Inverse Optimal Control for a Simple Stepping Task

John R. Rebula[1,2], Sean Mason[1,2], Stefan Schaal[1,2], Ludovic Righetti[1]

***Abstract*— We consider a method of characterizing cost functions underlying legged locomotion. We use a trajectory optimization approach, allowing a separation of the constraints associated with walking, which walkers must follow, from various proposed cost components, which reflect preference. Here we present results on synthetic walking data, where we generate stepping motions using known cost functions, and use an iterative trajectory optimization process which requires only a single demonstrated optimal trial to infer the cost function weightings that were used to generate it.**

I. Introduction

While walking is a complex whole-body task, some underlying costs govern overall gait. Optimization approaches to understanding behavior assume that locomotion arises from trying to achieve a certain goal (forward motion) while minimizing some measure of cost. A range of potential costs of human motion have been propose, including metabolic cost (e.g. [1]). This may include costs for using muscle [2] to produce work or force. Kinematic jerk has also been proposed as a cost that humans tend to minimize, and may be related to costs for muscle activation. Stability characteristics may be another consideration. A combination of cost terms is generally beneficial for predicting and describing natural motion. In this work, we will consider locomotion control as an optimization problem with the goal of characterizing the underlying cost function used to generate the motion. Given a demonstrated behavior which we consider optimal, we will estimate a cost function that may have given rise to that behavior. This process, known as inverse optimal control, has been used to study human locomotion in terms of overall path characteristics [3]. We use an inverse optimal control framework, treating the given task (taking a step), as an optimization problem minimizing some combination of costs while satisfying walking constraints, defined by a simple physical model of walking. We plan to apply this method to, for example, human subjects experiments where investigators design experimental conditions hypothesized to re-weight the underlying costs (e.g. testing if a condition with perturbations increased a subject's preference for stable motions).

II. Our Approach

To test our overall approach for performing inverse optimal control we test our ability to characterize a known cost function used in trajectory optimization of a simple model of walking. While various approaches have been proposed for inverse optimal control [IOC examples, pi2, that optimization paper], an important consideration in walking is the handling of active physical constraints (e.g. unilateral ground contact). Additional constraints may be assumed based on experimental conditions, such as by controlling the speed of a treadmill on which a subject walks. Practically, we wish to separate, where possible, aspects of the behavior are determined by physical constraints from aspects determined by subject preference. We therefore consider a general trajectory optimization method for inverse optimal control, incorporating both costs and hard constraints. In this work, we first show the ability of an inverse optimal control method to characterize a cost function of walking under ideal conditions. We generate an optimal motion using a simple trajectory optimization procedure, and then attempt to rediscover the cost weights that were used in the optimization, using the single demonstrated trajectory (Figure 1).

A. Trajectory Optimization

We consider a linear inverted pendulum model stepping to a given point, with constrained step timing. The states in the optimization ($\mathbf{x}$) procedure include the trajectories of the center of mass (COM), center of pressure (COP), and feet. These trajectories are discretized throughout the step, consisting of a stance phase, a left foot swing phase, and a final double support phase (1 second per phase). The cost function $c(\mathbf{x}) = \mathbf{w}^T \cdot \phi(\mathbf{x})$ consists of a weighted (with weights $\mathbf{w}^*$) combination of squared cost features ($\phi(\mathbf{x})$): foot placement error, COP acceleration, swing foot acceleration (assuming a trivial massless point foot model). Constraints include the model physics, such as the integrated equations of motion and a convex hull base of support constraint for the COP location. For this task we also constrain the final COM position to be directly over the final swing foot location. An example optimal trajectory ($\mathbf{x}^*$) is found for a desired step forward of 0.3m. This problem is solved using a general sequential quadratic programming implementation, IPOPT.

