Research Article

A Novel Automatic Tracking Method of Moving Image Sequence Marker Points Uses Kinect and Wireless Network Technology

Guoliang Chen and Lin Wei

Luohe Medical College, Luohe City, Henan, China 462000

Correspondence should be addressed to Guoliang Chen; 0342@lhmc.edu.cn

Received 3 February 2022; Revised 20 February 2022; Accepted 21 February 2022; Published 26 March 2022

1. Introduction

In recent years, detecting, identifying, and tracking interested human parts from sports image sequences are one of the hot topics in the field of image processing and computer vision, and it has a very broad application prospect. Sports science is a new science to study sports, and sports biomechanics is one of its branches. With the rapid development of mass sports, information technology, represented by computer technology, has also made rapid progress, with the rapid popularization of personal computers and the rapid improvement of computing speed. People are amazed at its intelligence and low price, and at the same time, they have made new explorations in sports assistance [1, 2]. Using various technical means to collect various sports biomechanical parameters and variables is an important content of sports biomechanics research. The current sports education system has a long way to go in comparison to the ideal requirements. Currently, the majority of existing systems use human motion measurement and analysis methods based on high-precision video capture and analysis. Humans’ primary means of obtaining information from nature is through vision. According to statistics, humans acquire about 60% of their information visually, and images are the primary means of acquiring visual information [3]. The so-called “graph” [4] refers to the distribution of light transmitted or reflected by an object. “Image” is the impression or recognition that the human visual system receives the information of the picture in the brain. Target tracking is a key technology in the field of computer vision, such as intelligent monitoring, traffic statistics, and video analysis of sports information. By accurately tracking the motion trajectory of the target, a high-level semantic analysis and understanding of a sports video can be carried out [5, 6]. In video tracking, firstly, the image sequence of moving objects should be acquired, and the analysis of continuous video moving objects in reality should be transformed into the analysis of target points in discrete image sequences. The analysis process includes the following main steps: (1) extracting moving objects from the image background, (2) mark and locate the position of the target point in the moving image sequence, and (3) track these target points. Among them, the
identification, positioning, and tracking of moving targets are a key step and the basis of image analysis and understanding.

As a low-cost and reliable unmarked deep motion capture technology, Microsoft Kinect sensor has attracted the attention of motion research circles. Combined with the advantages of Kinect collector of Microsoft in image acquisition, this paper constructs a general sports auxiliary training system and designs some functions of the system in detail [7, 8]. So far, as a general depth camera, Kinect has been applied in the field of sports training to realize human-computer interaction. Scholars at home and abroad have also done a lot of research in these aspects, mainly including recognition [9], bone tracking, and depth image. The Kinect somatosensory camera is used for bone tracking to obtain the spatial coordinate information of human nodes in the process of teaching, and the teaching data of imitation learning is obtained after preprocessing. In order to overcome camera jitter and background jitter, symmetrical difference and neighborhood filtering are used to locate moving targets. Finally, the moving target region is extracted by projection algorithm, and the clutter block removal and region merging are adopted to make the extracted moving target region more accurate and continuous [10].

The joints captured by Kinect are used to recognize human posture, so as to remotely monitor patients’ yoga posture. This paper discusses the application of somatosensory interaction in national physical fitness test and self-training with Kinect sensor [11, 12]. Using the depth image collected by Kinect, the height of the horizontal bar is determined according to the field of view and the depth value. Using bone tracking technology to determine the position of the jaw, the bending angle of the arm is calculated according to the three joint points of the arm. Count and score this movement according to the distance between jaw and crossbar and the angle of arm bending. Combine Kinect cheap and powerful somatosensory equipment with mass sports to create a brand-new sports teaching method and promote the rapid development of mass sports [13, 14]. We firmly believe that through our persistent efforts, we will definitely realize a high-quality sports teaching system that meets the ideal conditions. Using Kinect to collect the information of bone points, establish a virtual coordinate system, get the joint vectors from the node information, calculate the included angle between the vectors, and then get the control angle of each joint, so as to obtain the motion information of each joint and provide data support for the representation learning [15] of imitation.

