

OstrichRL: A Musculoskeletal Ostrich Simulation to Study Bio-mechanical Locomotion

Vittorio La Barbera^{*1} Fabio Pardo^{*2} Yuval Tassa³ Monica A. Daley⁴
 Christopher T. Richards¹ Petar Kormushev² John R. Hutchinson¹

¹Royal Veterinary College, London, ²Imperial College London

³DeepMind, London, ⁴University of California, Irvine

{vlabarbera, jhutchinson, ctrichards}@rvc.ac.uk,
 {f.pardo, p.kormushev}@imperial.ac.uk,
 tassa@deepmind.com, madaley@uci.edu

Abstract

Muscle-actuated control is a research topic that spans multiple domains, including biomechanics, neuroscience, reinforcement learning, robotics, and graphics. This type of control is particularly challenging as bodies are often overactuated and dynamics are delayed and non-linear. It is however a very well tested and tuned actuation mechanism that has undergone millions of years of evolution with interesting properties exploiting passive forces and efficient energy storage of muscle-tendon units. To facilitate research on muscle-actuated simulation, we release a 3D musculoskeletal simulation of an ostrich based on the MuJoCo physics engine. The ostrich is one of the fastest bipeds on earth and therefore makes an excellent model for studying muscle-actuated bipedal locomotion. The model is based on CT scans and dissections used to collect actual muscle data, such as insertion sites, lengths, and pennation angles. Along with this model, we also provide a set of reinforcement learning tasks, including reference motion tracking, running, and neck control, used to infer muscle actuation patterns. The reference motion data is based on motion capture clips of various behaviors that we preprocessed and adapted to our model. This paper describes how the model was built and iteratively improved using the tasks. We also evaluate the accuracy of the muscle actuation patterns by comparing them to experimentally collected electromyographic data from locomoting birds. The results demonstrate the need for rich reward signals or regularization techniques to constrain muscle excitations and produce realistic movements. Overall, we believe that this work can provide a useful bridge between fields of research interested in muscle actuation.

1 Introduction

In nature, vertebrates produce movement by the contraction of skeletal muscles that pull on the bones, creating torque at the joints. Understanding muscles is of interest in many different fields. First, from a biological point of view, it is desirable to understand how animals use their bodies to perform various behaviors. Muscles are also relevant in computer graphics to obtain accurate skin deformations in virtual characters and to produce more natural-looking gaits for films and video games (Angles et al., 2019; Abdrashitov et al., 2021; Modi et al., 2021). In sports medicine, musculoskeletal simulations

^{*}Equal contribution. The models, tasks and data are available at <https://github.com/vittorione94/ostrichrl> and visualizations at <https://sites.google.com/view/ostrichrl>.

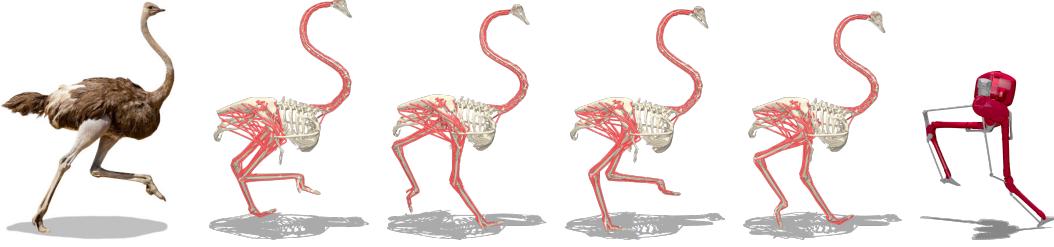


Figure 1: The proposed musculoskeletal model is based on data collected from real ostriches, while the Cassie robot provides a robotic approximation of the morphology of an ostrich.

can be used to understand sports injuries (Bulat et al., 2019) or create effective training and recovery strategies (Coste et al., 2017). Unfortunately, in robotics, with the exception of soft robots, muscles have been studied relatively little despite their interesting energetic properties and the sophisticated movements they enable (Wang et al., 2021; Cotton et al., 2012).

Biomechanics is the study of how biological systems produce motion. Biological motion is affected by many factors such as the musculoskeletal geometry, internal states of the muscles, and force-length-velocity curves. Understanding how complex biological systems such as humans and other animals use their muscles to move is quite challenging. Obtaining data from living animals is difficult and not neutral to animal welfare. To overcome this problem, biomechanics researchers often use computer-based simulations combined with numerical optimization to estimate how muscles are used. However, most of the research that uses optimization techniques does not leverage recent progress made in reinforcement learning.

In this work, we propose a new ostrich musculoskeletal model, illustrated in Figure 1, that uses a fast physics engine with reinforcement learning tasks to infer muscle actuation patterns. Ostriches are interesting to study because they are fast, economical bipedal runners (Alexander et al., 1979). To the best of our knowledge, we are the first to apply RL to such a realistic animal model, solving locomotion tasks. With the proposed open-source computational tools, we hope to open new opportunities for researchers interested in accurate and fast muscle simulation provided by MuJoCo combined with RL.

2 Related work

2.1 Building musculoskeletal models



The NeurIPS conference has been hosting recurrent competitions to bridge the gap between biomechanics and reinforcement learning², where the goal was to make a musculoskeletal human model walk. The challenge used the OpenSim (Delp et al., 2007; Seth et al., 2018) simulator. However, as documented in many of the solutions (Pavlov et al., 2018; Zhou et al., 2019; Akimov, 2019; Kolesnikov & Khrulkov, 2020), the engine was too slow to properly use traditional deep RL methods, which often require millions or billions of samples to find good solutions. Moreover, the model used in the latest challenge was in many ways quite simplistic, actuating only the legs with 22 muscles in total and removing the arms.

Simulating musculoskeletal animal models is not common outside the field of biomechanics. Most of the efforts are focused on humans because of available resources in anatomical atlases and interest from the entertainment industry and sports. Available full-body animal musculoskeletal models include, for example the chimpanzee (Sellers et al., 2013) and dog (Stark et al., 2021). Moreover, for human models, researchers tend to have access to much higher quality motion capture data than for animals, but animal research often have superior empirical measurements of physiological function.

²<https://www.crowdai.org/challenges/nips-2017-learning-to-run>

<https://www.crowdai.org/challenges/nips-2018-ai-for-prosthetics-challenge>

<https://www.aicrowd.com/challenges/neurips-2019-learning-to-move-walk-around>

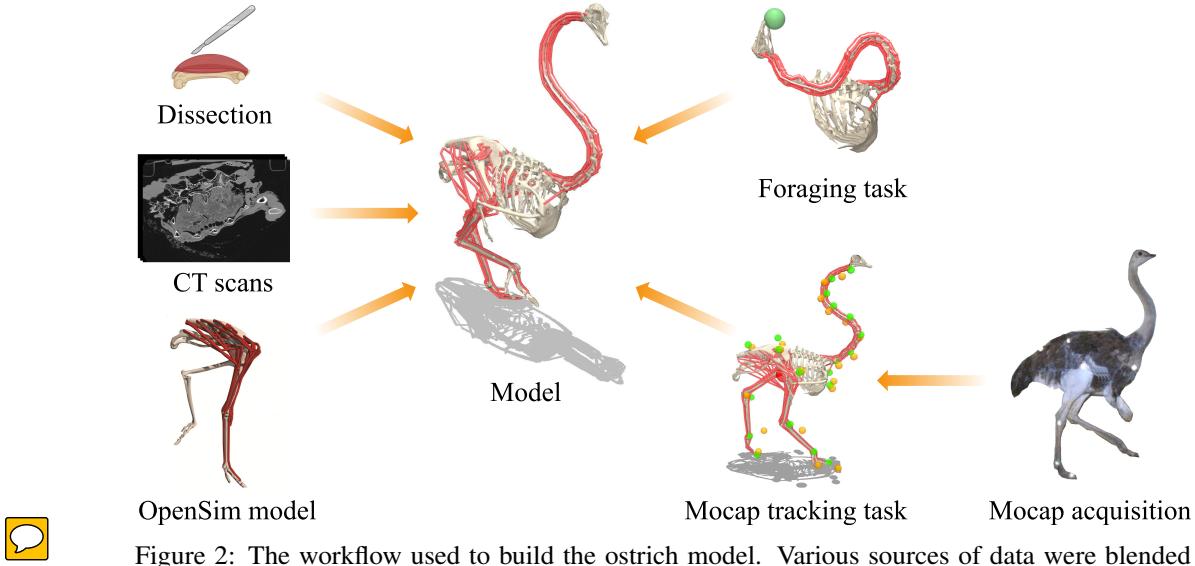


Figure 2: The workflow used to build the ostrich model. Various sources of data were blended together for the modeling part, and tasks helped in fine-tuning muscle strength and joint limits.

