Estimating Contact Dynamics

Marcus A. Brubaker Leonid Sigal David J. Fleet
Department of Computer Science
University of Toronto
{mbrubake,ls,fleet}@cs.toronto.edu

Abstract

Motion and interaction with the environment are fundamentally intertwined. Few people-tracking algorithms exploit such interactions, and those that do assume that surface geometry and dynamics are given. This paper concerns the converse problem, i.e., the inference of contact and environment properties from motion. For 3D human motion, with a 12-segment articulated body model, we show how one can estimate the forces acting on the body in terms of internal forces (joint torques), gravity, and the parameters of a contact model (e.g., the geometry and dynamics of a spring-based model). This is tested on motion capture data and video-based tracking data, with walking, jogging, cartwheels, and jumping.

1. Introduction

Motion and interaction with the environment are fundamentally intertwined. The motion of an object is determined in part by its contact with the environment, and conversely, motion is a rich source of information about contact, much like the locations of people are informative about the ground plane [8, 9]. Prior knowledge of an inelastic ground plane has been incorporated in physics-based models for people tracking with encouraging results [4, 5, 18]. The inference of surface contact from motion is, however, unexplored in computer vision.

We formulate a general physics-based model of motion and contact for articulated bodies. Our principal results are general, but our primary concern is with human pose tracking. We show how one can explain motion and contact by decomposing the net forces acting on a body in terms of external forces (contact and gravity) and internal forces (muscle actuations at joints). We explore the intimate relation between internal and external forces, and we present a method to simultaneously recover both from observed motion. At the same time we show how one can estimate the parameters of a damped, elastic model of surface contact.

The resulting approach provides information about the timing and location of contact. This includes, but it is not restricted to, contact with the ground plane. Similarly, our model explicitly allows for contact at arbitrary locations over the surface of body, e.g., as someone leans on a table, falls down, or performs a cartwheel (see Fig. 5). The parametric contact model also provides information about material properties such as stiffness and damping; these are useful for prediction and control, and of course for understanding intrinsic surface properties.

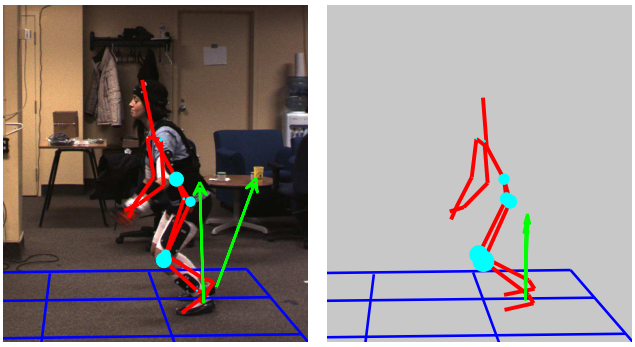


Figure 1. **Recovery of Contact Forces and Joint Torques:** These two figures show the skeleton of the subject (red), the joint torques (cyan disks), a planar contact surface (blue grid), and the ground reaction forces (green arrows) acting on the body as estimated from synchronized video (left) and motion capture data (right) for a jumping motion. The radii of the cyan disks are proportional to joint torques, and the green arrows are proportional to the ground reaction forces acting on the body.

In the process of recovering contact properties, our formulation effectively decomposes the forces acting on a body into external forces and internal joint torques (e.g., see Fig. 1). Such external and internal forces are valuable for biomechanical research on human locomotion, and for clinical applications where expensive and cumbersome force plates are the principal source of existing data. Internal joint torques should be useful for developing physics-based models of human motion for tracking, and they may also form a useful basis for identifying motion and scene interpretation, like inferring that a person is carrying a heavy object.

We demonstrate the approach on motion capture data and video-based 3D pose tracking. We consider contact on both hands and feet, and with several different activities, including walking, jogging, jumping (Fig. 1), and gymnastics.

2. Related Work

Context is important for detecting and tracking people in images. It has been shown, for example, that prior knowledge of scene geometry significantly improves people detection, and the detection of people is useful for estimating scene geometry, assuming prior information about human heights and that people are supported by the ground plane (e.g., [8, 9]). With prior knowledge of foot contact on the peddles of a bicycle Rosenhahn et al [15] showed how to enforce kinematic constraints to improve 3D pose tracking.

The interplay between motion and contact is naturally expressed in multi-body dynamics. Interaction and contact are inherent in physics-based models. So one might hope that they would facilitate the simultaneous inference of motion and interaction. Recent physics-based methods for 3D people tracking incorporate an explicit representation of the ground plane and contact dynamics [4, 5, 18]. Nevertheless, rather than inferring contact properties (e.g., ground geometry and elasticity) during pose inference, they assume that these properties are known a priori.

While not extensive, there is other related work in computer vision and in computer graphics that has inspired our research. At a high-level, physics-based models and contact have been used for image interpretation of simple scene domains [3, 12, 17]. At a lower level, using modal analysis Pentland and Williams considered the inference of material properties from two non-rigid colliding bodies, assuming that the time-varying shapes of the two bodies are given [13]. Bhat et al [2] estimate physical properties of rigid objects in free flight but do not address the issue of contact.

