IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 January 2024, accepted 29 January 2024, date of publication 1 February 2024, date of current version 12 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3361281

RESEARCH ARTICLE

Reconstruction of 3D Non-Rigid Moving Human Targets Based on Reliable Estimation of Contour Deformation Degree

YAN ZHANG[1], MOHAMED BAZA[2], AND HANI ALSHAHRANI[3], (Member, IEEE)

[1]Sports Department, Zhongnan University of Economics and Law, Wuhan 430000, China
[2]Department of Computer Science, College of Charleston, Charleston, SC 29424, USA
[3]Department of Computer Science, College of Computer Science and Information Systems, Najran University, Najran 61441, Saudi Arabia

Corresponding author: Mohamed Baza (bazam@cofc.edu)

This work was supported by the Deanship of Scientific Research at Najran University for funding this work under the Research Groups Funding Program Grant Code NU/RG/SERC/12/27.

ABSTRACT The human body is a typical non-rigid object, and its 3D reconstruction is a classic problem in the field of computer vision. Due to the inherent complexity and dynamism of the human body, it is not suitable for existing non-rigid 3D motion reconstruction algorithms that assume that the number of shape bases of non-rigid bodies is known. The number of shape bases is very important for 3D reconstruction methods. If the number of shape bases is estimated incorrectly in contour deformation estimation, it can lead to unreliable or even complete failure of the reconstruction algorithm. Therefore, this paper designs a 3D non-rigid motion human object reconstruction algorithm based on reliable estimation of contour deformation degree. This algorithm leverage Scale Invariant Feature Transform (SIFT) to obtain non-rigid moving human target features. Firstly, the contour appearance model of the moving human sequence is used to extract the contour feature sequence, which is preprocessed based on the contour appearance depth feature; Furthermore, the deformation degree of the contour is reconstructed and the calculation process of the number of shape bases was optimized, which is no longer simply defined as known. This method optimizes and solves the problem of missing data, improves the reliability of estimating the degree of contour deformation, and completes target reconstruction. The experimental results show that the three-dimensional reconstruction algorithm can accurately reconstruct the changes in the movements of athletes' shots; The accuracy of 3D reconstruction can reach 95.98%; Moreover, PSNR, SSIM, and MSE indexes performed well with smaller fluctuation range, and the distribution of three-dimensional reconstructed scattering points is very close to the three-dimensional position distribution of real scattering points, and the three-dimensional reconstruction effect is good with strong reliability.

INDEX TERMS 3D reconstruction, reliability, non-rigid motion, exercise the human body, bilateral filtering, outline appearance.

I. INTRODUCTION

The three-dimensional image reconstruction of non-rigid objects is a classic problem in computer vision [1], and it is also a fundamental problem in constructing artificial intelligence vision systems. The research results of target reconstruction algorithms for moving human bodies can endow them with the ability to capture dynamic objects, thereby achieving true environmental interaction. In addition, the research on 3D non rigid motion human target reconstruction has important value in many application scenarios. Through three-dimensional non rigid motion human target reconstruction, precise tracking and recognition of human targets in monitoring scenes can be achieved, improving the intelligent performance of video monitoring systems. It can quantitatively analyze the process of human movement,

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaojun Steven Li.

help athletes improve their skills, prevent sports injuries, and be used in daily rehabilitation, posture assessment, etc. in the medical field. At the same time, it can achieve real-time recognition and tracking of human posture and movements, for applications such as human-computer interaction, gesture control, and virtual character control. It can achieve motion capture and restoration of actors, and be used for character animation production and special effects design in movies, TV dramas, games, etc. And with the continuous development of research and technology, there may be more new application scenarios in the future. Due to the high complexity of human motion and the strict quality requirements for 3D reconstruction in related industries, the reliable reconstruction of 3D human bodies has always been a challenging research hotspot in the fields of computer graphics and computer vision. Therefore, the research on 3D reconstruction algorithms for non-rigid moving targets has important theoretical significance and application value in the field of pattern recognition.

Numerous scholars have conducted extensive research on issues in related fields. Li, C, and others conducted research on 3D reconstruction algorithms based on the ant hill model [2]. However, this method uses a dynamic radius sphere to roll along the fiber axis to construct a closed 3D surface, which suffers from unstable path tracking and results in inaccurate 3D reconstruction results. Liu, Z et al. studied a multi degree of freedom human motion pose multi-objective image reconstruction algorithm [3]. However, when this method uses the roulette wheel selection method to achieve chromosome hybridization and mutation, unreasonable hybridization and mutation strategies can lead to strong local search ability, which limits the improvement of image reconstruction ability. Coenen et al. proposing a stereo image vehicle pose estimation and 3D reconstruction based on subcategory perception shape priors [4]. However, this method is affected by factors such as lighting and complex background segmentation, and the reconstruction reliability is not strong. Kim et al. proposed a two-dimensional TEM image reconstruction method to find the optimal affine transformation between images [5]. Due to the involvement of multiple continuous images in registration, this method also adopts hierarchical registration, but the operation of this method is relatively complex, which affects the reliability and applicability of the algorithm. Mejri proposed a 3D motion recovery method for non-rigid objects based on noise measurement [6]. This method assumes that the motion of the object is not fully known and considers it as an unknown input to the perspective dynamic system, which leads to blurring in the reconstruction effect of this method. The 3D reconstruction effect is not ideal and the reliability is reduced. Han et al. proposed a method for reconstructing the three-dimensional structure of blood cells based on two non-orthogonal phase images [7]. They developed a four-step processing process that includes registration, detection, segmentation, and reconstruction to achieve the 3D reconstruction. However, the matching process of this method has too many constraints, and the reliability of image feature reconstruction is low. Knudsen et al. proposed a direct regularization reconstruction method for three-dimensional features [8], which directly performs reconstruction through inverse imaging. This method is robust and reliable against small perturbations in data, but its depth and accuracy decrease with increasing distance, and its reliability also decreases.

Based on the shortcomings of the above methods, this paper designs an object reconstruction algorithm for 3D non-rigid moving human based on reliable estimation of contour deformation degree. The specific research path is as follows:

(1) Extract features of non rigid moving human targets using scale invariant feature transformation methods.

(2) Construct a contour appearance model of a moving human body sequence, extract a two-dimensional sequence of contour features, and use it to recognize and reconstruct human motion targets.

(3) Convert the 2D contour appearance depth data information into 3D point cloud data, extract 3D features, and then use bilateral filtering methods to filter them.

(4) Simultaneously considering noise and data loss factors to improve the shape reconstruction effect of motion image sequences.

(5) By solving the problem of missing data and optimizing the reliability of estimating the degree of contour deformation, the three-dimensional reconstruction of the target has been completed.

II. METHOD DESIGN FOR RECONSTRUCTING 3D NON-RIGID MOVING HUMAN OBJECTS

A. ACQUISITION OF HUMAN MOTION SEQUENCE DATA

1) OBTAINING FEATURE EXTRACTION MAP OF MOVING HUMAN OBJECTS BASED ON SIFT

Using SIFT to recognize moving human target image databases and extract features from image sequences. The SIFT algorithm can extract key feature points and local feature descriptors from images, which can reflect the shape and structural information of human moving targets in different poses. By extracting these features, the appearance of the motion posture contour of human motion targets can be described and represented, and the shape, posture, rotation and other features of the targets can be captured. The moving image sequence with this optical feature can show the characteristics of constant translation, constant scale, and constant selection [9], which enables better results to be achieved when matching moving human targets from various perspectives. This feature also can identify the target object, which is determined by selecting the extreme points of the Gaussian difference function in the scale space. The process of extracting the feature is as follows:

At the initial stage of the reconstruction algorithm, the Gaussian scale space of the input image is established, and the original image is obtained through the Gaussian function

convolution to obtain the Gaussian difference function:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, k\sigma) \tag{1}$$

where, k represents the scaling factor, $L(x, y, k\sigma)$ represents a human motion image, where $I(x, y)$ represents the Gaussian function $G(x, y, k\sigma)$ result of convolution is:

$$L(x, y, k\sigma) = G(x, y, k\sigma) * I(x, y) \tag{2}$$

Among them, "$*$" represents the convolution operator.

$$G(x, y, k\sigma) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma^2}\right\} \tag{3}$$

