

Research Article Distributed Soccer Training Smart Sensors for Multitarget Localization and Tracking

Jian Jiang 🕞 and Zhiqun Qiu 🕒

Department of Humanities, Gannan University of Science and Technology, Ganzhou, Jiangxi 341000, China

Correspondence should be addressed to Jian Jiang; 9320050232@jxust.edu.cn

Received 16 June 2022; Revised 16 July 2022; Accepted 19 July 2022; Published 5 August 2022

Academic Editor: Gengxin Sun

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

This paper presents an in-depth study and analysis of the localization and tracking of multiple targets in soccer training using a distributed intelligent sensor approach. An event-triggered mechanism is used to drive the acoustic array sensors in the distributed acoustic array sensor network, which solves the problem of increased communication load caused by frequent communication of microphones and effectively reduces the communication load between microphones as well as the energy consumption of the acoustic array sensor network. By designing a suitable state estimation equation for the acoustic source target and fully utilizing the measurement and state estimation information of its nodes as well as the state estimation information of neighboring nodes, the next moment state of the acoustic source target can be accurately predicted. A correlation filtering tracking algorithm based on multiscale spatial co-localization is proposed. In the proposed algorithm, the tracker contains a total of several subfilters with different sampling ranges. Then, this paper also proposes a collaborative discrimination method to judge the spatial response of the target image samples of each filter and jointly localize the target online. Based on this, this paper further explores the potential of correlation filter tracking algorithms in complex environments and proposes a robust correlation filter tracking algorithm that fuses multiscale spatial views. The cross-view geometric similarity measure based on multiframe pose information is proposed, and the matching effect is better than that based on single-frame cross-view geometric similarity; to solve the problem of player appearance similarity interference, a graph model-based cross-view appearance similarity measure learning method is further proposed, with players in each view as nodes, player appearance depth features as node attributes, and connections between cross-view players as edges to construct a cross-view player graph. The similarity obtained by the graph convolutional neural network training is better than the appearance similarity calculated based on simple cosine distance.

1. Introduction

The detection and tracking of motion targets in sports video are one of the most fundamental research areas in sports video analysis. This fundamental work is of great importance for further implementation of higher semantic level event detection and technical and tactical analysis. However, this fundamental work has not been completely solved due to many challenges in sports video [1]. For the former question, considering the significant improvement in processing speed and accuracy of human pose estimation technology in recent years, this chapter intends to adopt the appearance feature representation method of pose alignment. In recent years, with the advent of the era of deep learning, many visual detection and tracking methods based on convolutional neural networks have been widely used in sports videos with good results. Since the ball belongs to a single target and the player generally belongs to multiple targets, their detection and tracking methods are different, so this section will introduce the development status of their detection and tracking methods for the ball and player, respectively. Distributed acoustic array system networks are widely used in acoustic source-target tracking because of their high tracking accuracy, high multitarget resolution, and large error tolerance compared to singlearray networks [2]. The distributed acoustic array system network consists of multiple array sensors working in a distributed and cooperative manner, with each node receiving acoustic signals from the source target for processing in real time, as well as receiving measurement information and estimated state information from neighboring communication nodes.

Due to the periodic and frequent communication, the sensor network communication volume inevitably increases, thus wasting the bandwidth resources of the acoustic array sensor network, which may result in information delay and packet loss, channel fading, and other phenomena. Therefore, we consider a filter tracking algorithm based on the event mechanism and introduce event-triggered variables to filter the signal transmission through the nodes [3]. When the event triggered, the acoustic array nodes only transmit information to neighboring nodes, which reduces the communication load and energy consumption between microphone nodes while accurately tracking the source target trajectory. During the motion of the target object, its appearance will change in real time, and the tracker needs to adapt to the changes in the appearance of the object while also dealing with the interference caused by the surrounding contextual information on the target object. Usually, the object is in motion, along with the object's motion or the camera's movement, the object presents a different pose on the video screen, which produces a change in the appearance of the object; for such changes, the tracker needs to adapt to it by some method, and most tracking algorithms use a template update strategy to update the model [4]. However, in such a process, the object may also receive interference such as occlusion, which also causes changes in the appearance of the object, and if the tracker does not handle such changes well, it will easily absorb the interference information of the wrong pair and lose the target object when it recovers in the video frame.