B. Inverse Optimal Control

The inverse optimal control method consists of generating a set of suboptimal feasible trajectories (X^-) around the optimal demonstration, and finding cost weights ($\mathbf{w}$) which ensure the overall cost function is lower for the demonstrated trajectory than for any of these suboptimal trajectories. In particular, we first sample N normally distributed trajectories around $\mathbf{x}^*$, $x_i^r, i \in [1 \dots N]$. As there are no guarantees of feasbility with the problem constraints for trajectories generated randomly, we generate feasible versions of these trajectories. For each randomized trajectory x_i^r a trajectory optimization is formulated with all of the constraints (dynamics, task, etc) as the original optimization above, but using a cost function penalizing the distance of the resulting feasible trajectory $x_i^{r,f}$ from the original x_i^r. Our initial set of suboptimal trajectories is then $X^- = x_i^{r,f}, i \in [1 \dots N]$ Given a set of suboptimal trajectories X^-, we find cost function weights $\mathbf{w}^c$ consistent with the optimality of the demonstrated trajectory $\mathbf{x}^*$ by solving a constraint satisfaction problem,

$$\mathbf{w}^{cT} \cdot \phi(\mathbf{x}^*) < \mathbf{w}^{cT} \cdot \phi(\mathbf{x}^-), \forall \mathbf{x}^- \in X^- \& ||\mathbf{w}^c||_2 = 1. \quad (1)$$

The 2-norm of the weights is constrained to 1 to avoid the trivial solution of $\mathbf{w}^c = \mathbf{0}$. As multiple $\mathbf{w}$ may satisfy these constraints, we also minimize the 3-norm of the weight vector $||\mathbf{w}||_3$, encouraging use of all possible weights. While potentially this procedure could result in the optimal cost weight $\mathbf{w}^*$, we found it difficult to produce trajectories with, for example,

[1]Max Planck Institute for Intelligent Systems, Tuebingen, Germany
[2]University of Southern California, Los Angeles, USA rebula@usc.edu

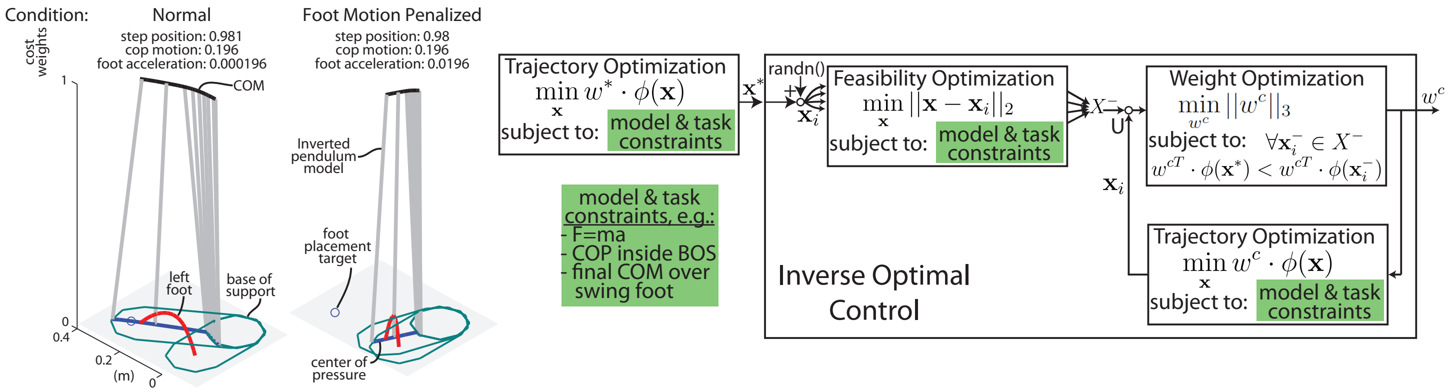

Fig. 1. Left: example optimal trajectories arising from different choices for weight costs. Right: an overview of our approach. Trajectory optimization is used to generate an optimal trajectory with chosen weights. The inverse optimal control algorithm then generates random trajectories based on the demonstration and optimizes these trajectories to ensure that they are consistent with the constraints. A weight optimization finds a set of weights consistent with relative optimality of the demonstrated trajectory. The weights are then used to generate a new trajectory, which is added to the set of suboptimal states. The process then repeats until the trajectory matches the demonstration.

