Deep Reinforcement Learning for Modeling Human Locomotion

Xue Wu, *Computer Science,the University of Alabama,*
Shengting Cao, *Electrical and Computer Engineering,the University of Alabama,*

***Abstract*—It is difficult to measure and interpret the activity of the millions neurons involved in human activity related to motor control. A neuromechanical simulation framework has been introduced to produce the correct motions of a musculoskeletal model. Current control model could simulate and explain many basic locomotion behaviors like walking, running and climbing stairs. But they can't explain some locomotion of diseases such as stroke patients with unbalanced walking gait or just because of old age with unbalanced walking gait. Reinforcement learning has been rarely applied in simulations to model human control for rehabilitation training to avoid disease in different ages, especially with support of wearable health platform. We would like to use utilize the mHealth (mobile health) with the neuromechanical simulation to investigate the cue for unblanced gait phases. We propose to use reinforcement learning to predict muscleskeleton motion of different gait type from camera and wearable sensor.**

***Index Terms*—Reinforcement Learning, Rehabilitation, Human Locomotion, Wearable Technology, neuromechanical simulation.**

I. INTRODUCTION

PRedictive neuromechanical simulations can produce the motion without directly use the experimental motion data[1]. When the predictive motion data match the human experimental data, such predictive simulations can be used to accelerate research in rehabilitation or psychical training. Although we can model such a motion with musculoskeletal modeling and physics engine, understanding and modeling human motor control is still difficult. In particular, it is difficult to measure and interpret the biological neural circuits that underlie the human motor.

Recently there are two simulation environments: Mujoco[2] and OpenSim[3]. These two platforms will represent different state in complex controlled environment with a definition of a policy of reinforcement learning. For setting up Mujoco environment, a camera may need to get real-time physical locomotion and collect real physical measured sensor data from wearable devices. Right now all the sensor data will be collected by a fossil LTE watch. All the simulation will be performed by Linux or MacOS with installation of Conda and Tensorflow.

II. RELATED WORK

A. OpenSim

OpenSim-RL[3] is a an open source software to create and analyze dynamic simulation of physical movements. It is a good platform to develop model for biomechanical simulations that can be tested, improved, analyzed and exchanged with musculoskeletal reinforcement learning environment.

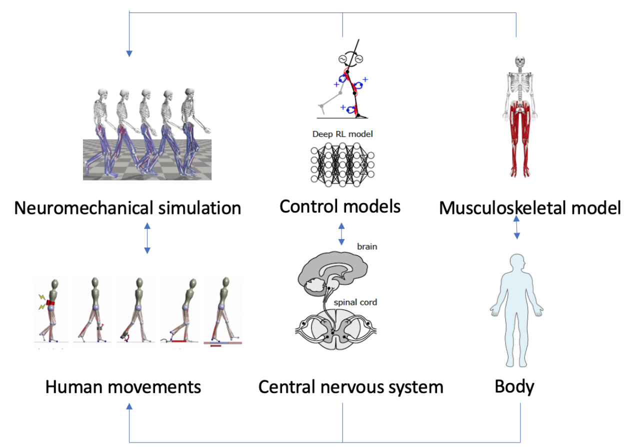

Fig. 1. Neuromechanical simulation

B. MuJoCo

Gym package provide a platform for reinforcement learning. Mujoco is one of the package, it is a new physical engine tailored to model-based control. New efficient algorithm could be used for connnection user model and real physical pose. In order to alleviate optimal control applications differently and particularly, different states and controls could be evaluated for dynamics parallely. Around 400,000 dynamics evaluations per second are possible on a 12-core machine, for a 3D homanoid with 18 dofs and 6 ativate contacts[2].

III. METHOD

We plan to use reinforcement learning to predict muscle-skeleton motion of different age's human gait type from video and wearable sensor. A novel approach could also be used to model human motion with physics simulation and use imitation learning to learn a video-conditioned control policy for ego-pose estimation[4]. There are some simulations at open source gym package and we plan to see if our model could imitate movements from real physic motion.

First, we should will build a reinforcement learning framework to predict the walking of the skeleton and collect simulated data, then we will collect real human motion data

Xue Wu, Computer Science, The University of Alabama, Tuscaloosa, USA xwu33@crimson.ua.edu.

