ISSN 13303651 (Print), ISSN 18486339 (Online)
https://doi.org/10.17559/TV-20250411002583
Original scientific paper

Research on Object Detection and Tracking in Sports Competitions using Two-Dimensional Fuzzy Semantic Algorithm

Linlin WANG

Abstract: The research object of this paper is the sports game video. The athlete's motion state is clearly displayed in the video, and the target detection and tracking for the specified athlete is realized. A T-S fuzzy semantic interactive maneuvering target tracking algorithm based on two-dimensional fuzzy semantic reduction is proposed. Firstly, a knowledge system is composed of the weights of each model and the parameters of the previous ones, and the redundant models are eliminated by neighborhood fuzzy intensive reduction. At the same time, excessive fuzzy intensive reduction will consume a lot of computing power and lose effective information. An adaptive reduction judgment algorithm is proposed according to each moment and each model residual. Through residual monitoring and tracking, feature re-reduction and feature re-reduction can be adaptive. The experimental results show that the 2D fuzzy semantics can effectively reduce the features and improve the real-time performance of the algorithm. The test results show that the proposed method can effectively improve false alarm rate and tracking accuracy, and has higher robustness than traditional target detection and tracking in real complex competition venues.

Keywords: interactive maneuvering target tracking; reduction judgment algorithm; sports competitions; target detection and tracking; two-dimensional fuzzy semantic algorithm

1 INTRODUCTION

Various consumers have varied demands when it comes to sports game video, thus it is important to analyze and handle game videos accordingly. Competitions in sports tend to garner more coverage in sports videos and a larger audience than other types of sporting events. Consequently, there is a great deal of practical use for methods that can recognize, extract, position, and track moving objects in sporting event films [1]. With the rapid development of computer vision technology, moving object detection has been applied more and more widely in the field of sports training [2]. As the scene of sports training video is more complex and the sports form is more diverse, the traditional sports object detection method is often difficult to achieve the ideal effect [3], and the intelligent analysis of sports training video can not only help coaches accurately evaluate the performance of athletes, but also provide personalized training suggestions for athletes. Thus, it can help athletes improve their training effect more comprehensively and effectively and enhance their competitive level [4].

Target tracking is to find the position movement of the detected object in all sequence images of the video [5]. The method of target tracking has gone through the development process from traditional feature extraction and machine learning to target tracking combined with neural network. Moving target tracking algorithms can be divided into two types: generative pattern type and discriminant pattern type. The kernel function tracking algorithm Mean Shift is jointly proposed, and its main idea is to carry out template matching [6]. The color histogram is obtained by weighted calculation of Mean Shift function. Finally, the color histogram is used as the template to match the moving target [7], but it cannot cope with the change of the target size. Based on the multi-scale theory, Collin proposed a multi-scale image feature completed by Gaussian filter, which uses the local maximum value of the scale space filter to judge which target scale is currently used, and solves the problem that Mean Shift cannot cope with the change of target scale [8]. In the learning model, a method based on online Boosting is proposed [9], and Boosting target tracking method integrates several weak classifiers into a strong classifier to track specified targets, with significant improvement in effect. Thus, the target tracking method using machine learning is widely used. However, in the field of target recognition and tracking in sports competitions, different detection methods adapt to different detection scenes. How to select the correct algorithm for the current scene is a key link to be further studied.

This paper proposes an adaptive feature reduction method combining neighborhood fuzzy sets and two-dimensional fuzzy semantics to enhance the performance of the T-S fuzzy algorithm in athlete target detection and tracking. The method involves: 1) Constructing a dynamically updated hybrid feature set (incorporating multi-model weights and historical parameters) and using a sliding time window to capture system time-varying characteristics; 2) Defining an adaptive neighborhood radius based on feature standard deviation for initial feature reduction using neighborhood fuzzy sets; 3) Iteratively eliminating redundant features via a forward greedy strategy by calculating conditional attribute dependency and setting thresholds; 4) Innovatively introducing a residual-monitoring-based adaptive re-reduction mechanism: tracking model residuals in real-time, building a residual variation index, and dynamically adjusting reduction granularity to balance computational efficiency with information integrity; 5) Finally integrating the two-dimensional fuzzy semantic algorithm to optimize parameter updates. Experiments demonstrate that this method significantly enhances the algorithm's real-time performance while maintaining tracking accuracy.

2 RELATED WORK

The first step of a good maneuvering target tracking system is to reasonably model a highly appropriate mathematical model according to the displacement characteristics and maneuvering characteristics of the maneuvering target. On the basis of this mathematical model of motion, filter is adopted to filter the observed data information to eliminate interference, so as to realize the target's motion state (position, speed, speed, etc. acceleration, etc.) for accurate estimation. In practice, the maneuvering target often changes its original motion state suddenly, that is, the maneuvering target occurs. The maneuvering target tracking model in the target tracking system is to build a reasonable mathematical model of the maneuvering target's motion state, which is a basic and difficult problem in the field of target tracking. After years of development, a variety of preferential motion models have emerged, such as: uniform motion model (CV), uniform acceleration motion model (CA), turning motion model (CT), "current" statistical model [10], Singer model [11], Jerk model [12], etc.

The uniform motion model regards the target as a mathematical modeling model with uniform motion, while the uniformly accelerated motion model regards the target as uniformly accelerated. When the target is not maneuvered or weakly maneuvered, the filter can usually get good results by using these two kinds of filters. In practice, there are few such targets, and the use of these two kinds of motion models for target tracking often produces relatively large errors. Singer motion model was born. Singer motion model is a first-order time-dependent target motion model that considers that the velocity change rate of target motion is consistent with zero mean and obeys the assumption of uniform distribution. In practice, such cases are also relatively rare [13]. The "current" statistical model is based on the Singer model, and the acceleration of the target is counted at different times and described by certain mathematical functions [14]. Based on various modeling methods, the Interactive Multiple Models (IMM) algorithm was usually adaptive and could effectively adjust the probabilities of Interacting models. The interactive multiple model target tracking algorithm consisted of multiple models, each with their own filters and different motion models [15]. In IMM algorithm, various models track the target through interaction, that is, IMM believes that the moving target can be transferred to each other in the motion state described in each model, and the transfer between the models is determined by the Markov probability transfer matrix. Generally speaking, the richer the model library of IMM, the more powerful the description of the model library will be. However, the huge motion model library will inevitably lead to an explosive increase in the computational amount of the algorithm, and two adjacent motion models may cause competition between models in the model probability adjustment stage of IMM algorithm, resulting in the degradation of algorithm performance [16].

