

Research Article

Feature Extraction Method of Snowboard Starting Action Using Vision Sensor Image Processing

Tao Zhang,^{1,2} Zhipeng Li,³ Myungsoo Shin,⁴ Chunxia Wang,⁵ Wenli Song ,² and Lei Lui⁶

¹Graduate College, Harbin Sport University, Harbin 150008, China

²College of Physical Education and Training, Harbin Sports University, Harbin 150009, China

³Winter Olympic College, Harbin Sports University, Harbin 150008, China

⁴Soongsil University College of Physical Education, Seoul, Republic of Korea

⁵Mudanjiang Normal University, Mudanjiang 1157011, China

⁶Heilongjiang Institute of Technology, Harbin 150050, China

Correspondence should be addressed to Wenli Song; songwenli@hrbipe.edu.cn

Received 5 October 2021; Revised 31 October 2021; Accepted 6 January 2022; Published 19 January 2022

Academic Editor: Jinan Fiaidhi

Copyright © 2022 Tao Zhang 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.

There is a lot of noise in the snowboard starting action image, which leads to the low accuracy of snowboard starting action feature extraction. We propose a snowboard starting action feature extraction using visual sensor image processing. Firstly, the overlapping images are separated by laser fringe technology. After separation, the middle point of the image is taken as the feature point, and the interference factors are filtered by laser. Secondly, the three-dimensional model is established by using visual sensing image technology, the action feature images are input in the order of recognition, and all actions are reconstructed and assembled to complete the action feature extraction of snowboard. The interference factors are filtered by laser, the middle part of the action image is extracted according to the common features of multiple images, and its definition is described. The movement change and moving distance are used to count the most features and clarity. Finally, the edge recognition effect of snowboard starting action image and the action recognition effect under multiple complex images are taken as experimental indexes. The results show that the method has a good effect on image edge extraction, the extraction effect is as high as 95%, and the accuracy is as high as 2.1%. In addition, under multiple complex images, the action feature recognition rate is also high, which can prove that the method studied has better accuracy in snowboard starting action feature extraction.

1. Introduction

In order to improve the professionalism of snowboard athletes, it is necessary to analyze the athletes' starting action and side rotation turnover action and find the deficiencies in time [1]. In analyzing athletes' starting action of snowboard, it is difficult to accurately analyze the features of such action if only observed with the naked eye [2, 3].

Physical training and special training for the starting stage are two necessary links in the training of snowboard team athletes. At present, the physical fitness training of snowboard is still dominated by barbells and other traditional instruments [4, 5]. However, due to the inertia participation of counterweight in the exercise process, the resistance in the

training process is extremely unbalanced. The greater the speed, the more unbalanced the inertia. The instantaneous impulse is very easy to increase the risk of sports injury. Based on extracting features of the starting action of snowboard and image of the traditional starting action of snowboard, this paper proposes a method to extract features of the starting action of snowboard based on visual sensing image processing. Then, laser is used for filter processing of the interference factors, and the Hough method is adopted to obtain several parameters. The results of experiments conducted on two true data sets show that compared with snowboard action features extraction, snowboard action features extraction is better than other literature methods in terms of stability, error rate, and accuracy.

The main contributions of this paper are as follows: (1) The overlapping images are segmented by laser fringe technology. After segmentation, the middle point of the image is taken as the feature point, and the interference factors are filtered by laser. (2) A three-dimensional model based on visual sensing image technology is proposed. (3) The three-dimensional model of visual sensing tracking is used to input the action feature images in the recognized order, reconstruct and assemble all the actions, carefully enlarge and capture the edge and center of the image, extract the feature points of the starting action, and complete the spatial construction of action features.

2. Related Works

Many scholars have carried out relevant research. There are many existing methods to extract the starting action features of snowboard [6, 7]. Literature [8] proposed a human action image recognition method under high-intensity motion. This method segmented the image in advance and then extracted the action features, so as to obtain the Gaussian distribution model of the action image background and realize the extraction of human action, but the feature extraction clustering of this method is poor. Literature [9] proposed a human motion recognition method based on 3D CNN, which encodes the motion information to recognize and extract the motion. However, the stability of motion feature extraction is low. Literature [10] proposed a human skeleton behavior recognition method based on spatio-temporal weighted pose motion features. This method extracts the space and information of different joint points of the human body to represent the action sequence and then realizes the extraction of motion features. However, the feature extraction error of this method is large. Literature [11] proposed the research on the feature extraction method of football players' foul action based on machine vision. Machine vision is a branch of the development of artificial intelligence, which uses machines instead of human eyes to judge and measure. It mainly converts the target into image signals through machine vision products, transmits them to a special image processing system, and converts the color, brightness, and distribution of image pixels into digital signals. The target feature information is extracted by operation. However, the accuracy of foul action feature extraction is low. Literature [12] proposed a data feature extraction method based on deep learning. Firstly, based on the flexible setting of convolution kernel, the branch structure of neural network is introduced to extract the depth features of original data at multiple scales; then, the data features obtained from each branch are fused and used as the input of the next convolution layer. However, this method takes a long time. Although the features of the starting action of snowboard can be extracted, there are also such problems as the low accuracy. Visual sensing technology, as one of the most commonly used image detection technologies, can segment the target into several subimages and then detect and segment them one by one [13]. Thus, the technology has a good recognition effect and is applied to extracting features of the starting action of snowboard.

