Development of Three-dimensional Human Motion Modeling Framework and Its Computational Techniques

Chenyu Li

Culver Academies, Culver, IN, United States of America

Chenyu.li@culver.org

Abstract. The research of virtual human motion modeling and simulation technology is a challenging subject involving physics, robotics, biomechanics, and computer graphics. Virtual human motion simulation based on a physical model is a research hotspot of computer graphics and animation technology, but the algorithm of kinematics and dynamics calculation has not been deeply studied. In this paper, we developed a protocol where the joints are removed in the reverse order of the connection stage in the segmentation stage, and the constraint forces and joint accelerations of the complete kinematic chain are solved. And the hierarchical control strategy of the human reach movement was discussed, and the experimental prototype of the human reach movement was constructed to verify the correctness and effectiveness of the research results.

Keywords: Human motion modeling, computational techniques, optimal control

1 Introduction

In the field of computer vision, the study of human motion analysis has broad application prospects. Due to the complexity of human motion, existing research methods impose many constraints on the human body of the research project. In this paper, we present a new method to study various types of human motion information to mainly discuss the reconstruction of the three-dimensional motion skeleton of the human body [1,2]. The basic idea is to build an image of each image based on calibration, applying knowledge of the 3D human models and motion continuity [3]. Based on obtaining binarized moving images, automatic labeling of joint feature points is realized. First, the target contour is extracted using the operator, the appropriate human contour shape is obtained using the vector decomposition algorithm [4]. The position of each joint point is determined according to the joint proportional structure, and the joint feature points are automatically marked [5]. Three-dimensional human motion tracking technology is studied. The marked joint points are tracked by the feature optical flow algorithm, and the joint points of the tracking error are corrected by the Kalman filter [6]. The automatic labeling method of the joint points of the human

H. Wu et al. (Eds.): CSIEDE 2022, ACSR 103, pp. 58–66, 2022.
https://doi.org/10.2991/978-94-6463-108-1_8

body model reduces the manual intervention of the system to a certain extent [7]. Three-dimensional human motion tracking technology is studied.

Three-dimensional reconstruction technology is a mathematical process and computer technology that only uses two-dimensional projections or images to restore the three-dimensional information of objects. The 3D reconstruction in computer vision mainly refers to the process of reconstructing the original 3D information by using sensor devices (such as cameras) to obtain images of the real world to obtain depth information of the real scene. Multi-view stereo vision uses the output of SFM to compute depth and normal maps for each pixel in an image. The fusion of depth and normal maps from multiple images in 3D produces a dense point cloud of the scene. The 3D surface geometry of the scene is recovered using algorithms such as depth and normal maps of the fused point cloud, as well as Poisson surface reconstruction.

3D modeling has always been an important research topic in computer vision [8,9]. As an integral part of computer-human simulation, 3D human modeling first appeared in the human-machine system of the aerospace industry [10]. Since the birth of interactive computer graphics, some scholars have been exploring computer-human modeling technology. With the continuous development of computer technology, three-dimensional human modeling is now widely used in scientific research, animation, computer games, clothing design, industry, and other fields. comprehensive application.

Human body motion analysis is to obtain the motion information of the human body from the image sequence or video and analyze and identify it. Because this research has a very broad application prospect, it is being paid more and more attention by researchers as the images are very useful in sports analysis, medical analysis, animation production, etc. In addition, automatically identifying and understanding human motion can be widely used in surveillance, human-computer interaction, video retrieval, and other fields. Researchers usually directly use traditional robotics motion modeling and simulation control techniques. However, due to the huge degree of freedom and complex kinematic chain structure of the human body model, the motion control technology of the human body model is much more difficult than the robot system.

3D reconstruction refers to the establishment of a mathematical model suitable for computer representation and processing of 3D objects. It is the basis for processing, operating and analyzing its properties in a computer environment.