With the gradient and direction of the human motion image L at different scale levels after convolutional difference calculation, image features of different important parts are extracted [10], [11]. The gradient size M_{ij} and direction R_{ij} of different pixels L_{ij} obtained by image differentiation are:

$$M_{ij} = \sqrt{(L_{ij} - L_{i+1,j})^2 + (L_{ij} - L_{j+1,i})^2} \tag{4}$$

$$R_{ij} = \tan^{-1} \frac{L_{ij} - L_{i+1,j}}{L_{i+1,j} - L_{ij}} \tag{5}$$

2) DESIGN OF CONTOUR APPEARANCE MODEL FOR CONSTRUCTING MOVING HUMAN TARGET SEQUENCES

Based on the features of human motion targets obtained in the previous section, extract contour feature sequences using the contour appearance model of the moving human sequence [12], [13]. The linear time series model mainly focuses on the changes in data over time, ignoring the spatial information in the image. For moving images, the pixel positions in each frame of the image will change over time, and linear time series models cannot meet the changes in spatial relationships. Therefore, non-linear time series methods are used for modeling and analysis. The nonlinear time series of the motion image is as follows:

$$x_{i+1} = h(x_i) + u \tag{6}$$

where, $x_{i+1} = (x_1, x_2, \cdots, x_T)$ represents the observed sequence of human motion images; h represents a nonlinear function; u represents the noise model. Modeling the current dynamic model through kernel functions by using $x_{i+1}^{\varphi} = (\varphi(x_1), \varphi(x_2), \cdots, \varphi(x_T))$ on human motion sequence x_{i+1} for investigation, the mapping of features from R^d to H is represented based on Eq.7.

$$x_{i+1}^{\varphi} = T^{\varphi} X_1^{\varphi} + \mu^{\varphi} + v^{\varphi} \tag{7}$$

Then the kernel function of the relevant dynamic kernel model is obtained, where T^{φ} represents the state transition matrix, μ^{φ} represents a translation vector, $v^{\varphi} \in H$ represents the Gaussian white noise.

The maximum probability direction for detecting human motion contour targets is at the extreme value of the correlation probability density function $f(P, x, y, A, B)$. Therefore, the position between the hidden variable and the observed variable is set as the probability density function $f(P, x, y, A, B)$. The direction of the human motion target contour is described by x [14], the direction and speed of detecting human motion are described by y, the appearance mold of the human motion target is represented by A_i, and the detected appearance mold is represented by B_i, This method is similar to graph model learning. The use of probability graph models in human motion target detection is to establish probability relationships between different variables and provide a formal representation method for modeling and inferring human motion targets. The construction of the probability graph model is shown in Figure 1.

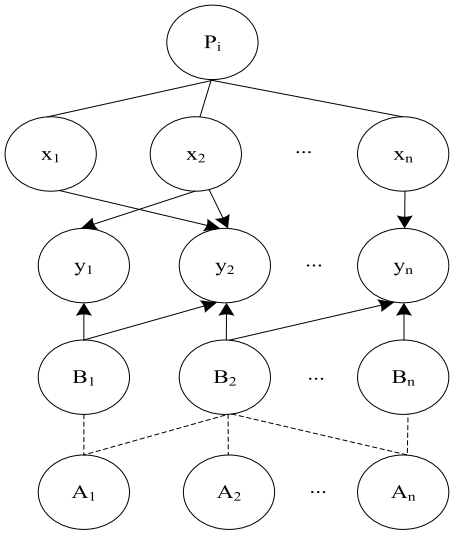


FIGURE 1. Probability model of moving human target sequence network.

The overall variable P and the direction and speed of human body movement ($x_i, i \neq k$) determine the direction and speed of the k -th action x_k. The direction and speed of human body movement x_k affect the detected direction y_j, and the appearance mold A_i of the human body movement determines B_j. The joint probability formula obtained from the probability model is shown in Eq.8.

$$f(P, x, y, A, B) = f(y|x, B) f(B|A) f(A) f(P) f(x|P) \tag{8}$$

where, $f(P)$ represents the observation probability of different human body movements. The human body during the movement process does not have prior knowledge, so the probability of the existence of different human body movements should be set to be non-interference with each other, as follows:

$$f(P) = \prod_{i=1}^{M} f(P_i) = \prod_{i=1}^{M} P_i^{\sigma_i} (1 - P_j)^{1-\sigma_i} \tag{9}$$

Among them, P_i represents the SIFT feature probability of the i^{th} motion action in the human motion image frame, and the value of P_i is obtained through learning data $f(x|P) = \prod_{i=1}^{M} f(x_i|P_i)$; σ_i represents the mean square deviation. Due to the limitations of mutually independent settings, setting motion labels is uniformly assigned $f(x_i|P_i = 1) = 1/N$,

where N is the preset threshold. If no motion action I is detected, the value of x_i is ignored, and $f(x_i|P_i = 0) = 1/N$ is selected, so we have Eq.10.

$$f(x|P) = (1/N)^M \tag{10}$$

The Gaussian distribution is the detection of its orientation and velocity modeling to obtain the appearance model of human motion contour, with the formula:

$$f(B|A) = g(b_{SD} = a_D)(1/V)^{N-|f|} \tag{11}$$

Among them, 1/V is the foreground detection threshold; $g(b_{SD} = a_D)$ is the Gaussian distribution weight; D is the update rate; SD is the pixel; b is the Gaussian distribution priority; a is a user-defined parameter that completes the acquisition of appearance sequences.

B. DEEP FEATURE EXTRACTION AND PREPROCESSING OF MOVING HUMAN TARGET SEQUENCES

Based on the two-dimensional contour sequence of the moving human target obtained in section II-A2, convert it into three-dimensional features, and use bilateral filtering method to filter the pixel interference of the collected contour.

In order to separate the background from the human body, the two-dimensional contour appearance depth data information is first converted into three-dimensional point cloud data [15]. Set the coordinates of a point in the contour depth map of a moving human body as (u, v), and its depth value is d(u, v). After passing through the internal parameters K of the depth sensor mapped to the three-dimensional spatial coordinate V, the formula is expressed as:

$$V(x, y, z) = d(u, v)K^{-1}[u, v, 1]^T \tag{12}$$

where, $K = \begin{bmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ represents the camera parameter matrix, (u_0, v_0) represents the center coordinate of the image, f is the camera focal length, d_x and d_y are the unit pixel size of the camera sensor on the u and v axes, respectively. After converting all the depth data, the point cloud data of the depth map can be obtained. As the point cloud data contains a large amount of background data, which is not what this article wants, it is necessary to filter the point cloud to remove the background and extract the human body part.

The characteristic of bilateral filtering is that it can preserve the edge features of moving human target images, and the filtering process is as follows:

Given the pixel (x, y, z) and its pixel value f(x, y, z), the output pixel value g(x, y, z), depends on the weighted sum of neighboring pixel values. If the intensity value of a pixel is similar to the intensity value of surrounding pixels, the corresponding weight is larger; On the contrary, if the intensity value of a pixel differs significantly from that of surrounding pixels, the corresponding weight is smaller. The formula can be expressed as:

$$g(x, y, z) = \frac{\sum_{k,l} f(x, y, z)\, w(u, v, l, k)}{\sum_{k,l} w(u, v, l, k)} \tag{13}$$

where, (l, k) represents the neighboring pixels of (u, v), w(u, v, l, k) represents the filtering kernel function, and w(u, v, l, k) depends on the domain kernel functions x(u, v, l, k) and y(u, v, l, k), which can be obtained as follows:

$$w(u, v, l, k) = w_x(u, v, l, k) * w_y(u, v, l, k) \tag{14}$$

$$w_x(u, v, l, k) = \exp\left(-\frac{(u-k)^2 + (v-l)^2}{2\sigma_x}\right) \tag{15}$$

$$w_y(u, v, l, k) = \exp\left(-\frac{(f(u, v) - f(k, l))^2}{2\sigma_y}\right) \tag{16}$$

Among them, σ_x and σ_y are the Gaussian kernel standard deviation, $w_x(u, v, l, k)$ represents the Gaussian filtering kernel function, $w_y(u, v, l, k)$ represents the range kernel function.

Due to the filtered depth image features obtained contain depth data of the background, and for the reconstruction of non-rigid moving human targets, without depth data about the human body, it is naturally useless. Therefore, it is necessary to separate the human body from the depth map containing the background. This article uses threshold segmentation method for background removal processing.

