Human Motion Attitude Tracking Method Based on Vicon Motion Capture Under Big Data

Ze-guo Liu(✉)

College of Information Engineering, Xizang Minzu University, Xianyang, China
liuzeguo316@163.com

Abstract. Aiming at the problem that the human body motion posture cannot be correctly and quickly marked in the conventional method, a human body motion attitude tracking method based on Vicon motion capture under big data is proposed and designed. Under the motion capture filtering algorithm, the human body weight measurement function is constructed by the combination of color, edge and motion features, and different images are selected according to the occlusion between limbs to establish a constrained human motion model, and the model is based on Vicon action. The tracking calculation of the capture realizes the tracking process of the human body motion posture. The effectiveness of the method is determined by the method of experimental argumentation analysis. The results show that the method can track the motion posture of the human body quickly and accurately, and the robustness is better. The tracking accuracy is 13.87% higher than the conventional method.

Keywords: Big data technology · Vicon motion capture · Human body movement posture · Tracking method

1 Introduction

At present, the tracking and recognition of human motion based on machine vision is a popular research direction in the world. It has universal applications in video surveillance, video retrieval, human-computer interaction, virtual reality and sports training. Using machine vision to track and identify human motion information in video can greatly reduce the original labor cost and complete the recognition process of human motion more quickly and accurately [1]. In order to improve the tracking rate of human body motion posture, it is extremely necessary to establish a recognition model based on attitude features. The human body motion posture features are divided into external features and internal features. External features include shape, contour, etc. Internal features include color, texture, etc. Since the human body motion posture is mainly composed of the above features, the tracking method based on human motion capture is one of the most important methods at present.

The Vicon motion capture method can capture the reflective sphere moving in the scanning space according to different cameras at the same sampling time, and obtain

G. Gui and L. Yun (Eds.): ADHIP 2019, LNICST 301, pp. 383–391, 2019.
https://doi.org/10.1007/978-3-030-36402-1_41

the three-dimensional coordinates of the reflective sphere at that moment [2]. Based on these coordinates, kinematics and dynamics analysis can be performed. The displacement, velocity, acceleration, and the variation of physical quantities such as momentum and kinetic energy are obtained. At the highest resolution, the Vicon motion capture method has a sampling frequency of 482 *spf* and a maximum sampling frequency of 10,000 *spf*. Therefore, Vicon's matching motion capture software is used to track the human body's motion posture, and its efficiency and accuracy are extremely high.

Although the conventional tracking method can capture the human body's motion posture, but the tracking time is long, and the capturing accuracy is not ideal. Based on this, this paper proposes and designs a human body motion attitude tracking method based on Vicon motion capture, and demonstrates effectiveness of our proposed method through experiments. The method of analysis verifies the effectiveness of the method. The results show that the human body motion attitude tracking method based on Vicon motion capture can improve the tracking speed of human motion posture and ensure the accuracy of tracking results, which is extremely robust.

2 Design of Human Body Motion Attitude Tracking Method

2.1 Reference Motion Capture Filter

Vicon motion capture filter provides a probabilistic verification of non-Gaussian, non-linear and multi-module observation data, while human pose tracking is a tracking problem in high-dimensional space [3], which is solved by using motion capture method. In the case of a problem, the high-dimensional state space is usually decomposed into two or more subspaces to apply dynamic processes and resampling processes in each subspace.

Consider the tracking problem in the 2-dimensional state space $s = (x, y)$ consisting of two 1-dimensional subspaces $S_A(S_A \in R)$ and $S_B(S_B \in R)$. In order to avoid searching for the correct image state in the entire 2-dimensional space, block sampling divides the search process into two steps [4]. That is, the state x is searched in the subspace S_A, and then the state y is searched in the subspace S_B. Since the survival factor of each step is a, the survival factor for 2-dimensional state tracking in the 2-dimensional space s is $2a$, and the required capture path is $2D_{\min}/a$, and the 2-dimensional state is directly performed in s using the motion capture filtering method. The survival factor of the trace is a^2, and the number of particles needs to be $D_{\min}/a^2$. Among them, represents the minimum acceptable survival diagnosis required for successful tracking [5]. Since a is much smaller than 1, the motion capture filtering method can greatly reduce the number of required capture paths and improve tracking efficiency.

