Real-time Animation and Motion Retargeting of Virtual Characters Based on Single RGB-D Camera

Ning Kang
Beihang University
Beihang University Qingdao Institute

Junxuan Bai
Beihang University

Junjun Pan*
Beihang University
Beihang University Qingdao Institute

Hong Qin†
Stony Brook University

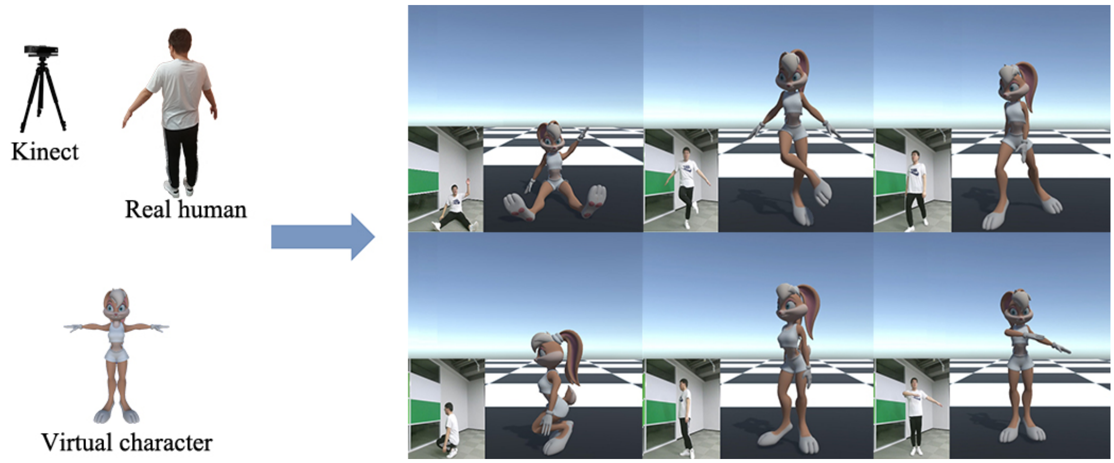

Figure 1: Our system captures the motion of real human using single Kinect, and then retargets the motion to a virtual character to produce animation.

ABSTRACT

The rapid generation and flexible reuse of characters animation by commodity devices are of significant importance to rich digital content production in virtual reality. This paper aims to handle the challenges of current motion imitation for human body in several indoor scenes (e.g., fitness training). We develop a real-time system based on single Kinect device, which is able to capture stable human motions and retarget to virtual characters. A large variety of motions and characters are tested to validate the efficiency and effectiveness of our system.

Index Terms: Character Animation; Motion Capture; Retargeting; RGB-D Camera; Adaptive filter; Inverse Kinematics

1 INTRODUCTION

Nowadays the animation of 3D characters has been widely applied in CG movies and video games. Since commodity depth camera became popular, it is possible to capture the motion of real human and generate an animated character rapidly. Although some toolkit developments, such as [1], have facilitated the basic method to capture the motion of human body, the quality of raw data is limited to be used directly, due to the data vibration and noises in the recorded motions. The motion retargeting process adapts the motion from real human to virtual character. Early researchers [3, 4] mainly focused on the motion transformation between characters with different sizes or topologies of the skeletons. Recent researchers [2, 5] tended to compute the transformation through an existing database, which contains the matches of the corresponding joints. These retargeting results are unavoidably affected by the content of the database. In this paper, we design a real-time system for producing stable character animation using single Microsoft Kinect device. As shown in Fig. 1, our system consists of two stages: capturing the human motion and retargeting the motion to a 3D virtual character. We combine and improve several the state of art techniques in our framework. To reduce the computation, we take the joints extracted by Kinect as input, then we retarget the captured motion to a character using quaternions. The actions of turning around and actions in side view are very common during in regular scene that occlusion usually appears and leads to failure. In order to reduce the artifacts, we integrate an adaptive filter and several inverse kinematics (IK) constraints into our system to refine the result.

*e-mail: pan_junjun@buaa.edu.cn
†e-mail: qin@cs.stonybrook.edu

2 SYSTEM DESCRIPTION

According to our observation, there are two main cases when capturing the motion: (1) the participant is facing the camera; (2) the participant is turning around, and the camera obtains a side view. It is difficult to handle the second case due to occlusion. In our system, we solve this problem using several methods. When the body turns around, the filter estimates the rotation angle of root, according to the length between projections of left and right hips. During the motion retargeting stage, we set up a dictionary that stores the skeleton hierarchy. The skeleton of virtual character is animated according to the hierarchy. For input motion sequence, using quaternions, we transfer the joints rotation of real human to 3D virtual character to generate animation. Finally, in order to avoid the posture distortions caused by the loss of skeleton, Inverse Kinematics (IK) constraints are employed to refine the positions during capturing.

2.1 Robust Motion Capture of Human Turning and Side View

Given a sequence of human motion data, we first calculate the 2D projection length between the left and right hip joints to draw a

2019 IEEE Conference on Virtual Reality and 3D User Interfaces
23-27 March, Osaka, Japan
978-1-7281-1377-7/19/$31.00 ©2019 IEEE

motion curve that includes information about whether the body rotates and how fast he moves. Through threshold segmentation, the curve of human hips projection length can be divided into three status: face to camera, side to camera and moving. The projection length of hip bone L can reach two peaks: the maximum value at 0 degree (face to camera) and the minimum value at 90 degrees or -90 degrees (side to camera). Then we introduce a gradient preserving method to smooth the motion curve when the person turning the body. Gradient retention aims at removing noise points from the motion curve and reducing vibration. Because the angle value of body turning is based on the bone length mapping, if the vibration of curve is not processed, the model will tremble when it rotates. We classify the overall motion trend into five phases by linear filtering. Finally, a mapping function is used to convert the hips spacing value into the angle of human body turning.