Physics-based animation with spring-based contact models is common in computer graphics [11, 14, 20]. In this paper we adopt a similar class of models but, rather than hand-specifying the contact geometry, the times of contact, and the spring parameters for individual classes of motion, we automatically infer the contact model parameters from the observed kinematics.

3. Motivating Example

As a motivating example, Fig. 2(top) depicts a video of a ball dropped onto a surface. The height of the ball is tracked, as shown in the first row of plots in Fig. 2. By measuring the accelerations, the net force acting on the ball (up to mass) is determined by Newton's second law of motion. These forces can be decomposed as the sum of forces due to gravity and to contact (shown in the middle row of plots in Fig. 2). This experiment was done with the same ball dropped onto a hard table and then onto a soft mouse pad (respectively the left and right plots in Fig. 2). In both cases the occurrence of contact is clearly evident by virtue of the spike in external forces. The somewhat smaller magitude and broader temporal duration of the contact forces on the right plot are consistent with the greater compliance and damping of the softer surface.

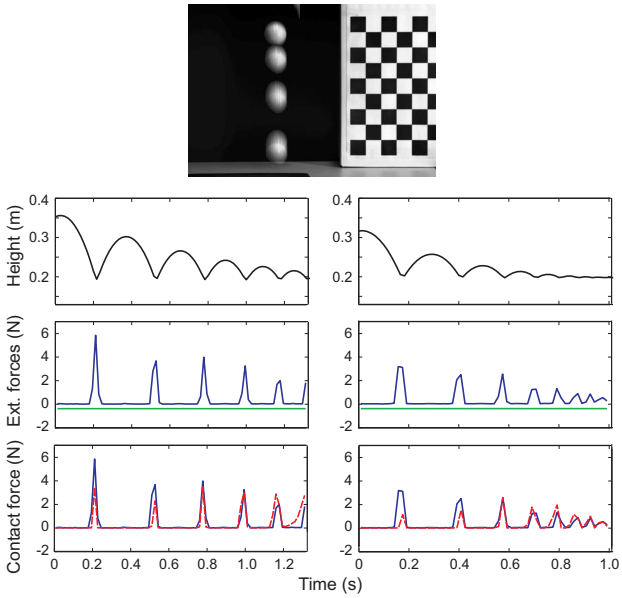


Figure 2. The image depicts a ball dropping onto a table. The plots show results for a ball bouncing onto a hard surface (left) and onto a soft mouse pad (right). The top plots show height as a function of time as a ball bounces. The net forces are then decomposed into gravitational forces (green) and contact forces (blue) in the middle two plots. The bottom plots compare contact forces to those predicted by a damped eleastic contact model (dashed red).

Based on these forces one can infer properties of a simple contact model comprising a surface of unknown height which, through a sigmoidal non-linearity, modulates a linear spring of unknown stiffness and damping (e.g., see Fig. 3). The model parameters are optimized to minimize the difference between the measured net contact forces and those produced by the model (Fig. 2(bottom) plots). For the two surfaces in Fig. 2, the optimization yields stiffness values of approximately 24 and 15 N/m (Newtons per meter), indicating that the table top is considerably harder (stiffer) than the mousepad. The damping for the soft surface was found to be marginally greater, and the heights of the two surfaces were extremely close to ground truth.

This example demonstrates that motion contains information about surface contact. Below we generalize this idea to surfaces acting on articulated human motion. Obviously, coping with human motion is much more challenging than a bouncing ball. Far from a simple point mass, the human body is a complex articulated body for which the dynamics are the result of forces and torques on each body part, which are constrained by rotational joints. The net force on the body must be explained in terms of internal forces (e.g., joint torques) in addition to external forces (e.g., gravity and contact). Finally, unlike the model of the ball, contact between a person and the environment can occur at one or more points over the entire the surface of the body.

4. Physics of Motion and Contact

Consider an articulated body consisting of P parts with N degrees of freedom (DoF) comprising $N-6$ joint angles and 6 DoFs for the global position and orientation of the root of the kinematic tree (usually the pelvis). A Lagrangian formulation expresses the configuration of the body in terms of its *generalized coordinates*, $\mathbf{q} \in \mathbb{R}^N$, and N second-order differential equations that govern its motion:

$$\mathcal{M}(\mathbf{q})\,\ddot{\mathbf{q}} \;=\; \mathcal{F}(\mathbf{q},\dot{\mathbf{q}}) + \mathcal{A}(\mathbf{q},\dot{\mathbf{q}}) \qquad (1)$$

where $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ denote the first and second time derivatives of $\mathbf{q}$, $\mathcal{M}$ is called a generalized mass matrix, $\mathcal{F}$ denotes a vector of generalized forces acting on the N DoFs (including contact, gravity and joint torques), and $\mathcal{A}$ comprises all other terms including those necessary to enforce joint constraints. These equations can be derived in different ways, e.g., the TMT method described in [19, 20].