Let each joint point of the human motion correspond to a filter set $\gamma_i(i = 1, 2, \cdots, n)$, and let $\gamma_i = D_{\min}/a^2$ update the joint image of the human body after the filtering calculation is finished. The position of the corresponding joint point in the image at time

t is obtained by calculating the mean of each joint in the filter set. The filtering process at each joint point is as follows:

Since the number of images in each filter set is small, the sampling process is easily degraded [6]. Therefore, we perform "re-sampling" at each time step to complete the selection process of the human body motion image.

Considering that the human body pose changes between two frames of images is small, the method of this paper does not use motion models for prediction. At the time t, the tracking result of the time $t-1$ is used as the mean value, and N filters having a Gaussian distribution are generated. Among them, the variance $\varpi = \gamma w_0$ of the Gaussian distribution, by adjusting the value of γ, changes the filter distribution area of the new image, and finally obtains the prediction result of the image Gaussian distribution.

The measurement process of the human body is realized by the combination of color, edge and motion characteristics. The measurement process of the head uses only the color features, and the measurement process of the limbs uses the fusion of the three features. The occlusion and occluded limbs are measured using different features to complete the measurement of each initial parameter [7]. By calculating the filtering state in the filter set corresponding to each joint point, the mean value is accurately estimated, and the human body weight measurement function is obtained:

$$s_k = \sum_{i=1}^{N} \pi \varpi^2 \tag{1}$$

Where s_k represents the weight measure of the human body motion; k is a constant, $k = 1, 2, \cdots, N$.

After the above definition, the estimation function of the human body motion posture is obtained, and the function is analyzed and backed up to prepare for the tracking calculation of the human body motion posture.

2.2 Establish a Constrained Human Motion Model

A rectangular model is used to establish a human body model, and the constraint relationship between the limbs is represented by a line connecting the nodes. Both frame models use nodes to describe the limbs [8], and the connections between the nodes represent the relationship between the limbs. The model structure is shown in Fig. 1.

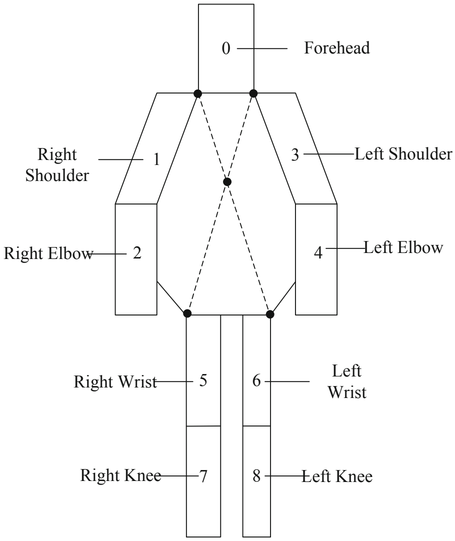

Fig. 1. Schematic diagram of the node relationship of the human body model

Analysis of Fig. 1 shows that the constrained human body appearance model designed in this paper consists of 9 "cardboards" connected by joints. Among them, there are 14 nodes, 1 virtual node and the connection between nodes. For the virtual node, Since it does not play a fixed role, it does not need to be tracked [9]. The nodes in the human body model correspond to the real joint points, the solid lines represent the real limbs, the dashed lines and arcs represent the constraints between the two adjacent joint points, and the dashed arrows indicate the one-way constraints that only affect the joint points indicated by the arrows.

It is assumed that the spatial state of the human pose is represented by X, where X is the Cartesian product $X = X^1 \times X^2 \times \cdots X^{14}$ of 14 sub-state spaces, and each sub-state space is a 2-dimensional vector space. Thus, the tracking problem in a 2-dimensional state space translates into tracking problems in two planar state spaces.

