PHYSICS-BASED PRIORS FOR HUMAN POSE TRACKING

by

Marcus Brubaker

A thesis submitted in conformity with the requirements
for the degree of Master of Science
Graduate Department of Computer Science
University of Toronto

Abstract

Physics-Based Priors for Human Pose Tracking

Marcus Brubaker

Master of Science

Graduate Department of Computer Science

University of Toronto

2006

This thesis presents a physics-based prior model of human motion suitable for use in pose tracking. This prior uses abstract physical models to capture salient aspects of human walking such as contact dynamics and the pendular nature of body motion. Because of the low dimensionality of the model, simple control strategies are successful in controlling its motion. Two different physical models are used in experiments to explore the trade-off between the complexity of the dynamics and the effectiveness of the prior. The ability of the prior to work with non-trivial walking sequences and in the presence of occlusion is demonstrated.

Dedication

To my parents, for raising me to think critically and believe in myself.

Acknowledgements

I would like to first thank and acknowledge my supervisor, David Fleet. He has been an incredibly patient mentor and astutue collaborator. His direction and support was instrumental in the origins of this thesis as well as its successful completion.

Aaron Hertzmann also played an important role in motivating and directing this research. For this, as well as for serving as the 2nd reader, I would like to thank him for this time and input.

I would also like to thank Sven Dickinson. As my research supervisor during the end of my undergraduate years he was an enthusiastic and supportive role model. Without him, I doubt I would have had the perspective, the motivation or the confidence to pursue graduate studies and a career in academics.

On a personal note I would like to first thank my best friend and beloved partner, Liv. She has been more than patient with the many long nights and weekends which were necessary to complete this work. Her sainthood has been confirmed, not only by her ability to suffer my moodiness when tired or stressed, but also by her ability to smile and laugh in spite of it. Her continued support and encouragement allows me to pursue my dreams for which I am deeply indebted.

To my family and friends I would like to say thanks for keeping me sane through it all. A special thanks to my bellydancing girls. Hanging with you is always a blast; maybe one day I really will get those sultan pants. To Danielle, thanks for the constant stream of empathy and encouraging words and, of course, the brain. To Rolf, thanks for the coffee, fantastic parties and golf outings. We will make it to the course again!

Contents

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

The problem of human pose tracking has been of interest for many years and remains a challenging problem. In computer graphics, pose tracking has been used to capture animations, typically by using multiple cameras and markers applied to known body locations. In computer vision, pose estimation and tracking remains an illustrative example of the difficulties of extracting complex, 3D information from impoverished 2D views. Outside of computer science, pose tracking is of interest in the general study of human motion, in areas such as ergonomics, physiotherapy, biomechanics and sports training.

The difficulty of the problem arises from several places, such as the limitations of data sources and the nonlinearity of the 3D model. The 2D nature of most data sources (e.g., images) leaves many parameters of pose unconstrained such as scale. Even with good data, the space of possible poses is typically over 30 dimensional and can be extremely challenging to search.

One approach to dealing with this issue is to frame the problem in terms of inference in a probabilistic model. In such a framework, the pose is an unobserved random variable and the problem determining determining the probability distribution over poses given some observed data. This allows for prior information to be integrated in the inference problem naturally. For instance, distributions over limb lengths can be specified to help

resolve scale ambiguities. It is also possible to express a prior over motions, called pose evolution priors, which can be critical for effective tracking results.

In the past, pose evolution priors have typically been expressed simply. Most models have been either simple hand specified models or learned from kinematic data [21, 34, 17, 46, 35, 49]. Unfortunately all of these models suffer from similar problems in their tracking results such as footskate and out-of-plane rotations. Footskate, which occurs when a planted foot appears to slide along the ground, is due to many factors such as measurement noise, inaccurate pose evolution models and depth-scale ambiguities. For monocular tracking the depth scale ambiguity is most prominent. This ambiguity admits a one-dimensional family of consistent observations where the scale of the object is related to its depth relative to the camera. That is, a 2D observation of an object is consistent with a small object nearby or a large object far away. Out-of-plane rotations occur when a pose is rotated about an axis which lies within the image plane of the camera. This occurs naturally when a person leans or turns away from the camera. It can also happen due to noisy observations. Out-of-plane rotations, up to a few degrees, cause only small variations in 2D observations.

These kinds of ambiguities, where multiple states are equally consistent with the observations, need to be resolved with prior models. While the works cited employ a range of prior models, they are too weak to sufficiently constrain the space of possible motions. One observation about problematic tracking results is that they violate natural expectations about the physics of the world. A planted foot, with sufficient friction, should not be able to slide around. Similarly, a person should not be able to rotate freely in space without loosing their balance. Thus, in the context of physics, these motions are highly unlikely.

This observation suggests that a prior model of pose dynamics should somehow express known physical properties of people and the world. A natural way to do this is to explicitly include physical dynamics in the specification of the pose dynamics. This

thesis presents a dynamic[1] model of human walking. This prior aims to characterize key aspects of human locomotion without undue complexity. As such it is both an abstraction of the true physics of the human body as well as a specific model of walking. The model represents an initial step towards using physically realistic models in tracking. It is hoped that physically realistic priors can better generalize to new physical situations. Priors learned from motion capture data are typically able to generalize well to motions which are similar to the data (e.g., Urtasun et al. [49]). In contrast, dynamic priors generalize to motions which are close to the physics of the underlying model (e.g., Liu et al. [23]). It is hoped that this is a more natural and useful class of generalization.

The remaining portion of this chapter will review related works in computer vision and computer graphics and provide an introduction to the biomechanics which will provide the physical basis of the model.

Chapter 2 will introduce the model, providing details about the physical models used. Along with the physical models, stochastic control models will be presented which implicitly define a distribution of motion. A kinematic model will then be presented which connects the physical dynamics to a more complete kinematic model. Finally, a model of anthropometric parameters is presented which models several sources of variability in the length of an individual.

Chapter 3 will present the details of probabilistic inference in the model. Based on standard sequential importance sampling, a method is developed which handles the state dependencies induced by the anthropometric parameters.

Chapter 4 describes the experimental setup, including the likelihood function used and the summary statistics which are used to evaluate model performance. Experimental results are presented which demonstrate the models ability to handle occlusion and generalize to motions which are not explicitly modeled.

---

[1]Throughout this thesis the use of the word "dynamic" is restricted to its physics-based meaning. For the more general meaning of the word "dynamic" as changing over time, the term "evolution" is used.

Finally, Chapter 5 concludes and discusses potential directions for future work.

1.1 Related Work

There has been a large amount of previous work which relates to this thesis. It would be impractical to review all of it here. Instead, a selection of representative work, which is most directly relavent, will be presented.

Character Animation

In character animation, the problem of realistically modeling physical motion has been studied from several different directions.

In their seminal paper, Witkin and Kass [53] introduced the notion of spacetime optimization. They formulated the problem of animating an object in a physically realistic fashion in terms of an optimization where the equations of motion of an object are viewed as constraints called spacetime constraints. Then an objective function, such as the integral of applied forces, is optimized subject to the spacetime constraints and other, potentially user-specified, constraints. Using this basic technique they were able to generate a range of motions for the classic Pixar character, Luxo, Jr..

Unfortunately spacetime optimization has had difficulty scaling up in terms of character complexity. Achieving successful convergence within a reasonable time has proven a formidable challenge for spacetime constraint based methods. Many people have worked on addressing these issues such as Liu et al. [26], Liu and Popović [22], Fang and Pollard [10], Safonova et al. [45]. In one particularly relevant work Popović and Witkin [37] attempted to address this issue in the context of motion editing by introducing the idea of motion transformations. Here an existing complex motion is transformed into a motion of a simplified physical model. Modifications are then made to the motion of the simplified model and spacetime optimization is used to generate a new motion. This

new motion is then mapped back on to the more complex model. This technique, similar in spirit to the prior model presented here, was successful in allowing realistic edits of motion capture data.

Liu et al. [23] applied a different technique to a 35 degree-of-freedom human body model with 147 individual style parameters. They viewed a motion as the result of a spacetime optimization for a given set of parameters. Then the problem becomes determining the set of parameters which generated a given motion. By approximately optimizing an objective function they are able to determine the style parameters from motion capture data. Further, by changing the constraints they are able to use these style parameters to generate new motions.

Using a different approach, Raibert and Hodgins [41] and Hodgins et al. [15] used hand designed controllers to animate characters. Raibert and Hodgins [41] animated simple models of a biped, a quadruped and a kangaroo. Hodgins et al. [15] focused on a detailed human model and created controllers which were able to generate running, cycling and vaulting animations. However, simply requiring that all actions be physically possible is not sufficient for the generation of natural looking motions. As Raibert and Hodgins [41] point out in their paper it "...seems that animals move with a smoothness and coordination that is not required by physical realism alone."

Grzeszczuk et al. [13] attempted to learn controllers rather than hand design them. Unlike other approaches, which used generate-and-test techniques to learn controllers, they were able to compute gradients of an objective function by approximating the physical dynamics with a neural network. This allowed them to compute the gradient of the simulation which was required for the gradient of the overall objective function.

People Tracking

Motion models in people tracking have tended to be simple. Very few have attempted to model physical dynamics at all and none have done so with non-trivial models. A few

representative papers will be discussed here as a backdrop for the physics-based model which is introduced in this thesis.

The use of physics-based models to motivate optimizations and other procedures has been widely used. For instance, Metaxas and Terzopoulos [32] and Terzopoulos and Metaxas [48] used physical models to motivate a method for fitting deformable shape models. While this is a novel and interesting approach, the use of physics as a metaphor is not of particular relevance to this work.

Kakadiaris and Metaxas [17] used silhouettes from multiple known camera views to estimate and track human pose. Their motion model was a simple second-order white Gaussian noise model. They were able to recover both shape and pose using their method however, multi-camera setups significantly reduce the difficulty of the pose inference problem by removing many of the ambiguities discussed above.

North et al. [33] used an auto-regressive process model of motion. They applied it to the problem of hand tracking and learned the parameters from data. However, auto-regressive process models do not scale well with the dimensionality of the hidden states. The number of parameters for such models are generally large difficult to effectively learn.

In [35] human pose parameters were treated with a 2nd order Markov model which assumed zero-mean noisy accelerations. Hybrid Monte Carlo filtering was used to do inference and help find high probability states efficiently. This model is very generic but suffers from the classic tracking problems mentioned above such as footskate and out-of-plane rotations.

Urtasun et al. [50] used a Gaussian Process latent variable model to learn a mapping between the high-dimensional pose space and a low-dimensional latent space. Urtasun et al. [49] extended this work by simultaneously learning a Gaussian Process of the dynamics in the latent space. Trained with motion capture data both systems were able to track successfully using limited observations but still suffer from the common problems of footskate and out-of-plane rotations.

While not strictly tracking people, the motion model used by Isard and Blake [16] was physically based. They tracked a bouncing ball with a physical model which assumed that acceleration was due to gravity plus some Gaussian noise. This model was interrupted by ground collisions which were elastic due to the nature of the ball. An edge-based image likelihood was used with the Condensation algorithm to perform inference and track the motion of the ball.

Pavlovic et al. [34] attempted to model the non-linearities of motion by using a switching linear dynamic system (SLDS). The SLDS parameters were fit to hand-labeled walking and running data using an EM-like procedure. The resulting parameters were used to classify and track new motions as well as generate novel motions.

Bissacco and Soatto [1] used a switching auto-regressive process model of pose dynamics for the purposes of motion analysis and identification. The switch in their linear system is used to model the change in dynamics at ground contact points. They also developed a probabilistic distance metric for their model to allow comparison and between motions. While not actually used for tracking, their model of motion could have been used to learn a prior for Bayesian tracking.

1.2 Physiology, Biomechanics and Engineering

Over the years, there have been many models of human motion proposed for varying purposes. Engineers have tried to characterize bipedal motion in terms of stability criteria which are used to restrict motion planning in robots. Researchers in physiology and physiotherapy have been interested in understanding the energetics and characteristics human motion in order to help the rehabilitation of people with pathological gaits. Biomechanicists have been interested in the human gait as an example of efficient bipedal locomotion. All of these areas have contributed to the range of models available for explaining human gait. It would be impossible to provide a complete overview of each of

these areas however a brief overview will be given here.

1.2.1 Robotics

The Honda humanoid robot, *Asimo*, is probably the most famous bipedal robot and is exemplary of the state of the art in robotics engineering. As Hirai et al. [14] explain, *Asimo* combines two techniques to generate motions. First, they use the common stability criterion known as the zero moment point (ZMP) to define a class of balanced states. Second, they use recordings of human motions to suggest desired motion trajectories. Together, their planning algorithm attempts to mimic the recorded motion trajectories while always maintaining balance.

The ZMP is defined to be the point on the ground plane where the angular momentum of the robot about the forward and side axes is zero. If the ZMP is located within the convex hull of ground contact points, the robot is considered to be stable. The ZMP has a long history and was first formally introduced for the study of human motion by Elftman [9] under a different name. Vukobratovic and Juricic [52] suggested its use in robotics and named it the "Zero Moment Point." Unfortunately, robots which use ZMP based stability criteria have a characteristically inhuman gait [38] and are highly inefficient. One estimate for *Asimo* indicates that its dimensionless cost of transport[2] is more than ten times that of a human [6]. While inefficiency may not be of particular concern for tracking, the abnormality of gait of ZMP based control strategies makes them a poor starting point for a model of human motion. While other stability criterion exist, they are less widely used and tend to be qualitatively similar to the ZMP [36].

(a) The Monopode

(b)   The   Anthropomorphic
Walker
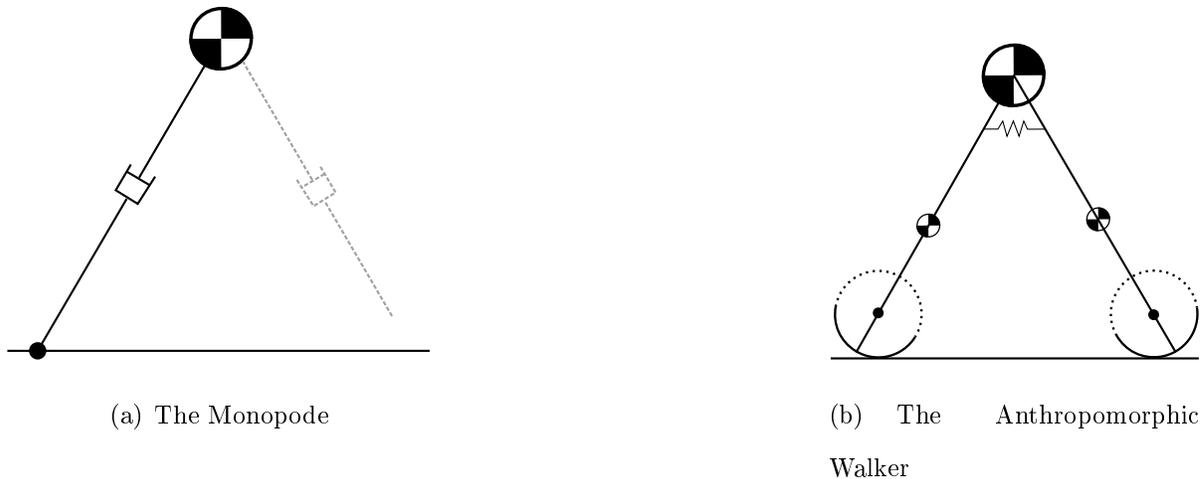
Figure 1.1: Two Abstract Models of Walking. The hatched circles represent the physical centers of mass.

1.2.2   Abstract Models of Locomotion

Physiologists have developed an extremely simple, yet highly general model of biological motion known as the monopode. The monopode, illustrated in Figure 1.1(a), is a highly abstracted model where the entire body is represented by a point-mass located approximately at a hip and the legs are represented by massless, prismatic joints. When a leg is in contact with the ground it is able to apply force through the prismatic joint and influence the motion of the body. When a leg is not in contact with the ground it is able to instantaneously be in any orientation and length due to its massless nature. Blickhan and Full [2] compared the monopode to the motion of a wide class of animals, including humans, horses, kangaroos and cockroaches. Such a widely applicable model is extremely appealing for the creation of a more general prior model of motion. More recently, Ruina et al. [44] studied the monopode as a collisional model of locomotion. Building on that work, Srinivasan and Ruina [47] determined that an approximately passive control strategy with impulsive collisions results in human like gait patterns for both running and walking. The monopode has been physically manifested as a robot in the work of

---

[2]The cost of transport is the amount of energy required to move a unit amount of mass a unit distance.

Raibert [40].

A more complicated model is the anthropomorphic walker [27, 20]. This model, shown in Figure 1.1(b), features two straight legs with a torso mass attached at the hip. The "feet" are circles of fixed radius which roll as the model moves. The stability and dynamics of this model were studied extensively as a passive model of motion by McGeer [31]. This model features walking as a natural limit cycle of its dynamics and provides a basis for understanding efficient walking. The analysis performed by McGeer has been the basis of many bipedal robots which exhibit natural gaits [6] and is a potentially natural basis for a prior model. The anthropomorphic walker can also be powered in order to walk on level ground. One such variant has a spring between the legs and impulses are applied along the back leg prior to support transfers. This model has been studied by Kuo [19, 20] and has been able to predict the preferred speed-step length relationship found in human walking. Several other variations on the anthropomorphic walker exist such as a version with a kneed swing leg [28], another with an active torso [30] and even a variant capable of running [29].

The principles of these models have been successfully applied in the creation of efficient control strategies for bipeds [39] as well as in the creation of real robots [5]. These robots have relatively realistic gaits and a human-like cost of transport [6].

Chapter 2

A Model of Human Locomotion

Both the monopode and the anthropomorphic walker capture some of the more salient features of human walking. Figure 2.1(a) shows the trajectory of a human during normal walking and illustrates several of these features. The most prominent feature is the pendular nature of the torso trajectory. Similar to the torso, the supporting leg follows a pendular arc. The support leg can be easily identified by this feature, as its arc peaks at the same point that the torso peaks. Slightly more subtle is the inverse pendular arc of the swing leg. This corresponds to a dip in the trajectory of the leg, typically with its lowest point occurring around the same position as the highest point of the hip and stance leg.