In the actual tracking of the target, it is necessary to use various observation tools (sensors, etc.) to observe various state values of the target and obtain various motion state values of the target. In practice, the observed quantity is usually noisy (that is, observed noise), which needs to be filtered by filtering technology to make it close to the true value. The filtering technology in the field of target tracking is often closely related to optimization theory. Haar-like is used as face feature, and this core idea is improved and applied in face recognition [17]. Haar represents the grayscale change of an image, which is regarded as a feature of part of the image and is usually combined with Adaboost strategy for target detection and recognition. The feature of target is represented by the histogram feature of directional gradient, and the gradient mainly exists at the edge [18]. Current target detection methods rely on building features by calculating gradient direction histograms (e.g., HOG) in local image regions, which effectively describe object shape. However, these traditional approaches suffer from significant drawbacks: excessive preparatory work, vulnerability to environmental interference during detection, and susceptibility to errors. This results in low accuracy and slow speed for multi-target detection, ultimately failing to meet the basic requirements of modern video content analysis.

The inherent ambiguity of the real world makes precise object categorization challenging, as many everyday concepts are fundamentally vague and lack clear boundaries. Objects often exist in transitional states between categories rather than binary either-or classifications, representing a continuum of quantitative to qualitative change. Classical set theory cannot adequately model this uncertainty. Higher education institutions have consequently engaged in long-term research to address these recognition and tracking challenges in sports video analysis. For the purpose of analyzing the variations in illumination and light occlusion in the stadium and spectator regions, and to apply tracking methods, a target identification and tracking approach based on optical flow recognition is proposed in literature [19]. Although this approach works well for indoor sporting events, early on there is a lack of clarity in the tracking due to the method's huge calculation and long detection time. Often used for sporting events with basic sceneries, the adjacent-frame difference technique to detect moving targets is introduced in literature [20]. It has a minimal amount of computing, but produces imperfect detection results and easily loses moving target points during tracking. The background difference approach, as described in literature [21], is used in the time domain of sports video. It is effective for basic scenes and works well for detecting and monitoring occurrences in sports. Fuzzy decision tree algorithm is proposed by applying fuzziness to decision tree algorithm [22]. Fuzzy is applied to the traditional ID3 (Iterative Dichotomiser 3) algorithm, and the formula of ID3 algorithm is improved by fuzzy information entropy [23]. Since then, many improved algorithms of fuzzy decision tree algorithm have appeared, such as: Fuzzy decision tree induction method, which reduces classification fuzziness in the tree building process by adding fuzzy evidence, each fuzzy evidence is knowledge about specific attributes, each fuzzy evidence is selected based on its contribution to reducing classification fuzziness, and is more robust in dealing with uncertain knowledge and missing information [24]. Based on the fuzzy ID3 algorithm based on the interaction information between attributes, the test attributes selected by this algorithm not only have larger interaction information with the class, but also have smaller interaction information with the used attributes on the ancestor node [25]. Neuro-fuzzy ID3 algorithm, which applies algebraic learning method to adjust fuzzy rules, considers part of expert knowledge, and prevents the formed tree from being too simple [26]. Multi-fuzzy decision tree induction method based on fuzzy sets, which makes full use of the information provided by each fuzzy attribute reduction, is a soft computing method combining fuzzy sets and fuzzy sets [27]. Soft decision tree method, which combines the tree generation process and pruning process to determine the structure of the soft decision tree, and uses shape modification and trimming to improve the performance of the generated decision tree [28]. Fuzzy induction learning method, which applies fuzzy set theory to the learning process of rule induction, solves the problem that most decision tree induction methods are not ideal when dealing with boundary cases [29]. The automatic two-dimensional K-Means (A2DKM) algorithm is proposed [31], eliminating the need for users to determine the number of clusters. A2DKM combines the local and spatial information of the data into the cluster analysis. Qualitative and quantitative comparisons were made with traditional clustering algorithms. By generating more uniform segmentation results, it outperformed these algorithms. However, unlike static targets, most of the targets in video exist in dynamic form, and their patterns are more complex. Because the video lens will change with the demand, the size of the objects and people in the video will also be scaled, this sudden scaling will also lead to tracking and detection failures. Occlusion is always a difficulty in moving target detection and tracking. It is inevitable that the moving target will pass through some occlusion, which will cause the target to be missed, that is, the tracked target will go through a short disappearing state when tracking the target. How to reasonably deal with the moving target in this short disappearing state is the difficulty in solving the occlusion problem during tracking.

3 RESEARCH METHOD OF OBJECT DETECTION AND TRACKING IN SPORTS COMPETITION BASED ON TWO-DIMENSIONAL FUZZY SEMANTIC ALGORITHM

3.1 Research Methods of Target Detection

Recognizing sporting events is a crucial component of object detection systems designed for use in sporting scenarios. To accomplish human body tracking and motion detection, we integrate a background difference approach with spatio-temporal feature extraction. For more precise object-from-background extraction, we pick one input picture frame every seven frames from the movie. Fig. 1 displays a few video picture samples. If there is no change from one frame to the other, then the point is either not moving at all or is in the backdrop of the video. This being the case, we may denote these background points as $S_j(x, y)$, where $1 \leqslant j \leqslant M$, and the reconstructed backdrop as $B(x, y)$. In the longest static video clip, $M_j(x, y)$ is the middle frame, $ST_j(x, y)$ is the beginning frame, and $EN_j(x, y)$ is the end frame. From this, we may derive:

$$M_j(x,y) = ST_j(x,y) + EN_j(x,y) \tag{1}$$

First, by calculating the difference between the background and the current frame, an image of the human body can be obtained. We set a threshold to select the moving region from the image, and obtain a more accurate image of the moving human body by selecting a suitable threshold value for the differential image and filtering out most of the remaining static background. Then, morphological methods are used to further eliminate the effects of other noises, so that human motion images can be extracted from the video stream.

