



A high-precision correction method in non-rigid 3D motion poses reconstruction

Cuihong Fan, Weina Fu & Shuai Liu

To cite this article: Cuihong Fan, Weina Fu & Shuai Liu (2022) A high-precision correction method in non-rigid 3D motion poses reconstruction, Connection Science, 34:1, 2845-2859, DOI: [10.1080/09540091.2022.2151569](https://doi.org/10.1080/09540091.2022.2151569)

To link to this article: <https://doi.org/10.1080/09540091.2022.2151569>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 04 Jan 2023.



Submit your article to this journal [↗](#)



Article views: 160



View related articles [↗](#)



View Crossmark data [↗](#)



A high-precision correction method in non-rigid 3D motion poses reconstruction

Cuihong Fan^a, Weina Fu^b and Shuai Liu^{b,c,d}

^aCollege of Education and Sports Sciences, Yangtze University, Jingzhou, People's Republic of China; ^bCollege of Computer Science and Engineering, Hunan Normal University, Changsha, People's Republic of China;

^cSchool of Educational Sciences, Hunan Normal University, Changsha, People's Republic of China; ^dKey Laboratory of Big Data Research and Application for Basic Education, Hunan Normal University, Changsha, People's Republic of China

ABSTRACT

Occlusion, rotation and other factors affect human motion structure because of the incomplete acquired image sequence, resulting in poor performance of non-rigid three-dimensional (3D) motion pose reconstruction. A non-rigid 3D reconstruction and high-precision correction method for motion pose are studied in this paper. A non-rigid imaging model is designed to obtain 3D moving images. According to the frame difference and morphological processing, the background of image is separated and denoised. Combined with motion analysis, 3D motion pose features are extracted as identification of non-rigid 3D motion error actions in a hybrid Convolution Neural Network-Hidden Markov Model to train the correction coefficients, which are used to adjust the pose in 3D motion reconstruction and realise correction. Experimental results show that this method has high precision reconstruction and correction of non-rigid 3D motion pose.

ARTICLE HISTORY

Received 21 September 2022
Accepted 19 November 2022

KEYWORDS

Non-rigid; 3D motion; pose correction; motion reconstruction; correction coefficient

1. Introduction

At present, the tracking and recognition of three dimensional (3D) human motion has become a research hotspot in image processing and computer vision (Wang et al., 2021). Related methods utilise single or multiple cameras to detect, recognise, track, and understand human bodies from frame sequences or videos containing human motion (Dhiman et al., 2021). A State-of-the-arts (SOTA) tracking method of human movement is 3D feature reconstruction, which mainly includes pattern recognition, image processing, computer vision, and shape feature expression analysis (Jahren et al., 2021). Because the human motion is in a non-rigid motion state set with complex and uncertain background, the reconstructed human pose mainly shows a high volatility and low accuracy.

CONTACT Shuai Liu ✉ liushuai@hunnu.edu.cn College of Computer Science and Engineering, Hunan Normal University, Changsha 410081, People's Republic of China, School of Educational Sciences, Hunan Normal University, Changsha 410081, People's Republic of China, Key Laboratory of Big Data Research and Application for Basic Education, Hunan Normal University, Changsha 410081, People's Republic of China

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many scholars have proposed a verity method in this area. Włodarczyk-Sielicka et al. designed an automatic refinement reconstruction algorithm for human body based on limited angular depth data (Kulikajewas et al., 2021). Aiming at the limitation that most of the current researches use a single static human body as the reconstruction object, a new three-stage deep adversarial neural network structure is constructed to remove the noise of the initial collected 3D image and use it as the input of the depth sensor to achieve reconstruction of human pose. The algorithm can accurately measure the moving distance and the chamfering distance on a synthetic 3D image data set, and can reconstruct the deformed objects in the dynamic scene without external influence. Kim et al. proposed a method for 3D object reconstruction from cross-sectional data (2022). The method is based on the AC (Allen-Cahn) equation with the source term for the shape transformation of the target object. The governing equation was explicitly decomposed into three steps by operator splitting method, which was solved by Euler method and variable separation method. Finally, the method was applied to the multi-layer data of human body posture to complete 3D reconstruction. Although the current method has achieved good application results, the effect of 3D motion pose reconstruction should be further optimised.

Also, Yu et al. proposed a high-precision 3D model for 3D reconstruction of large low texture scenes (2020), a mesh motion statistical feature matching algorithm, and calculated the number of matching points in the neighbourhood. The image sequence is divided into several subsequences, and each subsequence is allocated and optimised independently, which solves the verification problem of camera attitude drift when the cumulative error gradually increases. Hasenberger et al. proposed a 3D morphological object reconstruction method (2020). AVIATOR algorithm was used to analyse the shape in the projection image, and 3D reconstruction method was used to reconstruct its model.

In this paper, a high-precision reconstruction and correction method for non-rigid 3D motion pose is designed. The method is experimented on the false movement recognition and correction of dance pose, and the good application results are validated using the effectiveness of the proposed method.