All of these features are cyclic and separated by discontinuities. There are two sources of these discontinuities which both occur at approximately the same time. One is the collision of the swing leg with the ground. This collision is relatively inelastic and results in a sudden change in velocities. The other source is a "toe-off" which usually begins just before the swing leg hits the ground. Toe-off is when the supporting leg is used to push off of the ground. The mechanical purpose of toe-off is primarily to minimize energy loss from the collision of the swing leg with the ground by redirecting the momentum of the body. Analysis of simple bipedal models has shown that the optimal time for this
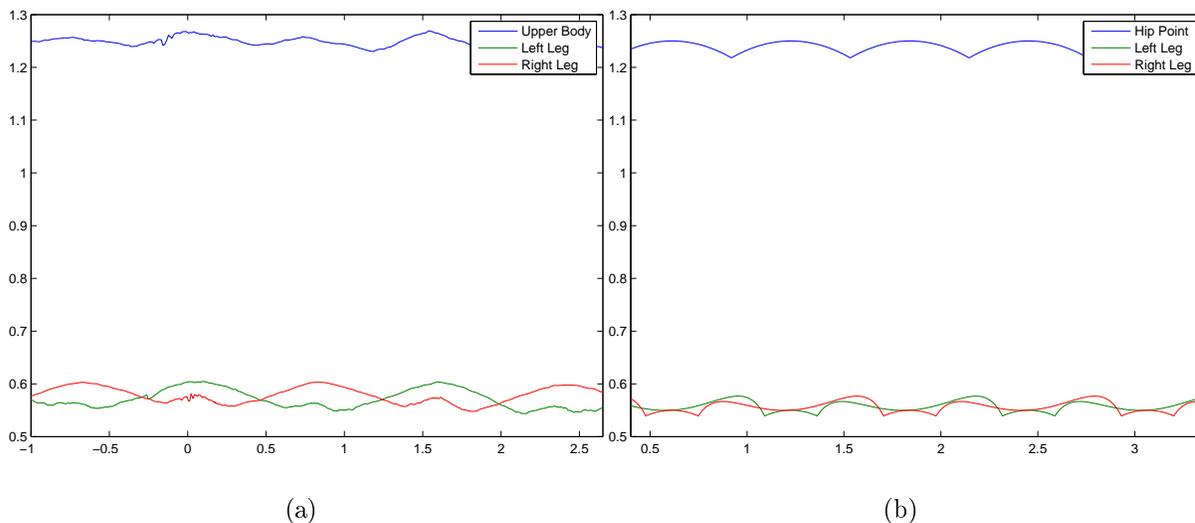
Figure 2.1: Characteristics of Human Walking. (a) shows the trajectory of the center of mass of upper body and legs of a human during normal walking. This plot is based on motion capture data from the CMU Motion Capture Database (`http://mocap.cs.cmu.edu`). (b) shows the trajectory of the point masses on the hip and legs of the anthropomorphic walker during a passive gait cycle. Notice the regular, pendular like paths in both plots which are interrupted by sudden discontinuities when a leg strikes the ground and support is transferred.

push-off is immediately before the swing leg hits the ground [31]. If toe-off does not occur until after impact, the energy has already been lost; if it occurs too soon before impact, energy is wasted pushing against gravity.

Each of these features of walking is captured by the anthropomorphic walker. A plot of its centers of mass is presented in Figure 2.1(b) for comparison with the human data. In this plot, it is walking passively (i.e., without any added forces except gravity) down a slight grade in a stable, cyclic gait. The pendular trajectories of the hip and legs can be clearly seen as well as the discontinuities. The trajectories of the legs show one aspect of human walking which is not well modeled, the knee. Because the anthropomorphic walker does not have a knee, the swing leg is typically too high and takes a seemingly unusual trajectory towards the end of its swing. This illustrates one of the challenges of using abstract motion models in tracking: the connection between the model and real data is not always clear.

A physical model provides a set of dynamics which can be simulated in a deterministic fashion. To create a distribution over motions some form of stochastic dynamics must be expressed. The remaining sections of this chapter present the details of the model. The deterministic and stochastic dynamics of the two models are presented in Section 2.1. The relationship of the models to a human pose, along with the evolution of unconstrained pose parameters, is presented in Section 2.2. Finally, a model of anthropometric parameters is presented in Section 2.3.

2.1   Walking Model Dynamics

For a physical model, a set of equations of motion can be derived which define the motion of the model as a function of applied forces [12]. These equations are typically second order, ordinary differential equations which need to be integrated to find the motion of

the model. In general terms, then, a physical model defines the equation

$$\ddot{q} = h(q, \dot{q}, f)$$

where f is an external force vector, q is a vector of the model coordinates, $\dot{q}$ are their velocities and $\ddot{q}$ are their accelerations. Given an initial pose q_0 and velocity $\dot{q}_0$ and an appropriately defined force function $f(t)$, the equations of motion can be integrated to find

$$
\begin{aligned}
q(t) &= \int_{-\infty}^{t} \dot{q}(s)ds \\
&= \int_{-\infty}^{t} \left(\int_{-\infty}^{s} h(q(\tau), \dot{q}(\tau), f(\tau))d\tau \right) ds \\
&= q_0 + t\dot{q}_0 + \int_{0}^{t} \int_{0}^{s} h(q(\tau), \dot{q}(\tau), f(\tau))d\tau ds
\end{aligned}
$$

the state of the system at time t. In general, these integrals are intractable and numerical approximations must be used.

While a large number of numerical integration techniques exist (see [18] for an introduction), the classical fourth order Runge-Kutta method is used, as it is reasonably fast and accurate. Almost any numerical integration technique could be used for the equations of motion presented here as they are not particularly stiff.

However, many equations of motion can be stiff and will not behave well with standard integration techniques. For instance, the standard application of Newton's laws on an articulated system enforces ground and joint constraints with springs [8]. This introduces a trade-off between the accuracy of constraint enforcement and the stiffness of the equations. In early experiments for this work, it was determined that spring-based constraint enforcement rarely produced satisfactory results and such equations of motion should be avoided to the extent possible. To this end, other methods of deriving equations of motion should be used which explicitly enforce constraints. One general method is the use of Lagrange's equation. However, for articulated rigid bodies a more obvious method can be used known as the TMT method [51]. While roughly equivalent to using

the Lagrangian, the TMT method provides a straightforward recipe for finding equations of motion for articulated bodies. When combined with an impulsive collision model, the TMT method produces well-behaved equations of motion and stable simulations.

The numerical integration of physical equations of motion corresponds to the deterministic portion of a this model. While a general approach to making the dynamics stochastic would be to add noise to the results of an integration step, the net result of this is unclear and moves towards a model where physically unrealistic things are possible. A better approach is to add noise to the forces $f(t)$ which affect the system. This ensures that the resulting motion is always physically possible and generally realistic, provided that the forces are reasonable. With this approach it is straightforward to sample from the stochastic dynamics.

Each physical model generally has several parameters such as masses and abstract limb lengths which are meant to be fixed. These parameters are described and given in the text and are fixed for all subjects. While the use of fixed parameters may seem restrictive, many of them are redundant and changes in others would make only a small difference in the resulting dynamics. For instance, absolute mass is irrelevant because it is always possible to rescale the forces applied to a body to take in to account any absolute changes in mass. Relative mass between segments, on the other hand, is crucial but does not seem likely to vary much between subjects and thus is reasonable to set to a fixed value. Reasonable variations in length parameters have only a minimal impact on overall dynamics due to the connection with gravity. Specifically, the accelerations of bodies may be slightly off but in the course of the present work it has not been found to be a significant problem.
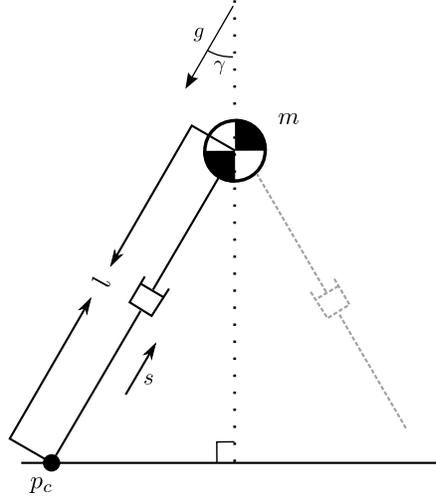
Figure 2.2: The Monopode

2.1.1   The Monopode

For the monopode, the model's degrees of freedom, $q = [x, y]^T$, are the coordinates of the point mass. The equations of motion are

$$m\ddot{q} = mg \begin{bmatrix} \sin\gamma \\ -\cos\gamma \end{bmatrix} + f(t)s \tag{2.1}$$

where m is the mass of the point, $f(t)$ is the force applied along the stance leg, g is the acceleration due to gravity, γ is the orientation of the ground with respect to gravity, $p_c = (x_c, y_c)$ is the stance leg's ground contact point, $l = \sqrt{(x - x_c)^2 + (y - y_c)^2}$ is the leg length and $s = \frac{1}{l} [x - x_c, y - y_c]^T$ is the stance leg direction vector. These values are illustrated in Figure 2.2.

The leg has a maximum length $l^{\max}$ which is treated as a hard constraint. The constraint is enforced at integration boundaries by assuming collisions with the constraint to be both impulsive and inelastic. This results in an instantaneous change in velocity with the post-collision velocity

$$\dot{q}^+ = \dot{q}^- - \max(0, s \cdot \dot{q}^-)s \tag{2.2}$$

and the post-collision position

$$q^+ = p_c + l^{\max}s \tag{2.3}$$

where $\dot{q}^-$ is the pre-collision velocity. These equations simply satisfy the length constraint and remove any velocity which would violate it.

With a force function $f(t)$, generating a motion from the monopode then consists of integrating Equation (2.1) and, when $l > l^{\max}$, enforcing the leg length constraint with Equations 2.2 and 2.3. Because the monopode does not have a meaningful representation of the swing leg, it must be externally decided when support transfer happens and where the new stance position is located. Srinivasan and Ruina [47] choose these quantities and optimize a force trajectory which would cause the motion to be cyclic and minimize the work performed. The resulting force trajectory showed a constant force exerted throughout the stance with nearly impulsive forces at the beginning and end of a stride. This result justifies a simple, constant force function

$$f(t) = c \tag{2.4}$$

with matched impulses ι applied at support transfer. Matched impulses result in an instantaneous change of velocity by ιs^- before transfer and by ιs^+ after transfer where s^- and s^+ are the pre- and post-transfer support leg directions. Hence the post-transfer velocity will be

$$\dot{q}^+ = \dot{q}^- + \iota(s^+ + s^-)$$

where $\dot{q}^-$ is the pre-transfer velocity.

These equations provide the basis to deterministically simulate with the monopode model. However, as mentioned earlier, a probabilistic model is needed. To achieve that the forces are assumed to have some stochastic element. To create a stochastic model of force it is first assumed that only a single stride is of interest and that the constant c is known. Then Equation (2.4) is modified to become

$$f(t) = c + \eta_f$$

where η_f is zero mean IID Gaussian noise with a variance of σ_f^2. Thus

$$P(f(t)|c) = \mathcal{N}(f(t); c, \sigma_f^2)$$

is the probability density of $f(t)$ where

$$\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

is the probability density function of a Gaussian distribution with mean μ and variance σ^2.

In general, more than a single stride is of interest and c is not known. To address this, a model of how c changes over time is needed. Let $s(t)$ be the stride index at time t and $c(s(t))$ be the force constant at time t. Because c is a manifestation of the speed and style of a stride $c(s)$ is assumed to change slowly from stride to stride and remain close to some "mean" value c_0. To model this, the equation

$$c(s) = \alpha c(s - 1) + (1 - \alpha)c_0 + \eta_c$$

is used where $\alpha \in [0, 1]$ controls how close $c(s)$ stays to the "mean" value c_0 and η_c is mean zero IID Gaussian noise with variance σ_c^2. Thus

$$P(c(s)|c(s - 1)) = \mathcal{N}(c(s); \alpha c(s - 1) + (1 - \alpha)c_0, \sigma_c^2)$$

is the probability density of $c(s)$ given $c(s - 1)$. The impulse ι is modeled as a gamma distribution such that

$$P(\iota) = \mathcal{G}(\iota; s_\iota, \lambda_\iota)$$

where

$$\mathcal{G}(x; s, \lambda) = \frac{\lambda^s x^{s-1}}{\Gamma(s)} e^{-\lambda x}$$

is the gamma density function with shape s and scale λ.

The physical model parameters used can be found in Table 2.1 and the stochastic parameters are in Table 2.2.

Parameter	Value	Description
m	1	Mass of the monopode.
l^{max}	1	Maximum leg length.
g	9.81	Gravitational acceleration.
γ	0	Ground slope angle.

Table 2.1: Physical Parameters of the Monopode

Parameter	Value	Description
σ_f	2.0	Force sample noise.
σ_c	1.0	Force constant mean process noise.
c_0	7.0	Mean force constant.
α	0.5	Mean preference factor.
s_ι	$\frac{(1.0)^2}{(0.5)^2}$	Impulse shape parameter.
λ_ι	$\frac{1.0}{(0.5)^2}$	Impulse scale parameter.

Table 2.2: Stochastic Monopode Parameters. The impulse parameters are set so that $E[\iota] = 1.0$ and $V[\iota] = (0.5)^2$.
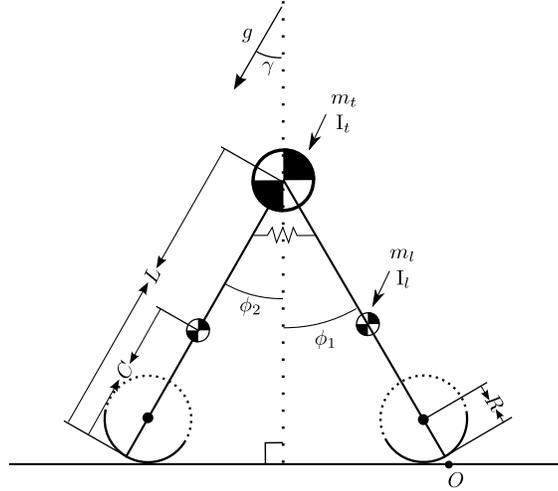
Figure 2.3: The Anthropomorphic Walker. L is the length of the leg, R is the radius of the foot, m_l is the mass of a leg, m_t is the mass of the torso, I_l is the moment of inertia of a leg, I_t is the moment of inertia of the torso and C is the distance from the end of a leg to its center of mass.

2.1.2 The Anthropomorphic Walker

For the anthropomorphic walker [27, 11], the equations of motion are somewhat more complicated. To enforce ground and joint constraints generalized model coordinates $q = [\phi_1, \phi_2]^T$ are used where, as shown in Figure 2.3, ϕ_1 is the global orientation of the stance leg and ϕ_2 is the global orientation of the swing leg. Derived in Appendix A, the equations of motion can be summarized here as

$$\mathbf{T}^T M \mathbf{T} \ddot{q} = f(t) + \mathbf{T}^T M (G - \mathbf{g})$$

where

$$
\mathbf{T} = \begin{bmatrix}
-R - (C_1 - R)\cos\phi_1 & 0 \\
-(C_1 - R)\sin\phi_1 & 0 \\
1 & 0 \\
-R - (L - R)\cos\phi_1 & (L - C)\cos\phi_2 \\
-(L - R)\sin\phi_1 & (L - C)\sin\phi_2 \\
0 & 1
\end{bmatrix}
$$

is the first order kinematic transfer matrix,

$$
M = \begin{bmatrix}
m_1 & 0 & 0 & 0 & 0 & 0 \\
0 & m_1 & 0 & 0 & 0 & 0 \\
0 & 0 & I_1 & 0 & 0 & 0 \\
0 & 0 & 0 & m_l & 0 & 0 \\
0 & 0 & 0 & 0 & m_l & 0 \\
0 & 0 & 0 & 0 & 0 & I_l
\end{bmatrix}
$$

is the mass matrix,

$$
\mathbf{g} = \begin{bmatrix}
\dot{\phi}_1^2(C_1 - R)\sin\phi_1 \\
-\dot{\phi}_1^2(C_1 - R)\cos\phi_1 \\
0 \\
\dot{\phi}_1^2(L - R)\sin\phi_1 - \dot{\phi}_2^2(L - C)\sin\phi_2 \\
-\dot{\phi}_1^2(L - R)\cos\phi_1 + \dot{\phi}_2^2(L - C)\cos\phi_2 \\
0
\end{bmatrix}
$$

is the convective acceleration,

$$
G = g \begin{bmatrix}
\sin\gamma \\
-\cos\gamma \\
0 \\
\sin\gamma \\
-\cos\gamma \\
0
\end{bmatrix}
$$

is the gravitational acceleration vector,

$$m_1 = m_l + m_t$$

is the combined mass of the stance leg,

$$C_1 = \frac{(Cm_l + Lm_t)}{m_l + m_t}$$

is the location along the leg of the center of combined mass,

$$I_1 = I_l + I_t + (C_1 - C)^2 m_l + (L - C_1)^2 m_t$$

is the combined moment of inertia and $f(t)$ is the generalized force vector. L, R, C, m_l, m_t, I_l and I_t are illustrated and described in Figure 2.3.

As before, with a force function $f(t)$, these equations can be integrated to determine a motion. Unlike the monopode, the anthropomorphic walker has a meaningful swing leg model which can indicate when a collision should occur. In particular, the end of the swing leg is even with the ground when $\phi_1 = -\phi_2$. Potential collisions can thus be found by detecting zero crossings of the function $\mathcal{C}(\phi_1, \phi_2) = \phi_1 + \phi_2$. Zero crossings of $\mathcal{C}$ will occur both when the swing foot hits the ground and, if it was already beneath the ground, when it exits the ground. Because the swing leg does not have a knee which could prohibit stubbing, where the foot hits the ground too early in the stride, zero crossings of $\mathcal{C}$ are only considered an actual collision when $\phi_1 < 0$ and $\mathcal{C}$ is decreasing.

Once it has been determined that a collision has occurred the result of that collision must be determined. By assuming ground collisions to be impulsive and inelastic the result can be determined by solving a set of equations for the post-collision velocity. To model toe-off before such a collision, an impulse along the stance leg is added. The post-collision velocities can then be solved for using

$$\mathbf{T}^{+T} M \mathbf{T}^+ \dot{q}^+ = \mathbf{T}^{+T} (S + M \mathbf{T} \dot{q}^-)$$

where $\dot{q}^-$ and $\dot{q}^+$ are the pre- and post-collision velocities, $\mathbf{T}$ is the pre-collision kinematic transfer matrix specified above,

$$
\mathbf{T}^+ =
\begin{bmatrix}
(L-C)\cos\phi_1 & -R-(L-R)\cos\phi_2 \\
(L-C)\sin\phi_1 & -(L-R)\sin\phi_2 \\
1 & 1 \\
0 & -R-(C_1-R)\cos\phi_2 \\
0 & -(C_1-R)\sin\phi_2 \\
0 & 1
\end{bmatrix}
$$

is the post-collision kinematic transfer matrix, M is the mass matrix as above and

$$
S = \iota
\begin{bmatrix}
-\sin\phi_1 \\
\cos\phi_1 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
$$

is the impulse vector with magnitude ι. Finally, the origin of the model will have shifted forward by a distance equal to $2(R\phi_2 + (L-R)\sin\phi_2)$. After the collision has occurred the swing and stance leg should then switch roles and, accordingly, ϕ_1 and ϕ_2 should be swapped along with their velocities.