Visual sensing image processing is used to recognize and detect the features, which provides a basis for improving the action standard of Chinese athletes.

In Literature [14], Yu et al. proposed a human motion capture method based on a single RGB-D sensor. This method realizes motion tracking refinement at the body part level through semantic tracking loss calculation, which can improve the tracking accuracy in the case of severe occlusion and fast motion. However, the motion accuracy of this method is low. In literature [15], Maruyama et al. proposed a MoCap capture method based on inertial measurement unit. This method does not need any optical equipment and can measure the motion attitude even in the outdoor environment. However, the motion delay of this method is too long. Literature [16] proposed to use a series of precise physiological and biochemical instruments to detect the physical function of athletes, so as to master the law of the change of athletes' mechanical energy in sports training, scientifically organize and arrange the training and practical activities of snowboarding skills on veneer U-shaped venues, and reasonably guide athletes' diet. This can effectively prevent sports injury, prolong sports life, and improve sports ability. However, due to the loss of motion in this method, it is impossible to accurately capture snowboarding athletes' skiing motion.

3. Preprocessing of Extracting the Starting Action of Snowboard

In the preprocessing of snowboarding starting action feature extraction, this paper mainly carries out image segmentation and image filtering and completes the feature extraction process.

3.1. Image Segmentation. Before extracting the action features, the laser fringe technique is used to separate the overlapping images, which is both simple and practical. Although there will be some difficulties in the methods of segmentation, a slight error will lead to the failure of segmentation. Therefore, the MeanShift algorithm is generally used for laser segmentation. Suppose there exists the space d which includes multiple simulation points n , and the segmentation point is x_i and $i = 1, n$, so the segmentation form of point x is

$$M_h(x) = \frac{1}{k} \sum_{x_i=sh} (x_i - x), \quad (1)$$

where k is the coefficient, and it meets the condition $Sh(x) = \{y: (y - x)^2 (y - x) \leq h^2\}$, so it can be seen from the area $Mh(x)$ that the displacement of the two points x and x_i will affect the feature extraction. The shorter the distance, the more accurate the feature recognition. Considering the changes of the distance, the distance of the algorithm can be expressed as

$$M(x) = \frac{\sum_n^{i=1} G_H(x_i - x)_w(x_i)(x_i - x)}{\sum_n^{i=1} GH(x_i - x)_w(x_i)}, \quad (2)$$

where $G(H)$ refers to the coordinate of the single feature, and H refers to the square matrix, i.e., $H = \text{diag}[h_1^2, \dots, h_d^2]$. To obtain the gray area in the center of the image, it is necessary to correctly distinguish the area and side length of the gray part [17] and take the position of the center as the coordinate center, which is taken as $[p, q]$, and the formula of the central position is

$$X_k = \frac{\sum_q^{i=p} I_i \cdot I_i}{\sum_q^{i=p} I_i}, \quad (3)$$

where I_i is the gray index of the pixel point i , p and q are points of two directions in the region, and k is the coefficient. This method can not only accurately identify the central area of the image, but also is not disturbed by other factors.

In the segmentation of overlapping image [18, 19], changes of the laser should be used to determine the number of overlapping images. All images are segmented into multiple modules along the direction of the initial image and establish the matrix H as

$$H(x, y) = \begin{bmatrix} \frac{\partial^2 g(x, y)}{\partial x^2} & \frac{\partial^2 g(x, y)}{\partial x \partial y} \\ \frac{\partial^2 g(x, y)}{\partial x \partial y} & \frac{\partial^2 g(x, y)}{\partial y^2} \end{bmatrix} \otimes Z(x, y), \quad (4)$$

where $\partial^2 g$ refers to the fuzzy function, $Z(x, y)$ refers to the pixel value of all points, ∂x refers to the feature distribution sequence, ∂y refers to action function, and ∂ refers to the action angle.

3.2. Image Filtering Processing. Before extracting the image feature points, we must first ensure that the extraction process is not disturbed by other factors, so as to maintain the stability of the sensing image. Therefore, the middle point of the image is taken as the feature point, the interference factors are filtered by laser [20, 21], and several parameters are obtained by Hough method. Suppose there are two critical points in the graph, i.e., b_1 and b_2 , the image's center coordinates are substituted into the line equation.

$$\left\{ F_e = \frac{\sum_{b_1}^{i=b_1} i \times [y_p(i) - y_v(i)]}{\sum_{N_2}^{i=1} [y_p(i) - y_v(i)]}, \quad (5)$$

where $y_p(i)$ refers to the horizontal axis coordinate. $y_v(i)$ represents the vertical axis coordinates. The size of graphic structure area is the specific form of system tracking, and the formula of the area range S is

$$S = \sum_{b_2}^{i=b} i \times [y_p(i) - y_v(i)], \quad (6)$$

where $y_p(i)$ refers to the feature set of correlation and $y_v(i)$ refers to the analytic factor. The above processing process can be completed quickly in a short time, and the information left by the laser can be detected accurately, which provides a basis for starting action feature extraction.

