Golf Swing Correction Based on Deep Learning Body Posture Recognition

Gaoyue, Sun*
The High School Affiliated to Renmin University of China
gfm8@163.com

ABSTRACT

With the rapid development of society and economy, people pay more and more attention to sports. Among them, golf has become more and more popular in China. However, golf swings are highly refined and usually require to be instructed by professional coaches. Actually, the time and economic cost of finding coaches have caused some degree of inconvenience to fast-paced urbanites. And sometimes wrong actions are difficult to be detected with naked eyes. This problem can be effectively solved by applying computerized human body recognition technology to the correction of golf swing. Based on the existing visual solutions for human pose recognition. This paper proposed a method of key frame detecting in video streams, and proposed a posture restoration based on pseudo angular velocity for the error detection problem of the existing Openpose. By quickly detecting key frames in the video stream, not only key skeleton information can be quickly extracted for golf action comparison, but also the amount of calculation can be reduced.

CCS CONCEPTS

• **Human-centered computing** → Visualization; Visualization design and evaluation methods; Visualization; Empirical studies in visualization; Visualization; Visualization application domains.

KEYWORDS

human body gesture recognition, key frame extraction, key point re-estimation, motion correction

ACM Reference Format:
Gaoyue, Sun. 2021. Golf Swing Correction Based on Deep Learning Body Posture Recognition. In *2021 3rd International Conference on Pattern Recognition and Intelligent Systems (PRIS 2021), July 28–30, 2021, Bangkok, Thailand*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3480651.3480713

1 INTRODUCTION

The use of deep learning methods to deal with computer vision problems has become popular in recent years. While effective sports training often requires the guidance of professional coaches, however, such coaches are limited in number and hence expensive. This not only increases the cost of training, but also causes a certain degree of inconvenience to athletes.

Existing Golf Auxiliary Equipment

(1) Record with cameras and observe with the naked eyes by slow down the video.
(2) Test the ratio of the centre of gravity between player's feet.
(3) Detect data of ball when it was hitting, such as, backspin, side spin, speed of the ball, and the speed of the player's swing. [5]

This paper applies gesture recognition technology to golf training, with the aim of allowing athletes to check and compare their actions more easily, as well as to enable coaches to provide more scientific guidance to athletes. In order to measure whether the athlete's actions are standard, the key frames of golf swing are extracted to compare with the standard action with the use of human body gesture recognition technology.

Based on the research on the existing auxiliary golf training equipment, this paper mainly examines how to use deep learning methods to detect and extract human poses and to recover the required bone key point coordinate information, as well as how to extract key frames from a video stream for pose comparison. The main innovations of this article are as follows:

(1) Proposal of a method to quickly locate key frames in the video stream.
(2) Proposal of a posture restoration method based on pseudo angular velocity.

2 HUMAN POSTURE DETECTION TECHNOLOGY

OpenPose is a high-precision multi-person gesture recognition method that adopts the bottom-up process. After an image is input, it uses a convolutional network to extract features that are then divided into two parts. One is used to develop a partial credibility graph, while the other is used to create a partial correlation field that represents the correlation score across the network. Matching these two pieces of information helps to determine the associated values for each point, which can be connected as the joints of a person to form an accurate image of the human skeleton.

3 GOLF SWING REPAIR AND CORRECTION

3.1 Introduction

The OpenPose method used in this paper has been applied in many fields. OpenPose uses a correlation field segment to represent relevant correlation values (PAFs). [4] However, the requirements for gesture recognition for golf actions are higher. There are some specific challenges to work in this field: (1) the number of persons, and the size and location of persons in a picture are often uncertain; (2) some bones and joints will be hidden during golf swing, which may lead to the OpenPose algorithm being unable to accurately

PRIS 2021, July 28–30, 2021, Bangkok, Thailand

ACM ISBN 978-1-4503-9039-2/21/07…$15.00
https://doi.org/10.1145/3480651.3480713

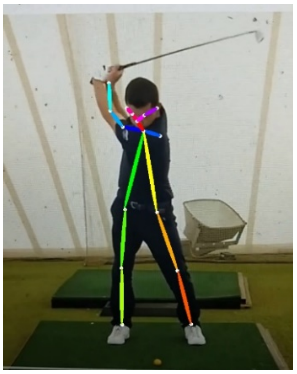

Figure 1: Position misdetection. (https://haokan.baidu.com/v?vid=993055588405337035&pd=bjh&fr=bjhauthor&type=video)

Figure 2: Length misdetection. (https://haokan.baidu.com/v?vid=993055588405337035&pd=bjh&fr=bjhauthor&type=video)

recognise posture information for the person; and (3) due to the rapid movement pictures of joints will be blurred in some frames.

