Frontiers in Art Research
ISSN 2618-1568 Vol. 5, Issue 16: 73-77, DOI: 10.25236/FAR.2023.051613

Virtual Shadow Puppet Play Generation Based on Alphapose

Zhengyu Lin[1,a,#], Kaiyi Tong[1,b,#], Tianyi Yang[1,c,#]

[1]Beijing 21st Century International School, Haidian, Beijing, China
[a]lzy1278497552@outlook.com, [b]2443622426@qq.com, [c]yangtianyi510@163.com
[#]Co-first author

Abstract: *As a precious art in China, shadow play has been spreading for more than 2000 years. But with the development of society, shadow play is no longer popular and attractive to the younger generation. Therefore, digital shadow play plays a very important role in the protection and dissemination of shadow play. In spite of this, there are still some deficiencies in this way of spreading shadow play culture today. It is usually used by placing the shadow puppet model in a designated position in a pre-built place and making adjustments. Of course, this method takes a lot of time, and compared with traditional shadow play, this kind of shadow play is not flexible and vivid. This paper proposes to use motion capture technology to make digital shadow play. The method adopted in this paper is based on the open source code of AlphaPose, Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time. First of all, the action of the characters is captured by the live action and high-definition camera, and finally the shadow puppet animation is generated according to the captured action.*

Keywords: *Motion capture, Virtual reality, Traditional culture*

1. Introduction

Shadow puppet play is a traditional drama form widely known in China. As the representative of Chinese culture, shadow play contains great cultural value and artistic value. In 2011, Chinese shadow play was selected into the Representative list of the intangible cultural Heritage of Humanity. However, with the development of society and the advent of the information age, science and technology have changed people's way of life and behavior. With the emergence of more and more new forms of entertainment, traditional arts such as shadow puppetry have gradually withdrawn from people's view due to the complex production process and long performance practice time. Especially the younger generation, they know but don't understand aboutshadow puppet play. The conservation and dissemination of shadow play is facing challenges.The development of digital media, especially the development of virtual reality technology and human-computer interaction technology, provides a new idea for the future development of traditional shadow play. The combination of modern digital technology and shadow puppetry presents shadow puppetry in a virtual orm, which not only retains the artistic characteristics of shadow puppetry, but also makes the creation of shadow puppetry easier, provides more means of communication, and injects new vitality into shadow puppetry.So we used Alphapose to bring shadow play into the virtual world(1) Realizing the real-time Pose Estimation and Tracking of actors, and establishing the joint model by The open source code of Alphapose is used to . (2)Culculating the rotation Angle of the joint model.(3) Generating shadow puppet play animation based on the above data.

2. Related work

In this section, we review the literature on multi-person pose estimation and motion capture, as well as provide background knowledge for our production of a virtual shadow theater.

Gkiox [1] USES k-poselet to detect a person in a cooperative manner and predicts the position of a person from the weighted average of all active poselets. Pishchulin et al. [2] first tested all parts of the entire car body and then used deepcut's integral linear programming method to mark, filter, and assemble these parts. Insafutdinov et al. [3] propose a stronger partial detector called DeeperCut [4] and a better incremental optimization policy on a resnet-based basis. In Openpose [5], we introduced

PAF (Part Affinity Fields) and decomposed the fit problem into two parts of the fit problem to solve the fit problem. Newell et al. [6] learns for each detected part an identification tag that indicates which individual it belongs to. This is called related embedding. Cheng et al. [7] USES a powerful multiresolution network [8] as the backbone and USES a high-resolution characteristic pyramid to learn scale-sense representations.Simple structured networks with resnet[4,9] as the backbone and a few deconvolution layers as the top sampling head [10] show efficient and competitive results. sunetal[11] constructs a powerful high-resolution network. In the first stage, high-resolution subnetworks are constructed, and in the subsequent stage, high- resolution to low-resolution subnetworks are added one by one in parallel, and multiscale feature fusion is repeated. "Cue.? Bertasius et al. [12] proposed a method to extend from images to dynamic images and learn the arc of posture on sparsely labeled dynamic images. Jin et al. [13] proposed zone, which uses roialign to trim hand and face areas on the feature map and predict key points on the resizing feature map. yang et al. [14] predicted the current pose based on the past pose sequence and merged it with the pose estimation results of the current frame. In particular, [15] it is recommended to use the re-id feature to solve the tracking problem. To address this issue, this approach explicitly employs human re-id feature quantities.Liu et al. [16] established the local correspondence between optical motion data and kinect motion data, and established the relationship with Gaussian process. The pose is reconstructed by taking direct input from the local gp model. Mousas et al. [17] extended the formal structure of the hidden Markov model (HMM) to enable motion datasets containing human dance movements to learn HMM. The system predicts dance movements based on the user's current location, allows the user to dance with a virtual character wearing a motion-capture combo, and can make a head screen (HMD). Constraint-based approaches allow transformations to be implemented through a set of geometric or physical constraints. Restrictions, such as joint angles, the position of the finished effect or dynamic Settings [18-19]. [20] Salbin method calculates the natural posture of constrained optimization problems[20]. Lv et al. [21] Reference to pose data and dynamics data led to the formalization of reverse dynamics. Motion was successfully reconstructed using pre-recorded motion data. In this case, the ik restriction applies. Parallel resolution of multiple joints.

