Research Article

Knee Bending Motion Extraction Algorithm for Ball Motion Image Using Vision Sensor

Zhanshu An, Zhigang Yuan, Xu Cao, and Shijia Song

College of Physical Education and Training, Harbin Sport University, Harbin 150008, China

Correspondence should be addressed to Zhigang Yuan; yuanzhigang@hrbipe.edu.cn

Received 7 February 2022; Revised 25 February 2022; Accepted 26 February 2022; Published 8 March 2022

Academic Editor: Mohammad Farukh Hashmi

Copyright © 2022 Zhanshu An et al. 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.

To improve the extraction accuracy of knee bending motion in ball motion image and reduce the extraction distance error and time consumption, a knee bending motion extraction algorithm using visual sensor is proposed. The visual sensor model is constructed based on the ball motion frame image, the trigger data is output through differential and logical judgment, and these data are normalized to generate the visual sensor sample set of the ball motion frame image. The sample set is used as the input of the convolution neural network (CNN) and the sample basis of the motion energy model. The CNN extracts the features of the sample set in the convolution layer, the motion energy model is combined with the local binary pattern to extract the features of the sample set, the weighted summation method is used to fuse the two features, and the Softmax classifier is used to classify and extract the knee bending motion. The results show that the proposed algorithm has good ball motion image collection effect, the knee bending motion extraction accuracy is always maintained at about 98%, the distance error is low, and the time consumption of ball motion feature extraction is only 2.65 s, which has high application value in the field of sports.

1. Introduction

With the vigorous development of China’s sports industry, all kinds of ball motion have made gratifying achievements in competitive competitions. However, the body injury of athletes is also becoming more and more serious. Therefore, collecting the features of athletes’ knee bending motion by scientific means is conducive to the relevant departments to formulate the training plan of ball players [1]. In the existing research, many scholars have made relevant research. Literature [2] proposed a motion extraction algorithm based on image length and Doppler broadening. This algorithm can extract motion, but the accuracy still needs to be further improved. Literature [3] takes the time channel of video as the research object to construct an action extraction algorithm. The algorithm can consider the features of video image and realize the extraction of motion action, but the extraction effect of knee bending action is not good. Literature [4] takes the video data set as the research sample to carry out motion extraction, but similar to the above algorithm, the accuracy of extracting detailed actions is not high and still needs to be further improved. Literature [5] uses the method of deep learning to extract motion, but the algorithm does not achieve good results in the image processing stage, and the image definition is insufficient, resulting in the impact of extraction accuracy. Literature [6] uses semisupervised learning to realize action extraction. The algorithm is still in the research stage and needs further verification. The vision sensor is actually a bionic device, which is conducive to the event driven action during motion extraction, and the sensor has high dynamic range and superior comprehensive performance. Based on the visual sensor, this paper studies ball motion and constructs a special algorithm for extracting knee flexion, which points out the direction for the formulation of athletes’ training plan in the future. In order to further improve the knee bending motion extraction accuracy of ball motion image and reduce the extraction distance error and time consumption, a knee bending motion extraction algorithm based on visual sensor is designed. The main contributions of this paper are as follows: (1) Vision sensor is a bionic device, which is conducive to the event driven action during motion extraction. The
sensor has high dynamic range and superior comprehensive performance, which can improve the efficiency and accuracy of ball motion image sample collection. (2) Using CNN to extract the features of ball motion images can lay a solid foundation for the subsequent knee flexion extraction. (3) Softmax classifier is used to classify and extract knee bending motion, which solves the problem of fuzzy category in the process of knee bending motion extraction of ball motion image.

2. Methodology

2.1. Construction of Sample Set Model Based on Visual Sensor. In order to accurately extract the knee bending action in the ball motion image, a visual sensor model is constructed based on the frames in the ball motion image. The imaging process of the model is shown in Figure 1.