In computer vision, 3D reconstruction refers to the process of reconstructing 3D information based on single-view or multi-view images. Since the information in a single view is incomplete, 3D reconstruction requires the use of empirical knowledge. Multi-view 3D reconstruction (like human dual target positioning) is relatively easy. The method is to calibrate the camera first, that is, to calculate the relationship between the camera's image coordinate system and the world coordinate system. Then use the information in multiple two-dimensional images to reconstruct three-dimensional information. A community-based rotation error averaging method was proposed in a global manner, which considers both the accuracy of peripolar geometry and the accuracy of pairwise geometry. Based on the estimated absolute rotational pose of the camera, the camera optical center position was estimated in an incremental manner. For each added camera, its rotation and intrinsic remain the same, while the optical center and scene structure are refined using improved BA.

With the development of computing power and the increase of people's needs, 3D human models and their animations can be widely used in film and television, animation, virtual fitting, virtual reality, video games and other fields. In some applications, people have a strong need to customize 3D human models according to their body size. For example, in virtual fitting, users need to customize their own 3D mannequins. Further, the user also needs to make the customized human body make movements to show the effect of dressing in different postures. The research idea of this paper is to first construct a customized static 3D human model based on 2D images, and then use the skeleton and joints to drive the deformation of the model to construct 3D human animation.

2 Methods and Techniques

Traditional human motion analysis methods are generally divided into two types. One is to add sensors to each joint part of the human body. During the movement of the human body, the sensor will continuously return the position of each joint in space to the computer, so that the computer can accurately obtain the movement information of the human body at each moment [12]. The second is to analyze the image sequence. The image sequence can be taken by a single camera, or it can be taken by multiple cameras simultaneously from various viewpoints. This is generally done in a three-step sequence. The first step is to extract low-level features from image frames, such as the position of each joint, and identify them. The second step is to establish correspondence between the feature of each frame. Finally, the three-dimensional human body structure and motion information are recovered from the feature correspondence [13].

Badler used the method of mapping the image to a block model for research. Hogg, Rohr's system is only for the human walking model with one degree of freedom. They extracted edge and corner features from the images and matched them to a 3D cylindrical mannequin. Chen and Lee composed a skeleton model with 17-line segments and 14 joints to represent the human body and added various constraints during the analysis. Bharatkumar et al also used skeleton models to describe human lower limbs, but they tried to build a general model for gait analysis of human walking. Almost all of these methods impose many limitations due to the special complexity of human motion. For example, the background is required to be single or even motionless, and the human body wears bodysuits, etc. In addition, the motion of the human body will inevitably become unnatural after adding sensors.

The iterative Closest Point (Iterative Closest Point) algorithm is a high-level registration method based on free-form surface point cloud data proposed by Becl and Mckay in 1992. After the corresponding nearest point set is determined from the measured point cloud, the new nearest point set is calculated by the method proposed by Faugera and Hebert. This method is used for iterative calculation until the value of the objective function formed by the residual sum of squares remains unchanged, and the iterative process is ended [14]. The ICP algorithm is widely used to solve the problem of 3D point cloud registration. As a nonparametric spline, TPS (Thin-Plate-Spline) is widely used in flexible coordinate transformation due to its rich physical meaning and closed-form solution [15].

The movement of the human body in a certain space can be recorded according to the methods of video recognition and tracking, but such methods cannot establish an intuitive dynamic three-dimensional human body model [16]. Existing methods that only rely on skeleton information for model deformation cannot match the real human pose well [17]. Some modeling methods use the method of treating the human body as multiple rigid body parts, and deformities or model nesting will occur at joints such as elbows and knees, which cannot naturally display the real posture of the human body [18].

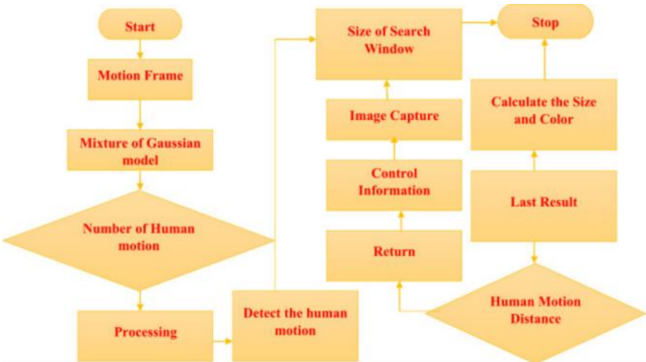

Fig. 1. Established image processing, simulation, and processing protocol developed for human motion analysis