The current target tracking algorithms generally consider the learning of classifier as a binary classification problem in the online training process; however, the scenario faced by the tracker is not a binary classification scenario, and there may be multiple states of the target object, so the learning of classifier should be multivariate classification and treated differently. The current target tracking algorithm mainly follows the classical binary classification sample in the sampling space context sampling problem, which still cannot truly reflect the actual state of the target object in the tracking process. How to use the spatial information around the target object to achieve better learning of the tracker is an urgent problem to be solved. This type of approach is equivalent to an extension of the single-target tracking algorithm, but it cannot handle the case of new target emergence. Detection-based data association tracking does not require manual initialization of target positions, but it requires detection of each frame using a target detector, and then data association of the detection results of each frame to finally obtain the motion trajectories of all targets of interest. This method can automatically discover new targets and automatically terminate disappearing targets, but the method is more dependent on the detection effect of the target detector.

2. Related Jobs

Razzaq et al. proposed a dynamic cluster membership selection algorithm for multitarget tracking and constructed a comprehensive performance metric function based on network energy consumption and tracking accuracy, which can measure WSN tracking performance to some extent [5]. However, the algorithm does not consider the energy consumption of individual nodes, the residual energy, and the contribution of the selected nodes' measurement information to target localization. In the sensor scheduling problem of WSN multitarget tracking, Goldhoorn et al. used Mahalanobis distance based on the predicted location of the target as a measure of the observation information of sensor nodes, designed optimization metrics considering tracking accuracy and residual energy distribution of sensor nodes, and proposed an adaptive node selection algorithm for energy efficiency optimization [6]. For the multitarget tracking problem with a known and fixed number in WSN, Bhat et al. propose a fully distributed event-triggered measurement and communication strategy that enables each sensor to achieve a balance between estimation error and energy consumption without the need for global information [7]. Tracking based on trajectory prediction needs to use the trajectory prediction method to predict the position of the tracking target in the current frame and then match the detection frame of the target in the current frame with the trajectory prediction frame. The most used algorithm for multitarget tracking is the detection-based multitracking algorithm, which consists of two parts: target detection and data association. The detector provides the detection results, and the association part assigns the same identity to the detection results belonging to the same target in consecutive video frames based on the detection results and keeps them stable and unchanged in the subsequent video sequences [8]. Meanwhile, thanks to the continuous development in the direction of single-target tracking, many single-target tracking methods are also used to assist in multitarget tracking tasks.

In detection-based multitarget tracking algorithms, high accuracy of target detection is the basis for ensuring tracking accuracy, and the good or bad detection results directly determine the tracking performance of this type of multitarget tracking paradigm [9]. Nowadays, the challenge of target detection algorithms is mainly whether the detector can completely and comprehensively detect the target of interest in complex environments, where the research difficulty is small target detection. With the continuous development of deep neural networks, convolutional neural networks have achieved good performance in target detection. Nowadays, most of the target detection tasks are implemented based on convolutional neural networks [10]. Liu et al. proposed a selective search method, which improves the idea of undifferentiated computation for each location in sliding windows, but uses pixel clustering to merge regions, generates candidate regions according to the merging, and finally converts these regions into fixed-size images and sends them to the convolutional neural network for classification [11]. The mainstream single-target tracking algorithm nowadays

builds a robust appearance model and updates this model over time so that the object can be tracked in new video frames even if the target is occluded, deformed, etc., and can continue tracking after the occlusion ends [12].

In WSNs, sensor nodes are much weaker than ordinary computers in terms of computing, storage, and communication capabilities, which is because to meet the practical needs, sensor nodes generally have a small size, less energy, less power consumption, and low cost. These make the sensor node operating system protocol level design requirements simple and efficient. And the impact of the environment or in the manual handling of more troublesome conditions, many nodes are difficult to replace the battery promptly once it is exhausted, which means that sensor nodes are generally one-time use, so we should consider designing the WSN to be as low energy consumption mode as possible. This is mainly since the contextual features of the players are not constant, but change with the player's position at different moments. The network can start with some initial deployment, and then the nodes can be deployed to collect information. The information collected by the mobile nodes can communicate with another mobile node when they are within a distance of each other. Another key difference is the data distribution. In a static wireless sensor network, data can be distributed in a way that uses fixed routing or flooding, but the way applied to a mobile wireless sensor network is dynamic routing.

3. Distributed Smart Sensor Design

When the target moves, each sensor node with sufficient energy can monitor the target within its sensing area and transmit the sensed information to the aggregation center. However, to achieve global optimization of tracking performance, only some of the sensors dispatched for monitoring the target at each step. The aggregation center fuses the received measurements, updates the state estimate for the current step, predicts the target state and energy consumption for the next step, and bases these predictions on the information. Since the measurements obtained from the sensors sensing the target are interspersed with various noises, the information obtained from the sensors needs to be filtered to reduce the effect of noise and improve the tracking accuracy.