Figure 1 Examples of sports images

In this paper, an algorithm based on bone detection is used to detect the key points of the human body. In order to realize the recognition of human behavior, a continuous sequence of postures over a period of time is selected to represent human behavior. We define the athlete's behavior as $F = (G1, G2, ..., Gn)$. Gi is a vector that describes the athlete's posture. We establish the human action similarity index (ASIM) to quantify the separation between the reference action template and the test action template, with the goal of identifying distinct human action F. The referenced operation template is defined as $R = R(1), R(2), ..., R(\mathrm{M})$. An action template for testing can be defined as follows: $T=T(1), T(2), ..., T(N)$. When the characteristic dimension of $R(M)$ and $T(N)$ is equal. Similarly, if we consider the cumulative distortion of the eigenvectors $T(n_i)$ and $R(m_i)$ as $D[T(n_i), R(m_i)]$ we get:

$$D[T(n_i), R(m_i)] = [T(n_i), R(m_i)]\, s.t. 1 \tag{2}$$

where: $d[T(n_i), R(m_i)]$ is the distortion degree of the feature vector:

$$\begin{aligned} &D[T(n_i), R(m_i)] = \\ &= d[T(n_i), R(m_i)] + [T(n_{i-1}), R(m_{i-1})] \end{aligned} \tag{3}$$

Therefore, the similarity of sports behavior in this paper can be defined as:

$$d(H_1, H_2) = \sum_{i=1}^{N} \frac{(H_2(i) - H_1(i))^2}{H_1(i) - H_2(i)} \tag{4}$$

Here, H_1 and H_2 stand for the paper's 24-dimensional feature vectors, and iteration allows one to determine the minimal cumulative distortion between the reference action template and the test action template. The test template is considered to be part of the class with the least cumulative distortion if it matches all reference templates sequentially.

3.2 Interactive UKF Target Tracking Algorithm Based on Two-Dimensional Fuzzy Semantics

2D fuzzy semantic modeling establishes the fuzzy rules based on the object's motion properties. Some objects' motion models serve as parameters for these rules, which in turn serve as parameters for the fuzzy division of motion characteristics. These rules have the ability to change and interact with one another, just like the classic interactive multi-model. Different fuzzy rules in two-dimensional fuzzy semantics will self-transform and interact with each other based on the cross-degree fuzzy membership function between them.

Suppose k, b is the set of maneuvering targets that are fuzzy and separated by the GTH motion characteristic at k. The degree of transfer across rules is directly proportional to the degree $c_{k,g}$ to which their fuzzy sets of parameters are comparable to one another. What follows is the definition of the transition probability from fuzzy set i to fuzzy set h when the time is k:

$$p(c_{k,g}=l\mid c_{k-1,g}=h)=\frac{\vartriangle(l,h)}{\sum\limits_{j=1,G}\vartriangle(l,h)} \tag{5}$$

Considering that there are a total of J features to be selected in an object, and each feature is divided into n_G fuzzy sets, assuming that the total number of fuzzy rules is M, then:

$$M=\prod_{G=1}^{j}n_G \tag{6}$$

Then the transition probability π can be calculated by the following formula:

$$\pi_{i,j}=\min\{p(c_{k,g}=l\mid c_{k-1,g}=h)\} \tag{7}$$

Further, the mixed probability and interaction results are calculated according to the following steps: First, the predicted model probability is calculated:

$$u_k^i=\sum_j \pi_{i,j}u_{k-1}^j \tag{8}$$

μ_k represents the mixed probability of converting the total other fuzzy rules to the ith fuzzy rule.

Then, the mixing probability is normalized as follows:

$$u_{k-1}^j=\frac{\pi_{j,i}u^j{}_{k-1}}{u_k^j} \tag{9}$$

Finally, according to the mixing probability, the target state and covariance of the j fuzzy rule are interactively mixed as follows:

$$x_{k-1}^i=\sum_{j=1}^{m}x_{k-1}^j u_{k-1}^j \tag{10}$$

In this article, UKF (Unscented Kalman Filter) is used to verify the subsequent parameters of fuzzy rules. UKF state estimation is implemented as follows:

Initialize and compute sigma points:

$$x_{k-1}^i=\hat{x}_{k-1}+\sqrt{\lambda\hat{P}_{k-1}} \tag{11}$$

Measurement update:

$$z_{k,j}^i=h(x_{k,j}^i) \tag{12}$$

For measuring new information, it is divided into two fuzzy sets (Small and Large). The heading Angle is divided into three fuzzy sets, so six fuzzy rules can be divided. For each fuzzy rule, two subsequent parameters are set: turning rate ω and process noise standard deviation σ. Specific parameter Settings are shown in Tab. 1.

Table 1 Lists of the specific parameters

Rule	Innovation	Course Angle difference	Turning rate ω	Process noise covariance σ
1	S	NL	−0.123	0.00001
2	L	NL	0.23	0.0011
3	S	S	0	0.0007
4	L	S	0.121	0.0081
5	S	PL	0.123	0.00001
6	L	PL	0.235	0.0011

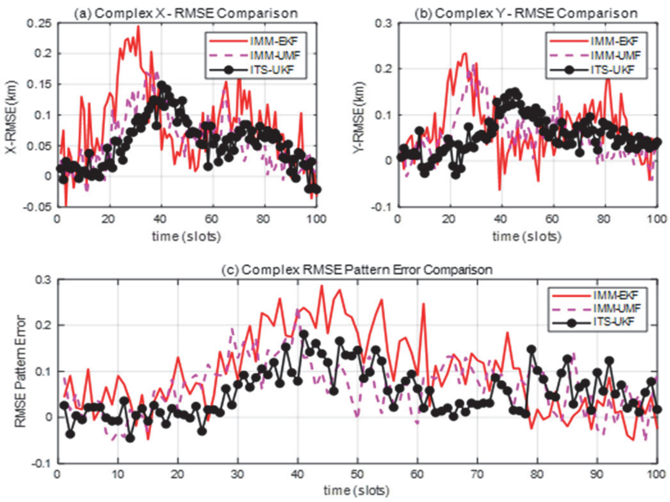

Figure 2 Tracking error performance of the three algorithms

Suppose that the model set of these two algorithms consists of three motion models: the motion linear motion model, the right turning motion model with the turning rate $\omega = 0.125$ rad/s, and the left turning motion model with the turning rate $\omega = -0.125$ rad/s. Two dimensional fuzzy semantics, IMMUKF, IMM-EKF three algorithms do not directly provide the correct turning motion model. The tracking error performance of the three algorithms is shown in Fig. 2.