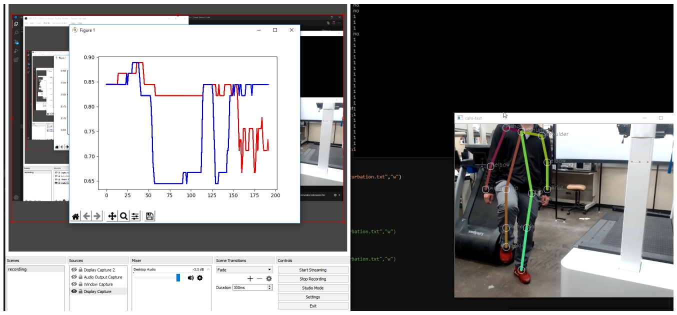

Fig. 2. place holder

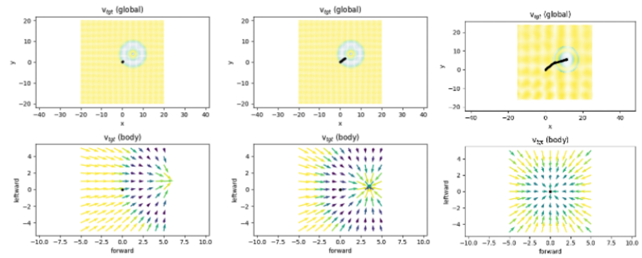

Fig. 3. Movement ink and direction

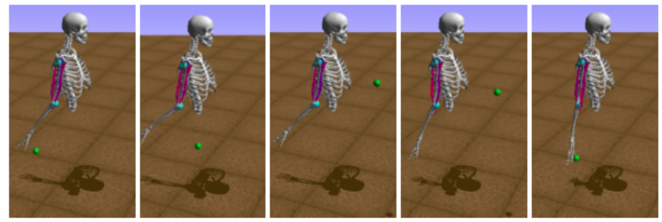

Fig. 4. Arm movements with a ball

using wearable sensor. Then I will plot the two different kind of data and calculate the mean square error.

A. *Simulation*

Some experiments could be conducted on age-related changes in human locomotion. 20-39 is young age, 50-64 is middle age and more than 65 is old age. Muscle reflexes will be different based on different ages. Data could be collected by testing different groups of people, for examples three dimensions data like accelerometer, gyroscope and magnetometer. Then we could figure out how these data will represent the status of the person and how to distinguish data when the person in different situation. Experiments will help us understand more about human rehabilitation and how to improve policy to train our model. The musculoskeletal is the agent and action is the movement of the musculoskeletal, the agent takes a reward and observation as input and trains a policy that outputs an action to achieve high cumulative rewards.

The proposed policy that mimics human motion capture data for rehabilitation training. A novel approach could also be used to model human motion with physics simulation and use imitation learning to learn a video-conditioned control policy for ego-pose estimation[4]. There are some simulations at open source gym package and we plan to see if our model could imitate movements from improved policy. And use a based policy learned via reinforcement leaning to predict and forecast 3D human pose[5] form camera and devices.

B. *Visualization*

We have a cluster based on CentOS Linux and laptop based on elementary linux system, we used laptop to get a visualization of the muscleskeleton walking, arm moving and leg moving right now.

1) RL for muscleskeleton walking: This is a ”Learn to Move Competition” that I run to let muscleskeleton walk and plan to use it to train with reinforcement learning and predict motion of different age with camera and wearable devices with support of Mujoco as well. Figures 7 shows the start position of walking. Figure 6 shows the ink of the motion trail and how direction changes over time. Figure 5 shows the walking phases that defined in physical therapy. A gait cycle starts from heel strike and end with terminal swing to show each phase transaction.

2) Arm moving: After training the arm of the muscleskeleton, right arm will tap the small ball to a position and move back to previous position. Figure 3 shows the one position of arm movements. In this platform, different joints will have a function that we could edit and control the movements. The animation we get is right arm could catch and throw the ball. We plan to see if there are other movements we could predict for dynamic gait type via reinforcement learning.

IV. Results

Experiments could be conducted on age-related changes in human locomotion. 20-39 is young age, 50-64 is middle age and more than 65 is old age. Muscle reflexes will be different based on different ages. Data could be collected by testing different groups of people, for examples three dimensions data like accelerometer, gyroscope and magnetometer. Then we could figure out how these data will represent the status of the person and how to distinguish data when the person in different situation. At beginning all the human movements are captured by camera and real human and then we find a way to simulate it and get information to analyze human movements by deep reinforcement learning.