smoother COP trajectories than the demonstration, due to our method of generating suboptimal trajectories through the use of added noise. This resulted in $\mathbf{w}^c$ which were independent of the COP and foot acceleration costs. To address this, we generate an additioanl suboptimal trajectory, by using the candidate weights $\mathbf{w}^c$ to perform a full trajectory optimization outlined in section II-A. This yields a new candidate trajectory x^c. If this trajectory is sufficiently close to the demonstrated trajectory $\mathbf{x}^*$, then $\mathbf{w}^c$ are determined to be consistent with the optimality of the demonstration. Otherwise, x^c is added to X^- and the constraint satisfaction problem (Equation 1) is solved again for a new $\mathbf{w}^c$, and this process iteratively repeats.

III. Results

Optimal trajectories are found for the given task with two sets of optimal weights, one that achieves a foot placement close to the desired, one that more heavily penalizes foot acceleration (Figure 1, left). The inverse optimal control procedure converges to the actual weights after about 20 iterations (Figure 2).

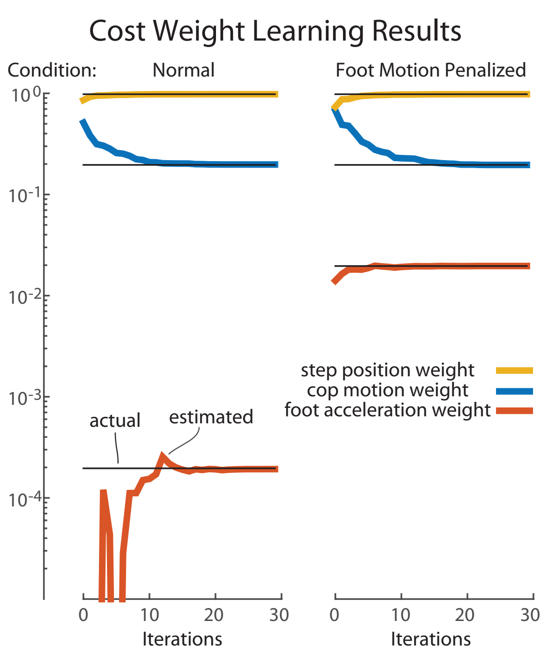

Fig. 2. The optimal weights during the iterative IOC

IV. Discussion

We have shown an example of inverse optimal control infering cost weights for a model of walking which, while simple, includes important features such as contact switching and center of pressure constraints. We wish to extend our methodology to infer cost weights based on human motion when the full set of potential cost features is unknown, and when some proposed cost features may not in fact influence the subject's behavior. We therefore will investigate the performance of our procedure in the presense of distractor features, where we consider features in our inverse optimal control that are not considered in the forward optimal control, as well as missing features where insufficient cost components are proposed to reproduce the demonstrated Behavior. We will also extend this approach using a more complete model of locomotion, including joint level dynamics as opposed to lumped mass dynamics. In addition, human data may afford the collection of many near optimal trajectories (e.g. from treadmill walking), which could potentially be incorporated as multiple optimal demonstrations. Alternatively, they could be interpreted as each slightly suboptimal and therefore ease the requirement of suboptimal trajectory generation in our current formulation.

Acknowledgments

This research was supported in part by National Science Foundation grants IIS-1205249, IIS-1017134, EECS-0926052, the Office of Naval Research, the Okawa Foundation, the Max-Planck-Society and the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No 637935). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding organizations.

References

[1] J. M. Donelan, R. Kram, and A. D. Kuo, "Mechanical work for step-to-step transitions is a major determinant of the metabolic cost of human walking," *The Journal of experimental biology*, vol. 205, no. Pt 23, pp. 3717–3727, Dec. 2002.

[2] A. D. Kuo and J. M. Donelan, "Dynamic Principles of Gait and Their Clinical Implications," *Physical Therapy*, vol. 90, no. 2, pp. 157–174, Feb. 2010. [Online]. Available: http://ptjournal.apta.org/content/90/2/157

[3] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion–an inverse optimal control approach," *Auton. Robots*, vol. 28, no. 3, pp. 369–383, Apr. 2010. [Online]. Available: http://dx.doi.org/10.1007/s10514-009-9170-7