With the minimum mean square error as the best estimation criterion, the state-space model of the signal is used to update the estimates of the state variables using the estimates of the previous moment and the observations of the current moment to find the estimates of the current moment state. Many types of research in action recognition are based on the pose estimation information, and similarly, the player 3D pose estimation will also help the subsequent player action recognition [13]. However, after losing more than 3 frames, it is considered that the contextual characteristics of the player will also change due to the changes of the adjacent players, so the individual characteristics of the player are used to match. The players in sports videos often move vigorously with large pose changes, and there are also more occlusions, so it is not easy to get an accurate 3D pose estimation of players. In this paper, we also do some research on 3D pose estimation based on the implementation of 3D player tracking.

In the electromagnetic signal radiation source localization technology, firstly, the mutual correlation between the direct signal and the target scattering signal is processed so that the delay of the latter relative to the former can be obtained, and then the distance difference between the distance from the transmitting station to the receiving station via the target and the baseline can be obtained, which has a wide range of applications in various fields of modern society. Especially in military systems, the accurate localization of the radiation source target is very important, as shown in Figure 1.

Although it has a good localization effect, strong antiinterference capability, and good climate adaptation, however, the communication station is susceptible to the influence of indeterminate factors leading to failure of localization when collecting signals when the radiation source signal is known. The positioning accuracy of this target has a strong impact on the positioning of silent and stealthy targets. The existing techniques for target localization can usually be divided into active and passive localization.

But the use of active positioning systems, easy to expose themselves, especially the emergence of anti-radiation missiles on the survival of active equipment such as radar poses a serious challenge, this relies on the emission of high-power to achieve positioning technology, making this way very easy to reveal their location information; the enemy can launch an attack based on the exposed position. This attack has multiple lethality's, with soft kill by electronic jamming and hard kill by anti-radiation missiles, which in this case will affect the positioning accuracy and will not even guarantee the safety of the system itself.

Radiation source localization based on wireless sensor networks is not simple transplantation of traditional passive localization techniques on WSN hardware platforms, but a new cross-cutting research area. The application of radiation source passive positioning in wireless sensor networks with limited resources and capabilities is very valuable for research and practical difficulties. This shows that the performance of the NRSSI_DV-Hop algorithm has been effectively improved. Compared with the traditional base-station-based passive positioning system, this system has more reliability and finer localization. In addition, the deployment method is more flexible and covert.

In RSSI ranging, the receiving node uses the received known signal strength to analyze the correspondence between the received signal and the distance through signal propagation models such as the free-space propagation model, the log distance path loss model, the ha-it model, and the log-normal distribution model, based on the degree of signal consumption during transmission, to convert these propagation losses into the distance. In general, the average power of the received signal decays exponentially with the increase of distance [14].

Let *d* be the distance from the transmitter to the receiver; P(d) is the received signal strength in dBm at distance *d*; *np* is the path loss factor, which is used to indicate the rate of



FIGURE 1: Overall architecture of distributed positioning system.

path loss growth with distance, ranging from 2 to 4; d0 is called the trapping distance, with a typical value of d0 = 1 m (from the transmitter); P0 is the signal strength at the reference distance d0; ε denotes the error term, which is the masking factor:

$$P(d) = P_0 - 10n_p \ln \frac{d}{d_0} - \varepsilon.$$
(1)

The distance between the transmitting terminal and the receiving terminal has a certain connection with the absolute error formed by the fluctuation of the RSSI value, which increases when the distance increases. In the communication data transmission process, when the distance between two nodes is close, the signal attenuation is quite serious, and at a longer distance, the signal attenuation is gentler. When np is smaller, the signal attenuation in the process of propagation is also smaller, then the signal can propagate a longer distance, and when this propagation distance is farther, the accuracy of RSSI is also higher.

However, the signal interference is different in different environments, and there will be some errors in estimating the distance using this method. So, to get a more accurate blind node location, we also need to correct this error; specifically, the least-squares method can be used to correct the measured distance according to the specific environment.

$$p = \sum_{i} \left[(a - bx_i) + y_i \right]^2.$$
 (2)

However, these existing sensor scheduling methods are based on local performance optimization with finite time steps, which may be suboptimal from a global perspective, while pursuing global optimality can save resource consumption overall and promote the sustainable operation of the network. In EHWSN, the network life cycle can be theoretically infinite, so the development of energy harvesting theory and technology brings new challenges to obtaining the global performance optimal target tracking sensor scheduling on an infinite life cycle. The ADP is an effective method to solve the global performance optimization based on the infinitely long-life cycle, and it can obtain the approximate optimal control iteratively by continuously approximating the optimal value of the infinite step performance through the function approximation structure.

$$X[k] = \left[x[k], v_x(k), y_x(k), v_y(k)\right]^T.$$
(3)

These schemes provide important academic implications and applications for optimal control of EHWSN. However, these schemes are single-sensor scheduling schemes, while multiple sensors cooperatively monitoring targets can further improve the tracking performance and enhance the flexibility, stability, and fault tolerance of the tracking system. Therefore, for cooperative target tracking in EHWSN, the ADP-based multisensor scheduling approach is still an open and challenging problem to be solved.

