Data-driven Inverse Dynamics for Human Motion

Xiaolei Lv[⋆♯] Jinxiang Chai[†] Shihong Xia[⋆]
[⋆] Institute of Computing Technology, Chinese Academy of Sciences
[♯] University of Chinese Academy of Sciences [†]Texas A&M University

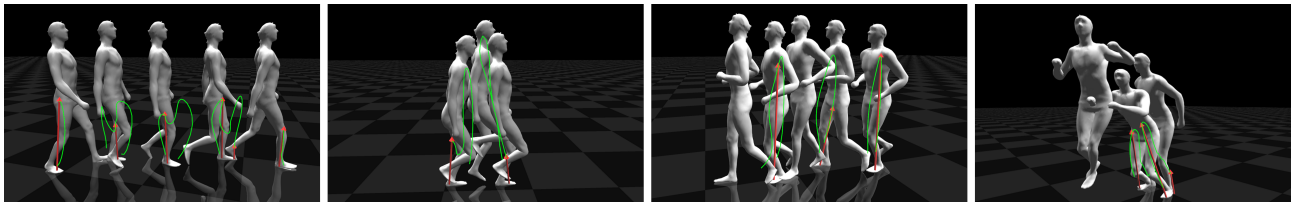

Figure 1: *Our data-driven inverse dynamics system accurately estimates biomechanically valid contact information and internal joint torques from input kinematic human motion data.*

Abstract

Inverse dynamics is an important and challenging problem in human motion modeling, synthesis and simulation, as well as in robotics and biomechanics. Previous solutions to inverse dynamics are often noisy and ambiguous particularly when double stances occur. In this paper, we present a novel inverse dynamics method that accurately reconstructs biomechanically valid contact information, including center of pressure, contact forces, torsional torques and internal joint torques from input kinematic human motion data. Our key idea is to apply statistical modeling techniques to a set of preprocessed human kinematic and dynamic motion data captured by a combination of an optical motion capture system, pressure insoles and force plates. We formulate the data-driven inverse dynamics problem in a maximum a posteriori (MAP) framework by estimating the most likely contact information and internal joint torques that are consistent with input kinematic motion data. We construct a low-dimensional data-driven prior model for contact information and internal joint torques to reduce ambiguity of inverse dynamics for human motion. We demonstrate the accuracy of our method on a wide variety of human movements including walking, jumping, running, turning and hopping and achieve state-of-the-art accuracy in our comparison against alternative methods. In addition, we discuss how to extend the data-driven inverse dynamics framework to motion editing, filtering and motion control.

Keywords: Character animation, data-driven, inverse dynamics

Concepts: •**Computing methodologies** → **Motion capture; Motion processing; Virtual reality; Physical simulation;**

1 Introduction

One of classic problems in computer graphics, robotics and biomechanics is how to accurately identify contact information and internal joint torques from input kinematic motion data. With such dynamic data, motion can be more accurately edited to meet new constraints [Sulejmanpašić and Popović 2005]. Those data can be utilized to generate new motions that react to external perturbations and changes in physical quantities such as frictions [Wei et al. 2011], can be physically simulated to allow external perturbations [Muico et al. 2009] and can serve as a basis for development of new control or synthesis algorithms for human characters [Sok et al. 2007; Liu et al. 2010].

Inverse dynamics is challenging because the problem is often noisy and ambiguous, particularly when double stances occur. Current solutions often utilize heuristic rules [Ren et al. 2008; Ko and Badler 1996] or "minimal principle" [Witkin and Kass 1988; Cohen 1992] to reduce the ambiguity of the solution space. However, none of the solutions can accurately produce biomechanically valid contact information and internal joint torques that are consistent with input kinematic data.

In this paper, we present a new inverse dynamics method that accurately estimates both contact information and internal joint torques from input kinematic human motion data (Figure 1). Our approach is data-driven because it utilizes kinematic and dynamic priors embedded in prerecorded kinematic and dynamic motion data. Specifically, we apply statistical modeling techniques to a set of pre-captured human motion data consisting of both human kinematic and dynamic motion data and construct a low-dimensional data-driven prior model to constrain the solution space for inverse dynamics. We formulate the data-driven inverse dynamics problem in a maximum a posteriori (MAP) framework by estimating the most likely contact information and internal joint torques consistent with input kinematic motion data. To construct a large database of human kinematic and dynamic motion data, we firstly combine a twelve-camera Vicon optical motion capture system with four Kistler force plates to capture both human kinematic and contact data. We then accurately reconstruct the internal joint torques, contact forces, torsional torques and center of the pressure from recorded kinematic and contact data.

We demonstrate the accuracy of our method on a wide variety of human movements including walking, jumping, running, turning and hopping. We show that our method advances the state of the art by comparing it against alternative methods, including both single frame optimization method [Ko and Badler 1996] and spacetime optimization method [Witkin and Kass 1988; Cohen 1992]. The evaluation shows that our method produces much more accurate results than alternative methods for all types of test data. We

SA '16 Technical Papers,, December 05-08, 2016, , Macao
ISBN: 978-1-4503-4514-9/16/12
DOI: http://dx.doi.org/10.1145/2980179.2982440

ACM Reference Format
Lv, X., Chai, J., Xia, S. 2016. Data-driven Inverse Dynamics for Human Motion. ACM Trans. Graph. 35, 6, Article 163 (November 2016), 12 pages. DOI = 10.1145/2980179.2982440
http://doi.acm.org/10.1145/2980179.2982440.

further show that the data-driven inverse dynamics model can be generalized to new people and new motions. Finally, we extend our method to motion editing, motion filtering and motion control. For example, we can generate physically realistic human motions that react to changes in physical parameters such as friction properties of environments.

1.1 Contributions

Our method is made possible by the following technical contributions:

- A novel data-driven method for accurately estimating the contact information and internal joint torques from kinematic human motion data.
- The first full-body kinematic and dynamic motion database that includes contact forces, center of pressure, torsional torques and internal joint torques for a wide variety of human motions including walking, jumping, running, turning and hopping.
- An efficient method for estimating physical quantities of human body, including center of mass and moments of inertia of each body segment, from prerecorded kinematic motion data and contact information.
- A new data-driven motion model that utilizes both kinematic and dynamic priors embedded in prerecorded kinematic and dynamic data.

2 Background

We introduce a data-driven inverse dynamics method by constructing a low-dimensional prior model on kinematic and dynamic human motion data, and integrate the model into previous motion synthesis model to generate natural-looking and physically-valid human motion. Therefore, we focus on models solving the inverse dynamics problem, the construction of dynamic motion database and physically valid data-driven motion models.

Inverse dynamics solution. Inverse dynamics is a basic step in physics-based human motion modeling. It is an indeterminate problem during the double support phase. Some scientists in biomechanics community presented a variety of assumption [Ren et al. 2008; Ko and Badler 1996] to solve the indeterminate problem. Those methods do not work for asymmetric human gait. Xiang et al. [2011] used an optimization-based predictive dynamics approach to calculate the joint torques and ground reaction forces during asymmetric gait motion. However, it dose not work well for a completely natural motion because of the usage of "minimal principle" objective function. Oh et al. [2013] proposed a two phases method to solve the inverse dynamics problem. The traditional method of Newtonian mechanics was used for the single support phase. An artificial neural network model is applied for the double support phase. Since the two phases are mutually disjointed, there can be inevitable discontinuity. Zell et al. [2015] proposed a combined representation of motion and physical parameters and used the k-nearest-neighbour and asymmetrical principal component analysis to deduce the ground reaction forces and joint torques directly from an input motion. The predicted contact information using the regression method may not be physically valid due to the lack of physical constraints. There are also some tools such as OpenSim [Delp et al. 2007], Anybody [Anybodytech 2016] used in biomechanics community to solve the inverse dynamics problem with estimated contact forces. In this paper, we combine the priors embedded in training data and physical constraints to solve the inverse dynamics problem.

The standard solution to the inverse dynamics problem over a motion sequence is spacetime formulation in computer graphics literature [Witkin and Kass 1988; Cohen 1992]. This optimization is subject to foot-ground contact constraints, friction limit constraints and the discretization of physics constraints determined by a finite difference scheme. The generated contact force and internal joint torques are physically valid. However, they may not match the true situation. In order to make the generated motion look natural, the "minimal principle" strategy is adopted, e.g. minimal energy, minimal torque, and so on. These heuristics work well for highly dynamic motions, but it remains challenging to model low energy motions and highly stylized human movements. Brubaker et al. [2009] used a 12-segment articulated body model to estimate contact dynamics. This method gives a good approximation concerning the mean value, but differs regarding temporal development. Liu et al. [2010; 2015] used randomized sampling to tackle control problems given a reference motion. Vondrak et al. [2008; 2012] used physical simulation for human motion tracking given monocular/multi-view imagery. These methods can be seen as another solution to the inverse dynamics problem. However, the tracking and open-loop control nature of the algorithm implies the algorithm can fail when the motion is too long. Meanwhile sporadic contact points and highly discontinuous pressure distribution often occur in collisions between rigid bodies [Jain and Liu 2011]. We instead use a data-driven model on the captured kinematic and dynamic motion data to solve the inverse dynamics problem.