2.2 Controlling musculoskeletal models

Some innovative research has used deep RL to solve complex locomotion tasks such as playing basketball (Liu & Hodgins, 2018), performing athletic jumps (Yin et al., 2021), or boxing and fencing (Won et al., 2021). These studies used motion capture data to build a rich reward signal during training (Merel et al., 2017, 2018; Hasenclever et al., 2020). Wang et al. (2012) was one of the first studies in the graphics community to use biological actuators. Geijtenbeek et al. (2013) used evolution-based algorithms to control muscular bipeds and also optimized muscle routing. Jiang et al. (2019) tried to bridge torque-actuated models and muscle dynamics, using neural networks to map muscle activations to torque commands and achieved natural looking gaits. Lee et al. (2018) used volumetric muscles and trajectory optimization for juggling. Lee et al. (2019) combined RL and motion capture clips to control a human musculoskeletal model solving a variety of tasks from walking to weight lifting, separating muscle coordination from trajectory mimicking through two different networks with privileged information and an intermediate proportional derivative (PD) controller.

3 Methods

This paper describes how a fast and realistic musculoskeletal model of an ostrich was built by combining several sources of information, including an existing OpenSim model of legs, a dissection, a CT scan, and reinforcement learning tasks that allowed us to iteratively improve properties of the model. This workflow is represented in Figure 2.

3.1 Anatomical data acquisition

Computed tomography scanning We first acquired multiple computed tomography (CT) scans of a 96.5 kg adult male ostrich specimen, donated by a local farm. The CT scans were then segmented (separating the bone from the rest of the x-ray slices) using the Mimics software (Materialise, Inc.) to obtain 3D models (polygonal meshes) for each bone from the ribcage to the head, including the wings. These 3D geometries were important for two reasons: firstly, they provided an accurate representation of the geometry of the bones, vital for defining muscle attachment points, and secondly, they allowed inference of the inertia of body segments. To obtain the inertia of body segments we followed the steps provided in Hutchinson et al. (2007). The idea is to wrap the skeleton parts within meshes representing flesh and assume constant density. From there, volumes, masses, and the inertia tensors were computed.

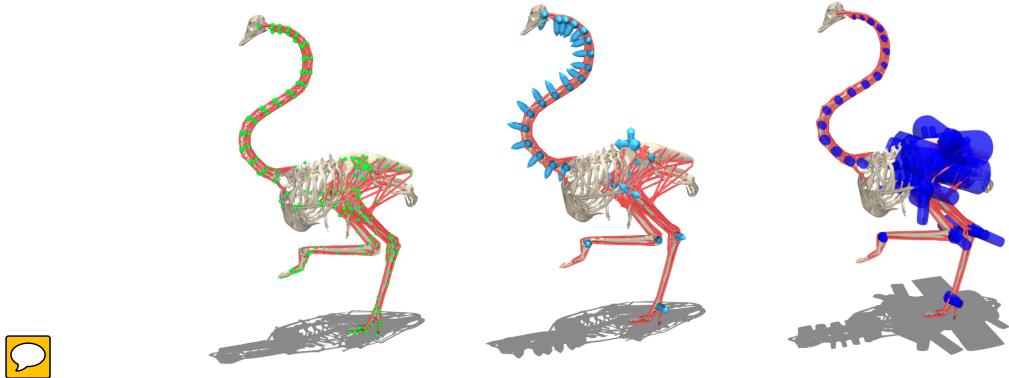


Figure 3: The skeleton of the model follows a tree structure of bodies and joints. The muscle routing is defined using waypoints, shown in green on the left. The joints are shown in light blue in the middle. The wrapping geometries are shown in blue on the right (some wrapping geometries have been omitted for clarity).



Muscle dissection Next, we performed a dissection in order to gather data for muscles of the neck and rib cage. Wings are not actuated in our model because they are mostly used when turning at high velocity and air friction is not considered in our simulations. Dissection is important to properly understand muscle routing, where the muscles start (origin) and where they end (insertion). During the dissection, we used anatomical descriptions from Böhmer et al. (2019), Tsuiiji (2005), and Tsuiiji (2007) to identify the different muscles. We also acquired muscle-specific data: muscle mass, pennation angle, tendon length, and muscle fiber lengths, used when simulating the muscle-tendon dynamics.

3.2 Modeling

OpenSim and MuJoCo Two computer modeling and simulation software packages are commonly used by the biomechanics and reinforcement learning communities respectively. The first one is OpenSim (Delp et al., 2007; Seth et al., 2018), an open source engine that provides benchmarked physics-based muscle simulations. It is relevant to our study not only for its popularity, but also because the legs of our ostrich model were originally modeled with this engine (Hutchinson et al., 2015). The same model was also used to run some tracking (inverse) simulations to estimate the forces, activations, and mechanical work of the muscles involved in solving locomotion tasks (Rankin et al., 2016). OpenSim is rather slow in simulating each time step, as documented in previous NeurIPS competitions that used it to simulate muscle dynamics (Kolesnikov & Khrulkov, 2020). The second simulator, MuJoCo (Todorov et al., 2012), has recently been open-sourced, and is used in multiple RL domains, because it provides fast and accurate rigid body dynamics. Specifically, tendon routing in OpenSim uses an iterative algorithm, while MuJoCo’s tendon routing is closed-form and, therefore, much faster. In a comparison made by Ikkala & Hämäläinen (2020), when calculating the average run time over 97 forward simulations, MuJoCo was roughly 600 times faster than OpenSim.



Musculoskeletal models Both OpenSim and MuJoCo define their models in a hierarchical tree structure. Muscles are attached to at least two bones, using two sites at the extremities, called the origin and the insertion sites. Other sites, called waypoints, and wrapping geometries, can also be specified to define elaborate muscle routings. Waypoints are sites through which the muscle must pass, which are useful to maintain an anatomically realistic 3D path for the muscles. Wrapping geometries are useful to prevent muscles from penetrating the bones, and are geometric primitives such as spheres or cylinders that muscles must wrap.



Proposed model After converting the OpenSim leg model to MuJoCo’s format, we completed the ostrich model by adding the ribcage, wings, neck, and head geometries. We then actuated the neck by adding the muscles, matching the routing from the dissection. We also modified the morphology of the feet, which were initially flat, to make their shape more realistic. We found that fine tuning the muscle lengths was necessary to improve stability. To find the new muscle ranges, we randomized

the model joint angles in their limits a large number of times and recorded the global minimum and maximum lengths for each muscle. In our final model, visible in [Figure 3](#), each leg contains 7 hinge joints: 3 for the hip, 2 for the knee, 1 for the ankle, and 1 for the metatarsophalangeal (mtp), while the neck contains 36 hinge joints, 2 for each pair of vertebrae and for the head. These 50 joints are actuated by 120 muscles, 68 in the legs and 52 in the neck, making this model very challenging to control. In comparison, the OpenSim model used for the NeurIPS 2019: Learn to Move - Walk Around challenge had only 22 muscles.

Simulation for RL MuJoCo has already been used in a number of popular control domains for reinforcement learning research, in particular dm_control ([Tassa et al., 2020](#)) and OpenAI Gym ([Brockman et al., 2016](#)). Most of these domains use simplistic models, often inspired by animal morphologies, made up of basic geometric bodies such as capsules, and torque-controlled multiaxial joints. While such models are fast and fairly easy to control, they do not provide a realistic basis for studying animal movement. On the other hand, simulations used in biomechanics typically model animals more accurately ([Sellers et al., 2013; Hutchinson et al., 2015; Stark et al., 2021](#)) but are slow and often used in conjunction with relatively simple control techniques from the trajectory optimisation repertoire, such as direct collocation.