2. Non-rigid 3D motion pose reconstruction and high-precision correction

2.1. Feature extraction of non-rigid 3D moving images

Assuming that the process of non-rigid motion imaging needs to calculate the relevant benchmark and evaluation conditions (Atkinson & Becker, 2021; Jensen et al., 2021), which can be expressed as:

$$\alpha n = \begin{bmatrix} HS & Hk \end{bmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} \quad (1)$$

where α is the depth factor, $n = \begin{bmatrix} v & u & 1 \end{bmatrix}^T$ is the homogeneous coordinate of an image point in the non-rigid 3D moving, $z = \begin{bmatrix} x & y & z \end{bmatrix}^T$ is the in-homogeneous coordinate of the point in 3D space, S and k are the camera's external parameter matrices, and H is the camera's internal parameter matrix.

Assuming that M is the number of 3D space points, the j -th non-rigid 3D motion frame can be calculated as follows:

$$\alpha N_j = [HS_j \quad Hk_j] \begin{bmatrix} z_j \\ 1 \end{bmatrix} \quad (2)$$

In Equation (2), $N_j = [\alpha_{j,1}n_{j,1} \quad \alpha_{j,2}n_{j,2} \quad \cdots \quad \alpha_{j,M}n_{j,M}]$ and $z_j = [z_{j,1} \quad z_{j,1} \quad \cdots \quad z_{j,M}]$ are the image matrix composed of the j -th non-rigid 3D moving image points and the depth factors, and the space matrix composed of all 3D space points in the j -th frame. When the target moves non-rigid, z_j is considered to be composed of K shape baselines in Equation (3).

$$z_j = \alpha \sum_{s=1}^K \beta_{j,s} R_s \quad (3)$$

where $\beta_{j,K}$ and R_K are the weight and shape bases. Introducing Equation (3) into Equation (2), we have Equation (4) as follows.

$$N_j = [HS_j \quad Hk_j] \begin{bmatrix} \alpha \sum_{s=1}^K \beta_{j,s} R_s \\ 1 \end{bmatrix} \quad (4)$$

Then, with expansion and organisation of Equation (4), we have Equation (5) to extract features of non-rigid 3D motion.

$$N_j = [\beta_{j,1}HS_j \quad \cdots \quad \beta_{j,K}HS_j \quad Hk_j] \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_K \\ 1 \end{bmatrix} \quad (5)$$

2.2. Denoising on non-rigid 3D moving frames

According to the features of the extracted 3D motion, the frame difference and morphological processing are used to denoise the extracted image. In the acquired N_j by Equation (5), there is background fused with foreground in the 2D background, such as lighting, equipment, etc., which increases the complexity of the reconstruction and correction of the target's action. Therefore, texture enhancement is used to obtain clearer features of foreground, thus reducing the difficulty of movement reconstruction. In this paper, the Gaussian frame difference method is first used to remove the interference background in the 2D moving frames (Ahmad et al., 2021), and then the denoising of depth information is considered next by considering the particularity of non-rigid 3D continuous frames (Ilesanmi & Ilesanmi, 2021).

According to the obtained two neighbouring frames in 3D movement frames N_j and N_{j+1} , the image texture of the 2D level is processed by Equation (6).

$$F(o) = \begin{cases} 0, & |N_j(o) - N_{j+1}(o)| \geq Thres \\ 1, & |N_j(o) - N_{j+1}(o)| < Thres \end{cases} \quad (6)$$

where F is an auxiliary binary image recording the value change of each point between the two frames N_j and N_{j+1} . The difference between the values of the pixel o at the same position on the j -th and the $j + 1$ -th frames is calculated, and binarised according to the absolute value of the difference and the pre-set threshold $Thres$. If the difference between the pixel values is greater than $Thres$, the point o is classified as the non-rigid moving target because the surface has changed; otherwise it is classified as the background. Then, the pixel value at the corresponding position in F is set to 1 or 0. At the same time, when the change between neighbouring frames is relatively gentle, the points with large difference between these two frame images are regarded as abnormal and removed.

Considering that the non-rigid 3D motion frames contain more noise, image morphological operations can be used to remove noise contained in two-dimensional images (Mahdaoui et al., 2022). In the auxiliary binarised image, the moving foreground is corresponding to a connected large area, and the unauthenticated pixels (Al-Otum & Ibrahim, 2021) with a small area are classified to noise. In addition, erosion and dilation operation are used to clear the frames with a structural element.

The motion map feature N_j is set (the part where the pixel value is equal to 1 is denoted as F_1 , and the part equal to 0 is denoted as F_0), the structuring element is A , and the erosion and dilation operation can be expressed as Equations (7)–(8)

$$N_j \odot A = \{o | F_1 \in A\} \quad (7)$$

$$N_j \oplus A = \{o | F_1 \cap A = \emptyset\} \quad (8)$$

Then, following operations are performed on F and N_j as Equation (9).

$$N'_j(o) = \begin{cases} N_j(o), & F_1 \\ 0, & F_0 \end{cases} \quad (9)$$

By adding 3D depth information, the calculated N'_j is the initial feature of the non-rigid 3D moving frame with background and noise removal. The depth information reflects the distance from the target to the optical centre of the depth camera. Meanwhile, the certain distance between the foreground and background shows a gap between the depth value of the pixel point on the target and background. In this way, a threshold is found to separate the depth data of the moving target and the background, and the target information in the background is deleted, the depth information of the moving target is kept.