The body segment parameter (BSP) is very essential to the inverse dynamics [Rao et al. 2006]. There are two main methods to estimate the BSPs. Predictive equation method is based on the density data from cadavers [Yeadon and Morlock 1989]. The advantage of this method is quick and easy. However, the predictive equations are limited by the measure techniques (e.g. uniform density) and moreover by the group on which they are based (e.g. small sample, elder males). The second method is to use the medical imaging techniques to measure BSPs directly on living subjects [De Leva 1996; Ho et al. 2013]. This method provides accurate results, however, it may not be possible for all researchers to estimate BSPs by scanning individual subjects. Our approach uses the force plate data and the captured kinematic data to compute the mass and the moment of inertia of each body segment. The result of optimized BSPs is consistent with the biomechanics literature [Ho et al. 2013].

Dynamic human motion database. Human motion data includes kinematic and dynamic motion data. The kinematic human motion data consists of joint kinematics, and the dynamic data contains internal joint torques, ground reaction force, and so on. In computer graphics community, the databases of human motion mainly contain kinematic data, for example, the CMU online mocap database. There are many works in computer graphics involving force or pressure captures [Aladdin and Kry 2012; Kry and Pai 2006]. In this paper, we focus on the capture of ground reaction forces. To obtain the dynamic motion data, force plates and pressure insoles are widely used to measure the ground reaction forces [Zhang et al. 2014; Adelsberger and Troster 2013; Ha et al. 2011; Yin and Pai 2003]. Unfortunately, the unmovable force platform can only be operated in a highly restricted environment. It is not suitable to capture highly dynamic human motion. The pressure insoles can be used to acquire motions that are difficult to capture in restrictive lab settings, such as highly dynamic motions that require a large amount of space. However, each pressure sensor in pressure insoles only records an analog resistance reading proportional to the applied pressure. As a result, a scaling parameter for each sensor and the horizontal contact force and torsional torque cannot be measured by the pressure insoles. Forner-Cordero et al. [2006] proposed a method to calculate the inverse dynamics from pressure insoles. However, there is no benchmark data (force plate data) to

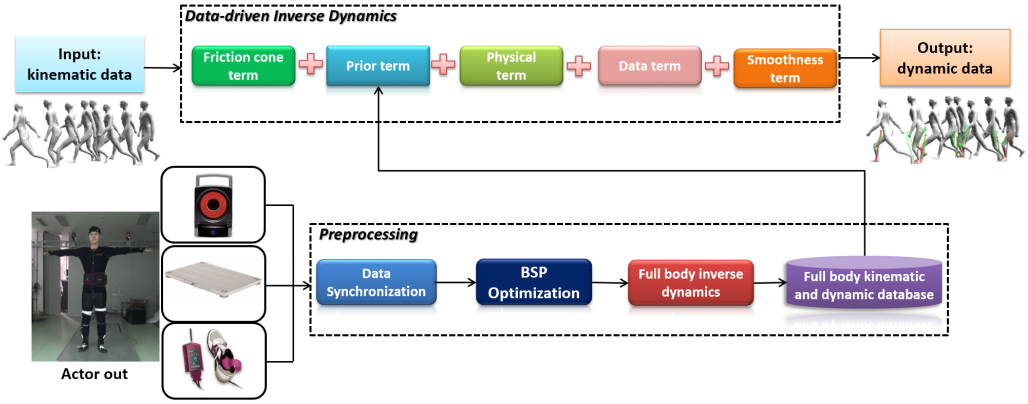

Figure 2: *System overview.*

calibrate the scaling parameters of pressure insoles. As a result, the noise in the kinematic data and pressure insoles data will affect the accuracy of the optimization. The work of [Zhang et al. 2014] uses three Kinect cameras and pressure insoles to capture highly dynamic motions. In this paper, the preprocessing method for constructing ground truth training data requires kinematics data captured by Vicon, as well as force plate/pressure data. The data-driven inverse dynamics method does not use any pressure or force sensor data required by [Zhang et al. 2014]. Instead, it utilizes priors embedded in training kinematics/dynamics mocap data to reduce ambiguity.

Physically valid data-driven motion models. A few works have been done using the data-driven techniques in physics-based human motion modeling. Ye et al. [2008] and Safonova et al. [2004] used the data to reduce the search space for physics-based optimization. It has been shown that a global subspace model for kinematic poses is not sufficient to model heterogeneous human actions [Wei et al. 2011]. In particular, Wei et al. [2011] proposed a physically valid statistical model for human motion generation. They choose to model the force field priors based on generalized forces rather than joint torques. As a result, the learned force field priors can only predict resultant forces of joint torques and contact forces. Different from the previous methods, our data-driven techniques utilize dynamics priors to predict contact information and internal joint torques accurately. This enables us to generalize the physically-valid data-driven motion models to human motion control.

3 Overview

We develop an efficient data-driven inverse dynamics method for accurately estimating both contact information and internal joint torques from input kinematic human motion data. The system consists of two main components (Figure 2): the *preprocessing* component and *data-driven inverse dynamics* component.

Preprocessing. Our data-driven inverse dynamics method utilizes priors embedded in large sets of kinematic and dynamic data for human motion modeling. Therefore, we discuss how to construct a full-body kinematic and dynamic motion database, including contact forces, center of pressure, torsional torques and internal joint torques, for a wide variety of human motions. In addition, we introduce an efficient method for estimating physical quantities of human body, including center of mass and moments of inertia of each body segment, from prerecorded kinematic motion data and contact information.

Data-driven Inverse Dynamics. We apply statistical modeling techniques to large sets of kinematic and dynamic motion data and construct a low-dimensional data-driven prior model to constrain the solution space for inverse dynamics. We formulate the data-driven inverse dynamics problem in a nonlinear optimization framework by estimating the most likely contact information and internal joint torques from input kinematic motion data. The optimization ensures the solution is not only consistent with input kinematic motion data but also physically correct and statistically valid.

Applications. We discuss how to extend our data-driven inverse dynamics method for motion editing, filtering and control.

We discuss those components in detail in the next sections.

4 Preprocessing

We construct the first full-body motion capture database that accurately captures contact forces, center of pressure, torsional torques and internal joint torques for varieties of human motions including walking, jumping, running, turning and hopping. In this section, we will discuss the construction of kinematic and dynamic motion database. We acquire kinematic and dynamic motion data using Vicon motion capture system, force plate and pressure insoles. Because the number of available force plates limits the number of consecutive gait cycles that can be analyzed [Hreljac and Marshall 2000], we use pressure insoles to expand our capture volume. The pressure insoles can only provide the vertical force and center of pressure information. In order to construct the whole contact information and expand our capture volume, we use Vicon device for full-body kinematic data capture, and pressure insoles and force plates for contact information. In this paper, the motion data are obtained mainly by wearable pressure insoles.

Data Acquisition. We performed a series of captures to create a large motion database (about 12 minutes per subject) using a Vicon optical motion capture system with twelve 120 Hz cameras [Vicon 2016], four Kistler force plates [Kistler 2016], a pair of pressure insoles with 99 sensors each [Novel 2016] (The accuracy is linear within $\pm$ 3% of full scale). The pressure insoles data are transmitted via a wireless Bluetooth connection at 100 fps. Six healthy subjects (Table 1) take part in the capture. The database consists of a wide variety of human actions, including walking with different speeds (quick, normal and slow), step sizes (big, medium and small), turning angles (45°, 90° and 180° left and right turning), running with different step sizes (big, medium and small), jumping (in place and forward with 0°, 90° and 180° turning angles),

hopping (in place and forward with 0° and 90° turning angles) and their transitions (run to walk, walk to run, jump to run, run to jump, walk to jump and jump to walk).