It can be seen from the figure that the tracking error performance of the three algorithms is generally the same when the target is in uniform linear motion stage (1 ~ 31 s, 51 ~ 70 s). From 31 to 50 s, the target began to make a right turn maneuver. Since the model sets of IMM-UKF and IMM-EKF lacked accurate right-turn motion models, both algorithms struggled to precisely track and estimate the target's movement during these maneuvers. Consequently, their tracking performance deteriorated, and the RMSE increased.

Due to the powerful fitting ability of T-S fuzzy model, 2D fuzzy semantic modeling is based on target motion features and the parameters of the previous rules are adjusted adaptively, which makes the performance of 2D fuzzy semantic modeling still very stable.

The experimental data are shown in Tab. 2.

Table 2 shows the tracking errors of the three algorithms

Algorithm	RMSE	X RMSE	Y RMSE
IMM-EKF	0.0859 km	0.0561 km	0.0561 km
IMM-UKF	0.0505 km	0.0365 km	0.0324 km
Two dimensional fuzzy semantics	0.0369 km	0.0265 km	0.0256 km

3.4 Interactive Two-Dimensional Fuzzy Semantic Modeling Maneuvering Target Tracking Algorithm Based on Fuzzy Intensive Reduction

Each feature is divided into m_i, $i = 1, 2, ..., n$ fuzzy sets, then the total number of fuzzy models is:

$$N_s = \prod_{i=1}^{n} m_i \tag{13}$$

When the fuzzy sets of features and their partition increase, N_s will show an explosive growth trend. In this case, the concept of fuzzy set is introduced to reduce the redundant features and reduce the computational load of the algorithm.

If given a knowledge base $K = (U, R)$, for each subset $X \subseteq U$ and an equivalence relation $R \in IND\ (K)$, the upper and lower approximation RX are defined as follows:

$$\overline{R}X = U\{Y \in U/R \mid Y \subseteq X\} \tag{14}$$

Visualize the concept of fuzzy sets, as shown in Fig. 3.

As shown in Fig. 3, "Interactive T-S Fuzzy Semantic multi-model maneuvering" and "fuzzy dense reduction" have a complementary optimization relationship in the algorithm framework. The former builds a multi-model interaction mechanism through the T-S fuzzy model and dynamically adapts to the maneuvering characteristics of the target relying on fuzzy rules. However, the number of models increases exponentially with the increase of feature dimensions, resulting in a sharp increase in computing load. The latter, as the core optimization method, quantifies the feature dependency degree based on the neighborhood fuzzy set theory. Through the adaptive residual monitoring mechanism, redundant models are dynamically eliminated, reducing the computational complexity by more than 40% while retaining the key motion features. The dependence degree of each feature is calculated by the theory of neighborhood fuzzy intensification. The dependence of characteristic $B1$ is 1. In other words, the weight of these rules depends to a large extent on feature $B1$. Therefore, rules 1 and 3 are retained, and rules 2 and 4 are reduced. The complete algorithm flow chart is shown in Fig. 4.

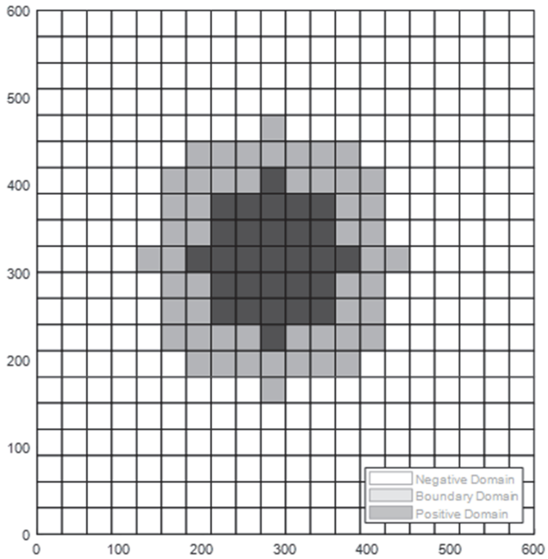

Figure 3 Graphical representation of the concepts of negative domain, boundary domain and positive domain in fuzzy sets

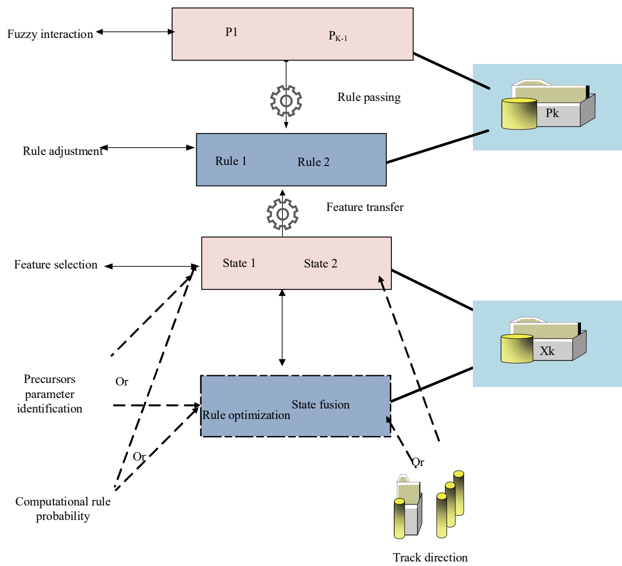

Figure 4 Flow of interactive T-S fuzzy semantic multi-model maneuvering target tracking algorithm based on fuzzy intensive reduction

Fig. 4 shows the process from policy initialization to data storage and subsequent processing. The policy initialization is input into the policy module, processed by the rule module, and then flows to the stage relying on the function module. The stage module also writes associated data. Subsequently, the stage information flows to the stage storage, involving interactions such as resource generation/operation and communication cost estimation. Finally, the stage storage connects to the physical infrastructure. Present the complete logic of the orderly transmission, processing of data among various modules of the system and the association with external facilities. Data presented in a sequential format ($x_1, x_2, ..., x_n$). The decoder receives input data encoded into a semantic vector e from the encoder, which contains extra characteristics. The decoder finishes the prediction by adding the output of the prior instant to the present moment as an additional input. Furthermore, while designing sports virtual sceneries, four issues must be taken into account: (1) A dynamically generated model is necessary for the designed sports virtual scenes to adapt to different users' needs. (2) As the number of models increases, there is a need for an effective mechanism to manage them. (3) The system needs to be updated and supplemented with new models in a timely manner. (4) The problem of dynamically generating models according to user requirements necessitates a method to apply them to an instructional software development environment. Various human motions are described using image processing methods in this paper's methodology. The target detection system uses the human body's important points to rebuild an action in response to a given human movement; this allows athletes to swiftly and easily adjust to new training techniques and tactics, allowing them to perform at a higher level. In Tab. 3 you can see the algorithm in action.