Unlike the greyscale image that has only 256 values, it is necessary to segment the depth value into fragments because the possible value range of the depth is large. Then, the number of pixels in each depth fragment is counted. The specific statistical method in this paper is: (a) the depth value is segmented by 100, which means a point with depth c is segmented into the fragment $[c/100]$. It should be pointed out that in the depth matrix obtained through the above preliminary processing; most points have a value 0, and the distribution of the depth histogram is discontinuous (Dharejo et al., 2021). Since the moving target has a certain speed, the gap between the depth of the target and the background can be detected (Mora-Martín et al., 2021). The continuous depth fragments with the smallest nonzero depth value are corresponding to the moving target, and the remaining fragments are corresponding to the background. The number of pixels in the fragments between them is 0, which is regarded as the gap between the foreground target and the background. Therefore, an appropriate value $C - Thres$ is selected within this range as the threshold to

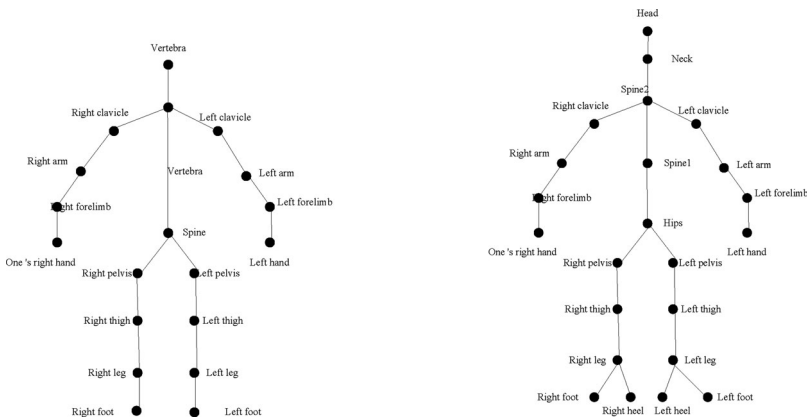


Figure 1. Limb vector structure and 3D human skeleton hierarchy.

segment the frames. The pixels remain unchanged if less than $C - Thres$, the pixels larger than $C - Thres$ are regarded as the background, and removed by setting its value to 0 by using Equation (10).

$$N'_j(o) = \begin{cases} N'_j(o), & N'_j(o) \leq C - Thres \\ 0, & N'_j(o) > C - Thres \end{cases} \quad (10)$$

2.3. Extraction of 3D motion pose feature

The premise of feature extraction is to classify the features with high accuracy (Kumarasamy & Rajendran, 2021). Before the 3D pose feature extraction in this paper, the human body motion is analysed first, so as to determine the type of pose action feature to be extracted.

2.3.1. Human action description model

(1) Define the initial pose

According to the ISB standard, the biomechanical motion recognition of the human body in $N'_j(o)$ is shown in Figure 1. The initial pose of different people will be different due to factors such as height and limb length. Therefore, this paper defines the human body structure as a 3D hierarchical structure to eliminate the structural differences of human body by constructing a 3D human skeleton hierarchy with 23 nodes.

(2) Define the joint motion plane

In order to describe the motion flexibility of the skeleton nodes of the 3D human skeleton hierarchy in $N'_j(o)$, the joint motion plane needs to be defined to ensure the completeness of the description. In this paper, the direction from one shoulder to the other shoulder of the human body in the initial pose is taken as the Horizontal direction, and the Vertical direction is the body vector at the anatomical position. Currently, the motion on different planes

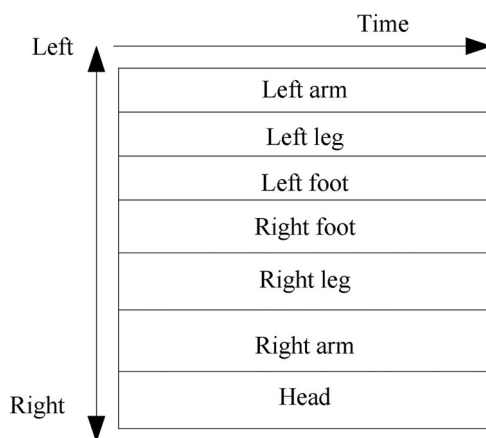


Figure 2. Human body structure.

results in different behaviours of joint mechanics and muscle movements, and biomechanical rules are not considered when calculating joint angles. To solve this problem, this paper describes non-rigid 3D motion gestures from 4 force effect attributes containing space, time, gravity, and fluidity based on ISB and Laban motion analysis theory.

2.3.2. 3D motion analysis

Through the analysis of the human motion description model, the extracted pose motion feature types are judged. Since current 3D image processing methods only judges the motion direction of the standard 2D plane (Sungheetha & Sharma, 2021), this paper introduces the Laban movement analysis theory to further analyse the motions by using human body structure, spatial orientation, and movement force effect.

(1) Analysis of human body structure

Human limbs are represented in the form of “columns” to record the movements of each part. Since the action body in a “column” may correspond to multiple joint nodes, it is necessary to redefine the descriptions of body when parsing motion capture data.

The motion capture data used in this paper is a multi-joint skeleton hierarchy. The body structure is divided by aggregating similar joint points, and the orientation changes of the limbs and head are mainly analysed. Figure 2 is the description of the human limb structure determined in this paper. The seven main body parts, from top to bottom are left arm, left leg, left foot, right foot, right leg, right arm and head, each row represents a limb part.