2. Related Work

The sports image sequence marker points are applied to the sports video analysis system, according to literature [16]. The first project involves using the moving target tracking method to automatically place joint points in human motion sequence images. Automatically determining the position of joint points during human movement by computer will greatly reduce the time spent by operators in video analysis, as well as avoid errors caused by operators’ analysis fatigue and a variety of other factors. According to the literature [17], if more detailed information about the shape or appearance of the sports image is available before tracking, an a priori model for the target can be created, and then, the target can be actively searched in the sequence of sports images to achieve the tracking goal. These models can be used to predict the target’s possible performance in the sequence, reduce the feature space to be searched, and improve the algorithm’s robustness in the target tracking process. Because the human body has uniform chromaticity and no obvious feature points and various joint parts of the human body deform to varying degrees during movement, the general sports image processing method applied to the recognition of human joint points has some difficulty in target recognition and position positioning, according to the research in reference [18]. The problem of human joint location has yet to be solved perfectly. Literature [19] proposed that active contour method is a unique algorithm, which is applied in many places of image processing, such as using snake for target detection and segmentation. Firstly, the initial contour line is marked manually near the boundary of the target, and the internal force, external force, and binding force acting on the contour line are defined. By minimizing the curve energy, the curve can actively move to the real contour of the target of interest. Reference [20] proposed a target detection algorithm scheme of automatic clustering detection and outlier detection. Sui Hua and others proposed a double difference target detection algorithm based on the detection and classification of moving targets in traffic monitoring system and the principle of maximum estimation; VSAM, a major visual surveillance project led by Carnegie Mellon University, has developed a hybrid algorithm combining adaptive background subtraction method and interframe subtraction method. Literature [21] shows that feature-based tracking algorithms generally adopt correlation algorithms. Unlike region-based tracking algorithms, region-based tracking uses the whole marker points of sports image sequences as the relevant objects, while feature-based tracking uses one or some local features of the target as the relevant objects. The balloon model of active contour is presented in the literature [22]. These models increase the initial contour’s capture area so that it does not have to be close to the object of interest’s real boundary, and they reduce the sensitivity to contour initialization [23]. The landmark points of sports image sequence are applied to the rapid feedback system of sports, which is primarily manifested in video image performance, using the method of big data analysis. In sports image processing technology, for example, segmentation techniques for moving objects include superposition comparison of different images and videos and superposition display of the same video. At present, all the fast feedback systems have the above functions. It is very helpful for athletes to understand and improve technical movements by visually displaying the technical differences of different athletes or the technical movements of the same athlete at different times through sports images. Literature [24] proposed that MSEA algorithm can quickly judge whether the candidate position needs matching operation by comparing the gray information of template with
that of local image. Literature [25] proposes that the combination of segmentation based on optical flow and depth information of video can realize the tracking of rigid and nonrigid targets. Lu et al. combined edge, optical flow, and shadow information to optimize the model in tracking and realized the tracking of human hand under changing conditions.

This paper studies the automatic tracking method of sports image sequence marker points based on Kinect. In the early stage of sports image research, the process of target point positioning and tracking is manual. Due to the large number of images to be processed, manual positioning and tracking take a lot of time, which limits the analysis of a large number of images. The process is very complex, and the positioning accuracy is low. It is not suitable for the analysis of moving video images. Therefore, the automatic fast positioning method can be adopted in the processing process, which will greatly speed up the processing speed.

3. Principle and Algorithm of Kinect

The template matching method is mainly used to realize fast positioning in the moving image sequence. Its main idea is that in the analysis process, the target point position of the first frame image is manually input as a priori knowledge, and the subsequent image processing uses the priori knowledge of the target position in the first frame image to accurately register the moving target points in the front and back frames. The accurate position of the moving target in the post frame image is obtained, and then, the intelligent computer automatic positioning of the moving target point in the image is realized to improve the processing speed and accuracy. The optical flow based tracking algorithm assumes that the gray values of the corresponding pixels of the target in the front and back frames are the same. Through the calculation of the optical flow field, the motion field of the pixels in the image can be obtained, which can be used to guide the next tracking. The feature point optical flow method obtains the optical flow vector at the feature point through feature matching. Compared with the global optical flow method, this algorithm has the characteristics of small amount of calculation, fast, and flexible. Optical flow method can track moving targets well in dynamic background, but due to the aperture and occlusion problems, the solution of estimating two-dimensional motion field by optical flow method is unstable, and additional hypothetical models need to be used to simulate two-dimensional motion field. The color image collected by color camera and the depth image collected by infrared camera are used. The Kinect algorithm is applied to motion comparison analysis. The steps of the proposed method are shown in Figure 1.