4. Action Feature Extraction Using Visual Sensing Image Processing

Based on the above processing, the starting action features of snowboard are extracted, and based on the capture of athletes' action images, the visual sensing and tracking technology is used to collect the features [22, 23]. Because the action is not standard, the presentation on the visual sensing image will be very uneven, and there will be some errors in laser detection. When the camera is at a long distance, the available laser becomes less, and the laser transmitted to the camera cannot achieve the expected effect, so the image display is not clear and impurities will appear. The specified detection principle is "when the action feature points are not detected [24], the search will continue in a certain area" [25]. If the fuzzy pixel points are identified, they are impurity points in the image. Ignore this pixel and continue to identify in the range till the completion. The removal process of inactive feature points is shown in Figure 1.

On this basis, the contour is extracted. Generally, the center of the extracted action feature image is clear and the edge is fuzzy. Therefore, the length of the edge feature of the action image is

$$E = \theta E^{LBF} + (1 - \theta) E^{LGF} + \nu L(\phi) + \mu P(\phi), \quad (7)$$

where E refers to the length range of image edge, μ is action fitting parameter, ν is the speed, L is the length, P is fitting index, ϕ is the rotation angle, and θ is the coefficient. Extract the middle part of the action image using the common features of multiple images:

$$L(\phi) = \int_a \delta(\phi) |\nabla \phi| dx, \quad (8)$$

where d is the distance, δ is the gray coefficient of the image, and a is the natural number. Then, adjust the size of the image into the 3×3 mode, and then conduct image recognition block by block to get the data about action features [26].

$$P(\phi) = \frac{1}{2} (\nabla \phi - 1)^2 dx, \quad (9)$$

where $\nabla \phi$ refers to the typical feature set of the image and dx refers to the fusion information of edge gradient. For the tracking action feature set, it just conforms to the normal distribution of the image pixel set $N(0, 1)$, so the intensity values of typical features of the image are

$$E^{LBF} = \lambda_1 \left[\int \left[\int K_e(x-y) |I - f_1(x)|^2 H(\phi) dy \right] dx + \lambda_2 \left[\int K_a(x-y) |I - f_2(x)|^2 (1 - H(\phi)) dy \right] dx, \quad (10)$$

where $\int K_e$ refers to the balance factor, x refers to the feature amount of the action, I refers to the edge pixel coefficient of dynamic image, and y represents fuzzy correlation. dy represents the fusion degree of feature extraction. f_1 and f_2 are 3D model of visual sensor tracking. Input the action feature image according to the recognition order, reconstruct and assemble all actions, carefully enlarge and capture the edge and center of the image, extract the feature points of

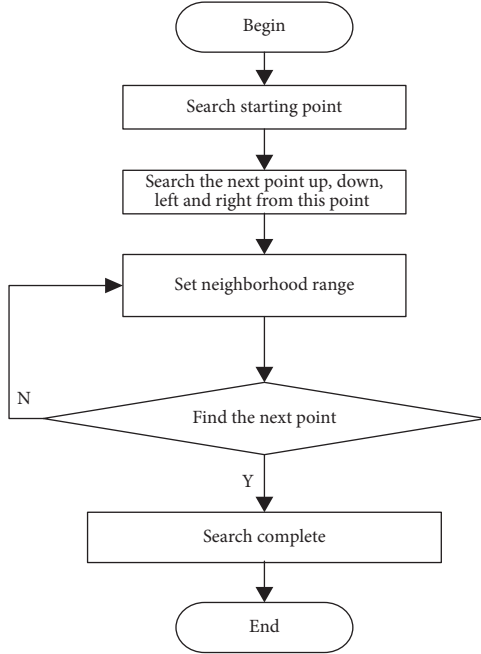


FIGURE 1: Removal process of inactive feature points.

the starting action, and complete the spatial construction of action features [27].

$$\frac{\partial \phi}{\partial t} = -\delta(\phi)\theta(\lambda_1 e_1^{LBF} - \lambda_2 e_2^{LBF}), \quad (11)$$

where λ and ∂ refer to 3D model index, t refers to time, and e refers to feature parameter.