According to data in Figure 1, the complementary metal oxide semiconductor in the vision sensor uses the vision sensor array to sense the pixel circuit through the photodiode and quantify the light intensity [7], so that the pixels of the ball motion image are converted into electrical signals. In Figure 1, path 1 and path 2 are the imaging process of the visual sensor. The imaging principle is to add the “address event” expression as logic after each pixel unit to realize circuit judgment. Build the data set of ball motion frame image according to path 1 and path 2. In the process of building this kind of data set, it is necessary to take the pixel value and pixel information of the output ball motion frame image as the research goal and fuse an “address event” logic judgment module to the end of the visual sensor array to obtain the relevant image data set.

The visual sensor itself has superior event triggering ability, but when the visual sensor triggers an event, it mainly uses the analog circuit in the visual sensor to quantify, accumulate, and release the light intensity. In this process, it can only ensure that each pixel unit in the image can work independently [8, 9]. However, there is no concept of frame and high delay. In order to solve these problems and ensure that the ordinary ball motion frame image can use the visual sensor, this study takes the ball motion frame image as the object, discretizes the pixel units in the continuous ball motion frame image through the visual sensor, and ensures that each pixel unit in the motion frame image is independent of each other. That is, the differential and logical judgment is only for the current frame and the initial state of the pixel. The logical judgment process of the visual sensor data set model of the ball motion frame image is to judge whether the light intensity change amplitude of each pixel unit is close to the trigger threshold \( T \), and the judgment is based on the current frame and the state of the previous frame of the pixel unit.

The premise of event triggering is that the change amplitude of spherical moving image pixel unit has approached the threshold \( T \), and Equation (1) is the expression of this case:

\[
I_c(x, y) = I_p(x, y) \times (1 \pm T), \tag{1}
\]

where \( I_c(x, y) \) and \( I_p(x, y) \) are used to express two light intensities (original and current frame), respectively. \((x, y)\) is used to describe pixel unit coordinates. \( I_p(x, y) \) represents the light intensity of the pixel in the current frame.

The visual sensor data set model method of ball motion frame image is constructed in this paper. When processing ball motion frame image, the event trigger does not occur immediately when the pixel light intensity changes close to the threshold. Generally, the event is triggered after the logical judgment of the next frame is completed. The specific process is as follows:

1. In the process of model construction, the first frame of the ball motion image is set as the pixel light intensity in the original state, and the difference is made, and the pixel light intensity in the original state is logically judged to be the pixel light intensity in the next frame. If the result does not reach the trigger threshold, there is no need to activate the pixel unit.
2. Continue the difference and logically judge the light intensity of the original state pixel and the third frame light intensity of the same pixel.
3. According to the above rules, continue to differentiate and logically judge the light intensity of each frame. When reaching the frame, the obtained pixel unit light intensity first realizes the trigger threshold and outputs the trigger state. The pixel light intensity of the frame becomes the pixel light intensity of the new original state after output. Make a logical judgment on the pixel light intensity of subsequent frames according to the above steps, and the change amplitude of pixel light intensity is \( \Delta I(x, y) \), as shown in Equation (2).

\[
|\Delta I(x, y)| = \begin{cases} 
\log_2 \left( \frac{I_c(x, y)}{I_p(x, y)} \right) - 1 & \text{if } \geq T, I_p = I_c, \text{ Activate output} \\
< T, I_p = I_p, \text{ No output} 
\end{cases} \tag{2}
\]
If the trigger threshold $T$ is equal to the change amplitude $\Delta I(x, y)$, the pixel light intensity of the current frame of the ball moving image and the pixel light intensity of the original state can be output, and the pixel activation information can be output at the same time [10].

If you want to obtain the sample set of knee bending motion visual sensor of ball motion frame image, you need to first obtain the knee bending motion trigger data from the massive frame images of ball motion continuous video stream and first disperse the continuous video stream according to the time features to obtain the ball motion frame image. After that, the ball motion frame image is classified according to the cluster, the dynamic difference and logical judgment are completed in the cluster, the dynamic active region image is obtained, and these regions are converted into time address/frame data [11, 12].

Extracting knee flexion from ball motion images requires CNN for classification, but the inputs of such networks are static images [13], rather than the dynamic active area images obtained above. Therefore, this study establishes a static image simulation motion model. Random motion simulates the knee bending action in the static image and simulates the dynamic state of knee bending action by moving or enlarging and reducing the image. The static image simulation motion model is used to expand the static image sample set, and the visual sensor data set model difference and logic judgment of the ball motion frame image are used to determine the static image to simulate the motion image, and finally, the visual sensor sample set of the ball motion frame image is obtained.