Each feature point direction can specify one or more direction parameters based on the gradient direction characteristics of the local image of feature points. Sample in the feature point neighborhood window first, then compute the gradient direction histogram of each feature point neighborhood pixel, and finally compute the contribution value of each neighborhood pixel. The main direction of feature points is determined by the gradient histogram’s peak value; the secondary direction is determined by the 80 percent peak value. First, the affine invariant vector construction method and the occlusion criterion based on it are presented. Then, using prior knowledge of the tracked target, a trajectory and occlusion area prediction method is discussed in detail. On this basis, the affine invariant vector-dependent occlusion criterion for moving targets is modified. It can also be divided into single target tracking and multitarget tracking depending on the number of tracking targets. The research object of sports image sequence marker processing is transferred from still image to sports image sequence marker using the Kinect algorithm. It is possible to know and analyze the dynamic process by analyzing multiple sports images, in order to obtain information that cannot be obtained by a single sports image sequence marker. Using the depth image collected by Kinect, the height of the horizontal bar is determined according to the field of view and depth value. For example, as far as tracking objects are concerned, there are tracking body parts such as the hands, faces, heads, and legs and tracking the whole human body. Automatic tracking method is to monitor the spatio-temporal changes of objects in the landmark points of sports image sequence, which includes the existence, position, size and shape of objects, and so on. The system uses Microsoft’s Kinect sensor to extract the information in the image for data analysis, count, and score the pull-up movement, so as to realize the functions of automatic test and self-training. The system mainly includes three modules: cross bar position recognition module, user bone tracking module, and counting and scoring module. At the same time, users can evaluate their actions according to the action playback and scoring of video, so as to realize the standardization of actions. The system block diagram is shown in Figure 2.

Kinect obtains the depth map by emitting near-infrared light source. Kinect will actively track large-shaped objects whether there is enough light or not. Then, the target detection algorithm will search the whole field of view for the marker points of the sports image sequence to remove the marker points of the background sports image sequence and extract multiple moving object regions, that is, generate a differential binary image with the foreground as the moving object region. The three-dimensional coordinate information of bone points is obtained by Kinect, and joint vectors are generated. The joint control angles are obtained by calculating the included angles of the joint vectors. For the joint angles of RElbowRoll and RElbowYaw, the bone points that Kinect needs to extract are Wrist_R, Elbow_R, Shoul-der-R, and Spine-Shoulder, and their three-dimensional coordinate information is \( (x_1, y_1, z_1) \) and \( (x_3, y_3, z_3) \), respectively.

\[
\begin{align*}
    P_2P_1 = P_1 - P_2 = (x_1 - x_2, y_1 - y_2, z_1 - z_2), \\
    P_2P_3 = P_3 - P_2 = (x_3 - x_2, y_3 - y_2, z_3 - z_2), \\
    P_3P_4 = P_4 - P_3 = (x_4 - x_3, y_4 - y_3, z_4 - z_3). 
\end{align*}
\]
Then, the joint angle RElbowRoll is

$$\theta_{\text{RElbowRoll}} = \arccos \left( \frac{P_2 \cdot P_3}{|P_2||P_3|} \right).$$

Note that the vector $P_3 \times P_2 = P_3 \times P_4 \times P_3$, then the

The joint angle RElbowYaw is

$$\theta_{\text{RElbowYaw}} = \arccos \left( \frac{P_3 \cdot P_2}{|P_3||P_2|} \right) + \frac{\pi}{2}.$$

For any degree of freedom, let the $j$th teaching data point

Figure 1: Flow chart of Kinect method.