Using the Vicon motion capture method, the length of the limb, the neck and shoulders, the distance between the shoulders, the sternum, the sternum and the hips are constrained [10]. Let ε denote the distance between two nodes, then the quantized constraint can be expressed as a Gaussian function:

$$P(\varepsilon) = \frac{s_k}{\sqrt{2\pi\vartheta}} \exp\left(\frac{-(\varepsilon - \varepsilon_0)^2}{2\vartheta^2}\right) \tag{2}$$

Where $P(\varepsilon)$ represents the constrained quantification result of the human body motion posture; ε_0 represents the distance between the two joint points initialized; ϑ is the variance of the Gaussian function, since the inflection point of the Gaussian function is located at $\varepsilon = \varepsilon_0 + \vartheta$, when $\varepsilon \in \varepsilon_0 - \vartheta$, the curve of $P(\varepsilon)$ is steep; When $\varepsilon \in \varepsilon_0 + \vartheta$, the curve of $P(\varepsilon)$ becomes gentle. It can be seen that the larger the ϑ value, the looser the corresponding constraint, and the larger the allowable variation range of ε; conversely, the smaller the ϑ value, the tighter the constraint, and the smaller the allowable variation range of ε.

At this point, the constraint design of the human body posture model is completed, and the corresponding limb length constraint is determined after the initialization calibration is completed.

2.3 Design of Human Body Motion Attitude Tracking Algorithm

The human body motion attitude tracking calculation adopts an adaptive method, and combined with the Vicon motion capture method, the spatial position of the human motion posture is defined by selectively using the likelihood functions corresponding to the three characteristics according to the occlusion condition between the limbs [11]. And determine the conflict set S_0, representing a collection of limb samples that are covered. It should be pointed out that since this article only considers the situation that the human body is facing the camera, we think that the big arm is blocked when the big arm and the small arm are blocked. In the calculation, only the strong relationship between the limbs is considered, that is, only the occlusion relationship between the upper limbs and the lower limbs is considered, regardless of the occlusion relationship between the head, the upper limbs and the lower limbs. We use the motion capture filter and the constrained human body model introduced above to determine the mutual occlusion relationship of the human body's motion posture, as shown in Fig. 2.

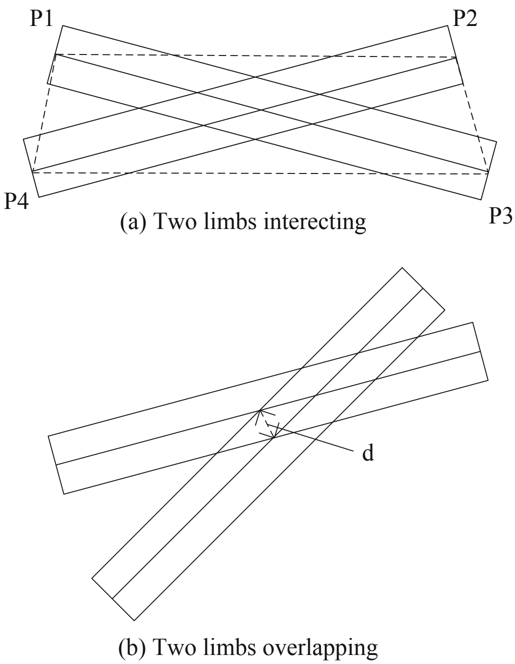

Fig. 2. Conflict between two limbs

According to the conflict relationship between the two limbs in Fig. 2, when the intersection P_0 of the mid-line P_1, P_2, P_3 and P_4 of the two limb samples is located in the quadrilateral $P_1P_2P_3P_4$, the two limb samples intersect, and the constrained quantitative result $P(\varepsilon)$ of the human motion posture will appear. Repeat the calculation phenomenon. In order to avoid such phenomena, this calculation introduces an adaptive method to bring the constraint relationship of the human body model into the adaptive calculation process, so that only the filter that meets $P(\varepsilon)$ can pass the calculation.

