

Animating Human Lower Limbs Using Contact-Invariant Optimization

Igor Mordatch¹

Jack M. Wang^{2,3}

Emanuel Todorov¹

Vladlen Koltun^{3,4}

¹University of Washington

²University of Hong Kong

³Stanford University

⁴Adobe Research

Abstract

We present a trajectory optimization approach to animating human activities that are driven by the lower body. Our approach is based on contact-invariant optimization. We develop a simplified and generalized formulation of contact-invariant optimization that enables continuous optimization over contact timings. This formulation is applied to a fully physical humanoid model whose lower limbs are actuated by musculotendon units. Our approach does not rely on prior motion data or on task-specific controllers. Motion is synthesized from first principles, given only a detailed physical model of the body and spacetime constraints. We demonstrate the approach on a variety of activities, such as walking, running, jumping, and kicking. Our approach produces walking motions that quantitatively match ground-truth data, and predicts aspects of human gait initiation, incline walking, and locomotion in reduced gravity.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: human motion, human simulation, optimization

Links:  DL  PDF

1 Introduction

Creating realistic human motion from compact high-level objectives is a fundamental problem in character animation. For example, we would like to generate realistic walking for a given human character by simply specifying the desired velocity, or realistic jumping by specifying the desired height. While animations can always be created by a skilled artist given sufficient time and effort, a technique that can generate custom realistic motions can accelerate the process by providing a high-quality animation that can be refined by the artist. Such a technique can also be used for assisting novice animators and for pre-visualization in production settings.

In this paper, we present a method for *de novo* synthesis of human motion. Our method produces animations for given sparse objectives: walking emerges when the objective specifies a velocity for the torso, running emerges when the specified velocity is higher, jumping emerges when a target height is given, and kicking is synthesized when the character’s foot must have certain velocity at a given point in space. The primary goal of our work is to enable automatic synthesis of realistic lower body motion for human activities that are driven by the lower body. Such activities include walking, running, jumping, and kicking.

To synthesize high-fidelity lower limb motion we use a trajectory optimization approach that is based on spacetime constraints. Trajectory optimization techniques have been applied to human characters, but have either required objectives that depend on task-specific motion capture data [Liu et al. 2005] or failed to produce realistic results [Erez and Todorov 2012]. The problem is challenging because spacetime constraints on three-dimensional humanoid models lead to high-dimensional search spaces and complex nonlinear constraints that are difficult to satisfy. Furthermore, ground contact forces that change abruptly with body kinematics lead to discontinuous objective functions prone to poor local minima.

Nevertheless, we show that high-fidelity lower limb motions for a range of activities can be generated by a single trajectory optimization formulation that does not rely on prior task-specific motion data. We achieve this using a reformulation of contact-invariant optimization (CIO) [Mordatch et al. 2012b]. CIO is an optimization framework that smoothes out discontinuities in the objective and allows a single continuous optimization to search over the space of possible motion trajectories and contact patterns. While CIO has been previously used on reduced character models, we apply the framework to human characters with equations of motion that capture the detailed dynamics of the body. In order to enable fine-grained optimization over contact timings, we generalize and simplify CIO by eliminating the quantization of contact patterns that previous formulations relied on.

Since the actuation mechanism of the humanoid model has significant impact on the realism of synthesized motion [Liu et al. 2005; Wang et al. 2012], we actuate the lower body joints with Hill-type musculotendon units (MTUs). The MTUs regularize the synthesized torque patterns and enable the use of a biologically-plausible effort model. Unlike prior work that utilized MTUs and biologically-inspired objectives to generate human locomotion [Wang et al. 2012], we do not rely on a task-specific control structure. As a result, our approach is not limited to steady-state walking and running.

We restrict our attention to the fidelity of lower limb motion since a simplified model of the musculoskeletal system for the lower body is readily available [Wang et al. 2012]. The ideas we present may also enable *de novo* synthesis of human motion in which the entire body moves with high fidelity to real-world human movement, but this likely requires developing computationally efficient musculoskeletal models for the full body, perhaps building on the work of Lee et al. [2009]. We leave such development to future work.

Our work demonstrates that a single trajectory optimization formulation can produce high-fidelity lower body motion for a range of human activities, without the need for prior motion data or task-specific control structures. We evaluate the presented approach by synthesizing high-fidelity locomotion in a variety of conditions, as well as jumping and kicking motions.

2 Related Work

One natural approach to human motion synthesis is to design controllers for humanoid models in forward dynamical simulations. Early efforts focused on hand-designing controllers for specific activities, resulting in characters with impressive movement reper-

toires [Hodgins et al. 1995; Faloutsos et al. 2001]. In recent years, significant efforts have been devoted to improving controller robustness and motion style for simulated locomotion [Yin et al. 2007; de Lasa et al. 2010; Coros et al. 2010; Wang et al. 2012]. Yet methods that produce the most human-like motions require significant customization of the target trajectory or the control structure [Coros et al. 2010; Wang et al. 2012], while more flexible control algorithms fall short of generating motions that appear natural [de Lasa et al. 2010]. For example, Wang et al. [2012] demonstrate that human-like walking and running can emerge from biologically-plausible modeling of the musculotendon units in the lower body. However, their work relies on a detailed control structure that was designed specifically for steady-state locomotion and produced only fixed-velocity walking or running.

We adopt an alternative approach, which casts human motion synthesis in terms of trajectory optimization. This approach simultaneously solves for the motion, the actuation forces, and the contact forces. Specific requirements by the animator can be incorporated as objectives or constraints. Trajectory optimization had shown early promise with simple character models [Witkin and Kass 1988; Cohen 1992], but applying it to full three-dimensional human models proved challenging. Traditional trajectory optimization formulations have failed to produce high-fidelity human motion de novo, both due to the difficulty of the optimization problem, and the difficulty of formulating general constraints or objectives that characterize “human-like” motions. Successful applications have often relied on using motion capture or prior animation data to constrain the problem, or focused on ballistic motions that are determined by relatively simple mechanical principles [Popović and Witkin 1999; Liu and Popović 2002; Fang and Pollard 2003; Safonova et al. 2004; Sulejmanpašić and Popović 2005; Liu et al. 2005]. Wampler and Popović [2009] presented impressive results for imaginary animal locomotion, but their work did not produce high-fidelity human gaits. Al Borno et al. [2013] explored a wide range of humanoid motions, but their work relied on many task-specific objectives and produced results of limited fidelity.