Kuo [19, 20] analyzed this model and determined that the impulsive push-off combined with a torsional spring between the legs permits the generation of a range of walking motions on level and slanted ground. Thus the generalized force vector is

$$
f(t) = \kappa(\phi_2(t) - \phi_1(t))
\begin{bmatrix}
1 \\
-1
\end{bmatrix}
\tag{2.5}
$$

where κ is the spring stiffness.

The stochastic force function for the anthropomorphic walker is modelled in a similar way as the monopode's. Let $s(t)$ be the stride number at time t. Then κ in Equation

(2.5) is an IID sample $\kappa(t)$ from a Gaussian distribution with sample variance σ_κ^2 and stride specific mean $\bar{\kappa}(s)$. Thus,

$$P(\kappa(t)|\bar{\kappa}(s(t))) = \mathcal{N}(\kappa(t); \bar{\kappa}(s(t)), \sigma_\kappa^2)$$

is the probability desnity of $\kappa(t)$. The mean spring constant, $\bar{\kappa}(s)$, is modeled as a slowly-changing value over multiple strides which remains close to some mean value. Therefore, its evolution is modeled by

$$P(\bar{\kappa}(s)|\bar{\kappa}(s-1)) = \mathcal{N}(\bar{\kappa}(s); \alpha\bar{\kappa}(s-1) + (1-\alpha)\bar{\kappa}_0, \sigma_{\bar{\kappa}}^2)$$

where $\sigma_{\bar{\kappa}}^2$ is the process variance. Also like the monopode, the impulse ι is modeled with a gamma distribution such that

$$P(\iota) = \mathcal{G}(\iota; s_\iota, \lambda_\iota)$$

where s_ι and λ_ι are the shape and scale parameters.

The physical parameters used can be found in Table 2.3 and the stochastic parameters are in Table 2.4.

2.2 Kinematic State Evolution

To connect the abstracted physical dynamics to the motion of real people, a kinematic model needs to be specified. This model needs to be constrained by the physical dynamics to the extent possible. The kinematic model used is a simple model of the lower body and is illustrated in Figure 2.4. Because the kinematic model is more complex, state evolution models need to be specified for aspects of the kinematic model which are not constrained by the dynamics.

There are two kinds of quantities in the kinematic model which need to be specified to have a complete pose. The joint angles specify the orientation of the joints between segments. In contrast, segment lengths specify the distance between these joints. Joint

Parameter	Value	Description
R	0.3	Radius of the foot.
L	1	Leg length.
C	0.645	Leg center of mass offset.
m_l	0.161	Leg mass.
I_l	0.017	Leg moment of inertia.
m_t	0.678	Torso mass.
I_t	0.167	Torso moment of inertia.
g	9.81	Gravitational acceleration.
γ	0	Orientation of gravity.

Table 2.3: Physical Parameters of the Anthropomorphic Walker. Mass and inertial parameters are based on Dempster's body segment parameters found in [43]. L, R and C are taken from [20].

Parameter	Value	Description
σ_κ	1.0	Spring stiffness sample noise.
$\sigma_{\bar{\kappa}}$	0.75	Spring stiffness mean process noise.
$\bar{\kappa}_0$	2.0	Mean spring stiffness.
α	0.5	Mean preference factor.
s_ι	$\frac{(0.4)^2}{(0.15)^2}$	Impulse shape parameter.
λ_ι	$\frac{0.4}{(0.15)^2}$	Impulse scale parameter.

Table 2.4: Stochastic Anthropomorphic Walker Parameters. The impulse parameters are set so that $E[\iota] = 0.4$ and $V[\iota] = (0.15)^2$.
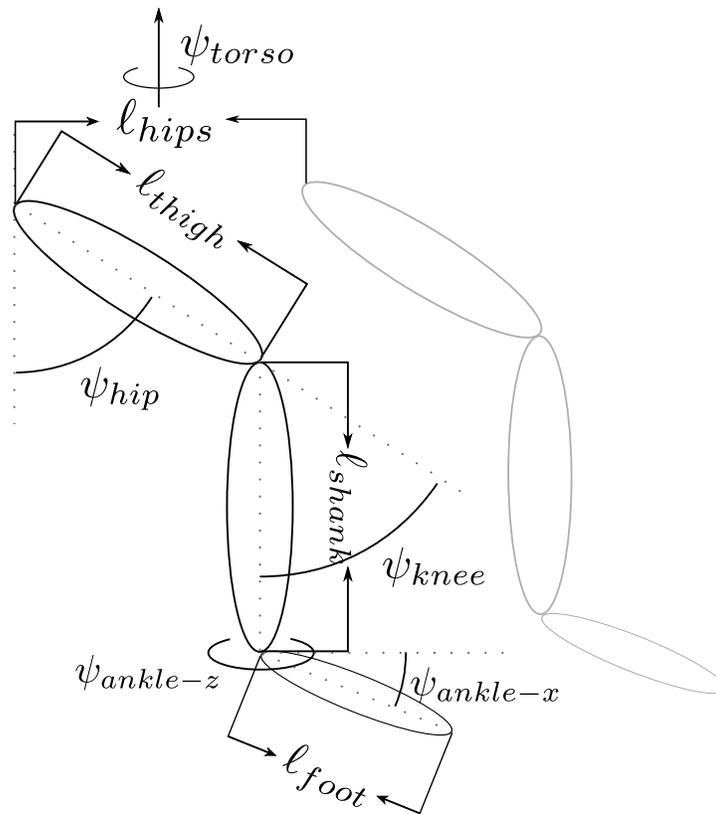
Figure 2.4: Kinematic Model of the Lower Body

$$\psi_{hip} = \phi$$

$$\psi_{hip} = \phi - \psi_{knee}$$
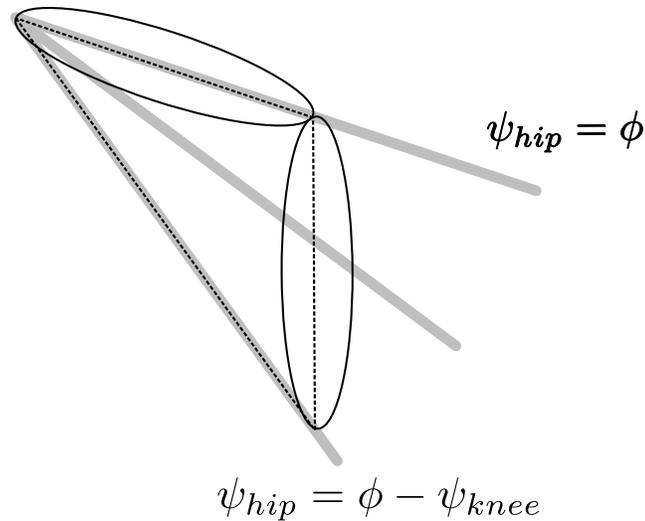
Figure 2.5: The Connection Between Dynamics and Kinematics. The dashed lines are the convex hull which contain the center of mass for the leg. The light grey lines are three possibilities for the orientation of the underlying dynamics given the kinematics which are consistent with the convex hull for the center of mass.

angles, as well as global position and orientation, are partly constrained by the dynamics and are otherwise expected to evolve smoothly. Anthropometric parameters such as segment lengths are expected to be relatively constant over time for a given person. Because of this difference these quantities are modeled separately with the evolution of joint angles and global coordinates being specified next and the modeling of anthropometric parameters being specified in Section 2.3.

The physical models provide a single leg angle ϕ. The kinematic model however, has hip, knee and ankle angles which need to be constrained. One way to view the leg angle ϕ is as the orientation of the center of mass of the leg about the hip joint. Because the center of mass of a combined body is a convex sum of the individual centers of mass then, assuming that the centers of mass of the thigh and shank are contained within their imagined geometry, the combined center of mass must be within the convex hull of the geometries. Thus, for a physical leg angle ϕ and kinematic knee angle ψ_{knee}, the hip

angle ψ_{hip} must satisfy the inequalities

$$
\begin{cases}
\phi \le \psi_{hip} \le \phi - \psi_{knee} & \text{for } -\pi \le \psi_{knee} \le 0 \\[2mm]
\phi - \psi_{knee} \le \psi_{hip} \le \phi & \text{for } 0 \le \psi_{knee} \le \pi
\end{cases}
$$

which is illustrated in Figure 2.5.   While ψ_{hip} can be solved for exactly using inverse kinematics, a trivial approximation is simpler to use.   The hip angle is approximated as

$$
\psi_{hip} \approx \phi
$$

for given a leg angle ϕ.   Several approximations were tried and this was found to work best.   Conceptually, this assumes that the underlying dynamics represent the upper leg and that the lower leg is generally irrelevant to the motion of the upper leg.

This leaves the knees, ankles and global orientation unconstrained by the dynamics. The evolution of these joints is modeled as second-order Markov.   To avoid actually having a second-order model the state space is augmented to include both the joint angles ψ and the joint velocity $\dot{\psi}$.   The model includes a form of damping by only using a fraction α of the velocity at the previous time to update the current joint angle.   It also includes a "preferred orientation" ψ^0 which the joint is pulled towards with a strength κ.

Including all these terms and some amount of Gaussian noise, a joint angle ψ evolves such that

$$
\psi(t) = \psi(t-1) + \tau(\alpha\dot{\psi}(t-1) + \kappa(\psi^0 - \psi(t-1))) + \eta_\psi
$$

where τ is the size of the timestep and η_ψ is IID Gaussian noise with mean zero and variance σ^2.   This model is equivalent to a damped spring model with noisy accelerations where κ is the spring constant, ψ^0 is the rest position, α is related to the damping constant and η_ψ is noisy acceleration.

To account for joint limits which require that $\psi^{\max} \le \psi \le \psi^{\min}$ η_ψ is truncated so that the inequality is always satisfied.   This gives a truncated normal distribution for

$\psi(t)$

$$P(\psi(t)|\psi(t-1), \dot{\psi}(t-1)) = \mathcal{N}(\psi(t); \mu, \sigma^2, [\psi^{\min}, \psi^{\max}]) \tag{2.6}$$

where

$$\mu = \psi(t-1) + \tau(\alpha\dot{\psi}(t-1) + \kappa(\psi^0 - \psi(t-1)))$$

is the deterministic mean component,

$$\mathcal{N}(x; \mu, \sigma^2, [\mu^-, \mu^+]) = \left(\Phi\left(\frac{\mu^+ - \mu}{\sigma}\right) - \Phi\left(\frac{\mu^- - \mu}{\sigma}\right) \right)^{-1} \left(2\pi\sigma^2\right)^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

is the density function of a Gaussian distribution with mean μ and variance σ^2 truncated to lie in the range $[\mu^-, \mu^+]$ and

$$\Phi(x) = \int_{-\infty}^{x} \mathcal{N}(x; 0, 1)dx$$

is the cumulative density function of the standard normal distribution. Once the new joint angle has been chosen, the velocity is then deterministically updated to be

$$\dot{\psi}(t) = \tau^{-1}(\psi(t) - \psi(t-1))$$

where the multiplication by τ^{-1} allows $\dot{\psi}$ to be interpretted in units of radians per second. Some sample plots of Equation (2.6) can be found in Figure 2.6 on page 30.

To help further constrain the kinematic evolution some special properties are added to this general model. The knee of the swing leg has a resting length ψ^0 which is dependent on the orientation of the stance leg.

$$\psi^0_{knee-swing}(t) = \min(0, -\psi_{hip-stance}(t-1))$$

This serves to model the bend of the swing knee immediately after support transfer which is followed by a gradual straightening. The ankle of the stance leg also has a variable resting length which is dependent on the orientation of the other stance leg joints.

$$\psi^0_{anklex-stance}(t) = -(\psi_{knee-stance}(t-1) + \psi_{hip-stance}(t-1))$$
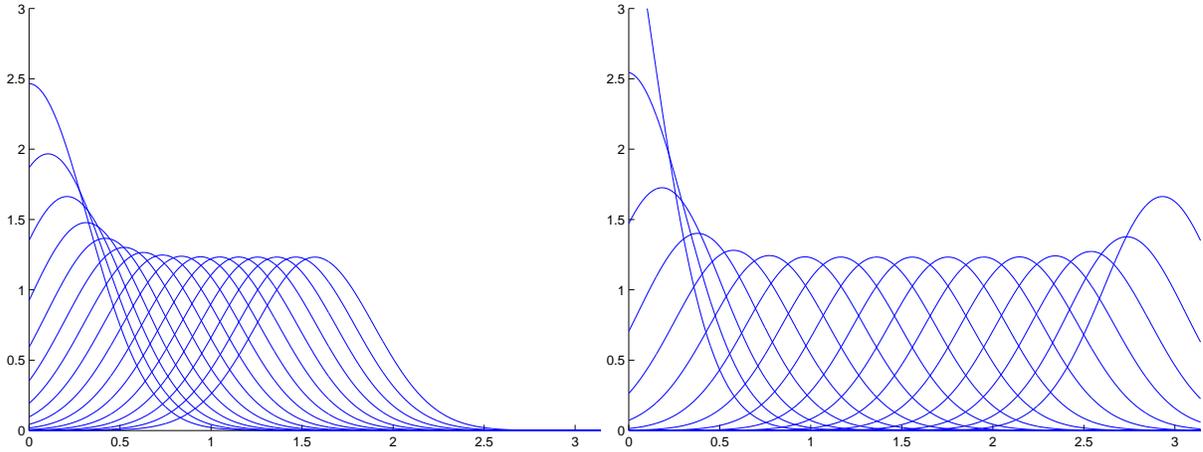
(a) $P(\psi_t)$ given $\psi_{t-1} = \frac{\pi}{6}$ and $\dot{\psi}_{t-1} = -5\pi, \ldots, 8\pi$ (b) $P(\psi_t)$ given $\dot{\psi}_{t-1} = -2\pi$ and $\psi_{t-1} = 0, \ldots, \pi$

Figure 2.6: Truncated Normal Distributions over Joint Angles.

This helps keep the foot level with the ground on the stance leg. The Z-axis rotation of the stance leg ankle is fixed such that the foot remains pointing the direction it was pointing when it hit the ground, ensuring that the foot is not allowed to "slip" while it is supporting the body.

The global orentation is assumed to be parallel to the ground plane normal with a rotation about this axis which is treated as other joint angles. The global position of the walker is indirectly constrained by the dynamics of the physical models. Both models require that the stance foot remain in contact with the ground. Precisely where the ground contact point is on the body is not clear but a simple assumption is that the heel of the foot is the ground contact point. Then the global position of the root node of the kinematic tree is set such that these points are coincident after the joint angles have been updated. When ground contact is detected the location of the contact point on the kinematic model is updated to the point on the ground closest to the heel of the new stance leg.

The values of the parameters σ, α, κ and ψ^0 for each joint can be found in Table 2.5.

Joint	Angle Noise $\tau^{-1}\sigma$	Velocity Decay α	Spring Stiffness κ	Resting Orientation ψ^0	Joint Limits $\psi^{\min}, \psi^{\max}$
Torso	0.2	1	0	N/A	None
Hip	π	0.75	0	N/A	None
Knee	$\frac{\pi}{4}$	1	2	0/See Text	$-\pi, 0$
Ankle-X Axis	$\frac{\pi}{4}$	0.75	2	0/See Text	$-\frac{\pi}{2}, \frac{\pi}{2}$
Ankle-Z Axis	$\frac{\pi}{4}$	0.75	5	0/See Text	$-\frac{\pi}{2}, \frac{\pi}{2}$

Table 2.5: Joint Model Parameters.

2.3   Anthropometric Parameters

A model of anthropometric parameters, such as segment lengths, is needed as well. These quantities are strictly positive and should be relatively constant over time. Unfortunately, the assumption made with the kinematic model that the distance between joint centers is constant is not entirely realistic. The connective tissue between bones is elastic which provides small variations in segment length. Thus a distribution over segment lengths is needed which can provide a small variance around a mode which is close to the mean. To this end, a gamma distribution over segment length $\ell(t)$ at time t is used

$$P(\ell(t)|s_\ell, \lambda_\ell) = \mathcal{G}(\ell(t); s_\ell, \lambda_\ell) \tag{2.7}$$

where

$$\mathcal{G}(x; s, \lambda) = \frac{\lambda^s x^{s-1}}{\Gamma(s)} e^{-\lambda x}$$

is the gamma density function with shape s and scale λ. Then $\ell(t)$ has mean

$$E[\ell(t)|s_\ell, \lambda_\ell] = \frac{s_\ell}{\lambda_\ell}$$

variance

$$V[\ell(t)|s_\ell, \lambda_\ell] = \frac{s_\ell}{\lambda_\ell^2}$$

coefficient of variation

$$CV[\ell(t)|s_\ell, \lambda_\ell] = \frac{1}{\sqrt{s_\ell}}$$

and mode located at

$$\arg\max_{\ell(t)} P(\ell(t)|s_\ell, \lambda_\ell) = \frac{s_\ell - 1}{\lambda_\ell}$$

for $s_\ell > 1$. For large s_ℓ (or, equivalently, a small coefficient of variation) the mode is approximately equal to the mean. Intuitively, s_ℓ and λ_ℓ should be fixed for a given segment of a specific person.

To account for variations between people a prior distribution over the parameters can be specified. Because the variability over people is one of length, or scale, it seems

sensible to place the prior over λ_ℓ. For convenience a gamma distribution is again used. Therefore,

$$P(\lambda_\ell) = \mathcal{G}(\lambda_\ell; \bar{s}_\ell, \bar{\lambda}_\ell)$$

where $\bar{s}_\ell$ and $\bar{\lambda}_\ell$ are hyper-parameters which define the distribution of scales over a population of people.

Unfortunately, it isn't obvious how to set the parameters s_ℓ, $\bar{s}_\ell$ and $\bar{\lambda}_\ell$. To better understand these parameters the unknown scale parameter λ_ℓ can be integrated out. This gives the posterior distribution over $\ell(t)$

$$
\begin{aligned}
P(\ell(t)) &= \int_0^\infty P(\ell(t)|\lambda_\ell)P(\lambda_\ell)d\lambda_\ell \\
&= \mathcal{B}'(\ell(t); s_\ell, \bar{s}_\ell, \bar{\lambda}_\ell)
\end{aligned}
\tag{2.8}
$$

where

$$\mathcal{B}'(x; a, b, \lambda) = \frac{\lambda^{-a}}{\beta(a, b)} \frac{x^{a-1}}{(1 + \frac{x}{\lambda})^{a+b}}$$

is the density function of a scaled beta prime distribution with shape parameters a and b and scale parameter λ. (Note that the fixed parameters s_ℓ, $\bar{s}_\ell$ and $\bar{\lambda}_\ell$ are implicitly conditioned on but are not included in equations for notational clarity.) Then ℓ has mean

$$E[\ell(t)] = \lambda_s \frac{s_\ell}{\bar{s}_\ell - 1}$$

and variance

$$E\left[(\ell(t) - E[\ell(t)])^2\right] = \lambda_s^2 \frac{s_\ell(s_\ell + \bar{s}_\ell - 1)}{(\bar{s}_\ell - 1)^2(\bar{s}_\ell - 2)}$$

The derivation of these equations can be found in the Appendix, Sections B.3 and B.4.