(2) Spatial orientation analysis

The orientation describes the action area of the limbs. In this paper, the orientation description method is determined by the Laban movement space, and the movement orientation analysis method is determined by dividing the spatial orientation of the human limb in $N'_j(o)$. The components in vertical orientation of space are divided according to the

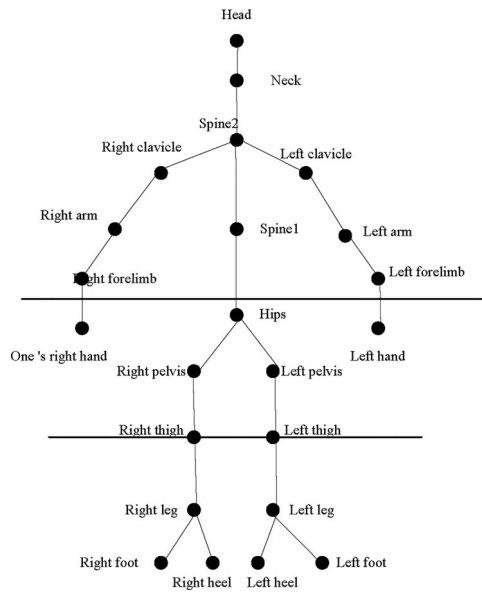


Figure 3. Vertical orientation Division.

height between the centre of gravity of the human body and ground. The division result is shown in Figure 3.

The spatial horizontal orientation is to divide the motion space of the human body into 8 orientations with an interval of 22.5 in each vertical component.

(3) Action force effect

The purpose of motion force analysis is to dig out the emotional connotation expressed by the action. This paper calculates the force feature of the action according to the spatiotemporal features, and realises the further reconstruction of the action feature.

2.3.3. Action feature reconstruction

In order to extract the azimuth, force effect and other features of the motion in the non-rigid 3D motion frame $N'_j(o)$, this paper converts the rotation angle in the motion capture into the trajectory position information of the joint nodes based on the Euler angle principle (Shabana, 2021). First, Euler Angle is transformed with a matrix, which is converted from the rotation angle data are L (roll angle), Q (pitch angle), U (yaw angle). The relative orientation of the node position is calculated by the rotation matrix through the intermediate node between the node and the root. Then the position information in the world coordinate system of the joint is obtained through the process of de-rotation and de-translation.

The movement of human body in space causes the deviation between the world coordinate of bone node in x direction and the initial acquisition position, making spatial orientation analysis cannot be carried out. Therefore, after obtaining the world coordinates, it is necessary to calculate the relative coordinates of the reference point with the coordinates (x_r, y_r, z_r) of the root $Root$. After the coordinates of each node are obtained, the feature

parameters describing the biomechanics of human motion, such as joint pair distance, bone pair angle, human orientation, spatial orientation, weight and smoothness in force effect, are calculated from the motion capture data. From the perspective of limb structure, the joint pair distance and bone pair angle is considered. The joint pair distance δ reflects the speed of movement, and the Euclidean distance is used to calculate the distance between joint pairs (Wahyudi et al., 2022), expressed in Equation (11). The included angle of bone pair ϕ represents the bending state between adjacent bones, expressed in Equation (12).

$$\delta = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2} \quad (11)$$

$$\phi = \arccos(\phi_1 \cdot \phi_2 / (||\phi_1|| \times ||\phi_2||)) \quad (12)$$

In the Equations, the coordinate of joint point i is (x_i, y_i, z_i) , ϕ_1 and ϕ_2 represent the distance between joint points 1 and 2 respectively. The eigenvalues of human body orientation, spatial orientation, and the weight and fluency of force effects are considered from the perspective of spatial orientation. The orientation of the human body e can be expressed as:

$$e = \vec{\varphi} \times \vec{\gamma} \quad (13)$$

where $\vec{\varphi}$ and $\vec{\gamma}$ are the plane normal vectors composed of different skeleton points, respectively.

In the non-rigid 3D motion frame $N'_j(o)$, spatial orientation g is divided into vertical and horizontal orientations. The vertical direction divides the human skeleton into the upper, middle and lower layers. The horizontal direction divides the horizontal components into front, left front, left, left rear, rear, right rear, right front, and home position, and each horizontal component is divided at 22.5° intervals.

The eigenvalue force V includes weight and fluency. The weight Ω represents the integration of the curve on y-axis, denoting the joint point, from the start to the end of the time (x) axis, which represents the rise and fall of the action. The reconstructed characteristic equation is:

$$\Omega = \int p dt \quad (14)$$

where p is the fluency; Fluency p expresses the expansion and contraction of space through distance.

2.4. Non-rigid 3D motion pose and action correction with hybrid CNN-HMM

Various non-rigid 3D motion pose features extracted in Section 2.3 are used as recognition samples in a depth learning hybrid CNN-HMM (Castillo et al., 2021). The model structure is shown in Figure 4.

