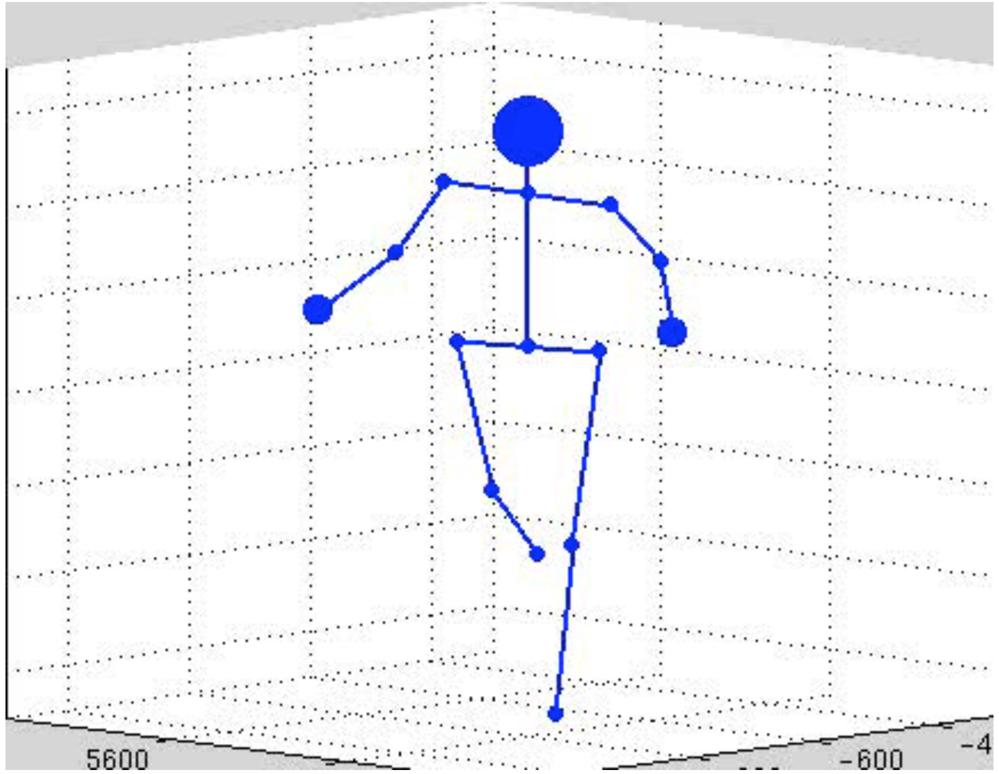
Human Pose Tracking III: Dynamics

David Fleet
University of Toronto

CIFAR Summer School, 2009

Interactions with the world are fundamental

Implausible motions



[Poon and Fleet, 01]

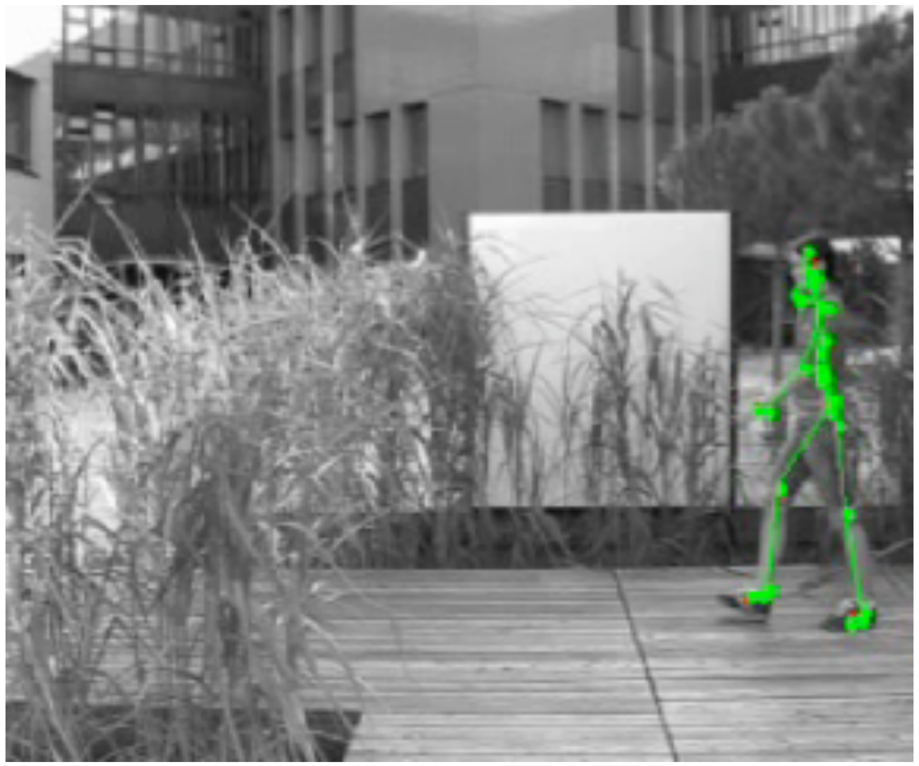
- Kinematic Model: damped 2nd-order Markov model with Beta process noise and joint angle limits
- Observations: steerable pyramid coefficients (image edges)
- Inference: hybrid Monte Carlo particle filter