3. Virtual shadow puppet play generation base on Alphapose

In this part,we will talk about the process we make the virtual shadow puppet. There are three parts to this process.(1)Use alphapose to capture motion and calculate the turn angle data(2)Use spine to generate a Bone binding model(3)Drive the model and generate the animation (*Figure 1)*.

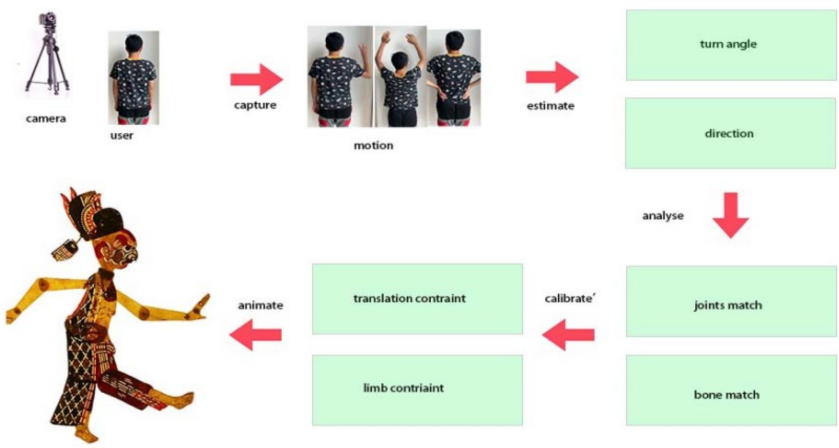

Figure 1: Overall flow chart

3.1. Motion capture base on alphapose

Alpha Posture is a system that allows whole-body posture estimation and tracking of movement in real-time. Alpha contains several new techniques. Symmetric Integral Keypoint Regression (SIKR) facilitates fast and accurate localization, and Parametric Posture Non-Maximum Suppression (P-NMS) reduces unnecessary human detection and posture-aware identity embedding, which helps in posture estimation and tracking. The human function can be the same person identified in the environment.

ISSN 2618-1568 Vol. 5, Issue 16: 73-77, DOI: 10.25236/FAR.2023.051613

Extracted-id features for each bounding box generated by the object detector. MSIM algorithm is used to store the previous detection and tracking results in the pool Pl. The whole inference process is divided into 5 stages to deal with the pipeline. Each module contains a separate process that stores the results in a queue, and subsequent modules can get the results directly from the queue. This allows the whole process to run at high speed. The Fast Pose skeleton is used to make the estimation more accurate and efficient (*Figure 2)*.