The 3D tracking system of a ball usually includes at least 3 cameras, and the cameras are paired two by two to obtain multiple 3D coordinate points by triangulation algorithm [15]. In the second part, the real nodes are arranged in the university playground, and the radiation source is in the real environment. How to further fuse these multiple 3D points into one 3D point is also the problem to be solved in this section. Considering that the 2D tracking results in each camera may still be noisy, it is not advisable to directly average the 3D coordinates. Here, the accuracy of each 3D coordinate calculated with the triangulation algorithm is first evaluated by introducing the inverse projection error, and then the 3D coordinates with higher accuracy are selected and averaged to obtain the fused 3D coordinate positions, as shown in Figure 2.

For continuous dynamic systems, the optimality principle also holds. Taking a two-dimensional continuous-time dynamic system as an example, suppose there is a state variable trajectory from a known initial state x(t,) to x(t,) that is optimal, as shown in Figure 2, then the control corresponding to this trajectory necessarily makes the following performance index functions minimal.

$$J(x(t_1)) = \int_{t_1}^{t_f} U(x(\tau), e(\tau)) d\tau.$$
 (4)

Different energy harvesting sensor nodes have different sensing target capabilities, and they can transmit the sensed target information (such as received signal strength and the distance between the target and itself) to the convergence center with information processing capability when energy allows [16]. Considering the harm of radiation to human beings, the radiation intensity used in the experiment is very small. At the same time, each energy harvesting sensor has an energy harvesting module, which can convert the collected solar energy into electrical energy and store it in the energy storage device to provide energy for the subsequent work of sensing targets, sending information, and receiving information. Since the capacity of the energy storage device of each energy harvesting sensor node is not infinite, here we assume that the maximum storable energy of the energy storage device of the ith energy harvesting sensor is E.

4. Analysis of Multiobjective Localization and Tracking Algorithms for Soccer Training

For player appearance features in team sports games, they mainly influenced two aspects. On the one hand, large changes in player posture often cause impure global appearance features. On the other hand, the fact that players of the same team wear the same color jerseys makes them sometimes difficult to be distinguished by appearance features. For the former problem, considering the dramatic improvement in processing speed and accuracy of human pose esti-



FIGURE 2: Schematic diagram of dynamic planning for continuoustime dynamic systems.

mation techniques in recent years, this chapter proposes to adopt a pose-aligned appearance feature representation method [17]. The global depth features of the player detection frame are first extracted, and then a heat map containing the pose information is obtained by the pose estimation method, and the heat map is used as a mask to extract the appearance features near the player's human body joints so that the background interference outside the human body parts in the detection frame can be effectively shielded. The feature representation method has been effectively validated in a pedestrian re-identification task.

For the latter problem, which makes it difficult to distinguish between players due to similar appearance, this chapter proposes to further improve the representation of player appearance features by introducing contextual information. In real life, people also usually use contextual information to distinguish targets, for example, when they cannot identify a person in a crowd due to occlusion, they can also infer to identify the target person by recognizing its surrounding people or objects. The players in each perspective are used as nodes, the players' appearance, and depth characteristics are used as node attributes, and the connections between players across different perspectives are used as edges to construct a cross-view player graph model. This chapter intends to learn stronger appearance feature representation by building a contextual graphical model for the target player and its surrounding players, which in turn alleviates the tracking identity exchange problem between similar-looking players that often occurs in player tracking in sports videos.

The detection-based multitarget tracking mentioned above matches the detection frame of the target in the current frame with the tracking frame in the previous frame directly, while the track prediction-based tracking needs to predict the position of the tracking target in the current frame using the track prediction method and then match the detection frame of the target in the current frame with the track prediction frame, as shown in Figure 3.



FIGURE 3: Network distribution of unknown nodes and anchor nodes.

On the one hand, the depth features are used to match the detection target in the current frame with the tracking target in the previous frame; on the other hand, the Kalman filter is used to predict the motion state of the tracking target in the previous frame in the current frame, and the above matching based on depth features is further modified by calculating the intersection and ratio between the martingale distance between the target motion state and the target detection frame [18]. When deep sort is directly used for tracking multiple players in sports videos, the problem of player tracking identity exchange often occurs. The main reason for this is that players of the same team often wear the same color jerseys and their appearance features are very similar, so they are prone to false matching in the association matching stage of multitarget tracking, thus causing tracking identity exchange. In this regard, this chapter extracts more accurate depth features of player appearance utilizing pose guidance on the one hand and further improves the representation of player appearance features by exploring the contextual information of players on the other hand.

