Inverse Distance Weighting to Rapidly Generate Large Simulation Datasets

Kalyn M. Kearney[1], Joel B. Harley[2], Jennifer A. Nichols[1]
[1]J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, United States
[2]Department of Electrical and Computer Engineering, University of Florida, Gainesville, United States
Email: kalynkearney@ufl.edu, joel.harley@ufl.edu, jnichols@bme.ufl.edu

Summary

Advances in machine learning have increased demand for large, biomechanical datasets. Musculoskeletal simulation techniques, such as computed muscle control (CMC), have the potential to quell this demand at little expense. Yet, calculating one second of muscle activations using CMC can take minutes. We tested the accuracy of using inverse distance weighting (IDW) to approximate muscle activations across models scaled by mass and height. Our results suggest that IDW can accurately predict muscle activations with little cost.

Introduction

Given sufficient data, machine learning approaches can elucidate complex biomechanical relations. Physics-based simulations provide one approach for generating these large datasets. However, some simulation approaches have a high computational cost, thereby hindering the number and type of observations achievable in a reasonable time. For example, CMC, which computes muscle activations from target kinematics and forces [1], can take minutes to simulate seconds of movement. Across thousands of simulations, this can result in weeks of calculations. In contrast, interpolation methods, such as IDW, are computationally fast. The objective of this work was to investigate the utility of IDW for interpolating muscle activations from sparse CMC datasets.

Methods

We generated three CMC datasets with different levels of sparsity. Briefly, using OpenSim v. 4.1, we scaled a generic thumb model [2] across uniformly-spaced combinations of 30 masses and 10 heights chosen to represent 5th-95th percentile adults [3]. For each model, muscle activations to achieve a 40 N lateral pinch force were calculated using CMC. 117 CMC simulations ran to completion, providing muscle activations for a population roughly 63-103 kg and 1.72-1.88 m tall. This dataset was divided into three uniformly-spaced grids: 17 masses by 7 heights (high density), 9 masses by 4 heights (medium density), and 5 masses by 3 heights (low density).

To evaluate IDW's ability to approximate muscle activations, we selected 108 random mass-height pairs that did not correspond to the 117 computed grid vertices. We scaled the thumb model to each mass-height pair and identified the 3 closest grid vertices on a normalized scale. We used IDW with an inverse power weighting function [4] and power parameter of 1 to interpolate the 108 sets of activations. This approach preferentially weighted activations from closer grid vertices. We repeated this procedure for all three grid densities.

To compare against the interpolated muscle activations, we simulated lateral pinch using CMC for each randomly selected mass-height pair. We quantified the root mean square error (RMSE) between the interpolated activations and the CMC activations. We report the mean RMSE for each simulation (averaging across all muscles) as well as the mean RMSE for each muscle (averaging across simulations).

Results and Discussion

IDW approximated muscle activations in lateral pinch with minimal error and substantially less computational cost than CMC alone. Fig. 1 illustrates the low density and high density RMSE results. The RMSE associated with interpolated muscle activations never exceeded 4%. Increasing the grid density slightly decreased RMSE, which averaged 0.33%, 0.40%, and 0.45% for high, medium, and low density grids, respectively. At the grid boundaries, the high density grid had higher RMSEs. As CMC simulations outside the grid failed to run to completion, CMC solutions at the boundary may have been less stable. The *flexor pollicis longus* and *extensor carpi ulnaris* had the highest mean RMSE, peaking at 3.2% and 3.7%, respectively, on the low-density grid. No other muscles exceeded a mean RMSE of 1.5%. Furthermore, computation time for CMC simulations averaged 4.22 core-minutes, while IDW averaged 0.95 core-seconds per mass-height pair.

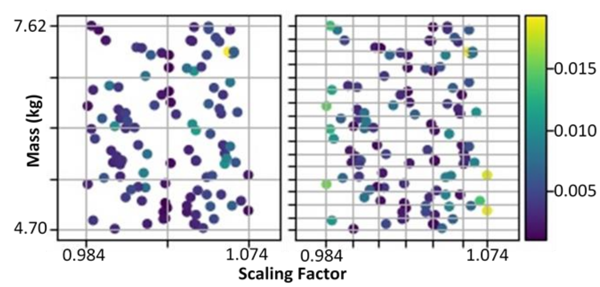

Figure 1. Mean RMSE of interpolated muscle activations across each randomly selected mass-height pair (colored). Muscle activations from CMC were calculated at the grid vertices (gray).

Conclusions

The present work proposed and tested a method for efficient and accurate estimation of muscle activations using a grid of CMC simulations and IDW. This work indicates IDW is a powerful approach for rapidly estimating muscle activations from sparse CMC datasets. Rapid generation of data will facilitate machine learning analyses of human movement.

References

[1] Thelen et al. (2003). *J Biomech*, **36**: 321-328.
[2] Nichols et al. 2017. *J Biomech.* **58**:97-104.
[3] Fryar et al. (2016). *Vital Health Stat*, **39**: 1-46.
[4] Shephard, D (1968). *Proc. of ACM Nat. Conf.*, 517-524.