Table 3 T-S fuzzy semantic multi-model maneuvering target tracking algorithm based on fuzzy intensive reduction

1) Initialize $k = 0$
a) Set the initial component parameters and component structure
The object features are selected and fuzzy partition is performed to establish the initial fuzzy membership function.
According to the feature division of each object, fuzzy rules are established.
2) For $k = 1,2,...$
a) Conduct the nearest neighbor data association for observation data and exclude false alarms.
b) Fuzzy interaction
Calculate the intersection degree $\Delta(i, h)$ between uzzy sets.
Calculation of transition probability matrix and mixing probability.
Calculate the state and covariance after the mixed interaction.
c) Parameter identification of T-S fuzzy model
Based on UKF, the parameters of the latter part are identified, and the parameters of the former part are identified.
d) Model update probability
The weights of each fuzzy rule and the normalized weights are calculated.
e) Target feature reduction based on fuzzy sets
The dependence of each weight on each feature is calculated.
The residuals between the predicted states and the observations of each rule are calculated.
Select feature reduction or feature reselection according to the residual results.
f) State and covariance calculation
Calculate state estimates and covariance.

4 SIMULATION VERIFICATION

We can observe that the athletes' performance increased substantially after implementing the target detection suggested in this article, as shown in Tab. 4. With the help of the target recognition technology, players may show off a wide range of technical maneuvers that complement certain strategies and reduce the risk of harm.

Table 4 compares the results before and after using the object detection system

Movementtype	Rebound	Shoot at thebasket	Pass	Excel	Steal
Before use	76.7%	81.8%	80.5%	73.2%	79.4%
After use	85.7%	91.7%	93.2%	87.9%	89.6%

The target detection system offers certain benefits and can deliver superior teaching outcomes compared to standard training techniques. This advantage is more readily depicted in Fig. 5. Athletes can benefit from enhanced retention and comprehension of the technology through this mixed-modal approach that combines explanation and demonstration.

Traditional models generally fail to recover some of the arm postural features of sports, such as severe occlusion, high movement speeds, sudden changes of direction, and a lot of physical confrontation between players. These characteristics pose challenges to the accuracy of the detection efficiency of individual players and teams. In the experiment, the author combined the background difference algorithm and the skeleton detection algorithm to detect the key points of the human body. This combination can effectively identify the behavior of sports players. The author compared the performance of different algorithms in the existing literature, as shown in Tab. 5. It can be seen that the author's method shows very good performance in all data sets, and without additional annotation in the data set construction, sports play membership can be classified, so that the technical movements of sports players can be more accurately identified.

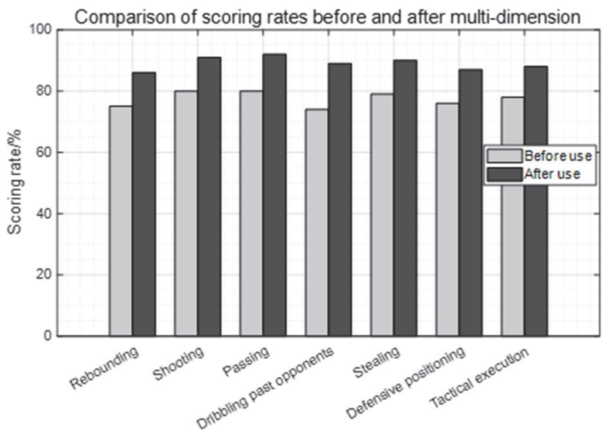

Figure 5 Performance of different sports activities before and afte rusing the object detection system

Table 5 Performance of different algorithms in recognizing human movements

Data set algorithmclass	HMDB51	UCF	Hollywood2	Olympic	Meanvalue
KNN	0.6234	0.6626	0.5717	0.6137	0. 6026
Co-training	0.5612	0.6179	0. 5872	0. 6391	0.5913
MCM	0.6750	0.5395	0.6031	0.5174	0.5877
Schuldt	0. 5147	0.6627	0. 6582	0.5328	0. 5780
Dollar	0.7438	0.7032	0.6747	0. 6753	0.7184
Niebles	0. 7681	0.7165	0. 6938	0.7104	0.7246
3D CNN	0.7583	0.7309	0. 6401	0.7320	0. 7213
Textualalgorithm	0.7687	0.7543	0.7689	0.7521	0. 7610

There are two moments: at $k-1$, use all the fuzzy rules to track the target, get all the feature information, and reduce it to select the most suitable feature. The fuzzy motion model corresponding to these features is used at kk. Manual control input all features in 1 s, 30 s, 50 s, 70 s, while 2 s, 31 s, 51 s, 71 s. Run 1,000 Monte Carlo experiments on this. The experimental results are shown in Fig. 6.

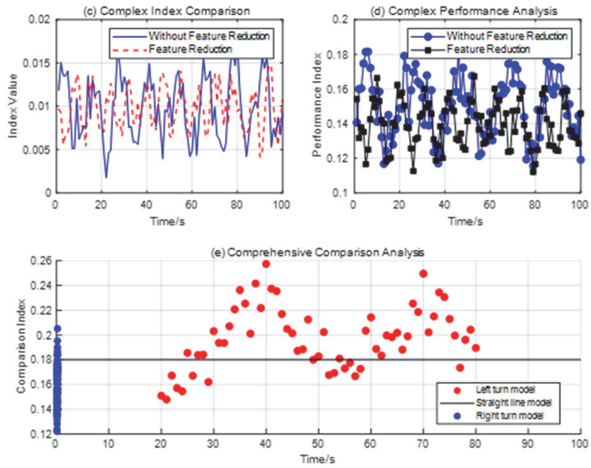

Figure 6 Simulation results of feature selection based on neighborhood fuzzy sets

As you can see from Fig. 6, not all time model sets have three models. In the interval of 3 ~ 30 s, there is only -"linear model" in the model library. From 31 to 50 s, there are "straight line model" and "right turn model" in the model library, but "left turn model" is not selected into the model library. "Right turn model" was not selected into the model library in the period of 71 - 90 s. This is also consistent with the motion state of the actual target, indicating that the neighborhood fuzzy set can select the appropriate model base according to the motion state of the target, corresponding to the UKF maneuvering target tracking algorithm based on T-S fuzzy semantics, that is, the features of the fuzzy model can be reduced to select the appropriate feature configuration rules.

