



Algorithm based Approach for Precise Sports Training Data Extraction without Intervention

Bowen Xiao, Wei Xiao

College of Physical Education, Fuyang Normal University, Fuyang, Anhui. China

XBW289381025@163.com

ABSTRACT

This paper explores a Sports movement trajectory data capturing technique based on the mean shift algorithm to overcome the challenges of collecting information in complex environmental conditions and high speed movements. The method constructs a framework with 51 degrees of freedom and 16 joints to collect and analyze trajectory information during training and performs corresponding dimensionality reduction for better results. To reduce the dependency of the mean shift algorithm on environmental parameters, the probability density function in the gradient iteration estimation algorithm is chosen, and the color information of the target is used as a feature for trajectory data capturing. Experimental results demonstrate that this method effectively captures the joint activities of athletes, making Sports movement training trajectory data more accurate without any external parameter intervention.

Keywords: Sports Training Trajectory, Data Capturing, Dimensionality Reduction, Mean Shift Algorithm, Gradient Iteration

Received: 19 March 2025, Revised 29 May 2025, Accepted 26 June 2025

Copyright: with Authors

1. Introduction

To truly grasp the core of physical games and highly skilled sports, it is vital to examine their dynamic traits. To effectively manage its movement path, we can employ the mean shift algorithm to capture its dynamic data accurately. Methods such as image feature analysis and edge feature detection can also help in identifying crucial information. Utilizing machine vision techniques, including edge contour detection and corner detection, can proficiently capture subtle Sports movements, which are essential for enhancing overall Sports performance [1]. With the ongoing progress of globalization, the need for intensified research in sports science has become increasingly pressing. VSAM (Virtual Storage Access Method) is a sophisticated motion target recognition technology [2] collaboratively developed by Carnegie Mellon University and other researchers, extensively utilized in military applications [3]. Researchers have introduced a knee joint motion capture measurement technique

based on MATLAB, which effectively addresses the limitations of conventional knee joint motion analysis. By programming the pose coordinate data of lower limb markers, it establishes the coordinate systems for the femur and tibia and executes coordinate transformations to compute knee joint extension and assess volunteer walking, significantly lowering measurement costs while improving accuracy and reliability [4]. However, this approach still faces challenges in gathering comprehensive trajectory running information. This paper proposes a novel method for capturing sports training trajectories that employs the mean shift algorithm to transform human motion states into visual representations and utilizes the color of the target object as a feature, thereby achieving effective trajectory capture for sports movements.

2. Related Work

Typically, target tracking involves extracting significant objects from sequential image frames and efficiently following them. This procedure entails initially conducting target detection followed by effective monitoring, which is more intricate and demanding than traditional target detection [5]. Early tracking techniques relied on linear Gaussian processes, where a Gaussian correlation existed between the target's motion equation and the observation equation, resulting in a posterior distribution primarily dictated by their square root matrix and covariance matrix. To better capture this state, Kalman filtering (or more broadly, extended Kalman filtering) can be utilized [6]. While Kalman filters can adeptly capture intricate movements between the camera and the target, reduce noise in images, enhance lighting conditions, and minimize background interference, they may not function optimally in diverse environments, which restricts the tracking efficacy of Kalman filters [7]. In recent years, advancements in technology have led to the widespread application of particle filter techniques across various domains, extending beyond traditional tracking methods to include innovative techniques developed in recent times. Researchers have made notable strides in utilizing particle filters to capture complex human motion. Some have integrated multiple rectangular features from grayscale images with particle filter techniques and employed mixture factor analysis (MFA) technology to construct different target types, achieving commendable tracking performance [8]. By linking high dimensional and low dimensional data, scholars have applied the I condensation method to investigate motion equations and attain efficient dynamic tracking [9]. Through the use of mean error, template tracking techniques, and Mean Shift technology, researchers have conducted in depth learning on objects with complex shapes and utilized the diversity of kernel functions to perform affine transformations, allowing for more accurate object position capturing [10].