Matching is performed using a similarity matrix based on individual player features, which is mainly because the contextual features of players are not constant, but change with the change of player positions at different moments [19]. This basic work is of great significance for the subsequent further realization of event detection and technical and tactical analysis at a higher semantic level. Therefore, when matching players in two frames that are far apart, the contextual features of players often change a lot and can no longer be used as a basis for matching, and individual player features need to be used instead.

Finally, a cost matrix can be obtained by subtracting the similarity matrix from 1. So far, the association matching problem of current frame detection and existing trajectory can be converted into a binary assignment problem about the cost matrix, which can be solved effectively by the classical Hungarian algorithm, and the cascade matching strategy to deal with this association matching problem, as shown in Figure 4.

Cascade matching prioritizes the trajectories with smaller age values to match with the detection in the current frame, that is, the trajectories with age from 0 to a preset threshold are matched with the detection in the current frame one by one, and the trajectories that have not been lost are matched first, and the trajectories that have been lost for a long time are matched later. By processing in this way, the obscured targets can be retrieved again, and the number of identity exchanges for targets that reappear after being obscured can be reduced. In addition, the matching is mainly based on the four types of input information mentioned above. This is because if a trajectory loses a certain number of frames, its trajectory prediction value is less reliable, so only trajectories that have not been lost are matched with Marxian distance and intersection ratio [20]. For lost trajectories (age > 0), contextual features preferred to calculate the similarity matrix, but after losing more than 3 frames, it is considered that the player's contextual features will also change due to the change of neighboring players, so individual player features are used to match instead.

By varying the parameters, such as the number of unknown nodes, the number of anchor nodes, and different node deployment areas, the two algorithms are simulated and compared on three performance metrics. The specific formulas for localization error and localization energy consumption are given below:

$$e = \frac{\sum_{i=0}^{N} \left| X_i + \widehat{X}_i \right|}{N}.$$
(5)

In the RSSI ranging process, the larger the RSSI value sensed by the terminal device represents the smaller the distance between the anchor node and the unknown node, the smaller the signal attenuation when propagating in



FIGURE 4: Schematic diagram of association matching and trajectory management strategy.

space, and the less interference in that environment; when the positions of three anchor nodes form an approximately equilateral triangle, the more accurate the localization information. Therefore, this section will introduce the development status of detection and tracking methods for balls and players, respectively. The nodes with large RSSI values and other nodes forming equilateral triangles are used to locate the unknown nodes in a certain order according to the perceived RSSI values.

5. Analysis of Results

5.1. Distributed Sensor Positioning Accuracy Performance Analysis. From the localization error results, the localization error changes greatly as the proportion of anchor nodes increases, and when the proportion of anchor nodes increases to a certain value (around 20%), the localization error of unknown nodes changes little, because the proportion of anchor nodes is not a decisive factor affecting the localization error of the DV-Hop algorithm. The superiority of the NRSSI_DV-Hop localization algorithm can be seen from these results. Firstly, this is because the algorithm locates the unknown nodes within one hop from the anchor node using the RSSI technique, and the locating of these unknown nodes has higher accuracy compared with the original algorithm; secondly, the NRSSI_DV-Hop locating algorithm upgrades those unknown nodes located by RSSI measurement technique to anchor nodes for the locating of subsequent unknown nodes. This approach increases the proportion of anchor nodes in the network, resulting in higher localization accuracy. Therefore, the NRSSI_DV-Hop localization algorithm is effective in reducing the algorithm localization error as shown in the comparison of the node localization error within one hop and multihop nodes in the previous section.

Figure 5 depicts the curve of the average localization error of the two algorithms in multiple trials with the change of the communication radius by varying the node communication radius in a random scenario where the nodes are deployed with a constant number of nodes 150, a constant number of anchor nodes, and the same environment deployment. While reducing the communication load and energy consumption between the microphone nodes, it can accurately track the sound source target trajectory. In this case, the communication radius is taken as R = 10, 20, 30, 40, and 50 m, and comparative simulation experiments are conducted to obtain the node localization error of both algorithms for the NRSSI_DV-Hop algorithm and the DV-Hop algorithm, respectively.

From the experimental results, in wireless sensor networks, when the communication radius of the nodes to be located gradually becomes larger, the accuracy of the unknown nodes increases accordingly. Since the large communication distance can increase the possibility of communication between the unknown node and other nodes and increase the accuracy of the transmitted information, thus reducing the localization error, the NRSSI_DV-Hop algorithm has obvious advantages in localization accuracy, and the localization error of NRSSI_DV-Hop algorithm is smaller under each communication radius compared to



FIGURE 5: Unknown node localization error under different communication radius.