Our approach is based on the contact-invariant optimization (CIO) framework [Mordatch et al. 2012b]. CIO incorporates contact locations into the optimization problem, allows these contact locations to apply virtual forces to the character while the optimization is in progress, and makes these forces dynamically consistent as the optimization converges. This has the effect of reshaping the difficult objective function into one that is smooth and easier to optimize. So far, CIO has only been applied to simplified models with massless limbs [Mordatch et al. 2012b] and fingers [Mordatch et al. 2012a]. In these prior applications, the models were largely non-physical and joint angles were computed using inverse kinematics. We extend CIO to synthesize human motion using a fully physics-based humanoid model. In the process, we simplify and generalize the CIO formulation. Traditional CIO relies on introducing auxiliary contact activation variables into the optimization to break the trajectory up into piecewise constant contact phases for the purpose of accurately modeling rigid contact forces. We show that these contact phases are not necessary and that contact activation variables do not need to be explicitly represented. In contrast to prior applications of CIO, our formulation enables continuous optimization over contact timings.

In movement science, trajectory optimization with abstract models had been used to study arm and saccadic eye movements [Harris and Wolpert 1998], as well as human gait selection [Srinivasan and Ruina 2006]. Anderson and Pandy [2001] used dynamic optimization with a detailed musculoskeletal model of the lower body to synthesize a single walk cycle. However, they dealt with a much more restricted solution space, in which the first configuration in the trajectory was specified based on human walking data and the

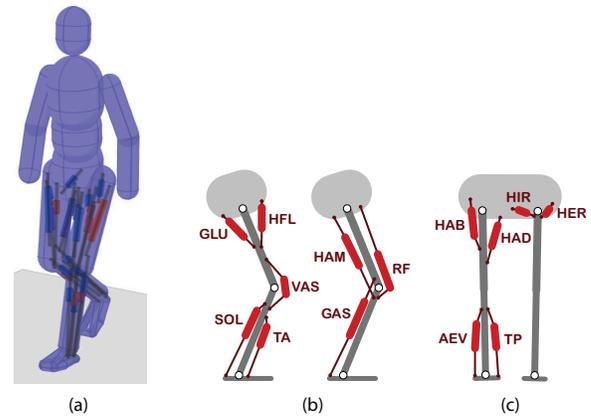


Figure 1: Humanoid model. (a) A visualization of the model during a walk. Red and blue cylinders in the legs correspond to the currently active and inactive MTUs. (b) Sagittal plane MTUs, defined following Wang et al. [2012]. (c) Additional MTUs that provide control in the coronal plane (HAD, HAB, AEV, TP) and the transverse plane (HER, HIR).

motion was constrained to be cyclical. The same model was also used to simulate a vertical jump [Anderson and Pandy 1999], again with a fixed initial configuration. We likewise use a musculoskeletal model for the lower body, which enables better modeling of energy expenditure and restricts internal forces to be biologically plausible. However, our optimization framework is considerably more powerful and supports motion synthesis for a broader range of activities.

3 Overview

Our method is based on the spacetime constraints framework [Witkin and Kass 1988]. The input is a set of sparse objectives, such as a target COM velocity for locomotion or a target foot position and velocity for a kick. The output is a kinematic trajectory for the humanoid model along with corresponding muscle activations, contact locations, ground reaction forces, and other parameters that describe the state of the simulated model over time. We provide an overview of both the model and the objective terms in this section.

3.1 Humanoid Model and Forces

The kinematic trajectory is a sequence of pose vectors over time. Each pose \mathbf{q} contains 36 degrees of freedom (DOFs), including 6 for the root translation and orientation and 30 for the joints. We employ a model with mass distributions approximating a 180 cm, 70 kg male [Wang et al. 2012]. A physically consistent kinematic trajectory for the model is determined by external forces due to gravity and ground contact, as well as internal forces generated by the model. The internal forces actuate the humanoid model through torques generated at its 30 joint DOFs.

Our model of force generation for the lower body joints is motivated by properties of human musculotendon units (MTUs). Physiological properties of MTUs shape the torque patterns that can be produced by the human body. Given identical MTU control signals, the actual torque output at a particular joint can differ greatly depending on the current joint moment arms, muscle fiber length, and contraction velocity. Moreover, biarticular MTUs generate torques for pairs of joints, which leads to correlations in the torque patterns. Consequently, the model’s capacity to generate a particular torque pattern is highly dependent on its kinematic state. We capture this dependency using Hill-type MTU models, which serve to increase

the fidelity of the synthesized motion.

In particular, the lower body joints (hip, knee, and ankle) of our humanoid model are actuated exclusively by Hill-type MTUs. We use the MTUs that were used in the work of Wang et al. [2012], as well as additional ones that support torque generation in the coronal and transverse planes. Our model is illustrated in Figure 1. Four MTUs in each leg generate torque in the coronal plane: the hip adductor and abductor muscle groups (HAD/HAB), the ankle evorator group (AEV), and the tibialis posterior (TP), which generates ankle inversion torques. We also add the hip internal and external rotation muscle groups (HIR/HER) that generate hip torque in the transverse plane. The choice of MTUs and their parameters are based on simplifying a more detailed musculoskeletal model [Delp et al. 1990].

While we do not model force generation in the upper body with MTUs, we treat active motor torques and passive torques separately. Passive torques provide a coarse approximation of the effects of musculoskeletal structure by applying spring-damper forces to the upper body joints as a function of kinematics [Liu et al. 2005; Wampler and Popović 2009]. In particular, we assume that damping forces are applied to all upper-body joints, while the neck and back have weak tendencies to stay upright. The primary function of this treatment is to better model the effort of actively moving the upper body (Section 4.5).

Ground reaction forces are applied to the system when specific body parts come into contact with the ground. Since we are interested in capturing detailed pressure distributions over the feet, our model has six potential contact points on each foot, as illustrated in Figure 2. Two contact points are on the toe edge, two are on the ball of the foot, and two are on the heel. Ground reaction forces and foot locations are treated as free variables in the optimization. Their specific forms are described in Section 4.2.

When the motion of the system is physically consistent, the quantities described above are related by the equations of motion (EOM) for an articulated rigid body system:

$$M(\mathbf{q})\ddot{\mathbf{q}} + C(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{f}_{\text{GRF}} + \mathbf{f}_{\text{MTU}} + \mathbf{f}_{\text{passive}} + \mathbf{f}_{\text{motor}}. \quad (1)$$

Here M is the system inertia matrix, C is the matrix of Coriolis, centrifugal, and gravitational terms, and \mathbf{f}_{GRF} , \mathbf{f}_{MTU} , $\mathbf{f}_{\text{passive}}$, and $\mathbf{f}_{\text{motor}}$ are generalized forces due to ground reaction, musculotendon units, passive structures, and joint motors. The matrices M and C are determined by the mass distribution and the kinematics of the humanoid model. We discuss individual terms in the EOM in more detail in Section 4.