This paper instead proposes a dynamics-based repair method based on the characteristics of golf action, which ensures that repairing posture in a video is more efficient and faster than repairing the posture of a single image.

3.2 OpenPose false detection

Due to the overlapping of the arms or legs of a person, incorrect bone information may be detected, [1] and this can also be caused by the rotation of the body or even the person's clothes. OpenPose may thus not be able to accurately detect bone and joint information in a given position. It may detect misalignment or even fail to detect at all, leading to a need to repair these incorrectly detected postures..

3.2.1 Position misdetection. Joint points may be identified to the wrong location, which will result in false detection problems. In figure 1, the left shoulder was mistakenly detected as being in a normal position; the correct position should be under the chin. The right shoulder was also detected in the normal position for the left shoulder, while the correct position should be behind the head.

3.2.2 Length misdetection. Misdetection of the position of a key point causes problems by implying limbs that are too short or too long. In figure 2, misdetection of the left wrist has caused problems with the length of the left forearm

3.2.3 Undetected joints. Sometimes problems arise where a joint is undetected: in figure 3, the right wrist is not detected.

Human body posture recognition will detect the bone joint point information, but it will inevitably contains some errors. OpenPose may not be able to accurately detect the bone and joint information at these positions. It may detect misalignment or fail to detect. In order to repair these incorrectly detected postures, this paper proposes a posture restoration method based on pseudo angular velocity.

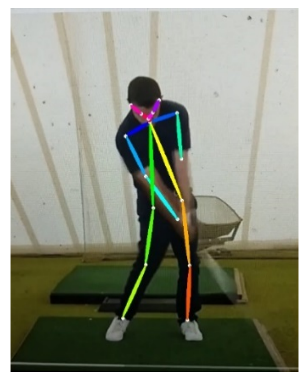

Figure 3: Undetected joint. (https://haokan.baidu.com/v?vid=993055588405337035&pd=bjh&fr=bjhauthor&type=video)

3.3 Assessment of incorrect human bone information

Within the process of gesture recognition, some skeletal joint information is generally incorrectly detected. It is thus necessary to judge whether the detection of human bone information is correct in each case. If the information of a joint is not detected correctly, this must be repaired on outputting the index and information for each frame of bone information.

3.3.1 Position misdetection judgment. Position misdetection is assessed in such a way that when the distance change exceeds a certain threshold, a misdetection is assigned.

There are two types of math equations: the *numbered display math equation* and the *un-numbered display math equation*. Below are examples of both.

3.3.2 Length misdetection judgment. The Euclidean formula is to calculate the initial length of each limb, which is compared with the joint length detected in each frame to determine whether the

a Unrepaired Bone Length

b Repaired Bone Length

Figure 4: Comparison of error detection information before and after repair. (https://haokan.baidu.com/v?vid=9930555884053370358&pd=bjh&fr=bjhauthor&type=video)

length has been misdetected. If the difference exceeds a certain threshold, the estimation of the joint point is deemed to be wrong.

3.3.3 Joint undetected judgment. If the information for a joint is not correctly detected, its joint point index will be output as minus one; this allows this joint point to be marked as not having been detected

3.4 Repair of incorrect skeleton information

3.4.1 Restoration of location misdetection. Human body movements are coherent and continuous, and all limbs are moved around certain joints; the movement of the human body tends to rely on the fact that one joint moves around another joint to create "circular motion". This thus allows the determination of speed according to the positions of the front and rear limbs, allowing repair of a location that has been misdetected.

This can be done using formulas 1 and 2, as derived below:

$$R\_last = (x1 - x2)^2 \times (x1 - x2)^2$$

$$Angel\_all = \omega = (x3 - x2) / (y3 * r - y2 * r)$$

$$angel = tan^{-1}((y3 - y1) / (x3 - x2))$$

$$angel\_change = angel - Angel\_all$$

Applying (4) and (5) simultaneously,

$$x4 = x1 + r * sin(angel\_change) \tag{1}$$

$$y4 = y1 + r * sin(angel\_change) \tag{2}$$

where R represents the radius of the circular motion, $x1$ represents the x-axis coordinate of the centre joint point, $y1$ represents the y-axis coordinate of the centre joint point, $x2$ represents the x-axis coordinate of the joint point two frames before the joint point was missed, $y2$ represents the y-axis coordinate of the joint point two frames before the joint point was missed, $x3$ represents the x-axis coordinate of the joint point one frame before the joint point was missed, and $y3$ represents the y-axis coordinate of the joint point one frame before the joint point was missed. The angular velocity of the joint points and the specific coordinates of the missing joint points can then be calculated.