The threshold segmentation method, as the name suggests, is based on a set threshold for segmentation. If it is greater than the threshold, it is retained, and if it is not, it is removed. The threshold values set for the three coordinate axes x, y, z of point cloud data form a range box. If the point is within the box, it is retained, and if not, it is removed [16]. For the collected data, since the extracted contour appearance pixels have no other objects, the threshold segmentation method is very suitable for processing the contour appearance data of the moving human body. The process is as follows:

If a three-dimensional point is p, the threshold is ∂, and T(p) represents point cloud data, then it can be determined whether to retain the point.

$$T(p) = \begin{cases} p, & f(p) \le \partial \\ 0, & f(p) > \partial \end{cases} p \in R(p) \tag{17}$$

where, f(p) represents the threshold judgment function for the point.

C. RELIABILITY ESTIMATION OF NON-RIGID BODY DEFORMATION DEGREE

The feature points in motion image sequences exhibit motion and deformation, especially the shape of non rigid moving human targets may undergo drastic changes. Based on the filtered features, while considering noise and data loss factors, to improve the shape reconstruction effect of motion image sequences.In the process of non-rigid motion reconstruction, the coordinates of the moving image are transformed so that the origin of the image coordinate system is located at the centroid of the object, eliminating the translation vector, and the measurement matrix $\bar{W} = M_{2F\times 3K} B_{3K\times P}$. Obviously, the rank of the measurement matrix $\bar{W}$ is not greater than

3K (assuming $2m > 3K, n > 3K$). Therefore, the K is of great significance for shape reconstruction of motion image sequences and is an important parameter reflecting the degree of non-rigid body shape deformation [17]. The reconstruction of the degree of deformation of non-rigid bodies is an estimation of the K. The presence of noise causes random changes in the position of measured feature points, which in turn affects the perception of K. In addition, when feature point data is lost, it is necessary to restore the lost feature points in order to further estimate the K. Otherwise, the K cannot be estimated, which will affect the reliability of the algorithm. To address these issues, a K estimation method is proposed that considers both noise and data loss factors, and the reconstruction process is completed as follows:

The purpose of reconstructing the noise covariance matrix is to better handle and estimate shape deformation. The presence of noise can cause random changes in the position of measurement feature points, which can affect the accurate estimation of shape deformation. By reconstructing the noise covariance matrix, it is possible to more accurately understand and predict the impact of noise on the position of measurement feature points, thereby estimating shape deformation more accurately. The coordinates representing the shape of an object can be regarded as a random process [18], [19]. Representing the coordinates of all feature points in the i^{th} frame of the image as a column vector $\hat{W}_i = [\bar{u}_{i1}, \ldots, \bar{u}_{iP}, \bar{v}_{i1}, \ldots, \bar{v}_{iP}]^T$. So, there we have:

$$\hat{W}_i^T = \left[\omega_{i1}\overline{R}_i^1, \ldots, \omega_{iK}\overline{R}_i^1, \omega_{i1}\overline{R}_i^2, \ldots, \omega_{iK}\overline{R}_i^2\right] \begin{bmatrix} S_1 & & \\ \vdots & 0 & \\ S_K & & \\ & S_1 & \\ 0 & \vdots & \\ & S_K & \end{bmatrix} + \xi^T \tag{18}$$

Further simplify equation (18) for ease of calculation and optimization. In complex motion reconstruction processes, simplifying matrices can reduce computational complexity and improve algorithm efficiency. expression:

$$\hat{W}_i = [M'_{1\times 6K} B'_{6K\times 2P}]^T + \xi = B'^T M'^T + \xi \tag{19}$$

where, $\overline{R}_i^1$ and $\overline{R}_i^2$ are the first and second lines of $\overline{R}_i$ respectively, ξ is the noise of characteristic points, which can be seen as a random process with zero mean value. It is obvious that the shape basis B' of the image sequence is unchanged, which is also the root cause of the reduced reliability. Therefore, the correlation coefficient matrix of $\hat{W}$ can be reconstructed as follows:

$$R_{\hat{W}} = E\left[\hat{W}\hat{W}^T\right] = B'^T E\left[M^T M\right] B + C_\xi \tag{20}$$

where C_ξ is the noise covariance matrix.

Then, How to reconstruct this matrix is discussed. In principle, the noise covariance of the feature points is a function of the tracking algorithm and its parameters and the brightness change characteristics near the tracking feature points. The tracking point position error introduced by the tracking algorithm is non-uniform and correlated, which depends on the structural characteristics of the local image. For example, u_{ij} and v_{ij} tracking of corner j has high reliability [20], [21]; For point j on a line, the tracking reliability in its gradient direction is very high (normal flow), and the tracking reliability in its tangent direction is very low, which is directional error. For x_{ij} of a given I_i array image, we have Eq.21.

$$x_{ij} = \begin{bmatrix} u_{ij} \\ v_{ij} \end{bmatrix}, i = 1, \cdots, F, j = 1, \cdots, P \tag{21}$$

The noise covariance of the j^{th} feature point in the i^{th} frame image can be estimated by the inverse of the Hessian matrix, that is

$$Q_{ij} = E\left[\Delta x_{ij} \Delta x_{ij}^T\right] = \begin{bmatrix} \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial x^2} & \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial x \partial y} \\ \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial y \partial x} & \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial y^2} \end{bmatrix}^{-1} = \begin{bmatrix} \sigma_{ij_1}^2 & \sigma'_{ij} \\ \sigma'_{ij} & \sigma_{ij_2}^2 \end{bmatrix} \tag{22}$$

The elements of the Hessian matrix are the second derivative and partial differential of the image brightness in the x and y directions [22]. Formula (22) multiplied by a scale factor is approximately equal to the actual noise covariance, and the scale factor will not have any impact on the application. Therefore, the noise covariance matrix can be obtained by the following formula [23]:

$$C_\xi = \frac{1}{F}\sum_{i=1}^{F} \begin{bmatrix} \sigma_{1i_1}^2 & & 0 & & \sigma'_{1i} & & 0 & \\ & \ddots & & & & \ddots & & \\ 0 & & \sigma_{Pi_1}^2 & & 0 & & \sigma'_{Pi} & 1 \\ \sigma'_{1i} & & 0 & & \sigma_{1i_2}^2 & & 0 & \\ & \ddots & & & & \ddots & & \\ 0 & & \sigma'_{Pi} & & 0 & & \sigma_{Pi_2}^2 & \end{bmatrix}_{2P\times 2P} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} \tag{23}$$

The inverse C_ξ^{-1} reconstruction of the noise covariance matrix is as follows:

$$C_\xi^{-1} = \begin{bmatrix} (c_1 - c_2 c_4^{-1} c_3)^{-1} & -c_1^{-1} c_2 (c_4 - c_3 c_1^{-1} c_2)^{-1} \\ -c_4^{-1} c_3 (c_1 - c_2 c_4^{-1} c_3)^{-1} & (c_4 - c_3 c_1^{-1} c_2)^{-1} \end{bmatrix} \tag{24}$$

Combining (22) and (23), it can be concluded that:

$$R_{\hat{W}} C_\xi^{-1} = B'^T E\left[M^T M\right] B C_\xi^{-1} + I_{2P\times 2P} \tag{25}$$

where, $R_{\hat{W}} = \frac{1}{F}\sum_{i=1}^{F} \hat{W}_i \hat{W}_i^T$.