Figure 2: Kinect system framework diagram.
be $\xi_j = \{\xi_{ij}, \xi_{ij}\}$ and $j = \{1, 2, \cdots, N\}$, where $n$ is the number of data points contained in a single teaching, $\xi_{ij}$ is the joint angle, and $\xi_{ij}$ is the time value. It is assumed that each data point EE follows the following probability distribution

$$p(\xi_j) = \sum_{k=1}^{K} p(k)p(\xi_j|k),$$  \hspace{1cm} (4)

where $P(k)$ is a priori probability and $p(\xi_j|k)$ is a conditional probability distribution and obeys Gaussian distribution. Therefore, the whole teaching data set can be represented by Gaussian mixture model, and $K$ is the number of Gaussian distributions constituting Gaussian mixture model

$$p(k) = \pi_k,$$  \hspace{1cm} (5)

$$p(\xi_j|k) = N(\xi_j; \mu_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^d|\Sigma_k|}},$$  \hspace{1cm} (6)

where $d$ is the dimension of GMM encoding teaching data. Therefore, the parameter to be determined in Gaussian mixture model is $\{\pi_k, \mu_k, \Sigma_k\}$, which, respectively, represents the prior probability, expectation, and variance of the $k$th component. The EM algorithm is used to estimate the parameters of GMM, and the maximum likelihood estimation of parameters is found in the probability model to learn the parameters.

After the feature matching of the action sequence by Kinect algorithm, the key point images of two groups of action sequences are obtained. According to the characteristics of the action, we use two ways to segment the action, which are as follows:

(1) Equal Time Interval Action Segmentation Method. Segment according to the time sequence of two images matched by the characteristics of two groups of action sequences, that is, each action sequence matching image is divided into five parts on average. Specifically, the first image is evenly distributed according to the horizontal coordinate interval and divided into five parts. The division of the second image is the same as above

(2) Action Segmentation Method with Equal Key Point Proportion. Segment according to the principle that the proportion of the corresponding segment key points of the two images matched by the characteristics of the two groups of action sequences is the same. The first image is evenly distributed and divided into five parts using a specific method based on the horizontal coordinate interval. Calculate the proportion of the number of key points in each segment in the first image to the total key points found, multiply this proportion by the total key points found in the second image to get the boundary coordinates of each segment in the second image, and then divide the segmentation of the second image by the corresponding key point interval

4. Research on Automatic Tracking Method of Sports Image Sequence Markers

4.1. Kinect-Based Automatic Tracking Method of Sports Image Sequence Markers. Automatic tracking method of sports image sequence landmarks based on Kinect is an important research topic of machine vision, which is widely used in intelligent monitoring, video coding, and artificial intelligence. Among the environmental information perceived by people, visual information accounts for a very large proportion, among which dynamic visual information is the main component, and a large number of meaningful visual information are included in sports. Based on Kinect algorithm, the research object of sports image sequence marker processing is transferred from still image to sports image sequence marker. By analyzing multiple sports images, it is possible to know and analyze the dynamic process, so as to obtain information that can not be obtained by a single sports image sequence marker. In order to solve the problem that objects and objects, objects and scenes, and scenes and scenes are often occluded in image sequences under complex background, this paper presents an occlusion detection and tracking method for multiple moving objects based on affine invariants. Firstly, the construction method of affine invariant vector and the occlusion criterion based on it are given. Then, combined with the prior knowledge of the tracked target, a prediction method of trajectory and occlusion area is discussed in detail. On this basis, the occlusion criterion of moving target which depends on affine invariant vector is modified. As far as tracking angles are concerned, there are single angles corresponding to a single camera, multiple angles corresponding to multiple cameras, and omni-directional angles. Kinect algorithm, as a concise nonparametric density estimation method, has been widely used in the field of real-time target tracking in recent years because of its high efficiency in feature space search.

The rapid advancement of the Kinect algorithm in hardware technology has resulted in significant improvements in the storage and processing speed of sports image sequence marker points, as well as strong support for sports image sequence marker point analysis. To begin, the input sports image sequence marker point data is preprocessed in order to improve the quality of the markers. The target detection algorithm will then search the entire field of view for marker points from the sports image sequence in order to remove marker points from the background sports image sequence and extract multiple moving object regions, resulting in a differential binary image with the foreground as the moving object region. The author proposes a new method for extracting moving target regions from image sequences. To begin, the image sequence is differentiated; the domain value for binarizing the difference image is automatically selected based on an analysis of the histogram of the difference result image. To locate moving targets, symmetrical difference and neighborhood filtering are used to overcome camera jitter and background jitter. Finally, using a projection algorithm, the moving target region is extracted, and clutter block removal and region merging are used to make the extracted...
moving target region more accurate and continuous. The target continuous tracking stage is primarily investigated using the prediction matching correction algorithm framework. To begin, the trajectory prediction algorithm is used to forecast the target’s possible position in the next frame. The target template obtained during the detection and positioning stage is then used for location matching within the prediction-based search range. The optimal result obtained by matching is used to determine the target’s position in this frame. The corresponding parameters must be adjusted in order to avoid affecting the effect of target optimal position matching. Furthermore, when using a position prediction algorithm based on trajectory fitting, if the target is mobile throughout the movement process, the linear prediction algorithm is preferable; otherwise, the square prediction algorithm can be used. As a result, the specific algorithm scheme for target trajectory prediction must be chosen in conjunction with the above optimal selection principles, based on the actual application scenario and target motion characteristics.