Subject	Height(cm)	Weight(kg)	Age(year)
A	177	66	28
B	175	70	22
C	178	71	21
D	173	60	25
E	166	55	24
F	176	75	32

Table 1: *Subject characteristics*

Data Synchronization. Each pressure sensor of pressure insoles records an analog resistance reading proportional to the applied pressure, which is then converted to a digital value. The relationship of the analog resistance reading R_m, the digital pressure force value P_m returned and the scaling parameter k_m is defined as follows:

$$P_m = k_m / R_m \tag{1}$$

Data from different devices are synchronized by aligning contact events to the timeline of the Vicon camera. The Kistler force plate data are synchronized to Vicon motion capture data by hardware. The pressure insoles are not synchronized to the Vicon system or Kistler force plate by hardware. We have designed an algorithm to synchronize the Kistler force plate and Novel pressure insoles. We compute the time shift δt between Kistler force plate and Novel pressure insoles and scaling parameter k_m by solving the following optimization problem

$$\min_{\delta_t, k_m} \sum_t \| F_z^{t+\delta_t} - \sum_m \frac{k_m}{R_m^t} \|^2 + w_1 \sum_t \| cop^{t+\delta t} - \sum_m Pos_m^t \frac{k_m}{R_m^t} \|^2 \tag{2}$$

where $F_z^{t+\delta_t}$ is the vertical component of the contact force of the force plate at time $t+\delta t$, $cop^{t+\delta t}$ is the center of pressure of the force plate at time $t+\delta t$, Pos_m^t is the position of the mth pressure sensor at time t, and R_m^t is the reading of the mth pressure sensor at time t. The weight w_1 is set to 100. The algorithm provides very accurate synchronization, usually less than 1 frame in our experiments. This accuracy is sufficient for the synchronization between the force plate and the pressure-sensing shoes. The Kistler force plate are running at a much higher frame rate (1000fps), hence picking the frame with the closet time stamp for alignment usually gives satisfactory results. In Figure 3, we shows the result of data synchronization.

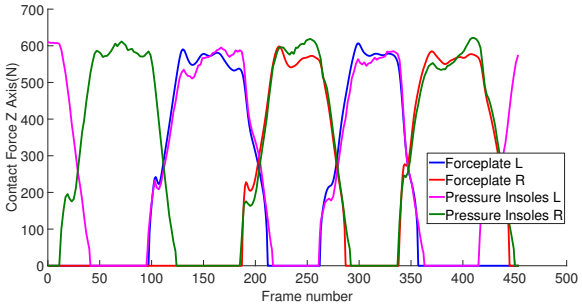

Figure 3: *Data synchronization of force plate and the pressure insoles. The L and R represent the left foot contact and right foot contact. The first two steps are not on the force plates.*

Body Segment Parameter (BSP) Optimization. Acquiring full-body dynamic motion data requires computing joint torques, given physical quantities of each bone segment and contact forces. We can use the force plate and pressure insoles to capture the contact forces. However, the measurement of BSP is difficult because the human body parts are not as easily described as rigid bodies with defined joints and inflexible geometry, daily varying mass, etc. Our dynamic model approximates human motion with a set of rigid body segments using the Vicon Plug In Gait model in Nexus [Vicon 2016]. The skeleton fit is done by Vicon Nexus. We describe a full-body character pose with a set of independent joint coordinates $q \in R^{44}$, including absolute root position, orientation and the relative joint angles of 17 joint. These bones are pelvis (6 DOF), head (3 Dof), thorax (3 Dof) and left and right shoulders (2 Dof), arms (3 Dof), forearms (2 Dof), hands (2 Dof), upper legs (3 Dof), lower legs (1 Dof) and feet (3 Dof).

We use Newtonian dynamic equations to optimize the physical properties of human bone segment. The Newtonian dynamic equations for full body movement can be defined as follows:

$$M(q)\ddot{q} + C(q,\dot{q}) + G(q) = \tau + J_F^T F + J_{\tau_z}^T \tau_z \tag{3}$$

where q, $\dot{q}$ and $\ddot{q}$ represent the joint angle poses, velocities and accelerations, respectively. The quantities $M(q)$, $C(q,\dot{q})$ and $G(q)$ are the joint space inertia matrix, centrifugal/Coriolis and gravitational forces, respectively. The vector τ, F and τ_z are internal joint torques, contact force and torsional torque respectively. J_F and J_{τ_z} are contact force Jacobian matrix and torsional torque Jacobian matrix respectively. Human muscles generate torques about each joint, leaving global position and orientation of the body as unactuated joint coordinates. The movement of global position and orientation is controlled by contact forces F. We have captured the movement of global position and orientation by the Vicon motion capture system and contact force by Kistler force plate. As a result, the redundancy introduced by the use of whole body kinematic data and force plate data can be used to improve the accuracy of the estimated body segment parameters. To achieve this goal, we initialize each subject's body node approximated by a cuboid with the same density. The dimensions of cuboids are set by Plug In Gait models. Assuming the weight of each subject is known, we can use it to compute the physical quantities of each bone segment as the initial value of the body segment parameter optimization.

For an input motion, we select K frames with equally spaced, and formulate the body segment parameter optimization as follows:

$$\begin{aligned}\min_{m,com,I} \; & w_1 \sum_{i=1}^{K} \| (M(q)\ddot{q} + C(q,\dot{q}) + G(q) - J_F^T F - J_{\tau_z}^T \tau_z)_{1:6} \|^2 \\ & + w_2 \sum_{j=1}^{n} ||I_j - I_{j\text{init}}||^2 + w_3 \sum_{j=1}^{n} ||m_j - m_{j\text{init}}||^2 \\ & + w_4 \sum_{j=1}^{n} ||com_j - com_{j\text{init}}||^2 \\ \text{subject to } & m_{\text{total}} = \sum_{j=1}^{n} m_j .\end{aligned} \tag{4}$$

where m, com, I represent the set of mass, center of mass and the inertia of body nodes, m_j, com_j, I_j represent the mass, center of mass, mass inertia of the jth body node, $m_{j\text{init}}$, $com_{j\text{init}}$, $I_{j\text{init}}$ is the initial mass, center of mass and mass inertia value of the jth body node. m_{total} is whole mass of human body. The subscript 1:6 means taking the first six elements of the vector. We optimize using the Levenberg-Marquardt programming method [Bazaraa et al. 1993]. The weights w_1,w_2,w_3 and w_4 are set to 0.1, 10, 5 and 10, respectively.

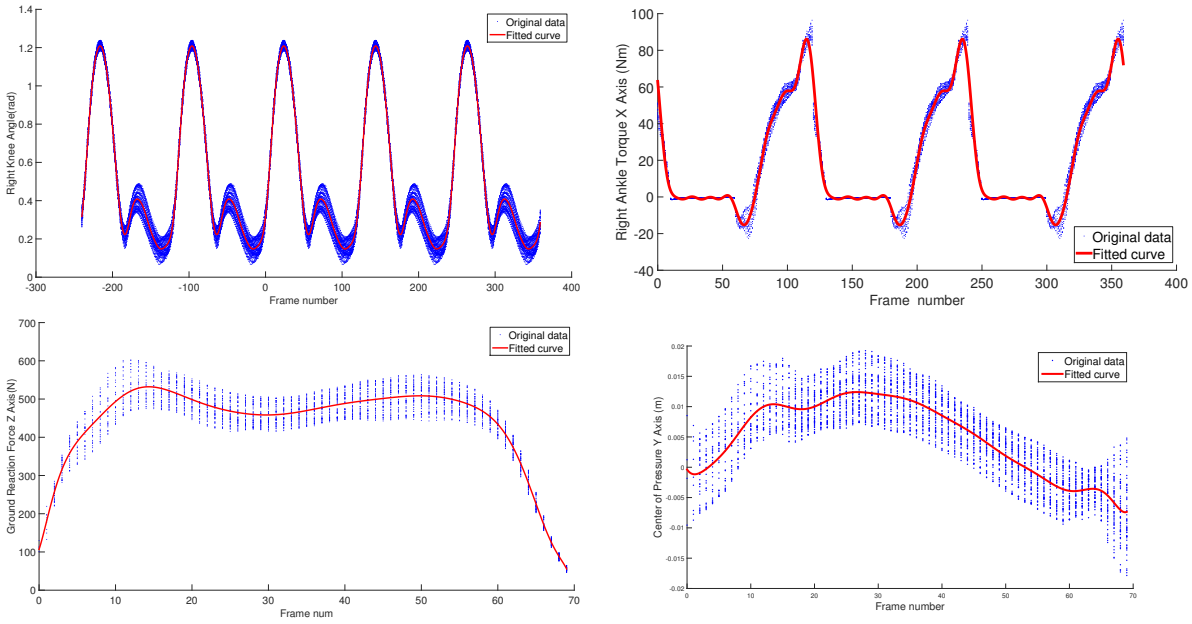

Figure 4: *Part of kinematic and dynamic patterns (joint angle, internal joint torque, contact force and center of pressure) obtained by temporally aligning and averaging 120 captured walking sequences.*

We do the optimization for 100 different motion capture sequences (walking, running, jumping and hopping) of the same subject, and get the average value of mass, center of mass and mass inertia as the final physical properties of the subject. The statistics about the estimated physical properties can be found in the supplementary material. The resulting body parameters are in the reasonable ranges of human data as compared with the biomechanics literature [Ho et al. 2013].