As shown in **Figure 1**, a total of 20 human skeleton points were obtained, which are the hip, abdomen, neck, head, left shoulder, left elbow, left wrist, left hand, right shoulder, right elbow, right wrist, right hand, left hip, left knee, left ankle, left foot, right hip, right knee, right ankle, right foot. Correspondingly, the three-dimensional positions of 20 skeleton points are also marked on the standard human body template. According to the positions of the 20 skeleton points, the standard human body template is divided into seven parts: abdomen, chest, head, left arm, right arm, left leg, and right leg. In the large model part, the limb models of arms and legs are divided into three small model parts, for example, the left arm is divided into the left upper arm, the left forearm, and the left hand. A tree structure is formed between the seven large model parts, and the child model inherits the deformation matrix from the parent model; the limb model also has its tree structure, the child model inherits the deformation matrix from the parent model, and the parent model receives from the child model.

The deformation method adopted by the limb model is a deformation method based on the inheritance and feedback of the rotation matrix. Because the situation of the limb model is the same, the left arm model is used as an example for illustration. There are four joint points in the left arm model: left shoulder, left elbow, left wrist, and left hand. The human arm movement is more complex because there are two arm bones in the human forearm: the ulna and the radius. The two ends of the forearm are the elbow joint and the wrist joint, and the two joints are different. For the elbow joint, if both the shoulder joint and the wrist joint keep themselves from rotating in all directions, the arm can only perform the curved arm action driven by the biceps, that is, the elbow joint has only one degree of freedom. The up and down swing of the forearm is accomplished by the twisting of the shoulder joint to drive the forearm, not the elbow joint. In the model of the present invention, the shoulder, elbow, and wrist of the arm model of the template model are strictly along the X-axis direction.

The rotation matrix of the big arm and the forearm around the X-axis uses the same rotation matrix, and the shoulder and elbow joints included in the big arm can only determine the Ry and Rz of the big arm. The rotation matrix R x around X must be determined by the forearm, and then fed back to the big arm, which is the process of matrix feedback; on the other hand, due to the rotation constraint of the forearm (for example, the elbow cannot be broken), it is necessary to inherit the rotation matrix from the arm to adjust the attitude first, and then determine the rotation angle and rotation matrix, which is the inheritance of the matrix.

The rotation of the real human wrist is much more complicated than that of the elbow, but because the collection of the skeleton is more inaccurate as it goes to the end, and the specific rotation of the wrist cannot be analyzed only through the collected point cloud data, and the analysis of the overall action behavior of a human being and identification, the reference value of wrist motion is limited, and the rotation of the wrist and the elbow is regarded as the same situation, the difference is that only the degree of freedom of rotation of the wrist around the Z-axis is given, and only the rotation of the forearm is inherited without feedback to the forearm.

3 Results

System biomechanical simulation of the musculoskeletal system requires a representation of realistic muscle geometry. Most previous musculoskeletal models are based on multibody dynamics simulations that simplify muscles into one-dimensional (1D) line segments without considering the larger area of muscle attachment, the spatial fiber arrangement within the muscle, and the distance between the muscle and surrounding tissue. In previous musculoskeletal models with three-dimensional (3D) muscles, muscle contractions were input, which hindered the predictive power of these models. To address these issues, a finite element musculoskeletal model with the ability to predict 3D muscle contraction was developed. Muscles with realistic 3D geometry, taking up space for muscle fiber arrangement and muscle-muscle and muscle-skeletal interactions. Actively contractile stress of 3D muscles is determined based on the kinematic principles of gait lower extremity and ground force through an effective optimization method. The model also provides muscle stress and strain, as well as muscle-muscle contact mechanisms and muscle-bone interactions. The total knee contact force predicted by the model was in good agreement with in vivo measurements. Contact and wrapping between the muscle and surrounding tissue are evident, indicating that a 3D contact model of the muscle needs to be considered. This modeling framework (see **Figure 2.**) serves as a methodological basis for the development of a musculoskeletal modeling system in a finite element approach incorporating 3D muscle deformable contact models.