The simulation results of the algorithm are shown in Fig. 7.

Statistics: the correct target number is 387, the correct rate is 96.99%. Total running time: 0.3189 s.

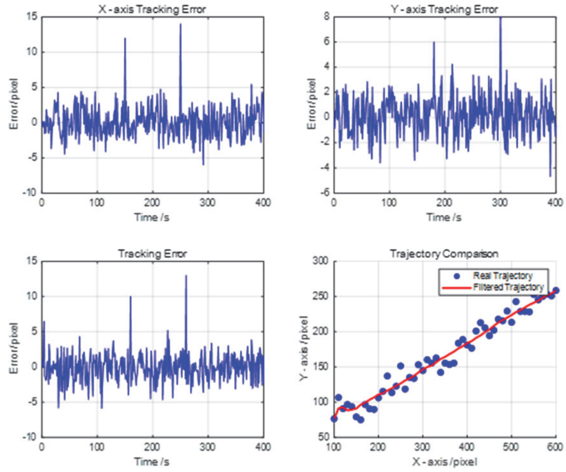

Figure 7 Simulation results of two-dimensional fuzzy semantic algorithm

Tab. 6 lists the comparison results tested on the data set. It is found that the method proposed in this paper has the smallest root-mean-square error, and can effectively detect the cause of video distortion, and use the corresponding method to evaluate the quality of video modeling. At the same time, the results of target detection in this paper can make up for the problem of missed detection of target detector to a certain extent. In addition, the performance of target detection can be further improved by feeding the detection results back to the target detector for on-line fine-tuning. It can be seen that the research of joint target detection and tracking algorithm is the key to further improve the video target tracking performance, and it is also the development trend of the future target tracking field.

Table 6 Comparison of evaluation results of sports competitions at different levels

Evaluation method		Type of competition					
		College sports competition	CBA	AsianGames	World championship	OlympicGames	NBA
BLIINDS－II	SROCC	0.9258	0.8682	0.8814	0.8228	0.9088	0.8638
	KROCC	0.7847	0.7958	0.8158	0.8857	0.8957	0.8592
	PLCC	0.9347	0.8745	0.8774	0.8794	0.9257	0.9173
	RMSE	2.4562	3.6175	5.9623	4.9561	4.9626	6.9789
DIIVINE	SROCC	0.9241	0.8258	0.8345	0.8793	0.8431	0.8694
	KROCC	0.7393	0.7924	0.7957	0.8451	0.7365	0.7927
	PLCC	0.9375	0.8694	0.8274	0.8735	0.8877	0.8636
	RMSE	3.2930	4.0761	4.0915	5.9741	3.5194	4.9639
BRISQUE	SROCC	0.9248	0.8148	0.8848	0.9228	0.8765	0.9117
	KROCC	0.7851	0.7285	0.8628	0.8649	0.8124	0.8795
	PLCC	0.9219	0.8637	0.8964	0.9255	0.9214	0.9473
	RMSE	1.4635	3.2176	3.8766	3.8461	3.4842	3.2983
Textualalgorithm	SROCC	0.9311	0.8725	0.8531	0.9286	0.9266	0.9356
	KROCC	0.8133	0.8285	0.8596	0.8859	0.8783	0.8974
	PLCC	0.9317	0.9138	0.8872	0.9822	0.9732	0.9386
	RMSE	1.0336	2.4766	3.3161	2.1146	2.4123	3.1671

In practice, the models in the general model library may not exactly match the movement of the target. Now keep other conditions unchanged, model 2 and model 3 in the model library are changed into turning models with turning speed of 20 °/s and −20 °/s respectively, and the simulation results are shown in Fig. 8.

The experimental sample consisted of a sports league movie that had been retrieved using multi-mode image capture. A computer with two processor cores, four gigabytes of random access memory (RAM), Windows 7, and other specifications was used to process and show the photos, data, and icons. Using Visual Studio 2010 with CV

2.0 for data programming. For the purpose of detecting and tracking moving objects (such as sports trajectories and athlete behavior), this research proposes three methods: optical flow detection, two-dimensional fuzzy semantic modeling, and an enhanced two-dimensional fuzzy semantic modeling approach. Fig. 9 shows the results of monitoring moving objects, which are used to collect verification data of interference processing capabilities of various approaches.

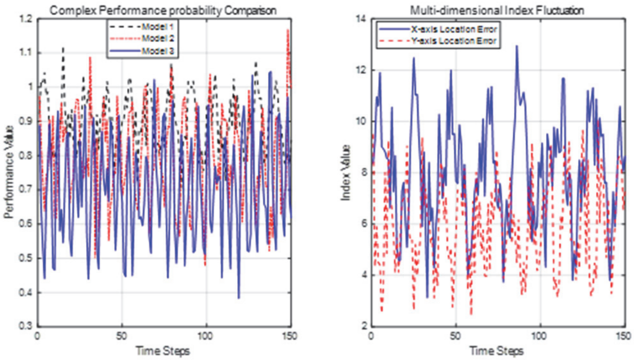

Figure 8 Simulation results of unsatisfactory model library algorithm

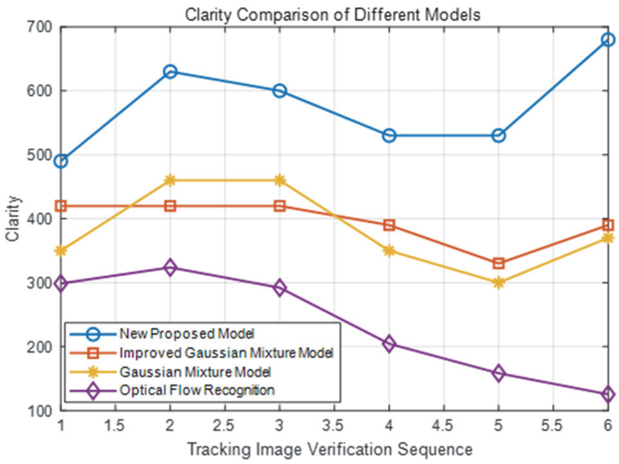

Figure 9 Verification results of interference processing capability

Fig. 9 shows that compared to the previous two-dimensional fuzzy semantic modeling model, the improved method achieves the highest level of picture clarity, keeps the details of moving objects, and effectively removes noise from non-moving objects. This suggests that the method has strong interference processing ability and achieves good tracking results.