Full-body Inverse Dynamics Optimization. Because the number of available force plates limits the number of consecutive gait cycles, we use pressure insoles to expand our capture volume. The dynamic data are obtained mainly by pressure insoles. The pressure insoles can only provide center of pressure and the vertical force information with the help of scaling parameter which has been solved in the data synchronization part. In this part, we use the BSP information to reconstruct the internal joint torques, contact forces, torsional torques and center of the pressure from kinematic motion data and pressure data.

We formulate the optimization for every frame as follows:

$$\min_{F,cop,\tau_z} w_1||(M(q)\ddot{q}+C(q,\dot{q})+G(q)-J_F^TF-J_{\tau_z}^T\tau_z)_{1:6}||^2 + w_2||F_z-F_{z\text{init}}||^2+w_3||cop-cop_{\text{init}}||^2 \quad (5)$$

where F,F_z,cop and τ_z are the contact force, the vertical component of contact force, center of pressure and torsional torque respectively, $F_{z\text{init}}$ and cop_{init} are the captured vertical contact forces and center of pressure using pressure insoles. We initialize the optimization using pressure insole data with $F_x=0, F_y=0, \tau_z=0$ and optimize using the Levenberg-Marquardt programming method [Bazaraa et al. 1993]. The optimization often took several seconds to converge. The weights w_1,w_2 and w_3 are set to 1.0, 1.0 and 100, respectively.

In the Data Synchronization step, we synchronize the force plate data and pressure insole data and get the accurate scaling parameters of pressure insoles. In the Body Segment Parameter Optimization step, we use the force plate data to optimize for the BSP of specified subject. In the Full-body Inverse Dynamics Optimization step, we use the optimized scaling parameters and BSP to optimize the contact information and internal joint torques for pressure insole data. We evaluate the accuracy of contact information optimized from pressure data by comparing it against contact information measured by force plates. We also compare our method with [Zhang et al. 2014]. For a fair comparison, the same BSP parameter is used. The result of our method and [Zhang et al. 2014] compared with the force plates data can be found in the supplementary material.

We have captured 120 sequences of normal walking for subject E and used the proposed techniques to get the kinematic and dynamic data. We use dynamic time warping techniques [Myers and Rabiner 1981] to automatically register each sequence which is visualized in Figure 4. In order to visualize the pattern of center of pressure, we aligned the foot to the same coordinate. We repeat the captured motion 3 times and extract the joint angle and joint torque patterns by temporally aligning and averaging 120 sequences (red curve in Figure 4). The pattern results are consistent with the biomechanics literature [Bergmann et al. 2001].

5 Data-driven Inverse Dynamics

Given an input motion $y=\{(q^t,\dot{q}^t)|t=1,..,T\}$, the problem of data-driven inverse dynamics is to predict x, $x=\{(\chi^t,\tau^t)|t=1,...,T\}$, where χ^t is the six dimensional contact information of each foot, $\chi^t=\{F_x,F_y,F_z,cop_x,cop_y,\tau_z\}$, and τ^t is the internal joint torques at frame t. $F_x,F_y,F_z,cop_x,cop_y,\tau_z$ are three dimensional contact force of different axis, center of pressure on the ground plane and torsional torque respectively.

We formulate the data-driven inverse dynamics problem in a maximum a posteriori (MAP) framework by estimating the most likely contact χ and joint torque τ from the input motion y:

$$\arg\max_x pr(x|y) \propto \arg\max_x pr(y|x)pr(x) \quad (6)$$

In our implementation, we minimize the negative logarithm of the posteriori probability density function $pr(x|y)$, yielding the following energy minimization problem:

$$\arg\min_{x} \underbrace{-\ln pr(y|x)}_{E_{\text{physical}}} + \underbrace{-\ln pr(x)}_{E_{\text{prior}}} \tag{7}$$

where the first term is the likelihood function measuring how well the generated contact information χ and internal joint torques τ match the input motion. The second term is the prior distribution function.

The data-driven inverse dynamics problem can now be solved by nonlinear optimization methods, given an input kinematic motion. The optimization computes internal joint torques and contact information by minimizing the following objective function

$$\min_{x} \lambda_1 E_{\text{physical}} + \lambda_2 E_{\text{prior}} + \lambda_3 E_{\text{data}} + \lambda_4 E_{\text{friction}} + \lambda_5 E_{\text{smooth}} \tag{8}$$

where E_{physical} ,E_{prior}, E_{data}, E_{friction} and E_{smooth} represent the physical term, prior term, data term, friction cone term and smoothness term, respectively. The weights $\lambda_1, ..., \lambda_5$ control the importance of each term and are experimentally set to 2, 2, 100, 100, 1. We describe details of each term in the following subsections.

5.1 Physical term

The term E_{physical} measures how well the generated contact information and internal joint torques satisfy the physical constraints. Mathematically we have

$$E_{\text{physical}} = -\ln pr_{\text{physics}} \tag{9}$$

$$pr_{\text{physics}} \propto pr(\ddot{q}^t, \dot{q}^t, q^t | \tau^t, \chi^t) \tag{10}$$

Because of simplified dynamics [Sok et al. 2007]/contact models [Muico et al. 2009], discretization of physics constraints and approximate modeling of BSPs, dynamic models are often inconsistent with observed data in physics-based modeling. Accordingly, the Newtonian dynamic equations (Equation 3) are often not satisfied precisely. Similar to the work of [Wei et al. 2011], we assume Newtonian dynamic equations are distributed by Gaussian noise of a standard deviation of σ_{physics}:

$$pr_{\text{physics}} \propto \exp \frac{\|M(q)\ddot{q} + C(q,\dot{q}) + G(q) - J_F^T F - J_{\tau_z}^T \tau_z - \tau\|^2}{-2\sigma_{\text{physics}}^2} \tag{11}$$

where the standard deviation σ_{physics} shows our confidence of physics-based dynamic models.

5.2 Prior term

E_{prior} measures how well the generated contact information and internal joint torques match the contact and joint torque patterns of specific motion. To compute inverse dynamics for each motion, we need to segment the input kinematic motion into different contact configurations (i.e., motion primitives). We segment motion examples into four distinctive motion primitives automatically according to the contact information (right foot contact, left foot contact, double foot contact, and fly state). For walking, there are three distinctive contact configurations (right foot stance, left foot stance, and double stance). Running has three (right foot stance, left foot stance, and flight). This ensures that all the motion segments within the same motion primitive are contact consistent.

$$E_{\text{prior}} = -\ln pr_{\text{prior}} \tag{12}$$

In practice, most human behaviors are intrinsically low dimensional with legs and arms operating in a coordinated way. A nonlinear manifold is embedded in a high-dimensional space. We automatically learn a series of low-dimensional linear models to approximate this high-dimensional manifold. To build a local low-dimensional model, we search the kinematic and dynamic database for examples that are close to the current pose, velocity and acceleration $(q^t, \dot{q}^t, \ddot{q}^t)$. These examples are then used as training data to learn a simple linear model via Principal Component Analysis (PCA) [Bishop et al. 1995]. A new local model is created to optimize the internal joint torques and contact information for each frame.

In our implementation, we represent the root translation in the ground plane and the rotations about the vertical axis at the current frame with respect to the root coordinate system at the previous frame in order to eliminate the effect of absolute positions in the ground plane and the rotations about the vertical axis. The query metric, for each example $(\tilde{q}, \tilde{\dot{q}}, \tilde{\ddot{q}})$ in the database is

$$\sum_i w_i(\alpha\|q_i^t - \tilde{q}_i\|^2 + \beta||\dot{q}_i^t - \tilde{\dot{q}}_i||^2 + \gamma||\ddot{q}_i^t - \tilde{\ddot{q}}_i\|^2) \tag{13}$$

The subscript i means the ith degree of freedom. The weight w_i may be chosen to assign more importance to ceratin degree of freedom. The weights α, β and γ control the importance of each term and are set to 1.0, 0.1,and 0.01, respectively.

After subtracting the mean, p^t, from the K closest examples in the local region we apply PCA to the covariance matrix of these K examples $\{(\tilde{q}_k, \tilde{\dot{q}}_k, \tilde{\ddot{q}}_k, \tilde{\chi}_k, \tilde{\tau}_k)|k = 1, ..., K\}$. We set the value of K to 200. We choose the pose q, angular velocity $\dot{q}$, angular acceleration $\ddot{q}$, internal joint torques τ and contact information $\chi = \{F_x, F_y, F_z, cop_x, cop_y, \tau_z\}$ as the feature space of PCA model. In order to make the priors to be character independent, we scale the internal joint torques, contact force, torsional torque and center of pressure by the following equation:

$$\tau_{\text{feature}} = \frac{\tau}{ghm_{\text{total}}} \tag{14}$$

$$F_{\text{feature}} = \frac{F}{gm_{\text{total}}} \tag{15}$$

$$\tau_{z\text{feature}} = \frac{\tau_z}{ghm_{\text{total}}} \tag{16}$$

$$cop_{\text{feature}} = \frac{cop - hip}{len} \tag{17}$$

where hip is the hip joint position and len is the length of leg. h is the height of the character and m_{total} is the whole mass of the character. g is the gravitational acceleration.

