Reinforcement Learning Framework for Active Muscle Control of the Head and Neck

Sayak Mukherjee, Daniel Perez-Rapela, Jason Forman, Matthew B. Panzer

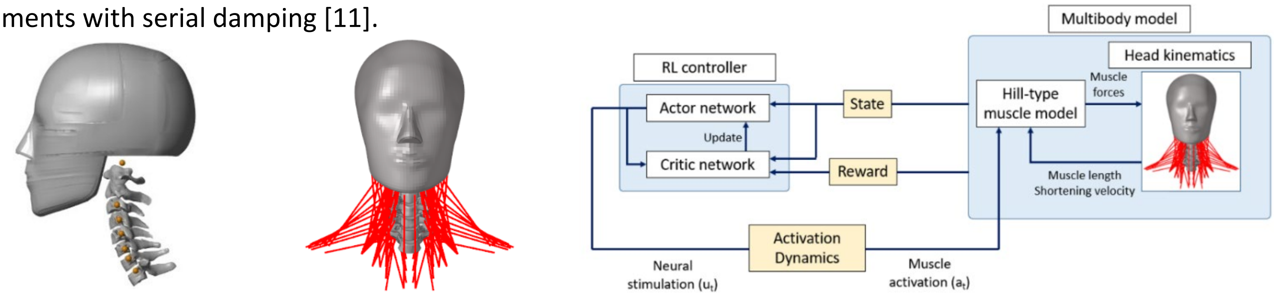
I. INTRODUCTION

The head-neck region is a highly redundant mechanical structure with several non-linear joints and complex muscle orientation. Constant stabilisation of the head under gravity and voluntary movements of the head require a coordination of activated neck muscles that is not easily understood [1]. Due to the complex anatomy of the head-neck region, it is difficult to associate muscles corresponding to specific head movements [2-5]. The coordinated activations of the neck muscles for voluntary head movements are difficult to replicate in human body models (HBMs), with recent muscle control studies being limited to just stabilisation of the head under gravity and external loads [6-8]. The goal of the present study was to overcome these limitations using a muscle control framework based on reinforcement learning (RL) to synthesise fast, goal-directed head rotations in the three anatomical planes. The RL muscle actuation control (RLMAC) framework was previously used to stabilise the head under external perturbations [9] and to generate extension-flexion motion at the elbow joint [10]. In both the control studies using RLMAC, the muscles were activated independently without any prior assumption on agonist and antagonist grouping. In this study, the RLMAC was extended to generate voluntary kinematics using a multibody (MB) model of the head and neck region with the anatomy of a 50thpercentile male.

II. METHODS

Multibody model of the head and neck region

Multibody model of the head and neck region was developed in Matlab and used previously in muscle control studies [9]. The MB model consists of rigid vertebrae (C1–T1) with the head graphics represented as Hybrid-III head form (Fig. 1(a)). Forty-six neck muscles were incorporated into the model (Fig. 1(b)) as HIll-type muscle elements with serial damping [11].



(a) *(b)* *(c)*

Fig. 1. (a) Head-neck model with location of vertebral joints, (b) neck muscles included in the head-neck model, (c) RLMAC framework for neck postural control.

Muscle control framework

Twin-delayed deep deterministic policy gradient (TD3) [12] is used as the agent for RLMAC (Fig. 1(c)) in the present study. The TD3 agent uses three neural networks – one actor and two critics. The actor network maps states to control outputs (actions) whereas the critic networks evaluate the actions based on a reward function. The states of the control environment were defined for both head in space stability (vestibulocollic reflex or VCR) and head on trunk stability (cervicocollic reflex or CCR). The VCR parameters include the translation and rotation head kinematics – position and velocity. The spinal joint kinematics corresponding to the muscle spindles and joint proprioceptors were added to the state as CCR. The reward function (Eqn. 1) penalised the RLMAC proportional to the orientation error (error (ε) between the target angle and current angle). The reward function also penalised the agent for linear and rotational head velocities to dampen any disturbance to the head at the target position.

S. Mukherjee is a PhD Student, J. Forman is a Principal Scientist and M. Panzer (e-mail: panzer@virginia.edu; tel: +1 434.297.8062) is an Associate Professor at the Center for Applied Biomechanics in the Department of Mechanical and Aerospace Engineering, University of Virginia, USA. D. Perez-Rapela is a research engineer at Insurance Institute for Highway Safety / Highway Loss Data Institute, USA.

Internal joint energy was used as the cost function to stabilize the spine. A symmetry term was included in the reward function to enable the RLMAC to encourage a symmetric activation policy for the left and right muscles.

$$\text{Reward} = (1 - \text{Sin}(2\varepsilon_X))^2 + (1 - \text{Sin}(2\varepsilon_y))^2 + (1 - \text{Sin}(2\varepsilon_Z))^2 - \alpha\,(V_x + V_y + V_z) - \beta\,(\omega_x + \omega_y + \omega_z) - \eta \sum \text{Joint energy} + \mu\,(\text{symmetric factor}) \quad (1)$$

where α, β, η, and μ are positive coefficients of the components of the reward function for training.