2.2. Knee Bending Action Extraction. In order to extract the knee bending motion from the visual sensor samples of the ball motion frame image, while increasing the accuracy of extraction, the two parts of the CNN and the motion energy model algorithm are merged, and finally, the knee bending motion extraction of the ball motion image is completed.

2.2.1. Feature Extraction Based on CNN. The CNN includes an input layer, a convolution layer, a pooling layer, a fully connected layer, and an output layer. The visual sensor sample set of the ball motion frame image obtained is the input of the input layer, and the feature extraction is mainly in the convolution. The CNN used in this paper contains 10 convolutional layers in total. Ball motion images are randomly selected from the visual sensor sample set of ball motion frame images. Each image contains multiple neurons. A feature of each image is extracted through convolution filter in each convolution layer [14–16]. The convolution process is as follows:

Equation (3) describes the input of the second neuron in layer $l$:

$$a^{(l)} = f\left(w^{(l)} * a^{(l-1)} + b^{(l)}\right),$$

where $b$ represents bias and $w$ represents filter. The expression of cross entropy loss function is as follows:

$$J(y, f(x, \theta)) = -\sum_{i=1}^{C} y_i \log f_i(x, \theta),$$

where $y$ is used to express the forecast target category, $C$ is the set number of target categories, and $y_i = 1$ and $f(x, \theta)$ represent the probability of each category.

The gradient descent method is used for iterative solution, and the gradient of the convolution layer is calculated by Equation (5):

$$\frac{\partial J(W, b; x, y)}{\partial b^{(l)}} = \sum_{ij} \left(\delta^{(l+1)}\right)_{ij},$$

where $x$ and $k$ represent constant and dimension, respectively, $W$ represents weight, and $i$ and $j$ represent different gradient descent parameters.

In this study, a total of 10 convolution layers are used, and the pool layer is between each two convolution layers. According to the principle of local correlation, the pool processing of the images in the data set obtained can reduce the computational complexity and ensure that the original ball motion image will not be deformed after feature extraction. The full connection layer including 4107 neurons is used in this paper, and several different types of features are fused. The Softmax function is used in classification. This function realizes probability mapping in multiple categories. The $x$ input (visual sensor sample set of ball motion frame image) is described by $K$-dimensional vector. The mapping form is described by Equation (6):

$$z = \text{softmax}(x) = \frac{\exp(x)}{\sum_{i=1}^{K} \exp(x)}.$$
the ball motion frame image is represented by \( \text{map}_v(r, s, t) \). The motion energy model is as follows:

\[
\text{DMM} = \sum_{i=a}^{\text{end}} |\text{map}_i - \text{map}_i+1|.
\]

(7)

When using the motion energy model to extract features, it is necessary to segment the image [20], so as to obtain the depth energy of several different frames:

\[
\text{DMM}' = \frac{n}{L} \text{DMM},
\]

(8)

\[
n = 1, 2, 3, \cdots, L - 1,
\]

(9)

where \( \text{DMM}' \) represents the energy of the ball moving image of \( i \) frame and \( L \) stands for frame image sequence. The motion energy model cannot describe the local information of the moving image. Therefore, it is necessary to use the local binary mode LBP to extract the local features of the moving image [21, 22]. LBP operator can only cover part of the sphere moving image, but not all the frequency textures. Therefore, the extended LBP neighborhood is a circular neighborhood with \( R \) as the radius, in which there are many pixels. Suppose there is a pixel \( a \) coordinate is \( (x, y) \). The neighborhood coordinates of the pixel are described as Equation (10):

\[
(x_i, y_i) = \left( x + R \cos \left( \frac{2\pi i}{n} \right), y - R \sin \left( \frac{2\pi i}{n} \right) \right).
\]

(10)

Taking \( a \) as the central pixel, set the pixel as the threshold, compare the threshold with the gray value of adjacent pixels, and obtain the \( n \)-bit binary number. The value is calculated by Equation (11).

\[
\text{LBP}_{n,R} = \sum_{i=0}^{n-1} T_h(f(x_i, y_i) - f(x, y)) \ast 2^i,
\]

(11)

\[
T_h(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}
\]

(12)

where \( T_h(\cdot) \) represents the image gray threshold and \( f(x_i, y_i) \) and \( f(x, y) \) represent the gray image and the original image, respectively.