3.2 Objective Function

Since the input to our method is a set of sparse task objectives (J_{task}), a key requirement for our approach is to synthesize realistic movement for a variety of activities while keeping J_{task} as compact as possible. For example, the user only needs to change the target velocity to synthesize locomotion gaits appropriate for different speeds. Of course, in the absence of additional constraints, many trajectories \mathbf{q} can satisfy J_{task} , including ones that are highly implausible. We must therefore use additional objectives that prioritize physical and biomechanical consistency. These *task-independent* objectives are universal for all motions we synthesize and include an EOM term, a contact consistency term, a body integrity term, and a musculotendon actuation term.

The EOM term (J_{EOM}) serves to enforce the equations of motion (1), which capture the requirement for gross dynamical consistency: bodies should move in ways that are consistent with their mass properties and forces in the system. The contact consistency term (J_{CIO}) governs the behavior of contact forces, so that forces are

only applied when contact locations on the corresponding surfaces align. The body integrity term (J_{body}) enforces joint limits and prevents self-collision (e.g., legs going through each other). The musculotendon actuation term (J_{MTU}) requires control forces generated by musculotendon units to be consistent with muscle physiology. Finally, to resolve the remaining ambiguities, we assume a preference for economy of effort (J_{effort}). The complete objective is a linear combination of these terms:

$$E = \int_{t_i}^{t_f} \sum_k w_k J_k, \quad (2)$$

where $k \in \{\text{task, EOM, CIO, body, MTU, effort}\}$. The objective is in general nonconvex and we seek solutions through gradient-based optimization. The full state vector is defined in Section 5, which also describes the optimization algorithm. All the motions we synthesize use the same set of weights w_k .

Note that ideally all terms except for J_{effort} should be satisfied fully. Accordingly, we assign high weights to these terms so that they converge to near-zero values at the end of the optimization. In contrast, J_{effort} is inherently a soft constraint that disambiguates otherwise admissible motions that minimize all other terms. The effort term is particularly important for low-energy activities such as walking: there are many physically and biomechanically consistent ways to move forward at low speed. For more energetic motions such as jumping and kicking, where the other terms sufficiently constrain the motion, J_{effort} does not influence the solution significantly.

4 Objective Terms

4.1 Equations of Motion

The synthesized trajectory is physically consistent if it respects the equations of motion (1). In practice, we compute all quantities in (1) except for the active motor torques $\mathbf{f}_{\text{motor}}$, which is recovered by rearranging the equation:

$$\mathbf{f}_{\text{motor}} = M(\mathbf{q})\ddot{\mathbf{q}} + C(\mathbf{q}, \dot{\mathbf{q}}) - \mathbf{f}_{\text{GRF}} - \mathbf{f}_{\text{MTU}} - \mathbf{f}_{\text{passive}}. \quad (3)$$

As discussed previously, only the upper body DOFs are actuated using joint motor torques, hence elements of $\mathbf{f}_{\text{motor}}$ that correspond to the root and lower body DOFs (denoted by Q_{root} and Q_{lower}) should vanish for physical consistency. We define the EOM term to drive the residuals for these DOFs towards zero:

$$J_{\text{EOM}} = \sum_i (f_{\text{motor}}^i)^2, \quad (4)$$

where $q^i \in Q_{\text{root}} \cup Q_{\text{lower}}$. The matrices M and C are computed using MuJoCo [Todorov et al. 2012]. We discuss how the other generalized forces on the right hand side of (3) are computed in the rest of this section.

4.2 Contact Consistency

In principle, since \mathbf{f}_{GRF} is a function of the body kinematics (\mathbf{q}) and contact geometry, one could naively employ a procedure that detects collisions and applies contact forces accordingly. In practice, however, such an approach results in intractable optimization problems, since hard dynamics constraints and contact discontinuities break the state space up into a jagged landscape with poor local minima. Traditional formulations of spacetime constraints avoid the issue by requiring user-specified fixed contact locations, which significantly restrict the range of motions that can be explored by the optimizer.

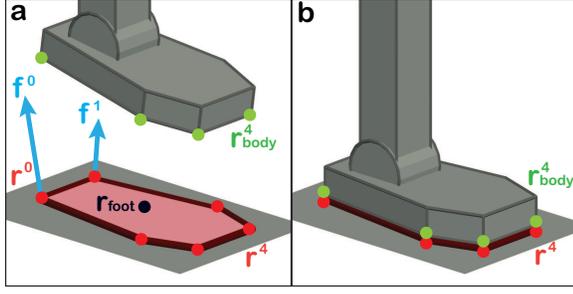


Figure 2: Contact-invariant optimization (CIO). Left: Although the contact points $\mathbf{r}_{\text{body}}^i$ on the foot are not coincident with the ground, CIO can still employ ground reaction forces \mathbf{f}^i originating at points \mathbf{r}^i . Such “virtual” contacts are penalized by the contact consistency term J_{CIO} , but serve to smooth out the objective and assist the optimization. Right: A valid contact state is achieved when J_{CIO} is minimized.

We instead adopt contact-invariant optimization (CIO), which incorporates contact positions and forces into the optimization, and treats contact consistency as an objective term to be optimized alongside others. Figure 2 illustrates our treatment of contact positions and forces. Each foot has a virtual contact location with the ground at a planar world-space position \mathbf{r}_{foot} and orientation θ_{foot} . Six individual contacts \mathbf{r}^i are calculated relative to this virtual frame. The corresponding points on the body are denoted by $\mathbf{r}_{\text{body}}^i$. The ground applies a reaction force \mathbf{f}^i at contact point \mathbf{r}^i . For each foot, the contact location $[\mathbf{r}_{\text{foot}} \ \theta_{\text{foot}}] \in \mathbb{R}^3$ and individual contact forces $\mathbf{f}^i \in \mathbb{R}^3$ are free variables in the optimization. Contacts contribute to (3) through the total contact force:

$$\mathbf{f}_{\text{GRF}} = \sum_i \mathcal{J}(\mathbf{q}, \mathbf{r}^i)^\top \mathbf{f}^i, \quad (5)$$

where $\mathcal{J}(\mathbf{q}, \mathbf{r}^i)$ is the Jacobian matrix mapping generalized velocities $\dot{\mathbf{q}}$ to contact force origin \mathbf{r}^i .