3.3 Muscle dynamics

Muscle excitation and activation In biomechanics and neuroscience, researchers separate muscle excitation and activation. A neural excitation u , produced by the nervous system, is responsible for the contraction of muscle fibers via an intermediate state called activation a ([Zajac, 1989](#)). This intermediate state converts an electrochemical signal to mechanical force output. In MuJoCo, this conversion is modelled as a first-order nonlinear filter:

$$\frac{\partial a}{\partial t} = \frac{(u - a)}{\tau(u, a)}$$

where τ is defined as:

$$\tau(u, a) = \begin{cases} \tau_a(0.5 + 1.5a) & u > a \\ \tau_d/(0.5 + 1.5a) & u \leq a \end{cases}$$

τ_a , τ_d are time constants for activation and deactivation, equal by default to 10 ms and 40 ms, respectively.

Contraction dynamics The contraction dynamics are implemented using the Hill-type model as shown in figure [Figure 4](#). This mechanical model computes the total force generated by a muscle by adding together the active contractile and passive properties of muscles. The contractile element models the muscle contraction force, scaled by the activation signal a . The passive properties are modeled like springs: if a muscle is stretched, it will generate a passive force to return to its rest position.

The active contraction dynamics can be summarized by the formula:

$$f(l, \dot{l}, a) = a \cdot f_l(l) \cdot f_v(\dot{l}) + f_p(l)$$

where f_l is the active force as a function of muscle length, f_v is the active force as a function of velocity, and f_p is the passive force which is always present regardless of activation. The force-length and force-velocity curves are showed in [Figure 4](#). These functions are built into MuJoCo, capturing the contraction dynamics for basic purposes, and have been validated by the biomechanics community for some conditions ([Millard et al., 2013](#)).

All muscle-related quantities in MuJoCo are scaled by the muscle-tendon actuator's resting length. The advantage of this representation is that all muscles behave similarly. MuJoCo does not allow specification of the muscle resting length L_0 and tendon slack length LT directly, in contrast to OpenSim. This is due to the fact that MuJoCo does not include a stateful elastic component in the muscle model. This is a major compromise in MuJoCo when compared to OpenSim, that allows it to run faster. Instead, the actuator length range LR needs to be specified, which is the minimum and maximum of the sum of muscle and tendon lengths, along with a range R in units of L_0 . In other words, the actuator length range LR defines the interval of values that the actuator is allowed to use during the simulation and the range R is a percentage expressing how much the actuator (here simply called "muscle") can shorten or extend.

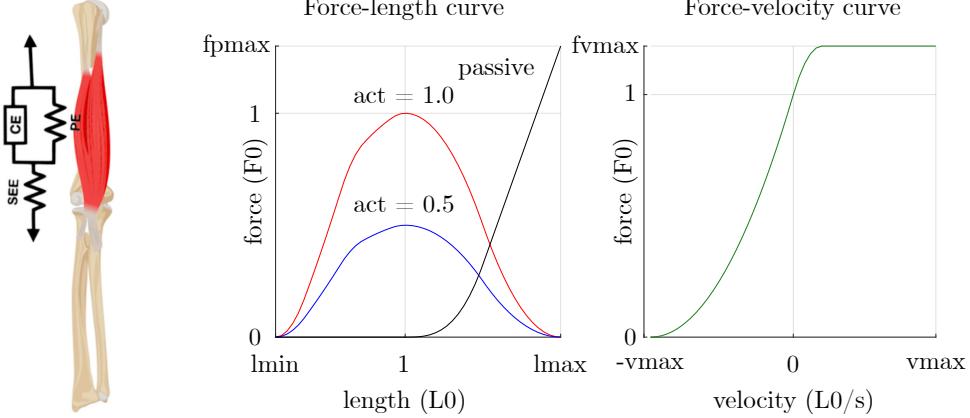


Figure 4: On the left a diagram of the Hill-type muscle model comprising the 3 components: a contractile element, a passive element and an elastic element. On the right two plots showing the curves for the force-length and force-velocity functions used in the contraction dynamics.

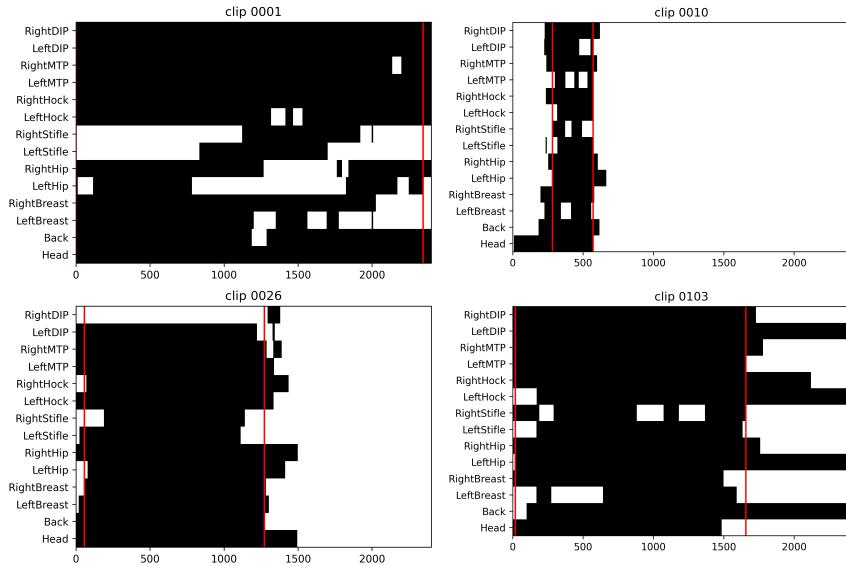


Figure 5: Binary masks showing, in black, the available data for the 14 markers, on 4 clips. The red lines indicate the intervals we used. The remaining missing markers had to be predicted.

At model compile time, MuJoCo solves these two equations where the unknowns are the muscle rest length L_0 and tendon length LT :

$$(LR_{min} - LT)/L_0 = R_{min}$$

$$(LR_{max} - LT)/L_0 = R_{max}$$

3.4 Motion capture data

Original mocap data We did not find any open-source motion capture database of ostrich movements, but after contacting the authors of Jindrich et al. (2007), we obtained the data originally used to study joint kinematics of ostriches performing cutting maneuvers. The provided dataset was composed of 82 clips, recorded at 240 Hz, containing the 3D coordinates of 14 markers distributed as follows: 1 the head, 1 on the spine, and 1 on each breast, hip, knee, ankle, mtp, and toe.

Cleaning of the mocap clips We selected the largest part of the clips where at least 10 markers were simultaneously present for at least 1 second, typically disregarding the beginning and the end of

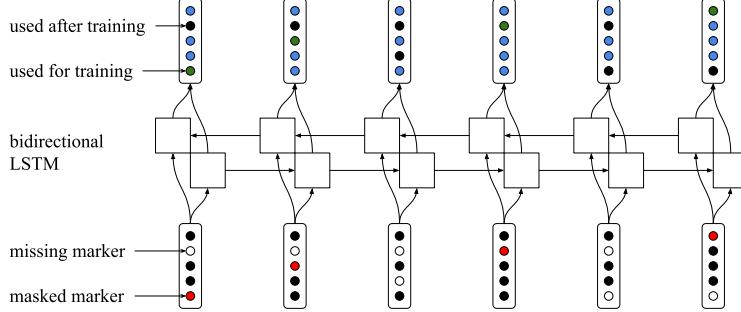


Figure 6: Illustration of the bidirectional LSTM used to predict the missing markers. The sequence of masked coordinates are encoded, then a sequence of decoded coordinates are produced, including predictions for the masked ones.



Figure 7: Visualization of the mocap prediction process. Segments are drawn between pairs of existing markers. Left: The original incomplete data. Right: The complete data after predicting the missing markers with a bidirectional LSTM.

the clips when the ostrich was not in the field of view of the cameras. This selection left 35 clips. However, even in those clips, we found that a large number of markers were periodically missing, probably due to feathers and limbs occluding them from the cameras, as can be seen in Figure 5. To overcome this issue, we used a bidirectional LSTM trained in a similar way as a denoising autoencoder, as shown in Figure 6. With a probability of 0.1, we randomly masked some markers and asked the model to predict their value given the context of a segment of 100 steps from the same clip. We then used this model to predict the actual missing markers. Since the model was trained and then used on the same, relatively small dataset, we anticipated some overfitting issues, but playing the clips with the predicted markers in place of the missing ones gave surprisingly good results, as shown in Figure 7.