For a desired mean μ and variance σ^2 of $\ell(t)$ these equations can be solved to yield

$$\bar{s}_\ell = \frac{2s_\ell \frac{\sigma^2}{\mu^2} + s_\ell - 1}{s_\ell \frac{\sigma^2}{\mu^2} - 1}$$

$$\bar{\lambda}_\ell = \mu \frac{\bar{s}_\ell - 1}{s_\ell}$$

Segment	Shape $CV[\ell(t)|\lambda_\ell]$	Mean μ	Standard Deviation σ
Thigh	0.025	0.4 m	0.1 m
Shank	0.025	0.5 m	0.2 m
Hips (Width)	0.03	0.3 m	0.2 m
Foot	0.01	0.2 m	0.1 m

Table 2.6: Segment Length Parameters. μ and σ are the mean and standard deviation over an entire population which determine the hyper-parameters $\bar{s}_\ell$ and $\bar{\lambda}_\ell$. $CV[\ell(t)|\lambda_\ell]$ is the coefficient of variation of an individual segment which determines the shape parameter s_ℓ.

where $s_\ell > \frac{\mu^2}{\sigma^2}$ remains a free parameter which tunes the variance of the length samples for a particular segment. While s_ℓ can be set directly, it is more intuitive to set the coefficient of variation $CV[\ell(t)|s_\ell, \lambda_\ell]$ instead and set

$$s_\ell = \frac{1}{CV[\ell(t)|s_\ell, \lambda_\ell]^2}$$

appropriately.

The parameters used for individual segments can be found in Table 2.6. Graphs of Equation (2.7) for different values of $CV[\ell(t)|s_\ell, \lambda_\ell]$ and a fixed λ_ℓ can be seen in Figure 2.7(b). Graphs of Equation (2.8) for different values of $CV[\ell(t)|s_\ell, \lambda_\ell]$ and a fixed value of μ and σ^2 can be seen in Figure 2.7(a).

(a)   (b)
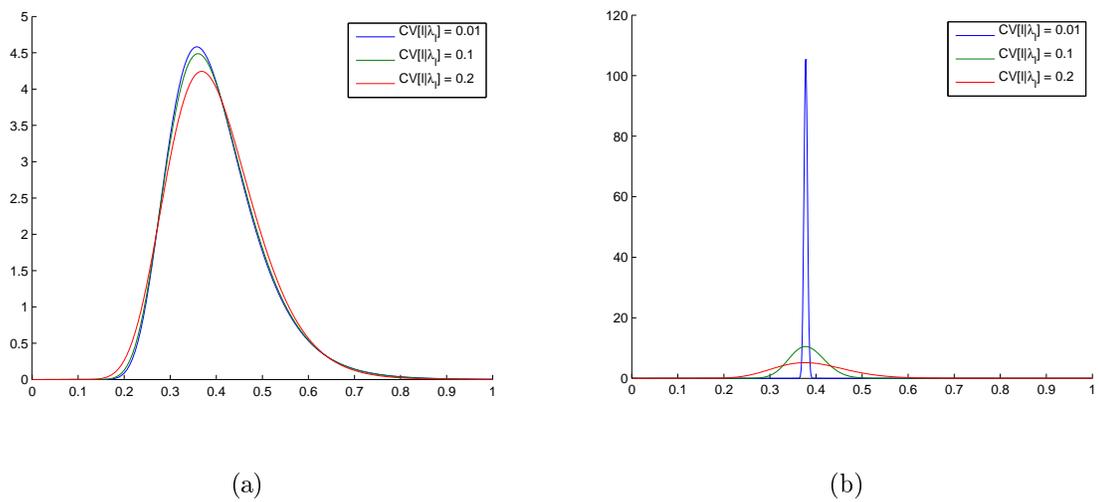
Figure 2.7: Segment Length Distributions. (a) is a plot of the beta-prime distribution $P(\ell(t))$ for different values of $CV[\ell(t)|\lambda_\ell]$ with $\mu = 0.4$ and $\sigma = 0.1$ (b) is a plot of the gamma distribution $P(\ell(t)|\lambda)$ for different values of $CV[\ell(t)|\lambda_\ell]$ with a fixed value of λ_ℓ. Figure (a) is the distribution of length of a segments over a population. Figure (b) is the distribution of lengths of a single segment of an individual

Chapter 3

Model Inference with Sequential Monte Carlo

The problem of people tracking is formulated here as a Bayesian inference problem where tracking is equivalent to infering the distribution of some hidden variables given incomplete or noisy observations. Let x_t represent the hidden state at time t. This includes the kinematic and dynamic states at time t, as well the anthropometric segment lengths at time t. Let y_t be the observations at time t. The observations could be almost anything and their particular form is described in Chapter 4 in the context of the experiments run. Finally, there are some hidden global parameters θ which don't change over time but affect the evolution of the hidden states x_t. θ includes the segment scale parameters λ_ℓ and the mean spring constants and force values which are used to generate the forces at each time.

The model is specified by the state evolution distribution $P(x_t|x_{0:t-1}, \theta)$, the observation likelihood $P(y_t|x_t)$ and the prior distribution over the global parameters $P(\theta)$. It is assumed that the hidden state evolution is a first-order Markov process given θ such that $P(x_t|x_{0:t-1}, \theta) = P(x_t|x_{t-1}, \theta)$, and that the observations are conditionally independent given the hidden states. This model can be represented by the graphical model in Figure
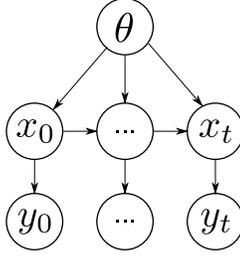
Figure 3.1: Bayesian Tracking as a Graphical Model

3.1.

Unfortunately, the posterior distribution $P(x_{0:t}|y_{0:t})$ is generally impossible to specify analytically and an approximation must be used. The following description of sequential Monte Carlo sampling follows the excellent review by Doucet et al. [7].

The distribution $P(x_{0:t}|y_{0:t})$ is approximated with a set of weighted importance samples $\{(x_{0:t}^{(i)}, w_t^{(i)}) : i = 1, \ldots, N\}$ where $x_{0:t}^{(i)}$ are samples drawn from some proposal distribution $\pi(x_{0:t}|y_{0:t})$ and

$$w_t^{(i)} = \frac{P(y_{0:t}|x_{0:t}^{(i)})P(x_{0:t}^{(i)})}{\pi(x_{0:t}^{(i)}|y_{0:t})} \tag{3.1}$$

are the importance weights. These weighted samples can be used to approximate the expectation of a sufficiently smooth function of state $f(x_{0:t})$ as follows

$$
\begin{aligned}
E_{x_{0:t}}[f(x_{0:t})|y_{0:t}] &= \int f(x_{0:t})P(x_{0:t}|y_{0:t})dx_{0:t} \\
&\approx \sum_{i=1}^{N} f(x_{0:t}^{(i)})\hat{w}_t^{(i)}
\end{aligned}
$$

where

$$\hat{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^{N} w_t^{(i)}}$$

is the normalized weight of the ith sample.

3.1 Sequential Importance Sampling

Sequential importance sampling exploits the sequential structure of the model to allow a set of samples from time t to be efficiently updated to a set of samples at time $t + 1$.

To do this, a proposal distribution is chosen which can be factored as

$$\pi(x_{0:t+1}^{(i)}|y_{0:t+1}) = \pi(x_0^{(i)}|y_0) \prod_{j=1}^{t+1} \pi(x_j^{(i)}|x_{0:j-1}^{(i)}, y_{0:j})$$

Next, notice that there is a recursive form to

$$P(y_{0:t+1}|x_{0:t+1})P(x_{0:t+1}) = P(y_{t+1}|x_{t+1})P(y_{0:t}|x_{0:t})P(x_{t+1}|x_{0:t})P(x_{0:t})$$

Combining these two equations together with Equation (3.1) gives a recursive formula for the weights

$$
\begin{aligned}
w_{t+1}^{(i)} &= \frac{P(y_{0:t+1}|x_{0:t+1}^{(i)})P(x_{0:t+1}^{(i)})}{\pi(x_{0:t+1}^{(i)}|y_{0:t+1})} \\
&= \frac{P(y_{t+1}|x_{t+1}^{(i)})P(x_{t+1}^{(i)}|x_{0:t}^{(i)})P(y_{0:t}|x_{0:t}^{(i)})P(x_{0:t}^{(i)})}{\pi(x_0^{(i)}|y_0) \prod_{j=1}^{t+1} \pi(x_j^{(i)}|x_{0:j-1}^{(i)}, y_{0:j})} \\
&= \frac{P(y_{t+1}|x_{t+1}^{(i)})P(x_{t+1}^{(i)}|x_{0:t}^{(i)})}{\pi(x_{t+1}^{(i)}|x_{0:t}^{(i)}, y_{0:t+1})} \frac{P(y_{0:t}|x_{0:t}^{(i)})P(x_{0:t}^{(i)})}{\pi(x_0^{(i)}|y_0) \prod_{j=1}^{t} \pi(x_j^{(i)}|x_{0:j-1}^{(i)}, y_{0:j})} \\
&= \frac{P(y_{t+1}|x_{t+1}^{(i)})P(x_{t+1}^{(i)}|x_{0:t}^{(i)})}{\pi(x_{t+1}^{(i)}|x_{0:t}^{(i)}, y_{0:t+1})} w_t^{(i)} \qquad\qquad (3.2)
\end{aligned}
$$

This gives rise to the basic sequential importance sampling (SIS) method.

Unfortunately, samples based on basic SIS are known to become degenerate as t increases. To avoid this, occasional resampling is performed. Intuitively, resampling discards unlikely samples and keeps extra copies of more likely samples. To perform simple, random resampling given a set of samples $\{(x_{0:t}^{(i)}, w_t^{(i)}) : i = 1, \ldots, N\}$, a new set of samples $\{(x_{0:t}'^{(j)}, w_{0:t}'^{(j)}) : j = 1, \ldots, M\}$ can be created by choosing $x_{0:t}'^{(j)} = x_{0:t}^{(i)}$ with probability $\hat{w}_{0:t}^{(i)}$ and setting $w_{0:t}'^{(j)} = \frac{1}{M}$ for $j = 1, \ldots, M$. A slightly better approach to resampling is known as residual resampling. In residual resampling $k_i = \left\lfloor M\hat{w}_{0:t}^{(i)} \right\rfloor$ copies of $x_{0:t}^{(i)}$ are automatically kept for $i = 1, \ldots, N$ where $\lfloor \cdot \rfloor$ is the floor operator. The remaining $M - \sum_{i=1}^{N} k_i$ samples are selected according to simple, random resampling with weights proportional to $w^{(i)} = \hat{w}_{0:t}^{(i)} - \frac{k_i}{M}$.

While resampling can be done at any time, if it is done too frequently, good samples may be thrown out prematurely. Alternately, if it is not done frequently enough, the

---

Algorithm 1 Sequential Importance Resampling

For each time t given a set of samples $\{x_{0:t}^{(i)}, w_t^{(i)}\}$:

1. Compute the effective sample size

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N}(\hat{w}_t^{(i)})^2}$$

and, if $\hat{N}_{eff} < N_{thres}$, resample $\{x_{0:t}^{(i)}, w_t^{(i)}\}$ using either simple or residual resampling.

2. For $i = 1, \ldots, N$:

    (a) Sample $x_{t+1}^{(i)} \sim \pi(x_{t+1} | x_{0:t}^{(i)}, y_{0:t+1})$

    (b) Update the importance weights

$$w_{t+1}^{(i)} = \frac{P(y_{t+1}|x_{t+1}^{(i)})P(x_{t+1}^{(i)}|x_{0:t}^{(i)})}{\pi(x_{t+1}^{(i)}|x_{0:t}^{(i)}, y_{0:t+1})} w_t^{(i)}$$

3. For $i = 1, \ldots, N$ compute the normalized weights

$$\hat{w}_{t+1}^{(i)} = \frac{w_{t+1}^{(i)}}{\sum_{j=1}^{N} w_{t+1}^{(i)}}$$

---

sample set will degenerate. A good heuristic is to resample when the estimated effective sample size

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N}(\hat{w}_t^{(i)})^2}$$

is less than some threshold $\hat{N}_{thres}$. This method, known as sequential importance resampling (SIR), is summarized in Algorithm 1.

3.2   Proposal Distribution for the Dynamic Model

To use either SIS or SIR, an appropriate proposal distribution $\pi(x_{t+1}|x_{0:t}, y_{0:t+1})$ must be chosen. For the prior model described in Chapter 2, this choice is complicated by two factors. The first is that it is not generally possible to evaluate the transition density $P(x_{t+1}|x_t, \theta)$. Because the stochastic model of dynamics adds noise to the forces rather than the integration, it would be necessary to invert the dynamics and precisely determine the forces applied in order to determine the probability of a transition. In order to avoid this, the proposal distribution must be the prediction density

$$\pi(x_{t+1}|x_{0:t}, y_{0:t+1}) = P(x_{t+1}|x_{0:t})$$

so that Equation (3.2) becomes

$$w_{t+1}^{(i)} = P(y_{t+1}|x_{t+1}^{(i)})w_t^{(i)}$$

due to cancellation.

The second complicating factor is the global variables θ which induce a dependence such that the hidden states are no longer a first-order Markov process. This means that entire trajectory $x_{0:t}$ may need to be kept and evaluated for each sample in order to draw from the prediction distribution $P(x_{t+1}|x_{0:t})$. One solution would be to perform importance sampling on the posterior $P(\theta, x_{0:t}|y_{0:t})$. This would allow the prediction density to be

$$P(x_{t+1}|x_{0:t}, \theta) = P(x_{t+1}|x_t, \theta)$$

but would suffer from poor sample density in θ, as sample depletion would result in samples with only a single value of θ. In the case of the dynamic prior, another solution is possible.

Let the hidden state at time t be $x_t = (\hat{x}_t, \hat{\theta}_t)$ where $\hat{\theta}_t$ are the variables of x_t which directly depend on θ and $\hat{x}_t$ are the remaining variables. In particular, $\hat{\theta}_t$ are the segment lengths $\ell(t)$ and force parameters at time t and $\hat{x}_t$ is the dynamic state and kinematic

---

Algorithm 2 Sampling from the kinematic and dynamic transition density.

Given segment lengths and force function parameters at the current time $\hat{\theta}_t$ and kinematic and dynamic state at the previous time $\hat{x}_{t-1}$ $\hat{x}_t$ can be drawn as follows

1. Use $\hat{\theta}_t$ to determine the forces to apply to the system. For the monopode this is simply using $f(t)$ directly. For the anthropomorphic walker this consists of evaluating Equation (2.5) using the spring constant $\kappa(t)$.

2. Integrate the equations of motion using the dynamic state given in $\hat{x}_{t-1}$ as the starting condition and $f(t)$ as the applied forces.

3. Sample from the kinematic joint angle model described in Section 2.2 using the new dynamic state to get a new kinematic state.

---

pose. Then the prediction density can be written as

$$
\begin{aligned}
P(x_t|x_{0:t-1}) &= P(\hat{x}_t, \hat{\theta}_t|x_{0:t-1}) \\
&= P(\hat{\theta}_t|x_{0:t-1})P(\hat{x}_t|x_{0:t-1}, \hat{\theta}_t)
\end{aligned}
$$

using the chain rule of probability. If efficient methods could be found to sample from $P(\hat{\theta}_t|x_{0:t-1})$ and $P(\hat{x}_t|x_{0:t-1}, \hat{\theta}_t)$ they could be used to sample from the prediction distribution.

The variables which constitute $\hat{\theta}_t$ in Chapter 2 are the segment lengths ℓ and noisy force value $f(t)$ for the monopode and the noisy sprint constant $\kappa(t)$. For both models, $\hat{x}_t$ consists of the kinematic and dynamic states. Because of this,

$$
P(\hat{x}_t|x_{0:t-1}, \hat{\theta}_t) = P(\hat{x}_t|\hat{x}_{t-1}, \hat{\theta}_t)
$$

which can be sampled from in a straight forward manner. An overview of how sampling from this is done is given in Algorithm 2.

In order to sample from $P(\hat{\theta}_t|x_{0:t-1})$ assume that it can be written as

$$
P(\hat{\theta}_t|x_{0:t-1}) = \mathcal{P}(\hat{\theta}_t; \xi_t)
$$

Algorithm 3 Sampling $\hat{\theta}_t$

Given the previous state x_{t-1} and the previous summary statistics ξ_{t-1}

1. Compute $\xi_t = f(\xi_{t-1}, x_{t-1})$ recursively.

2. Sample $\hat{\theta}_t$ from the density $\mathcal{P}(\hat{\theta}_t; \xi_t)$. This includes drawing samples of $\ell(t)$ for each segment length for both models. For the anthropomorphic walker it includes drawing the spring stiffness $\kappa(t)$ and for the monopode it includes drawing $f(t)$.

---

where $\mathcal{P}$ is a density function with parameter

$$\xi_t = f(\xi_{t-1}, x_{t-1})$$

which can be recursively updated and has limited dimension. If $\mathcal{P}$ can be sampled from efficiently then $\hat{\theta}_t$ can be sampled efficiently using the method described in Algorithm 3.

In the case of a segment length parameter $\ell(t)$

$$\mathcal{P}(\ell(t); \xi_t) = \mathcal{B}'(\ell(t); s_\ell, S_t, \Lambda_t)$$

where s_ℓ is the shape parameter and S_t and Λ_t are the summary statistics which compose ξ_t. They can be updated by

$$S_t = S_{t-1} + s_\ell$$

and

$$\Lambda_t = \Lambda_{t-1} + \ell(t-1)$$

and are initialized to $S_0 = \bar{s}_\ell$ and $\Lambda_0 = \bar{\lambda}_\ell$ which are the population parameters specified in Table 2.6 on page 34. The derivation of this can be found in the Appendix, Section B.3.

The force parameters are similar, but $\mathcal{P}$ is Gaussian. Using the spring constant $\kappa(t)$ of the anthropomorphic walker as an example

$$\mathcal{P}(\kappa(t); \xi_t) = \mathcal{N}(\kappa(t); M_t, s_t^2 + \sigma_\kappa^2)$$

Algorithm 4 Sampling from the prediction density $P(x_t|x_{0:t-1})$.