Three conditions should be satisfied for the combination of contact locations and forces to be physically plausible. First, the contact force should lie within the friction cone, which we enforce using bound constraints of the optimizer. Second, when a contact is not active, the contact force should be zero. Third, active contacts on the body and the ground should be touching and not sliding.

We now explain how the second condition is satisfied by the optimization. Consider a variable $c^i \in [0, 1]$ whose value correlates with the validity and activity of contact i . Specifically, $c^i = 1$ if the contact is active and $c^i = 0$ if it is not. Unlike previous formulations of CIO, we do not maintain $\{c^i\}$ as separate free variables. Instead, we define them in terms of the normal contact force:

$$c^i = 0.5 \tanh(k_1 \|\mathbf{f}_{\perp}^i\| - k_2) + 0.5. \quad (6)$$

This forces the contact to deactivate when the normal force magnitude is zero. The parameters k_1 and k_2 govern the physical realism of the contact model. The closer (6) is to a step function, the more realistic the contact. However, if (6) is too sharp, the smoothing effect of CIO is reduced and the optimizer may fail to find a good solution. In practice, we first run the optimizer with $k_1 = 10, k_2 = 3$ and increase the sharpness to $k_1 = 20, k_2 = 2$ in a second optimization pass.

The function (6) is incorporated into the contact consistency term J_{CIO} , which also enforces the third condition:

$$J_{\text{CIO}} = c^i (\|\mathbf{r}^i - \mathbf{r}_{\text{body}}^i\|^2 + \|\dot{\mathbf{r}}_{\text{foot}} - \dot{\theta}_{\text{foot}}\|^2). \quad (7)$$

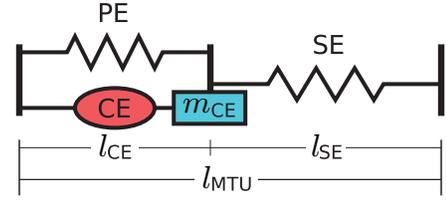


Figure 3: Hill-type musculotendon model. We employ an acceleration model [Millard and Delp 2012] that regulates the acceleration of l_{CE} by introducing a point mass (m_{CE}) between the contractile (CE) and the serial-elastic (SE) elements.

The key idea of this formulation is that during the optimization contact points on the body may invoke ground reaction forces at a distance, through “virtual” contacts with points in the environment. That is, the optimization is allowed to explore intermediate solutions in which active “virtual” contacts ($c^i > 0, \mathbf{r}^i \neq \mathbf{r}_{\text{body}}^i$) generate non-zero ground reaction forces (\mathbf{f}_{GRF}) to reduce J_{EOM} . However, minimizing J_{CIO} minimizes such “virtual” ground reaction forces.

As mentioned above, we found it unnecessary to treat contact activations $\{c^i\}$ as additional variables in the optimization. Furthermore, by not explicitly modeling contact activation trajectories using piecewise constant functions, we remove the assumption that motions are broken up into discrete contact phases [Mordatch et al. 2012a]. Thus the precise timing of the contacts can be adjusted by the optimizer.

4.3 Musculotendon Actuation

Our Hill-type model is illustrated in Figure 3. It consists of a contractile element (CE), a serial-elastic element (SE), and a parallel-elastic element (PE). Let l_{CE} and l_{SE} denote the lengths of the elements connected in series. The total MTU length is then given by $l_{\text{MTU}} = l_{\text{CE}} + l_{\text{SE}}$. Conceptually, CE corresponds to the muscle fiber and SE corresponds to the tendon. Note that while the MTU length l_{MTU} is fully determined by the body pose \mathbf{q} , the length l_{CE} of the contractile element must be recovered separately and is included as a variable in the optimization.

The contractile element is responsible for generating active force F_{CE} given the control signal $a \in [0, 1]$:

$$F_{\text{CE}} = F_{\text{max}} f_l(l_{\text{CE}}) f_v(\dot{l}_{\text{CE}}) a, \quad (8)$$

where f_l, f_v are the force-length and force-velocity relations and F_{max} is an MTU-specific maximum isometric force parameter, which captures the strength of the muscle.

The serial-elastic element exerts a unidirectional nonlinear spring force that is a function of l_{SE} :

$$F_{\text{SE}} = F_{\text{max}} 0.5 \ln(1 + \exp(4(F'_{\text{SE}} - 0.5))), \quad (9)$$

where $F'_{\text{SE}} = 0.04^{-1}(l_{\text{SE}}/l_{\text{slack}} - 1)$, and l_{slack} is an MTU-specific tendon slack length constant. Note that (9) is a smooth approximation to the typical form: $F_{\text{SE}} = F_{\text{max}} \max(0, F'_{\text{SE}})^2$, where force generation can change abruptly (e.g., when the tendon becomes slack) and create difficulties for gradient-based optimization. The approximation is based on Tassa and Todorov [2010].

Following Millard and Delp [2012], the length of the contractile element l_{CE} is governed by its own set of equations of motion, which is derived by assuming that a point mass m_{CE} lies in between the contractile element and the tendon. The point mass (and therefore l_{CE}) is accelerated by the net force acting on it:

$$F_{\text{SE}} - F_{\text{CE}} - F_{\text{PE}} = m_{\text{CE}} \ddot{l}_{\text{CE}}, \quad (10)$$

where F_{PE} models the passive force generated by the MTU, which is also a function of the l_{CE} . We empirically set $m_{CE} = 0.1$.

Auxiliary variables introduced by the i -th MTU into the optimization are the control signal a^i and the length l_{CE}^i of the contractile element. The total force applied by the MTUs on the body is

$$\mathbf{f}_{MTU} = \sum_i \mathbf{m}_i(\mathbf{q}) F_{SE}^i, \quad (11)$$

where \mathbf{m}_i is the moment arm vector of the MTU.