Generally, for 3D motion, the maximum rank of $B'^T E[M^T M]B$ is 6K. Then, $\mu_i(H)$ is used to represent the i^{th}

eigenvalue of matrix H, then

$$\begin{cases} \mu_i(H) = \mu_i\left(B'^{T}E\left[M^{T}M\right]BC_{\xi}^{-1}\right)+1, i=1,\cdots,6K \\ \mu_i(H) = 1, i=6K+1,\cdots,2P \end{cases} \tag{26}$$

Therefore, matrix H has 6K eigenvalues greater than 1. If the number of eigenvalues of the matrix greater than 1 is counted, the estimated value of K can be obtained by dividing by 6. K is the dimension of the shape space representing the deformation feature point sequence and the number of shape bases indicating the simulation feature point sequence. Therefore, K can be used to reconstruct the deformation degree of the shape sequence. Therefore, the reconstruction of the deformation degree of the 3D scene after translation and rotation can be defined as:

$$\text{Deformation Degree} = \frac{\text{Count(eigenvalues of H} > 1)}{6} \tag{27}$$

D. IMPROVEMENT OF DATA LOSS

Based on the image sequence considering noise and data loss factors mentioned above, a filter is used to restore its feature points to improve the phenomenon of data loss.For the actual image sequence, some feature points will be occluded, resulting in data loss, which affects the reliability of the algorithm. Due to inertia and high acquisition speed, the motion of feature points between frames of sequence images is usually smooth, so this paper uses $\alpha-\beta-\gamma$ filter to track the movement of feature points [24]. The $\alpha-\beta-\gamma$ filter is a constant coefficient Kalman filter, which be used to simulate and predict the movement of the target when the velocity and acceleration are constant within the sampling interval. The mathematical expression of this filter is as follows:

$$\begin{cases} x_p(t) = x_s(t-1) + Tv_s(t-1) + \frac{T^2}{2}a_s(t-1) \\ v_p(t) = v_s(t-1) + Ta_s(t-1) \\ a_p(t) = a_s(t-1) \\ x_s(t) = x_p(t) + \alpha(x_m(t) - x_p(t)) \\ v_s(t) = v_p(t) + \frac{\beta}{T}(x_m(t) - x_p(t)) \\ a_s(t) = a_p(t) + \frac{\gamma}{2T^2}(x_m(t) - x_p(t)) \end{cases} \tag{28}$$

where, $x_p(t)$ is the location of the target at the predicted time t, $v_p(t)$ is the speed at the predicted time t, $a_p(t)$ is the acceleration at the predicted time t, $x_s(t)$ is the location at the smoothed time t, $v_s(t)$ is the speed at the smoothed time t, $a_s(t)$ is the acceleration at the smoothed time t, $x_m(t)$ is the location at the measured time t. T is the measurement time interval, α is the position filter parameter, β is the speed filter parameter, γ is the acceleration filter parameter. filter α, β, γ parameter optimization selection:

$$\begin{cases} \beta = 2(2-\alpha) - 4\sqrt{1-\alpha} \\ \gamma = \frac{\beta^2}{\alpha} \end{cases} \tag{29}$$

The $\alpha-\beta-\gamma$ filter is used to predict the next "x" and "y" coordinates of the specified feature point. The tracking process is the continuously executing the cycle of "predict match update". At time t, the position of each tracking point at time $t+1$ is predicted. Next, check whether there are matching feature points near the predicted location. If the matching is successful, update the parameters of the filter [25]. After the motion of feature points is tracked by the $\alpha-\beta-\gamma$ filter, the motion trajectory of feature points can be obtained. When a feature point data x_{ij} is lost, use the position prediction value to recover its position information and define the covariance of the recovered feature point as:

$$Q_{ij} = \frac{Q_{ij-1}Q_{ij+1}}{2} \tag{30}$$

III. EXPERIMENTAL ANALYSIS

In order to verify the effectiveness of the algorithm proposed in this paper, an experimental system was established to test the aforementioned 3D reconstruction algorithm. The experiment selected the MPII Human Pose public dataset to obtain human motion image sequences, with a resolution of 768 × 576 and a frame rate of 25 frames per second, to test the above algorithm. There are a total of 20000 pieces of data. There are three training sets in total, among which Training Set 1 contains a large number of shooting actions, involving coordinated movements of arms and wrists. In the training set, you can find shooting action image sequences from different angles and postures, which are used to train the model to recognize the features of shooting postures. Training set 2 includes a large number of passing movements, involving hand and arm movements. The training set contains image sequences of different passing methods, such as chest passing, overhead passing, etc., which are used to identify the passing posture of the training model. Training set 3 includes a lot of dribbling exercises, involving control of hands and arms. The training set contains image sequences of different dribbling methods, such as two handed dribbling, one handed dribbling, etc., used to train the model to recognize dribbling postures. And divide it in an 8:2 ratio to obtain the training set, validation set, and test set. Based on this data environment, a simplified camera model was used, assuming that the center of gravity of all points is at the center of the image, and that the internal parameters of the camera remain unchanged during the projection process. The partial reconstruction results in the dataset are shown in Figure 2. This image displays the original images of three images, as well as the reconstruction results of the corresponding front view, side view, and top view. From Figure 2, it can be seen that this method can accurately reconstruct the three-dimensional structure of non rigid bodies in dynamic scenes and accurately reflect the changes in human motion. The reconstruction results can be used for image recognition and other purposes.

As can be seen from Figure 2, the operation process of 3D reconstruction of athletes' related actions through the algorithm in this paper is relatively simple, and the

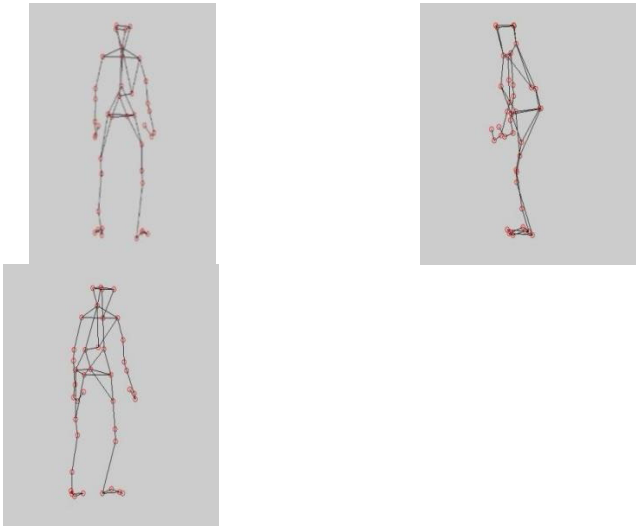
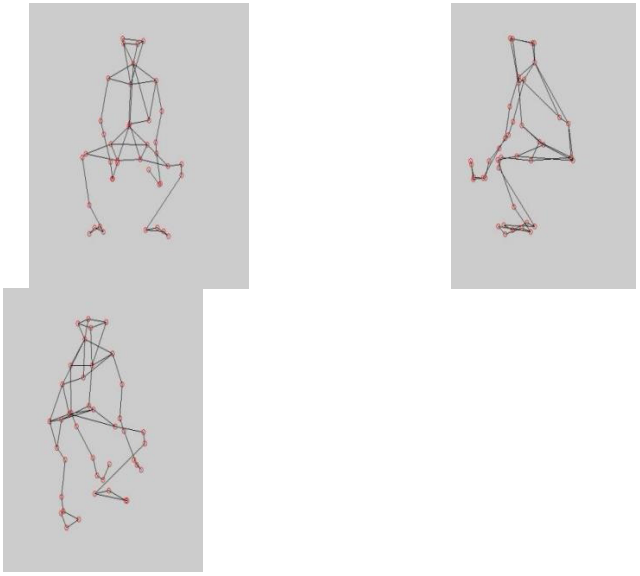


FIGURE 2. 3D reconstruction of basketball players' shooting actions.



a b c

FIGURE 3. Analysis of pre-treatment effect.

3D reconstruction effect is ideal. It can accurately track the changes of athletes' actions without blurring and other phenomena. The reconstruction results are clear and robust. Randomly select a squat image for basketball training, apply the algorithm in this paper to filter it and preprocess the segmentation of foreground moving human objects, and the 3D reconstruction results are shown in Figure 3.

From Figure 3, it can be seen that the image obtained after basketball training squat filtering is based on the actions shown in Figure 2. Through comparison, it can be found that the algorithm in this article has significantly improved the reliability of the image detail information processed. There is no shape deformation or missing shape information in the image, indicating that the filtering process effectively helps to restore the shape of the image and fill in the missing information. And highlight the detailed information in the image.

A. RECOGNITION ERROR (%) OF 3D RECONSTRUCTION RESULTS

In order to verify the optimization performance of the algorithm in this paper for the three-dimensional reconstruction of non-rigid moving human objects, we collected about 40 basketball body sequences of 2015 freshmen from a sports college, with 4 basketball positions sampled for each subset, and 10 positions for various positions. We used the algorithms in [4], [5], [6], [7], and [8] as the comparison algorithms of this algorithm, and used six algorithms to reconstruct the human body targets of basketball players in four movements, namely shooting, running, jumping, and passing. Combined with the convolutional neural network action recognition model, the reconstruction results of different algorithms are identified, and the action recognition errors of the reconstruction results of different algorithms are counted. The expression is:

$$R = \sqrt{\frac{1}{N}\sum \|P - P'\|^2} \tag{31}$$

where, N is the number of samples, P is the position vector of the true results, P′ is the position vector of the predicted outcome.

The test results are shown in Table 1.

TABLE 1. Recognition error (%) of 3D reconstruction result of athletes'human body targets using diffenent methods.