Fig. 10 shows how the position coordinates of the right moving target in the detection results are used to verify the detection range. With a wide detection range and excellent detection effect, the suggested approach successfully detects moving targets in sports videos. The findings are also complete in contour, meaning there are no empty places.

The comparative analysis of experimental data in Fig. 9 and Fig. 10 shows that 3D VR reconstruction of multiple image information feature points under the constraint of triangle gravity center can achieve a high accuracy of regional reconstruction, and has a great improvement in the stability of the algorithm.

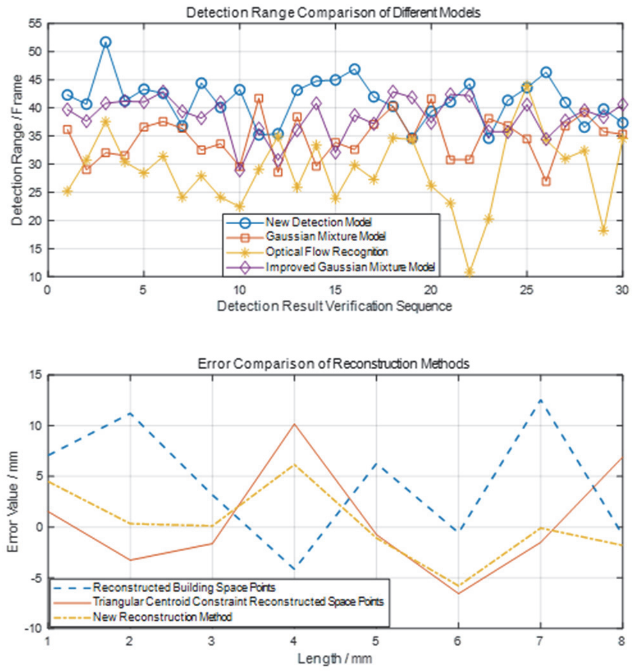

Figure 10 Reconstruct the line length relative error line

6 CONCLUSION

This study proposes a T-S fuzzy semantic multi-model maneuvering target tracking algorithm based on fuzzy intensive reduction to address computational overload and real-time degradation caused by model explosion when processing excessive target features in 2D fuzzy semantic algorithms. By consolidating model weights and parameters into a unified knowledge system, it leverages neighborhood fuzzy sets to eliminate redundant models and conserve computing power.

An adaptive residual-based reduction algorithm is introduced to dynamically trigger re-reduction by monitoring prediction-observation residuals, determining when feature/model re-reduction is needed through specialized diagnosis. Simulation results demonstrate this approach achieves a favorable trade-off: while slightly reducing tracking accuracy compared to standard 2D fuzzy semantic methods, it significantly enhances real-time performance through efficient computation management. Good tracking effects can still be achieved when the requirements for the details of the foreground are not high. Once the requirements for the details of the foreground are raised, it is often difficult to achieve good results. If multiple cameras can be adopted to shoot the moving target from multiple angles and integrate the correlations among the cameras to track the moving target, then the robustness and stability of the algorithm will be greatly improved. In future research, the development of multi-source information fusion should be strengthened.

7 REFERENCES

[1] Zeng, J. & Fu, J. (2024). Basketball robot object detection and distance measurement based on ROS and IBN-YOLOv5s algorithms. *PLoS ONE*, *19*(11), 1-13. https://doi.org/10.1371/journal.pone.0310494

[2] Zhao, J., Bao, W., & Mo, L. (2025). Design of tomato picking robot detection and localization system based on

deep learning neural networks algorithm of Yolov5. *Scientific Reports*, *15*(1), 80-86. https://doi.org/10.1038/s41598-025-90080-6

[3] Feng, Y., Ma, W., & Tan, Y. (2024). Approach of Dynamic Tracking and Counting for Obscured Citrus in Smart Orchard Based on Machine Vision. *Applied Sciences-Basel*, *14*(3), 14-36. https://doi.org/10.3390/app14031136

[4] Shen, X., Yuan, G., & Wang, H. (2024). High-strength synergic-calibration attention system in YOLO for underwater object detection application. *Multimedia Systems*, 2024, *30*(6), 1-28. https://doi.org/10.1007/s00530-024-01545-z

[5] Xu, Y. (2025). Research on Positioning and Tracking Method of Intelligent Mine Car in Underground Mine Based on YOLOv5 Algorithm and Laser Sensor Fusion. *Sustainability*, *17*(2), 542-556. https://doi.org/10.3390/su17020542

[6] Xu, Z., Meng, Y., &Yin, Z. (2024). Enhancing autonomous driving through intelligent navigation: A comprehensive improvement approach. *Journal of King Saud University - Computer and Information Sciences*, *36*(6), 102108-102123. https://doi.org/10.1016/j.jksuci.2024.102108

[7] Wei, B., Liu, J., & Li, A. (2024). Remote Distance Binocular Vision Ranging Method Based on Improved YOLOv5. *IEEE Sensors Journal*, *24*(7), 14-31. https://doi.org/10.1109/JSEN.2024.3359671

[8] Susanto, R., Silitonga, J. A., & Analia, R. (2023). Tiny-YOLO distance measurement and object detection coordination system for the BarelangFC robot. *International Journal of Electrical & Computer Engineering* (*2088-8708*), *13*(6), 6926-6939. https://doi.org/10.11591/ijece.v13i6.pp6926-6939

[9] Wang, S., Chen, M., & Ratnavelu, K. (2025). Online classroom student engagement analysis based on facial expression recognition using enhanced YOLOv5 for mitigating cyberbullying. *Measurement Science and Technology*, *36*(1), 15419-15432. https://doi.org/10.1088/1361-6501/ad8a80