In order to obtain the texture information of ball motion image, binary mode needs to be used, and LBP feature is used to encode the motion energy model [23]. The detailed process is as follows: (1) the motion energy model template is divided into uniform grids, and the local binary mode features of each grid are counted; (2) concatenating each local binary pattern feature to obtain a local binary pattern feature descriptor; and (3) the row vector model is used to connect multiple motion energy model feature descriptors in series to obtain the final feature.

2.2.3. Algorithm of Feature Fusion Knee Bending Motion Extraction. In this study, the algorithm combines CNN and motion energy model and finally realizes the knee bending motion extraction of ball motion. Figure 2 shows the process of the proposed algorithm.

According to Figure 2, there are four key steps to realize knee bending action extraction:

1. **CNN Feature Extraction.** Take the visual sensor sample set of ball motion frame image as the input, set the size of each image in the sample set as \( 224 \times 224 \), feature extraction is carried out in the convolution layer, set \( 3 \times 3 \) as the convolution kernel size, and use back propagation and cross entropy loss function when optimizing CNN parameters.

2. **Feature Extraction of Motion Energy.** The energy of human for ball motion is obtained by using three-dimensional depth information, the three-dimensional coordinate system is constructed based on the visual sensor sample set of ball motion frame image, the motion energy model is constructed, and the features are obtained by using local binary mode.

3. **Fuse the Features of the Two Algorithms.** The weighted summation method is used in feature fusion. Set \( Z = (X_1, X_2, \cdots, X_n) \) to describe the weight vector, and Equation (8) describes the feature fusion \( Y_{\text{fusion}} \) model:

4. **The visual sensor sample set of ball motion frame image is divided into training set, verification set, and test set. The training set is used to train the feature fusion model, and the training of the completed model is obtained.**

5. **In the model, the verification set and test set are input, respectively, to obtain the model parameters and determine the model structure.**

3. Experimental Analysis and Results

3.1. Data Set. Two data sets are selected in the experimental analysis process of this paper. Data set 1 is a ball motion data set under complex background, from UCF Sports Action. The database contains 150 video sequences with a resolution of \( 720 \times 480 \), with a frame rate of 10 fps and an average of about 6.39 s for each video. The database contains a large number of images of behavior video, scene, and perspective changes. The data set 2 is a ball motion data set under a single background, which is mainly obtained in the network and captured by web crawlers. The data volume is 324 g. This paper uses the data in these two data sets to carry out the knee bending action extraction experiment of ball motion images. The dataset for evaluation is divided in such a way that 80% of the data is used for training, and 20% of the data is used for testing.

3.2. Experimental Index. The extraction algorithm based on image length and Doppler broadening (literature [2] algorithm), the extraction algorithm based on combining
practical channel information (literature [3] algorithm), the extraction algorithm based on the video data set (literature [4] algorithm), the extraction algorithm based on deep network (literature [5] algorithm), the extraction algorithm based on semisupervised learning (literature [6] algorithm), and the algorithm in this paper are used as experimental algorithms to verify the actual application effects of different algorithms by verifying different index.
(a) Literature [2] algorithm extracts the results of data set 1 and data set 2.

(b) Literature [3] algorithm extracts the results of data set 1 and data set 2.

(c) Literature [3] algorithm extracts the results of data set 1 and data set 2.

(d) Literature [4] algorithm extracts the results of data set 1 and data set 2.