Method	Shoot at the basket	Running	Jumping	Pass
Ref.4	28.14	27.46	28.96	28.36
Ref.5	27.62	26.37	27.37	27.05
Ref.6	27.98	26.82	27.21	26.82
Ref.7	25.37	25.79	25.07	25.15
Ref.8	25.19	25.17	25.45	25.03
The proposed method	21.22	20.51	20.97	21.08

It can be seen from Table 1 that after 3D reconstruction of the human body targets of athletes performing four kinds of basketball movements by different algorithms, the recognition error of the reconstruction results of the algorithm in Ref.4 is the highest. When the basketball athletes jump, the recognition error rate reaches 28.96%, and the recognition effect is the worst. It is said that the algorithm in Ref.4 has poor reconstruction effect on the human body targets performing jump movements, the convolutional neural network action recognition model is difficult to accurately identify actions; The recognition effect of the reconstruction results of the algorithm in Ref.5 and the algorithm in Ref.6 is not much different. Although the recognition error rate is lower than that of the algorithm in Ref.4, there is still a high error recognition rate, which indicates that there is little difference in the reconstruction effect of the moving human body target between the two

algorithms; The recognition error rate of the reconstruction results of Ref.7 algorithm and Ref.8 algorithm is about 25%, and the recognition effect is improved. However, when compared with the recognition error rate of the reconstruction results of this algorithm, it can be clearly seen that the recognition error rate of the reconstruction results of this algorithm is the lowest. When basketball athletes are running, the error recognition rate of the reconstruction results is only 20.51%. The experiment uses the convolutional neural network action recognition model to verify the three-dimensional non-rigid reconstruction of the human target of the algorithm in this paper. The implementation results show that the algorithm can accurately reconstruct the movement of different basketball movements, which is convenient for the action recognition model to accurately identify specific actions, and has good reliability.

B. COMPARATIVE TEST OF 3D RECONSTRUCTION ACCURACY

1) PIXEL FEATURE MATCHING ACCURACY TEST

In order to prove that the 3D reconstruction of non-rigid moving human objects in this algorithm has better pixel feature matching accuracy, six algorithms are used to match the feature pixels in Figure 4. The comparison results are shown in Table 2.

TABLE 2. Comparison results of 3D reconstruction accuracy.

Method	Pixel matching degree /(%)			Reconstruction Accuracy
	(1)	(2)	(3)	
Ref.4	65.28	68.77	70.25	68.05
Ref.5	72.39	79.65	80.37	75.94
Ref.6	71.62	78.43	82.46	76.23
Ref.7	83.58	81.69	89.72	84.76
Ref.8	83.99	82.31	88.29	85.47
The proposed method	95.68	93.47	96.53	95.98

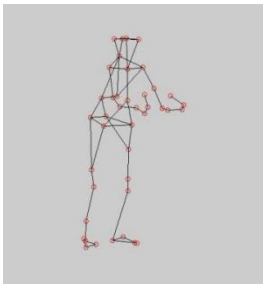


FIGURE 4. Reference image.

It can be seen from the analysis of Table 2 that the pixel feature matching degree of the Ref.4 algorithm is low, and the 3D reconstruction accuracy of the non-rigid moving human target is low, and the reliability is poor; The reconstruction accuracy of the algorithms in Refs.5-6 is 75.94%, which is obviously low. The 3D reconstruction effect of non-rigid moving human objects is not ideal, and the reliability is poor; The pixel matching degree of the algorithm in Refs.7-8 are more than 81.69%, and the reconstruction accuracy is about 85%. The 3D reconstruction effect of the non-rigid moving human target has improved, while the pixel matching degree of the algorithm in this paper can reach 96.53% at the highest, and the reconstruction accuracy is 95.98%, which shows that the 3D reconstruction effect and reliability of the algorithm in this paper are the best for the non-rigid moving human target.

2) ALGORITHM COMPARISON RESULTS OF IMAGES FROM DIFFERENT FRAMES

In the acquired sequence images of basketball movement, the images of the 50th frame, 100 frame and 150 frame were selected, and the different methods in the literature and the algorithm were compared to test the reconstruction results of different methods, and the experimental results were obtained as shown in Figures 5-7.

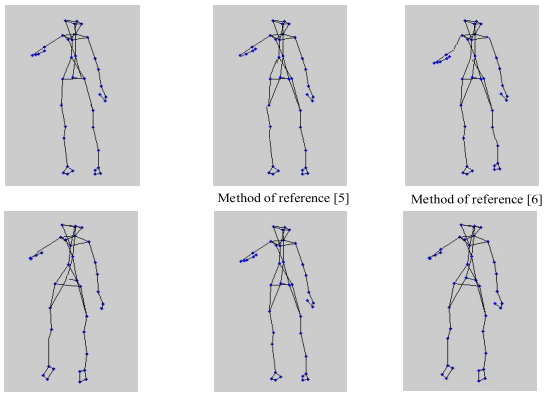


FIGURE 5. Reconstruction results of 50th frame in action "Pass."

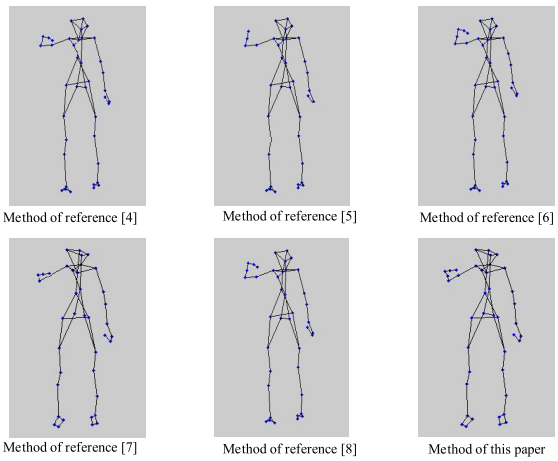


FIGURE 6. Reconstruction results of 100th frame in action "Pass."

The experimental results from Figures 5-7 reveal that under the conditions of reconstructing different frames of images using different methods, the approach in Reference [4] frequently exhibits the phenomenon of missing information in certain body parts. This may be attributed to the method's inability to completely capture certain body part information during rapid human movement or complex postures, leading

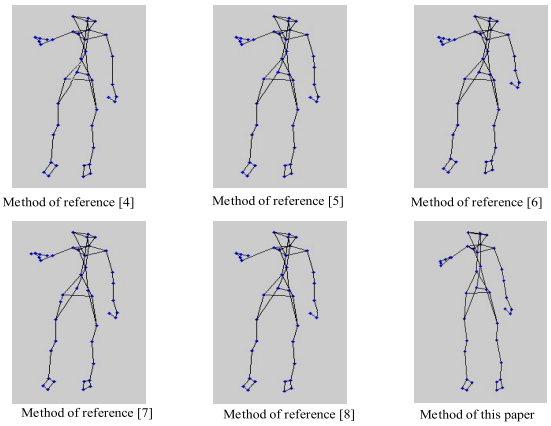


FIGURE 7. Reconstruction results of 150th frame in action "Pass."

to occlusion or partial information loss and consequently resulting in poor reconstruction effectiveness.

The method in Reference [5] encounters the issue of non-rigid deformation, as the human body typically presents non-rigid changes during motion, such as joint bending and muscle stretching. This non-rigid deformation poses a challenge to the reconstruction of three-dimensional nodes, and it is possible that the method in Reference [5] did not consider non-rigid change scenarios, resulting in deformation and deviation in the reconstruction results.

The method in Reference [6] may have registration issues due to the assumption that the object's motion is not completely known, leading to inaccurate matching of feature point positions from different perspectives or the failure to accurately integrate information from certain regions into the overall reconstruction process, thus affecting the accuracy of the three-dimensional reconstruction.

The method in Reference [7] produces incongruent or unreasonable reconstruction results due to the presence of numerous constraints during the matching process. Particularly, in the reconstruction results of the 150th frame, it is evident that this method may have been affected by noise interference, resulting in biased reconstruction nodes and decreased reconstruction accuracy.

Furthermore, the method in Reference [8] exhibits scattered points or deviation during the reconstruction process, possibly due to its heavy reliance on inverse imaging, which introduces excessive nodes and leads to overfitting, resulting in unnecessary details in the reconstruction and lowered algorithm accuracy.