Given the previous hidden state $x_{t-1} = (\hat{x}_{t-1}, \hat{\theta}_{t-1})$ and the statistics ξ_{t-1} which summarize the trajectory $x_{0:t-2}$ then $x_t = (\hat{x}_t, \hat{\theta}_t)$ can be drawn from $P(x_t|x_{0:t-1})$ as follows

1. Using x_{t-1} update ξ_{t-1} to ξ_t and draw $\hat{\theta}_t$ according to Algorithm 3.

2. Using $\hat{x}_{t-1}$ and $\hat{\theta}_t$ draw $\hat{x}_t$ according to Algorithm 2.

---

where σ_κ^2 is the variance of the spring constant noise and M_t and s_t^2 are the summary statistics which compose ξ_t. These parameters are the parameters of the distribution

$$P(\bar{\kappa}(s(t))|x_{0:t-1}) = \mathcal{N}(\bar{\kappa}(s(t)); M_t, s_t^2)$$

and can be updated by

$$M_t = \begin{cases} \alpha(s_{t-1}^{-2} + \sigma_\kappa^{-2})^{-1}(s_{t-1}^{-2}M_{t-1} + \sigma_\kappa^{-2}\kappa(t-1)) + (1-\alpha)\bar{\kappa}_0 & \text{if } s(t) \neq s(t-1) \\ (s_{t-1}^{-2} + \sigma_\kappa^{-2})^{-1}(s_{t-1}^{-2}M_{t-1} + \sigma_\kappa^{-2}\kappa(t-1)) & \text{otherwise} \end{cases}$$

$$s_t^2 = \begin{cases} \alpha^2(s_{t-1}^{-2} + \sigma_\kappa^{-2})^{-1} + \sigma_{\bar{\kappa}}^2 & \text{if } s(t) \neq s(t-1) \\ (s_{t-1}^{-2} + \sigma_\kappa^{-2})^{-1} & \text{otherwise} \end{cases}$$

where α is the smoothing parameter, $\bar{\kappa}_0$ is the global mean spring constant, $\sigma_{\bar{\kappa}}^2$ is the process variance on spring constant means and $s(t)$ is the stride index at time t. These values are initialized to $M_0 = \bar{\kappa}_0$ and $s_0^2 = \sigma_{\bar{\kappa}}^2$ as specified in Table 2.2 on page 19. The motivation of these equations are discussed in the Appendix, Section B.2.

Algorithms 3 and 2 can then be combined to sample from the prediction density $P(x_t|x_{0:t-1})$ as described in algorithm 4.

Chapter 4

Experimental Results

Inference based on the prior discussed in Chapter 2 is done as described in Chapter 3 using sequential importance resampling with residual resampling. The observations are hand-labeled image points from two different monocular video sequences. The image points correspond to known points in the kinematic geometry. The intrinsic and extrinsic camera parameters are known along with the location and orientation of the ground plane. An example frame can be seen in Figure 4.1 with the ground plane overlayed using the camera calibration.

The observational likelihood, which is about to be presented, was not included in Chapter 2 to make clear that the model is not dependent on the choice of likelihood. Many other likelihoods could be used with the same underlying model of motion. This particular likelihood was chosen for its simplicity and clarity in evaluating the underlying model of motion.

Now, given the camera projection operator $\mathbb{P}$, the generative model of the 2D image location of the ith kinematic point is

$$\hat{o}_t^i = \mathbb{P}(k^i(x_t)) + \epsilon^i$$

where $k^i(x_t)$ is the world coordinate location of the ith kinematic point and ϵ^i is IID 2D
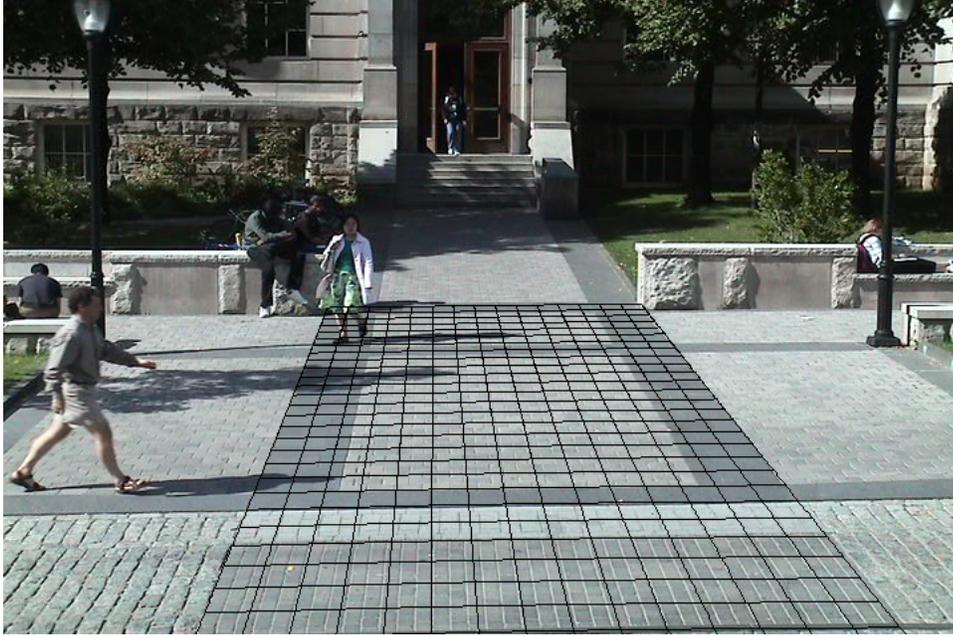
Figure 4.1: The ground plane is overlayed on a single frame using the camera calibration.

isotropic Gaussian noise with a variance of σ_i^2 and

$$p_m(\epsilon^i; \sigma_i^2) = (2\pi\sigma_i^2)^{-1} e^{-\frac{1}{2\sigma_i^2}\left\|\epsilon^i\right\|^2}$$

is the density function of the measurement error. Therefore, the likelihood of a set of M 2D image observation $y_t = (o_t^1, \ldots, o_t^M)$ is

$$P(y_t|x_t) = \prod_{i=1}^{M} p_m(o_t^i - \mathbb{P}(k^i(x_t)); \sigma_i^2)$$

where x_t is the hidden state. All sequences are hand initialized by randomly sampling a small region around an estimate of the starting pose. The parameters used are those specified in the tables in Chapter 2. The variances used for the different points can be found in Table 4.1.

The result at time t of the approximate inference algorithm is a set of N weighted samples $\{(x_{0:t}^{(i)}, w_t^{(i)}) : i = 1, \ldots, N\}$ which approximate the distribution $P(x_{0:t}|y_{0:t})$. Displaying all the samples is impractical so three summary statistics are displayed. The first

Observation Point	σ_i
Hip	7 pixels
Knee	5 pixels
Foot	5 pixels

Table 4.1: Observation Variances

is the mean state at time t which is

$$E[x_t|y_{0:t}] \approx \sum_{i=1}^{N} \hat{w}_t^{(i)} x_t^{(i)}$$

where scalar multiplication and addition is done in the space of kinematic joint angles and anthropometric lengths. This produces a "mean pose" which can be illustrative when the distribution is fairly peaked around a single pose.

When the distribution is more diffuse or when there are multiple peaks the mean pose becomes relatively meaningless. In such a case, a display which represents the kinematic points can be useful. Specifically,

$$\begin{aligned} \mu^i &= E[\mathbb{P}(k^i(x_t))|y_{1:t}] \\ &\approx \sum_{j=1}^{N} \hat{w}_t^{(i)} \mathbb{P}(k^i(x_t^{(j)})) \end{aligned}$$

is the mean of the projection of the ith kinematic point and

$$\begin{aligned} \Sigma^i &= V[\mathbb{P}(k^i(x_t))|y_{1:t}] \\ &\approx \sum_{j=1}^{N} \hat{w}_t^{(i)} (\mu^i - \mathbb{P}(k^i(x_t^{(j)})))(\mu^i - \mathbb{P}(k^i(x_t^{(j)})))^T \end{aligned}$$

is the covariance of the projection of the ith kinematic point. An ellipse can be plotted around μ^i with the axes defined by Σ^i for each kinematic point i to summarize the distribution in cases where the mean pose is not illustrative.

Finally, while these representations are indicative of the state at time t, given the observations up to time t, it is interesting to ask what the previous states look like,

given an entire sequence which ends at time T. One way to answer this is to look at the maximum a posteriori (MAP) trajectory.[1] A crude approximation of the MAP trajectory is the trajectory $x_{0:T}^{(i)}$ such that $w_T^{(i)} = \max_j w_T^{(j)}$. This display has some noteworthy advantages. First it represents the motion of a single particle over time which ensures that a "valid" state is always being displayed. In contrast the mean pose can result in poses which are not possible for an individual sample. The MAP trajectory also allows a "filling-in" effect where uncertainty in the past is reduced by observations in the future. This particular feature is most interesting in the cases of missing data.

In the displayed images, the hand-labeled data points are displayed as circles with a radius proportional to their standard deviation. The red circle is the hip point and the green and blue circles are the knee and foot points on the left and right legs respectively. In the mean and MAP trajectory images, the hip segments are displayed in red, the thigh segments are displayed in green, the shank segments are displayed in blue, and the feet are displayed in red. For the distribution images the hip point ellipses are in red, the knee point ellipses are in green and the feet point ellipses are in blue.

Each experiment was run using both the anthropomorphic walker and the monopode for dynamics. For the anthropomorphic walker $N = 2500$ particles were used with a resampling threshold of $N_{thres} = 500$. Because the monopode is a weaker dynamical model more particles were required and values of $N = 5000$ and $N_{thres} = 2000$ were used. With these parameters the inference algorithm is able to run at real-time speeds (Approximately 22 frames per second for the anthropomorphic walker and 10 frames per second for the monopode) on a modest Intel Pentium M 1.6GHz laptop. Movies of the experimental results can be found online at `http://www.cs.toronto.edu/~mbrubake/ permanent/msthesis`.

---

[1]Note that the MAP trajectory is different than the MAP state. A state of the MAP trajectory can only be computed once all the observations of a sequence have been presented.
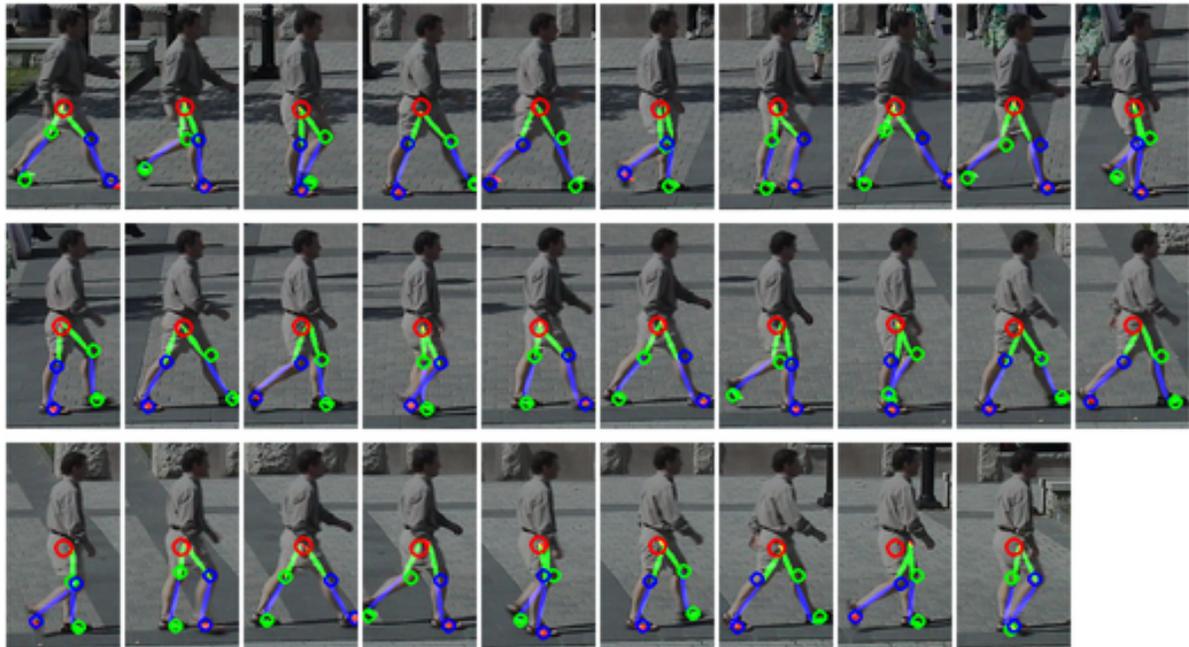
Figure 4.2: Anthropomorphic Walker Experiment #1. The mean pose of every 4th frame is shown.

Experiment #1: Simple Walking

In this experiment the subject walks in a straight line which is approximately parallel to the camera plane. This simply illustrates the most basic functionality of the tracker and the inference algorithm. Both models are able to infer the lower body pose for the entire sequence.

The mean pose for the anthropomorphic walker and monopode are in Figures 4.2 and 4.3 respectively.

Experiment #2: Occlusion

In this experiment the same data is used as in experiment 1. However for 2 seconds (60 frames) in the middle of the sequence occlusion is simulated by removing the points on the knees and feet. Based only on the torso point the distribution must represent the uncertainty in the orientations of the thigh, shank and feet.
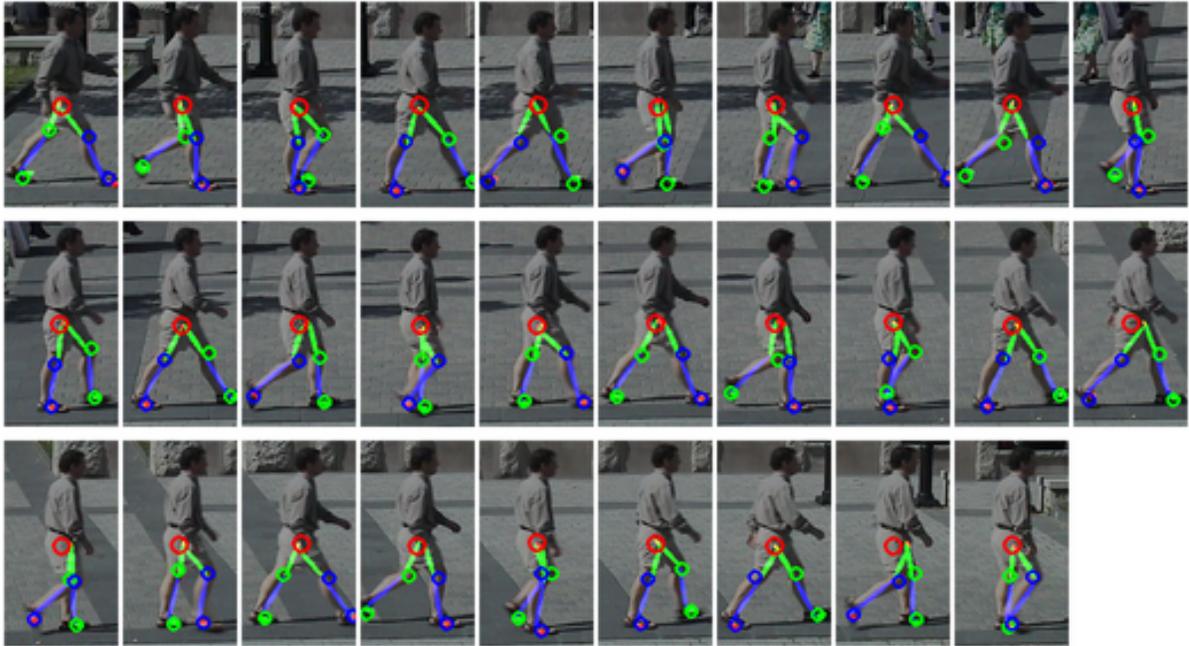
Figure 4.3: Monopode Experiment #1. The mean pose of every 4th frame is shown.
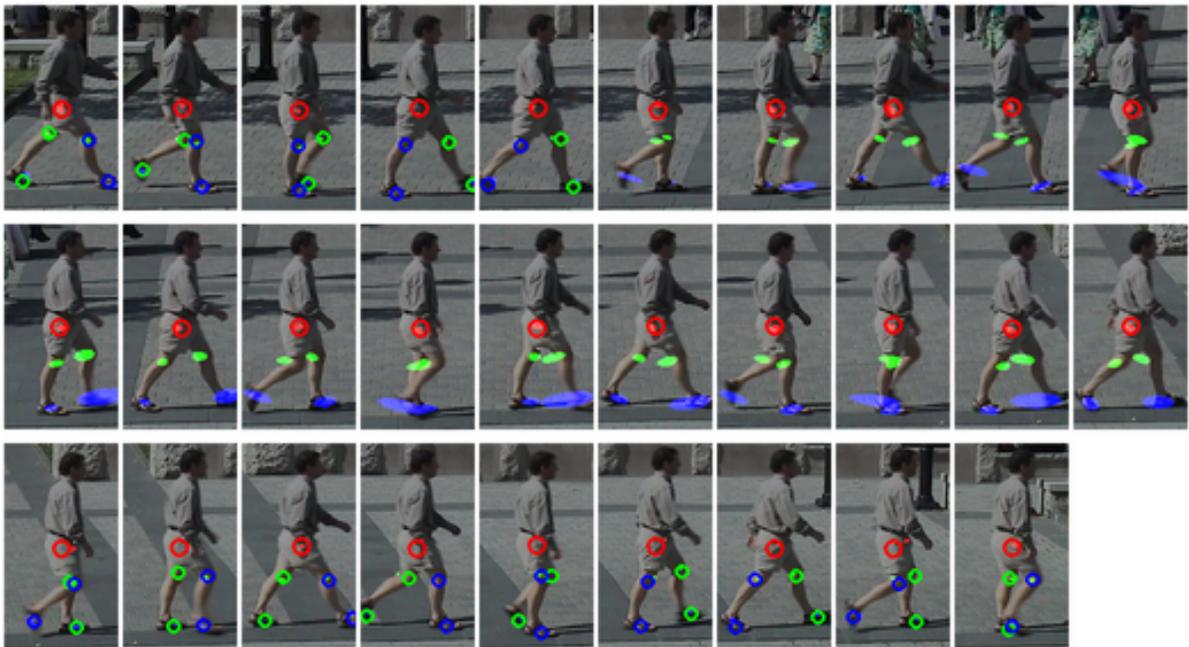
Figure 4.4: Anthropomorphic Walker Experiment #2. The distribution of tracking points for every 4th frame is shown. While the data is present the variances are small and the ellipses may be hard to see.
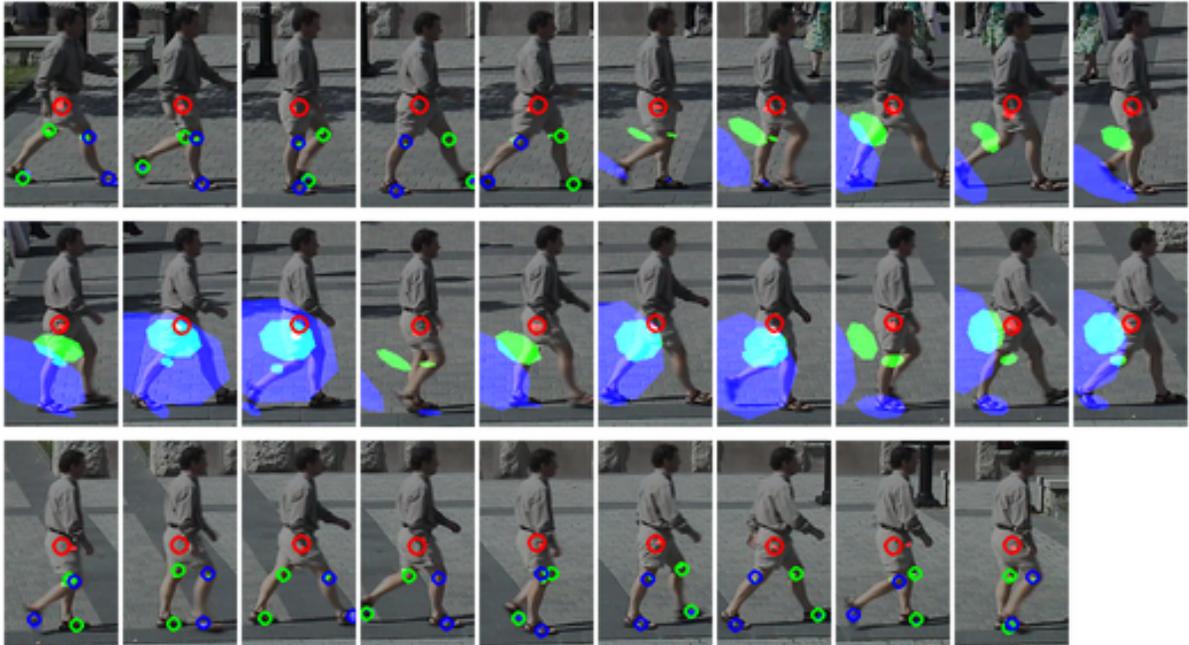
Figure 4.5: Monopode Experiment #2. The distribution of tracking points for every 4th frame is shown. While the data is present the variances are small and the ellipses may be hard to see.
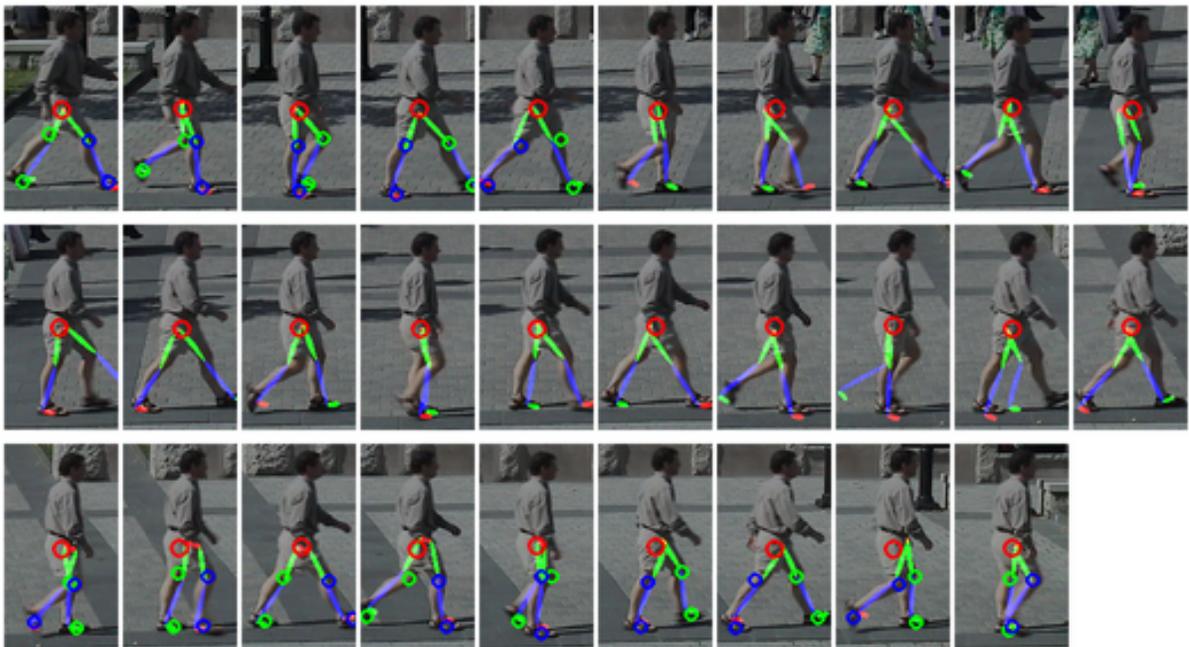