Figure 4: Continued.
Motion image collection effect: for two data sets, use the algorithm in this paper to construct a visual sensor sample set of ball motion frame images. The sample set collection needs to be obtained through the visual sensor data set model of the ball motion frame image. In order to verify the anti-interference of the image sample, it is necessary to change the image light intensity through gamma change to verify the image acquisition effect of the algorithm in this paper:

\[ I_{c}(x, y) = \left( \frac{I_c(x, y)}{255} \right)^{\lambda} \times 255, \quad (13) \]

where \( I_c(x, y) \) represents the collected original image and \( \lambda \) represents the image scale.

Knee bending action extraction effect: for ball motion images, the higher the image quality, the lower the false extraction of knee bending motion features, and the better the extraction effect.

Knee flexion extraction accuracy: accuracy is an important index to evaluate the effect of knee flexion extracted by the algorithm. The images in data set 1 belong to complex background images, and the effect of the algorithm can be better analyzed by extracting such complex background images. Therefore, the experiment is mainly aimed at dataset 1. In the analysis process, \( K \) is used to represent the number of scenes in data set 1. The calculation process of extraction accuracy \( P \) is shown in Equation (14):

\[ P = \frac{\sum_{i=1}^{k} C_i}{K} \times 100\%, \quad (14) \]

where \( C_i \) represents the knee bending action to be extracted in the \( i \)th image.

Distance error: distance error is an important index to measure whether there is an error between the extraction result of knee motion in ball motion image and the actual position of knee motion change in the image. The calculation equation of this index is as follows:

\[ G = |k(x, y) - z(x, y)|, \quad (15) \]

where \( k(x, y) \) and \( z(x, y) \) represent the standard coordinates of the same knee bending position and the coordinates calculated by different algorithms, respectively.

Extraction time consumption: it covers a wide range of motions. It is verified that each algorithm extracts the changes of time consumption of knee flexion before and after application for different ball motions. The calculation formula of extracted time consumption is as follows:

\[ T = t_e - t_s, \quad (16) \]
where \( t_s \) and \( t_e \) represent the end and start time of knee flexion extraction, respectively.

### 4. Results and Discussion

It can be seen from Figure 3 that the images in the sample set constructed using the method in this paper have high definition. Ball motion image sample set is constructed in complex background or single background. The design model has strong anti-interference ability for complex background changes and high light intensity and can retain the basic information and original details in the image. The image effect after using the model in this paper is shown in Figure 3.

At the same time, the knee bending action of ball motion image in the data set is extracted. The comparison results of each comparison algorithm and the algorithm in this paper are shown in Figure 4.

It can be seen from Figure 4 that there are great differences in the knee bending action extraction effect in the ball motion image of each comparison algorithm. The image processing in Figures 4(c) and 4(d) has poor quality and insufficient clarity and brightness. Therefore, there is also a large error in the knee bending action extraction result. Instead of accurately extracting the knee bending action in ball motion, the nonknee area is extracted; the probability of missing extraction and false extraction is high. The cleaning angle after image processing of the algorithms in Figures 4(a), 4(b), and 4(e) is moderate, but the effect of extracting the knee bending action is still not very ideal. For example, the extraction algorithm in literature [5] in Figure 4(a) cannot distinguish between the foreground and the background, and some pixels in the complex background are also extracted. The false extraction is serious, but for the image with a single background, there are extraction errors; that is, the false recognition is serious, so it still needs to be further improved. This algorithm has been able to obtain high-quality images in the sample construction stage. Therefore, the knee bending motion extraction effect is also more accurate; there is no false extraction or missing extraction and has a better extraction effect. Calculate the different situations of the number of images in the data set. Each algorithm extracts the accuracy rate of the knee bending action, and the calculation result statistics are shown in Figure 5.

It can be seen from Figure 5 that although the five comparison algorithms can achieve the extraction of knee bending motions, as the amount of data increases, the extraction accuracy gradually decreases. The extraction accuracy of the algorithm in this paper is always 98%. Therefore, the extraction accuracy of the knee bending motion of the ball motion image of the algorithm in this paper is higher, and the effect is better.

After each algorithm processes the ball motion image, the distance error statistics are shown in Figure 6.