We obtain a linear model for the current frame.

$$z^t = p^t + B^t u^t \tag{18}$$

where B^t is constructed from the eigenvectors corresponding to the largest eigenvalues of the covariance matrix of the local examples. u^t is a low-dimensional representation of the current state z_t, $z^t = (q^t, \dot{q}^t, \ddot{q}^t, \tau^t, \chi^t)^T$. We automatically determine the dimensions of subspaces by keeping 95% of the original energy. In our implementation, we normalize the different components prior to the PCA. The dimensions of subspaces vary from 5 to 15. We model the probability distribution pr_{prior} with a Gaussian mixture model (GMM). The parameters of the Gaussian mixture model are automatically estimated using an Expectation-maximization algorithm [Bishop et al. 1995].

5.3 Data term

Thus far, we have discussed how to incorporate the contact and joint torque priors into the data-driven inverse dynamics optimization framework. Note that in the prior modeling step, we performed dimensionality reduction analysis on feature space of $(q, \dot{q}, \ddot{q}, \chi, \tau)$ and keeping 95% of the original energy. However, for the current frame, $z^t = (q^t, \dot{q}^t, \ddot{q}^t, \tau^t, \chi^t)^T$, the $(q^t, \dot{q}^t, \ddot{q}^t)^T$ are known before. In order to make optimized contact information and internal joint consistent with the given input motion, we need to add some soft constraints.

$$E_{\text{data}} = \|Pz^t - (q^t, \dot{q}^t, \ddot{q}^t)^T\|^2 \tag{19}$$

where P is the selection matrix for the dimension of $(q^t, \dot{q}^t, \ddot{q}^t)$. We set a higher weight to the data term because PCA subspace constraints here are enforced as soft constraints rather than hard constraints.

5.4 Friction cone term

Similar to the work [Wei et al. 2011; Zhang et al. 2014] in physics based character animation, we use Coulomb's friction model to compute the forces caused by the friction between the character and environment. A friction cone is defined to be the range of possible forces satisfying Coulomb's function model for an object at rest. We ensure the contact forces stay within a basis that approximate the cone with nonnegative basis coefficients. We use multiple contact points $m = 1, \ldots, M$ to model the contact between two surfaces. The contact force can be defined as follows:

$$f(w_1, \ldots, w_M) = \sum_{m=1}^{M} B_m w_m \quad \text{subject to } w_m \geq 0 \tag{20}$$

where the 4×1 vector w_m represents the nonnegative basis weights for the mth contact force and the matrix B_m is a 3×4 matrix consisting of 4 basis vectors that approximately span the friction cone for the mth contact force. The coefficient of friction is set to 0.8 in the data acquisition environment. We can edit the coefficient to generate the motion on different surfaces (Section 6.1).

In our original captured dynamic data, our ground contact model is represented by ground reaction force, center of pressure and torsional torque. Therefore, there needs some conversion relation between the two different contact models.

$$E_{\text{friction}} = \|F - \sum_{m=1}^{M} B_m w_m\|^2 + \alpha \|\tau_z - \sum_{m=1}^{M} (pos_m - cop) \times B_m w_m\|^2 \tag{21}$$

where F,cop and τ_z represent contact force, center of pressure and torsional torque respectively. pos_m represents the global position of the mth contact points. The weight α is set to 100.

5.5 Smoothness term

E_{smooth} measures the smoothness of the generated contact information. The smoothness term is as follows

$$\begin{aligned} E_{\text{smooth}} = &\|F^{t+1} - 2F^t + F^{t-1}\|^2 + \alpha \|\tau_z^{t+1} - 2\tau_z^t + \tau_z^{t-1}\|^2 \\ &+ \beta \|cop^{t+1} - 2cop^t + cop^{t-1}\|^2 \end{aligned} \tag{22}$$

where F,cop τ_z are the generated contact force, center of pressure and torsional torque. The weights α and β are set to 5 and 100, respectively.

In our implementation, we use the backward difference approximation to compute joint velocities and use the central difference approximation to compute joint accelerations with δ_t set to the acquisition frame rate (0.01s). We initialize the optimization with the closest example in the database which is optimized using the Levenberg-Marquardt programming method [Bazaraa et al. 1993]. We firstly drop the smoothness term E_{smooth} in the objective function and use the remaining objective function to optimize for every single frame. We then add the smoothness term to optimize the objective function across the entire motion sequence. The first step converges rapidly because of a good starting point and a low-dimensional optimization space. Each optimization often took from 5 minutes to two hours to converge without code optimization. All the experiments are run on a PC with Intel(R) Xeon(R) processor E5-1620 3.70GHz with 16GB of RAM.

6 Application of Data-driven Inverse Dynamics

In Section 5, the problem of data-driven inverse dynamics is to predict the contact information and internal joint torque $\{(\chi^t, \tau^t) | t = 1, .., T\}$, given the input motion $\{(q^t, \dot{q}^t) | t = 1, .., T\}$. However, in general cases, the users want to specify various forms of kinematic constraints throughout the motion to generate physically-valid natural looking motions. It is the problem of physics-based motion generation which can be seen as the application of data-driven inverse dynamics models. The physics-based motion generation is to generate a motion $\{(q^t, \dot{q}^t, \chi^t, \tau^t) | t = 1, .., T\}$, which satisfies the user constraints c, the physics constraints and prior model constraints. The key idea of our motion synthesis process is sampling the prior distribution function $pr(q, \dot{q}, \chi, \tau)$ to generate a physically-valid motion instance x that matches the users' input c. The prior distribution function $pr(q, \dot{q}, \chi, \tau)$ models the relation between the human pose, pose velocity, contact forces and internal joint torques. The formulation of physics based motion generation is defined as follows:

$$\min_x \lambda_1 E_{\text{physical}} + \lambda_2 E_{\text{prior}} + \lambda_3 E_{\text{c}} + \lambda_4 E_{\text{friction}} + \lambda_5 E_{\text{smooth}} \tag{23}$$

where E_{physical} ,E_{prior}, E_{friction} and E_{smooth} are defined in Section 5. E_{c} includes the various forms of kinematic constraints throughout the motion. In our experiment, we set the weights for E_{physical} ,E_{prior}, E_c, E_{friction} and E_{smooth} to 2, 2, 1000, 100 and 1, respectively. The weight of the constraint term is very large because we want to ensure the generated motion can match user constraints accurately.

6.1 Motion Editing

As noted in previous work [Wei et al. 2011], the incorporation of physics into probabilistic motion prior models significantly improves the generalization of statistical motion prior models. Our system uses the online local PCA to model the statistical motion priors. The experimental results show that the system can generate physically realistic motion that reacts to changes in physical parameters of interaction environments and human bodies.

Walking on slippery surfaces. We can edit the friction coefficients to generate various walking motion results on different surfaces. In Figure 5, we show a simulated character walking on a slippery surface by reducing the friction coefficient to 0.05. In the accompanying video, we show the generated motions on different surfaces. Specifically the friction coefficients are 0.8, 0.2, 0.1 and 0.05 respectively.