Conversion to joint values After rescaling the marker coordinates to match the size of our model, the next step was to transform the set of marker coordinates to joint positions. Fortunately, enough markers were present in the different parts of the ostrich body to limit the space of solutions, except for the neck, where only one marker was present on the head. We used gradient descent with the mean Euclidean distance between the reference markers and those produced by joint poses on two sets of parameters simultaneously. The first set described where to attach the markers on the body parts and is shared across all the clips. This was needed because we did not have the exact location of the markers used during the motion capture acquisition and how they would translate to locations with respect to the bones. The second set of parameters described the joint values at every step. To reduce the space of solutions for the neck shape, we added a regularization term, encouraging the neck to follow an S shape. We first tried to use a finite difference approximation of the gradient, but found this approach to be too slow to be practical. We then decided to implement the kinematics function in a differentiable way, mapping joint values to body locations to create marker locations. This function was implemented using differentiable operations from the TensorFlow library (Abadi et al., 2016). We found this approach to be particularly fast thanks to the possibility to simultaneously optimise for all the steps across all the clips using the batch dimension. Finally, the velocity of the joints was inferred using the rate of change between two consecutive steps.

Cyclic clip We also created a cyclic running clip that can be used as a reference for running over an extended period of time. We used the middle section of one of the clips, which we repeated several times, and smoothed to remove discrepancies at the boundaries. This clip helped us produce the

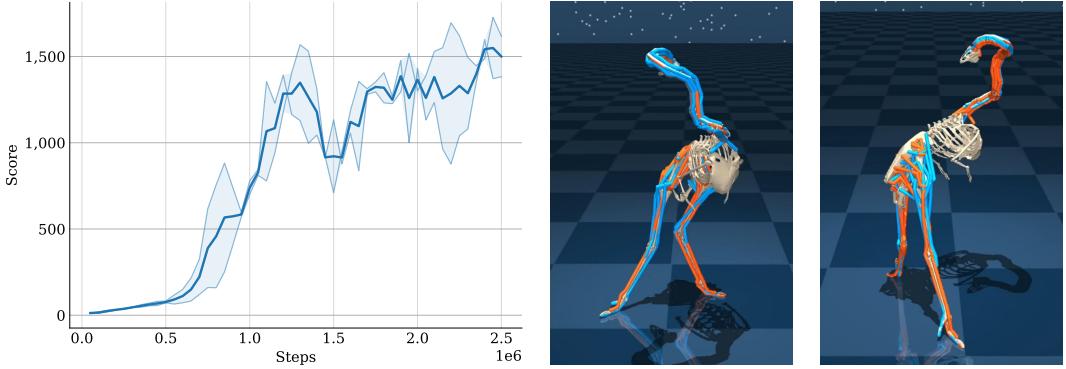


Figure 8: Performance on the “run forward” task. Sums of rewards are averaged over 2 seeds and smoothed with a sliding window of size 3. The agent manages to achieve satisfactory performance, but the gaits do not look natural as shown in the videos and images. Muscle excitations from -1 to 1 are represented by colors ranging from blue to orange respectively.

simulated biomechanical data for a cyclic gait that we compared to muscle excitations and lengths measured on emus in subsection 4.5.

4 Experiments

The proposed model was iteratively improved using a set of reinforcement learning tasks. The performance and visual quality of the solutions allowed us to select values for some muscle parameters such as muscle length ranges and forces, or to identify issues with unrealistic joint ranges. We focused on two types of tasks: locomotion and neck control. We found these to be particularly useful when designing the model, because they used the two main groups of muscles present in the model.

4.1 Running forward

The first task we experimented with had a simple “move forward” objective, rewarding agents for running as fast as possible. It was initially constrained to a vertical planar space, it used torque-driven joints, and the neck was kept rigidly attached to the model’s thorax. The goal was to ensure that the underlying model and simulation structure functioned as intended before adding muscles. The resulting gait was similar to that obtained when training on the 2D walker tasks of dm_control and OpenAI Gym. After adding muscles to the legs, we found that the same gait repeatedly emerged: the toes were unrealistically bent backwards (plantarflexed) and the ankles remained straight. After a series of model adjustments on the different tasks, we were able to achieve more diverse and efficient policies on this one, even after relaxing the model constraint from 2D to full 3D space and actuating the neck. The final “running forward” task is described below, and the results can be seen in Figure 8.

Initialization The model is initialized in an upright pose with small random perturbations added to the leg and neck joints. Concretely, for these joints, the initial values are sampled uniformly in 1/5th of their respective intervals, centered around a default pose.

Observations The agent perceives the height of the head, pelvis, and feet, the joint positions except for the x-axis, the joint velocities, the muscle forces, activations, lengths and velocities, and the forward velocity of the center of mass on the x-axis.

Reward and termination The reward is simply the forward velocity of the center of mass on the x-axis. Episodes end if the head height falls below 0.9 meters, if the pelvis height falls below 0.6 meters, or if the torso rotation around the y-axis falls below -0.8 or exceeds 0.8 radians.



Figure 9: Examples of motion capture tracking taken from two different clips. Left: A locomotion clip. Right: A complex clip requiring to stand still, perform neck movements, and turn in place. The reference is shown in white next to the model.

4.2 Motion capture tracking

The previous “run forward” task showed that learning to control such a complex articulated model with a simple reward function does not sufficiently regularize the search space, producing policies that are both unnatural and suboptimal. A typical solution to better constrain the policy space is to design a more sophisticated reward function. Reference motion tracking is a particularly well-suited option that has been used with motion capture data in many studies to produce natural-looking motions (Liu et al., 2010; Peng et al., 2018; Chentanez et al., 2018; Merel et al., 2018; Bergamin et al., 2019; Peng et al., 2019; Hasenclever et al., 2020). As a proof of concept, we started with reference motion data that we labeled by hand, frame by frame, using videos of running ostriches found online. The task was initially planar and the model torque-driven. The quality of the results produced encouraged us to continue in this direction.

After producing a satisfactory dataset, described in subsection 3.4, we experimented with a number of variations for the task components. By performing large sweeps over the 35 clips and the cyclic one, we arrived at the mocap tracking task described below. Examples of tracking on two different clips is illustrated in Figure 9, the performance on every clip is shown in Figure 10 and some of the videos demonstrate the quality of the tracking. The result on short walking clips is very satisfactory. These are the curves that quickly converge. For the most challenging clips, including standing and abrupt changes in speed, learning takes more time, and it seems that longer training would be needed.

Initialization and step During training, an initial time-step is uniformly sampled over the length of the episode, ignoring the last 20 steps. The model is initialized in the pose corresponding to this step, and a small amount of Gaussian noise, with a scale of 0.02, is added to help diversify experiences. The velocity of the joints was set to the one of the reference without modification. Since the control frequency used in our experiments was 40 Hz, the task used every sixth data point of motion capture data (240 Hz) during tracking.

Observations The agent perceives the height of the pelvis and feet, the joint positions and velocities, the muscle forces, activations, lengths and velocities, and the time left in the clip to allow time-dependent policies and to deal with the finite horizon, as described in Pardo et al. (2018).

Reward and termination The tracking reward is defined as a product of Gaussian kernels $\exp(-w_p e_p) \times \exp(-w_r e_r)$ across all body parts. The quantity $e_p = \|\bar{p} - p\|$ measures the Euclidean distance between the coordinates of the center of mass p and the reference \bar{p} , while the quantity $e_r = \arccos((\text{tr}(\bar{R}R^T) - 1)/2)$ measures the angle of the difference rotation between the orientation matrix of inertia R and the corresponding reference \bar{R} . The weights w_p and w_r control the wideness of the Gaussians, accounting for the different magnitudes and importances. We used the values 0.2 and 0.1 respectively. The reward is bounded in (0, 1] and the multiplicative nature ensures that if one of the components is too far from the reference, the whole reward decays to 0. Episodes end when the reward falls below 0.01 or when the end of the clip is reached.