Suppose there is a point (i, j) in the image, the pixel value of the image at this moment can be expressed by $x(i, j)$, and the pixel difference of the action feature of the image can be expressed by M :

$$M_{ij} = \text{med}(X_{i-1,j-1} X_{i,j} L X_{i+1,j+1}), \quad (12)$$

where X_{i-1} refers to parameter distribution set. When the pixel value of the image's central area $M \times M$ is smaller than the gray value, the image does not achieve the predicted feature extraction. The fuzzy pixel values are compared to mark the damaged points on the edge. The marking range is expressed as

$$\begin{cases} (x) = \frac{K_\sigma(x) * [H_z(\phi(x))I(x)]}{K_\sigma(x) * H_z(\phi(x))}, \\ (y) = \frac{K_\sigma(x) * [(1 - H_s(\phi(x)))I(x)]}{K_\sigma(x) * (1 - H_s(\phi(x)))}, \end{cases} \quad (13)$$

where S refers to image area. According to the marking results of the above formula, the most features and clarity are counted with the change of action and the distance of movement [28].

The edge feature threshold of the starting action of single board at moment t is $\text{Ima}(x_i) = p(x_i | d_{0,1,2,\dots,t})$, and the specific pixel value of the feature n is

$$S_i = \{s'_i(x_i^j, w_t^j); j = 1, 2, L, N\}. \quad (14)$$

When $t = 0, 1, 2, L, k$, the total amount of features is $Z_k = \{z_0, z_1, L, z_i\}$, and Z is the variant. Suppose xk is all the parameters of the four angles in the image, and the function is $\delta(z) = (d/dz)H(z)$. Obtain the difference value of the two images and draw them in the order of imaging successively. The pixel difference obtained is expressed as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix}, \quad (15)$$

where ξ and η are two different vectors. The expression of the maximum gray value in the sensing area is

$$\text{pixel_A} = \max \left(\sum_s^{i=1} (Q - P) \right), \quad (16)$$

where Q is the minimum gray value. Connecting the different points will form a waveform data map, mark the details of the four corners, compare the final gray value with the initial gray value, and evaluate the image. The transformation displacement is

$$Econ(v_i) = |\text{Edge}(\text{Gray}(i) - \text{Gray}(i))|. \quad (17)$$

Finally, after the clear edge contour is obtained, the typical feature information is analyzed as [18]:

$$P(y_j | x_{v_i}, \theta, \beta) = \begin{cases} \frac{\sum_{n=1}^{i-1} |\theta_n^i(x, y) - \theta_n^j(x, y)|}{(2s + 1)^2}, \\ 0, m = n, \end{cases} \quad (18)$$

where β is action flip angle, m and n are two different sensors, P is fitting index, and S is area.

After the feature information is obtained, find the target image by relying on the original image, compare it with the image according to its typical features, confirm the usable action image, and finally determine the end-point pixel features.

Image segmentation method is adopted to extract the starting action of snowboard. Noise reduction is realized for the collected action feature segmentation method, the correlation spectrum feature quantity of the action is extracted, and the variance fusion model is obtained. The signal sampling node is defined as shown in Figure 2.

Input: according to the test sample set and training sample set of covariance array data, the clutter scatterer defining snowboard action features is obtained.

Output: the fusion result meeting the minimization target parameters.

Get the set of the first element; that is, the clutter parameters of the action characteristics are as follows: (1) the correlation spectrum's feature quantity of the output layer is analyzed, and the multithreshold decision and threshold detection method are adopted to obtain the estimated value of the maximum covariance parameter n . (2) The upper limit

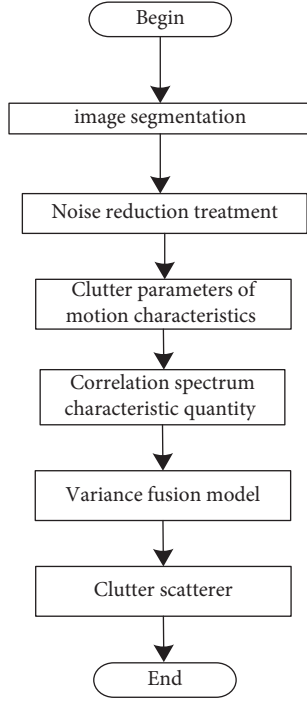


FIGURE 2: Process of feature extraction.

of noise reduction processing is estimated. (3) Combined with the image segmentation method, the snowboarding action feature fusion is realized to improve the recognition accuracy of action feature extraction.

5. Experimental Analysis and Results

In order to verify the application performance of this method in the feature extraction of snowboarding starting action, the experimental test and analysis are carried out.

5.1. Experimental Environment and Data Set. In order to ensure the applicability of the test results, the action feature extraction test room is set, three cameras are installed to obtain the action diagram, and the machine vision system is ov7670 camera.