C. PEAK SIGNAL-TO-NOISE RATIO CONTRAST RESULTS

Peak Signal-to-Noise Ratio (PSNR) is a specific form of signal-to-noise ratio that is commonly used to measure the quality of distorted images, with the maximum possible pixel value of the original image serving as a reference. The calculation formula is follows:

$$\mathrm{PSNR} = \frac{\left(\mathrm{d_x}, \mathrm{d_y}\right) \mathrm{f_{max}}(\mathrm{x}, \mathrm{y})}{\left[\mathrm{f}(\mathrm{x}, \mathrm{y}) - \mathrm{f}(\mathrm{x}^{\mathrm{i}+1}, \mathrm{y}^{\mathrm{i}+1})\right]^2} \tag{32}$$

where, $\mathrm{f}(\mathrm{x}, \mathrm{y})$ is the raw sequence images, $\mathrm{f_{max}}(\mathrm{x}, \mathrm{y})$ represents the maximum gray value of the image.

PSNR can be used to compare the differences between the reconstructed image and the original image, thereby assessing the fidelity of the reconstructed image. In general, a higher PSNR value indicates a smaller difference between the reconstructed image and the original image, and thus, a higher fidelity. Due to the fact that the magnitude of PSNR values can intuitively reflect the quality difference between the reconstructed image and the original image, it provides comparability and intuitiveness for the effects of different reconstruction algorithms and parameter settings. Therefore, PSNR is chosen as the indicator for comparative testing with different methods in the literature, and the experimental results were obtained as shown in Table 3.

TABLE 3. Comparison results of PSNR.

Method	PSNR		
	Frame 50	Frame 100	Frame 150
Ref.4	70.56	71.36	70.25
Ref.5	71.32	72.18	73.10
Ref.6	72.13	72.36	72.65
Ref.7	73.15	73.59	73.79
Ref.8	76.64	77.38	76.81
The proposed method	93.55	93.89	94.56

According to Table 3, the PSNR values of References 4-8 are not significantly different at the same frame rate, indicating that their reconstruction effects are relatively similar. The proposed method exhibits significantly higher PSNR values at all frame rates, indicating its significant advantage in image reconstruction. By comparing the PSNR values of different methods, it can be found that the proposed new method has significant advantages in image reconstruction quality. Therefore, it can be concluded that the proposed method has higher fidelity and better performance in image reconstruction. In practical applications, it is demonstrated that the proposed method can demonstrate better performance and higher quality in the fields of video reconstruction and image processing, which is of great significance for scientific research and engineering applications.

D. STRUCTURAL SIMILARITY CONTRAST RESULTS

Structural similarity index (SSIM) for the brightness, contrast and structure of the image, so that it can be effective to evaluate the different types of distortion may occur in the process of image reconstruction, SSIM is more sensitive to the loss of uniform and heterogeneous areas in the image, and can better reflect the details of the image. The calculation formula is follows:

$$\mathrm{SSIM} = \frac{\left[\left(\mathrm{I_x}, \mathrm{I_y}\right) + \mathrm{C_1}\right]\left[\left(\sigma_\mathrm{x}, \sigma_\mathrm{y}\right) + \mathrm{C_2}\right]}{\left(\mathrm{I_x}, \mathrm{I_y}\right)\left(\sigma_\mathrm{x}, \sigma_\mathrm{y}\right)} \tag{33}$$

where, $\left(\mathrm{I_x}, \mathrm{I_y}\right)$, $\left(\sigma_\mathrm{x}, \sigma_\mathrm{y}\right)$ represents the respective the brightness and contrast of pixels (x, y), and both $\mathrm{C_1}$ and $\mathrm{C_2}$ represent constant terms that ensure that the denominator is not zero.

The larger the SSIM, the higher the visual perception quality such as brightness, contrast, and linear correlation, indicating the better the reconstruction of the algorithm. The experimental results were obtained as shown in Table 4.

TABLE 4. Comparison results of SSIM.

Method	SSIM		
	Frame 50	Frame 100	Frame 150
Ref.4	70.34	71.12	69.98
Ref.5	72.23	72.34	72.65
Ref.6	76.92	76.53	75.15
Ref.7	78.59	78.25	78.28
Ref.8	78.56	78.15	79.35
The proposed method	92.96	93.57	93.78

According to Table 4, different methods exhibit different trends in SSIM values at different frame rates, while the proposed method exhibits significantly higher SSIM values at all frame rates, indicating better performance in visual perception quality such as image brightness, contrast, and linear correlation. The SSIM values of references 4 to 8 are not significantly different at the same frame rate, while the proposed method has significantly higher SSIM values at all frame rates than the references, indicating that the new method has significant advantages in evaluating image structure similarity. By comparing the SSIM values of different methods, it is evident that the proposed new method has significant advantages in image reconstruction quality.

E. MEAN SQUARE ERROR CONTRAST RESULTS

Mean Square Error (MSE) is a commonly used evaluation metric in image reconstruction, which can intuitively express the degree of difference between the reconstructed image and the original image, quantifying the differences between the reconstructed image and the original image. The calculation formula is follows:

$$\mathrm{MSE} = \frac{1}{(d_x, d_y)} f(x, y) - f(x^{i+1}, y^{i+1}) \tag{34}$$

Generally, a larger MSE usually indicates a significant difference between the reconstructed image and the original image.The experimental results were obtained as shown in Table 5.

TABLE 5. Comparison results of MSE.

Method	MSE		
	Frame 50	Frame 100	Frame 150
Ref.4	0.93	1.20	0.99
Ref.5	0.90	1.21	1.01
Ref.6	1.05	1.09	1.13
Ref.7	1.02	0.96	1.23
Ref.8	0.98	1.22	0.97
The proposed method	0.09	0.11	0.13

According to Table 5, it can be seen that the methods in references 4 to 8 have relatively large mean square error (MSE) values at different frame rates, indicating a significant difference between the reconstructed image and the original image. The MSE values of the proposed method are significantly lower than those of the reference at all frame rates, indicating that the new method can better maintain the consistency between the reconstructed image and the original image during the image reconstruction process. Different methods show significant differences in MSE values at different frame rates, while the proposed method has significantly lower MSE values at all frame rates than the reference, indicating its significant advantage in image reconstruction quality.

F. RELIABILITY OF THE 3D RECONSTRUCTION

In order to further verify the reliability of the algorithm in this paper for 3D reconstruction of non-rigid moving human objects, the deflection angle range of the center axis of the non-rigid moving human object relative to the radar line of sight is θ: 3 ∼180; Uniform sampling, the target axis deflection angle between two adjacent frames is $\Delta\theta$: 1 ∼8; Sampling frames: F > 3; SNR of image: 0dB∼30dB. The position distribution results of the reconstructed scattering points and the real scattering points are shown in Figure 8.

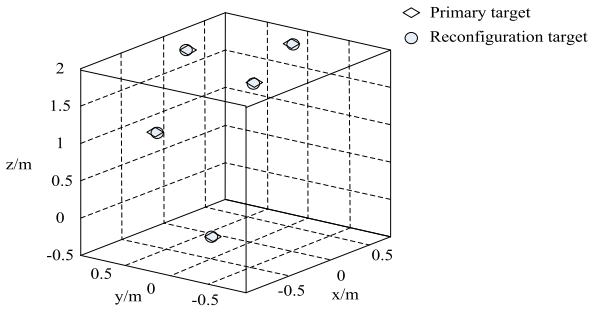


FIGURE 8. 3D reconstruction effect of non-rigid moving human object.

As can be seen from Figure 8, after 3D reconstruction of non-rigid moving human target by the algorithm in this paper, the best estimate of the center position of the reconstructed scattering points is obtained. The distribution of 3D reconstructed scattering points of non-rigid moving human target is very close to the 3D position distribution of real scattering points, which proves that the 3D reconstruction of non-rigid moving human target by the algorithm in this paper has good reliability.

The Procrustes distance (Pro-D) is used as the evaluation index for 3D reconstruction of non-rigid moving human targets in this algorithm, and the algorithm in this paper is tested under different imaging frames F and SNR, respectively. The impact of these two parameters on the 3D reconstruction performance of non-rigid moving human targets is statistically obtained. The smaller the Pro-D, the better the reliability of 3D reconstruction of non-rigid moving human targets. The verification results are shown in Figure 9.