The RLMAC outputs 46 independent actions representing the neural stimulation for each muscle. No assumptions were made regarding the grouping of muscles as agonist or antagonist for a specific axis. The actor network consists of an input layer, two hidden layers, and the final layer. Input layer and hidden layer outputs were activated with the rectified linear unit (ReLU). The output of the final layer was bound between 0 and 1 using a tanh layer followed by a scaling layer. The critic networks have identical architecture with two hidden layers between the input and the final layer. The output of the critic network is the Q-value, a scalar value that determines the efficacy of the action for the given state.

The RLMAC was trained to generate goal-directed rotations of the head in multiple anatomical planes. In each training iteration, model was stabilised in the neutral position (0^o angle) and then the model was asked to move the head in flexion, extension, axial rotation, or lateral bending (selected randomly) to a target angle (also selected randomly) while keeping the angle in the other planes at 0^o. Gravity was applied throughout the simulation and the T1 vertebrae was fixed. Each iteration was simulated for 500 ms (taking 10–15 minutes each) and RLMAC took 15,000 iterations to converge.

III. INITIAL FINDINGS

The trained agent was able to maintain the neutral posture of the head under gravity, as well as generate omnidirectional goal-directed head rotations. The RLMAC was able to stabilise the head in the first 400 ms and synthesise the targeted rotations when the target signal was applied as a step function (Fig. 2). The RLMAC was also able to carry the desired head rotations with minimum out-of-plane motion of the head. Head rotation under every prescribed target signal used the same trained RLMAC to achieve the desired head position.

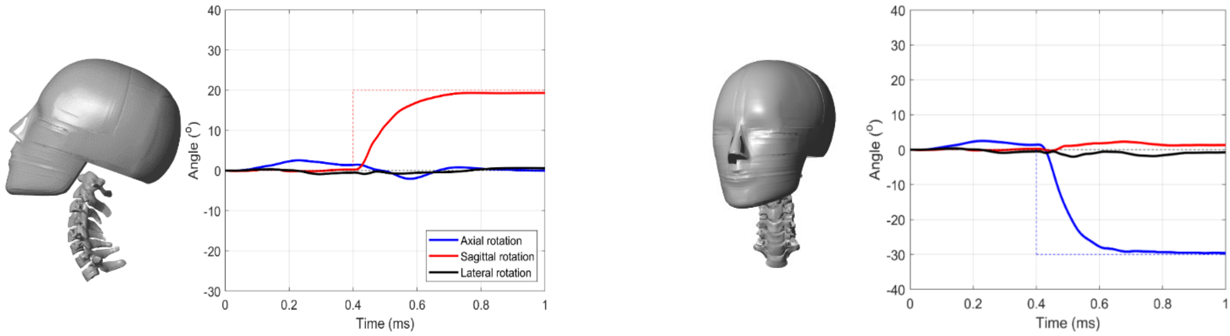


Fig. 2. Examples of test cases showing the final posture and response of the RLMAC for (a) 20^o Extension and (b) 30^o right axial. The dashed line indicates the prescribed target head angle for goal directed motion.

IV. DISCUSSION

The present study was an attempt to generate voluntary head kinematics in HBMs by activating the neck muscles using RL algorithms and preliminary results have been presented. While previous control studies were limited to stabilising the head, the RLMAC could synthesise the goal-directed head rotations along different individual anatomical planes. Future studies will focus on validating the response of the RLMAC using volunteer data sources as well as evaluating the applicability of the model for voluntary motion along multiple planes at once as well as under injurious loading environments.

V. REFERENCES

[1] Crowninshield, *et al.*, *J Biomech*, 1981.
[2] Lee, *et al.*, ASME, 1990.
[3] Vasavada, *et al.*, *J Biomech*, 2008.
[4] Fice, *et al.*, *J Neurophysiol*, 2018.
[5] Driess, *et al.*, *IEEE ICRA*, 2018.
[6] Happee, *et al.*, *J Biomech*, 2017.
[7] Zheng, *et al.*, *IEEE Neural Syst Rehabil Eng*, 2021.
[8] Correia, *et al.*, *Ann Biomed Eng*, 2021.
[9] Mukherjee, *et al.*, *IRCOBI*, 2021.
[10] Mukherjee, *et al.*, *J Biomech*, 2022.
[11] Haeufle, *et al.*, *J Biomech*, 2014.
[12] Fujimoto, *et al.*, *Int conf on ML*, 2018.