This model embeds the output probability matrix of CNN into the observation probability matrix of HMM. The input non-rigid 3D motion pose samples are denoted as $\sigma = \sigma_1, \sigma_2, \dots, \sigma_M$, where σ_M represents the M extracted features. Then the maximum posterior probability $Q(\Gamma|\sigma)$ output by CNN is selected, and the best-fitting sequence $\bar{\Gamma}$ of σ

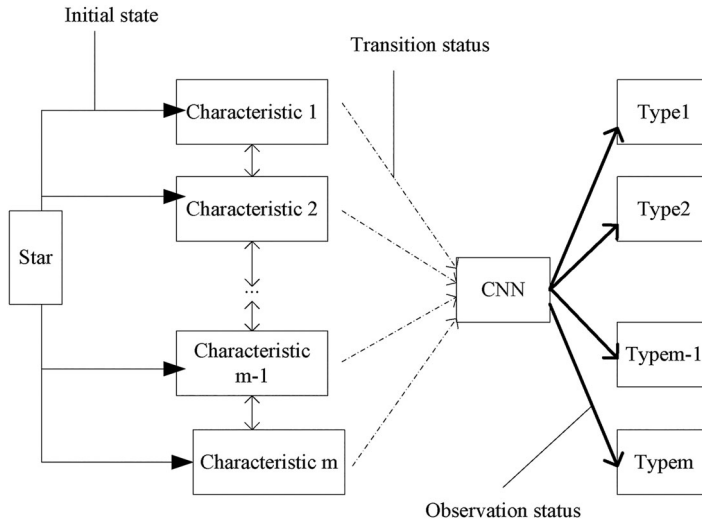


Figure 4. Hybrid CNN-HMM structure.

is obtained according to the Bayesian Decision (Williamson et al., 2022).

$$\bar{\Gamma} = \arg \max Q(\Gamma|\sigma) = \arg \max \frac{Q(\Gamma|\sigma)Q(\sigma)}{Q(\sigma)} \quad (15)$$

where $Q(\sigma)$ is the evidence factor of gesture recognition; Γ is the prior probability of gesture recognition. For a motion pose feature sample σ , Γ and $Q(\sigma)$ are irrelevant, so the posterior probability $Q(\Gamma|\sigma)$ is expressed as the product of the class prior probability and $Q(\Gamma|\sigma)$.

$$\bar{\Gamma} = \arg \max Q(\Gamma|\sigma) \cdot \Gamma \quad (16)$$

According to the input time changes, HMM (Khalili Sadaghiani & Forouzandeh, 2022 Stoltz et al., 2021;) is used to maximise $Q(\Gamma|\sigma)$ by using the first-order Markov hypothesis, which can be obtained as following Equation (17).

$$Q(\Gamma|\sigma) \approx \max \prod_{t=1}^T Q(\Gamma_t|\Gamma_{t-1})Q(\Gamma_t|\sigma) \quad (17)$$

where $Q(\Gamma_t|\Gamma_{t-1})$ is the transition probability between different non-rigid 3D motion poses in the HMM; $Q(\Gamma_t|\sigma)$ is the class conditional probability function. According to Bayes' Rule, $Q(\Gamma_t|\sigma)$ is converted to the likelihood value, which can be obtained as Equation (18).

$$Q(\Gamma_t|\sigma) = \frac{Q(\Gamma_t|\sigma)Q(\sigma_t)}{\Gamma_t} \quad (18)$$

After deleting the constant $Q(\sigma_t)$, the final pose recognition result of the hybrid CNN-HMM model can be obtained according to Equations (16)–(18).

$$\bar{\Gamma} = \arg \max \prod_{t=1}^T Q(\Gamma_t|\Gamma_{t-1}) \frac{Q(\Gamma_t|\sigma)}{\Gamma_t} \quad (19)$$

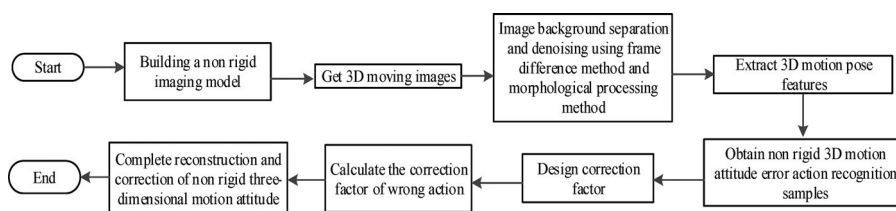


Figure 5. Flow of non rigid three-dimensional motion attitude reconstruction and high-precision correction method.

Action correction is mainly achieved through image feature comparison. Aiming at the process of motion reconstruction, the form of comparison template is used to achieve motion correction. Assuming that the plane of $N'_j(o)$ is classified after all the proposed processing steps, if the median of the k th plane is higher and exist certain features of the binary image, the pixels in the j and i rows are described as follows.

$$o_k(j, i) = \begin{cases} 0, & \text{otherwise} \\ 1, & \text{mod} \left[\frac{o_k(j, i)}{2_{k-1}} \right] = 1 \end{cases} \quad (20)$$

where $\text{mod}[]$ is a remainder function. All random pixel values $o(j, i)$ can be expressed with $o_k(j, i)$.

Comparing the action pixel features in $N'_j(o)$ with the standard pose action pixel features, the action correction coefficient can be obtained as Equation (21).

$$Ho = \frac{\partial^2 \pi}{\partial^2 x} - \frac{\partial^2 \pi}{\partial^2 y} - \left(\frac{\partial^2 \pi}{\partial x \partial y} \right) \quad (21)$$

where the correction scale is π ; x, y represent the action to be corrected in the x -axis, y -axis vector. Combined with the obtained correction coefficient, it is used to adjust the reconstruction effect of the real moving target, so as to realise the correction of the motion reconstruction result. Flow of the proposed method is shown in Figure 5.