It can be seen from Figure 6 that the distance error of the algorithm in literature [2] is about 20 pixels, the distance error of the algorithm in literature [3] is about 33 pixels, and the distance error of the algorithm in literature [4] is about 55 pixels, which is the highest among the six algorithms; the distance error of the algorithm in literature [5] is about 40 pixels. The distance error of the algorithm in literature [6] is about 23 pixels, while the distance error of the algorithm in this paper is about 10 pixels. The distance error of the algorithm in this paper is the smallest, which further verifies that the algorithm in this paper can more accurately extract knee bending motions in ball motion images.

Ball motions cover a wide range, verify that each algorithm is used for different ball motions, before and after application, and the time consumption of the knee bending action is extracted. The comparison test results are shown in Table 1.

It can be seen from Table 1 that for different ball motions, compared with the literature algorithm, the algorithm in this paper can effectively shorten the time required to extract the knee bending action, and the calculation process has high efficiency. For example, for volleyball, the time before the algorithm is applied is 11.59 s. In literature [2], the time consumption after the algorithm is applied is 6.14 s; in literature [3], the time consumption after the algorithm is applied is 6.66 s; in literature [4], the time consumption after the algorithm is applied is 7.42 s; in literature [5], the time consumption after the algorithm is applied is 7.51 s; and in literature [6], the time consumption after the algorithm is applied is 5.83 s. Compared with these algorithms, the time consumption of the algorithm in this paper is only 2.65 s.
This paper studies the knee bending motion extraction algorithm of ball motion images based on visual sensors and builds a sample set of ball sports images on the basis of visual sensors. Based on the sample set, the motion energy model and CNN are used to extract the features. The feature results are fused in the CNN, and finally, the knee bending motion is extracted. Experiments show that the algorithm has high accuracy and short extraction time. Compared with similar algorithms, it has better performance and better practical application effect. However, the algorithm in this paper still has some defects; that is, the amount of experimental data is less. Therefore, the future work needs to take this problem as the goal to test the application performance of the algorithm.

5. Conclusions and Future

This paper studies the knee bending motion extraction algorithm of ball motion images based on visual sensors and builds a sample set of ball sports images on the basis of visual sensors. Based on the sample set, the motion energy model and CNN are used to extract the features. The feature results are fused in the CNN, and finally, the knee bending motion is extracted. Experiments show that the algorithm has high accuracy and short extraction time. Compared with similar algorithms, it has better performance and better practical application effect. However, the algorithm in this paper still has some defects; that is, the amount of experimental data is less. Therefore, the future work needs to take this problem as the goal to test the application performance of the algorithm.

Data Availability

The data used to support the findings of this study are included within the article. Readers can access the data supporting the conclusions of the study from UCF Sports Action data set.

Conflicts of Interest

The authors declare that there is no conflict of interest with any financial organizations regarding the material reported in this manuscript.

Acknowledgments

This work was supported by the feasibility study on promoting the operation of snow volleyball competition in Heilongjiang Province of China under Grant No. BXQN2012.