This paper presents experiments and tests on two data sets (KTH and HMDB). KTH data set: it is also one of the data sets with high utilization rate in this paper. The data set contains 2391 groups of data, including 6 actions. Each action is completed by 25 characters in 4 different scenes. Therefore, there are 600 video sequences, and each video can be divided into 4 subsequences. The action of KTH data set is relatively standardized, and fixed shots are used at the same time. The number is also relatively rich for the current model training, so it can be said that it is a very useful data set for the task of simple action recognition. HMDB data set: there are 51 categories in total, with an average of 100–200 groups of data in each category. In terms of data volume and category, we can see that there are rich data, but this data set is mainly composed of videos taken by some movie shots and daily cameras, so the background is relatively complex,

and there are also videos with dynamic shots and switching shots. Therefore, this data set is more suitable for target recognition and target detection.

In this study, the data in each task of training and testing are strictly implemented in accordance with the standards. The above data are in the form of KTH data set and HMDB data set. In the training process, 36000 tasks are randomly selected for testing. It is shown in Table 1.

5.2. Experimental Indicators

5.2.1. Poor Clustering. Clustering analysis can deal with the classification determined by multiple variables, eliminate the classification of variables, and improve the quality of feature extraction. In order to verify the operation performance of this paper, feature extraction clustering is selected as the index for analysis. The clustering calculation formula is

$$X = M_v - C_N(r), \quad (19)$$

where M_v refers to clustering feature decomposition, and $C_N(r)$ refers to clustering feature integration.

5.2.2. Low Stability. This can ensure that the extraction process is not disturbed by other factors, so as to maintain the stability of the sensing image and improve the quality of action feature extraction.

$$\lambda = X - Di f(C_1, C_2). \quad (20)$$

5.2.3. Large Error. The action standard will be presented on the visual sensing image. The error rate will be detected by laser. The smaller the error, the clearer the action of feature extraction.

$$p = \lambda - \frac{x_1 + x_2}{Z}. \quad (21)$$

5.2.4. Low Accuracy. According to the accuracy comparison results of several methods, the better the accuracy is, the stronger the applicability is.

$$m_0(\omega) = \min P - m(x). \quad (22)$$

5.2.5. Time Consuming. Through the time-consuming comparison results of several methods, it can be seen that the shorter the time-consuming, the more accurate the action of feature extraction.

5.3. Results and Discussion. In order to verify the effectiveness of the method to extract features of the starting action of snowboard based on visual sensing image processing, experimental analysis is carried out. The purpose of this experiment is to verify the accuracy of action feature extraction, which is mainly measured by two experiments. In the first experiment, by specifying the snowboard action

TABLE 1: Basic information of data set.

Basic information	KTH	HMDB
Number of classes	50	36
Number of class samples	500	156
Number of training set classes	56	21
Number of validation set classes	16	85
Number of test set classes	20	150

image, the edge features of the image are extracted and the effect of edge feature extraction is compared; the second experiment mainly verifies the effect of snowboard action extraction under multiple complex images. In order to make the experimental results more illustrative, the methods of literatures [8], [9], [10], [11], and [12] are compared with the proposed methods. The experimental results are calculated by formula (5), and the experimental results are shown in Figure 3.

According to Figure 3, under multiple complex images, the recognition rates obtained by the feature extraction method in literatures [8]-[12] are low, so it can be verified that the action extraction effects of these four methods are poor, and the application effect is not as good as the method studied. The method to extract features of the starting action of snowboard in this method is good, and the recognition rate is more than 90%. This proves that the action feature extraction method in this study is less affected by the detection background and can effectively remove the noise in the image with good clustering, which can ensure the accuracy of snowboard action features. The stability test results of snowboard feature extraction are in Figure 4.

According to Figure 4, the feature extraction stability of the method in literature [8] is 0.41, the feature extraction stability of the method in literature [9] is 0.43, the feature extraction stability of the method in literature [10] is 0.47, the feature extraction stability of the method in literature [11] is 0.38, and the feature extraction stability of the method in literature [12] is 0.32, and the stability of feature extraction of the proposed method is 0.49. It can be seen that the snowboard action feature extraction using the method in this paper has better stability and practicability.

According to Figure 5, the feature extraction error rate of the method in literature [8] is 3.2%, the feature extraction error rate of the method in literature [9] is 2.5%, the feature extraction error rate of the method in literature [10] is 6.3%, the feature extraction error rate of the method in literature [11] is 5.6%, and the feature extraction error rate of the method in literature [12] is 6.1%, and the error rate of feature extraction of the proposed method is 0.49%. It can be seen that the snowboard action feature extraction error rate using the method in this paper is lower and the extraction level is high.