There are numerous methods for tracking research and classification available today. In terms of tracking objects, there are single angles corresponding to a single camera, multiple angles corresponding to multiple cameras, and omni-directional angles; in terms of tracking angles, there are single angles corresponding to a single camera, multiple angles corresponding to multiple cameras, and omni-directional angles. Because of its high efficiency in feature space search, the Kinect algorithm has become widely used in the field of real-time target tracking in recent years as a concise nonparametric density estimation method. The purpose of the automatic tracking method is to track the spatio-temporal changes of objects in the landmark points of a sports image sequence, such as their existence, position, size, and shape. However, because the background in a real-world test might be quite complicated, it is necessary to threshold the depth image, separate the human body from the background, and then analyze the human motion. To begin, consider using face recognition technology to determine the position of the mandible. However, when pulling up, the occlusion of the crossbar will prevent face recognition, and thus, the positional relationship between the mandible and the crossbar will be inaccurately judged when the human body reaches its highest point. The Kinect depth image can be used to calculate the distance between the user and the Kinect camera. However, because the background in a real-world test might be quite complicated, it is necessary to threshold the depth image, separate the human body from the background, and then analyze the human motion. The target tracking becomes searching for a corresponding target in the current frame, which minimizes the distance function between the target description centered on Y and the original target, using the template description of the target and candidate target of the sports image sequence and the criterion for measuring their similarity. The project uses bones to track the user’s joint points and compares the position of the user’s mandibular point with the position of the determined crossbar to judge whether the user’s mandibular crosses the crossbar due to the above problems with face tracking and depth image. It is counted once when the mandibular crosses the crossbar twice. The author proposes a new method for extracting moving target regions from image sequences. To begin, the image sequence is differentiated; the domain value for binarizing the difference image is automatically selected based on an analysis of the histogram of the difference result image. The starting point for target search is the position of the target in the previous frame, and it searches the area around it.

4.2. Experimental Results and Analysis. In this paper, the Kinect algorithm is used to simulate and compare the predictions of the two prediction algorithms under different landmarks of sports image sequences, and the applicable scenes and optimal selection principles of the algorithms are summarized according to the experimental results. Target features are extracted from the image, such as lines, curves, areas, and surfaces of the target. The advantage of perspective-based modeling lies in that, due to its mathematical generality, it is possible to deal with all kinds of variability mentioned above within a consistent framework, and its computational complexity is much smaller than that of 3D expression. This method of expressing targets can be used to track objects with linked structures, such as human hands or torso. The simulation experiment mainly uses 3-point linear predictor and 5-point square predictor to test the prediction algorithm based on trajectory fitting. In the experiment of Kalman filter prediction algorithm, the state vector estimation $x(0)$, the initial value of error covariance matrix $P(0)$, the transition matrix $A$, and the observation matrix $H$ can be determined in advance as described in Section 4.1.3 of this chapter. In the experiment, three simulation experiments were conducted to compare the three trajectories. The experimental results are shown in Figures 3–8, respectively.

From the above simulation results, it can be seen that the Kalman prediction algorithm has the best effect in the prediction in Figures 3–5. The predicted trajectory is smooth, and the relative error is small, which fully reflects its robustness. In the case of large signal-to-noise ratio, the absolute value of prediction error of linear prediction algorithm is less than 1, and its actual prediction effect can be compared with Kalman prediction algorithm. The effect of the square predictor is poor, especially when the signal-to-noise ratio is low, and its error is obvious. Experiments show that the square predictor fails to predict straight lines and target trajectories with significant random motion. The overall effect of the Kalman prediction algorithm is still good in Figures 6–8, but it is unstable in the initial prediction stage and returns to normal after the observed values are corrected. As a result, the accuracy of the Kalman prediction algorithm’s first few prediction points is low. The square predictor’s prediction error is currently lower than the linear predictor’s. It has been proven that the square predictor is better than the linear predictor at predicting conic motion. A signal similarity-based global motion estimation method between image sequences is proposed. This method divides the image using a new quadtree-like structure, then gradually focuses on segmenting the image reference structure that
Figure 3: Relative error of linear prediction.