It can be seen from the analysis of Figure 9 that under each signal to noise ratio, with the increase of the number of two-dimensional image frames F used for 3D reconstruction, the error distance of the obtained three-dimensional

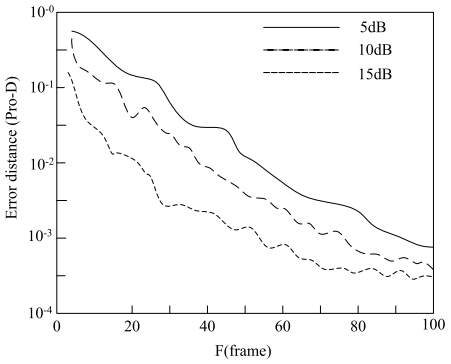


FIGURE 9. Influence of image frame number and SNR on 3D reconstruction.

reconstruction of the non-rigid moving human target shows a trend of decreasing; In the same case of F, with the increase of image SNR, the error distance of 3D reconstruction of non-rigid moving human objects also decreases, that is, the larger the F and SNR, the better the 3D reconstruction performance of non-rigid moving human objects. However, no matter what the SNR value is, when F is 100, the Pro-D value is below 10-3, indicating that the SNR value has little impact on the reconstruction effect of moving human objects in this algorithm, it has good reconstruction effect and reliability.

IV. RESULTS AND DISCUSSION

In this paper, a reliable 3D non-rigid motion human body target reconstruction method based on the estimation of contour deformation degree is proposed. In the experimental process, recognition error, reconstruction accuracy, PSNR, SSIM, MSE, etc., are used as evaluation metrics for algorithm performance to verify the performance results from multiple aspects. Based on the experimental results obtained, the following in-depth analysis will be conducted to explore and clarify the performance advantages of the proposed method in the process of non-rigid motion human body target reconstruction.

(1) From the reconstruction error results of the algorithm, it can be seen that compared to various methods in the literature, the reconstruction error of the algorithm in this paper is the lowest. In general, image reconstruction error is mainly related to image acquisition noise and data quality.

In the process of collecting image data in this paper, motion human body target features were obtained based on SIFT. In practical applications, for non-rigid motion human body targets, the movement of the body may lead to the appearance of features of different scales in the images. The scale of human body motion data is large, and due to the non-rigid nature of human body motion, the human body image targets in different frames may undergo rotation or positional changes. The SIFT algorithm has good stability and can be used to extract complex non-rigid targets, accurately matching different feature points, thereby ensuring that the original image features are extracted with as little noise interference as possible, ensuring the data quality in the images.

Meanwhile, in the subsequent processing, a bilateral filtering method is used to filter out the interference of the collected appearance contour pixels. This approach can retain the edge features of the motion human body target image while removing potentially interfering background data, further improving data quality.

(2) Using the accuracy of pixel feature matching as an evaluation criterion and comparing it with different methods in the literature, the reliability of the method designed in this paper is verified under different motion conditions.

Experiments have shown that the method designed in this paper achieves the highest accuracy in pixel feature matching after reconstructing images from different frames. The reconstructed nodes are reasonable, without deformation or deviation, resulting in good coherence in the reconstructed images, making them suitable for non-rigid deformation of human motion targets.

The accuracy and effectiveness of pixel feature matching are influenced by various factors, including noise and distortion in the images, occlusion, and partial visibility. In the process of 3D target reconstruction, this paper first uses threshold segmentation to remove background interference and eliminate irrelevant background data in the images, thereby reducing the impact of noise and distortion on the matching results and improving the authenticity of the images. This helps in retaining valid target-related features and reducing interference from irrelevant information in the matching process.

Moreover, considering noise and data loss factors, a K-value estimation method is proposed to address inaccurate matching resulting from noise and data loss. Estimating and correcting the data can improve the problems caused by data loss, ensuring the integrity and accuracy of the data.

Finally, by introducing non-uniform and correlated tracking point position errors, it is possible to better simulate complex scenarios in real situations, thus improving the robustness and usability of data processing. Continuously engaging in the cycle of ''prediction-matching-updating'' helps to improve the robustness and accuracy of the matching, enhancing the usability of the data, ultimately ensuring the precision of the image reconstruction.

(3) Metrics such as PSNR, SSIM, and MSE provide evaluations of image reconstruction effectiveness at different levels. Upon obtaining evaluation results based on these metrics, it was found that compared to methods in different literature, the performance indicators PSNR, SSIM, and MSE of the method in this paper were better, consistently optimal, and exhibited smaller fluctuations. This indicates that the stability of the designed algorithm is relatively high, resulting in good reconstruction effects. This is mainly attributed to the consideration of a reliable estimate of contour deformation levels in the method proposed in this paper.

Reliable estimation of contour deformation levels can assist in shape matching between different images, especially for objects or organizational structures with complex morphologies. The research target of this paper is the non-rigid motion human body, which exhibits dynamic deformation characteristics during motion with continuously changing positions. So reliable contour deformation levels can better capture the dynamic changes of the human body and provide a quantitative evaluation method to assist in the registration and alignment of dynamic human body target images. This enables better correspondence and matching of images collected from different viewpoints or at different times, thus facilitating the precise reconstruction of three-dimensional non-rigid motion human body targets.

Prior to reliability assessment, this paper constructs a probabilistic graphical model that takes into account the contour, orientation, and speed of human body motion to obtain observational probabilities for different human motion actions. This helps in achieving more accurate estimation of human body posture and motion states, understanding the spatial structure of human motion, capturing details and features during motion, and providing a fundamental guarantee for reliable estimation of contour deformation levels.

(4) In order to further verify the reliability of the algorithm for reconstructing non-rigid motion human body targets, the distribution of reconstructed scatter points and real scatter points was used to evaluate the reconstruction effectiveness. By comparing the distributions of the reconstructed scatter points and the real scatter points, analyzing the proximity can help to determine the effectiveness of the reconstruction algorithm. If the reconstructed results closely match the real distribution of points, this will demonstrate the high reliability and effectiveness of the algorithm in capturing the shape and position information of the target. Therefore, especially in dealing with non-rigid motion targets, analyzing the distribution of reconstructed scatter points and real scatter points helps identify errors and biases in the reconstruction process, determining the accuracy and realism of the reconstruction results.

The experiment shows that the algorithm proposed in this paper provided the best estimate of the center position of the reconstructed scattered points after performing 3D reconstruction of non-rigid moving human body targets. Furthermore, the distribution of the reconstructed scattered points closely matches the actual distribution of the scattered points.

This result confirms the favorable reliability of the algorithm for 3D reconstruction of non-rigid moving human body targets. It indicates that the algorithm can effectively capture and reconstruct the morphology and positions of body targets, thereby producing reconstruction results that closely resemble or match real scene.

Meanwhile, with an increase in the number of 2D image frames utilized for 3D reconstruction, the error distance of the reconstructed body targets show a continuous decrease. This suggests that with an increased number of 2D image frames, the 3D reconstruction effect for non-rigid moving targets will improve, as more perspectives and information can be utilized for reconstruction, thereby enhancing the accuracy and reliability of the reconstruction. This demonstrates that the algorithm can maintain good reconstruction results under relatively high signal-to-noise ratio conditions, exhibiting a certain degree of robustness.

(5) The acquired appearance contour sequence of the moving human body target is a 2D sequence, but this paper transforms the 2D contour appearance depth data into 3D point cloud data. Then, the bilateral filtering method is used to filter the collected appearance contour pixel interference, achieving the purpose of separating background from the human body.

In this process, recovering 3D spatial information from separate 2D information can help better capture the shape and details of the human body surface. At the same time, by using the bilateral filtering method to filter the interference of appearance contour pixels, noise and irrelevant information can be effectively removed, thereby achieving effective separation of the background from the human body, enabling more accurate modeling and reconstruction of the human body in subsequent processing.

Combining the methods section of the article and the above analysis, it can be seen that this paper is dedicated to reducing noise, avoiding interference, and improving data integrity. Furthermore, it considers non-uniformity and correlation when processing data, and designs a reliable estimation method based on the degree of contour deformation for reconstructing non-rigid 3D human motion targets. This method aims to enhance the accuracy of image feature matching and the authenticity of reconstruction results.

V. EXPERIMENTAL ANALYSIS

To obtain the features of human motion sequence data through SIFT, and to ensure the continuity of feature sequence extraction for moving human targets and the reliability of the algorithm, a probability model of contour appearance sequence for moving human targets is established; On the basis of completing image preprocessing, the improved algorithm for 3D reconstruction of non-rigid moving human targets proposed in this paper has a relatively simple process, which can effectively calculate the number of shape bases and accurately track the movements of athletes. The 3D reconstruction effect is good, with strong robustness and reliability. At the same time, the 3D reconstruction effect of non-rigid moving human targets in this paper is less affected by signal-to-noise ratio.