Moon walking. We can generate an animation that reacts to changes in gravity of interaction environment. In the accompanying video, we show the generated motions on moon by setting the gravity at $1.62m/s^2$

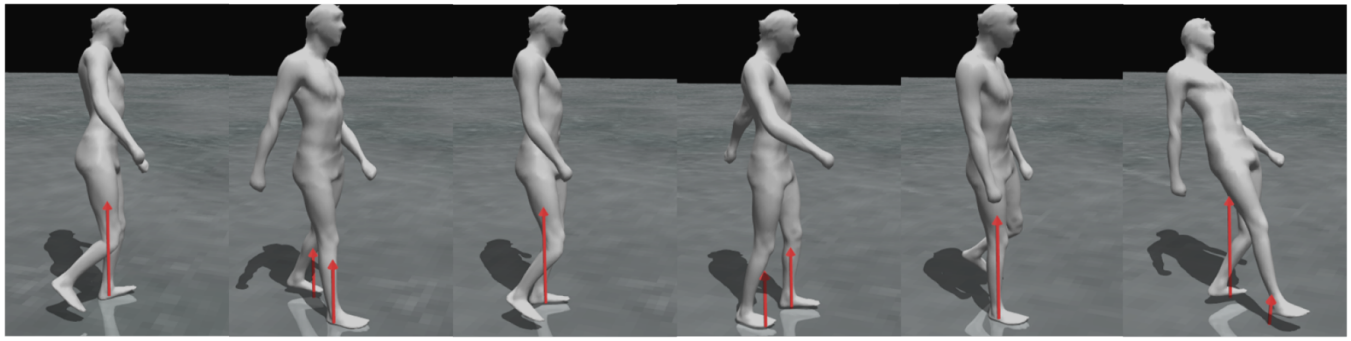

Figure 5: *Generating physically realistic motion that reacts to changes in friction properties of environments: walking on a slippery surface.*

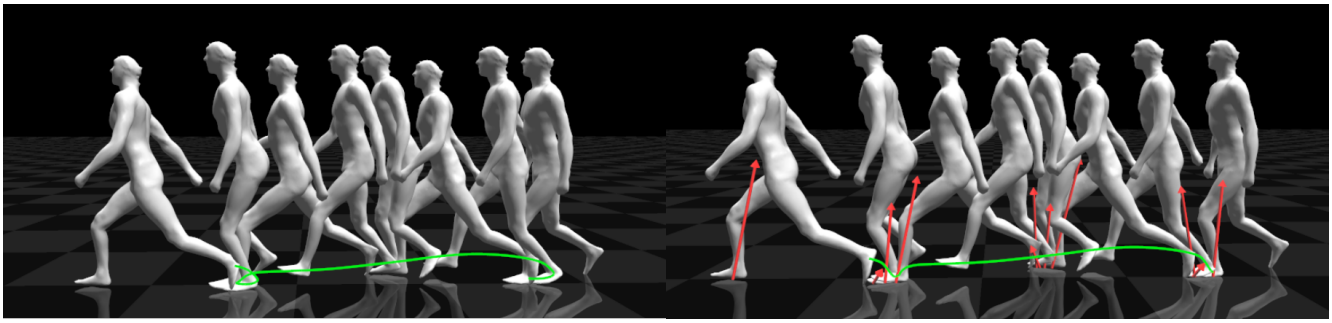

Figure 6: *Motion filtering of a noisy walking sequence: (left) the original walking sequence contains a significant number of foot-skating artifacts; (right) foot-skating removal result.*

Heavy foot. We can edit an animation by changing the physical quantities such as mass and inertias of human bodies. For example, we change the mass of the character by simulating a character wearing a 10 kilogram shoe. In the accompanying video, we show the generated motions of wearing 10 kilogram shoe.

6.2 Motion Filtering

One nice property of our data-driven inverse dynamics model is that they encode environmental contact information and physical constraints. With such a model, we could filter the noisy motion that violates environmental contact constraints. For example, foot-skating artifacts often appear during various motion data processing stages, such as motion editing, blending, warping, or synthesis. Our algorithm can be used to automatically remove footskating artifacts in an input walking sequence after specifying the contact states manually. Figure 6 shows the result of filtering a noisy walking sequence. The original walking sequence contains a significant number of footskating artifacts, in which, instead of remaining firmly in place, a character's foot slides on the ground after the character plants it.

6.3 Motion Control

Our system can be leveraged for motion control systems. Sampling based methods [Liu et al. 2010; Liu et al. 2015] can tackle control problems by using the randomized sampling given a reference motion. In practice, the author uses an offset technique to approximate feedforward torques which can be computed using inverse dynamics techniques. By using the data-driven inverse dynamics method, we can predict the internal joint torques for the reference motion as well. Moreover, our experiments show that we can greatly reduce the sampling number by using the torques predicted by data-driven inverse dynamics instead of the feedforward torques. We can track a normal walking motion using the parameter of nSample = 4000 and nSave = 500 with feedforward torques. Meanwhile we can generate a competitive motion with parameters nSample = 500 and nSave = 100 with the torque prior (torques computed by data-driven inverse dynamics) for the same reference motion. We used the same parameter settings (dynamic properties of character model, PD control parameters, sampling window size etc) in the two experiments. Please refer to the accompanying video for more detail.

7 Results

In this section, we test our algorithm on a wide variety of human actions, including walking, running, jumping and hoping (Section 7.1). We assess the quality of the generated ground contact information by comparing it with ground truth data. We evaluate the effectiveness of our algorithm by comparing it with single frame inverse dynamics and spacetime optimization (Section 7.2). We test the generalization ability of our data-driven inverse dynamics to new people and new motions (Section 7.3). We also evaluate the importance of contact force and joint torque priors and physical constraints for data-driven inverse dynamics (Section 7.4). Our results are best seen in the accompanying video.

7.1 Test on Real Data

We test our algorithm on a wide variety of human actions, including walking, running, jumping and hoping. The accompanying video shows the data-driven inverse dynamics result for a long motion sequence generated by using motion graph technique [Xia et al. 2015].

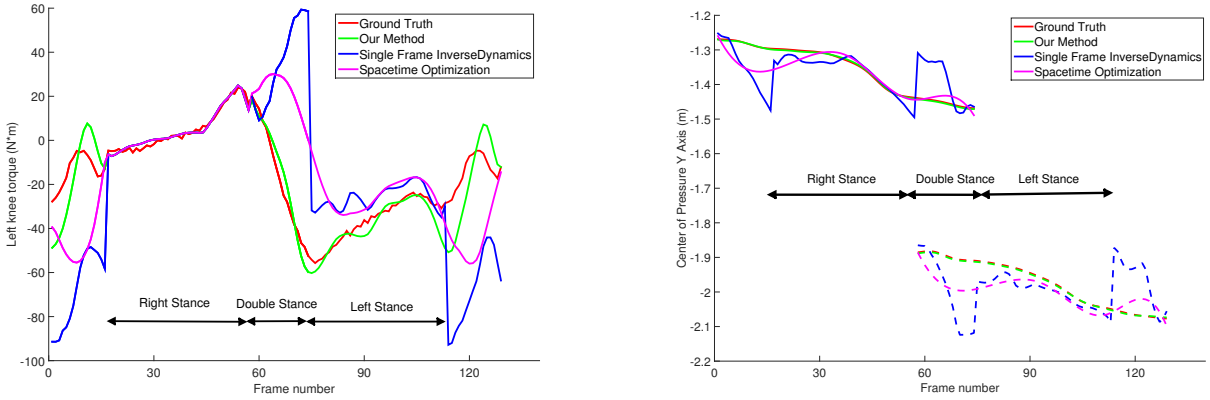

Figure 7: *Comparison our data-driven inverse dynamics against single frame inverse dynamics and spacetime optimization for a normal walking: (left) the comparison curves show the internal joint torque of left knee from all three methods and ground truth data; (right) the comparison curves show the center of pressure from all three methods and ground truth data, where the solid line and dash line represent the right foot contact and left foot contact respectively.*

7.2 Comparison against Ground Truth Data and Alternative Methods

We have assessed the quality of the generated contact information by comparing with ground truth data. We also evaluate the effectiveness of our data-driven inverse dynamics framework by comparing against alternative methods.

7.2.1 Comparison against ground truth data

We evaluate the performance of our algorithm via cross validation techniques. More specifically, we pull out an entire motion sequence in the training data as testing data, use it to extract the labels of foot contact information (double support, single support etc.), and apply the data-driven inverse dynamics algorithm to generate contact information that matches the labels. Figure 7 shows a comparison result of reconstructed left knee torque and center of pressure for normal walking motion. The accompanying video shows a side-by-side comparison between the ground truth contact information and the generated contact information. The average and variances of errors between estimated contact information and ground truth data on four kinds of motions are provided in the supplementary material.

We have also done experiments of leaving out categories of motions (rather than leaving out motions). Our data-driven inverse dynamics model requires the testing categories and the training categories are contact consistent. As a result, we cannot use the running motion categories to do the inverse dynamics for walking motion categories, because there is no double foot contact state for running motions. More specifically, we select three categories of walk (quick walk with medium step, normal walk with big step, slow walk with small step) for the experiments. Each category contains ten specific walking motions. We pull out one category as the testing data and use the other two as the training data. For each category, we have dropped the prior term and used the physical term only to calculate the results, which are available in the supplementary material.

7.2.2 Comparison against single frame inverse dynamics

Inverse dynamics allows us compute torques for the character with a single point of contact with the environment. For the closed loop formed during multiple support phase, however, the problem is undetermined. We use the approximation method proposed by [Ko and Badler 1996]. This method distributes contact force F according to the relative distances of the projection of the center of mass from the ankles. Figure 7 shows a comparison result of reconstructed left knee torque and center of pressure for normal walking motion. Further results are provided in the supplementary material showing the average and variances of errors between estimated contact information and ground truth data on four kinds of motions.