3. Sports Movement Trajectory Data Capturing Technique Based on Mean Shift Algorithm

3.1 Individual Modeling for Sports Movement Training

To represent motion data more clearly, we abstract the human model into a skeletal structure consisting of 16 joints, each with its own coordinate system, connecting them. In this skeletal structure, the motion state of child joints can be related to parent joints or adjusted freely based on their characteristics. In contrast, the motion of parent joints is entirely independent of external factors [11, 12]. To reduce the complexity brought by Euler angle representation, we use quaternions to describe the rotation state of each joint and define $m(t)$ as the motion data for each frame of sports training to reduce the occurrence of anomalies in the model.

$$M(t) = (p_1(t), q_1(t), q_2(t), \dots, q_n(t)) \quad (1)$$

In the above formula, M represents the number of joints that rotate during the training process, $p(t)$ refers to the specific position of the joint, $q_i(t) \in S_3$ refers to the orientation of the joint during rotation, and $q_n(t)$ refers to the rotation of the n th joint with respect to the parent joint. Through logarithmic mapping, we can map Quaternion to R_3 space, so we can define training dataset y as:

$$y_i = (p_1(t), \log(q_1(t)), \log(q_2(t)), \dots, \log(q_n(t))) \quad (2)$$

To better understand the motion state, we map Quaternion into R_3 space, and use a linear Time-invariant system to reduce the dimension of the data.

3.2 Track Data Capture

In sports training, due to the fast movement speed of the target and the complex background, this paper proposes a new tracking technique, the mean displacement method, which can effectively capture the target's motion trajectory without being affected by parameters. It can quickly calculate the probability density without being limited by parameters [13,14]. The following is a detailed description of the algorithm's Kernel density estimation:

If $\{X\}$ A, as a set of n points, the kernel function density of X can be estimated as follows: $f(x)$:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (3)$$

In this equation, h denotes the window size, while K signifies a kernel function. We then employ the internal kernels of Epanechnikov and Chebyshev to minimise the average bias between the estimated density and the actual density. Upon identifying the target, we implement the mean shift algorithm to secure it [6]. When the object's position, orientation, and dimensions remain constant, its color does not influence the outcome, making it a crucial characteristic of the image. Hence, we utilize color histograms to depict this scenario and establish a set of N data points in the collection $X = \{X_1, X_2, \dots, X_N\}$, with their probability density function defined as $p(X) = n(x; o, v)$, where o indicates the mean and v signifies the covariance matrix. For elliptical training, initial adjustments to the current frame are necessary, where x denotes the height of the athlete's complete image, o represents its starting point, and T represents the duration of movement. With specified weight values and Nv values, we can accurately pinpoint the target and determine its location. Suppose a particular target is detected in a certain frame image. In that case, we will utilise these weight values and Nv values to contrast them with other specific frames and establish their precise position.

3.3 Feature Point Detection Method

The benefits of DoG are quite evident: firstly, it eliminates the need for extensive computational resources since it only requires a second order differential convolution kernel to capture all data effectively; secondly, it is capable of capturing data across various sizes. Consequently, when we aim to gather information at multiple data sizes, it does not demand significant computational power, as it can retrieve information for each data size separately [15, 16]. Thirdly, its spot detection exhibits greater stability in comparison to conventional point detection methods like DoH (Determinant of Hessian) and Harris, which makes its advantages hard to overlook. DoG features exceptional anti-interference capabilities, and its LoG structure can be nearly fully replicated, resulting in characteristics akin to LoG. The fundamental principle of the DoG algorithm is to construct an image made up of octaves ($\hat{I}$) and establish different levels within these octaves to streamline the convoluted convolution process. Furthermore, the technique of sub pixel downsampling can be employed to extract more information for enhanced

outcomes. The principle of the DoG algorithm can be derived from the Gaussian pyramid framework, where the core concept is to compare images at five Gaussian scales and multiply the comparison values for each level to ultimately compute the final DoG algorithm value, which can be utilized in spot detection algorithms, as illustrated in Figure 1. The second data set is enlarged twofold using the Gaussian scale space images from the first set to achieve more precise results, which applies to each level.