According to Figure 6, the feature extraction accuracy of the method in literature [8] is 1%, the feature extraction accuracy of the method in literature [9] is 0.62%, the feature extraction accuracy of the method in literature [10] is 0.29%, the feature extraction accuracy of the method in literature [11] is 0.99%, and the feature extraction accuracy of the

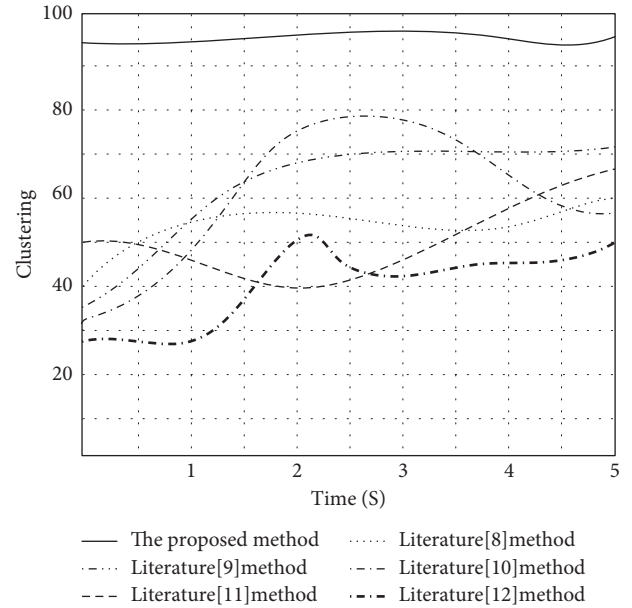


FIGURE 3: Comparison of clustering of snowboard actions.

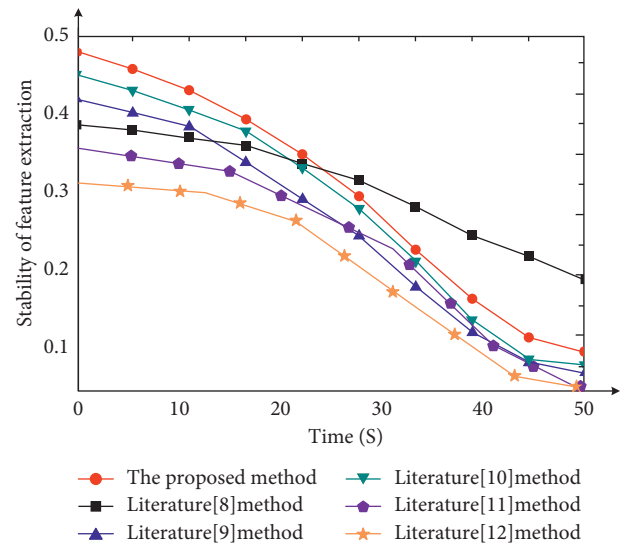


FIGURE 4: Comparison of stability of snowboard action feature extraction.

method in literature [12] is 1.19%, and the feature extraction accuracy of the proposed method is 2.1%. Therefore, the accuracy of snowboard action feature extraction using the method in this paper is lower.

According to Figure 7, the feature extraction time of literature [8] method is 59, the feature extraction time of literature [9] method is 65, the feature extraction time of literature [10] method is 80, the feature extraction time of literature [11] method is 108, the feature extraction time of literature [12] method is 110, and the feature extraction time of the proposed method is 33. Therefore, the snowboarding action feature extraction using this method takes less time.

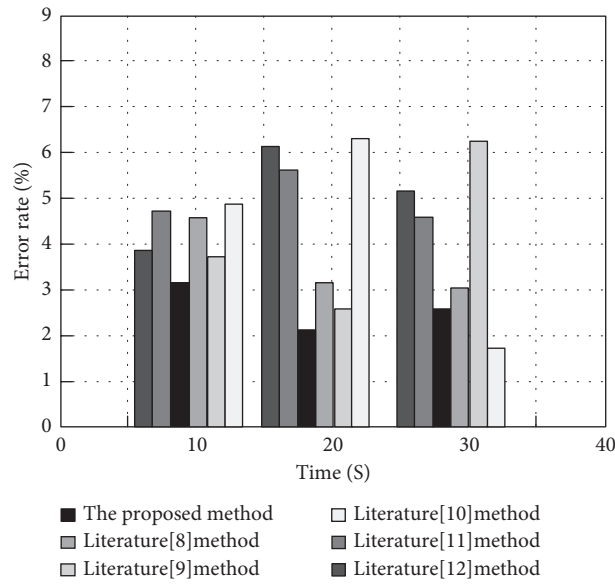


FIGURE 5: Comparison of the error rate of snowboard action feature extraction.

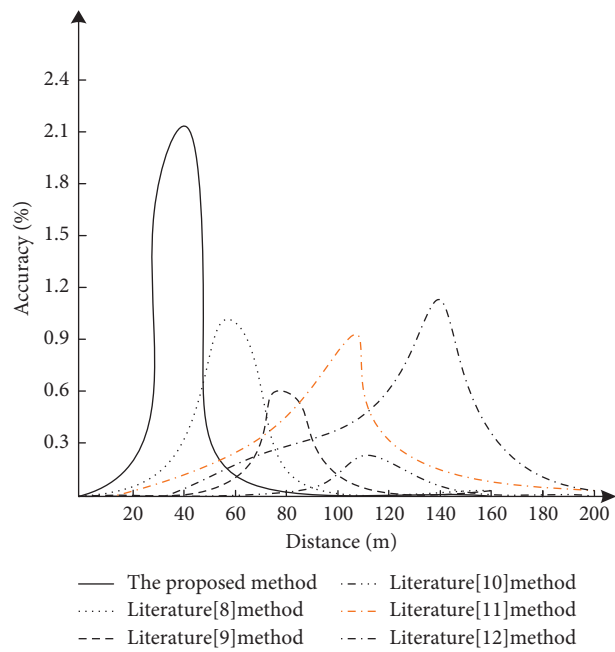


FIGURE 6: Comparison of the accuracy of snowboard action feature extraction.