7.2.3 Comparison against spacetime optimization

We follow the spacetime formulation in computer graphics literature [Witkin and Kass 1988; Cohen 1992]. Briefly, we minimize the deviation from prerecorded kinematic motion data as well as the sum of the squared torques. This optimization is subject to foot-ground contact constraints, friction limit constraints, and the discretization of physics constraints determined by a finite difference scheme. We use backward difference to compute joint velocities and use central difference to compute joint accelerations. We follow a standard approach of representing q^t using cubic B-splines. We solve the optimization problem using sequential quadratic programming (SQP) [Bazaraa et al. 1993], where each iteration solves a quadratic programming subproblem. Figure 7 shows a comparison result of reconstructed left knee torque and center of pressure for normal walking motion. We similarly provide the average and variances of errors between estimated contact information and ground truth data on four kinds of motions in the supplementary material.

7.3 Generalization Ability

We evaluate the generalization ability of our data-driven inverse dynamics in terms of different subjects and new motions which are not in the database.

7.3.1 Generalization to new people

We have tested our algorithm on different characters with different BSPs. To achieve this goal, we have constructed ground truth data for two subjects (A and B). We then used data-driven priors embedded in ground truth mocap data of subject A to solve the

inverse dynamics problem of subject B. The results can be found in the supplementary material. The subject characteristics for A and B are shown in Table 1. We have done 3D scans for Character A and B using Artec Eva 3D object scanner [Artec3D 2016], in order to show the differences in anthropometry (Figure 8). The accompanying video shows a side-by-side comparison between the ground truth contact information and the generated contact information using characters A and B's prior respectively. Please refer to the accompanying video for more detail.

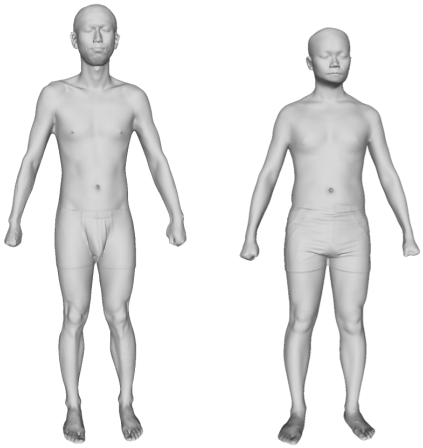

Figure 8: *Character A and B's meshes scanned by Artec Eva 3D object scanner.*

7.3.2 Generalization to new motions

We test the generalization ability of the data-driven inverse dynamics model to new motions. We have captured a Tai Chi motion sequence using Vicon, pressure insoles and force plates and used the preprocessing pipeline in Section 4 to construct the ground truth data. We use walking, running, hopping and jumping motions as training data. More details are available from the accompanying video and supplementary material.

We have also captured a walk upstairs motion sequence and used the same settings to construct the ground truth data. We test our model using the training data of walking, running, hopping and jumping motions. However, the results are not satisfactory. Even without the prior term, we can get a relatively better result than with the prior term. Please refer to the accompanying video and supplementary material for more detail.

7.4 Component Evaluation

In this section, we will evaluate the importance of key components of our data-driven inverse dynamics framework by dropping physical term E_{physical} and prior term E_{prior} in Equation 8.

The importance of the physical term. This experiment demonstrates the importance of physical constraints to our data-driven inverse dynamics. We remove the physical term in the objective function and use the remaining terms to optimize the contact forces and joint torques across the entire motion sequence. We select a sequence of lifting a heavy object with left hand (frame 0-100 is the left stance phase; frame 76-175 is right stance phase). The mass of the heavy object is 5kg. According to the Euler's equations in rigid body dynamics, the external moment of the coronal (frontal) plane is smaller when the supporting leg is on the same side with the heavy object (the left stance phase in this example) than that on the other side. With the physical term, the external moment of the coronal plane reacts appropriately to the added mass as noted in Figure 9. As expected, without the physical term the external moment of the coronal plane dose not respond to the added mass, because there are only normal walking motions in the motion database, in which the external moments of the coronal plane are nearly symmetric. This can be seen in Figure 9, in which the external moments of left stance and right stance are nearly symmetric. Please refer to the accompanying video for more detail.

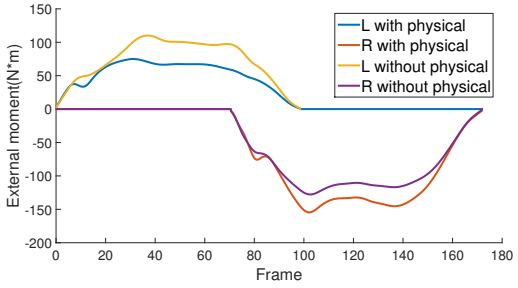

Figure 9: *The external moment of the coronal plane with the physical term and without the physical term. The L and R represent the external moment on the left foot and right foot.*

The importance of prior term. This comparison shows the importance of prior term for data-driven inverse dynamics. More specially, we test the data-driven inverse dynamics model on the 180 turning motions. Further evaluation of the contribution of the prior term is presented in the supplementary material. As shown in the accompanying video, the use of the prior term greatly improves the accuracy of estimation of contact force and center of pressure.

The choice of prior models. The data-driven models are primitive-dependent. In other words, for each action, we first segment the whole motion sequence based on contact states and then build data-driven models for each primitive. We use the online local PCA method to model the high-dimensional manifold for every single frame. We can also model the priors using a global PCA method (global prior model) which do the PCA according to the primitive with the same contact states. See the supplementary material for further results evaluating the performance of choosing different prior models. Please refer to the accompanying video for more detail. The online local PCA model avoids the problem of finding an appropriate structure for a global model, which would necessarily be high-dimensional and nonlinear.

8 Discussion and Conclusion

In this paper, we present an efficient data-driven inverse dynamics method that accurately estimates both contact information and internal joint torques from input kinematic motion data. We demonstrate the accuracy of our method on a wide variety of human movements including walking, jumping, running, turning and hopping and achieve state-of-the-art accuracy in our comparison against alternative methods. In addition, we discuss how to extend the data-driven inverse dynamics framework to motion editing, filtering and motion control.

With the use of pressure insoles, we have expanded our capture volume. The pressure insoles can only provide the vertical force and center of pressure information. We capture the force plate data and

the pressure insoles' data simultaneously, and synchronize the data by aligning contact event and compute the scales of the pressure insoles. Our method can estimate the scales of pressure sensors accurately. After the synchronization and optimization of BSP, we do the full-body inverse dynamics optimization to get the ground truth data. We visualize the data to get clear kinematic and dynamic patterns which are consistent with biomechanics literature. We believe these data will provide insights into designing controllers for simulated virtual humans and biped robots. In particular, the captured kinematic and dynamic data could be leveraged for many applications in human motion processing, analysis and synthesis, such as motion registration and physics-based motion control and optimization. We show an example of using the data-driven inverse dynamics model in the sampling based motion control.

We combine the Kistler force plate data and Vicon motion capture data to optimize the BSP for the specified character. However, the measurement of BSP is difficult because the human body parts are not as easily described as rigid bodies with defined joints and inflexible geometry, daily varying mass, etc. Joint location and rotation axes are very difficult to extract with confidence and the configuration of each joint can make a huge difference to inverse dynamics results. In this paper, we adopt the well-known biomechanical solution of human motion capture and human body modeling using Plug In Gait Model in Nexus [Vicon 2016] in which joint locations and rotation axes are known perfectly. The subject-specific mesh model [Corazza et al. 2010] may improve the accuracy of body segment parameter computation. In the future, we plan to use the scanned mesh model and make joint parameters part of the optimization of BSP for specific subject.

We have tested the generalization ability of our model to new people and new motions. The results show that the data-driven inverse dynamics can generate reliable contact information. The data-driven model can achieve better results than the physical constraints even for different category of motions. When the motions move further away from the database, the data will provide decreasing benefits. However, for the walking upstairs motions, the data-driven inverse dynamics get worse results than with the physical term only, which means the prior term may result in worse estimates. This is partly because scaling rule is too simple. The scaling rule is designed for the conditions that both contact feet are in the same plane. A more complex and well-designed scaling rule may enhance the performance of our algorithm. There are some basic assumptions in this paper. The database focuses on foot ground contact only on flat ground. As a result, it is difficult to generalize it to sloped terrain or steps. There are without nothing but foot ground contact in the database. Input motion should not involve any contacts other that these. In our formulation of data-driven inverse dynamics, we assume the Newtonian dynamics equation are distributed by Gaussian noise. As a result, our data-driven inverse dynamic model cannot guarantee successful forward simulations.