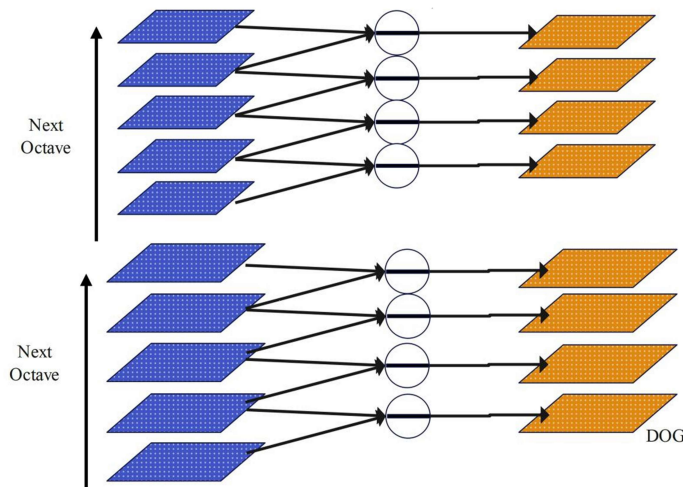


Figure 1. Gaussian Pyramid and DoG Implementation

Considering three dimensional search, the first and last layers of each group cannot be the current level. Therefore, the search process begins with the second layer of each group, which is taken as the current level. It then extracts 3x3 cubes from non border points in the second layer DoG image. In this way, the search results not only contain pose coordinates (in DoG image coordinates) but also include spatial scale coordinates (layer coordinates). After the first search is completed, the third layer is used as the current level, and the process is repeated to find similar points in the third layer. This process continues, and it can be seen that when $S = 3$ and $S_{mn} = 0$, each group generates four layers of DoG, and each group point only undergoes two searches. The above search for extremal points is carried out in the discrete spatial domain, so the number of extremal points detected is not the actual number of extremal points in real situations [17,18]. Figure 1 demonstrates the difference between the number of extremal points obtained in the discrete spatial domain and the continuous spatial domain for a two dimensional function. The method of using known discrete spatial points to interpolate and get the number of extremal points in the continuous spatial domain is called sub pixel interpolation.

4. Experimental Design and Result Analysis

4.1 Experimental Design

Utilizing the Epanechnikov kernel, segmenting each RGB subspace into 16 bins, and merging them into 16x16x16 bins, along with setting the maximum iteration frequency for each frame to 10 times, allows us to evaluate the performance of this approach effectively. The Laser5.0 video capture device is employed to modify the *fmri* frequency to 800KHz to assess if this method can accurately simulate Sports training movements in visual representations. With a pixel resolution of 200x400, a weak transmittance projection of 0.68, and template feature matching of 0.68, we can successfully capture the dynamics of Sports movements within the same simulation context. After numerous enhancements, the automatic Sports shooting trajectory capture technology, based on the background difference method, has seen significant improvement. To confirm the reliability of this

technology and ascertain if it aligns with current standards, we establish an operational environment and conduct further assessments of its accuracy. Ultimately, we will present the precision and dependability of this technology based on a series of testing outcomes. This experiment seeks to gather and analyze trajectory images utilizing cutting edge sensing technology, high definition imaging, and sophisticated software, documenting them in format and high quality AVI format for routine analysis and research. We will also modify the parameters for all participants based on varying environments to guarantee the accuracy and reliability of the experiment. We have selected the Intel(R) Core(TS) i3-2350T as our data processing unit, which boasts robust performance and is equipped with 50GB of operational memory. Our operating platform will be WIN10, and we will develop an application in C++ to analyze our simulated data. Through practical applications, we have determined that our motion trajectory capture technology can precisely represent athletes' movements across two distinct image datasets. We evaluate the accuracy of the detector to measure the reliability of this technology.