3. Experimental analysis

The computer CPU used in this experiment is: Intel Core i7 4702MQ with main frequency 2.2GHz and memory capacity 8GB. MATLAB 2021b simulation software is used to build a 3D reconstruction effect map. In order to test the applicable effect of the proposed method, the experimental method is used to scan the ballet course scene in dance studio with a handheld depth camera, recognise and correct the non-rigid 3D motion posture for ballet students, and analyse the effectiveness of the proposed method. In the acquired non-rigid 3D moving images of ballet students, since students are learning in the dance studio, the background in the image will affect the extraction and recognition of the students' pose features, and there is still noise in the image. With the background separation and denoising method of non-rigid 3D moving image and morphology processing is used to separate and denoise the background, 3D reconstruction effect diagram is shown in Figure 6.

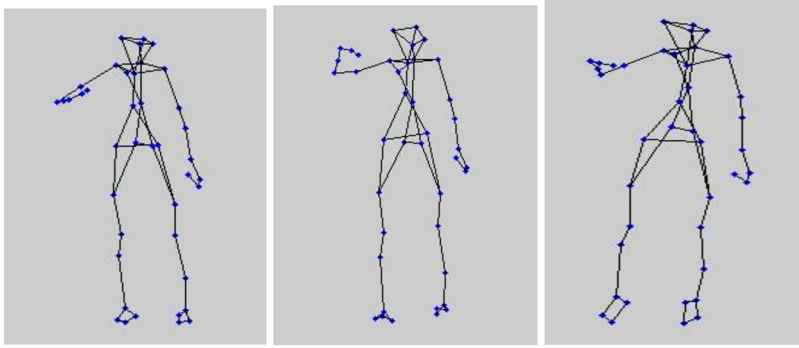


Figure 6. Non rigid 3D moving image background separation and de-noising rendering.

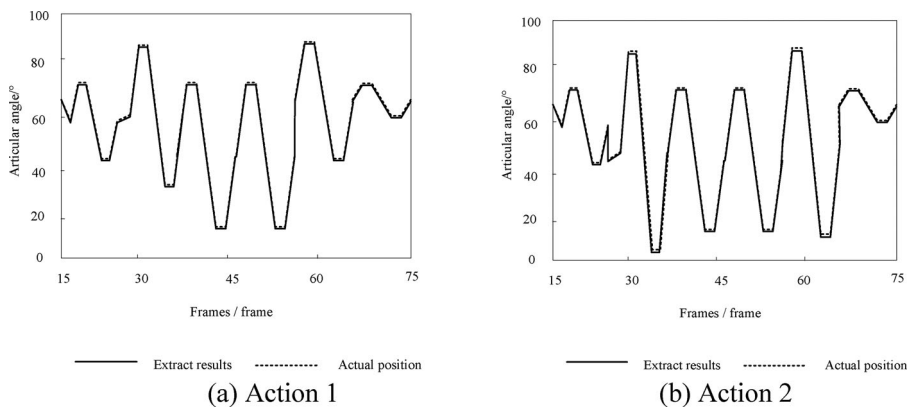


Figure 7. Recognition effect of the proposed method on 3D motion postures of dance students.

In Figure 6, the background separation and denoising method using frame difference and morphological processing effectively separates the non-rigid 3D moving image background well, without the background interference factors after separation. The image quality is significantly improved, and the posture of ballet students can be clearly observed from the figure. It can be concluded the effectiveness of the denoising method in this paper.

The extraction effect of the proposed method on the 3D pose features of the two dances shown in Figure 6 is tested. Taking joint angles as an example, the test results are shown in Figure 7.

From the experimental results in Figure 7, it can be seen that the post feature extraction results of dance actions 1 and 2 using the proposed method are consistent with the actual joint angle position. It can be concluded that this method can more accurately detect the movement posture features of dance students.

After the pose features are extracted, the proposed method is used to recognise whether there is any error in the motion poses of dance students. After identifying the two 3D motion poses shown in Figure 6, the wrong actions are directly marked in the image, and the result is shown in Figure 8.

As shown in Figure 8, the proposed method believes that there is a hand motion error in action 2 of dance students' movement posture. According to the ballet standard, it is

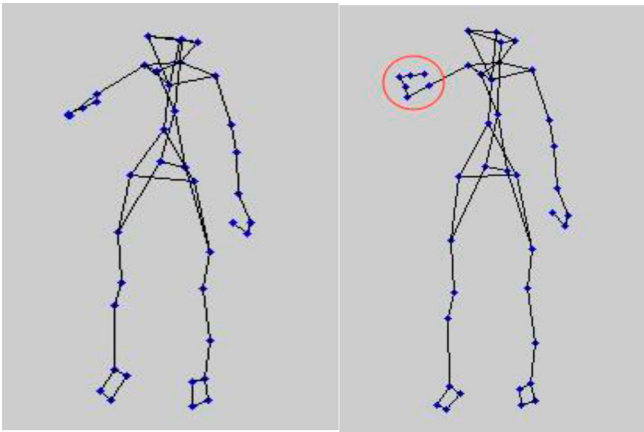


Figure 8. The recognition effect of this method on non-rigid 3D motion post of dance students.