Figure 4: Relative error of square prediction.
meets the uniqueness requirements, integrating the reference structure selection and image rough matching, so that the reference structure selection process is independent of the image segmentation result and avoids additional overhead. The optimal selection principle of target trajectory prediction algorithm can be summarized as follows: if it can

Figure 5: Relative error of Kalman prediction.

Figure 6: Linear trajectory prediction.
meet the real-time requirements of the system, the Kalman filter prediction algorithm will have better prediction accuracy than the other two algorithms, especially in the case of high randomness of target motion or large scene noise interference. If the real-time requirement of the system is high, the prediction algorithm based on trajectory fitting with small amount of calculation, simple structure, and easy implementation is mainly considered.
5. Conclusions

It covers a wide range of topics, including computer science, optics, mathematics, cognitive science, and control science, among others. One of the most fundamental and difficult topics in computer vision is moving object detection and tracking in image sequences. The study of it is both practical and theoretically important. The feature extraction of sports image sequence landmark points constructed in this paper has certain feasibility of automatic tracking, can extract the joint information of the moving human body, and can compare the differences between standard movements and sports movements with the help of an auxiliary system in the process of extracting human sports image sequence landmark points. Bone tracking is used to determine the user’s jaw position. Count and score using the distance between the mandible and the crossbar, as well as the arm’s bending angle, as per the national physical fitness test standard. First, the marking points of a sports image sequence are differentiated; then, based on an analysis of the difference result image’s histogram, the Kinect algorithm automatically selects a threshold value to binarize the difference image. Finally, the Kinect algorithm uses a projection algorithm to extract moving target areas and then uses clutter removal and region merging to improve the accuracy and consistency of the extracted moving target areas. As a result, the specific algorithm scheme for target trajectory prediction must be chosen in conjunction with the above optimal selection principles, based on the actual application scenario and target motion characteristics.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

All the authors do not have any possible conflicts of interest.

Acknowledgments

This work was supported by the Countermeasures for the Cultivation of college students’ exercise habits—a case study of Luohe Medical College Research Project of Innovation, Entrepreneurship, Development capacity Improvement Engineering Education and Teaching Reform of Luohe Medical College in 2019 (Project No. 2019-LYZJGYB020).