DV-Hop algorithm. It can also be observed that when the node communication radius is 30 m, the effect of communication radius has more influence on the localization accuracy for both the NRSSI_DV-Hop algorithm and DV-Hop algorithm, but the localization accuracy of the former is always higher than the localization accuracy of the latter, which indicates that the NRSSI_DV-Hop algorithm is effectively improved in performance.

The total energy consumed by both protocols decreases with the increase of the time interval of sending data, because the smaller the time interval of sending data, the more frequently the data is sent, and the energy consumption is naturally larger, and the larger the time interval, the less the number of sending data per second, and the less the energy consumed. For such changes, the tracker needs to adapt to it through some methods, and most tracking algorithms use a template update strategy to update the model. When the time interval of sending data is small, the communication volume of the network is very high, and the energy consumed by the MC_MAC protocol is about between 50 and 60 mJ, while the original protocol consumes much more energy than the MC_MAC protocol at this stage, between 80 and 100 mJ. As the interval time increases and the network traffic gradually decreases, the MC_MAC protocol slowly slows down the amount of energy consumption reduction, and the original protocol reduces energy consumption faster than the MC_MAC protocol, but still higher than the improved protocol. So, in general, the energy-saving effect of the MC_MAC protocol is much better than the original MAC protocol.

In addition, as all nodes in the WSN between the transmission of message packets are communicated through wireless signals, according to the signal transmission process, electromagnetic waves are sent for the exchange of message data, thus achieving the communication of message packets between all nodes in the network. In the traditional wireless communication process, the message data is transmitted in the channel, and the collision between the message packets causes the channel to be congested, making the message packets easily lost, or in the communication process, without any security measures program, the transmitted information is easily eavesdropped by the attackers present in the network, thus causing a series of damages to the message data, so in the packet communication process, various reasons lead to information leakage, although wired communication may also be subject to eavesdropping attacks, it is easy to be detected, as shown in Figure 6.

Based on the above existing packet eavesdropping, being analyzed packets, thus leading to message leakage, to avoid the above message leakage, this paper adopts a precodingbased anti-eavesdropping scheme, the cluster head nodes and intermediate nodes for the corresponding linear coding process, to ensure the security of data packets from being easily obtained decoding, mainly through the complete fusion of data and coding matrix, to obtain node data communication security, then the data after a series of encoding for network transmission, and finally to ensure that the data information in the communication process is not easily accessible.

To verify the localization accuracy of this wireless sensor network radiation source localization system, the first part is tested by system simulation of localization to derive the system localization results under different situations. In the second part, real nodes are arranged in a university playground



FIGURE 6: Comparison of node dose measurements at different distances.

to locate the radiation source in a realistic environment, and the intensity of the radiation value used in the experiment is very small considering the damage of radiation to humans.

Detection-based data association tracking does not need to manually initialize the target position, but it needs to use the target detector to detect each frame, and then perform data association on the detection results of each frame. At the same time, the single-channel MAC protocol with the LEACH algorithm is applied in the common wireless sensor network radiation source localization system, and the simulation result name is the original radiation source localization system representation, and the two radiation source localization systems are experimented separately for omnidirectional and localization two transmission comparisons. As mentioned above, we tried to ensure that the conditions of the two systems chosen for the experimental scenario were identical. Localization results were evaluated for seven nodes randomly distributed in our playing field, each with a maximum transmit power of 3.6 dBm and with a primary node of 5 m. The nodes were located at about 5 m from each other. The actual interval between nodes did not exceed 2000 ms.

5.2. Results of Soccer Training Multiobjective Localization Tracking Application. Players in different camera views can usually be matched by polar line constraints. Suppose there are two camera views with overlapping regions, and a point in space can be found as a homonymous image point in camera 1 and camera 2, respectively, then one of the homonymous points should be located in the polar line corresponding to the other homonymous point. In turn, if it is necessary to determine whether two points in different cameras are homonymous, it can be converted to determine whether the distance between the homonymous point and the polar line is less than a threshold value, and the cross-view player matching based on the geometric constraint of the polar line is based on this principle. In other words, if the corresponding points on two players are the same name points, the probability that these two players belong to the same player is higher.

Considering that the above cross-view geometric distances are only calculated based on the current frame, and since the two-dimensional trajectories of each player have been obtained, the cross-view geometric distances of two players in multiframe periods can be further considered based on the player's trajectory information. Use the estimated value of the previous moment and the observed value of the current moment to update the estimation of the state variable, and obtain the estimated value of the state at the current moment. In this way, a cross-view geometric similarity matrix based on multiframe pose information can be obtained.