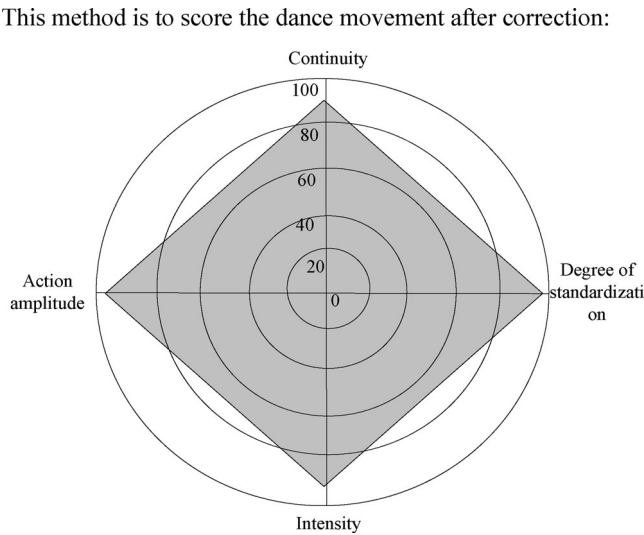


Figure 9. The performance of dance students after the use of this method.

shown that the shoulders should be relaxed and should be pulled back slightly, especially the shoulder blade behind the back. Therefore, combined with the visual effects, the proposed method can correctly identify dance students' movement gestures and actions 2. After identifying the wrong action and correcting them, the proposed method was used to test the performance of dancing students. The results are shown in Figure 9.

Figure 9 shows that after the wrong movements are identified by the proposed method, the dance students can complete the movement correction according to the correction coefficient. Among the scores of dance students, the lowest score is 90 points. The highest score is 98 points, which proves that the proposed method can achieve high-precision correction of non-rigid 3D movement posture taking dance students as examples, thereby improving the standardisation level of dance movements.

4. Conclusions

As an important research direction in AI, human action recognition technology is a hot topic, which has been successfully applied in many fields such as human–computer interaction, video surveillance, and sports training (Zhang et al., 2021). Taking ballet in physical training as an example, this paper studies the high-precision correction method of non-rigid 3D movement posture of ballet students, in order to provide assistance for the problem of dance movement specification. The method in this paper mainly solves the problems of background separation, denoising, pose feature extraction and recognition, and wrong action correction of non-rigid 3D moving images. The use effect is tested through experiments, and the experiments show that the proposed method can effectively separate the background of non-rigid 3D moving image, and the image quality is significantly improved. After identifying the wrong action, the method can correct the action according to the correction coefficient. To sum up, it is confirmed that the proposed method has useful value in the identification and correction of non-rigid 3D postures of dance students. In the follow-up work, the research will focus on different areas, and further improve the method in this paper to adapt to more areas (Zhang et al., 2020).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National Natural Science Foundation of China [grant number 62207012]; Natural Science Foundation of Hunan Province [grant numbers 2020JJ4434, 2020JJ5368]; National Social Science Foundation of China [grant number AEA200013]; Key Scientific Research Projects of Education Department of Hunan Province [grant number 22A0058]; Key Research Project on Degree and Graduate Education Reform of Hunan Province [grant number 2020JGZD025].

ORCID

Weina Fu  <http://orcid.org/0000-0002-4302-6505>

Shuai Liu  <http://orcid.org/0000-0001-9909-0664>

References

- Ahmad, N., Asif, H. M. S., Saleem, G., Younus, M. U., Anwar, S., & Anjum, M. R. (2021). Leaf image-based plant disease identification using color and texture features. *Wireless Personal Communications*, 121(2), 1139–1168. <https://doi.org/10.1007/s11277-021-09054-2>
- Al-Otum, H. M., & Ibrahim, M. (2021). Color image watermarking for content authentication and self-restoration applications based on a dual-domain approach. *Multimedia Tools and Applications*, 80(8), 11739–11764. <https://doi.org/10.1007/s11042-020-10368-9>
- Atkinson, D., & Becker, T. H. (2021). Stereo digital image correlation in MATLAB. *Applied Sciences*, 11(11), 4904. <https://doi.org/10.3390/app11114904>
- Castillo, M., Soto, R., Crawford, B., Castro, C., & Olivares, R. (2021). A knowledge-based hybrid approach on particle swarm optimization using hidden Markov models. *Mathematics*, 9(12), 1417. <https://doi.org/10.3390/math9121417>
- Dharejo, F. A., Zhou, Y., Deeba, F., Jatoi, M. A., Du, Y., & Wang, X. (2021). A remote-sensing image enhancement algorithm based on patch-wise dark channel prior and histogram equalisation with colour correction. *IET Image Processing*, 15(1), 47–56. <https://doi.org/10.1049/ipr2.12004>