References

[1] M. Eltoukhy, J. Oh, C. Kuenze, and J. Signorile, “Improved Kinect-based spatiotemporal and kinematic treadmill gait assessment,” Gait & Posture, vol. 51, pp. 77–83, 2017.
[2] H. Novatchkov and A. Baca, “Artificial intelligence in sports on the example of weight training,” Journal of sports science & medicine, vol. 12, no. 1, pp. 27–37, 2013.
[3] J. Zhang, W. Feng, T. Yuan, J. Wang, and A. K. Sangaiha, “SCSTCF: Spatial-Channel Selection and Temporal Regularized Correlation Filters for visual tracking,” Applied Soft Computing, vol. 118, article 108485, 2022.
[4] Y. Ding, X. Zhao, Z. Zhang, W. Cai, and N. Yang, “Graph sample and aggregate-attention network for hyperspectral image classification,” IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, 2022.
[5] J. Jianwei, P. Shijie, S. Parker, and J. Zhao, “3D occlusion target tracking based on all-optical image sequence,” Engineering, vol. 57, no. 6, pp. 061604.1–061604.8, 2018.
[6] M. Oshita, “Generating animation from natural language texts and semantic analysis for motion search and scheduling,” The Visual Computer, vol. 26, no. 5, pp. 339–352, 2010.
[7] J. Lamerca, S. Parashar, and A. Bartoli, “DeSLAM: tracking and mapping of deforming scenes from monocular sequences,” IEEE Transactions on robotics, vol. 37, pp. 291–303, 2021.
[8] S. Pinet, G. S. Dell, and F. X. Alario, “Tracking keystroke sequences at the cortical level reveals the dynamics of serial order production,” Journal of Cognitive Neuroscience, vol. 31, no. 7, pp. 1030–1043, 2019.
[9] J. Kong, H. Wang, X. Wang, X. Jin, X. Fang, and S. Lin, “Multi-stream hybrid architecture based on cross-level fusion strategy for fine-grained crop species recognition in precision agriculture,” Computers and Electronics in Agriculture, vol. 185, article 106134, 2021.
[10] Y. Ren, B. Xu, P. Zhu, M. Lu, and D. Jiang, “A multiclass visual tracking algorithm using multi-task particle swarm optimization for low-contrast image sequences,” Applied Intelligence, vol. 45, no. 4, pp. 1129–1147, 2016.
[11] X. Qian, Y. Wang, and H. Lei, “An object tracking method based on guided filter for night fusion image,” Infrared Physics and Technology, vol. 74, pp. 38–43, 2016.
[12] L. Wei, “Coastal land use planning and beach motion image recognition based on high-resolution remote sensing images,” Arab Journal of Geosciences, vol. 14, no. 11, pp. 1–14, 2021.
[13] X. Zhang and F. Q. Yang, “Machine learning-based multitarget tracking of motion in sports video,” Complexity, vol. 2021, Article ID 5533884, 10 pages, 2021.
[14] C. Duan, “Deep learning-based multitarget motion shadow rejection and accurate tracking for sports video,” Complexity, vol. 2021, no. 1, Article ID 5973531, p. 11, 2021.
[15] Y. Zhang, K. Xia, Y. Yang et al., “Multi-modality fusion & inductive knowledge transfer underlying non-sparse multi-kernel learning and distribution adaption,” IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. PP, p. 1, 2022.
[16] D. Rozumnyi, J. Kotera, F. Roubek, and K. Aimee, “Tracking by degating,” International Journal of Computer Vision, vol. 1, pp. 1–22, 2021.
[17] Z. Zhao, X. Liu, and X. She, “Tracking model of competitive sports function training based on artificial intelligence,” Journal of Intelligent Fuzzy Systems, vol. 40, no. 1, pp. 3347–3359, 2021.
[18] M. Feigin, A. Bhandari, S. Izadi, C. Rhemann, M. Schmidt, and R. Raskar, “Resolving multipath interference in Kinect: an inverse problem approach,” IEEE Sensors Journal, vol. 16, no. 10, pp. 3419–3427, 2016.
[19] X. Chang, Z. Ma, M. Lin, Y. Yang, and A. G. Hauptmann, “Feature interaction augmented sparse learning for fast Kinect motion detection,” IEEE Transactions on Image Processing, vol. 26, no. 8, pp. 3911–3920, 2017.
[20] R. A. Hauser, S. A. Factor, S. R. Marder et al., “KINECT 3: a phase 3 randomized, double-blind, placebo-controlled trial of valbenazine for tardive dyskinesia,” The American Journal of Psychiatry, vol. 174, no. 5, pp. 476–484, 2017.

[21] J. R. Terven and D. M. Córdova-Esparza, “Kin2. A Kinect 2 toolbox for MATLAB,” Science of Computer Programming, vol. 130, pp. 97–106, 2016.

[22] P. Plantard, H. P. H. Shum, A. S. Le Pierres, and F. Multon, “Validation of an ergonomic assessment method using Kinect data in real workplace conditions,” Applied Ergonomics, vol. 65, pp. 562–569, 2017.

[23] Y. Sun, C. Li, G. Li et al., “Gesture recognition based on Kinect and sEMG signal fusion,” Mobile Networks and Applications, vol. 23, no. 4, pp. 797–805, 2018.

[24] I. J. Ding and Z. G. Wu, “Combinations of eigenspace and GMM with Kinect sensor-extracted action gesture features for person identity recognition,” Engineering Computations, vol. 33, no. 8, pp. 2489–2503, 2016.

[25] R. Spoliansky, Y. Edan, Y. Parmet, and I. Halachmi, “Development of automatic body condition scoring using a low-cost 3-dimensional Kinect camera,” Journal of Dairy Science, vol. 99, no. 9, pp. 7714–7725, 2016.