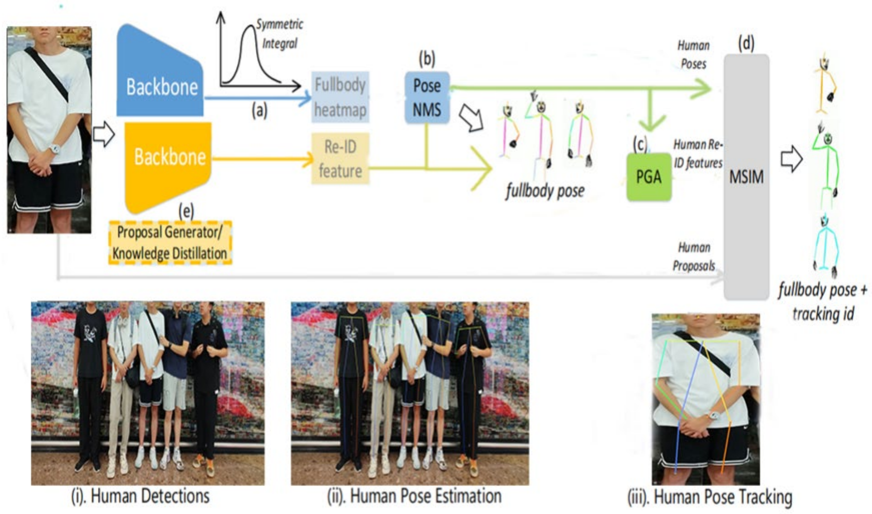

Figure 2: Motion capture flow chart

3.2. Calculate angle of model

The tool I use to calculate the angle is a Python function that uses the numpy library to calculate the angle between three 2D points. This function first calculates two vectors (BA and BC), then uses the dot product and magnitude of these two vectors to calculate the angle between them.

This function works based on the relationship between the dot product of vectors and the magnitude (or length) of the vector. The dot product formula is:

$$A*B=|A||B|\cos(\theta) \tag{1}$$

where θ is the angle between the two vectors rearranging, we get:

$$\cos(\theta)=A*B/|A||B| \tag{2}$$

Then, use the np.arccos function to calculate the angle.

In the example you provided, the coordinates of points A, B and C are defined and the calculate_angle function is used to calculate the angle between these three points. Finally, output the angle.

3.3. Make virtual model of shadow puppet play base on spine

Making animation is a complicated thing. In this project, we need two shadow puppet model, so I choosed Mokey Sun and Sun quan as shadow puppet model. At first, I cut out the limbs of each model through PS. Then I converted the PSD file format to JSON format with a script called PS to Spine. After that I opened this model in the Spine, and I created skeleton for each limb of the two models, and bind each bone to every part of the model*(Figure 3)*.

ISSN 2618-1568 Vol. 5, Issue 16: 73-77, DOI: 10.25236/FAR.2023.051613

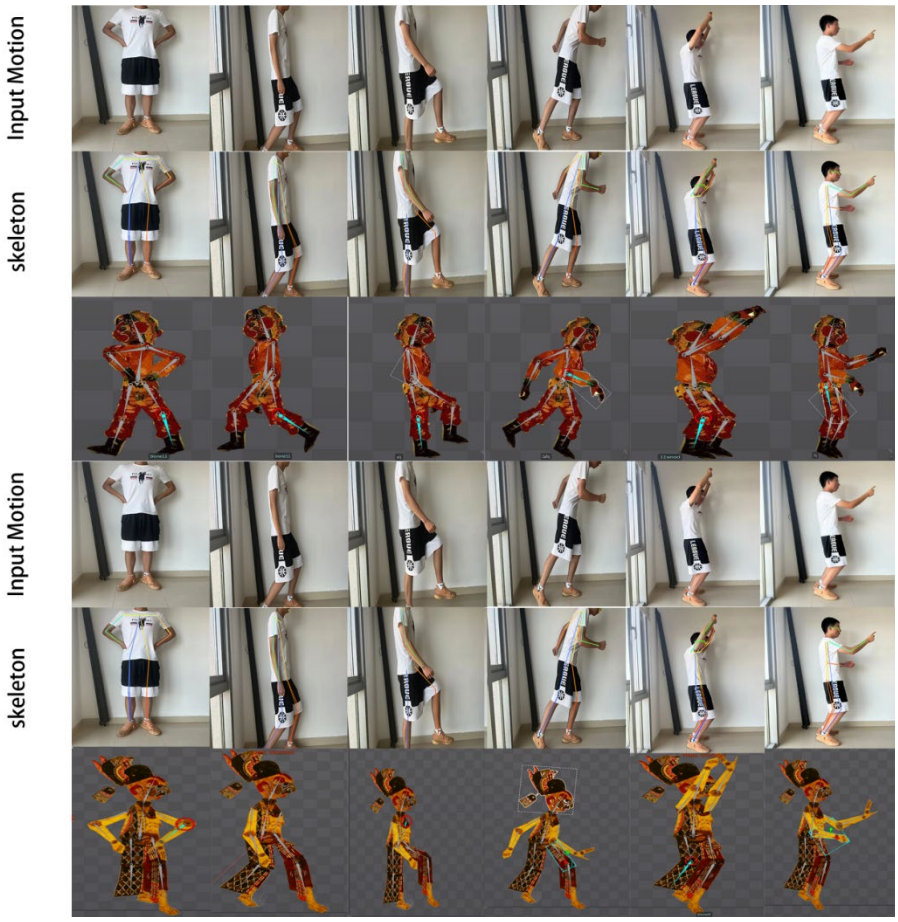

Figure 3: Output result

4. Conclusion

In this paper, we propose to use motion capture technology to make virtual shadow puppets. In addition, we also introduced relevant production methods. Based on the Alphapose system, we can quickly and accurately capture the movements of characters. We used spine to make the skeleton of shadow puppets, and then calculated the motion Angle based on the motion capture data to activate the motion of shadow puppets and realize the animation effect.