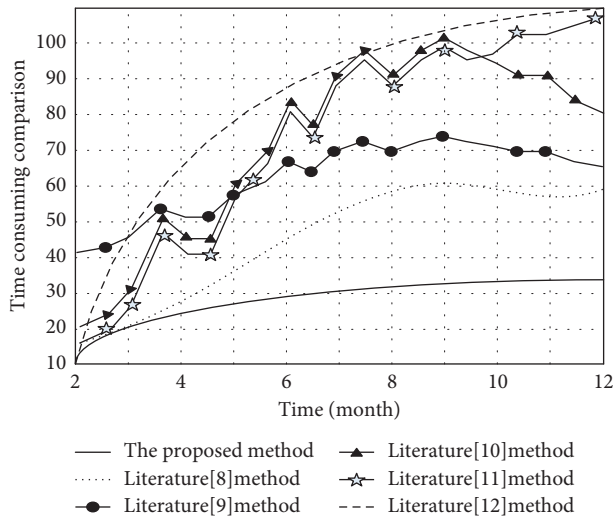


FIGURE 7: Comparison of time-consuming for feature extraction of snowboarding.

6. Conclusions and Future Works

To sum up, due to the large amount of noise in the snowboard starting action image, the snowboard starting action feature extraction is not so accurate, the error is large, and the feature extraction takes a long time. Therefore, this paper proposes a snowboard starting action feature extraction based on visual sensing image processing. The innovation of the method in this paper is that it uses the visual sensing image processing technology to delete the redundant images and can classify, extract, and segment the same features. In this way, each part of the image can be analyzed in more detail, the interference factors in the image can be filtered by laser, and the middle part of the action image can be extracted according to the common features of multiple images to obtain its definition description. The movement change and moving distance are used to count the most features and clarity. In the future, we also need to conduct in-depth research on the feature extraction of starting action of snowboard, so as to further improve the accuracy of action feature extraction and better the extraction effect.

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 KTH data set and HMDB data set.

Conflicts of Interest

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

Acknowledgments

This work was supported by the Foundation of General Project of Higher Education and Teaching Reform in

Heilongjiang Province China (SJGY20200406), General Project of the China University Sports Association (202013508), Research on the construction of the ski competition system for college students in my country, Discipline Echelon Project of Harbin Institute of Physical Education (XKB04), and Research on the Cultivation of Student Sports Core Accomplishment from the Perspective of Sports and Education Integration, and Innovative Research and Practice of Skills Training Methods for Normal Students in Normal Universities (20-XJ21009).

References

- [1] J. Yu, M. Tan, H. Zhang, D. Tao, and Y. Rui, "Hierarchical deep click feature prediction for fine-grained image recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 2, pp. 563–578, 2019.
- [2] S. Mushtaq and A. H. Mir, "Image copy move forgery detection: a review," *International Journal of Future Generation Communication and Networking*, vol. 11, no. 2, pp. 11–22, 2018.
- [3] C. Ding, K. Liu, L. I. Guang, L. Yan, B.-Y. Chen, and Y.-M. Zhong, "Spatio-temporal weighted posture motion features for human skeleton action recognition research," *Chinese Journal of Computers*, vol. 43, no. 1, pp. 29–40, 2020.
- [4] S. Memon, M. Ahmed, S. Narejo, U. Ahmed baig, B. Shankar chowdry, and M. Rizwan anjum, "self-driving car using lidar sensing and image processing," *International Journal of Grid and Distributed Computing*, vol. 13, no. 2, pp. 77–88, 2021.
- [5] D. Kim, "On-street parking vacancy and vehicle detection in video using computer vision technique," *International Journal of Transportation*, vol. 7, no. 2, pp. 11–18, 2019.
- [6] J. Yu, D. Tao, M. Wang, and Y. Rui, "Learning to rank using user clicks and visual features for image retrieval," *IEEE Transactions on Cybernetics*, vol. 45, no. 4, pp. 767–779, 2015.
- [7] M. Ghavidel, P. Bayat, and M. E. Farashiani, "Evaluation of boxus destruction using satellite image processing techniques in the northern forests of Iran," *International Journal of Future Generation Communication and Networking*, vol. 13, no. 4, pp. 11–22, 2020.
- [8] F. Lu, "Adaptive recognition method of aerobics decomposition action image based on feature extraction," *Science Technology and Engineering*, vol. 19, no. 7, pp. 148–153, 2019.
- [9] V. A. Chenarlogh 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.
- [10] P. Svermova and M. Cernik, "Corporate social responsibility of companies producing PFOA containing waxes for cross-country skiing," *Sustainability*, vol. 12, no. 12, p. 5141, 2020.
- [11] B. D. Da Silva, P. C. Bernardes, P. F. Pinheiro, E. Fantuzzi, and C. D. Roberto, "Chemical composition, extraction sources and action mechanisms of essential oils: natural preservative and limitations of use in meat products," *Meat Science*, vol. 176, no. 3, Article ID 108463, 2021.
- [12] A. P. Crupi, D. P. Gregovich, and K. S. White, "Steep and deep: terrain and climate factors explain brown bear (*Ursus arctos*) alpine den site selection to guide heli-skiing management," *PLoS One*, vol. 15, no. 9, 2020.
- [13] J. Jun Yu, Y. Dacheng Tao, and D. Tao, "Click prediction for web image reranking using multimodal sparse coding," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 2019–2032, 2014.