4.2 Result Analysis

There is no universal standard for selecting feature points; it should be customized to the specific issue at hand. In the context of object tracking, factors such as tracking accuracy, rotation invariance, illumination invariance, resistance to blur, and time complexity are especially vital. As illustrated in Figure 2, SIFT exhibits the highest tracking accuracy, along with superior rotation invariance, illumination invariance, and blur resistance. SURF also demonstrates good tracking accuracy, rotation invariance, and illumination invariance, along with commendable blur resistance. BRIEF, however, lacks rotation invariance. LAZY shows a moderately high performance across all indicators. ORB and RIFF exhibit relatively lower levels of illumination invariance and blur resistance. Hence, after a thorough comparison, this study selects SIFT as the feature for target tracking

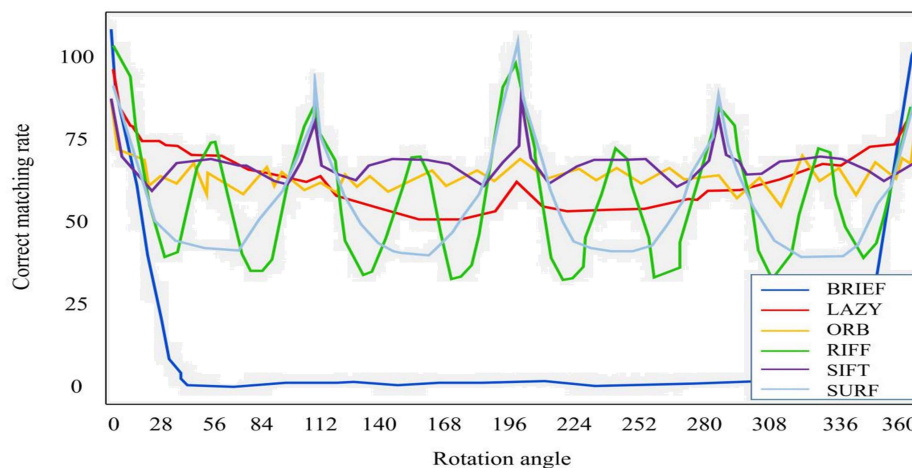


Figure 2. Performance comparison of Sports motion feature points (BRIEF, LAZY, ORB, RIFF, SIFT, SURF)

Next, a fusion improved mean shift tracking algorithm is proposed, which consists of an improved MS tracking algorithm (for initial positioning) and the SIFT tracking algorithm. Based on the results of the improved MS tracking, the SIFT features are extracted, matched, and computed, followed by the calculation of the gravitational offset for feature points. The final tracking result is obtained by integrating the improved MS and SIFT tracking results using a linear weighting method. Sports motion images are used as subjects, and Sports motion training actions are collected and feature extracted in the Gaussian blurred bio-space. The wavelet upgrading technique is used for blur recognition processing of Sports motion training actions, resulting in the training action collection effect shown in Figure 3.

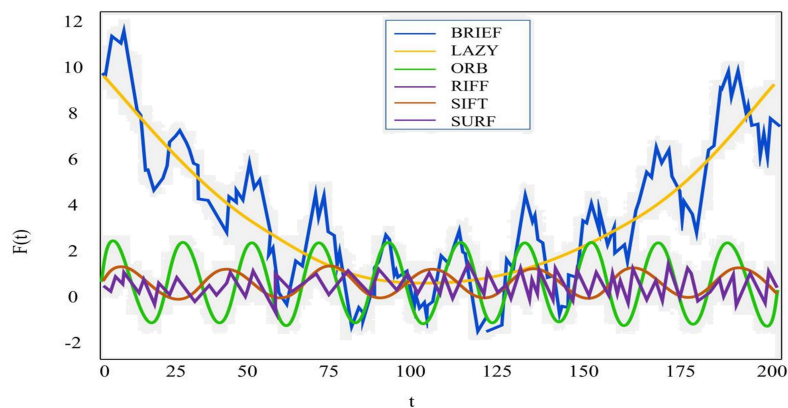


Figure 3. Comparison of detection rate and false alarm rate for different training trajectory data methods

By examining Figure 3, we discovered that this approach demonstrates strong recognition and precision in capturing the dynamic aspects of Sports movements. When articulating this information, the overall trend centers around a significant core area. In contrast, our alternative capture methods exhibit lower clustering efficiency and encounter challenges in conveying this information. By implementing our innovative clustering algorithm, we can significantly enhance the accuracy of tracking Sports shooting paths, thereby preventing data loss associated with this issue. In comparison to conventional estimation capture algorithms, our novel algorithm excels in precision and dependability.