Our moment arms are computed as the derivative of corresponding MTU lengths (l_{MTU}) with respect to q [An et al. 1984]. The MTU length model is based on Geyer and Herr [2010]. While effective for locomotion, their model for MTU-joint pairs with variable moment arms lead to unrealistic negative values for non-locomotion poses (e.g., knee extension angle in a crouch position). To prevent non-physical values while still capturing how MTU length changes with respect to the joint angles, we define

$$l_{MTU} = l_{slack} + l_{opt} + r_0 \rho (P(q - q_{max}) - P(q_{ref} - q_{max})), \quad (12)$$

where $P(x) = 2.0(1 + \exp(-2.1x))^{-1}$ is a scaled logistic function; l_{opt} is a MTU-specific optimal CE length; r_0 and ρ are MTU-specific attachment parameters; q_{ref} is the angle at which the l_{MTU} is in its default length, and q_{max} is the angle at which the moment arm is maximized.

The value of \mathbf{f}_{MTU} is valid when Equation 10 is satisfied. We therefore define the musculotendon force objective in terms of the residuals of (10):

$$J_{MTU} = \sum_i (F_{SE}^i - F_{CE}^i - F_{PE}^i - m_{CE} \ddot{l}_{CE}^i)^2. \quad (13)$$

The contraction dynamics model (10) accounts for the influence of the force generated by the MTU on the acceleration of the muscle mass. While the commonly used equilibrium models have been reported to be more efficient in forward simulations [Millard et al. 2013], we found that explicit modeling of acceleration is effective both for encouraging smooth l_{CE} trajectories and for guiding the optimization to biomechanically consistent solutions.

4.4 Body Integrity

Incorporating reasonable joint limits and preventing collisions between different segments of the body is clearly important for motion realism. Without a collision term, for example, the character can adopt gaits in which the legs pass through each other. We define

$$J_{body} = J_{limit} + J_{collision} \quad (14)$$

$$J_{limit} = \sum_i \{q^i - q_{hlim}^i\}_+^2 + \{q_{lim}^i - q^i\}_+^2 \quad (15)$$

$$J_{collision} = \sum_{a,b \in \mathcal{B}} \{\text{dist}(a,b) - (\text{rad}(a) + \text{rad}(b))\}_+^2, \quad (16)$$

where $\{x\}_+ = 0$ if $x < 0$ and $\{x\}_+ = x$ otherwise. q_{hlim}^i and q_{lim}^i are upper and lower joint limit values based on previous work [Delp et al. 1990; Wang et al. 2012]. \mathcal{B} is the set of rigid segments in our model, all of which are geometrically represented as capsules; $\text{dist}(a,b)$ is the shortest distance between the line segments representing the major axes of a and b , and $\text{rad}(a)$ is the radius of capsule a .

4.5 Economy of Effort

Our effort model consists of two components: a lower body term and an upper body term:

$$J_{effort} = J_{lower} + J_{upper}. \quad (17)$$

The effort J_{lower} exerted by the MTUs in the lower body is modeled using the metabolic energy expenditure model first presented by Anderson [1999], which is the sum of heat released and mechanical work done by the MTUs. In particular, at every time step, the total rate of metabolic energy expenditure is given by

$$J_{lower} = \dot{A} + \dot{M} + \dot{S} + \dot{W}, \quad (18)$$

where \dot{A} is the muscle activation heat rate, \dot{M} is the muscle maintenance heat rate, \dot{S} is the muscle shortening heat rate, and \dot{W} is the positive mechanical work rate. The specific terms are defined in detail by Anderson [1999].

The upper body term J_{upper} penalizes large active motor torques and accelerations generated by the upper body DOFs. As discussed in Section 4.1, active torques are obtained by rearranging the EOM after calculating other generalized forces in the system. We have already defined the contact force (\mathbf{f}_{GRF}) and the MTU force (\mathbf{f}_{MTU}). It remains to specify $\mathbf{f}_{passive}$: forces that capture passive effects of the musculoskeletal structure in the upper body. These are spring-damper forces for the set of upper body DOFs (Q_{upper}) [Wampler and Popović 2009]. These passive forces are designed to be a property of the humanoid model, and are task independent.

Let $f_{passive}^i$ denote the i -th element of $\mathbf{f}_{passive}$. For $q^i \in Q_{upper}$, we define $f_{passive}^i = -k_p q^i - k_d \dot{q}^i$. For all other DOFs, $f_{passive}^i = 0$. Specifically, we set $k_p = 50$, $k_d = 1$ for the neck and back, capturing a weak tendency for the body to stay straight, and $k_p = 0$, $k_d = 1$ for all other upper body joints.

We can now recover \mathbf{f}_{motor} through (3), and define the upper body effort term as follows:

$$J_{upper} = w_m \sum_i (f_{motor}^i)^2 + \ddot{\mathbf{q}}^\top \mathbf{W}_a \ddot{\mathbf{q}}, \quad (19)$$

where $q^i \in Q_{upper}$. This penalizes large active torques and accelerations. \mathbf{W}_a is a diagonal weight matrix, set to 10^{-5} for the neck DOFs, 10^{-4} for the back DOFs, and 10^{-6} for the rest of the DOFs in the upper body. Note the difference between J_{upper} and J_{EOM} : the upper body effort term encodes a preference for motions that require smaller active torques, while the EOM term approximates a hard constraint that there should not be any non-zero active torque in the root and lower body DOFs. Accordingly, we place a much higher weight on the EOM term.

5 Trajectory Optimization

The full set of variables that are optimized for each time step is

$$\mathbf{s}_t = [\mathbf{q}_t \ \mathbf{r}_t \ \mathbf{f}_t \ \mathbf{a}_t \ \mathbf{l}_t],$$

which includes the generalized coordinate values of the humanoid (\mathbf{q}_t), contact locations on the ground ($\mathbf{r}_t = [\mathbf{r}_t^1, \dots, \mathbf{r}_t^{10}]$), ground reaction forces ($\mathbf{f}_t = [\mathbf{f}_t^1, \dots, \mathbf{f}_t^{10}]$), MTU activation signals ($\mathbf{a}_t \in \mathbb{R}^{28}$), and MTU contractile element lengths ($\mathbf{l}_t \in \mathbb{R}^{28}$).

By concatenating the state vectors and their time derivatives over M time steps, we obtain the vector $\mathbf{x} = [\mathbf{s}_1 \ \dot{\mathbf{s}}_1 \ \dots \ \mathbf{s}_M \ \dot{\mathbf{s}}_M]$. The full trajectory is encoded with cubic splines, which serve to interpolate \mathbf{x} and improve computational efficiency. At any time t , the state and its velocities and accelerations can be calculated as $A_t \mathbf{x}$, where A_t is the spline encoding matrix. We thus numerically integrate the objective (2):

$$E(\mathbf{x}) = \sum_t \sum_k w_k J_k(A_t \mathbf{x}). \quad (20)$$

For all experiments, we set $w_{task} = 10$, $w_{EOM} = 10^2$, $w_{CIO} = 10^3$, $w_{body} = 10^3$, $w_{MTU} = 10$, $w_{effort} = 10^{-2.5}$, and discretize the integral at 30 Hz.