Since the APIDIS soccer dataset only gives the player's 2D coordinate position, but not the 3D coordinate and posture information, a 3D trajectory is generated based on the known real 2D coordinates using the method proposed in this chapter, and this is used as the real value for evaluation. The center point between the two ankles of the player is used as the 3D coordinate position of the player.

Figure 7 shows the 3D tracking results of each method on the APIDIS dataset and the Campus dataset. In addition, the 3D tracking results are projected back into each 2D camera plane to further verify their tracking effectiveness. From the comparison results, the tracking results based on the improved similarity matrix are significantly better than those based on the improved similarity before, both in 3D space and in 2D camera planes. For the apparent similarity, the original method based on the simple cosine similarity metric is almost difficult to complete the cross-view target



FIGURE 7: Continued.



FIGURE 7: Comparison of 3D tracking results.

matching, while the method based on the graph model similarity metric learning proposed in this chapter has significantly improved the tracking effect.

Figure 8 shows the results of the multiplayer 3D tracking and 3D pose estimation when the proposed method in this chapter is applied to the Campus dataset and the APIDIS dataset, respectively. Given the image frames in multiple camera views, the method in this chapter first uses the multiplayer tracking method proposed in the previous chapter for tracking and 2D pose estimation of players in the 2D camera plane, then uses the two similarity measures proposed in this chapter for cross-camera matching, and finally achieves 3D tracking and 3D pose estimation of multiple players or players based on 2D trajectory information and cross-view matching. From the experimental effect figure, we can see that the method proposed in this paper is effective.

In addition, as mentioned in the introductory section of this paper, there are two main ideas of multicamera multitarget tracking methods: One is reconstruction followed by



FIGURE 8: Effect of player 3D tracking and 3D pose estimation.

tracking, and the other is tracking followed by reconstruction. This technology, which relies on transmitting high power to achieve positioning, makes it easy to leak its own location information, and the enemy can launch attacks based on the exposed location. This paper belongs to the reconstruction-then-tracking approach, that is, firstly, the POM algorithm is used to fuse the two-dimensional targets in multiple camera views into a probability occupancy map in the stadium plane, and then the probability occupancy map at different moments is used as input to construct a spatiotemporal network flow graph, which converts the multitarget tracking problem into a linear programming problem based on the minimum cost flow and adjusts the edge weights in the network flow graph by introducing the jersey color and jersey number information. Finally, the K shortest path method is used to solve the problem. This method uses offline tracking, which takes a longer time in the reconstruction phase and is not conducive to real-time player tracking. In contrast, this method adopts the framework of tracking first and reconstruction later, i.e., firstly, 2D detection, tracking, and pose estimation are performed in each 2D camera plane, then cross-view player matching is completed by using the similarity measure proposed in this chapter, and finally, 3D player tracking is completed based on 2D tracking and cross-view matching.

The experimental results on two public datasets show that considering both proposed cross-view player similarities can effectively improve the accuracy of player 3D tracking and 3D pose estimation. In addition, the experiments also show that the method in this chapter applies not only to sports scenes but also to pedestrian 3D tracking and 3D pose estimation in smart surveillance scenes.

6. Conclusion

To solve the problems of missed detection, false detection, and tracking drift that often occurs in the ball tracking process in sports video, this paper proposes a 3D ball tracking framework based on small target detection and multiview fusion, which mainly consists of four stages: 2D detection, 2D tracking, 3D coordinate fusion, and 3D trajectory

smoothing of the ball. And transmit the perceived information to the aggregation center. But to achieve global optimization of tracking performance, only part of the sensors is scheduled to monitor the target at each step. At the twodimensional level, for the problem that the size of the ball is too small to cause missed detection and false detection, multiscale depth features are used to improve the twodimensional detection accuracy of the ball; on the other hand, the introduction of cross-view information based on the polar line constraint and the detection-based model update strategy solve the ball tracking drift problem. The cross-view geometric similarity metric based on multiframe pose information is proposed, and the matching effect is better than the cross-view geometric similarity based on a single frame; to solve the problem of player appearance similarity interference, the cross-view appearance similarity metric learning method based on graph model is further proposed, with the players in each view as nodes, the depth features of player appearance as node attributes, and the connections between cross-view players as edges to construct the crossview player graph. The similarity obtained by the graph convolutional neural network training outperforms the appearance similarity calculated based on the simple cosine distance. The experimental results show that the proposed two kinds of cross-view player similarity can effectively improve the accuracy of player 3D tracking and 3D pose estimation by considering the proposed cross-view player similarity at the same time.

Data Availability

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

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Department of Humanities, Gannan University of Science and Technology.