5. Conclusion

While reviewing training footage, we observed that the athletes' body positions may experience substantial changes and could be concealed. For these reasons, we found it quite difficult to track the athletes' walking paths. To tackle this challenge, we employed the mean shift algorithm and leveraged color data for analysis. Our tests revealed that the algorithm achieves high accuracy with a minimal error rate. By utilizing image enhancement methods, we can significantly boost motion capture during Sports training, attaining precision, speed, and reliability. Further more, this technique shows excellent adaptability, effectively improving motion accuracy.

References

- [1] Movafegh, M. S., Dehghan, S. M. M., Zardashti, R. (2021). Three-dimensional guidance and control for ground moving target tracking by a quadrotor. *Aeronautical Journal – New Series*, 1–28.
- [2] Saeed, K., Pouya, B., Hassan, M. (2021). Designing a hierarchical model-predictive controller for tracking an unknown ground moving target using a 6-DOF quad-rotor. *International Journal of Dynamics and Control*, (3), p 9.
- [3] Yang, J., Liu, X., Sun, J., et al. (2022). Sampled-data robust visual servoing control for moving target tracking of an inertially stabilized platform with a measurement delay. *Automatica*, (12), 137.
- [4] Li, B., Jing, Q. (2021). Fast moving and deformational target tracking approach based on heterogeneous features fusion. *Transactions of the Institute of Measurement and Control*, 43(3), 612–622.

- [5] Guo, Y. (2021). Moving target localization in sports image sequence based on optimized particle filter hybrid tracking algorithm. *Complexity*, 2021(7), 1–11.
- [6] Zang, W., Yao, P., Song, D. (2022). Standoff tracking control of underwater glider to moving target. *Applied Mathematical Modelling*, 102, 1–20.
- [7] Xiaoli, W., Weixin, X., Liangqun, L. I. (2022). Labeled Multi-Bernoulli maneuvering target tracking algorithm via TSK iterative regression model. *Chinese Journal of Electronics*, 31(2), 227–239.
- [8] Zhen-Taohu, Shi-Boyang, Yu-Meihu, et al. (2022). Distributed fusion target tracking based on variational Bayes. *Acta Electronica Sinica*, 50(05), 1058–1065.
- [9] Zhang, W., Sun, W. (2021). Research on small moving target detection algorithm based on complex scene. *Journal of Physics: Conference Series*, 1738(1), 03–04.
- [10] Yu, H., Sharma, A., Sharma, P. (2021). Adaptive strategy for sports video moving target detection and tracking technology based on mean shift algorithm. *International Journal of System Assurance*, 30(2), 19–21.
- [11] Yu, H., Sharma, A., Sharma, P. (2021). Adaptive strategy for sports video moving target detection and tracking technology based on mean shift algorithm. *International Journal of System Assurance Engineering and Management*, 1–11.
- [12] Liu, N., Liu, P. (2022). Goaling recognition based on intelligent analysis of real-time sports image of Internet of Things. *The Journal of Supercomputing*, 78(1), 123–143.
- [13] Hu, M. C., Chang, M. H., Wu, J. L., et al. (2010). Robust camera calibration and player tracking in broadcast sports video. *IEEE Transactions on Multimedia*, 13(2), 266–279.
- [14] Yoon, Y., Hwang, H., Choi, Y., et al. (2019). Analyzing sports movements and pass relationships using real-time object tracking techniques based on deep learning. *IEEE Access*, 7, 56564–56576.
- [15] Haritaoglu, I., Flickner, M. (2001). Detection and tracking of shopping groups in stores. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001)* (Vol. 1, pp. I–I).
- [16] Nelikanti, A., Venkata Rami Reddy, G., Karuna, G. (2021). An optimization-based deep LSTM predictive analysis for decision making in cricket. In *Innovative Data Communication Technologies and Application: Proceedings of ICIDCA 2020* (pp. 721–737). Springer Singapore.
- [17] Huang, Y., Liu, X., Zhang, X., et al. (2016). A pointing gesture-based egocentric interaction system: Dataset, approach and application. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 16–23).
- [18] Cardenas, E. J. E., Chavez, G. C. (2020). Multimodal hand gesture recognition combining temporal and pose information based on CNN descriptors and histogram of cumulative magnitudes. *Journal of Visual Communication and Image Representation*, 71, 102772.