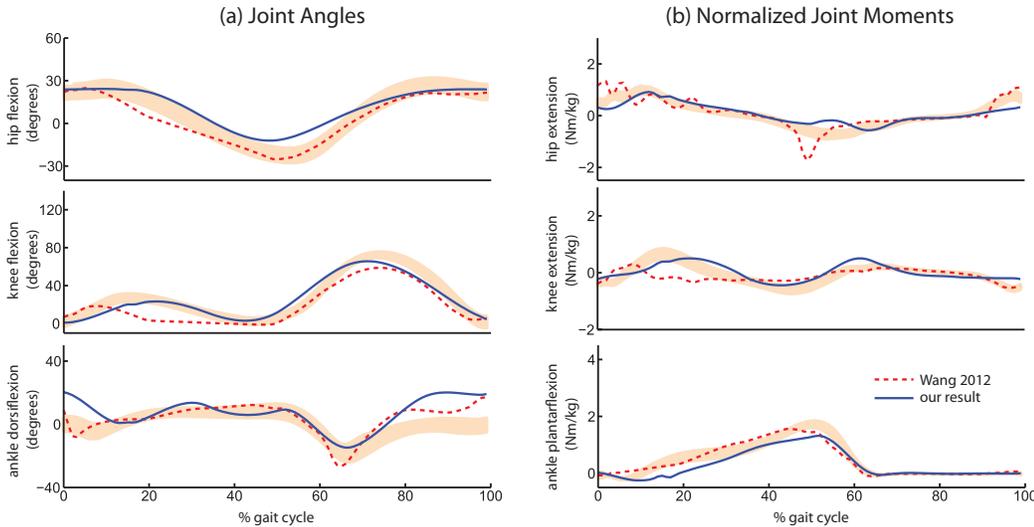


Figure 4: Locomotion at 1.5 m/s. The orange shaded areas represent one standard deviation of ground truth human data. Dashed lines represent walking data synthesized by the approach of Wang et al. [2012]. Solid lines represent our approach. Although our approach does not assume a locomotion-specific control structure, it produces lower limb motion that is quantitatively closer to the ground truth in average standard error. In particular, our approach yields better predictions for the kinematics of the knee.

We use L-BFGS for optimization with numerically approximated gradients. We found the accuracy of the gradients obtained by standard finite differencing to be insufficient, causing premature termination. To obtain more accurate gradients, we use finite differences with complex arithmetic for the approximation [Martins et al. 2003]: $\frac{dJ_k}{ds_n}(\mathbf{s}) \approx \frac{\text{Im}[J_k(\mathbf{s} + i h \mathbf{e}_n)]}{h}$, where h is an extremely small step size (we use 10^{-20}) and \mathbf{e}_n is the n th standard basis vector.

We use seven spline knots per DOF for walking and running results, and 14 for jumping and kicking. The resulting dimensionality of the problem is over 2000, but the optimization does not require example-specific initialization. All examples are initialized to a default “zero” pose throughout the motion duration, where all elements of $\mathbf{q}_t, \mathbf{r}_t, \mathbf{f}_t, \mathbf{a}_t$ are set to zero. The MTU contractile element lengths (l_t) are set to their respective optimal values (l_{opt}). From this initialization, we then proceed to run two optimization passes with different contact model settings. The first pass employs a smooth version of (6) with $k_1 = 10, k_2 = 3$. The result of the first pass is used to initialize a second pass with a sharper setting for (6): $k_1 = 20, k_2 = 2$. Both passes are run for a maximum of 3000 iterations.

Note that the objectives J_k are defined to be functions of the state vector at a single time instant, which means that the objective (20) and its gradient can be evaluated for different time steps in parallel. The independent nature of these terms enables rapid computation of numerical derivatives. As a result, despite the high dimensionality of the problem, all of the presented examples converge within 6 to 10 minutes on a machine with two Intel Xeon X5680 processors.

6 Results

6.1 Basic Locomotion

Locomotion is a particularly common human activity that is driven by the lower body. In the study of human locomotion gaits [Perry and Burnfield 2010], the legs and pelvis are referred to as the “locomotor system” and the upper body is referred to as the “passenger unit.” The passenger unit is “virtually a passive entity that is carried by the locomotor system”; it is also often referred to as the HAT (head, arms, and trunk/torso) to emphasize its secondary role in locomotion [Perry and Burnfield 2010].

We synthesize locomotion by specifying task terms that capture the goal of moving the model’s trunk forward with a target velocity

and a target upright head orientation. Additionally, we require the motion to be periodic and symmetric. The periodicity constraint is implemented by requiring the last spline knot to be identical to the first. This constraint implicitly specifies the duration of the step. We select human-like durations given our desired target velocity based on relations between locomotion velocity and stride length derived from empirical observations [Alexander 2003]. The symmetry constraint restricts the variables associated with the left side of the body to be the same as the corresponding variables from the right side a period later.

Figure 4 compares the sagittal hip, knee, and ankle kinematics and joint moments of a 1.5 m/s walk synthesized by our approach to motion generated by a prior approach that relies on a task-specific control structure [Wang et al. 2012]. Our results match the human ground truth more closely than the results of the prior approach, without reliance on a locomotion-specific control structure. (The joint angle and moment curves we compare to here are part of a dataset consisting of 20 healthy subjects walking and running at multiple speeds on an instrumented treadmill, see Wang et al. [2012] for details.) Specifically, we obtain an average standard error of 1.37 compared to 1.53 for Wang et al. [2012]. Notably, our results better capture the knee trajectories. One likely reason is that the control structure assumed by Wang et al. was overly restrictive in defining the set of allowable control torque patterns. On the other hand, our hip kinematics indicate a lack of hip extension, which results in a shorter stride length relative to human data.

As demonstrated in the supplementary video, running gaits emerge when we specify larger velocity and corresponding durations, without making any other modifications. However, we observe significant differences in knee and ankle angles during swing. In particular, our results exhibit both a lack of knee flexion and a slight over-plantarflexion during early swing phase. A closer look at the gait reveals that the tip of our foot model stays almost parallel over time and hugs the ground plane during the majority of the swing phase. A likely cause of this behavior is the lack of modeling of robustness and stability in our approach. Taking robustness into account would likely yield a gait with higher ground clearance and increased knee flexion during the swing phase, as observed in human running.