Our goal is to *explain* the N generalized accelerations in $\ddot{\mathbf{q}}$. To begin we first express $\mathcal{F}$ in terms of the $N-6$ internal torques, $\tau_{int} \in \mathbb{R}^{N-6}$, induced by muscle actuations at the joints, and the external forces acting on the body:

$$\mathcal{F}(\mathbf{q},\dot{\mathbf{q}}) \;=\; A_{int}\tau_{int} \;+\; \tau_{ext}(\mathbf{q},\dot{\mathbf{q}}) \qquad (2)$$

where the matrix A_{int} maps the joint torques into the vector of N generalized forces (e.g., $A_{int} = [\mathbf{I}_{N-6}\ \mathbf{0}]^T$). Given just $N-6$ linear DoFs for the joint torques in (2) one cannot fully model the generalized forces in (1). That is, with only joint torques the model is underactuated and will not be able to reproduce $\ddot{\mathbf{q}} \in \mathbb{R}^N$ in general. External forces must be taken into account. Indeed, estimates of internal torques depend strongly on the external forces (e.g., knees are passive when a person hangs freely by their hands, but stiff while standing).

4.1. External Forces

A natural and convenient way to parameterize external forces is through forces (torques) acting on (about) the centers of mass of each body part. This is straightforward as there is a linear mapping from part-specific forces and torques to generalized forces. We can further decompose external forces into those due to gravity $\mathbf{f}_g$, and other, as of yet unexplained forces $\mathbf{f}_e$:

$$\tau_{ext}(\mathbf{q},\dot{\mathbf{q}}) \;=\; F(\mathbf{q})\,[\,\mathbf{f}_g \;+\; \mathbf{f}_e(\mathbf{q},\dot{\mathbf{q}})\,] \qquad (3)$$

where $\mathbf{f}_g$ and $\mathbf{f}_e$ are vectors in $\mathbb{R}^{6P}$, comprising 3 forces and 3 torques for each of P body parts. The state dependent Jacobian matrix F maps the forces (torques) on parts into generalized forces. Finally, note that $\mathbf{f}_e$ is, in general, a (non-linear) function of $\mathbf{q}$, $\dot{\mathbf{q}}$ and scene parameters (e.g., the locations of contact surfaces).

Contact Forces: In this paper we assume that the contact forces arise due to contact between the body and fixed surfaces in the scene. For many hard surfaces contact is effectively inelastic and velocity is discontinuous at contact (e.g., [4]). While such models are appealing in their realism, they are challenging computationally; they require explicit detection of contact events, and often result in difficult, mixed discrete-continuous optimization problems. In contrast, here we adopt a continuous contact model, similar to those employed in space-time optimization (e.g., [11]). As a result we estimate the contact model parameters using efficient, gradient-based optimization techniques.

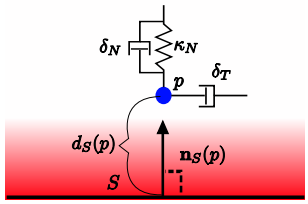


Figure 3. **Continuous Model of Contact:** Springs are modulated by two sigmoids, one of distance from the surface and the other of force. The distance sigmoid is illustrated here as a gradient with brighter red indicating sigmoid activation.

Our model for the force at a point p on the body, due to contact with surface S, is a damped, linear spring modulated by two sigmoidal functions. One sigmoid prevents forces from being applied when p is far from the surface S. The other sigmoid prevents forces from pulling points on the body towards the surface (i.e., sticky ground forces). As depicted in Fig. 3, the model requires $d_S(p)$, the signed shortest distance (positive for outside/above, negative for inside/below, in meters) from p to S, and $\mathbf{n}_S(p)$ the unit normal of S at the point on S closest to p. The model contact force acting on p, denoted $f_c \in \mathbb{R}^3$, is given by

$$f_c(p,\dot{p},\theta_S) = h(-60\,d_S(p))\,h(5n_c(p))\,[n_c(p)\mathbf{n}_S(p)+\mathbf{t}_c(p)] \qquad (4)$$

where $h(x) = \frac{1}{2}(1+\tanh(x))$ is the sigmoidal function, $n_c(p)$ is the signed magnitude of the normal force due to the linear spring alone, and $\mathbf{t}_c(p)$ is the tangential force of the frictional damper. The normal spring force is given by

$$n_c(p) \;=\; -\kappa_N\,(d_S(p)-1) - \delta_N\dot{p}^T\mathbf{n}_S(p) \qquad (5)$$

where κ_N denotes stiffness, and δ_N denotes the normal damping constant. The tangential force is given by

$$\mathbf{t}_c(p) \;=\; -\delta_T\,(\dot{p} - (\mathbf{n}_S(p)^T\dot{p})\mathbf{n}_S(p)) \qquad (6)$$

where δ_T is a damping constant, and $\dot{p}$ is the velocity of p. Finally, θ_S denotes the vector of surface parameters (e.g., the position and orientation of a plane, the spring stiffness κ_N, and the damping constants, δ_N and δ_T). The remaining constants in the model are somewhat arbitrary but the same values have worked well in all of our experiments.

The non-linear spring described above is applied independently at a set of contact points defined over the surface

of the articulated body. When a force is applied to a contact point on the body, it induces both a force at, and an angular torque about, the center of mass of the corresponding part. The net external force caused by contact between P contact points and S surfaces, denoted $\mathbf{f}_s \in \mathbb{R}^{6P}$, can be written as

$$\mathbf{f}_s(\mathbf{q}, \dot{\mathbf{q}};\, \theta) \;=\; \sum_{j=1}^{S} \sum_{k=1}^{P} A_k(\mathbf{q})\, f_c(p_k(\mathbf{q}), \dot{p}_k(\mathbf{q}, \dot{\mathbf{q}}), \theta_j) \quad (7)$$

where $\theta = \{\theta_j\}_{j=1}^{S}$ are the parameters of the surfaces and $A_k(\mathbf{q})$ maps the force applied at point k into a force and torque on the part containing point k.