Figure 4.6: Anthropomorphic Walker Experiment #2. The MAP trajectory at every 4th frame is shown.
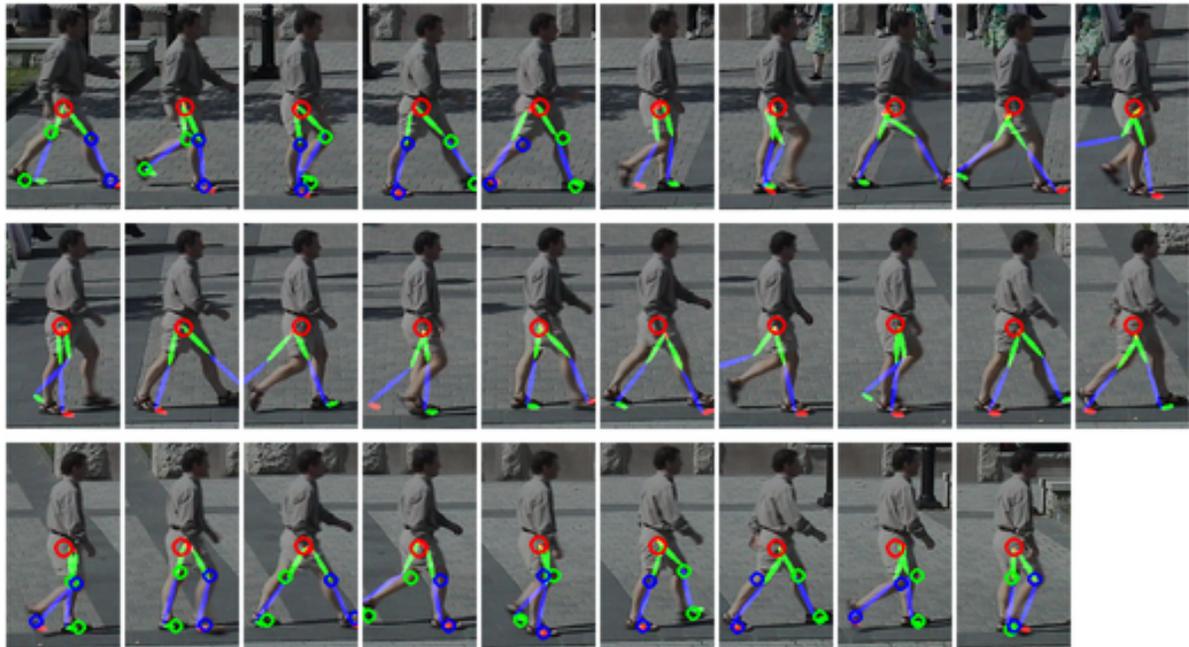
Figure 4.7: Monopode Experiment #2. The MAP trajectory at every 4th frame is shown.

Three very nice properties of the prior model are demonstrated in this experiment. First the uncertainty in the distribution grows naturally as more time is spent without the data. Second, and more interesting, is that the subtle cues in the torso movement are correctly interpretted and the uncertainty in the distribution decreases when the sudden change in velocity due to foot contact is detected. The prior model not only enables the interpretation of this information but also allows an accurate inference about the location of the new stance foot position. Finally, while the uncertainty is high while the data is missing, once the data has returned the model is able to fill-in what most likely happened during the time of missing data which can be seen in the MAP trajectory.

The weakness of the monopode is clearly seen in this experiment as the uncertainty of the swing leg during occlusion is much higher than with the anthropomorphic walker. Despite this weakness, the monopode is still able to reasonably infer the placement of the stance feet during the occlusion.

The distribution of kinematic points for the anthropomorphic walker and the monopode

are in Figures 4.4 and 4.5 respectively. The MAP trajectory for both models are displayed in Figures 4.6 and 4.7.

Experiment #3: Turning

Though the underlying dynamics are not designed to handle 3D features such as turning they still provide meaningful information about human gait during turns. In this experiment a sequence where the subject makes several turns is tracked. The anthropomorphic walker is able to keep track during the entire sequence but the monopode looses track during the last turn and is unable to recover.

When the subject is walking away from the camera, ambiguities in depth cause the posterior distribution over the recent past to broaden out. Because of this many more samples are necessary to represent the uncertainty effectively. As the subject turns back to parallel to the camera the posterior distribution over the recent past collapses to a much sharper distribution. However, if there were an insufficient number of samples to represent the uncertainty it may no longer have any samples near the peaks of the collapsed posterior and subsequently fail as the monopode did. With a smaller sample size similar modes of failure can be seen with the anthropomorphic walker.

The mean pose of the anthropomorphic walker and the monopode can be see in Figures 4.8 and 4.9 respectively.
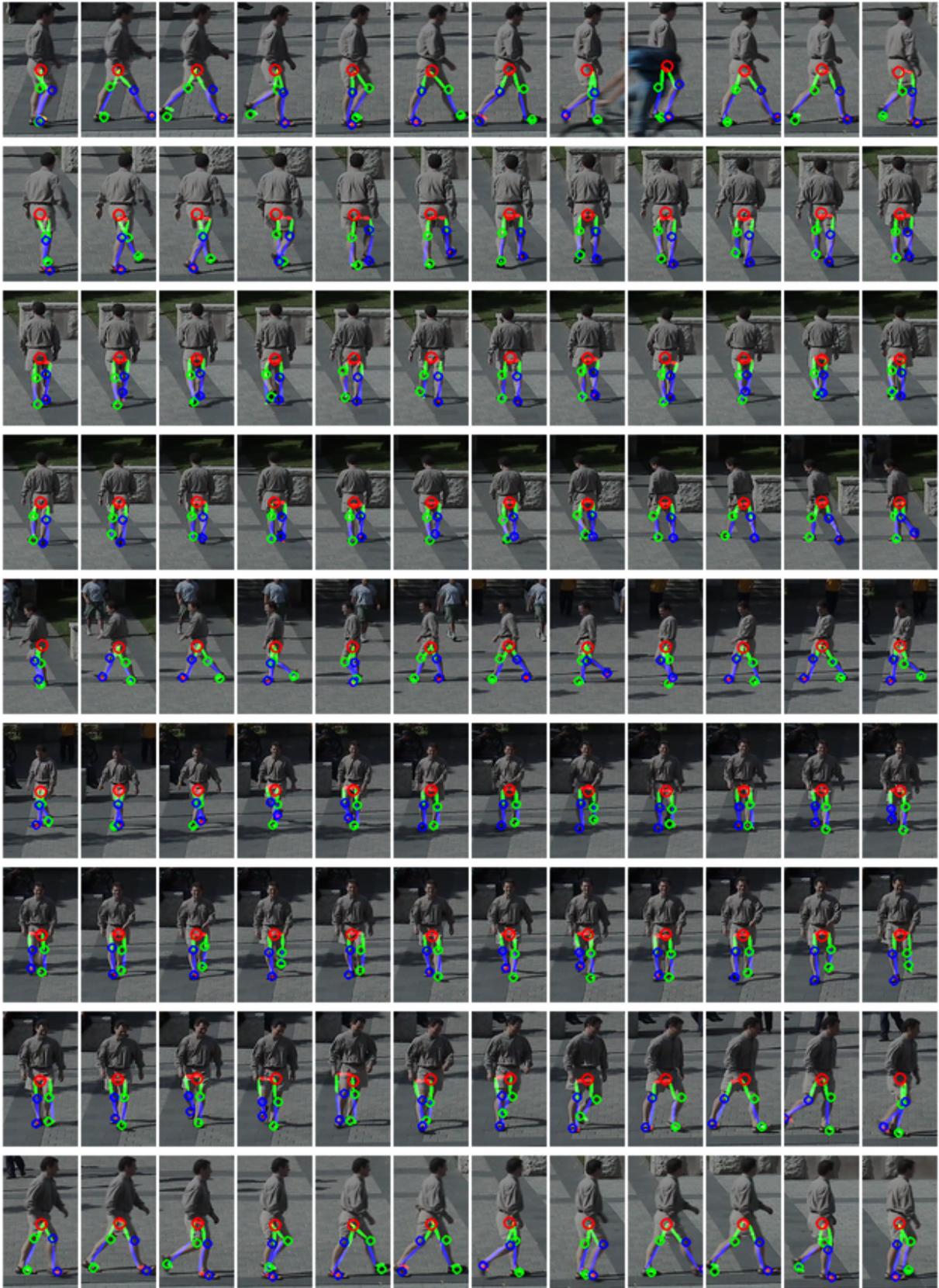
Figure 4.8: Anthropomorphic Walker Experiment #3. The mean pose of every 4th frame is shown.
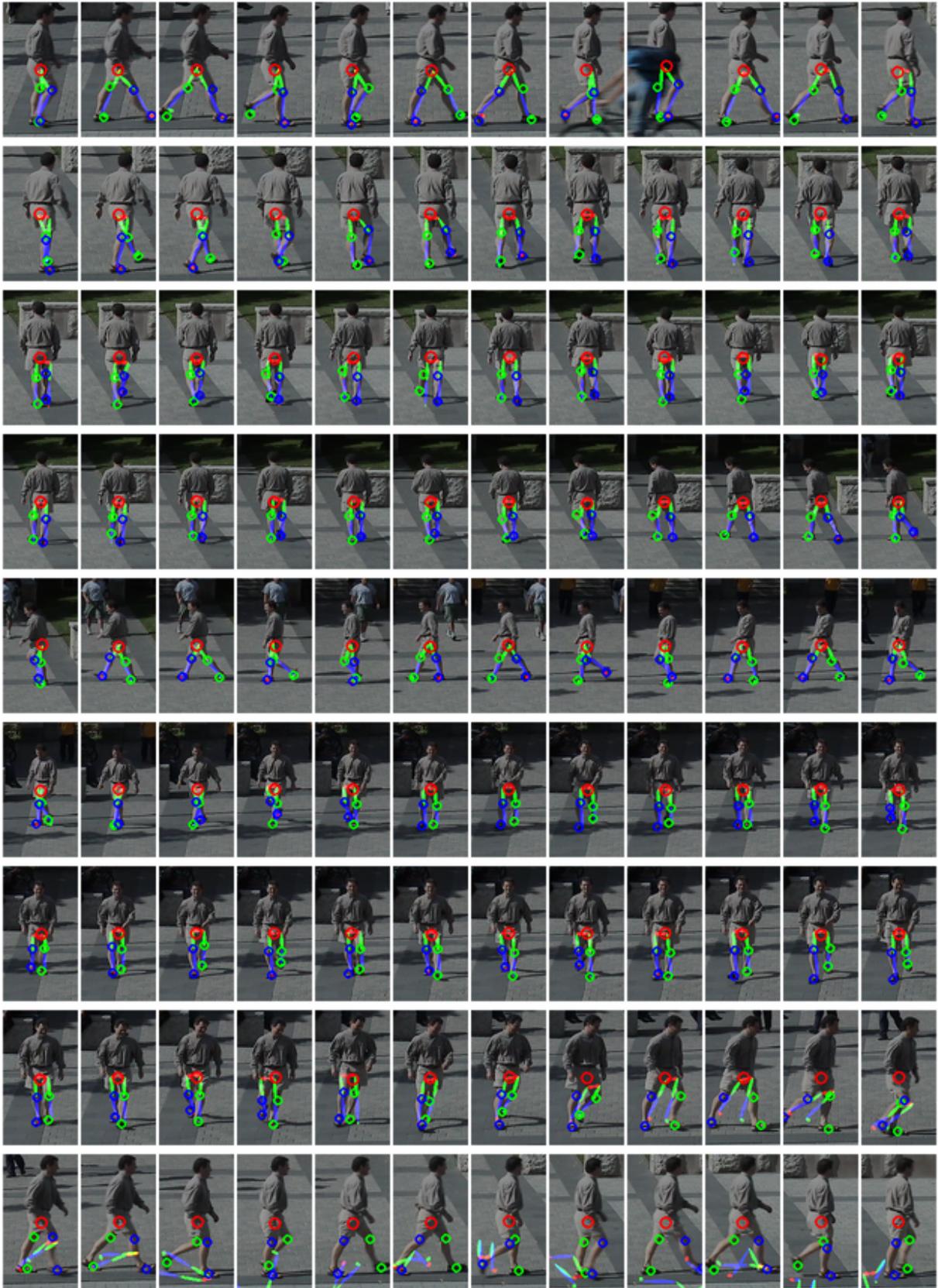
Figure 4.9: Monopode Experiment #3. The mean pose of every 4th frame is shown.

Chapter 5

Conclusions and Future Work

This thesis has presented two dynamical prior models for human motion. The experimental results presented in Chapter 4 indicate that the priors presented are capable of the challenging task of pose inference with limited, monocular data. The successful use of physical models in this context shows that dynamics can be managable if the models are wisely choosen. Conventional wisdom holds that dynamic character models are difficult to control, particularly in a stochastic context. In constrast, this thesis suggests dynamical models can be managed and that the key to their success is the use of abstraction.

Despite the successful results, this is preliminary work which demonstrates the potential of dynamical models and presents many possible directions for future work. Perhaps the most obvious direction is the use of more complex dynamical models to capture other aspects of human locomotion.

Improved Dynamics The abstract models used here, particularly the anthropomorphic walker, are capable of capturing most of the dynamics of basic walking. However they will fail to capture many salient aspects of walking on steep slopes and may have difficulty with abnormal gaits. Further, other motions, such as standing, running and jumping, are completely unmodelled.