REFERENCES

[1] P. Liang, W. Zheng, and X. Xu, "Three-dimensional reconstruction method based on jitter optimization," *J. Phys., Conf. Ser.*, vol. 1914, no. 1, May 2021, Art. no. 012006.

[2] C. Li, G. Dong, and R. Wang, "A three-dimensional reconstruction algorithm of nonwoven fabric based on an anthill model," *Textile Res. J.*, vol. 92, nos. 11–12, pp. 1876–1890, Jun. 2022.

[3] Z. Liu and G. Huang, "Simulation of multi object image reconstruction of human motion pose with multi degree of freedom," *Comput. Simul.*, vol. 38, no. 7, pp. 194–197, 2021.
[4] M. Coenen and F. Rottensteiner, "Pose estimation and 3D reconstruction of vehicles from stereo-images using a subcategory-aware shape prior," *ISPRS J. Photogramm. Remote Sens.*, vol. 181, pp. 27–47, Nov. 2021.
[5] S.-Y. Kim, J.-S. Kim, J. H. Lee, J. H. Kim, and T.-S. Han, "Comparison of microstructure characterization methods by two-point correlation functions and reconstruction of 3D microstructures using 2D TEM images with high degree of phase clustering," *Mater. Characterization*, vol. 172, Feb. 2021, Art. no. 110876.
[6] Z. Mejri, L. Sidhom, and A. Abdelkrim, "Three-dimensional structure from motion recovery of a moving object with noisy measurement," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 10, no. 1, p. 117, Feb. 2020.
[7] H. Han, Y. Wang, Y. Zou, J. Liao, and Y. Xu, "Three-dimensional substructure imaging of blood cells using maximum entropy tomography based on two non-orthogonal phase images," *Opt. Laser Technol.*, vol. 136, Apr. 2021, Art. no. 106799.
[8] K. Knudsen and A. K. Rasmussen, "Direct regularized reconstruction for the three-dimensional calderón problem," *Inverse Problems Imag.*, vol. 16, no. 4, pp. 871–894, 2022.
[9] A. Afanasev, J. J. Kingsley-Smith, F. J. Rodríguez-Fortuño, and A. V. Zayats, "Nondiffractive three-dimensional polarization features of optical vortex beams," *Adv. Photon. Nexus*, vol. 2, no. 2, Jan. 2023, Art. no. 026001.
[10] S. Huang, W. Fu, Z. Zhang, and S. Liu, "Global-local fusion based on adversarial sample generation for image-text matching," *Inf. Fusion*, vol. 103, Mar. 2024, Art. no. 102084.
[11] S. Liu, S. Huang, S. Wang, K. Muhammad, P. Bellavista, and J. Del Ser, "Visual tracking in complex scenes: A location fusion mechanism based on the combination of multiple visual cognition flows," *Inf. Fusion*, vol. 96, pp. 281–296, Aug. 2023.
[12] G. Wang, X. Qin, J. Shen, Z. Zhang, D. Han, and C. Jiang, "Quantitative analysis of microscopic structure and gas seepage characteristics of low-rank coal based on CT three-dimensional reconstruction of CT images and fractal theory," *Fuel*, vol. 256, Nov. 2019, Art. no. 115900.
[13] C. Fan, W. Fu, and S. Liu, "A high-precision correction method in non-rigid 3D motion poses reconstruction," *Connection Sci.*, vol. 34, no. 1, pp. 2845–2859, Dec. 2022.
[14] S. Xi, W. Li, J. Xie, and F. Mo, "Feature point matching between infrared image and visible light image based on SIFT and ORB operators," *Infr. Technoiogy*, vol. 42, no. 2, pp. 168–175, Feb. 2020.
[15] Y. Zihao, L. Jin, Y. Haima, Z. Pengcheng, and C. Yi, "Three-dimensional surface reconstruction based on edge detection and reliability sorting algorithm," *Laser Optoelectronics Prog.*, vol. 57, no. 24, 2020, Art. no. 241020.
[16] J. Huo and X. Yu, "Three-dimensional mechanical parts reconstruction technology based on two-dimensional image," *Int. J. Adv. Robotic Syst.*, vol. 17, no. 2, Mar. 2020, Art. no. 172988142091000.
[17] X. P. Wang, H. Lei, Y. Liu, and N. Sang, "Balanced functional maps for three-dimensional non-rigid shape registration," *J. Electron. Technol.*, vol. 19, no. 4, pp. 369–378, 2021.
[18] S. Liu, P. Chen, and Y. Zhang, "A multiscale feature pyramid SAR ship detection network with robust background interference," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 9904–9915, 2023.
[19] H. Liu, Z. Jin, Y. Xiang, and K. Ji, "High-resolution solar image reconstruction based on non-rigid alignment," *Res. Astron. Astrophys.*, vol. 22, no. 9, Sep. 2022, Art. no. 095005.
[20] M. Jian, J. Dong, M. Gong, H. Yu, L. Nie, Y. Yin, and K.-M. Lam, "Learning the traditional art of Chinese calligraphy via three-dimensional reconstruction and assessment," *IEEE Trans. Multimedia*, vol. 22, no. 4, pp. 970–979, Apr. 2020.
[21] S. Liu, X. Xu, Y. Zhang, K. Muhammad, and W. Fu, "A reliable sample selection strategy for weakly supervised visual tracking," *IEEE Trans. Rel.*, vol. 72, no. 1, pp. 15–26, Mar. 2023.
[22] S. Yang and M. Wang, "Reconstruction of ablation morphology of TiC/tungsten-based composites under digital image processing," *Ordnance Mater. Sci. Eng.*, vol. 45, no. 5, pp. 92–96, 2021.
[23] S. Kamali, T. Amraee, and M. Khorsand, "Intentional power system islanding under cascading outages using energy function method," *IET Gener., Transmiss. Distrib.*, vol. 14, no. 20, pp. 4553–4562, Oct. 2020.
[24] D. Wang, X. Jin, J. Zhao, Y. Wang, L. Rong, and J. J. Healy, "Continuous-wave terahertz diffraction tomography for measuring three-dimensional refractive index maps," *Chin. Opt. Lett.*, vol. 19, no. 12, 2021, Art. no. 123701.

YAN ZHANG was born in Beijing, in 1975. She received the B.S. degree from the Wuhan Sports Institute, in 2003.

From 1988 to 1999, she was a professional basketball player in Hubei Province. Since 2003, she has been a full-time Sports Teacher with the Sports Department, Zhongnan University of Economics and Law, Wuhan, China. During her time in school, she has published five public articles. She have been a PI/Co-PI for many governmental projects. Her research interests include intelligent physical education, physical data analysis, and scientific fitness.

Dr. Zhang was a recipient of many CUBA Female's Basketball High-Level Group in China. She also led the Basketball Team to Win the Runner Up and Third Place multiple times in the CUBA, respectively, and the Runner Up and Champion of the Provincial University Group. She received the "Outstanding Teacher" of the School, in 2017, "Outstanding Female Teacher" of the School, in 2018, and "Fair, Beautiful, Virtuous and Virtuous Teacher" of the School, in 2021.

MOHAMED BAZA received the B.S. and M.S. degrees in electrical and computer engineering from Benha University, Egypt, in 2012 and 2017, respectively, and the Ph.D. degree in electrical and computer engineering from Tennessee Tech University, Cookeville, TN, USA, in December 2020.

Since August 2021, he has been with the Department of Computer Science, College of Charleston, Charleston, SC, USA. He is the author of numerous articles published in major IEEE journals, such as IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Vehicular Technology, IEEE Transactions on Network Science and Engineering, and IEEE Systems Journal, and major IEEE conferences, such as IEEE Wireless Communications and Networking Conference, IEEE International Conference on Communications, and IEEE Vehicular Technology Conference. His research interests include blockchains, cybersecurity, machine learning, smart grids, and vehicular ad-hoc networks.

HANI ALSHAHRANI (Member, IEEE) received the bachelor's degree in computer science from King Khalid University, Abha, Saudi Arabia, the master's degree in computer science from California Lutheran University, Thousand Oaks, CA, USA, and the Ph.D. degree from Oakland University, Rochester, MI, USA.

He is currently an Associate Professor of computer science and information systems with Najran University, Najran, Saudi Arabia. His research interests include smartphones, the IoT, crowdsourcing security, and privacy.