References

- L. Sun and Y. Li, "Multi-target pig tracking algorithm based on joint probability data association and particle filter," *International Journal of Agricultural and Biological Engineering*, vol. 14, no. 3, pp. 199–207, 2021.
- [2] M. Yosra, L. M. Najeh, J. Imed, and B. Tahar, "Proposal of a player multi-target tracking algorithm applied in teams sport," *International Journal of Advanced Scientific Research and Innovation*, vol. 5, no. 1, pp. 16–27, 2022.
- [3] J. Li, W. Zhan, Y. Hu, and M. Tomizuka, "Generic tracking and probabilistic prediction framework and its application in autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3634–3649, 2020.
- [4] C. Zhu, "Exploring the role of sports APP in (campus fitness) intelligent solutions using data fusion algorithm and internet of things," *International Journal of Grid and Utility Computing*, vol. 13, no. 1, pp. 40–48, 2022.
- [5] M. A. Razzaq, J. M. Quero, I. Cleland et al., "uMoDT: an unobtrusive multi-occupant detection and tracking using robust Kalman filter for real-time activity recognition," *Multimedia Systems*, vol. 26, no. 5, pp. 553–569, 2020.
- [6] A. Goldhoorn, A. Garrell, R. Alquézar, and A. Sanfeliu, "Searching and tracking people with cooperative mobile robots," *Autonomous Robots*, vol. 42, no. 4, pp. 739–759, 2018.
- [7] S. J. Bhat and K. V. Santhosh, "A localization and deployment model for wireless sensor networks using arithmetic optimization algorithm," *Peer-to-Peer Networking and Applications*, vol. 15, no. 3, pp. 1473–1485, 2022.
- [8] Q. Zhou, B. Zhong, Y. Zhang, J. Li, and Y. Fu, "Deep alignment network based multi-person tracking with occlusion and motion reasoning," *IEEE Transactions on Multimedia*, vol. 21, no. 5, pp. 1183–1194, 2019.
- [9] A. Prati, C. Shan, and K. I. K. Wang, "Sensors, vision and networks: from video surveillance to activity recognition and health monitoring," *Journal of Ambient Intelligence and Smart Environments*, vol. 11, no. 1, pp. 5–22, 2019.
- [10] C. Sun and D. Ma, "SVM-based global vision system of sports competition and action recognition," *Journal of Intelligent Fuzzy Systems*, vol. 40, no. 2, pp. 2265–2276, 2021.
- [11] S. Liu, L. Liu, J. Tang, B. Yu, Y. Wang, and W. Shi, "Edge computing for autonomous driving: opportunities and challenges," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1697–1716, 2019.
- [12] Y. Yang, "Research on basketball sports neural network model based on nonlinear classification," *Journal of Intelligent Fuzzy Systems*, vol. 40, no. 4, pp. 7567–7576, 2021.
- [13] K. Zhao, W. Jiang, X. Jin, and X. Xiao, "Artificial intelligence system based on the layout effect of both sides in volleyball matches," *Journal of Intelligent Fuzzy Systems*, vol. 40, no. 2, pp. 3075–3084, 2021.
- [14] C. Cuevas, D. Quilon, and N. García, "Techniques and applications for soccer video analysis: a survey," *Multimedia Tools* and Applications, vol. 79, no. 39-40, pp. 29685–29721, 2020.
- [15] D. Simoes, N. Lau, and L. P. Reis, "Exploring communication protocols and centralized critics in multi-agent deep learning,"

Integrated Computer-Aided Engineering, vol. 27, no. 4, pp. 333–351, 2020.

- [16] V. Koolwal and K. K. Mohbey, "A comprehensive survey on trajectory-based location prediction," *Iran Journal of Computer Science*, vol. 3, no. 2, pp. 65–91, 2020.
- [17] M. E. Atik, Z. Duran, and R. Özgünlük, "Comparison of YOLO versions for object detection from aerial images," *International Journal of Environment and Geoinformatics*, vol. 9, no. 2, pp. 87–93, 2022.
- [18] C. Lim, J. Gwak, and M. Jeon, "Configuration of a min-cost flow network for data association in multi-object tracking," *International Journal of Computational Vision and Robotics*, vol. 8, no. 6, pp. 572–590, 2018.
- [19] D. Xianpeng, "Visual recognition and performance prediction of athletes based on target tracking EIA algorithm," *Journal of Intelligent Fuzzy Systems*, vol. 40, no. 4, pp. 7233–7246, 2021.
- [20] Y. Yan, J. Ren, H. Zhao et al., "Cognitive fusion of thermal and visible imagery for effective detection and tracking of pedestrians in videos," *Cognitive Computation*, vol. 10, no. 1, pp. 94–104, 2018.