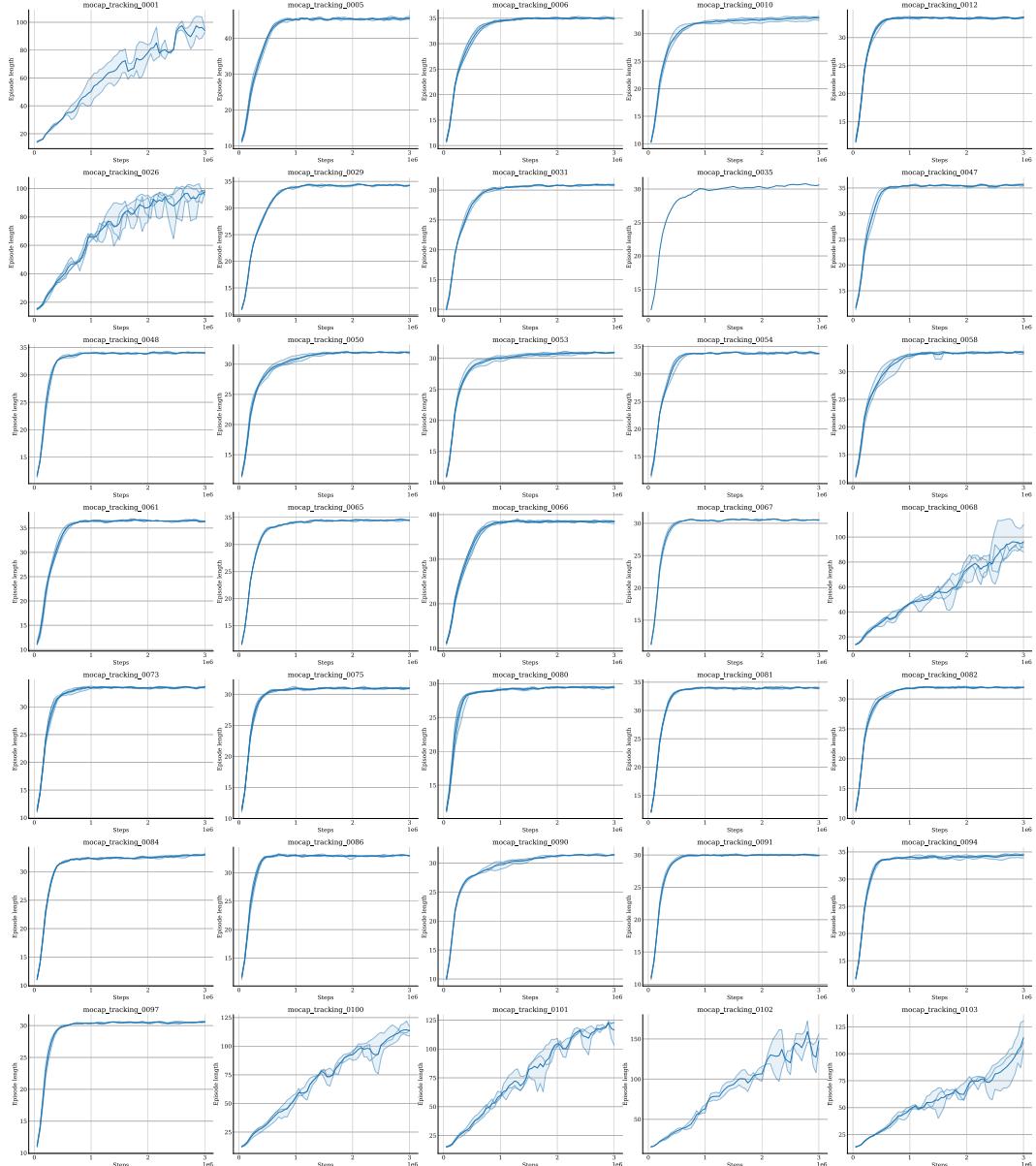


Figure 10: Average episode length for all the clips. As agents perform better at tracking, the episodes last longer. Values are averaged over 5 seeds and smoothed with a sliding window of size 3.

4.3 Neck control

The neck is composed of 17 vertebrae, which allows for a wide range of shapes. To achieve a particular head position and orientation, an infinite number of solutions exist, and during running, the weight and pose of the neck greatly influence balance. Controlling the neck with muscles that cross several vertebrae therefore requires a sophisticated policy.

To tune the neck portion of the model, we found that mocap tracking of the whole body was not ideal due to the large number of moving parts and the fact that the reference neck poses were artificial. Therefore, we created another task whose objective was to reach random targets with the beak. Only the rib cage, neck, and head were used and the rib cage was held fixed in space. Although effective policies were easily obtained, the initial neck shapes were very irregular, with many abrupt and unnatural changes. To reduce this phenomenon, we increased the stiffness of the neck joints until satisfactory smoothness was achieved. For the joint boundaries, we initially used the values of

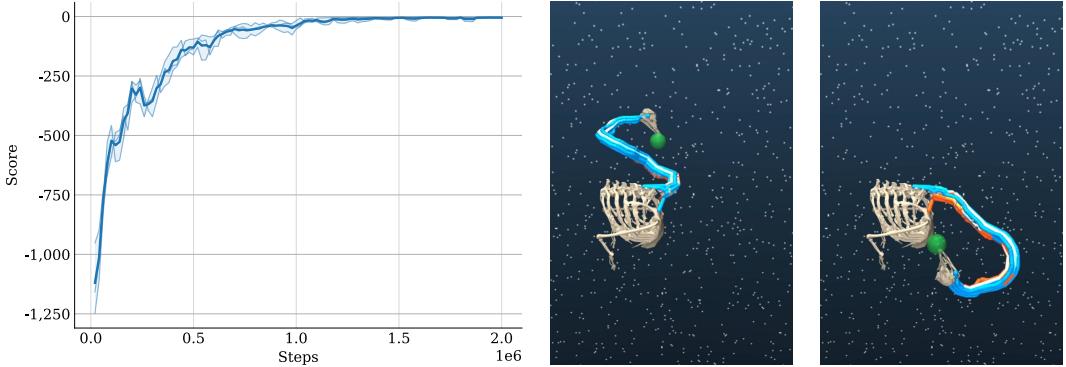


Figure 11: Performance on the “neck control” task. Sums of rewards are averaged over 3 seeds and smoothed with a sliding window of size 3. The agent manages to achieve satisfactory performance, and the neck shapes are fairly natural.

Dzemski & Christian (2007), but then realized that the boundaries needed to be increased to allow for tighter neck curves that ostriches can perform. The final “neck control” task is described below, and the results can be seen in Figure 11.

Initialization Every 100 episodes, and for the very first one, the neck is initialized with a default S shape. For all other episodes, the neck position and velocity are maintained from the last state in the previous episode. The target position is randomized to a feasible area. To do this, we use rejection sampling. Points are repeatedly sampled within a sphere centered at the base of the neck with a radius of 0.8 meters, slightly smaller than the length of the neck, and discarded when inside a second smaller sphere representing approximately the ostrich torso, with a radius of 0.6 meters.

Observations The agent perceives the joint positions and velocities, the muscle forces, activations, lengths and velocities, the coordinates of the beak and target, and the vector from the beak to the target.

Reward and termination The reward is simply the negative of the Euclidean distance between the beak and the target. When this distance falls below a threshold of 0.05 meters, episodes terminate.

4.4 Experimental details

To obtain the policies, we used a TD4 agent, proposed in Pardo (2020), a mixture of TD3 (Fujimoto et al., 2018), with its pair of critics, delayed actor training and target action noise, and D4PG (Barth-Maron et al., 2018), with its distributional value function (Bellemare et al., 2017) and n-step returns. We also found that Ornstein Uhlenbeck exploration noise (Lillicrap et al., 2015) was significantly better than Gaussian noise, probably due to the importance of correlated excitation when controlling muscles. We chose to use the Tonic deep RL library, because it provided us with a state of the art agent and the flexibility we needed to quickly try, evaluate and visualize experiments while allowing custom dm_control tasks to be used. The different tasks can be visualized on the website.

For the experiments reported above, we used the following hyper parameters. The models are simple MLPs with 2 hidden layers of size 256, ReLU activations, and 51 atoms for the distributional heads. The replay buffer has a size of 1,000,000, batches of size 50 are sampled after the first 50,000 steps and then 50 times every 50 steps. We use 1-step returns, a discount factor of 0.99, a learning rate of 0.0001, a target action noise scale of 0.25, an Ornstein-Uhlenbeck action noise with scale 0.25 and 10,000 initial random steps.