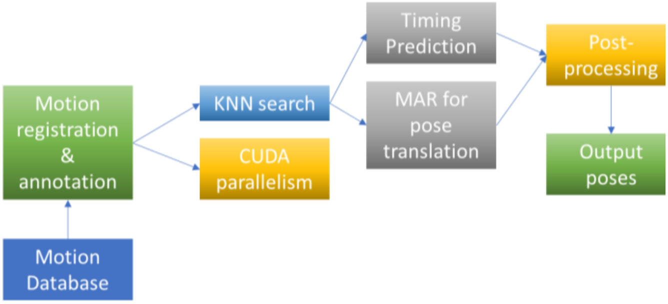

Fig. 2. 3D human motion modeling platform with motion registration and annotation as high-dimensional feature inputs

The current inverse kinematics solution algorithm has a high computational cost. The BHIK algorithm was implemented to transform the joint positions into a computational problem of finding a point on a straight line, calculating the joint positions through a bidirectional iterative process, and reducing the system error by adjusting a single joint angle each time. The algorithm decomposed the joint constraints into two parts: rotation and positioning and applies joint constraints in each step of the iterative calculation, to control the movement range of the joints. The BHIK algorithm avoided the rotation operation of the joints in the iterative calculation and uses the two-dimensional constraint ellipse to simplify the solution of the constraint problem, which overcomes the problems of high time complexity and easy generation of singularities of the current inverse kinematics method. Rod connectivity method was proposed to use pointers to indicate the relationship between adjacent rods, parent-child rods, and virtual rods, which is convenient for dynamic recursive calculation. The virtual open chain structure was obtained by dividing the original closed chain into several joints, and then the joint angle Jacobian matrix was driven by the open chain structure, and the dynamic equations of the general closed kinematic chain and the open kinematic chain were given. This method effectively solved the problem that traditional dynamic calculation methods cannot directly solve.

A connection segmentation algorithm CCA was proposed to solve the problem that the unit vector method is too complex to solve. The CCA algorithm realized the connection and segmentation of the target kinematic chain by adding and removing joints one by one. In the initial state, all joints were removed, and the links were completely unconstrained. In the connection phase, the joints were added one by one to create the target chain.

4 Conclusions

3D human modeling could acquire information about the human body, including predicting the contractions of the muscle fibers, which could help know how to maximize the work of the muscle to produce the most force while using the least joules of energy possible. Therefore, 3D modeling could be used in other fields of study, like sports.

Take fencing as an example. In fencing, we want speed, because the faster the blade moves, the harder for the opponent to react and counteract. If the opponent does not react in time, we will get a point. During the DE (Direct Elimination) round in fencing, people tend to get tired since each round of DE is 9 minutes for 15 points, and fencers have to move from one end of the strip to the other end in an instant, then back and forth throughout the whole match. Approaching the end of the match, fencers are often extremely tired. This affects their speed and accuracy. Being tired can decrease a fencer's speed, making fencers slower when doing blade movements and footwork. This could result in a fencer attacking too slowly and giving the opponent the time to catch the blade and score. This could also affect the fencer to not react fast enough when the opponent attacks, which would also result in the opponent scoring. In addition, when a fencer is tired, the accuracy of the blade movement decreases. Fencers tend to miss the target much more often towards the latter rounds of the DE. Thus, being able to reserve energy while being fast and accurate can be extremely beneficial to fencers, especially during DE, and this is where 3D human modeling comes in.

The main innovation of this paper lies in the generation of 3D human animation. It proposes a deformation method based on scaling and a skeleton subspace deformation method based on adaptive weights and realizes 3D human animation realistically. Image-based 3D human modeling allows users to obtain customized 3D human bodies quickly and easily. This modeling method first needs to take pictures of the front and side of the user to extract the human body feature points. In this paper, the existing human body feature point extraction method is studied, and the feature points are extracted by using the direction of human body contour points. Then according to the existing standard 3D human edge point information, the human body size is extracted from the image. Compare it to standard human body size and get a range of scaling ratios. Since the edge points of the standard human body in different parts have different characteristics, for different parts of the standard human body, use the free deformation method or the direct scaling method to deform.