Occlusion may appear and give rise to failure as far as tracking results are concerned. In order to reduce the artifacts, the posture of character model is calibrated by IK. Given the original motion of human, we consider poses with 5 IK joints, which are reflected in different regions (the waist, hands and feet). Due to the difficulty of capturing at body turning process, we preset a turn-around animation and retarget to the target model based on the angle of body turning eventually, so that the turning motion of target model become more realistic.

2.2 Motion Retargeting Based on Quaternions

Our method of motion retargeting can be divided into three tasks: (1) matching the joint points into corresponding skeleton segments, (2) the calculation of quaternion based on the skeleton segments between current and previous frame, and (3) the deformation and optimization of virtual model.

Connecting the bone joints of virtual character is the first step of our method. We specify the parent node for each joint and connect them to form the skeleton segments. Because the number of joints and the structure of skeleton are vary from different models, we define a data dictionary that stores the skeleton structure. After defining the skeleton dictionary, the joints of target character are connected to the bones, through which the movement of target character is operated. For the input motion data, we also need to connect the human joints captured by the RGB-D camera, according to the same hierarchical relationship.

In each frame of the captured motion sequence, we extract the direction vector and angular amplitude of bones. The bones of human will perform a rotation around its base joint. The rotation motion of skeleton segment can be expressed as a quaternion. We calculate the quaternions of all the bones between human body and target model. Quaternion method only transfers the rotation motions of bones. Without the translate motions of joint position, it may contain the artifacts such as foot floating. Thus we define an offset for each joint based on the coordinate of root joint.

In order to make the generated animation more coherent and smooth, the rotation angles of two adjacent frames are optimized by interpolation method.

3 Conclusion and Future Work

In this paper, we design a real-time system for producing virtual character animation using single depth camera. Compared with existing methods, we can directly deal with the motion data captured by Kinect, as well as we improve the performance when occlusion take place. In Fig. 2, we apply captured motion to four different targets, which fully demonstrates the robustness of our system. Our framework is capable of handling complex human motions, like turning around and occlusion. The success of our framework is hinging upon the following key factors: (1) After the reconstruction of bones, the quaternion-based method for motion retargeting is helpful for reducing vibration and noises; (2) The adaptive filter is capable of estimating the rotation of the human body and recover the motion when turning around; and (3) The motion refinement adjusts the posture of the entire body through the IK technique which enhances the performance in side view.

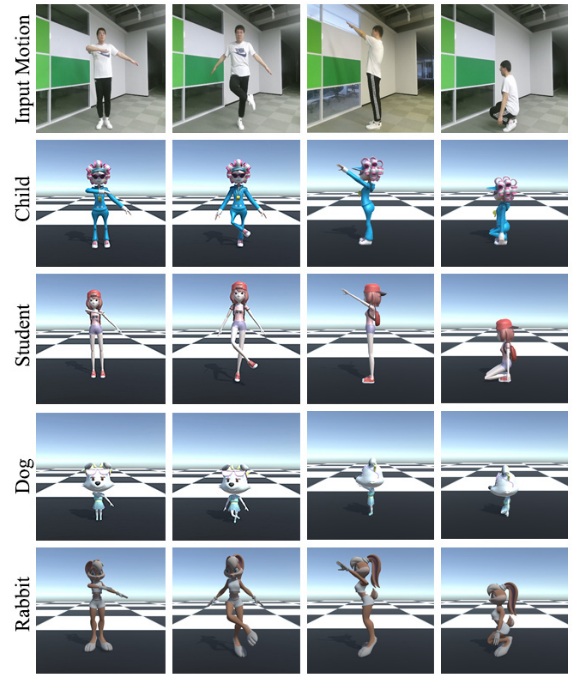

Figure 2: The animation results of 4 different virtual characters.

Meanwhile, our approach also has some limitations. Our method can only recognize turns from -90 to 90 degrees. For 360-degree turning, our method is not stable. The main reason is that our filter can not accurately decide the direction of turning when the camera gets a side view. We hope to improve our filter to increase the prediction. Another problem is that after inletting IK constraints into our system, the animation looks unnatural for some cases. In the future, we attempt to use a data-driven method to tackle the problem. At the same time, immersive device like Head Mounted Display (HMD) is considered to be applied to our system.

Acknowledgments

This research has been supported by National Key R and D Program of China (Grant No. 2018YFC0115100) and National Natural Science Foundation of China (Grant Nos. 61532002, 61672149, 61872020).

References

[1] Kinect - windows app development. `https://developer.microsoft.com/en-us/windows/kinect`.

[2] M. Abdulmassih, I. Yoo, and B. Benes. Motion style retargeting to characters with different morphologies. *Computer Graphics Forum*, 36(6):86–99, 2017.

[3] G. Baciu and B. K. C. Iu. Motion retargeting in the presence of topological variations. *Computer Animation and Virtual Worlds*, 17(1):41–57, 2006.

[4] J. Monzani, P. Baerlocher, R. Boulic, and D. Thalmann. Using an intermediate skeleton and inverse kinematics for motion retargeting. *Comput. Graph. Forum*, 19(3):11–19, 2000.

[5] R. Villegas, J. Yang, D. Ceylan, and H. Lee. Neural kinematic networks for unsupervised motion retargetting. *CoRR*, abs/1804.05653, 2018.