Based on this, this paper proposes an attitude recovery method based on pseudo angular velocity. From the first two frames of incorrect bone information, a pseudo angular velocity for the disappearing or incorrect joint movement around the normal joint can be calculated, and this angular velocity may then be used to calculate the correct positional coordinates of incorrectly detected joint.

Figure 4a shows a case with unrepaired bone length, while Figure 4b shows the same picture with repairs.

3.4.2 Restoration for length misdetection. As the joints of the body are constantly moving and may be impeded by each other, the detected length of a part of the human body may be shorter or longer than the visual image. The length as seen in the initial frame 0 is thus used as the standard length, and if the detected body part length is within a certain range of deviation, the detected length meets the standard. If it exceeds this range, the detected length in that frame is deemed wrong and requiring of restoration. Formula 3 is thus applied to achieve such restoration:

$$K = (y1 - y2) / (x1 - x2)$$

$$X3 = ((length)^2 / \left(\mathrm{k}^2 + 1\right))^{0.5}$$

$$Y3 = kx \tag{3}$$

where x1 represents the x-axis coordinate of the middle joint point.

In figure 5a, the x-axis coordinate of the right elbow is x1, while y1 represents the y-axis coordinate of the middle joint point; x2 represents the x-axis coordinate of the unrepaired end joint point. In figure 6.b, the x-axis coordinate of the right wrist is x2, and y2 represents the y-axis coordinate of the end joint point before repair; x3 thus represents the x-axis coordinate of the final joint point after repair, while y3 represents the y-axis coordinate of the final joint point after repair.

3.4.3 Restoration of missed joints. The same method of position restoration is used to repair joints where some joints are undetected.

3.5 Key frame detection in golf videos

There are some key actions that significantly affect golf swings. In this paper, the golf swing is divided into seven key actions: setup frame, back-swing frame, top frame, release frame, hit frame, fore-swing frame, and closing frame. The wrist and elbow on the arm

a Length of right arm repaired

b Length of right arm without repair

Figure 5: Comparison of arm length before and after repair.(https://haokan.baidu.com/v?vid=993055588405337035&pd=bjh&fr=bjhauthor&type=video)

are set as key joints, and the distance between them is thus used as the associated value for judging the key frame. Where the distance value is within a set threshold, the frame is identified as belonging to a certain key action; otherwise, the frame is judges to not be part of a key action, though it may be associated with the next key action [3]. Euclidean distance is used to calculate the associated value using formula 4

$$Dis = \sqrt{(y1-y2)^2 + (x1-x2)^2} \tag{4}$$

where x1 and x2 are the x-axis coordinates of the two key joints in the human body, and y1 and y2 are the y-axis coordinates.

All the distances between each of the left elbow, left wrist, right elbow, and right wrist are calculated, and the six associated values obtained are compared with the threshold values to determine the interactional relationships between these key joints to determine whether an image is a keyframe.

There are a certain number of key frames in each of the six key actions of a golf swing. This paper thus proposes the judgment of key frames be undertaken by using Euclidean distance to calculate the sum of the distances between 17 notable joints and the same joints as seen in the previous frame. If this sum is less than a certain threshold, this frame is a key frame; otherwise, it is not. When a frame is determined to not be a key frame, the interval of detecting key frames is multiplied by two, accelerating the number of frames viewed to make the search for key frames more rapid.

3.6 Comparison method for golf key actions

It is difficult to directly define standard golf movements, which are mainly composed of body and arm rotations. The angles and distances between joints are thus used to compare individual differences [2].

1) The characteristics of arm angle

Arm angle can be calculated using formula 5:

$$Agarm = cos^{-1}\left(\frac{(\mathrm{x1}-\mathrm{x2}) * (\mathrm{x2}-\mathrm{x3}) + (\mathrm{y1}-\mathrm{y2}) * (\mathrm{y2}-\mathrm{y3})}{\sqrt{(x1-x2)^2+(y1-y2)^2} + \sqrt{(x2-x3)^2+(y2-y3)^2}}\right. \tag{5}$$

where x1 represents the x-axis coordinate of the shoulder, x2 represents the x-axis coordinate of the elbow, and x3 represents the x-axis coordinate of the wrist, and y1 represents the y-axis coordinate of the shoulder, y2 represents the y-axis coordinate of the elbow, and y3 represents the y-axis coordinate of the wrist.

If the resulting angle is between 160° and 180°, the action meets the standard. Otherwise, the person's action needs to be improved.

The distance between the wrists can be calculated using formula 6:

$$dis_{wrist} = \sqrt{(y1-y2)^2 + (x1-x2)^2} \tag{6}$$

where x1 represents the x-axis coordinate of the right wrist, y1 represents the y-axis coordinate of the right wrist, x2 represents the x-axis coordinate of the left wrist, and y2 represents the y-axis coordinate of the left wrist.