The key to successfully modelling other motions is careful selection of dynamic features. For instance, when a person is walking on a slope one salient property is the lean of the upper body. Instead of adding an entire upper body a single representative mass could be added to the anthropomorphic walker. This provides a minimal dynamic addition which captures the orientation of the upper body and its impact on the overall dynamics without inducing the added complexity of arms and a head. Further, by augmenting a known model, it is likely that the stochastic techniques used to control the original model can be extended to the control of the more complex model. McGeer [31] has explored models along these lines and would be a natural starting point for future work.

Another example, running, provides two dynamic challenges. First, the knee plays a pivotal role in the dynamics of running and would need to be modelled. The second challenge is the "free-flight" stage of running when there is no contact with the ground. Modelling this stage, and the transitions to and from it, must be handled carefully. Naive take-off and touch-down models could cause unstable simulations and make the dynamics difficult to control. For instance, elastic ground contacts would create stiff equations of motion and unnecessarily complex dynamics when the foot is in contact with the ground. A better solution would be to somehow model touch-down and take-off impulsively and switch between one model when the foot is on the ground and a different model when the body is in flight.

A different direction to explore is the connection between abstract dynamic models and kinematic models. While one ad-hoc solution was presented in this thesis there is little more than experimental evidence to support its use. One potential solution is trying to fit the abstract model to motion capture data in a more principled fashion. This could enable the learning of a mapping from an abstract dynamic state to a full kinematic pose. Being able to match dynamics to real data could also provide insight into the structure of the control strategies of the abstract models when they are aiming to model real motions.

Popović and Witkin [37] have studied how to express the connection between abstract dynamics and high degree-of-freedom kinematic models and would be a valuable source of ideas for work along these lines.

Image-based Likelihoods The experiments used to evaluate the prior were based on tracking hand labeled data. This allowed for a direct evaluation of the prior model without concern for compounding factors relating to image data. Though the results indicate that the model can well explain simple human walking motions, it doesn't necessarily follow that it will succeed at tracking based on real image data. The nature of observational noise when tracking image data is not well understood and errors in the modelling of this noise can cause problems in trackers. Thus to fully determine the value of this prior in the context of tracking based on image data, a real image likelihood function needs to be used.

Inferring the World A subtle constraint of this prior model is that the world and camera geometry must be known a priori and a planar ground surface is assumed. However these are constraints of simplicity rather than strict requirements. It is possible that, with appropriate modelling, much of the world could be inferred along with the pose. For instance, the ground contacts provide strong hints about the geometry of the world. Pushing the inference even further light sources could be inferred using rough models of human geometry and exploiting the 3D representation of pose.

Improving Inference Inference was performed using basic sequential importance resampling of the usual posterior $P(x_{0:t}|y_{0:t})$. However given that many properties of dynamics are only observable over time other posteriors may be better. For instance, a time-lagged posterior $P(x_{0:t}|y_{0:t+\tau})$ will be more accurate and may allow the use of smaller numbers of particles. Other modifications to basic sequential importance resampling, such as rejection control [25] and local sampling using general Markov Chain Monte

Carlo techniques [24], could result in both more efficient and effective sampling.

Other Applications   The development of this prior model was motivated by the problem of people tracking. However the generative model could be applied to other problem domains. For instance, motion capture data typically has problems with dynamic realism even if it is visually realistic and the use of this prior when computing the kinematic component of motion capture data could help alleviate the problem. Another potential application is in the identification of gait. The dynamic control parameters, such as the mean spring constant in the anthropomorphic walker, provide a compact and potentially useful description of a stride.

Appendix A

Dynamics of Walking

This appendix discusses the methods used for deriving the equations of motion and includes the full equations of motion for the anthropomorphic walker.

A.1   The TMT Method

In order to derive the equations of motion for a kinematic model the TMT method is used. This method is similar to using the Lagrange equations in that it specifies the equations of the motion in terms of a generalized coordinate system. A brief derivation of the general method will be given and then applied to the anthropomorphic walker. See van der Linde and Schwab [51] for a more detailed explanation.

Starting with the principle of virtual work known as d'Alembert's Principle

$$\delta \dot{x} \cdot (F - M\ddot{x}) = \delta \dot{x}^T (F - M\ddot{x}) = 0 \tag{A.1}$$

where $\delta \dot{x}$ are virtual velocities, F are applied forces, M is the mass matrix and $\ddot{x}$ is the acceleration.

For a set of bodies assume that their centers of mass x can be expressed in terms of some set of independent generalized coordinates q such that

$$x = T(q) \tag{A.2}$$

where T is the kinematic transformation function. Differentiating gives

$$\begin{aligned}
\dot{x} &= \frac{\partial T}{\partial q}\dot{q} \\
\delta\dot{x} &= \frac{\partial T}{\partial q}\delta\dot{q}
\end{aligned} \tag{A.3}$$

where $\frac{\partial T}{\partial q}$ is the Jacobian of the kinematic transformation function and $\delta\dot{q}$ are the virtual velocities of the generalized coordinates.

Substituting Equation (A.3) into Equation (A.1) gives

$$\delta\dot{q}^T \left(\frac{\partial T}{\partial q}\right)^T (F - M\ddot{x}) = 0$$

which, because the generalized coordinates are independent (and thus, so are their velocities), gives

$$\left(\frac{\partial T}{\partial q}\right)^T (F - M\ddot{x}) = 0 \tag{A.4}$$

Writing the accelerations of the centers of mass by twice differentiating Equation (A.2) gives

$$\ddot{x} = \frac{\partial T}{\partial q}\ddot{q} + \frac{\partial}{\partial q}\left(\frac{\partial T}{\partial q}\dot{q}\right)\dot{q} \tag{A.5}$$

and substituting Equation (A.5) into Equation (A.4) gives

$$\left(\frac{\partial T}{\partial q}\right)^T \left(F - M\left(\frac{\partial T}{\partial q}\ddot{q} + \frac{\partial}{\partial q}\left(\frac{\partial T}{\partial q}\dot{q}\right)\dot{q}\right)\right) = 0$$

which can be rewritten as

$$\mathbf{M}\ddot{q} = \mathbf{F}$$

where

$$\mathbf{M} = \mathbf{T}^T M \mathbf{T}$$

is the reduced mass matrix,

$$\mathbf{F} = \mathbf{T}^T (F - M\mathbf{g})$$

is the reduced force vector,

$$\mathbf{T} = \frac{\partial T}{\partial q}$$

is the first order kinematic transfer function, and

$$\mathbf{g} = \frac{\partial}{\partial q} \left(\frac{\partial T}{\partial q} \dot{q} \right) \dot{q}$$

is sometimes known as the convective acceleration. Thus, for a given set of independent generalized coordinates q and a kinematic transformation function $T(q)$, deriving the equations of motion simply consists of finding $\mathbf{T}$ and $\mathbf{g}$.

A.2 Impulsive Collisions

A collision is defined by a function D of the generalized coordinates q. A collision is said to occur when $D(q) = 0$. The result of a collision can be modeled in terms of the coefficient of restitution such that

$$\frac{\partial D}{\partial q} \dot{q}^+ = -e \frac{\partial D}{\partial q} \dot{q}^- \tag{A.6}$$

where $\frac{\partial D}{\partial q} \dot{q}$ is the relative velocity with respect to the collision and the $+$ and $-$ superscripts denote pre- and post-collision respectively.

To fit this with the TMT method, contact forces λ are added to Equation (A.4) such that

$$\mathbf{T}^T \left(M\ddot{x} - F + \lambda \frac{\partial D}{\partial q} \right) = 0$$

which, after integrating and taking the limit, gives

$$\lim_{t^- \to t^+} \int_{t^-}^{t^+} \mathbf{T}^T \left(M\ddot{x} - F + \lambda \frac{\partial D}{\partial q} \right) dt = 0 \tag{A.7}$$

Notice that any non-impulsive forces applied to the system will go to zero in this equation. Thus, define

$$S = \lim_{t^- \to t^+} \int_{t^-}^{t^+} F dt$$

to be the net applied impulsive forces and

$$\frac{\partial D}{\partial q}\rho = \lim_{t^- \to t^+} \int_{t^-}^{t^+} \frac{\partial D}{\partial q}\lambda dt$$

to be the net impulsive contact forces. Putting this into Equation (A.7) gives

$$\lim_{t^- \to t^+} \int_{t^-}^{t^+} \mathbf{T}^T M \ddot{x} dt = \mathbf{T}^{+T}\left(S - \frac{\partial D}{\partial q}\rho\right)$$

Looking at the left side of that equation

$$\lim_{t^- \to t^+} \int_{t^-}^{t^+} \mathbf{T}^T M \ddot{x} dt = \mathbf{T}^{+T} M(\dot{x}^+ - \dot{x}^-)$$
$$= \mathbf{T}^{+T} M(\mathbf{T}^+ \dot{q}^+ - \mathbf{T}^- \dot{q}^-)$$

which, with rearranging, gives

$$\mathbf{T}^{+T} M \mathbf{T}^+ \dot{q}^+ + \mathbf{T}^{+T}\frac{\partial D}{\partial q}\rho = \mathbf{T}^{+T}S + \mathbf{T}^{+T} M \mathbf{T}^- \dot{q}^- \qquad (A.8)$$

Finally, combining Equations A.6 and A.8 forms a linear system of equations in the unknowns ρ and $\dot{q}^+$

$$\begin{bmatrix} \mathbf{T}^{+T}M\mathbf{T}^+ & \mathbf{T}^{+T}\frac{\partial D}{\partial q} \\ \frac{\partial D}{\partial q} & 0 \end{bmatrix}\begin{bmatrix} \dot{q}^+ \\ \rho \end{bmatrix} = \begin{bmatrix} \mathbf{T}^{+T}(S + M\mathbf{T}^-\dot{q}^-) \\ -e\frac{\partial D}{\partial q}\dot{q}^- \end{bmatrix} \qquad (A.9)$$

Note that the differentiation between $\mathbf{T}^+$ and $\mathbf{T}^-$ is significant as the kinematic transfer function can (and will in the case of the anthropomorphic walker) change at the time of collision. Finally, notice that if the collision is inelastic then Equation (A.9) reduces to

$$\mathbf{T}^{+T}M\mathbf{T}^+\dot{q}^+ = \mathbf{T}^{+T}(S + M\mathbf{T}^-\dot{q}^-)$$

A.3 The Anthropomorphic Walker

The generalized coordinates of this model are chosen to be the global orientation of the stance and swing legs. These coordinates are written as

$$q = (\phi_1, \phi_2)^T$$

where ϕ_1 is the orientation of the stance leg and ϕ_2 is the orientation of the swing leg. Assuming that the origin is located at the end of the stance leg when it is in a vertical position, the kinematic transformation function can be written as

$$T(q) = \begin{bmatrix} -R\phi_1 - (C_1 - R)\sin\phi_1 \\ R + (C_1 - R)\cos\phi_1 \\ \phi_1 \\ -R\phi_1 - (L - R)\sin\phi_1 + (L - C)\sin\phi_2 \\ R + (L - R)\cos\phi_1 - (L - C)\cos\phi_2 \\ \phi_2 \end{bmatrix}$$

the mass matrix can be written as

$$M = \begin{bmatrix} m_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & m_l & 0 & 0 \\ 0 & 0 & 0 & 0 & m_l & 0 \\ 0 & 0 & 0 & 0 & 0 & I_l \end{bmatrix}$$

and

$$m_1 = m_l + m_t$$

$$C_1 = \frac{(Cm_l + Lm_t)}{m_l + m_t}$$

$$I_1 = I_l + I_t + (C_1 - C)^2 m_l + (L - C_1)^2 m_t$$

where the parameters are illustrated in Figure A.1 on page 64.

There are several ways to define the mass parameters given the model in the figure. The equations given here assume that the torso mass is combined with the stance leg mass during the stance phase. This effectively says that the torso is supported entirely by the stance leg instead of being partly supported by the swing leg.
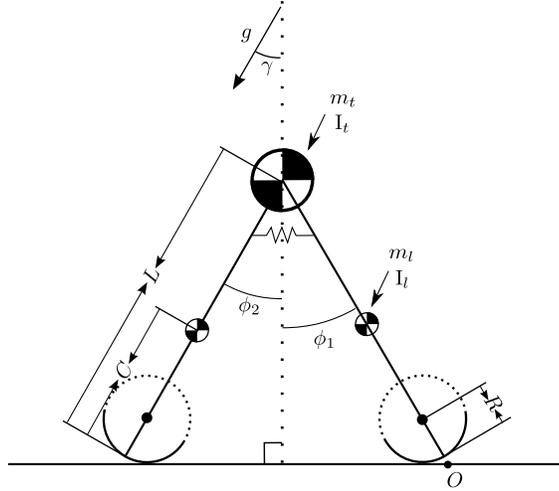
Figure A.1: Parameters of the Anthropometric Walker

Following the TMT method outlined above, the Jacobian of $T(q)$ is

$$
\mathbf{T} = \frac{\partial T}{\partial q}
$$

$$
= \begin{bmatrix}
-R - (C_1 - R)\cos\phi_1 & 0 \\
-(C_1 - R)\sin\phi_1 & 0 \\
1 & 0 \\
-R - (L - R)\cos\phi_1 & (L - C)\cos\phi_2 \\
-(L - R)\sin\phi_1 & (L - C)\sin\phi_2 \\
0 & 1
\end{bmatrix}
$$

and convective accelleration is

$$
\mathbf{g} = \frac{\partial}{\partial q}\left(\frac{\partial T}{\partial q}\dot{q}\right)\dot{q}
$$

$$
= \begin{bmatrix}
\dot{\phi}_1^2(C_1 - R)\sin\phi_1 \\
-\dot{\phi}_1^2(C_1 - R)\cos\phi_1 \\
0 \\
\dot{\phi}_1^2(L - R)\sin\phi_1 - \dot{\phi}_2^2(L - C)\sin\phi_2 \\
-\dot{\phi}_1^2(L - R)\cos\phi_1 + \dot{\phi}_2^2(L - C)\cos\phi_2 \\
0
\end{bmatrix}
$$

When collision happens in the model support is instantaneously transfered between legs. This means that the dynamics of the system must change accordingly and the impact equations need to be aware of this. Specifically, the kinematic transformation is different immediately after impact. That is:

$$
T^+(q) =
\begin{bmatrix}
-R\phi_2 - (L-R)\sin\phi_2 + (L-C)\sin\phi_1 \\
R + (L-R)\cos\phi_2 - (L-C)\cos\phi_1 \\
\phi_1 \\
-R\phi_2 - (C_1-R)\sin\phi_2 \\
R + (C_1-R)\cos\phi_2 \\
\phi_2
\end{bmatrix}
$$

and thus

$$
\begin{aligned}
\mathbf{T}^+ &= \frac{\partial T^+}{\partial q} \\[2mm]
&=
\begin{bmatrix}
(L-C)\cos\phi_1 & -R-(L-R)\cos\phi_2 \\
(L-C)\sin\phi_1 & -(L-R)\sin\phi_2 \\
1 & 1 \\
0 & -R-(C_1-R)\cos\phi_2 \\
0 & -(C_1-R)\sin\phi_2 \\
0 & 1
\end{bmatrix}
\end{aligned}
$$

Appendix B

Distributions and Priors

This appendix briefly reviews some properties of a few key probability distributions which are used in the thesis.

B.1 The Gaussian and Truncated Gaussian Distributions

A random variable x with a Gaussian distribution with mean μ and variance σ^2 has density function

$$\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

If the variable x is restricted such that $x \in [\mu^-, \mu^+]$ then it has the density function

$$\mathcal{N}(x; \mu, \sigma^2, [\mu^-, \mu^+]) = \left(\Phi\left(\frac{\mu^+ - \mu}{\sigma}\right) - \Phi\left(\frac{\mu^- - \mu}{\sigma}\right) \right)^{-1} \left(2\pi\sigma^2\right)^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where

$$\Phi(x) = \int_{-\infty}^{x} \mathcal{N}(x; 0, 1) dx$$

is the cumulative density function of the standard normal distribution. Trunctated normals can be efficiently sampled from using techniques developed by Robert [42].

B.2   The Hierarchical Gaussian Model

Let $x_{(i)}$ be a random variable with a Gaussian distribution with mean μ_i, variance σ_x^2 and density function

$$\mathcal{N}(x_{(i)}; \mu_i, \sigma_x^2) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x_i - \mu_i)^2}{2\sigma_x^2}}$$

and let $x_{(i),1:N_i}$ be a set of IID samples from $x_{(i)}$. Let μ_i be the ith variable from a first-order Markov sequence such that

$$P(\mu_i | \mu_{i-1}) = \mathcal{N}(\mu_i; \alpha\mu_{i-1} + (1-\alpha)\mu_f, \sigma_\mu^2)$$

is the transition density function and

$$P(\mu_1) = \mathcal{N}(\mu_1; \mu_0, \sigma_0^2)$$

is the initial probability. Then

$$P(\mu_1 | x_{(1),1:N_1}) = \mathcal{N}(\mu_1; \bar{\mu}_1, \bar{\sigma}_1^2)$$

and

$$P(\mu_2 | x_{(1),1:N_1}) = \mathcal{N}(\mu_2; \alpha\bar{\mu}_1 + (1-\alpha)\mu_0, \alpha^2\bar{\sigma}_1^2 + \sigma_\mu^2)$$

where

$$\bar{\mu}_1 = (\sigma_0^{-2} + N_1\sigma_x^{-2})^{-1}(\sigma_0^{-2}\mu_0 + \sigma_x^{-2}\sum_{i=1}^{N_1} x_{(1),i})$$

$$\bar{\sigma}_1^2 = (\sigma_0^{-2} + N_1\sigma_x^{-2})^{-1}$$

Then by induction,

$$P(\mu_i | x_{(1),1:N_1}, \ldots, x_{(i-1),1:N_{i-1}}) = \mathcal{N}(\mu_i; \bar{\mu}_i, \bar{\sigma}_i^2)$$

for $i > 1$ where

$$\bar{\mu}_i = \alpha\bar{\mu}_{i-1} + (1-\alpha)\mu_0$$

$$\bar{\sigma}_i^2 = \alpha^2\bar{\sigma}_{i-1}^2 + \sigma_\mu^2$$

For a set of observations X where $P(\mu_i|X) = \mathcal{N}(\mu_i; \hat{\mu}_i, \hat{\sigma}_i^2)$ then the probability of a new observation x of mean i given previous observations X is

$$P(x|X) = \mathcal{N}(x; \hat{\mu}_i, \hat{\sigma}_i^2 + \sigma_x^2)$$

These equations are the result of simple, if not tedious, algebraic manipulations which are omitted here.

