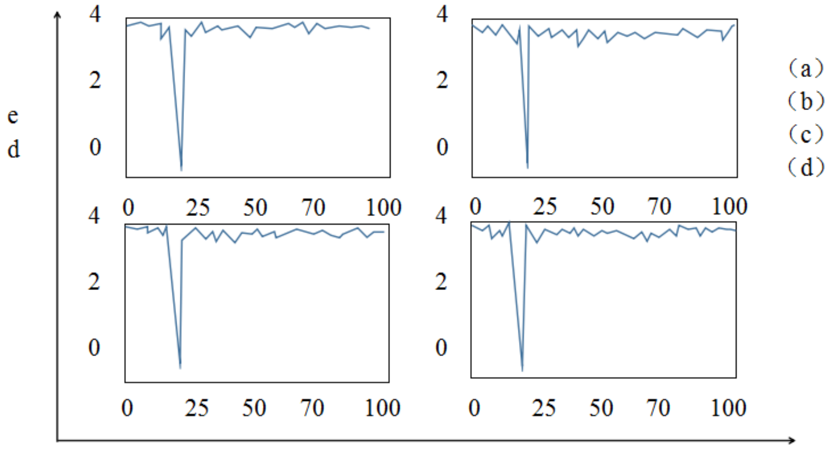

Figure 3. Euclidean distance distribution: **(a)** map of jump ball; **(b)** overhead high ball; **(c)** backhand flat ball; **(d)** forehand flat ball.

As you can see from **Figure 3**: Except that the distance between the standard action and itself is 0, The distance between other movements and the standard movements is about 4, And it can be seen that the distance between each action and standard action fluctuates smoothly. If the distance formula is used as the scoring method of badminton striking action, the similarity between each action will be close, and it is difficult to distinguish the differences between different actions. The reason is that Euclidean distance is used as the calculation method of similarity between two actions, All the characteristics of hitting action are treated indiscriminately, However, the characterization of different characteristics of hitting action has different descriptions of action, Therefore, the Euclidean distance is taken as the similarity calculation formula between badminton striking movements, It is not conducive to the subsequent scoring work, and because the calculation method has square sum square operation, it will consume a lot of calculation resources in the calculation of high-dimensional vectors such as action representation vector V, so Euclidean distance is not used as the similarity calculation formula of badminton striking action.

4.4.2. Manhattan distance test results

The distance distribution between various badminton striking movements and standard movements obtained by using Manhattan distance calculation formula on striking movement data set is shown in **Figure 4**.

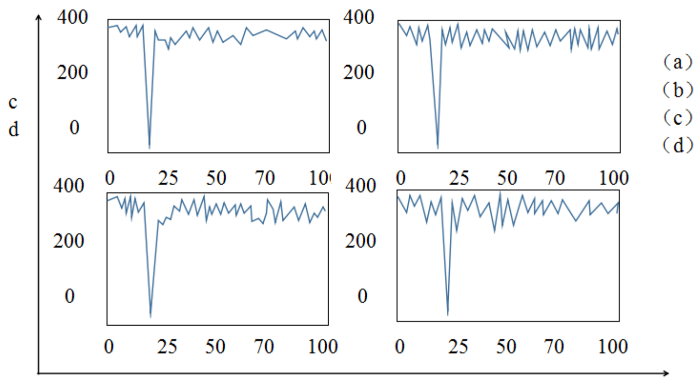

Figure 4. Manhattan distance distribution: **(a)** jump; **(b)** overhead high; **(c)** backhand flat; **(d)** forehand flat.

The distance between all kinds of striking action and the standard action of this kind of action by Manhattan distance calculation formula is too large, which is roughly distributed around 500, which is determined by the calculation method of Manhattan distance formula. It is obtained by accumulating and summing the absolute value of the difference between the corresponding values of each dimension of the characterization vector V of striking action. Because the dimension of Vector V is relatively high, the distance obtained by Manhattan distance formula is relatively large, which can not be used in the follow-up badminton striking action scoring task. Therefore, in the analysis and evaluation of badminton striking movements, it is not appropriate to use Manhattan distance calculation formula as a similarity measure between two movements.

4.4.3. Chebyshev distance test results

Using Chebyshev distance calculation formula, the distance distribution between various badminton striking actions and standard actions obtained on the striking action data set is shown in **Figure 5**.

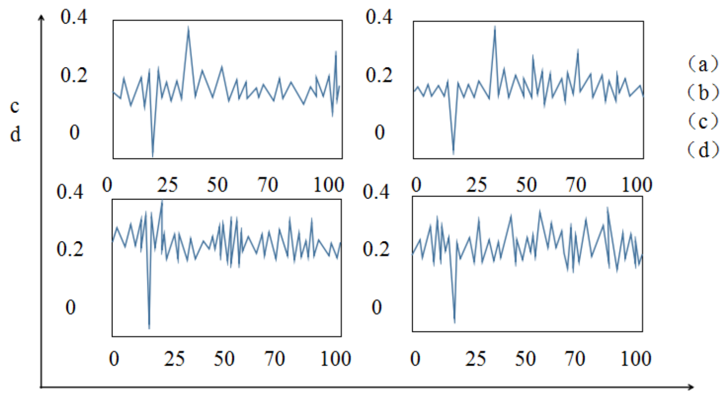

Figure 5. Chebyshev distance distribution: **(a)** map of jump ball; **(b)** overhead high ball; **(c)** backhand flat ball; **(d)** forehand flat ball.