According to the coverage relationship of the two limb samples in Fig. 2, it is assumed that l is the distance between the midpoints of the centerline of the two limb samples; d is the distance from the end point of the centerline of one limb sample to the centerline of the other limb sample. When both l and d are less than the threshold [12], the two limb samples are considered to be covered. The judgment is calculated as follows:

$$\kappa = \frac{P(\varepsilon)}{(l+d)^2} \tag{3}$$

Where κ represents the extent of coverage of the limb sample. $(l+d)^2$ is the judgment coefficient, $(l+d)^2 \in W + W_1/4$.

When the spatial positional relationship of all limb samples is judged, the occluded limb samples are placed in the filter set. For all limb samples, the likelihood of each limb sample is calculated using the likelihood function corresponding to the color, edge, and motion features. If the various key points of the human body have been completely occluded, the limb sample occluded with the limb sample is used for corresponding calculation, and a tracking model of the human body motion posture is obtained:

$$c = \frac{S_m/\gamma}{N\kappa} \tag{4}$$

Where c represents the mathematical expression of the tracking model.

Through the above definition, the human body posture weight measurement function obtained under the motion capture filtering algorithm is combined with the constrained human motion model to realize accurate and fast tracking of the human body motion posture.

3 Experimental Argumentation Analysis

In order to verify the effectiveness of the human body motion tracking method based on Vicon motion capture, three videos containing complex human motion were captured using the Cano IXUS 8015 digital camera. The video size was 640 × 480, and the average human height in the video was 313 pixels. The number of motion captures corresponding to the node is 20. There are 2 indoor video and 1 outdoor video. At the same time, in order to ensure the rigor of the experiment, the conventional tracking method was used to compare with the method.

The algorithm was implemented using VC++6.0, and the video clips were tested on an Intel Core2 computer with 2.0 GHz CPU. The average processing time was 5.3 spf, the processing time of each video and the human scale in the video and the joint points. The number of motion captures is proportional.

In this method, the number of captures of human joint movements and their distribution have an important influence on the performance of the method. Due to the interdependence of these two parameters, we simultaneously adjust the two parameter values and analyze them on a test video of 200 frames. The experimental performance, that is, the tracking accuracy and speed of the human motion posture, obtains the parameter value: the adjustment factor $\gamma = 0.6$ of the motion capture range, and the experimental results have the best performance when the number of motion captures $N = 25$, and the fixed values of the two parameters are used in the subsequent experimental tests.

3.1 Tracking Accuracy Comparison

The average tracking error of the two methods on the human body motion posture is shown in Table 1.

Table 1. Average tracking error of human joint points

Joint	Indoor 1	Indoor 2	Outdoor	Average tracking error of this method %	Average tracking error of conventional method %
Forehead	4.62	5.26	3.14	7.25	16.35
Neck	4.16	4.36	3.36	4.36	9.25
Shoulder	8.45	5.21	8.14	9.24	21.36
Elbow	6.32	3.95	5.16	1.36	7.25
Hip	6.58	5.32	3.24	5.27	15.26
Knee	4.85	6.14	4.48	7.15	14.15
Ankle	6.17	4.17	5.65	9.25	19.26

Because different sports have different technical requirements, the main parameters of the body's limb motion analysis should be determined according to the specific research of the sports project, with certain uncertainty, so there must be certain errors in the tracking results. It can be seen from the data in Table 1 that for the tracking result of the same test video, the average error of the method is smaller than the average error of the conventional method. The calculation shows that when the Vicon motion capture method is used to track the human body's motion posture, the tracking accuracy is 13.87% higher than the conventional method.

3.2 Tracking Speed Comparison

The comparison results of the average tracking speed of the two methods on the human body motion posture are shown in Fig. 3.