[10] Ren, L., Li, Z., & He, X. (2024). An Underwater Target Detection Algorithm Based on Attention Mechanism and Improved YOLOv7. *Computers, Materials & Continua*, *78*(2), 7028-7043. https://doi.org/10.32604/cmc.2024.047028

[11] Prabowo, A. B., Pratomo, A. H., & Ahmad, M. S. H. (2024). Humanoid robot control system utilizing cost-oriented automation (COA) and edge detection. *E3S Web of Conferences*, *501*(3), 10-32. https://doi.org/10.1051/e3sconf/202450101012

[12] Wang, X., Xue, G., & Huang, S. (2023). Underwater Object Detection Algorithm Based on Adding Channel and Spatial Fusion Attention Mechanism.*Journal of Marine Science & Engineering*, *11*(6), 1116-1137. https://doi.org/10.3390/jmse11061116

[13] Pei, B., Sun, Y., & Fu Y. (2025). Object detection in motion management scenarios based on deep learning. *PLOS ONE*, *20*(1), 5130-5146. https://doi.org/10.1371/journal.pone.0315130

[14] Chidi, E. U., Anoliefo, E., & Udanor, C. (2025). A blind navigation guide model for obstacle avoidance using distance vision estimation based YOLO-V8n. *Journal of Nigerian Society of Physical Sciences*, *7*(1), 2292-2305. https://doi.org/10.46481/jnsps.2025.2292

[15] Zhuo, J., Li, G., & He, Y. (2023). Research on Cone Bucket Detection Algorithm Based on Improved YOLOv5s. *World Electric Vehicle Journal*, *14*(10), 16-39. https://doi.org/10.3390/wevj14100269

[16] Yu, W., Li, J., & Wang, Y. (2023). Small ship detection based on YOLOX and modified Gaussian Wasserstein distance in SAR images. *Proceedings of SPIE*, *12552*(2), 8-27. https://doi.org/10.1117/12.2667277

[17] Wang, S., Chen, M., & Ratnavelu, K. (2025). Online classroom student engagement analysis based on facial expression recognition using enhanced YOLOv5 for mitigating cyberbullying. *Measurement Science and Technology*, *36*(1), 15419-15428. https://doi.org/10.1088/1361-6501/ad8a80

[18] Peng, Z., Han, J., & Tong, L. (2024). Research on improving the ranging accuracy of ships with stereo vision through Kalman filter optimization. *PLoS ONE*, *19*(11), 305714-305728. https://doi.org/10.1371/journal.pone.0305714

[19] Ma, J., Wang, X., & Xu, C. (2023). SF-YOLOv5: Improved YOLOv5 with swin transformer and fusion-concat method for multi-UAV detection. *Measurement & Control (0020-2940)*, *56*(7/8), 64126-64139. https://doi.org/10.1177/00202940231164126

[20] Wen, X., Wang, J., & Cheng, C. (2024). Underwater Side-Scan Sonar Target Detection: YOLOv7 Model Combined with Attention Mechanism and Scaling Factor. *Remote Sensing*, *16*(13), 2492-2506. https://doi.org/10.3390/rs16132492

[21] Sun, W., Shen, X., & Zhang, X. (2025). Lightweight Ship Object Detection Algorithm for Remote Sensing Images Based on Multi-scale Perception and Feature Enhancement. *Photogrammetric Engineering and Remote Sensing*, *91*(2), 111-122. https://doi.org/10.14358/PERS.24-00085R2

[22] Han, K. (2020). Motion Recognition Algorithm in VR Video Based on Dual Feature Fusion and Adaptive Promotion. *IEEE Access*, *8*, 201134-201146. https://doi.org/10.1109/ACCESS.2020.3023755

[23] Ge, X. & Yue, M. (2024). A detection strategy based on deep learning against sequential outages induced by false data injection attacks. *Electrical Engineering*, *106*(4), 2277-2295. https://doi.org/10.1007/s00202-024-02277-z

[24] Dadheech, A., Bhavsar, M., & Verma, J. P. (2024). Deep learning based smart traffic management using video analytics and IoT sensor fusion. *Soft Computing*, *28*(23), 13461-13476. https://doi.org/10.1007/s00500-024-10382-1

[25] Afroze, S., Hossain, M. R., & Hoque, M. (2022). DeepFocus: A visual focus of attention detection framework using deep learning in multi-object scenarios. *J. King Saud Univ. Comput. Inf. Sci.*, *34*, 10109-10124. https://doi.org/10.1016/j.jksuci.2022.10.009

[26] Zhang, Y. Y., Zhang, S., & Hui, M. (2024). Semantic-Consistency-guided Learning on Deep Features for Unsupervised Salient Object Detection. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *20*(6), 40816-40825. https://doi.org/10.1145/3640816

[27] Wu, W., Wen, C., & Yuan, Q. (2025). Construction and application of knowledge graph for construction accidents based on deep learning. *Engineering Construction and Architectural Management*, *32*(2), 255-278. https://doi.org/10.1108/ECAM-03-2023-0255

[28] Xu, W., Blancaflor, E., & Abisado, M. (2025). Performance and improvement of deep learning algorithms based on LSTM in traffic flow prediction. *Discover Applied Sciences*, *7*(4), 6702-6713. https://doi.org/10.1007/s42452-025-06702-1

[29] Li, M., Liu, C., & Pan, X. (2025). Digital twin-assisted graph matching multi-task object detection method in complex traffic scenarios. *Scientific Reports*, *15*(1), 87914-87922. https://doi.org/10.1038/s41598-025-87914-8

[30] Chen, G., Ding, X., & Gao, X. (2024). Dynamic scenario deduction analysis for hazardous chemical accident based on CNN - LSTM model with attention mechanism. *Canadian Journal of Chemical Engineering*, *102*(12), 25318-25327. https://doi.org/10.1002/cjce.25318

[31] Yusoff, I. A., Isa, N. A, M., & Hasikin, K. (2013). Automated two-dimensional K-means clustering algorithm

for unsupervised image segmentation[J]. *Computers & Electrical Engineering*, *39*(3), 907-917. https://doi.org/10.1016/j.compeleceng.2012.11.013

Contact information:

Linlin WANG
(Corresponding author)
School of Physical Education,
Zhoukou Normal University,
Zhoukou 466001, Henan, China
E-mail: wanglinlin_zknu@163.com