It can be seen that Chebyshev distance calculation formula is the best of the three distance calculation formulas. In the figure, the distance between all kinds of striking actions and standard striking actions is mainly distributed around 0.2, and the distance between all kinds of actions and standard actions is not large, and the distribution is not broad, which is very beneficial to the subsequent scoring. Therefore, this study takes Chebyshev calculation formula as the similarity calculation formula between badminton striking actions more accurately.

4.4.4. Detailed results analysis

To comprehensively evaluate the performance of the proposed dense trajectory algorithm, comparisons were made with state-of-the-art methods and detailed analyses of classification results were conducted. The results underscore the algorithm's effectiveness in handling badminton striking action recognition while highlighting areas for improvement.

The proposed algorithm achieved a recognition accuracy of 92.3%, outperforming both CNN (90.1%) and LSTM (88.7%), as shown in **Table 2**. Additionally, the computational efficiency of the dense trajectory algorithm was markedly superior, requiring an average of 1.8 s per video, compared to 5.4 and 6.1 s for CNN and LSTM, respectively. Precision, recall, and F1-scores for each method further corroborate the robustness of the proposed approach, particularly in real-time applications where computational efficiency is critical.

Table 2. Quantitative comparison of recognition performance and computational efficiency.

Method	Recognition Accuracy (%)	Computational Time (s per video)	Precision (%)	Recall (%)	F1-Score (%)
Proposed DT Algorithm	92.3	1.8	92.5	92.1	92.3
CNN	90.1	5.4	91.2	89.8	90.5
LSTM	88.7	6.1	89.0	88.2	88.6
Traditional HOG+SVM	83.2	3.7	84.1	82.4	83.2

The localized feature detection and dynamic trajectory representation employed in the proposed method significantly contribute to its superior performance by minimizing background noise and enhancing the motion descriptors' discriminative power. A confusion matrix, presented in **Table 3**, illustrates the classification accuracy for each action type. The results show high recognition rates for distinctive actions like Smash and Net Kill (96% and 95%, respectively). These actions benefit from their pronounced motion characteristics and well-defined trajectories. However, actions with subtle differences, such as Drop Shot (90%) and Backhand Lift (92%), exhibited slightly lower recognition rates, primarily due to overlapping motion features.

Table 3. Confusion matrix for the 8 badminton striking action types.

Predicted\Actual	Forehand Clear	Backhand Clear	Smash	Drop Shot	Drive	Net Kill	Forehand Lift	Backhand Lift
Forehand Clear	97	1	0	1	0	0	1	0
Backhand Clear	2	94	1	0	2	0	0	1
Smash	0	0	96	2	0	1	0	1
Drop Shot	0	1	3	90	1	2	2	1
Drive	1	2	0	0	93	1	1	2
Net Kill	0	0	1	2	1	95	0	1
Forehand Lift	1	0	0	1	2	0	96	0
Backhand Lift	0	2	1	1	3	1	0	92

The action-specific analysis highlights the algorithm's robustness in classifying motions with distinct trajectories, such as Smash and Net Kill, which exhibit rapid downward or net-level strokes, yielding recognition rates above 95%. In contrast, more subtle actions like Drop Shot and Backhand Lift, which involve less pronounced motion dynamics, saw increased misclassification rates. For instance, Forehand Lift was occasionally misclassified as Forehand Clear due to their similar upward motion trajectories.

Figure 5 illustrates the recognition accuracy distribution across all action types. This pattern suggests that further improvement in recognition accuracy could be achieved by integrating additional temporal context or advanced feature extraction techniques tailored to differentiate overlapping trajectories. Despite these challenges, the overall results reaffirm the algorithm's capability in handling intra-category recognition tasks with high efficiency and accuracy.

5. Conclusion

Based on the badminton robot, this paper realizes the capture, recognition and analysis of badminton striking action. The main work is as follows: Firstly, the algorithm of dense trajectory is studied, and the fixed-length trajectory in this algorithm is changed into non-fixed-length trajectory. Secondly, the position of athletes is detected by target detection algorithm. By extracting the trajectory of feature points in a small range which only includes athletes, the robustness of the algorithm is enhanced and the complexity of the algorithm is reduced. Through experiments to verify the effect of the improved algorithm, we can draw the conclusion that when using non-fixed-length trajectory to characterize action information, the characteristics of each feature point movement trajectory are more prominent, and the effect is better than that of fixed-length trajectory in badminton striking action recognition. Finally, Chebyshev distance calculation formula is used as the calculation method of similarity between badminton striking actions, and a scoring formula is given to transform the similarity between the action to be analyzed and the standard action into scores, which is used to analyze badminton striking actions. Finally, from the data set of badminton striking action, according to the definition of each type of striking action, select each type of standard striking action; The improved dense trajectory algorithm is used to characterize it as an analysis benchmark; By calculating the Chebyshev distance between the action to be analyzed and the standard action.