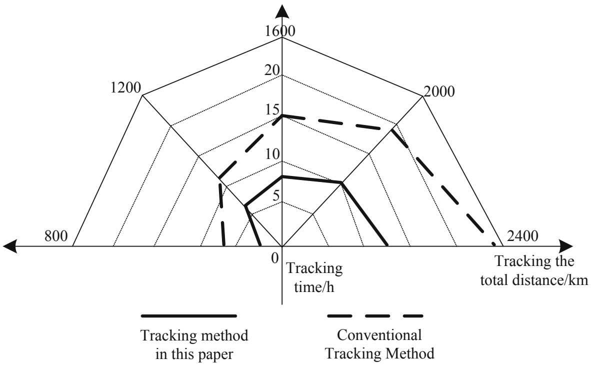

Fig. 3. Comparison of human body motion tracking speed

It can be seen from the experimental results that the proposed algorithm achieves a satisfactory tracking speed of human motion. The analysis shows that under the same tracking path, the tracking time of this method is lower than the conventional method, including complex background, human motion, clothing changes and occlusion between limbs, which can quickly capture and track the human body posture. The tracking speed of the conventional method is lower than the method of this paper. Therefore, the effectiveness of the proposed method can be determined. The results show that the human motion attitude tracking method based on Vicon motion capture can achieve satisfactory tracking results.

4 Conclusions

This paper proposes a human body motion attitude tracking method based on Vicon motion capture, which combines the color, edge and motion characteristics of human body image, combined with motion capture filter and human motion model to successfully realize the human body motion posture under big data. track. The experimental results show that the proposed method not only effectively reduces the capture path required for tracking, but also improves the operation efficiency, and greatly improves the accuracy of tracking, solves various forms of limb occlusion, and is satisfactory. Tracking Results.

Acknowledgements. On the digitalized 3D model of Tibetan Xianziwu based on 3D Motion Capture (Innovation-Supportive Project for Young Teachers in Colleges and Universities in Tibet Autonomous Region) QCZ2016—33.

References

1. Anonymous. Research on the motion characteristics of left and right ankle joints of Vicon motion capture system during human walking. Modern Inf. Technol. **2**(6), 15–17 (2018)
2. Wei, Y., Fan, X.: Tracking and recognition of human motion posture based on GMM. J. Beijing Inst. Fashion (Natural Science Edition) **100**(2), 47–55 (2018)
3. Wang, M., Hu, M., Yang, W.: Motion accuracy simulation of robot tracking gesture image. Comput. Simul. **34**(8), 346–350 (2017)
4. Gang, J., Nan, Z.: An athlete posture analysis model for tracking human joint points with multi-feature optical flow. Sci. Technol. Bull. **15**(12), 133–136 (2017)
5. He, K., Ding, X.: Human behavior tracking method based on low-dimensional manifolds. Comput. Eng. Des. **38**(5), 1361–1365 (2017)
6. Wang, Y., Li, H., Qi, H., et al.: A robust moving target tracking method inspired by three-stage memory mechanism of human brain. J. Electron. **45**(09), 12–17 (2017)
7. Man, W., Zhu, Z., Zhang, Z., et al.: Human motion posture analysis using synchronous extrusion wavelet transform. J. Xi'an Jiaotong Univ. **51**(12), 8–13 (2017)
8. Zhiping, W.: Pedestrian tracking algorithm based on human body feature recognition and Kalman filter. Electro-Optic Control **46**(11), 97–102 (2016)
9. Ying, Z., Ying, Z.: Research on motion target tracking algorithms based on computer vision. J. Chifeng Univ. (Nature Edition) **33**(4), 12–14 (2017)
10. Liu, S., Qiu, Z., Ding, Q., et al.: Research on head attitude tracking method based on adaptive filtering. Electro-optic and Control **91**(4), 33–36 (2016)
11. Wang, Q., Gong, L., Dong, C.: Passive attitude tracking control of Spacecraft Based on time-varying gain ESO. Control Decis.-Making **43**(2), 193–202 (2018)
12. Yin, C.: Multi-loop recursive attitude tracking control for capturing spacecraft. Space Control **12**(1), 54–56 (2018)