Substituting $\mathbf{f}_s$ for $\mathbf{f}_e$ in (3) we obtain a model for external forces in terms of contact and gravity. A natural way to estimate the joint torques and the contact model parameters is then to minimize the discrepancy between the observed motion and that generated by simulating the equations of motion. This is, however, extremely challenging due to noise and the existence of local minima. In our experience it is very difficult to obtain satisfactory results with this approach, even assuming a single planar surface for the ground plane. Accordingly, we consider alternative models.

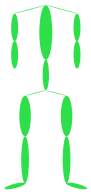
Root Forces: Imagine that arbitrary forces and torques could be applied to the root of the kinematic tree (or any other body part). This provides 6 independent DoFs which complement the $N{-}6$ internal joint torques. Then, the combined joint torques and root forces would be sufficiently rich to exactly account for the the N-dimensional accelerations. Accordingly, there should be no accumulated error in the output of a simulator that uses the estimated forces. This greatly simplifies the estimation problem by decoupling the estimation of the forces at each instant in time. This, therefore, avoids the need for optimization via simulation.

The problem with this model is obvious. It is not physically meaningful for almost all scenes of any interest.

Model of External Forces: The model of external forces we use below is a combination of surface contact (7), gravity and root forces, that is:

$$\tau_{ext}(\mathbf{q}, \dot{\mathbf{q}}) = \; F(\mathbf{q})\, [\, \mathbf{f}_g \;+\; \mathbf{f}_s(\mathbf{q}, \dot{\mathbf{q}}) \;+\; A_{root}\mathbf{f}_{root}\,] \quad (8)$$

where $\mathbf{f}_{root} \in \mathbb{R}^6$ is the root force vector, and matrix $A_{root} \in \mathbb{R}^{6P\times 6}$ maps the 6 components of the root forces into the forces and torques of the part to which root forces are applied. The addition of root forces allows us to decouple the estimation problem at different time steps. But the model is redundant; i.e., there are multiple ways to reproduce the accelerations. Our objective below is to explain as much of the accelerations as possible with the contact model. The root forces are only used to explain residual accelerations not accounted for by joint torques, gravity or the contact model; i.e., to model *noise* not accounted for by the contact model.



Joint	DoFs
Ankle	2
Knee	1
Hip	3
Pelvis-Torso	3
Shoulder	3
Elbow	1

Figure 4. 3D Articulated model of the human body.

4.2. Parameter Estimation

In the experiments below we assume a single planar contact surface parameterized by its normal and its distance from the origin. We seek to estimate the parameters θ that minimize the magnitude of the root forces, $\mathbf{f}_{root}$. Substituting (8) into (2) and subsequently (2) into (1) produces:

$$\hat{A}(\mathbf{q}) \begin{bmatrix} \tau_{int} \\ \mathbf{f}_{root} \end{bmatrix} = \mathcal{M}(\mathbf{q})\ddot{\mathbf{q}} - \mathcal{A}(\mathbf{q}, \dot{\mathbf{q}}) - F(\mathbf{q})[\mathbf{f}_g + \mathbf{f}_s(\mathbf{q}, \dot{\mathbf{q}}; \theta)] \quad (9)$$

where $\hat{A}(\mathbf{q}) = [A_{int},\; F(\mathbf{q})A_{root}] \in \mathbb{R}^{N\times N}$. This yields closed-form expressions for $\mathbf{f}_{root}$ and τ_{int}, as functions of θ, at every time step.

We solve for θ by mimimizing an objective function equal to the sum of root force magnitudes through time:

$$O(\theta) \;=\; \sum_{t} \|\mathbf{f}_{root}(t, \theta)\|^2 \quad (10)$$

where $\mathbf{f}_{root}(t, \theta)$ are the root forces at time t with contact model parameters θ. We also impose constraints on parameters $\kappa_N \in [1, 20]$ and $\delta_N, \delta_T \in [0.1, 20]$. Small values for these parameters produce an inactive contact model and large values are implausible given the data sampling rates. The objective $O(\theta)$ is differentiable with respect to θ, so we use the L-BFGS-B optimizer [21] to minimize (10) subject to the bound constraints. Once estimated, we use θ to compute the internal torques τ_{int} at each time.

The objective function requires $\mathbf{q}$, $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ at each time. To estimate $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ given a pose sequence $\mathbf{q}_1, \ldots, \mathbf{q}_T$ we use forward differences, $\dot{\mathbf{q}}_t = (\mathbf{q}_{t+1} - \mathbf{q}_t)/\Delta$ and $\ddot{\mathbf{q}}_t = (\dot{\mathbf{q}}_{t+1} - \dot{\mathbf{q}}_t)/\Delta$ for a time-step Δ. This choice of derivative estimator is consistent with the first-order Euler integration $\mathbf{q}_{t+1} = \mathbf{q}_t + \Delta\dot{\mathbf{q}}_t$, $\dot{\mathbf{q}}_{t+1} = \dot{\mathbf{q}}_t + \Delta\ddot{\mathbf{q}}_t$. Thus, forces that reproduce such accelerations will automatically reproduce the motion when integrated with this method.