Experiments will help me us understand more about human rehabilitation and how to improve policy to train our model. The muscleskeleton is the agent and action is the movement of the muscleskeleton, the agent takes a reward and observation as input and trains a policy that outputs an action to achieve high cumulative rewards.

In our experiment we use some figure to display our simulation. Figure 2 shows real human movements captured by camera, it imitates a person with stroke gait and different joints will also appear when capturing. We also simulate that by using a muscleskeleton with stroke, we pick some clips of our simulation experiment in gait phases figure 5 showed.

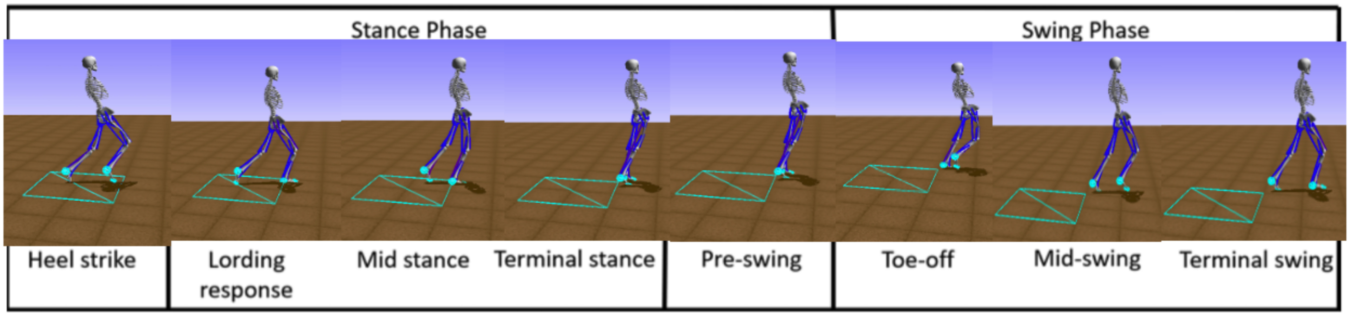

Fig. 5. original muscleskeleton

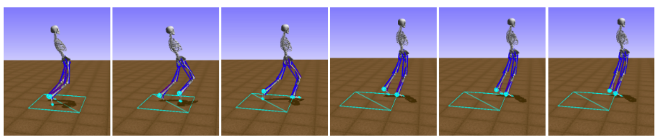

Fig. 6. Walking muscleskeleton

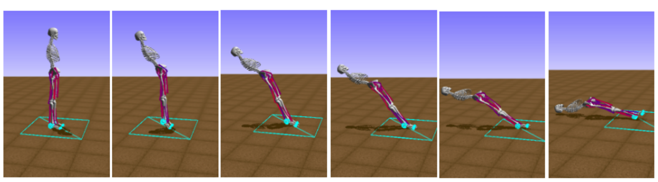

Fig. 7. original muscleskeleton

In this project we train the arm of the muscleskeleton and get the pre-trained walking muscleskeleton to imitate human with stroke. We could simulate different age of muscleskeleton will walk with different speeds.

For future we plan to simulate more different kinds of movements with this musclekeleton.

V. Conclusion

The opensim is a good framework to simulate the muscle skeleton of human body. Using the reinforcement learning method can simulate the human movement. We have build and trained a reinforcement learning network to make the muscleskeleton to simulate the human walking cycle.

Due to the time limitation, we are not able to finish the age analysis part for the project before deadline but we have pretty good results in terns of visualization .

Appendix B

References

[1] Seungmoon Song, Lukasz Kidziński, Xue Bin Peng, Carmichael Ong, Jennifer Hicks, Sergey Levine, Christopher G. Atkeson, and Scott L. Delp. Deep reinforcement learning for modeling human locomotion control in neuromechanical simulation. *bioRxiv*, 2020.

[2] Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. *IEEE International Conference on Intelligent Robots and Systems*, pages 5026–5033, 2012.

[3] Scott L. Delp, Frank C. Anderson, Allison S. Arnold, Peter Loan, Ayman Habib, Chand T. John, Eran Guendelman, and Darryl G. Thelen. OpenSim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering*, 54(11):1940–1950, 2007.

[4] Ye Yuan and Kris Kitani. 3D Ego-Pose Estimation via Imitation Learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11220 LNCS:763–778, 2018.

[5] Ye Yuan and Kris Kitani. Ego-pose estimation and forecasting as real-time PD control. *Proceedings of the IEEE International Conference on Computer Vision*, 2019-October(Iccv):10081–10091, 2019.