To evaluate the effect of musculoskeletal modeling and the fine-grained metabolic energy expenditure model, we replaced J_{lower} by the sum of squared torques of the lower body joints and optimized directly for joint torques in the lower body, without the use of MTUs. The weights on the squared torques were chosen so that

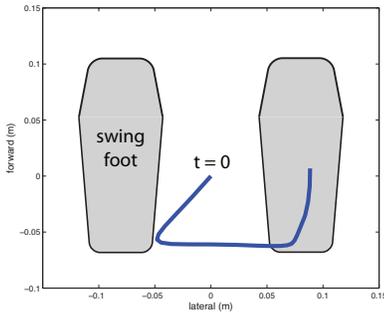


Figure 5: COP trajectory during simulated gait initiation. The shape of the trajectory matches human data.

the contribution of the effort term to the objective is comparable to J_{lower} for the MTU-driven model. The results are shown in the video. Walking synthesized by the purely torque-driven model suffers from severe crouching. The running gaits produced by the restricted model likewise suffer from dramatic artifacts.

The supplementary video also demonstrates walking results generated using the original CIO algorithm, which operates on a simplified character model with massless limbs [Mordatch et al. 2012b]. The original CIO formulation yields unnatural shuffling gaits in which the feet stay close to the ground. For running speeds, the original CIO algorithm fails to converge to reasonable solutions.

6.2 Locomotion Variations

Unlike the approach of Wang et al. [2012], our approach is not limited to fixed-velocity locomotion. By modifying the environment and the simple task objectives, our approach can generate a variety of motions that are appropriate to different scenarios.

Gait initiation. To animate walking initiation, we specify an initial key pose (standing) at time zero, a final key pose at time 1.5 s, and linearly interpolate the target torso velocity from 0 to 1.5 m/s. The final key pose is set to one of the walking poses synthesized by our method. We found the method to be effective regardless of the precise choice of the walking pose. One salient characteristic of human gait initiation is the movement of the center of pressure (COP) [Elble et al. 1994]. As walking is initiated, COP moves backwards towards the heels and laterally towards the swing foot. Then, before the swing foot leaves the ground, the COP shifts laterally to the stance foot and moves from heel to toe as the step is taken. As shown in Figure 5, the motion generated by our approach exhibits this stereotypical COP trajectory. In the accompanying video, we demonstrate this walking initiation animation, as well as similarly generated running initiation.

Incline walking. By varying the slope of the ground, we can synthesize appropriate incline walking gaits. In the video, we demonstrate 1.5 m/s walking for different incline angles (-40° , -15° , 15° , and 40°). A notable feature of human walking on inclined surfaces is the maximum hip extension angle, which corresponds to how far the stance leg stretches behind the pelvis before it enters the swing phase. This angle is largest when we walk on level ground, and decreases for both upslope and downslope walking [Lay et al. 2006]. As seen in the video, the gaits synthesized by our approach reproduce these changes in lower limb kinematics. The peak ankle joint moments in the synthesized motions also increase as slope increases, as observed in human data [Lay et al. 2006].

Locomotion in reduced gravity. Gravitational acceleration has a direct effect on locomotion. It is well-known that a change in gravitational force, as on the Moon, leads to a change in human

locomotion gait. Notably, at typical terrestrial walking speeds, humans adopt gaits that are closer to running when gravity is reduced [Margaria and Cavagna 1964; Cavagna et al. 1998]. Our approach predicts this behavior. We set gravity to $0.16g$, as on the Moon, and synthesized locomotion at 1.5 m/s. As shown in the supplementary video, the synthesized gait resembles the gaits adopted by astronauts on the Moon.

6.3 Jumping and Kicking

Our approach can be used to animate a range of motions. We demonstrate its generality by synthesizing jumping and kicking motions. To animate jumping, we specify the initial and final poses (standing upright) and the trunk’s height mid-jump. With these simple constraints, we synthesize forward, backward, and sideway jumps. As demonstrated in the video, features of human jumping such as knee flexion before lift-off emerge from the optimization. To animate kicking, we specify a target position and velocity for the ankle. As demonstrated in the video, a variety of different kicking motions can be generated by changing these simple constraints.

7 Discussion

We presented a trajectory optimization approach to animating human activities that are driven by the lower body. Our work demonstrates for the first time that a single optimization formulation can produce high-fidelity lower body motion for a range of human activities, without the need for prior motion data or task-specific control structures. This is achieved using a simplified and generalized formulation of CIO that optimizes over continuous contact timings and is applied directly to a detailed humanoid model in which the lower limbs are actuated by musculotendon units.

While effective, CIO does not completely prevent the optimization from converging to poor local minima. In particular, care must be taken to specify weights that trade off the importance of different terms in the objective. Visibly non-physical solutions that appear to “float” in the air can result from setting weights to EOM and CIO terms too low, while low weights on the MTU and effort terms tend to generate unnatural bent-knee gaits. Characterizing weights that can reliably generate a wide range of natural human motion is an interesting avenue for future work.

Like most methods based on spacetime constraints, the resulting control torques and activation signals generated by our work cannot be directly used in a forward simulation. This is both due to the spline encoding of the trajectories and the formulation of physical and biomechanical plausibility through soft constraints. Our trajectories could serve to initialize more complex optimization schedules designed to generate strictly physical trajectories, which would be important for applications outside of animation as well as for synthesizing motions that further approach human ground truth.

The focus of this work was on the fidelity of the synthesized lower body motion. The movement of the upper body in our results does not match human ground truth in general. A likely reason is that our simple spring-damper model for upper body actuation does not properly represent the passive effects of the musculoskeletal structure in the upper body. A natural direction for future work is to develop a musculoskeletal model for the entire body, including the upper body. The highly detailed model developed by Lee et al. [2009] provides a promising starting point, but some work is required to develop a simplified model that supports effective trajectory optimization. Such development would enable high-fidelity animation of activities that center on the upper body, such as reaching and pointing, as well as activities that require precise coordination of the entire body, such as ball sports, combat sports, and parkour.

Acknowledgements

We thank Matthew Millard, Tim Dorn, and the anonymous reviewers for helpful comments. This work was funded in part by NSF grant IIS-1017938, NIH grant R01NS073120, and an AWS research grant. This research was conducted in conjunction with the Intel Science and Technology Center for Visual Computing.