5. Experiments

Our approach to estimating internal torques and contact properties reduces to two steps: (1) Estimate velocity and acceleration; (2) Estimate the contact model parameters and internal torques by minimizing root forces. We have applied the algorithm to 3D mocap data and to the output of a 3D people tracking algorithm. In both cases we estimate ground contact properties and ground reaction forces with the 12-part, 23-DoF 3D articulated model depicted in Fig. 4.

Figure 5. **Cartwheel Sequence:** The motion capture and estimated forces are shown for a cartwheel (from right to left). Joint torques and ground reaction forces are indicated as in Fig. 1. Joint torques are small as the body initially stands comfortably and then as the the legs rotate (almost passively) over the torso. They are larger during landing as the feet collide with the ground and the body regains balance.

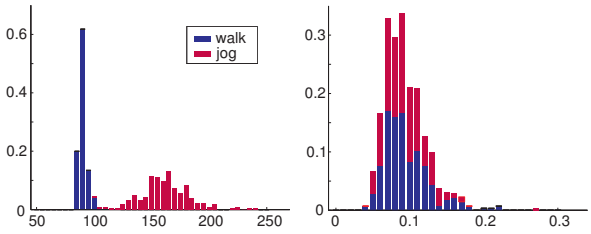


Figure 6. (left) Distribution over average root force magnitude per frame for 500 sequences of walking (blue) and jogging (red), when no contact model is present. External forces in jogging are much larger than those in walking. (right) Distribution over the fraction of root forces *not* explained by the contact model. For both motions the contact model explains approximately 90% of the root forces shown in the left plot. (Forces normalized by subject mass).

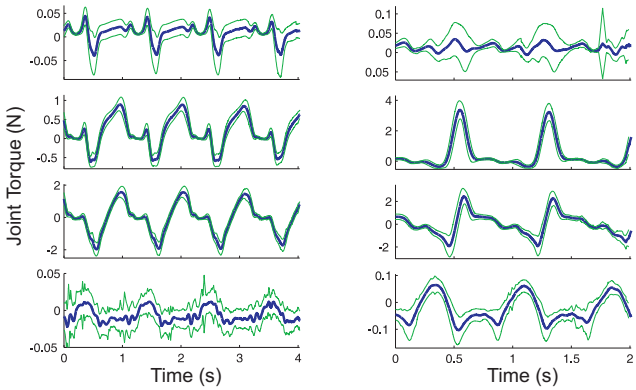


Figure 7. **Consistency in Walking and Running:** Estimated joint torques (in Newtons) for the ankle, knee, hip and shoulder (from top to bottom), based on 250 samples of walking (left) and 250 samples of running (right) from 115 subjects. Bold blue curves show mean torque (in Newtons) as a function of time (in seconds). Light green curves show one standard deivations. Despite variations in morphology, style, speed and steplength, the estimated torques are consistent.

Joint angles are represented with quaternions (see Appendix A for details). Body segment lengths are estimated from the mocap data for each subject, and then combined with standard biomechanical data [6] to determine mass and inertial properties. Eight contact points are placed around the end of each body segment, except for the feet, which have four contact points on the bottom.

5.1. Motion Capture Data

We have tested the algorithm on 120 subjects performing a wide range of activities, including walking, jogging, jumping, hopscotch, and cartwheels. The estimated ground forces and torques are illustrated in Fig. 1 for one such jumping motion (joint torques in cyan, ground reaction forces in green, ground plane in blue). Fig. 5 shows results on a cartwheel sequence.

Fig. 6(left) shows the distribution of average root force magnitudes per frame for several hundred walking and jogging motions when there are no contact model forces (i.e., remove $\mathbf{f}_s$ from (9) before solving for $\mathbf{f}_{root}$). Not surprisingly, these root forces for jogging are much larger than for walking. Fig. 6(right) shows the fraction of these root forces that remain aften the contact model is incorporated. For both walking and jogging, the contact model is explaining approximately 90% of root force magnitudes.

We also find that joint torque estimates are remarkably consistent over different subjects for running and walking. Based on approximately 3 trials of jogging and walking for each of 100 subjects, Fig. 7 shows the time-varying distribution of joint torques for the ankle, knee, hip and shoulder (mean in blue; one standard deviation in green). The contact models are also consistent. Over all walking and running data, the mean angle of the ground with respect to the mocap ground plane (our ground truth, defined as $Z = 0$), is $-0.058°$, with standard deviation $1.11°$. While the contact model does not explicitly define the location of the ground, the parameters do indicate its height. That is, we take the ground height to be that at which ground forces exactly cancel force due to gravity for a motionless subject. Relative to the mocap ground plane, the mean resting height is estimated to be 6.8cm with a standard deviation of 1.13cm.

While ground plane geometry is consistent across subjects and motions, the contact parameters are not. Fig. 8 shows a scatter plot of the estimated stiffness and the normal damping constants. Values for men and women are similarly distributed, but jogging (crosses) consistently produces higher stiffness and damping values than walking (circles). Stiffness and damping values are also correlated. We computed the ratio of the average jogging stiffness to the average walking stiffness for each subject and found an average ratio of 3.59 with a standard deviation of 1.55; i.e., jogging requires a consistently stiffer ground model.