Since our data-driven inverse dynamic model depends on whether the subject is in single stances, double stances or in the air, for each action, we should clearly divide the motion into different primitives corresponding to different contact states. In the future, we would like to use the "Deformable Motion Models" techniques [Min et al. 2009] to the prior construction encoding the environmental contact information. With this improvement, we can do the data-driven inverse dynamics optimization without specifying the contact states of the input motion.

Acknowledgements

We thank the anonymous reviewers for their helpful comments. This work was partially supported by the National Science Foundation under Grants No. IIS-1055046 and IIS-1065384, the National Natural Science Foundation of China under Grants No.61173055, and Foundation for Innovation of Institute of Computing Technology of Chinese Academy of Sciences under Grants No.20166040. Xiaolei Lv and Shihong Xia are affiliated with the Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology of Chinese Academy of Sciences (CAS).

References

ADELSBERGER, R., AND TROSTER, G. 2013. Pimu: A wireless pressure-sensing imu. In *Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on*, IEEE, 271–276.

ALADDIN, R., AND KRY, P. 2012. Static pose reconstruction with an instrumented bouldering wall. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, ACM, 177–184.

ANYBODYTECH, 2016. http://www.anybodytech.com/.

ARTEC3D, 2016. https://www.artec3d.com/.

BAZARAA, M. S., SHERALI, H., AND SHETTY, C. 1993. Nonlinear programming: theory and algorithms. *John Wiley&Sons, New York*.

BERGMANN, G., DEURETZBACHER, G., HELLER, M., GRAICHEN, F., ROHLMANN, A., STRAUSS, J., AND DUDA, G. 2001. Hip contact forces and gait patterns from routine activities. *Journal of Biomechanics 34*, 7, 859 – 871.

BISHOP, C. M., ET AL. 1995. Neural networks for pattern recognition.

BRUBAKER, M. A., SIGAL, L., AND FLEET, D. J. 2009. Estimating contact dynamics. In *Computer Vision, 2009 IEEE 12th International Conference on*, IEEE, 2389–2396.

COHEN, M. F. 1992. Interactive spacetime control for animation. In *ACM SIGGRAPH Computer Graphics*, vol. 26, ACM, 293–302.

CORAZZA, S., GAMBARETTO, E., MUNDERMANN, L., AND ANDRIACCHI, T. P. 2010. Automatic generation of a subject-specific model for accurate markerless motion capture and biomechanical applications. *Biomedical Engineering, IEEE Transactions on 57*, 4, 806–812.

DE LEVA, P. 1996. Adjustments to zatsiorsky-seluyanov's segment inertia parameters. *Journal of biomechanics 29*, 9, 1223–1230.

DELP, S. L., ANDERSON, F. C., ARNOLD, A. S., LOAN, P., HABIB, A., JOHN, C. T., GUENDELMAN, E., AND THELEN, D. G. 2007. Opensim: open-source software to create and analyze dynamic simulations of movement. *IEEE transactions on biomedical engineering 54*, 11, 1940–1950.

FORNER-CORDERO, A., KOOPMAN, H., AND VAN DER HELM, F. 2006. Inverse dynamics calculations during gait with restricted ground reaction force information from pressure insoles. *Gait & posture 23*, 2, 189–199.

HA, S., BAI, Y., AND LIU, C. K. 2011. Human motion reconstruction from force sensors. In *Proceedings of the 2011 ACM*

SIGGRAPH/Eurographics Symposium on Computer Animation, ACM, 129–138.

HO, W.-H., SHIANG, T.-Y., LEE, C.-C., AND CHENG, S.-Y. 2013. Body segment parameters of young chinese men determined with magnetic resonance imaging. *Medicine and science in sports and exercise 45*, 9, 1759–1766.

HRELJAC, A., AND MARSHALL, R. N. 2000. Algorithms to determine event timing during normal walking using kinematic data. *Journal of biomechanics 33*, 6, 783–786.

JAIN, S., AND LIU, C. K. 2011. Controlling physics-based characters using soft contacts. *ACM Trans. Graph. (SIGGRAPH Asia) 30* (Dec.), 163:1–163:10.

KISTLER, 2016. http://www.kistler.com.

KO, H., AND BADLER, N. I. 1996. Animating human locomotion with inverse dynamics. *Computer Graphics and Applications, IEEE 16*, 2, 50–59.

KRY, P. G., AND PAI, D. K. 2006. Interaction capture and synthesis. In *ACM Transactions on Graphics (TOG)*, vol. 25, ACM, 872–880.

LIU, L., YIN, K., VAN DE PANNE, M., SHAO, T., AND XU, W. 2010. Sampling-based contact-rich motion control. *ACM Transctions on Graphics 29*, 4, Article 128.

LIU, L., YIN, K., AND GUO, B. 2015. Improving sampling-based motion control. *Computer Graphics Forum 34*, 2.

MIN, J., CHEN, Y., AND CHAI, J. 2009. Interactive generation of human animation with deformable motion models. *ACM Transactions on Graphics (TOG) 29*, 1, 9.

MUICO, U., LEE, Y., POPOVIĆ, J., AND POPOVIĆ, Z. 2009. Contact-aware nonlinear control of dynamic characters. In *ACM Transactions on Graphics (TOG)*, vol. 28, ACM, 81.

MYERS, C. S., AND RABINER, L. R. 1981. A comparative study of several dynamic time-warping algorithms for connected-word recognition. *Bell System Technical Journal 60*, 7, 1389–1409.

NOVEL, 2016. http://www.novel.de.

OH, S. E., CHOI, A., AND MUN, J. H. 2013. Prediction of ground reaction forces during gait based on kinematics and a neural network model. *Journal of biomechanics 46*, 14, 2372–2380.

RAO, G., AMARANTINI, D., BERTON, E., AND FAVIER, D. 2006. Influence of body segments parameters estimation models on inverse dynamics solutions during gait. *Journal of Biomechanics 39*, 8, 1531–1536.

REN, L., JONES, R. K., AND HOWARD, D. 2008. Whole body inverse dynamics over a complete gait cycle based only on measured kinematics. *Journal of Biomechanics 41*, 12, 2750–2759.

SAFONOVA, A., HODGINS, J. K., AND POLLARD, N. S. 2004. Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. In *ACM Transactions on Graphics (TOG)*, vol. 23, ACM, 514–521.

SOK, K. W., KIM, M., AND LEE, J. 2007. Simulating biped behaviors from human motion data. In *ACM Transactions on Graphics (TOG)*, vol. 26, ACM, 107.

SULEJMANPAŠIĆ, A., AND POPOVIĆ, J. 2005. Adaptation of performed ballistic motion. *ACM Transactions on Graphics (TOG) 24*, 1, 165–179.

VICON, 2016. http://www.vicon.com.

VONDRAK, M., SIGAL, L., AND JENKINS, O. C. 2008. Physical simulation for probabilistic motion tracking. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, IEEE, 1–8.

VONDRAK, M., SIGAL, L., HODGINS, J., AND JENKINS, O. 2012. Video-based 3d motion capture through biped control. *ACM Transactions On Graphics (TOG) 31*, 4, 27.

WEI, X., MIN, J., AND CHAI, J. 2011. Physically valid statistical models for human motion generation. *ACM Transactions on Graphics (TOG) 30*, 3, 19.

WITKIN, A., AND KASS, M. 1988. Spacetime constraints. In *ACM Siggraph Computer Graphics*, vol. 22, ACM, 159–168.

XIA, S., WANG, C., CHAI, J., AND HODGINS, J. 2015. Realtime style transfer for unlabeled heterogeneous human motion. *ACM Transactions on Graphics (TOG) 34*, 4, 119.

XIANG, Y., ARORA, J. S., AND ABDEL-MALEK, K. 2011. Optimization-based prediction of asymmetric human gait. *Journal of Biomechanics 44*, 4, 683–693.

YE, Y., AND LIU, C. K. 2008. Animating responsive characters with dynamic constraints in near-unactuated coordinates. In *ACM Transactions on Graphics (TOG)*, vol. 27, ACM, 112.

YEADON, M. R., AND MORLOCK, M. 1989. The appropriate use of regression equations for the estimation of segmental inertia parameters. *Journal of biomechanics 22*, 6, 683–689.

YIN, K., AND PAI, D. K. 2003. Footsee: an interactive animation system. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*, Eurographics Association, 329–338.

ZELL, P., AND ROSENHAHN, B. 2015. A physics-based statistical model for human gait analysis. In *Pattern Recognition*. Springer, 169–180.

ZHANG, P., SIU, K., ZHANG, J., LIU, C. K., AND CHAI, J. 2014. Leveraging depth cameras and wearable pressure sensors for full-body kinematics and dynamics capture. *ACM Transactions on Graphics (TOG) 33*, 6, 221.