- Dhiman, C., Vishwakarma, D. K., & Agarwal, P. (2021). Part-wise spatio-temporal attention driven CNN-based 3D human action recognition. *ACM Transactions on Multimedia Computing Communications and Applications*, 17(3), 1–24. <https://doi.org/10.1145/3441628>
- Hasenberger, B., & Alves, J. (2020). AVIATOR: Morphological object reconstruction in 3D - An application to dense cores. *Astronomy & Astrophysics*, 633, A132. <https://doi.org/10.1051/0004-6361/201936095>
- Ilesanmi, A. E., & Ilesanmi, T. O. (2021). Methods for image denoising using convolutional neural network: A review. *Complex & Intelligent Systems*, 7(5), 2179–2198. <https://doi.org/10.1007/s40747-021-00428-4>
- Jahren, S. E., Aakvaag, N., Strisland, F., & Vogl, A. (2021). Towards human motion tracking enhanced by semi-continuous ultrasonic time-of-flight measurements. *Sensors*, 21(7), 2259. <https://doi.org/10.3390/s21072259>
- Jensen, S. H. N., Doest, M. E. B., Aanæs, H., & Bue, A. D. (2021). A benchmark and evaluation of non-rigid structure from motion. *International Journal of Computer Vision*, 129(4), 882–899. <https://doi.org/10.1007/s11263-020-01406-y>
- Khalili Sadaghiani, A.V., & Forouzandeh, B. (2022). Image interpolation based on 2D-DWT and HDP-HMM. *Pattern Analysis and Applications*, 25(2), 361–377. <https://doi.org/10.1007/s10044-022-01057-4>
- Kim, H., Lee, C., Kwak, S., Hwang, Y., Kim, S., Choi, Y., & Kim, J. (2022). Three-dimensional volume reconstruction from multi-slice data using a shape transformation. *Computers & Mathematics with Applications*, 113, 52–58. <https://doi.org/10.1016/j.camwa.2022.03.018>
- Kulikajėvas, A., Maskeliūnas, R., Damaševičius, R., & Włodarczyk-Sielicka, M. (2021). Auto-refining reconstruction algorithm for recreation of limited angle humanoid depth data. *Sensors*, 21(11), 3702. <https://doi.org/10.3390/s21113702>
- Kumarasuvamy, A. S., & Rajendran, R. S. (2021). Classification of remote sensing image scenes using double feature extraction hybrid deep learning approach. *Journal of Information Technology and Digital World*, 3(2), 133–149. <https://doi.org/10.36548/jitdw.2021.2.006>
- Mahdaoui, A. E., Ouahabi, A., & Moulay, M. S. (2022). Image denoising using a compressive sensing approach based on regularization constraints. *Sensors*, 22(6), 2199. <https://doi.org/10.3390/s22062199>
- Mora-Martín, G., Turpin, A., Ruget, A., Halimi, A., Henderson, R. K., Leach, J., & Gyongy, I. (2021). High-speed object detection with a single-photon time-of-flight image sensor. *Optics Express*, 29(21), 33184–33196. <https://doi.org/10.1364/OE.435619>
- Shabana, A.A. (2021). Frenet oscillations and Frenet–Euler angles: Curvature singularity and motion-trajectory analysis. *Nonlinear Dynamics*, 106(1), 1–19. <https://doi.org/10.1007/s11071-021-06798-1>
- Stoltz, M., Stoltz, G., Obara, K., Wang, T., & Bryant, D. (2021). Acceleration of hidden Markov model fitting using graphical processing units, with application to low-frequency tremor classification. *Computers & Geosciences*, 156, 104902. <https://doi.org/10.1016/j.cageo.2021.104902>
- Sungheetha, A., & Sharma, R. (2021). 3D image processing using machine learning based input processing for man-machine interaction. *Journal of Innovative Image Processing (JIIP)*, 3(01), 1–6. <https://doi.org/10.36548/jiip.2021.1.001>
- Wahyudi, M. I., Wibowo, E. W., & Sopiullah, S. (2022). Web-based face recognition using line edge detection and Euclidean distance method. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 135–142. <https://doi.org/10.29408/edumatic.v6i1.5525>
- Wang, S. H., Govindaraj, V. V., Górriz, J. M., Zhang, X., & Zhang, Y. D. (2021). COVID-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network. *Information Fusion*, 67, 208–229. <https://doi.org/10.1016/j.inffus.2020.10.004>
- Williamson, S. F., Jacko, P., & Jaki, T. (2022). Generalisations of a Bayesian decision-theoretic randomisation procedure and the impact of delayed responses. *Computational Statistics & Data Analysis*, 174, 107407. <https://doi.org/10.1016/j.csda.2021.107407>
- Yu, L., Fu, X., Xu, H., Xu, H., & Fei, S. (2020). High-precision camera pose estimation and optimization in a large-scene 3D reconstruction system. *Measurement Science and Technology*, 31(8), 085401. <https://doi.org/10.1088/1361-6501/ab816c>

- Zhang, Y. D., Dong, Z., Wang, S., Yu, X., Yao, X., Zhou, Q., Hu, H., Li, M., Jiménez-Mesa, C., Ramirez, J., Martinez, F. J., & Gorriz, J. M. (2020). Advances in multimodal data fusion in neuroimaging: Overview, challenges, and novel orientation. *Information Fusion*, 64, 149–187. <https://doi.org/10.1016/j.inffus.2020.07.006>
- Zhang, Y. D., Satapathy, S. C., Wu, D., Guttery, D. S., Górriz, J. M., & Wang, S. H. (2021). Improving ductal carcinoma in situ classification by convolutional neural network with exponential linear unit and rank-based weighted pooling. *Complex & Intelligent Systems*, 7(3), 1295–1310. <https://doi.org/10.1007/s40747-020-00218-4>