4.5 Electromyography comparison

To demonstrate the realism of the proposed simulation, we chose to compare the muscle excitations outputted by a policy trained on the cyclic mocap tracking task, to electromyography (EMG) data. On one side, EMG measures electrical activity in muscles, in response to a stimulation from motoneurons

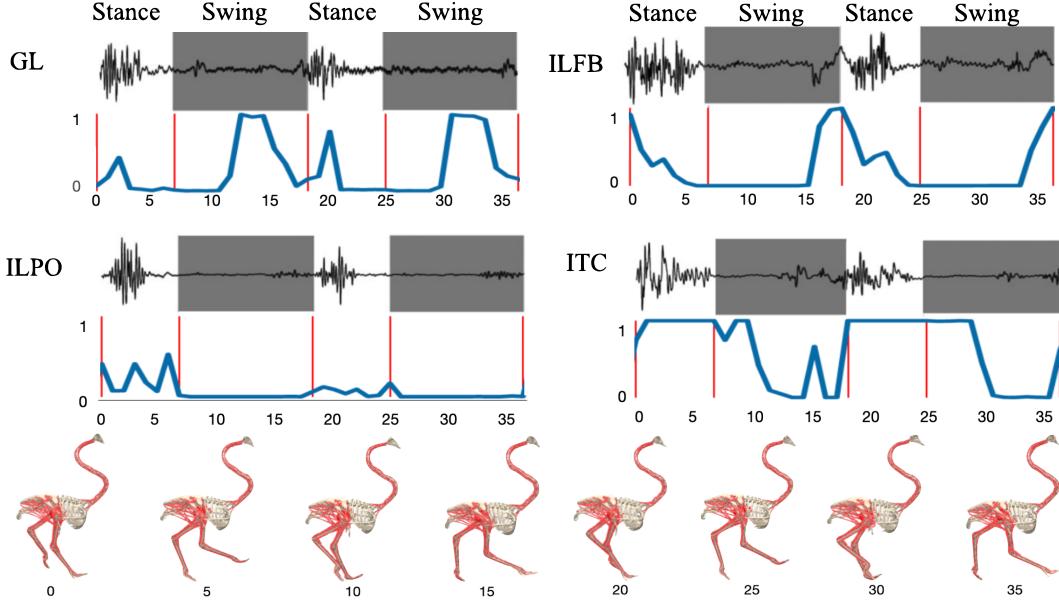


Figure 12: Comparison with experimental EMG data from emus. The first red vertical line indicates that the right foot was in contact with the ground (start of stance phase). The second one is when the foot leaves the ground (start of swing phase). EMG data from Cuff et al. (2019) are on top. Muscle excitations (ours) are below.

and on the other side, raw actions produced by a policy are used by MuJoCo to produce muscle activations, as explained in subsection 3.3.

Collecting EMG data in animals is challenging because their skin may be too thick for the electrodes and surgery may be required to implant them directly into the muscles. Because no ostrich EMG data have yet been collected, we decided to perform comparisons with data from emus, close relatives of ostriches. EMG data were originally acquired by Cuff et al. (2019) to study muscle recruitment patterns in different bird species and understand neuromuscular control from an evolutionary perspective. Their results showed that walking/running birds generally use their hind limb muscles in very similar ways, making it reasonable to compare emu EMG data with ostrich simulations.

To compare the data, following standard conventions in locomotion research, we divide the gait into two phases: stance (foot is on the ground) and swing (foot is off the ground). Once the gait has been segmented into these phases, we scale them accordingly between the two studies, to account for the difference in speed. As shown in Figure 12, we find broadly similar excitation patterns, except for the GL (gastrocnemius lateralis) which had an extra burst of swing phase excitation in these simulations, not usually found in any extant birds (Cuff et al., 2019). The ITC (iliotrochantericus caudalis) also had simulated excitation throughout more of the swing phase than in the EMG data, reminiscent of findings for simulated ostriches using the OpenSim model in Rankin et al. (2016). There are many more muscles than actual degrees of freedom in the legs, allowing many possible excitation patterns to match the kinematic data. In addition, because unnecessary excitations are not penalized, it is expected to see more excitations than in the experimental data.

4.6 Compatibility with Cassie

Cassie is a bipedal robot produced by Agility Robotics. It has undergone extensive work in trajectory optimization (Reher et al., 2019; Li et al., 2020; Apgar et al., 2018) and reinforcement learning (Xie et al., 2018, 2020; Li et al., 2021). Its morphology is similar to that of ostrich legs with short thighs, however, because Cassie’s joints are powered by electric motors, they are each limited to one degree of freedom. To create multiple degrees of freedom, several joints are added in series and an offset is necessary between them for mechanical reasons. This is a major difference between the ball-and-socket joints allowed by musculoskeletal bodies and traditional robots. In addition, Cassie couples the knee and ankle joints using rods. However, the similarity to the ostrich morphology

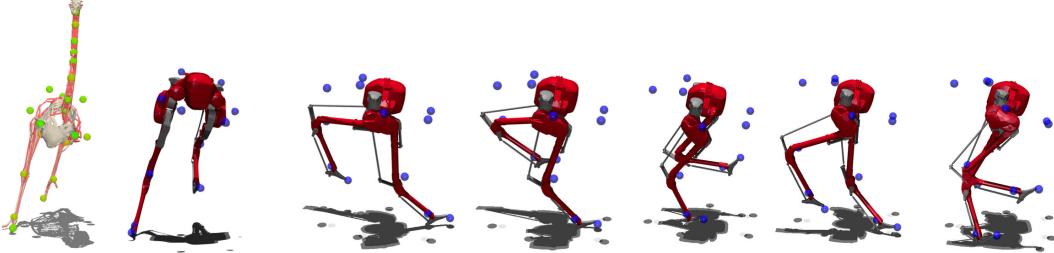


Figure 13: Ostrich and Cassie markers comparison, and poses from a motion capture clip. The morphology of the two models is similar enough to allow the motion capture data to be reused.

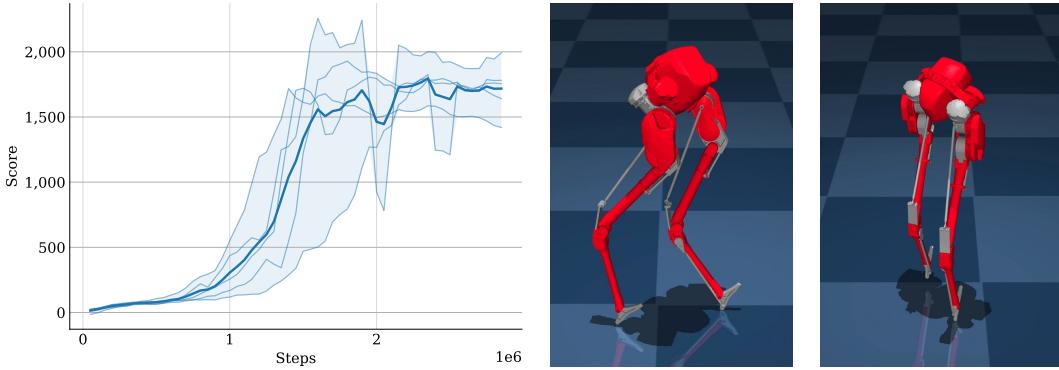


Figure 14: Performance on the “run forward” task with the Cassie model. Sums of rewards are averaged over 5 seeds and smoothed with a sliding window of size 3. The agent manages to achieve a slightly faster speed than with the musculoskeletal ostrich model.

seemed sufficient to attempt to apply some of our tasks to a model of this robot. This provides an interesting opportunity to compare muscle and torque actuation with two different real-world body models with fairly similar morphologies.