B.3 The Gamma Distribution

A positive random variable x with a gamma distribution with shape parameter $s > 0$ and scale parameter $\lambda > 0$ has density function

$$\mathcal{G}(x; s, \lambda) = \frac{\lambda^s x^{s-1}}{\Gamma(s)} e^{-\lambda x}$$

mean

$$E[x] = \frac{s}{\lambda}$$

variance

$$V[x] = \frac{s}{\lambda^2}$$

and coefficient of variation

$$CV[x] = \frac{1}{\sqrt{s}}$$

Gamma variates can be sampled efficiently using methods found in [3].

If λ is unknown but has a gamma prior distribution with shape parameter s' and

scale parameter λ' then the distribution over a set of samples $x_{1:N}$ has density

$$
\begin{aligned}
P(x_{1:N}) &= \int_0^\infty P(x_{1:N}|\lambda)P(\lambda)d\lambda \\
&= \int_0^\infty \left(\prod_{i=1}^N P(x_i|\lambda)\right) P(\lambda)d\lambda \\
&= \int_0^\infty \left(\prod_{i=1}^N \frac{\lambda^s x_i^{s-1}}{\Gamma(s)} e^{-\lambda x_i}\right) \frac{\lambda'^{s'} \lambda^{s'-1}}{\Gamma(s')} e^{-\lambda\lambda'} d\lambda \\
&= \left(\prod_{i=1}^N \frac{x_i^{s-1}}{\Gamma(s)}\right) \frac{\lambda'^{s'}}{\Gamma(s')} \int_0^\infty \lambda^{Ns+s'-1} e^{-\lambda(\lambda'+\sum_{i=1}^N x_i)} d\lambda \\
&= \left(\prod_{i=1}^N \frac{x_i^{s-1}}{\Gamma(s)}\right) \frac{\lambda'^{s'}}{\Gamma(s')} \int_0^\infty \lambda^{S_N-1} e^{-\lambda\Lambda_N} d\lambda
\end{aligned}
$$

where

$$
S_N = s' + Ns
$$

and

$$
\Lambda_N = \lambda' + \sum_{i=1}^N x_i
$$

Noticing the integrands similarity to the gamma density function it follows that

$$
\int_0^\infty \lambda^{S_N-1} e^{-\lambda\Lambda_N} d\lambda = \frac{\Gamma(S_N)}{\Lambda_N^{S_N}}
$$

which, after substitution, gives the general formula

$$
P(x_{1:N}) = \frac{\lambda'^{s'}}{\Gamma(s')} \frac{\Gamma(S_N)}{\Lambda_N^{S_N}} \prod_{i=1}^N \frac{x_i^{s-1}}{\Gamma(s)} \tag{B.1}
$$

for the joint density of a set of samples.

A special case of Equation (B.1), when $N = 1$, gives the density

$$
\begin{aligned}
P(x_1) &= \frac{\lambda'^{s'}}{\Gamma(s')} \frac{\Gamma(S_1)}{\Lambda_1^{S_1}} \frac{x_1^{s-1}}{\Gamma(s)} \\
&= \frac{\lambda'^{s'}}{\beta(s',s)} \frac{x_1^{s-1}}{(\lambda'+x_1)^{s'+s}} \\
&= \frac{\lambda'^{-s}}{\beta(s',s)} \frac{x_1^{s-1}}{(1+\frac{x_1}{\lambda'})^{s'+s}}
\end{aligned}
$$

which is a scaled beta-prime distribution with scale parameter λ', primary shape parameter s and secondary shape parameter s'. Equation B.1 can also be used to find the posterior density

$$
\begin{aligned}
P(x_N|x_{1:N-1}) &= \frac{P(x_{1:N})}{P(x_{1:N-1})} \\
&= \frac{\Gamma(S_N)}{\Gamma(S_{N-1})} \frac{\Lambda_{N-1}^{S_{N-1}}}{\Lambda_N^{S_N}} \frac{x_N^{s-1}}{\Gamma(s)} \\
&= \frac{\Gamma(S_{N-1}+s)}{\Gamma(S_{N-1})\Gamma(s)} \frac{\Lambda_{N-1}^{S_{N-1}} x_N^{s-1}}{(\Lambda_{N-1}+x_N)^{S_{N-1}+s}} \\
&= \frac{\Lambda_{N-1}^{-s}}{\beta(S_{N-1},s)} \frac{x_N^{s-1}}{(1+\frac{x_N}{\Lambda_{N-1}})^{S_{N-1}+s}}
\end{aligned}
$$

which is again a scaled beta-prime distribution but with scale parameter Λ_{N-1}, primary shape parameter s and secondary shape parameter S_{N-1}.

B.4   The Scaled Beta-Prime Distribution

A scaled beta-prime random variable x has density

$$
\mathcal{B}'(x;a,b,\lambda) = \frac{\lambda^{-a}}{\beta(a,b)} \frac{x^{a-1}}{(1+\frac{x}{\lambda})^{a+b}}
$$

with scale parameter λ and primary and secondary shape parameters a and b. The uncentered nth moment of x is

$$
\begin{aligned}
E[x^n] &= \int_0^\infty x^n \frac{\lambda^{-a}}{\beta(a,b)} \frac{x^{a-1}}{(1+\frac{x}{\lambda})^{a+b}} dx \\
&= \frac{\lambda^{-a}}{\beta(a,b)} \int_0^\infty \frac{x^{a+n-1}}{(1+\frac{x}{\lambda})^{a+b}} dx \\
&= \frac{\lambda^{-a}}{\beta(a,b)} \frac{\beta(a+n,b-n)}{\lambda^{-(a+n)}} \\
&= \lambda^n \frac{\Gamma(a+n)\Gamma(b-n)\Gamma(a+b)}{\Gamma(a)\Gamma(b)\Gamma(a+b)} \\
&= \lambda^n \frac{\Gamma(a+n)\Gamma(b-n)}{\Gamma(a)\Gamma(b)}
\end{aligned}
$$

which gives a mean of

$$
\begin{aligned}
E[x] &= \lambda \frac{\Gamma(a+1)\Gamma(b-1)}{\Gamma(a)\Gamma(b)} \\
&= \lambda \frac{a\Gamma(a)}{\Gamma(a)(b-1)\Gamma(b)} \\
&= \lambda \frac{a}{b-1}
\end{aligned}
$$

for $b > 1$ and a variance of

$$
\begin{aligned}
V[x] &= E[x^2] - E[x]^2 \\
&= \lambda^2 \frac{a(a+1)\Gamma(a)\Gamma(b-2)}{\Gamma(a)(b-1)(b-2)\Gamma(b-2)} - \lambda^2 \frac{a^2}{(b-1)^2} \\
&= \lambda^2 \left(\frac{a(a+b-1)}{(b-1)^2(b-2)} \right)
\end{aligned}
$$

for $b > 2$. Note that for values of $b \le 2$ $V[x]$ can still be computed, as can $E[x]$ for $b \le 1$, but the form of the equations are much messier and the constraint $b \ge 2$ is generally not that onerous.

To sample from a scaled beta-prime distribution simply draw y with probability $P(y) = \mathcal{B}(y; a, b)$ where

$$
\mathcal{B}(x; a, b) = \frac{x^{a-1}(1-x)^{b-1}}{\beta(a, b)}
$$

is the probability density function of the beta distribution with parameters a and b. Efficient sampling methods for beta distributions can be found in [4]. Then $x = \lambda \frac{y}{1-y}$ has density $\mathcal{B}'(x; a, b, \lambda)$ which can be confirmed using the method of moments

$$
\begin{aligned}
E[(\lambda \frac{y}{1-y})^n] &= \int_0^1 (\lambda \frac{y}{1-y})^n \frac{y^{a-1}(1-y)^{b-1}}{\beta(a, b)} dy \\
&= \frac{\lambda^n}{\beta(a, b)} \int_0^1 y^{a+n-1}(1-y)^{b-n-1} dy \\
&= \lambda^n \frac{\beta(a+n, b-n)}{\beta(a, b)} \\
&= \lambda^n \frac{\Gamma(a+n)\Gamma(b-n)}{\Gamma(a)\Gamma(b)}
\end{aligned}
$$

which shows that the uncentered moments of x and $\lambda \frac{y}{1-y}$ are the same and thus have the same distribution.

Bibliography

[1] Alessandro Bissacco and Stefano Soatto. Classifying human dynamics without contact forces. *CVPR*, 2:1678–1685, 2006. ISSN 1063-6919. doi: 10.1109/CVPR.2006. 75.

[2] R. Blickhan and R. J. Full. Similarity in multilegged locomotion: Bouncing like a monopode. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 173(5):509–517, November 1993. doi: 10.1007/BF00197760.

[3] R. C. H. Cheng. The generation of gamma variables with non-integral shape parameter. *Applied Statistics*, 26(1):71–75, 1977. ISSN 0035-9254.

[4] R. C. H. Cheng. Generating beta variates with nonintegral shape parameters. *Commun. ACM*, 21(4):317–322, 1978. ISSN 0001-0782. doi: 10.1145/359460.359482.

[5] S.H. Collins and A. Ruina. A bipedal walking robot with efficient and human-like gait. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 1983–1988, 2005.

[6] Steve Collins, Andy Ruina, Russ Tedrake, and Martijn Wisse. Efficient Bipedal Robots Based on Passive-Dynamic Walkers. *Science*, 307(5712):1082–1085, 2005. doi: 10.1126/science.1107799.

[7] Arnaud Doucet, Simon Godsill, and Christophe Andrieu. On sequential monte carlo

sampling methods for bayesian filtering. *Statistics and Computing*, 10(3):197–208, July 2000. doi: 10.1023/A:1008935410038.

[8] David H. Eberly. *Game Physics*. Morgan Kaufmann, 2004.

[9] H Elftman. The measurment of external force in walking. *Science*, 88:152–153, 1938.

[10] Anthony C. Fang and Nancy S. Pollard. Efficient synthesis of physically valid human motion. *ACM Transactions on Graphics*, 22(3):417–426, 2003. ISSN 0730-0301. doi: 10.1145/882262.882286.

[11] Mariano Garcia, Anindya Chatterjee, Andy Ruina, and Michael Coleman. The simplest walking model: Stability, complexity, and scaling. *Journal of Biomechanical Engineering*, 120(2):281–288, 1998.

[12] R. Douglas Gregory. *Classical Mechanics*. Cambridge University Press, 2006.

[13] Radek Grzeszczuk, Demetri Terzopoulos, and Geoffrey Hinton. Neuroanimator: fast neural network emulation and control of physics-based models. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 9–20, New York, NY, USA, 1998. ACM Press. ISBN 0-89791-999-8. doi: 10.1145/280814.280816.

[14] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka. The development of honda humanoid robot. In *Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on*, volume 2, pages 1321–1326 vol.2, 1998.

[15] Jessica K. Hodgins, Wayne L. Wooten, David C. Brogan, and James F. O'Brien. Animating human athletics. In *SIGGRAPH '95: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 71–78, New York, NY, USA, 1995. ACM Press. ISBN 0-89791-701-4. doi: 10.1145/218380.218414.

[16] M. Isard and A. Blake. A mixed-state condensation tracker with automatic model-switching. In *Computer Vision, 1998. Sixth International Conference on*, pages 107–112, 1998. doi: 10.1109/ICCV.1998.710707.

[17] L. Kakadiaris and D. Metaxas. Model-based estimation of 3d human motion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1453–1459, 2000. ISSN 0162-8828. doi: 10.1109/34.895978.

[18] David Kincaid and Ward Cheney. *Numerical Analysis: Mathematics of Scientific Computing*. Brooks/Cole, 3rd edition, 2001.

[19] Arthur D. Kuo. A simple model of bipedal walking predicts the preferred speed–step length relationship. *Journal of Biomechanical Engineering*, 123(3):264–269, June 2001. doi: 10.1115/1.1372322.

[20] Arthur D Kuo. Energetics of actively powered locomotion using the simplest walking model. *Journal of Biomechanical Engineering*, 124:113–120, February 2002. doi: 10.1115/1.427703.

[21] Michael Leventon and William Freeman. Bayesian estimation of 3-d human motion from an image sequence. Technical Report TR-98-06, MERL, Cambridge, MA, July 1998.

[22] C. Karen Liu and Zoran Popović. Synthesis of complex dynamic character motion from simple animations. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 408–416, New York, NY, USA, 2002. ACM Press. ISBN 1-58113-521-1. doi: 10.1145/566570.566596.

[23] C. Karen Liu, Aaron Hertzmann, and Zoran Popović. Learning physics-based motion style with nonlinear inverse optimization. *ACM Transactions on Graphics*, 24(3): 1071–1081, 2005. ISSN 0730-0301. doi: 10.1145/1073204.1073314.

[24] Jun S. Liu and Rong Chen. Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93(443):1032–1044, September 1998. ISSN 0162-1459.

[25] Jun S. Liu, Rong Chen, and Wing Hung Wong. Rejection control and sequential importance sampling. *Journal of the American Statistical Association*, 93(443):1022–1031, September 1998. ISSN 0162-1459.

[26] Zicheng Liu, Steven J. Gortler, and Michael F. Cohen. Hierarchical spacetime control. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 35–42, New York, NY, USA, 1994. ACM Press. ISBN 0-89791-667-0. doi: 10.1145/192161.192169.

[27] Tad McGeer. Passive dynamic walking. *International Journal of Robotics Research*, 9(2):62–82, 1990. ISSN 0278-3649.

[28] Tad McGeer. Passive walking with knees. In *IEEE International Conference on Robotics and Automation*, volume 3, pages 1640–1645, 1990.

[29] Tad McGeer. Passive bipedal running. In *Proceedings of the Royal Society of London*, volume 240 of *B, Biological Sciences*, pages 107–134, 1990.

[30] Tad McGeer. *Advances in Comparative and Environmental Physiology*, volume 11, chapter 4, pages 113–139. Springer-Verlag, 1992.

[31] Tad McGeer. Dynamics and control of bipedal locomotion. *Journal of Theoretical Biology*, 163:277–314, 1993.

[32] D. Metaxas and D. Terzopoulos. Shape and nonrigid motion estimation through physics-based synthesis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(6):580–591, 1993. ISSN 0162-8828. doi: 10.1109/34.216727.

[33] B. North, A. Blake, M. Isard, and J. Rittscher. Learning and classification of complex dynamics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22 (9):1016–1034, 2000. ISSN 0162-8828. doi: 10.1109/34.877523.

[34] V. Pavlovic, J.M. Rehg, Tat-Jen Cham, and K.P. Murphy. A dynamic bayesian network approach to figure tracking using learned dynamic models. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, 1999.

[35] E. Poon and D.J. Fleet. Hybrid monte carlo filtering: edge-based people tracking. In *Motion and Video Computing, 2002. Proceedings. Workshop on*, pages 151–158, 2002. doi: 10.1109/MOTION.2002.1182228.

[36] Marko B. Popovic, Ambarish Goswami, and Hugh Herr. Ground reference points in legged locomotion: Definitions, biological trajectories and control implications. *The International Journal of Robotics Research*, 24(12):1013–1032, 2005. doi: 10.1177/ 0278364905058363.

[37] Zoran Popović and Andrew Witkin. Physically based motion transformation. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 11–20, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co. ISBN 0-201-48560-5. doi: 10.1145/311535. 311536.

[38] Gill A. Pratt. Legged robots at MIT: what's new since Raibert? *Robotics & Automation Magazine, IEEE*, 7(3):15–19, 2000. ISSN 1070-9932.

[39] Jerry Pratt and Gill Pratt. Exploiting natural dynamics in the control of a 3D bipedal walking simulation. In *Proceedings of the International Conference on Climbing and Walking Robots (CLAWAR99)*, Portsmouth, UK, September 1999.

[40] Marc Raibert. *Legged Robots That Balance*. MIT Press, Cambridge, Mass., 1986.

[41] Marc H. Raibert and Jessica K. Hodgins. Animation of dynamic legged locomotion. In *SIGGRAPH '91: Proceedings of the 18th annual conference on Computer graphics and interactive techniques*, pages 349–358, New York, NY, USA, 1991. ACM Press. ISBN 0-89791-436-8. doi: 10.1145/122718.122755.

[42] Christian P. Robert. Simulation of truncated normal variables. *Statistics and Computing*, 5(2):121–125, June 1995. doi: 10.1007/BF00143942.

[43] Gordon E. Robertson, Graham Caldwell, Joseph Hamill, Gary Kamen, and Sandy Whittlesey. *Research Methods in Biomechanics*. Human Kinetics, 2004.

[44] Andy Ruina, John E.A. Bertram, and Manoj Srinivasan. A collisional model of the energetic cost of support work qualitatively explains leg sequencing in walking and galloping, pseudo-elastic leg behavior in running and the walk-to-run transition. *Journal of Theoretical Biology*, 237(2):170–192, November 2005. doi: 10.1016/j.jtbi.2005.04.004.

[45] Alla Safonova, Jessica K. Hodgins, and Nancy S. Pollard. Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. *ACM Transactions on Graphics*, 23(3):514–521, 2004. ISSN 0730-0301. doi: 10.1145/1015706.1015754.

[46] Hedvig Sidenbladh, Michael J. Black, and David J. Fleet. Stochastic tracking of 3d human figures using 2d image motion. In *ECCV '00: Proceedings of the 6th European Conference on Computer Vision-Part II*, pages 702–718, London, UK, 2000. Springer-Verlag. ISBN 3-540-67686-4.

[47] Manoj Srinivasan and Andy Ruina. Computer optimization of a minimal biped model discovers walking and running. *Nature*, 439(7072):72–75, January 2006. ISSN 0028-0836. doi: 10.1038/nature04113.

[48] D. Terzopoulos and D. Metaxas. Dynamic 3d models with local and global deformations: deformable superquadrics. In *Computer Vision, 1990. Proceedings, Third International Conference on*, 1990.

[49] R. Urtasun, D.J. Fleet, and P. Fua. 3d people tracking with gaussian process dynamical models. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 238–245, 2006. doi: 10.1109/CVPR.2006.15.

[50] Raquel Urtasun, David J. Fleet, Aaron Hertzmann, and Pascal Fua. Priors for people tracking from small training sets. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, pages 403–410, Washington, DC, USA, 2005. IEEE Computer Society. ISBN 0-7695-2334-X-01. doi: 10.1109/ICCV.2005.193.

[51] Richard Q. van der Linde and Arend L. Schwab. Lecture notes multibody dynamics B, wb1413, course 1997/1998. Laboratory for Engineering Mechanics, Delft University of Technology, 2002. URL `http://tam.cornell.edu/~als93/LinSch02.pdf`.

[52] M. Vukobratovic and D Juricic. Contributions to the synthesis of biped gait. *IEEE Transactions on Biomedical Engineering*, 16:1–6, 1969.

[53] Andrew Witkin and Michael Kass. Spacetime constraints. *SIGGRAPH Computer Graphics*, 22(4):159–168, 1988. ISSN 0097-8930. doi: 10.1145/378456.378507.