References

- AL BORNO, M., DE LASA, M., AND HERTZMANN, A. 2013. Trajectory optimization for full-body movements with complex contacts. *IEEE Trans. Vis. Comput. Graph.* 19, 8.
- ALEXANDER, R. M. 2003. *Principles of Animal Locomotion*. Princeton University Press.
- AN, K. N., TAKAHASHI, K., HARRIGAN, T. P., AND CHAO, E. Y. 1984. Determination of muscle orientations and moment arms. *Journal of Biomechanical Engineering* 106, 3.
- ANDERSON, F. C., AND PANDY, M. G. 1999. A dynamic optimization solution for vertical jumping in three dimensions. *Comput. Methods in Biomechanics and Biomedical Engineering* 2, 3.
- ANDERSON, F. C., AND PANDY, M. G. 2001. Dynamic optimization of human walking. *J. Biomech. Eng.* 123, 5.
- ANDERSON, F. C. 1999. *A dynamic optimization solution for a complete cycle of normal gait*. PhD thesis, U. of Texas at Austin.
- CAVAGNA, G. A., WILLEMS, P. A., AND HEGLUND, N. C. 1998. Walking on Mars. *Nature* 393, 6686.
- COHEN, M. F. 1992. Interactive spacetime control for animation. In *Proc. SIGGRAPH*.
- COROS, S., BEAUDOIN, P., AND VAN DE PANNE, M. 2010. Generalized biped walking control. *ACM Trans. Graph.* 29, 4.
- DE LASA, M., MORDATCH, I., AND HERTZMANN, A. 2010. Feature-based locomotion controllers. *ACM Trans. Graph.* 29, 4.
- DELP, S. L., LOAN, P., HOY, M. G., ZAJAC, F. E., TOPP, E. L., AND ROSEN, J. M. 1990. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Transactions on Biomedical Engineering* 37, 8.
- ELBLE, R. J., MOODY, C., LEFFLER, K., AND SINHA, R. 1994. The initiation of normal walking. *Movement Disorders* 9, 2.
- EREZ, T., AND TODOROV, E. 2012. Trajectory optimization for domains with contacts using inverse dynamics. In *Proc. IROS*.
- FALOUTSOS, P., VAN DE PANNE, M., AND TERZOPOULOS, D. 2001. Composable controllers for physics-based character animation. In *Proc. SIGGRAPH*.
- FANG, A. C., AND POLLARD, N. 2003. Efficient synthesis of physically valid human motion. *ACM Trans. Graph.* 22, 3.
- GEYER, H., AND HERR, H. 2010. A muscle-reflex model that encodes principles of legged mechanics produces human walking dynamics and muscle activities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18, 3.
- HARRIS, C. M., AND WOLPERT, D. M. 1998. Signal-dependent noise determines motor planning. *Nature* 394.
- HODGINS, J. K., WOOTEN, W. L., BROGAN, D. C., AND O'BRIEN, J. F. 1995. Animating human athletics. In *Proc. SIGGRAPH*.
- LAY, A. N., HASS, C. J., AND GREGOR, R. J. 2006. The effects of sloped surfaces on locomotion: a kinematic and kinetic analysis. *Journal of Biomechanics* 39.
- LEE, S.-H., SIFAKIS, E., AND TERZOPOULOS, D. 2009. Comprehensive biomechanical modeling and simulation of the upper body. *ACM Trans. Graph.* 28, 4.
- LIU, C. K., AND POPOVIĆ, Z. 2002. Synthesis of complex dynamic character motion from simple animations. *ACM Trans. Graph.* 21, 3.
- LIU, C. K., HERTZMANN, A., AND POPOVIĆ, Z. 2005. Learning physics-based motion style with nonlinear inverse optimization. *ACM Trans. Graph.* 24, 3.
- MARGARIA, R., AND CAVAGNA, G. A. 1964. Human locomotion in subgravity. *Aerospace Medicine* 35.
- MARTINS, J. R. R. A., STURDZA, P., AND ALONSO, J. J. 2003. The complex-step derivative approximation. *ACM Trans. Math. Softw.* 29, 3.
- MILLARD, M., AND DELP, S. L. 2012. A computationally efficient muscle model. In *ASME Summer Bioeng. Conf.*
- MILLARD, M., UCHIDA, T., SETH, A., AND DELP, S. L. 2013. Flexing computational muscle: Modeling and simulation of musculotendon dynamics. *Journal of Biomech. Eng.* 135.
- MORDATCH, I., POPOVIĆ, Z., AND TODOROV, E. 2012. Contact-invariant optimization for hand manipulation. In *Proc. SCA*.
- MORDATCH, I., TODOROV, E., AND POPOVIĆ, Z. 2012. Discovery of complex behaviors through contact-invariant optimization. *ACM Trans. Graph.* 31, 4.
- PERRY, J., AND BURNFIELD, J. 2010. *Gait Analysis: Normal and Pathological Function*. Slack Incorporated.
- POPOVIĆ, Z., AND WITKIN, A. 1999. Physically based motion transformation. In *Proc. SIGGRAPH*.
- SAFONOVA, A., HODGINS, J. K., AND POLLARD, N. 2004. Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. *ACM Trans. Graph.* 23, 3.
- SRINIVASAN, M., AND RUINA, A. 2006. Computer optimization of a minimal biped model discovers walking and running. *Nature* 439.
- SULEJMANPAŠIĆ, A., AND POPOVIĆ, J. 2005. Adaptation of performed ballistic motion. *ACM Trans. Graph.* 24, 1.
- TASSA, Y., AND TODOROV, E. 2010. Stochastic complementarity for local control of discontinuous dynamics. In *Proc. RSS*.
- TODOROV, E., EREZ, T., AND TASSA, Y. 2012. MuJoCo: A physics engine for model-based control. In *Proc. IROS*.
- WAMPLER, K., AND POPOVIĆ, Z. 2009. Optimal gait and form for animal locomotion. *ACM Trans. Graph.* 28, 3.
- WANG, J. M., HAMNER, S. R., DELP, S. L., AND KOLTUN, V. 2012. Optimizing locomotion controllers using biologically-based actuators and objectives. *ACM Trans. Graph.* 31, 4.
- WITKIN, A. P., AND KASS, M. 1988. Spacetime constraints. In *Proc. SIGGRAPH*.
- YIN, K., LOKEN, K., AND VAN DE PANNE, M. 2007. SIMBICON: Simple biped locomotion control. *ACM Trans. Graph.* 26, 3.