We started with Cassie’s MuJoCo model, provided by the Oregon State University Dynamic Robotics Laboratory³. The model had to be adjusted to increase its stability and some equality constraints had to be added to ensure that the parts were properly connected with the rods. In order to obtain the robot poses corresponding to the mocap clips, except for the neck, we used the mocap generation pipeline described in subsection 3.4. The only modification was the use of MuJoCo’s Jacobian function during the gradient descent step (Buss, 2004), to satisfy the equality constraints. Surprisingly, Cassie’s morphology is close enough to that of an ostrich that the clips can be reproduced fairly accurately, as shown in Figure 13. The resulting dataset provides an interesting set of behaviors achievable with the Cassie robot, including steps that are more natural than those typically found in the literature and demos using this robot.

Unfortunately, Cassie’s model remains quite unstable on the motion capture tracking task, probably due to constraints in the rods and more work would be needed to achieve a satisfactory level of tracking on all clips. However, we also adapted the “run forward” task with more success, as shown in Figure 14.

5 Discussion

We presented a new musculoskeletal model of an ostrich. This is the only existing complete model of the fastest biped on Earth, an appropriate choice for studying fast locomotion. The proposed MuJoCo simulation is significantly faster than most available musculoskeletal models, usually implemented with OpenSim. Anatomical data were used to ensure that the model was reasonably realistic. We also released a motion capture dataset of ostrich behaviors and open-sourced a set of reinforcement

³<https://github.com/osudrl/cassie-mujoco-sim>

learning tasks built with dm_control. The neck control task is novel and specific to the unique morphology and control aspects of the model. The motion capture tracking task, while similar to those found in the literature, has a unique tracking objective based on matching body part locations and rotations instead of relying on joints and markers. We found that this method produced better tracking results.

The unique combination of a fast and realistic musculoskeletal model with reinforcement learning tasks provides an interesting new benchmark for RL agents in general. Muscle-actuated systems are very different from torque-actuated systems and present specific challenges that we believe could be of interest to the RL community. These include over-actuation, filtered temporal dynamics, and nonlinear force curves. In addition, it is well known that when learning with a simple reward function to control complex articulated models, policies tend to produce odd and suboptimal behavior. In effect, we found that, compared to motion capture tracking, the running task produced very unnatural and inefficient policies. A major challenge for the RL community could therefore be to try to improve the policies discovered on tasks that do not rely on motion capture data. Research directions could involve more efficient exploration methods, new inductive biases in neural architectures, or other types of regularizations, including energy efficiency and fatigue (Potvin & Fuglevand, 2017). We believe that our model and task set are ideal for this type of research. In addition, new tasks can easily be created to study different problems, such as jumping or adapting to rough terrain, and we would be happy to support their inclusion in the repository.

Furthermore, this work is an interdisciplinary effort, with the goal of bridging several research communities. We believe that the interaction between fields such as biomechanics, reinforcement learning, graphics, robotics, and computational neuroscience can produce new understanding of biological bodies and advances in artificial ones. The use of reinforcement learning to improve our model can be more widely accepted for biomechanics research in place of more traditional techniques and can help produce synthetic muscle activation data. This data can in turn help fields such as sports science to create more effective training or recovery programs (Coste et al., 2017). In addition, the computer graphics community is striving for more true-to-life animations, and musculoskeletal simulations could be an important asset in producing more natural movements while introducing more accurate physical interactions and skin deformations when using volumetric muscles. Roboticists have attempted to implement robots powered by artificial muscles, and our work could be used to perform comparisons in energy efficiency and joint movement capabilities. Finally, in computational neuroscience research, models of motor neurons, motor primitives, and muscle synergies could be tested using our simulation.

Acknowledgements

The research presented in this paper has been supported by the Royal Veterinary College and Dyson Technology Ltd.

References

- Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: A system for large-scale machine learning. *USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2016.
- Rinat Abdashitov, Seungbae Bang, David IW Levin, Karan Singh, and Alec Jacobson. Interactive modelling of volumetric musculoskeletal anatomy. *arXiv preprint arXiv:2106.05161*, 2021.
- Dmitry Akimov. Distributed soft actor-critic with multivariate reward representation and knowledge distillation. *arXiv preprint arXiv:1911.13056*, 2019.
- R McN Alexander, GMO Maloiy, R Njau, and AS Jayes. Mechanics of running of the ostrich (*Struthio camelus*). *Journal of Zoology*, 1979.
- Baptiste Angles, Daniel Rebain, Miles Macklin, Brian Wyvill, Loic Barthe, Jp Lewis, Javier Von Der Pahlen, Shahram Izadi, Julien Valentin, Sofien Bouaziz, and Andrea Tagliasacchi. VIPER:

- Volume invariant position-based elastic rods. *Computer Graphics and Interactive Techniques (SIGGRAPH)*, 2019.
- Taylor Apgar, Patrick Clary, Kevin Green, Alan Fern, and Jonathan W Hurst. Fast online trajectory optimization for the bipedal robot Cassie. *Robotics: Science and Systems*, 2018.
- Gabriel Barth-Maron, Matthew W Hoffman, David Budden, Will Dabney, Dan Horgan, Dhruva Tb, Alistair Muldal, Nicolas Heess, and Timothy Lillicrap. Distributed distributional deterministic policy gradients. *arXiv preprint arXiv:1804.08617*, 2018.
- Marc G Bellemare, Will Dabney, and Rémi Munos. A distributional perspective on reinforcement learning. *International Conference on Machine Learning (ICML)*, 2017.
- Kevin Bergamin, Simon Clavet, Daniel Holden, and James Richard Forbes. DReCon: Data-driven responsive control of physics-based characters. *ACM Transactions On Graphics (TOG)*, 2019.
- Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI Gym. *arXiv preprint arXiv:1606.01540*, 2016.
- Muge Bulat, Nuray Korkmaz Can, Yunus Ziya Arslan, and Walter Herzog. Musculoskeletal simulation tools for understanding mechanisms of lower-limb sports injuries. *Current Sports Medicine Reports*, 2019.
- Samuel R Buss. Introduction to inverse kinematics with Jacobian transpose, pseudoinverse and damped least squares methods. *IEEE Journal of Robotics and Automation*, 2004.
- C Böhmer, J Prevoteau, O Duriez, and A Abourachid. Gulper, ripper and scrapper: Anatomy of the neck in three species of vultures. *Journal of Anatomy*, 2019.
- Nuttapong Chentanez, Matthias Müller, Miles Macklin, Viktor Makoviychuk, and Stefan Jeschke. Physics-based motion capture imitation with deep reinforcement learning. *Conference on Motion, Interaction and Games (MIG)*, 2018.
- Christine Azevedo Coste, Vance Bergeron, Rik Berkelmans, Emerson Fachin Martins, Ch'e Fornusek, Arnin Jetsada, Kenneth J Hunt, Raymond Tong, Ronald Triolo, and Peter Wolf. Comparison of strategies and performance of functional electrical stimulation cycling in spinal cord injury pilots for competition in the first ever CYBATHLON. *European Journal of Translational Myology*, 2017.
- Sebastien Cotton, Ionut Mihai Constantin Olaru, Matthew Bellman, Tim van der Ven, Johnny Godowski, and Jerry Pratt. FastRunner: A fast, efficient and robust bipedal robot. concept and planar simulation. *IEEE International Conference on Robotics and Automation*, 2012.
- Andrew R Cuff, Monica A Daley, Krijn B Michel, Vivian R Allen, Luis Pardon Lamas, Chiara Adami, Paolo Monticelli, Ludo Pelligand, and John R Hutchinson. Relating neuromuscular control to functional anatomy of limb muscles in extant archosaurs. *Journal of Morphology*, 2019.
- 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*, 2007.
- Gordon Dzemski and Andreas Christian. Flexibility along the neck of the ostrich (*Struthio camelus*) and consequences for the reconstruction of dinosaurs with extreme neck length. *Journal of Morphology*, 2007.
- Scott Fujimoto, Herke Hoof, and David Meger. Addressing function approximation error in actor-critic methods. *International Conference on Machine Learning (ICML)*, 2018.
- Thomas Geijtenbeek, Michiel Van De Panne, and A Frank Van Der Stappen. Flexible muscle-based locomotion for bipedal creatures. *ACM Transactions on Graphics (TOG)*, 2013.
- Leonard Hasenclever, Fabio Pardo, Raia Hadsell, Nicolas Heess, and Josh Merel. CoMic: Complementary task learning & mimicry for reusable skills. *International Conference on Machine Learning (ICML)*, 2020.