2) Characteristics of hand distance

The distance between the wrists should be close to zero in all key movements except for the setup frame. If the distance between the wrists exceeds a certain threshold, the swing needs to be improved.

3) Characteristics of head displacement

The head should rarely move during a golf swing. The displacement of nose between the back-swing frame and top frame is thus calculated, and if the displacement is within a set threshold, the head movement meets the required standard. Otherwise, the person needs to improve their swing. The displacement of the nose is calculated using formula 7:

$$dis_{noise} = \sqrt{(y1-y2)^2 + (x1-x2)^2} \tag{7}$$

where x1 represents the x-axis coordinate of the nose in the back-swing frame, y1 represents the y-axis coordinate of the nose in the back-swing frame, x2 represents the x-axis coordinate of the nose in the top frame, and y2 represents the y-axis coordinate of the nose in the top frame.

4) Characteristics of leg angle

In this paper, the angle between the left thigh and the left calf is calculated by means of leg detection. If the angle is less than 160° or larger than 180°, this will cause an excessive shift in the centre of gravity and affect the shot. This angle degree is calculated using

Table 1: Experimental Data

Experimental sample size: about 25,000	
Has it been corrected or not	accuracy
Before correction	98.7%
After correction	99.9%

formula 8:

$$Agleftleg = cos^{-1}\left(\frac{(x1-x2)*(x2-x3)+\ (y1-y2)*(y2-y3)}{\sqrt{(x1-x2)^2+(y1-y2)^2}+\sqrt{(x2-x3)^2+(y2-y3)^2}}\right) \quad (8)$$

This formula is same as the calculation formula for calculating the arm angle, except that (x1, y1) is used for the thigh, (x2, y2) for the knee, and (x3, y3) for the ankle.

If the angle is larger than 160° and less than 180°, the player's shift in centre of gravity is considered appropriate; otherwise, the player's swing needs to be improved.

5) Spine line features

Spine line displacement often causes instability in a shot; instead, the player should keep their spine stable. The gradient of the midpoint of both legs and hips is calculated to make a judgement on this factor: if the gradient exceeds a certain value, the swing needs to be improved. Formula 9 is used to calculate this gradient as follows:

$$G = \ (y1 - y2)\,/(x1 - x2) \quad (9)$$

where (x1, y1) are the coordinates of the nose and (x2, y2) are the coordinates of the midpoint of the legs.

4 EXPERIMENT AND ANALYSIS

4.1 Experimental materials and sources

The photos for this experiment were mainly taken from internet and some are taken by the researcher, with some materials downloaded from the teaching module provided by DaZheng Golf.

4.2 Experimental platform and method

The experimental platform used for this experiment was BitaHub, in its "brain-like computing center" version, 2.0.

In this experiment, OpenPose was initially used to analyse a single photo and then to analyse a full swing motion video. The relevant x and y axis coordinates, and the index and score of each joint point were then output into a three-dimensional matrix for subsequent calculation. After this matrix was obtained, the posture errors were repaired and the key frames detected and compared to the correct movements. This allowed production of the full experimental results.

4.3 Experimental results and analysis

In this experiment, OpenPose was found to correctly recognise most of the joint points; after further posture restoration, the posture recognition suffered very little from the problem of misdetection. The experimental data is stated in the Table 1.

5 CONCLUSION AND OUTLOOK

This paper combines human body posture recognition with traditional golf advice to provide guidance for golf players during their self-training. Posture restoration for golf actions, key frame detection in video streams, and key golf action comparison methods were then outlined. The paper explained the error detection problems inherent in OpenPose, and then introduced a method of judging detection errors so that the proposed method of golf posture restoration, based on pseudo angular velocity, can be applied. The paper then discussed the identification of golf key frames and a comparison method for golf actions. Finally, the experimental platform was introduced and the experimental results analyzed.

This paper mainly studied the correction of basic golf actions, the recognition of the side posture and golf clubs maybe be developed and added to this system.

ACKNOWLEDGMENTS

During the writing of the paper, there are too many teachers to thank. Thanks to Teacher Wan, Teacher Zhao and Teacher Xu for guiding me.

REFERENCES

[1] A. Bulat and G. Tzimiropoulos. Human pose estimation via convolutional part heatmap regression. In ECCV, 2016.
[2] Ji Yuepeng, A comparative analysis of golf swing based on video human posture estimation. 2019.
[3] Yang Wenlu, Yu Mengmeng, Xie Hong. Two-person interactive behavior recognition based on key poses [J/OL]. Computer Application.
[4] Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh, The Robotics Institute, Carnegie Mellon University:Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.2017.
[5] Zhu Yifeng:Design and Implementation of Golf Teaching Aided Analysis System Based on Target Tracking Technology. 2019