- [14] T. Yu, J. Zhao, Y. Huang, Y. Li, and Y. Liu, "Towards robust and accurate single-view fast human motion capture," *IEEE Access*, vol. 7, no. 99, Article ID 85548, 2019.
- [15] T. Maruyama, M. Tada, and H. Toda, "Riding motion capture system using inertial measurement units with contact constraints," *International Journal of Automation Technology*, vol. 13, no. 4, pp. 506–516, 2019.
- [16] C. Yi, B. Chen, S. Yuan, and B. Xu, "A method for multi-person motion capture based on multi-mode in virtual geographic environments," *Journal of Geo-Information Science*, vol. 021, no. 003, pp. 305–314, 2019.
- [17] A. Martínez, C. Snyder, S. R. Moore, and T. Stöggel, "A comprehensive comparison and validation of published methods to detect turn switch during alpine skiing," *Sensors*, vol. 21, no. 7, p. 2573, 2021.
- [18] L. Brea Alejo, J. Gil-Cabrera, A. Montalvo-Pérez, D. Barranco-Gil, J. Hortal-Fondón, and A. Navandar, "Performance parameters in competitive alpine skiing disciplines of slalom, giant slalom and super-giant slalom," *International Journal of Environmental Research and Public Health*, vol. 18, no. 5, p. 2628, 2021.
- [19] K. Gurucharan, S. S. Kiran, K. Babburu, and L. Vadda, "Computer vision based fruit recognition and classification system," *International Journal of Future Generation Communication and Networking*, vol. 13, no. 3, pp. 1–14, 2020.
- [20] C.-H. Lee and Y.-G. Jeon, "Biomechanical analysis concerning steepness application on ground training posture for snowboarding," *Korean Journal of Sports Science*, vol. 28, no. 4, pp. 1111–1120, 2019.
- [21] D. Kim and Y. Chang, "Traffic vision analysis with convolutional neural network model," *International Journal of Transportation*, vol. 8, no. 1, pp. 37–42, 2020.
- [22] V. Rosso, V. Linnamo, W. Rapp et al., "Simulated skiing as a measurement tool for performance in cross-country sit-skiing," *Proceedings of the Institution of Mechanical Engineers - Part P: Journal of Sports Engineering and Technology*, vol. 233, no. 4, pp. 455–466, 2019.
- [23] K. S. Ananda Kumar, R. Balakrishna, A. Y. Prasad, B. Worku, and K. Salih Siraj, "Development of integrated IoT application on vehicle tracking, traffic monitoring and vehicle theft," *International Journal of Future Generation Communication and Networking*, vol. 13, no. 4, pp. 1–10, 2020.
- [24] V. M. Reynier, B. J. Soulé, and J. M. Pabion-Mouriès, "Skiing and snowboarding injuries in snow parks: a study based on self-reported practitioners' statements," *Wilderness and Environmental Medicine*, vol. 31, no. 2, pp. 181–187, 2020.
- [25] S. S. Komissarov, "Mechanics of side-slipping in alpine skiing: theory of machining snow and ice," *Sports Engineering*, vol. 24, no. 1, pp. 1–9, 2021.
- [26] M. W. Dipl-Ing, P. Haegeli, and S. Fuchs, "Risk of death and major injury from natural winter hazards in helicopter and snowcat skiing in Canada," *Wilderness and Environmental Medicine*, vol. 30, no. 3, pp. 251–259, 2019.
- [27] M. Takeda, N. Miyamoto, T. Endo et al., "Cross-country skiing analysis and ski technique detection by high-precision kinematic global navigation satellite system," *Sensors*, vol. 19, no. 22, p. 4947, 2019.
- [28] Z. Xia, J. Xing, C. Wang, and X. Li, "Gesture Recognition Algorithm of Human Motion Target Based on Deep Neural Network," *Mobile Information Systems*, vol. 2021, Article ID 2621691, 12 pages, 2021.