- John R Hutchinson, Victor Ng-Thow-Hing, and Frank C Anderson. A 3D interactive method for estimating body segmental parameters in animals: Application to the turning and running performance of *Tyrannosaurus rex*. *Journal of Theoretical Biology*, 2007.
- John R Hutchinson, Jeffery W Rankin, Jonas Rubenson, Kate H Rosenbluth, Robert A Siston, and Scott L Delp. Musculoskeletal modelling of an ostrich (*Struthio camelus*) pelvic limb: Influence of limb orientation on muscular capacity during locomotion. *PeerJ*, 2015.
- Aleksi Ikkala and Perttu Hämäläinen. Converting biomechanical models from OpenSim to MuJoCo. *arXiv preprint arXiv:2006.10618*, 2020.
- Yifeng Jiang, Tom Van Wouwe, Friedl De Groote, and C Karen Liu. Synthesis of biologically realistic human motion using joint torque actuation. *ACM Transactions On Graphics (TOG)*, 2019.
- Devin L Jindrich, Nicola C Smith, Karin Jespers, and Alan M Wilson. Mechanics of cutting maneuvers by ostriches (*Struthio camelus*). *Journal of Experimental Biology*, 2007.
- Sergey Kolesnikov and Valentin Khrulkov. Sample efficient ensemble learning with Catalyst.RL. *arXiv preprint arXiv:2003.14210*, 2020.
- Seunghwan Lee, Ri Yu, Jungnam Park, Mridul Aanjaneya, Eftychios Sifakis, and Jehee Lee. Dexterous manipulation and control with volumetric muscles. *ACM Transactions on Graphics (TOG)*, 2018.
- Seunghwan Lee, Moonseok Park, Kyoungmin Lee, and Jehee Lee. Scalable muscle-actuated human simulation and control. *ACM Transactions On Graphics (TOG)*, 2019.
- Zhongyu Li, Christine Cummings, and Koushil Sreenath. Animated Cassie: A dynamic relatable robotic character. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.
- Zhongyu Li, Xuxin Cheng, Xue Bin Peng, Pieter Abbeel, Sergey Levine, Glen Berseth, and Koushil Sreenath. Reinforcement learning for robust parameterized locomotion control of bipedal robots. *arXiv preprint arXiv:2103.14295*, 2021.
- Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.
- Libin Liu and Jessica Hodgins. Learning basketball dribbling skills using trajectory optimization and deep reinforcement learning. *ACM Transactions on Graphics (TOG)*, 2018.
- Libin Liu, KangKang Yin, Michiel van de Panne, Tianjia Shao, and Weiwei Xu. Sampling-based contact-rich motion control. *ACM Transactions on Graphics (TOG)*, 2010.
- Josh Merel, Yuval Tassa, Dhruva TB, Sriram Srinivasan, Jay Lemmon, Ziyu Wang, Greg Wayne, and Nicolas Heess. Learning human behaviors from motion capture by adversarial imitation. *arXiv preprint arXiv:1707.02201*, 2017.
- Josh Merel, Leonard Hasenclever, Alexandre Galashov, Arun Ahuja, Vu Pham, Greg Wayne, Yee Whye Teh, and Nicolas Heess. Neural probabilistic motor primitives for humanoid control. *arXiv preprint arXiv:1811.11711*, 2018.
- Matthew Millard, Thomas Uchida, Ajay Seth, and Scott L Delp. Flexing computational muscle: Modeling and simulation of musculotendon dynamics. *Journal of Biomechanical Engineering*, 2013.
- Vismay Modi, Lawson Fulton, Alec Jacobson, Shinjiro Sueda, and David IW Levin. EMU: Efficient muscle simulation in deformation space. *Computer Graphics Forum*, 2021.
- Fabio Pardo. Tonic: A deep reinforcement learning library for fast prototyping and benchmarking. *arXiv preprint arXiv:2011.07537*, 2020.
- Fabio Pardo, Arash Tavakoli, Vitaly Levdk, and Petar Kormushev. Time limits in reinforcement learning. *International Conference on Machine Learning (ICML)*, 2018.

- Mikhail Pavlov, Sergey Kolesnikov, and Sergey Plis. Run, skeleton, run: Skeletal model in a physics-based simulation. *AAAI Spring Symposium Series*, 2018.
- Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. DeepMimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions on Graphics (TOG)*, 2018.
- Xue Bin Peng, Michael Chang, Grace Zhang, Pieter Abbeel, and Sergey Levine. MCP: Learning composable hierarchical control with multiplicative compositional policies. *arXiv preprint arXiv:1905.09808*, 2019.
- Jim R Potvin and Andrew J Fuglevand. A motor unit-based model of muscle fatigue. *PLOS Computational Biology*, 2017.
- Jeffery W Rankin, Jonas Rubenson, and John R Hutchinson. Inferring muscle functional roles of the ostrich pelvic limb during walking and running using computer optimization. *Journal of the Royal Society Interface*, 2016.
- Jacob Reher, Wen-Loong Ma, and Aaron D Ames. Dynamic walking with compliance on a Cassie bipedal robot. *European Control Conference (ECC)*, 2019.
- WI Sellers, L Margetts, KT Bates, and AT Chamberlain. Exploring diagonal gait using a forward dynamic three-dimensional chimpanzee simulation. *Folia Primatologica*, 2013.
- Ajay Seth, Jennifer Hicks, Thomas Uchida, A. Habib, Christopher Dembia, James Dunne, Carmichael Ong, Matthew DeMers, Apoorva Rajagopal, Matthew Millard, Sam Hamner, Edith Arnold, Jennifer Yong, Shrinidhi Lakshmikanth, Michael Sherman, and Scott Delp. OpenSim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. *PLOS Computational Biology*, 2018.
- Heiko Stark, Martin S Fischer, Alexander Hunt, Fletcher Young, Roger Quinn, and Emanuel Andrade. A three-dimensional musculoskeletal model of the dog. *Scientific Reports*, 2021.
- Yuval Tassa, Saran Tunyasuvunakool, Alistair Muldal, Yotam Doron, Siqi Liu, Steven Bohez, Josh Merel, Tom Erez, Timothy Lillicrap, and Nicolas Heess. dm_control: Software and tasks for continuous control. *arXiv preprint arXiv:2006.12983*, 2020.
- Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012.
- Takanobu Tsuihiji. Homologies of the transversospinalis muscles in the anterior presacral region of Sauria (crown Diapsida). *Journal of Morphology*, 2005.
- Takanobu Tsuihiji. Homologies of the longissimus, iliocostalis, and hypaxial muscles in the anterior presacral region of extant diapsida. *Journal of Morphology*, 2007.
- Jack M Wang, Samuel R Hamner, Scott L Delp, and Vladlen Koltun. Optimizing locomotion controllers using biologically-based actuators and objectives. *ACM Transactions on Graphics (TOG)*, 2012.
- Jiangxin Wang, Dace Gao, and Pooi See Lee. Recent progress in artificial muscles for interactive soft robotics. *Advanced Materials*, 2021.
- Jungdam Won, Deepak Gopinath, and Jessica Hodgins. Control strategies for physically simulated characters performing two-player competitive sports. *ACM Transactions on Graphics (TOG)*, 2021.
- Zhaoming Xie, Glen Berseth, Patrick Clary, Jonathan Hurst, and Michiel van de Panne. Feedback control for Cassie with deep reinforcement learning. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- Zhaoming Xie, Patrick Clary, Jeremy Dao, Pedro Morais, Jonanthan Hurst, and Michiel Panne. Learning locomotion skills for Cassie: Iterative design and sim-to-real. *Conference on Robot Learning (CoRL)*, 2020.

Zhiqi Yin, Zeshi Yang, Michiel Van De Panne, and KangKang Yin. Discovering diverse athletic jumping strategies. *ACM Transactions on Graphics (TOG)*, 2021.

Felix E Zajac. Muscle and tendon: Properties, models, scaling, and application to biomechanics and motor control. *Critical Reviews in Biomedical Engineering (CRB)*, 1989.

Bo Zhou, Hongsheng Zeng, Fan Wang, Yunxiang Li, and Hao Tian. Efficient and robust reinforcement learning with uncertainty-based value expansion. *arXiv preprint arXiv:1912.05328*, 2019.