References

[1] N. Wang, D. Mo, Z. Song, R. Wang, G. Li, and Y. Wu, “Feature extraction of precession target based on Doppler profile sequences by maximum likelihood estimation,” *Electronics Letters*, vol. 55, no. 9, pp. 554–556, 2019.
[2] B. Li, D. Mo, Z. Song et al., “Feature extraction algorithm of precession target based on image length and Doppler broadening,” *Applied Optics*, vol. 59, no. 29, pp. 9098–9103, 2020.
[3] J. Cai, J. Hu, S. Li, J. Lin, and J. Wang, “Combination of temporal-channels correlation information and bilinear feature for action recognition,” *IET Computer Vision*, vol. 14, no. 8, pp. 634–641, 2020.
[4] T. Zyer, D. S. Ak, and R. Alhajj, “Human action recognition approaches with video datasets—a survey,” *Knowledge-Based Systems*, vol. 222, no. 6, article 106995, 2021.
[5] S. Chaudhary and S. Murala, “Depth-based end-to-end deep network for human action recognition,” *IET Computer Vision*, vol. 13, no. 1, pp. 15–22, 2019.
[6] S. Sahoo, U. Srinivasu, and S. Ari, “3D features for human action recognition with semi-supervised learning,” *IET Image Processing*, vol. 13, no. 6, pp. 983–990, 2019.
[7] S. Yongmin, L. Shuntian, A. N. Panpan, GETHIK Group Co., Ltd, Hangzhou 311100, China, and School of Electronic Engineering, Xidian University, Xi’an 710071, China, “Memory tracking method of sensor location based on vision,” *Journal of Applied Optics*, vol. 42, no. 5, pp. 853–858, 2021.
[8] J. Xiaorong and Z. Huarui, “Secondary feature extraction algorithm of abnormal gait based on image,” *Computer Simulation*, vol. 37, no. 3, pp. 381–384, 2020.
[9] M. A. Li, J. F. Han, and J. F. Yang, “Automatic feature extraction and fusion recognition of motor imagery EEG using multi-level multiscale CNN,” *Medical & Biological Engineering & Computing*, vol. 59, no. 10, pp. 2037–2050, 2021.
[10] L. Yu, M. Jin, and K. Zhou, “Multi-channel biomimetic visual transformation for object feature extraction and recognition of complex scenes,” *Applied Intelligence*, vol. 50, no. 3, pp. 792–811, 2020.
[11] A. Sachan, W. Wei, and Y. S. Choo, “Feature recognition and strength estimation of chain links by 3D inspections,” *Ocean Engineering*, vol. 225, no. 1, article 108785, 2021.
[12] M. Mhamdi and D. Ziou, “3D object recognition through a size function resulting from an invariant topological feature,” *Pattern Recognition*, vol. 120, article 101831, 2021.
[13] Z. Xia, J. Xing, C. Wang, and X. Li, “Gesture recognition algorithm of human motion target based on deep neural network,” *Mobile Information Systems*, S.-B. Tsai, Ed., vol. 2021, Article ID 2621691, 12 pages, 2021.
[14] H. V. A. Chenarlog and F. Razzazi, “Multi-stream 3D CNN structure for human action recognition trained by limited data,” *IET Computer Vision*, vol. 13, no. 3, pp. 338–344, 2019.
[15] K. Gedamu, Y. Ji, Y. Yang, L. Gao, and H. T. Shen, “Arbitrary-view human action recognition via novel-view action generation,” *Pattern Recognition*, vol. 118, no. 10, article 108043, 2021.
[16] X. L. Tang, W. C. Ma, D. S. Kong, and W. Li, “Semisupervised deep stacking network with adaptive learning rate strategy for motor imagery EEG recognition,” *Neural Computation*, vol. 31, no. 5, pp. 919–942, 2019.

[17] F. Xu, F. Rong, Y. Miao et al., “Representation learning for motor imagery recognition with deep neural network,” *Electronics*, vol. 10, no. 2, p. 112, 2021.

[18] D. Gurve, D. Delisle-Rodriguez, M. Romero-Laiseca et al., “Subject-specific EEG channel selection using non-negative matrix factorization for lower-limb motor imagery recognition,” *Journal of Neural Engineering*, vol. 17, no. 2, 2020.

[19] G. Krishnan, R. Joshi, T. O’Connor, F. Pla, and B. Javidi, “Human gesture recognition under degraded environments using 3D-integral imaging and deep learning,” *Optics Express*, vol. 28, no. 13, pp. 19711–19725, 2020.

[20] A. Guzman-Pando, M. I. Chacon-Murguia, and L. B. Chacon-Diaz, “Human-like evaluation method for object motion detection algorithms,” *IET Computer Vision*, vol. 14, no. 8, pp. 674–682, 2020.

[21] X. Ma and X. Li, “Dynamic gesture contour feature extraction method using residual network transfer learning,” *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1503325, 11 pages, 2021.

[22] J. An, B. Choi, H. Kim, and E. Kim, “A new contour-based approach to moving object detection and tracking using a low-end three-dimensional laser scanner,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7392–7405, 2019.

[23] S. Javed, A. Mahmood, S. Al-Maadeed, T. Bouwmans, and S. K. Jung, “Moving object detection in complex scene using spatiotemporal structured-sparse RPCA,” *IEEE Transactions on Image Processing*, vol. 28, no. 2, pp. 1007–1022, 2019.