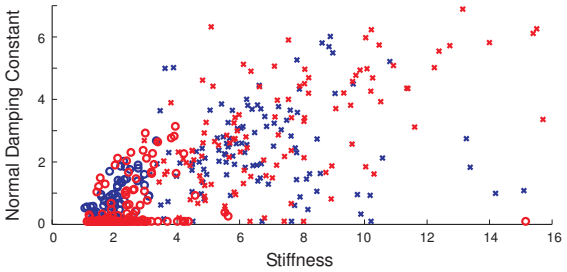


Figure 8. Scatter plot of stiffness and the (normal) damping parameter for men (blue) and women (red), walking (circles) and jogging (crosses). Parameters are normalized by body mass.

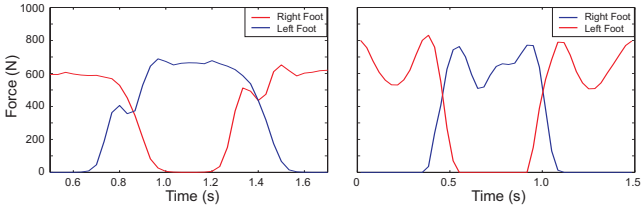


Figure 9. **Comparison of ground reaction force to force plate data:** (right) Force plate data for a walking motion. (left) Ground reaction forces estimated from mocap of a different subject.

Fig. 9(left) shows the estimated vertical (normal) ground reaction force on the feet for three strides of walking. Fig. 9(right) shows vertical ground reaction forces measured with a force plate (for a different subject). The timing and magnitudes are similar, but the shape of the curves differ. We believe this is due to the (fixed) steepness of the sigmoids in (4), and the placement of contact points only near the heel of the foot, making toe-off hard to express.

Finally, to test the generality of the method it was applied to gymnastic motions, namely, jumping, hopscotch (successive short jumps on one and two feet) and cartwheels. Fig. 5 depicts the cartwheel sequence, along with estimated joint torques and ground forces. Note that the ground reaction forces applied to the hands and feet have similar magnitudes. One can also see that the legs are nearly passive as they rotate over the body.

5.2. Video-Based Human Tracking

The algorithm can also be applied to 3D poses estimated from video. Our pose tracker used two views of a subject (one roughly sagittal and one roughly frontal). The cameras were stationary and calibrated with a mocap system to enable a comparison of estimated contact models and internal torques with those obtained using mocap (see Fig 11).

3D pose tracking was achieved with an Annealed Particle Filter (APF) [7] using the implementation of Balan et al. [1]. The likelihood used a probabilistic background model and the output of the 2D region-based WSL tracker [10]. The background model comprised the mean color image and intensity gradient, along with a single 5D covariance matrix (estimated over the entire image). Typical measurements from the WSL tracker are shown in Fig. 10, the likelihood for which was a truncated Gaussian (for robustness) on the 2D joint locations. The pose tracker did not employ a prior model other than weak joint limit constraints (learned from mocap) and interpenetration constraints. Following [7], the tracker used a first-order diffusion process whose variance was loosely learned from mocap (based on interframe differences in joint angles). All experiments used an APF with 200 particles per layer and 10 layers.

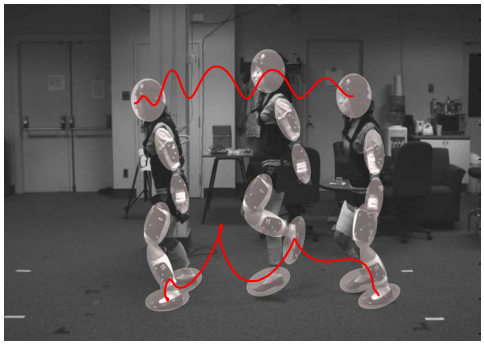


Figure 10. **WSL Tracks:** This depicts a cropped, time-lapse image sequence where 7 regions were tracked, for input to a 3D people tracker. Red curves depict 2D tracks for the head and left foot.

The performance of the tracker and the estimated dynamics are demonstrated in Fig 11, for walking, a long jump, and hopscotch (alternating jumps on one or two feet). While the tracker results are noisy they were sufficient to estimate the parameters of the contact model, as well as the internal torques in all cases. Rows 7 and 8 of Fig. 11 illustrate that the recovered ground reaction forces and internal joint torques correlate well with those recovered from the synchronized mocap. Not surprisingly, due to tracking noise, the joint torques are somewhat overpowered. Nevertheless their behavior is consistent with the mocap.

6. Discussion and Future Work

This paper descibes a method for recovering joint torques and a parametric contact model from motion. Experimental results demonstrate the validity, generality and robustness of the algorithm and contact model for a wide range of subjects and motions, from video and mocap data. There are, however, many interesting issues remaining for future work. These include the investigation of ambiguities when estimating contact with multiple surfaces, better models of human morphology to yield more accurate estimates of biomechanically interesting quantities, and the inclusion of different contact models (e.g., to allow active, or grasping, contact). Taken with previous work showing the benefits of physics-based pose tracking [4, 5, 18], it would also be natural to consider the recovery of contact properties and surface geometry during tracking.