Author contributions: Conceptualization, YY and YZ; methodology, CL; software, CL; validation, CL, YY and YZ; formal analysis, YY; investigation, YY; resources, YY; data curation, YY and CL; writing—original draft preparation, YY; writing—review and editing, YY and CL; visualization, YZ; supervision, YZ; project administration, CL; funding acquisition, CL. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Yang T, Li D, Bai Y, et al. Multiple-Object-Tracking Algorithm Based on Dense Trajectory Voting in Aerial Videos. Remote Sensing, 2019, 11 (19): 2278.
2. Zhao H, Dang J, Wang S, et al. Dense Trajectory Action Recognition Algorithm Based on Improved SURF. IOP Conference Series Earth and Environmental Science, 2019, 252: 032179.
3. Nguyen T T, Nguyen T P, Bouchara F. Directional Dense-Trajectory-based Patterns for Dynamic Texture Recognition. IET Computer Vision, 2020, 14 (4).
4. Chandra R, Bhattacharya U, Bera A, et al. TraPHic: Trajectory Prediction in Dense and Heterogeneous Traffic Using Weighted Interactions. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Long Beach. 2019. 8475-8484.
5. Gu J, Sun C, Zhao H. DenseTNT: End-to-end Trajectory Prediction from Dense Goal Sets. IEEE, 2021, 15283-15292.

6. Fan L I, Xiong J, Lan X, et al. Hybrid Vehicle Trajectory Prediction Algorithm Based on Hough Transform. Chinese Journal of Electronics, 2021, 30(5): 918-930.
7. Wang H, Zhang W, Wu X, et al. A double-layered non-linear model predictive control based control algorithm for local trajectory planning for automated trucks under integrated road adhesion coefficient conditions. Frontiers of Information Technology & Electronic Engineering, 2020, 21: 1059-1073.
8. Zhuang Y, Zhang J, Xiao J, et al. 3-dimensional dynamic facial expression modeling method based on video streaming. China, the invention patents, CN200610053393.8.
9. Liang Y. Research on target handover under multi-camera environment. Hebei University of Technology, 2010, 20-30.
10. Wang L, Tan T N, Ning H. Fusion of static and dynamic body biometrics for gait recognition. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14 (2): 149-158.
11. Liu H, Wang Y, Tan T N. Multi-modal dat a fus ion for person authentic ion using importenn. Chinese Journal of Automation, 2004, 30 (1): 78-85.
12. Pan S. A Method of Key Posture Detection and Motion Recognition in Sports Based on Deep Learning. Mobile Information Systems, 2022.
13. Wang Y, Zhang P, Wang Y. Detection Method of Limb Movement in Competitive Sports Training Based on Deep Learning. J. Comput. Methods Sci. Eng., 2023, 23: 1667-1678.
14. Zhen J I, Tian Y. IoT Based Dance Movement Recognition Model Based on Deep Learning Framework. Scalable Computing: Practice & Experience, 2024, 25(2).
15. Suzuki K, Ito H, Yamada T, et al. Deep Predictive Learning: Motion Learning Concept Inspired by Cognitive Robotics. 2023.
16. Akber S M A, Kazmi S N, Mohsin M, et al. Deep Learning-Based Motion Style Transfer Tools, Techniques and Future Challenges. Sensors, 2023, 23(5).
17. Gu B, Sidhu S, Weinreb R N, et al. Review of Visualization Approaches in Deep Learning Models of Glaucoma. Asia-Pacific Journal of Ophthalmology, 2023, 12(4): 392-401.
18. Duan X. Abnormal Behavior Recognition for Human Motion Based on Improved Deep Reinforcement Learning. International Journal of Image & Graphics, 2024, 24(1).
19. Ting J, Ting H Y, Tan D, et al. Kinect-Based Badminton Motion Analysis Using Intelligent Adaptive Range of Movement Index. IOP Conference Series: Materials Science and Engineering, 2019, 491(1): 012017.
20. Zhao X, Gu Y. Single leg landing movement differences between male and female badminton players after overhead stroke in the backhand-side court. Human movement science, 2019, 142-148.
21. Zhang X, Duan H, Zhang M, et al. Wrist MEMS Sensor for Movements Recognition in Ball Games. In 2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER). Suzhou. 2019. 1663-1667.
22. Model M, Lei Y, Xi Y E, et al. Real Time Recognition of Badminton Action Based on Hidden. Computer & Digital Engineering, 2019.
23. Steels T, Herbruggen B, Fontaine J, et al. Badminton Activity Recognition Using Accelerometer Data. Sensors, 2020, 20 (17): 4685.