This paper analyzes the effects of object rotation angle, skeleton position and skin width on the deformation results, and proposes a deformation method based on scaling and an adaptive weight skeleton subspace deformation method. Both of these methods avoid dents on the inside of the object's rotation. Through the layered skeleton model of the human body, the algorithm in this paper is applied to the rotation of the human elbow, and compared with some other animation algorithms, it is found that the two algorithms proposed in this paper can achieve better results. Finally, the adaptive weight skeleton subspace deformation method is applied to the human legs, and the human animation when walking is obtained.

References

1. Rallis, Ioannis, Ioannis Georgoulas, Nikolaos Doulamis, Athanasios Voulodimos, and Panagiotis Terzopoulos. "Extraction of key postures from 3D human motion data for cho-

reography summarization." In 2017 9th international conference on virtual worlds and games for serious applications (VS-Games), pp. 94-101. IEEE, 2017.
2. Zhang, Xin, Zhongqiu Xu, and Hongbo Liao. "Human motion tracking and 3D motion track detection technology based on visual information features and machine learning." Neural Computing and Applications (2022): 1-13.
3. Fu, Qiang, Xingui Zhang, Jinxiu Xu, and Haimin Zhang. "Capture of 3D Human Motion Pose in Virtual Reality Based on Video Recognition." Complexity 2020 (2020).
4. Wang, Xin, Qiudi Chen, and Wanliang Wang. "3D human motion editing and synthesis: A survey." Computational and Mathematical methods in medicine 2014 (2014).
5. Wei, Xiaolin, and Jinxiang Chai. "Videomocap: Modeling physically realistic human motion from monocular video sequences." In ACM SIGGRAPH 2010 papers, pp. 1-10. 2010.
6. Li, Jiaman, Ruben Villegas, Duygu Ceylan, Jimei Yang, Zhengfei Kuang, Hao Li, and Yajie Zhao. "Task-generic hierarchical human motion prior using vaes." In 2021 International Conference on 3D Vision (3DV), pp. 771-781. IEEE, 2021.
7. Shi, Guangtian, Yongsheng Wang, and Shuai Li. "Human motion capture system and its sensor analysis." Sensors & Transducers 172, no. 6 (2014): 206.
8. Howe, Nicholas, Michael Leventon, and William Freeman. "Bayesian reconstruction of 3d human motion from single-camera video." Advances in neural information processing systems 12 (1999).
9. Howe, Nicholas, Michael Leventon, and William Freeman. "Bayesian reconstruction of 3d human motion from single-camera video." Advances in neural information processing systems 12 (1999).
10. Manns, Martin, Stefan Mengel, and Markus Mauer. "Experimental effort of data-driven human motion simulation in automotive assembly." Procedia CIRP 44 (2016): 114-119.
11. Kakadiaris, Ioannis A., and Dimitri Metaxas. "Three-dimensional human body model acquisition from multiple views." International Journal of Computer Vision 30.3 (1998): 191-218.
12. Mündermann, Lars, Stefano Corazza, and Thomas P. Andriacchi. “The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications.” Journal of neuroengineering and rehabilitation 3.1 (2006): 1-11.
13. Taylor, Tad W., Haruka Okino, and Takami Yamaguchi. "Three-dimensional analysis of left ventricular ejection using computational fluid dynamics." (1994): 127-130.
14. Yan, Hongmei. Computations of fully nonlinear three-dimensional wave-body interactions. Diss. Massachusetts Institute of Technology, 2010.
15. Simo, Juan Carlos. "On a fully three-dimensional finite-strain viscoelastic damage model: formulation and computational aspects." Computer methods in applied mechanics and engineering 60.2 (1987): 153-173.
16. Ortiz, Michael, and Anna Pandolfi. "Finite-deformation irreversible cohesive elements for three-dimensional crack-propagation analysis." International journal for numerical methods in engineering 44.9 (1999): 1267-1282.
17. Gain, Arun L., Cameron Talischi, and Glaucio H. Paulino. "On the virtual element method for three-dimensional linear elasticity problems on arbitrary polyhedral meshes." Computer Methods in Applied Mechanics and Engineering 282 (2014): 132-160.
18. Lez-Perez, Alejandro, Rafael Sebastian, and Jose M. Ferrero. "Three-dimensional cardiac computational modelling: methods, features and applications." Biomedical engineering online 14.1 (2015): 1-31.