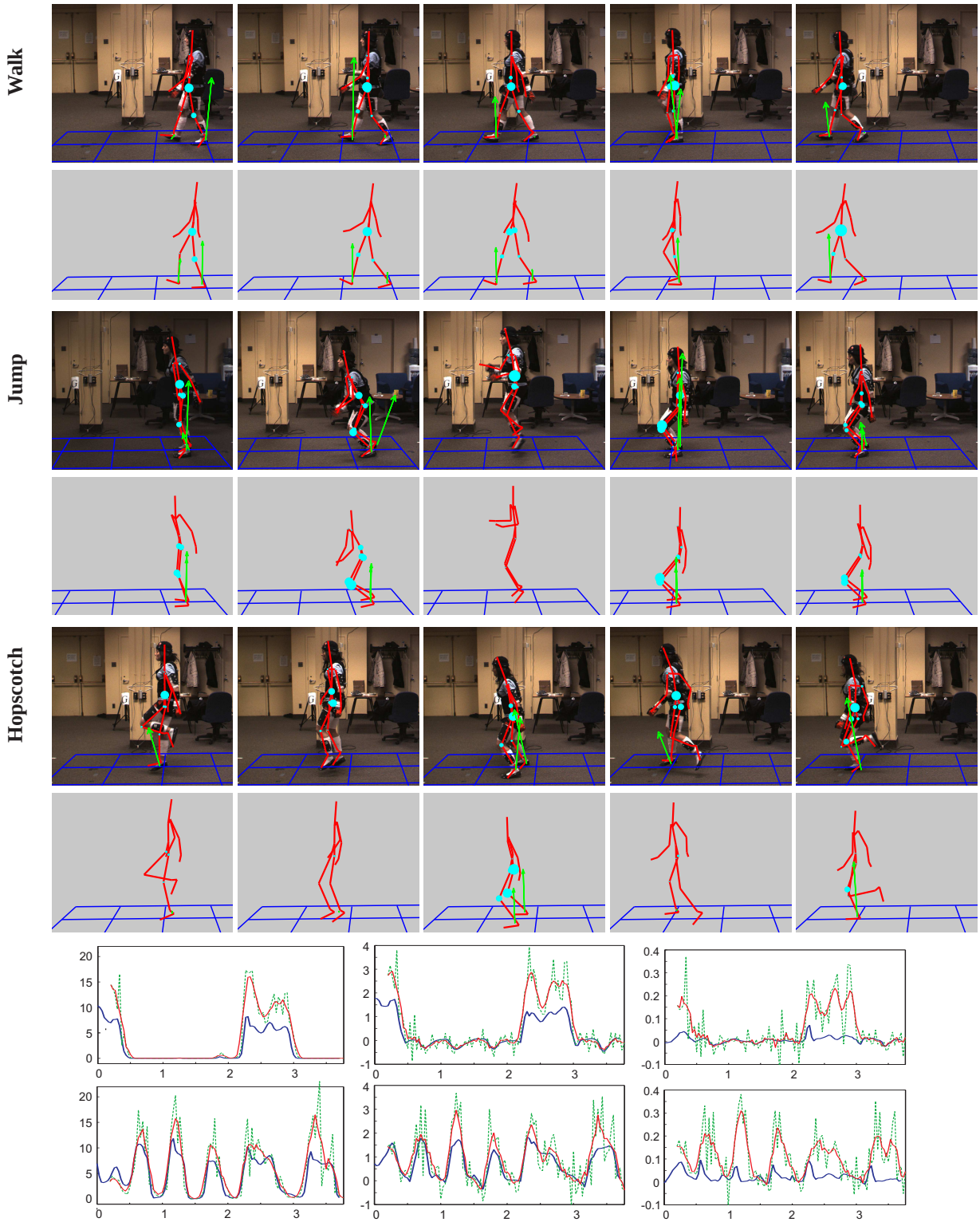


Figure 11. **Video-Based Estimation:** Rows 1, 3, and 5 show results for binocular tracking from two views (only one view shown). Rows 2, 4, and 6 show results on the corresponding mocap. Each figure shows the stick figure (red), the estimated ground plane (blue), the ground reaction forces (green) generated by the estimated model, and the magnitude of internal joint torques (diameter of the cyan disks). Plots on rows 7 and 8 compare the mocap (blue) and video tracking results (smoothed (red) and unsmoothed (green)) for the hopscotch sequence. Row 7 (8) shows, from left to right, ground reaction forces, knee torque and ankle torque for the left (right) leg with respect to time.

Acknowledgements: This work was supported in part by grants from NSERC Canada and CIFAR. We also thank Niko Troje for the MoCap and video data, and Yanshuai Cao for visualization software.

References

[1] A. Balan, L. Sigal, and M. Black. A quantitative evaluation of video-based 3d person tracking. *IEEE Workshop PETS*, pp. 349–356, 2005.

[2] K. Bhat, S. Seitz, J. Popović, and P. Khosla. Computing the physical parameters of rigid-body motion from video. *Proc. ECCV*, 2002.

[3] M. Brand. Physics-based visual understanding. *CVIU* 65(2):192–205, 1997.

[4] M. Brubaker and D. J. Fleet. The kneed walker for human pose tracking. *Proc. CVPR*, 2008.

[5] M. Brubaker, D.J. Fleet, and A. Hertzmann. Physics-based person tracking using simplified lower-body dynamics. *Proc. CVPR*, 2007.