References

[1] Gkioxari, G., Hariharan, B., Girshick, R., & Malik, J. (2014). Using k-poselets for detecting people and localizing their keypoints. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3582-3589).
[2] Pishchulin, L., Insafutdinov, E., Tang, S., Andres, B., Andriluka, M., Gehler, P. V., & Schiele, B. (2016). Deepcut: Joint subset partition and labeling for multi person pose estimation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4929-4937).
[3] Insafutdinov, E., Pishchulin, L., Andres, B., Andriluka, M., & Schiele, B. (2016). Deepercut: A deeper, stronger, and faster multi-person pose estimation model. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI 14 (pp. 34-50). Springer International Publishing.
[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
[5] Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7291-7299).
[6] Newell, A., Huang, Z., & Deng, J. (2017). Associative embedding: End-to-end learning for joint detection and grouping. Advances in neural information processing systems, 30.

ISSN 2618-1568 Vol. 5, Issue 16: 73-77, DOI: 10.25236/FAR.2023.051613

[7] Roth, D., Lugrin, J. L., Büser, J., Bente, G., Fuhrmann, A., & Latoschik, M. E. (2016, March). A simplified inverse kinematic approach for embodied VR applications. In 2016 IEEE Virtual Reality (VR) (pp. 275-276). IEEE.

[8] Villegas, R., Yang, J., Ceylan, D., & Lee, H. (2018). Neural kinematic networks for unsupervised motion retargetting. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 8639-8648).

[9] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu, "Rmpe: Regional multiperson pose estimation," in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2334–2343.

[10] Xiao, B., Wu, H., & Wei, Y. (2018). Simple baselines for human pose estimation and tracking. In Proceedings of the European conference on computer vision (ECCV) (pp. 466-481).

[11] Sun, K., Xiao, B., Liu, D., & Wang, J. (2019). Deep high-resolution representation learning for human pose estimation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 5693-5703).

[12] Bertasius, G., Feichtenhofer, C., Tran, D., Shi, J., & Torresani, L. (2019). Learning temporal pose estimation from sparsely-labeled videos. Advances in neural information processing systems, 32.

[13] Jin, S., Xu, L., Xu, J., Wang, C., Liu, W., Qian, C., ... & Luo, P. (2020). Whole-body human pose estimation in the wild. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16 (pp. 196-214). Springer International Publishing.

[14] Yang, Y., Ren, Z., Li, H., Zhou, C., Wang, X., & Hua, G. (2021). Learning dynamics via graph neural networks for human pose estimation and tracking. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8074-8084).

[15] Zhang, J., Zhu, Z., Zou, W., Li, P., Li, Y., Su, H., & Huang, G. (2019). Fastpose: Towards real-time pose estimation and tracking via scale-normalized multi-task networks. arXiv preprint arXiv:1908.05593.

[16] Liu, Z., Zhou, L., Leung, H., & Shum, H. P. (2015). Kinect posture reconstruction based on a local mixture of gaussian process models. IEEE transactions on visualization and computer graphics, 22(11), 2437-2450.

[17] Mousas, C. (2018, March). Performance-driven dance motion control of a virtual partner character. In 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (pp. 57-64). IEEE.

[18] Fang, A. C., & Pollard, N. S. (2003). Efficient synthesis of physically valid human motion. Acm transactions on graphics (tog), 22(3), 417-426.

[19] Lee, J., Chai, J., Reitsma, P. S., Hodgins, J. K., & Pollard, N. S. (2002, July). Interactive control of avatars animated with human motion data. In Proceedings of the 29th annual conference on Computer graphics and interactive techniques (pp. 491-500).

[20] Grochow, K., Martin, S. L., Hertzmann, A., & Popović, Z. (2004). Style-based inverse kinematics. In ACM SIGGRAPH 2004 Papers (pp. 522-531).

[21] Lv, X., Chai, J., & Xia, S. (2016). Data-driven inverse dynamics for human motion. ACM Transactions on Graphics (TOG), 35(6), 1-12.