[6] P. de Leva. Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters. *J. Biomech.* 29(9):1223–1230, 1996.

[7] J. Deutscher and I. Reid. Articulated body motion capture by stochastic search. *IJCV* 61(2):185–205, 2005.

[8] A. Ess, B. Leibe, and L. V. Gool. Depth and appearance for mobile scene analysis. *Proc. ICCV*, 2007.

[9] D. Hoiem, A. Efros, and M. Hebert. Putting objects in perspective. *IJCV*, 80(1), 2008.

[10] A. Jepson, D. Fleet, and T. El-Maraghi. Robust online appearance models for vision tracking. *PAMI* 25(10):1296–1311, 2003.

[11] K. Liu, A. Hertzmann, and Z. Popovic. Learning physics-based motion style with nonlinear inverse optimization. *ACM SIGRAPH*, 2005.

[12] R. Mann and A. Jepson. Toward the computational perception of action. *CVPR*, pp. 794–799, 1998.

[13] A. Pentland and J. Williams. Perception of non-rigid motion: Inference of shape, material and force. *Proc. IJCAI*, pp. 1565–1570, 1989.

[14] Z. Popovic and A. Witkin. Physically base motion transfer. *SIGGRAPH*, 1999.

[15] B. Rosenhahn, C. Schmaltz, T. Brox, J. Weickert, D. Cremers, and H.-P. Seidel. Markerless motion capture of man-machine interaction. *Proc. CVPR*, 2008.

[16] A. Schwab and J. Meijaard. How to draw Euler angles and utilize Euler parameters. *Proc. IDETC/CIE*, 2006.

[17] J. Siskind. Grounding and lexical semantics of verbs in visual perception using force dynamics and event logic. *J. AI Res.* 15, 2001.

[18] M. Vondrak, L. Sigal, and O.C. Jenkins. Physical simulation for probabilistic motion tracking. *CVPR*, 2008.

[19] M. Wisse, A. L. Schwab, and R. Q. V. Linde. A 3D passive dynamic biped with yaw and roll compensation. *Robotica* 19(3):275–284, 2001.

[20] A. Witkin and M. Kass. Spacetime constraints. *Proc. ACM SIGGRAPH*, V. 22, pp. 159–168, 1988.

[21] C. Zhu, R. Byrd, and J. Nocedal. L-BFGS-B. *ACM Trans. Math. Software*, 23(5):550–560, 1997.

A. Dynamics of Quaternions

Unit length quaternions form a convenient, singularity free parameterization of 3D rotations. Their use in dynamics requires modifications to the equations of motion and the simulation method to maintain their unit length. Failure to properly account for this yields quaternions that no longer represent rotations and simulate motion properly. (See [16] for further detail on using quaternions in dynamics.)

For a quaternion q the length constraint is written as a function, $c(q) \equiv \frac{1}{2}(\|q\|^2 - 1) = 0$. Further, since $c(q) = 0$ for q at all times, the first two temporal derivatives of $c(q)$ must also be equal to zero. This yields constraints

$$\dot{c}(q) = \dot{q}^T q = 0 \tag{11}$$

$$\ddot{c}(q) = \ddot{q}^T q + \dot{q}^T \dot{q} = 0\,. \tag{12}$$

Satisfying (12) can be done by augmenting the equations of motion (1) in two ways. First, virtual generalized forces in the direction of the constraints are added to the equations corresponding to that quaternion. That is,

$$\mathcal{M}_q(\mathbf{q})\ddot{\mathbf{q}} = \mathcal{F}_q(\mathbf{q}) + \mathcal{A}_q(\mathbf{q}, \dot{\mathbf{q}}) + \rho_q q \tag{13}$$

where the q subscript corresponds to the rows of the matrix corresponding to the quaternion q and ρ_q is the magnitude of the virtual force. Second, the constraint $q^T\ddot{q} = -\dot{q}^T\dot{q}$ is added to the system. The magnitude of the virtual force ρ_q is unknown but can be solved for along with the accelerations $\ddot{\mathbf{q}}$. This augmentation is done for each quaternion.

With exact integration of the augmented equations of motion and valid initial conditions, the two constraints above are always satisfied. However, numerical integration will violate the constraints. To address this, the quaternion and its time derivative are projected to satisfy the constraints at the end of each time step. Specifically, $q = \hat{q}/\|\hat{q}\|$ and $\dot{q} = \hat{\dot{q}} - (\hat{\dot{q}}^T q)q$ where $\hat{q}$ and $\hat{\dot{q}}$ are the quaternion and its time derivative after the integration step but prior to projection.

The projection changes the estimates of velocity and acceleration. Specifically, the quaternion q is observed but the result of the integration step $\hat{q}$ is unobserved. However, it is known that $\hat{q} = \alpha q$ for some unknown α. So the velocity can be written as $\dot{q}_t = (\alpha q_{t+1} - q_t)/\Delta$ and we solve for the value of α by constraining the recovered velocity $\dot{q}_t$ to satisfy (11). The same problem with quaternion velocity is solved by noting that the observed velocity $\dot{q}$ is related to $\hat{\dot{q}}$ by $\hat{\dot{q}} = \dot{q} + \beta q$. We solve for β while ensuring that the recovered acceleration $\ddot{q}_t$ satisfies (12).