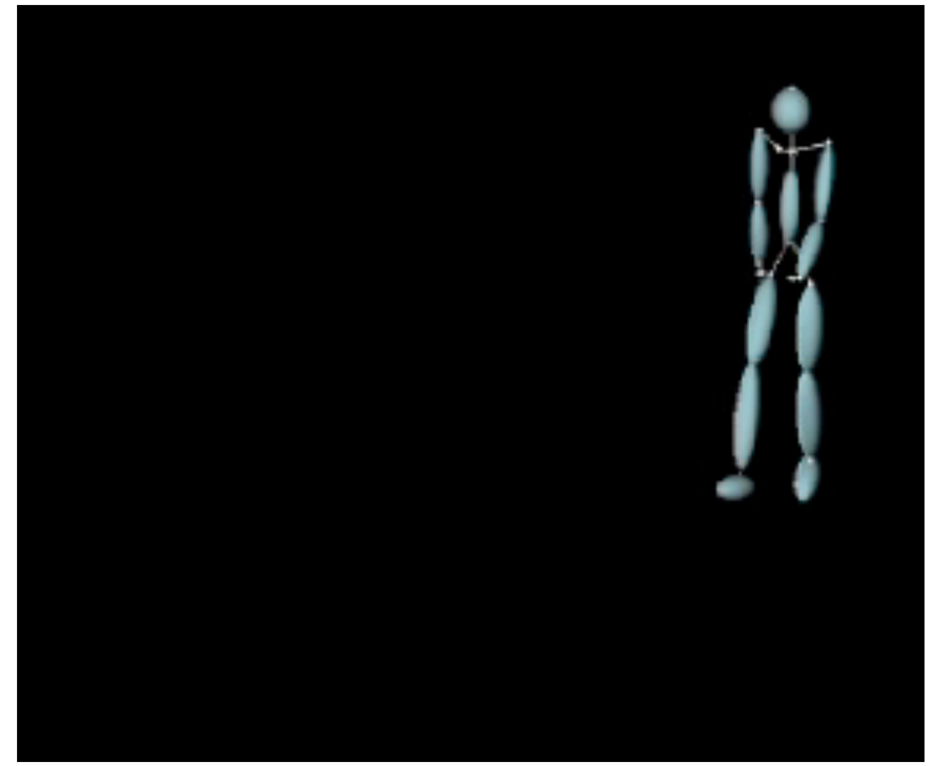
Implausible motions

[Urtasun et al. ICCV `05]

- Kinematic Model: GPLVM for pose, with 2nd-order dynamics
- Observations: tracked 2D patches on body (WSL tracker)
- Inference: MAP estimation (hill climbing)

Implausible motions





[Urtasun et al. CVPR `06]

- Kinematic Model: Gaussian process dynamical model (GPDM)
- Observations: tracked 2D patches on body (WSL tracker)
- Inference: MAP estimation (hill climbing) with sliding window

Will learning scale?

Problem: Learning kinematic pose and motion models from motion capture data, with dependence on the environment and other bodies, may be untenable …

Physics-based models

Physics specifies the motions of bodies and their interactions in terms of inertial descriptions and forces, and generalize naturally to account for:

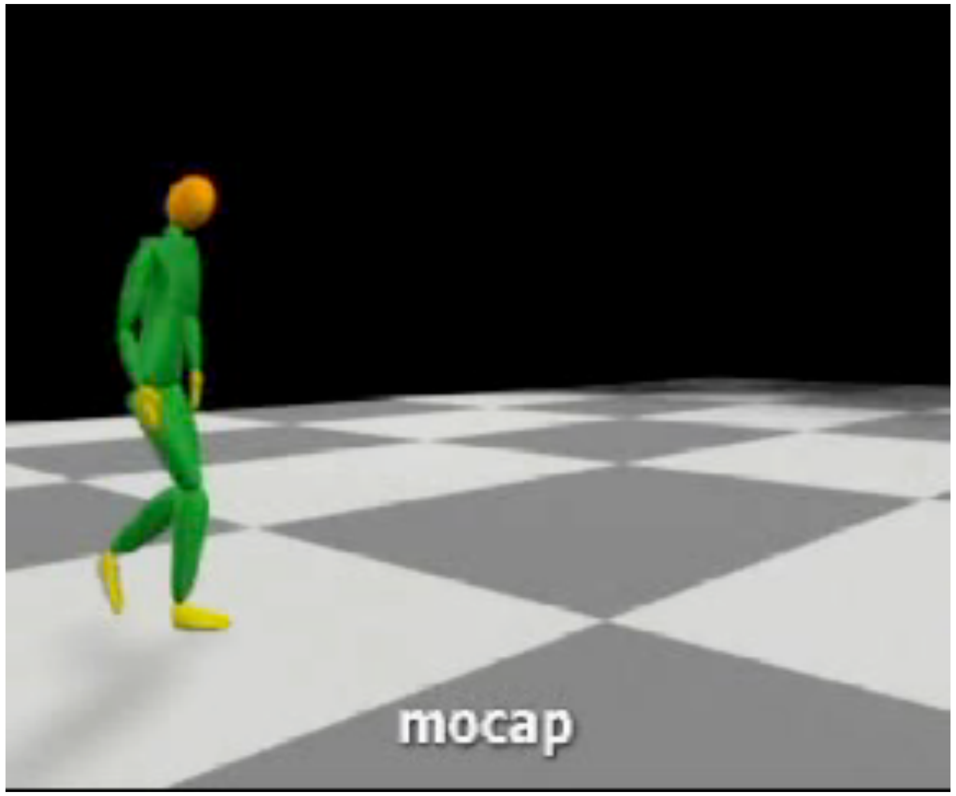
- balance and body lean (e.g., on hills)
- sudden accelerations (e.g., collisions)
- static contact (e.g., avoiding footskate)
- variations in style due to speed and mass distribution (e.g., carrying an object)
- …

Physics-based models for pose tracking

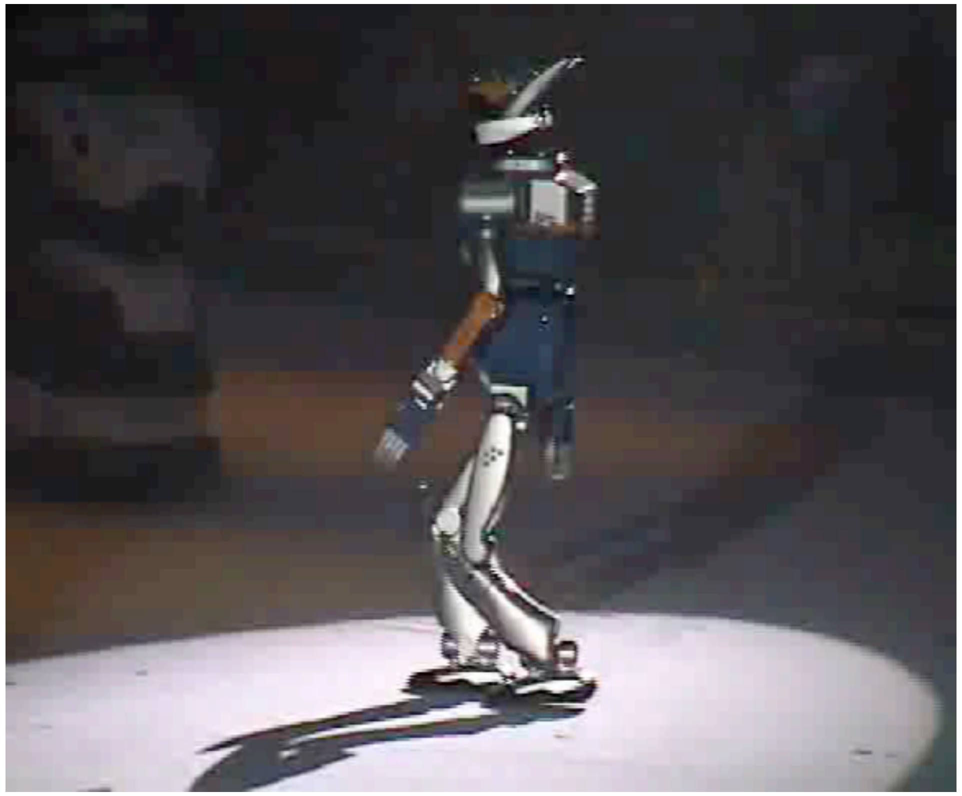
How should we incorporate physical principles in models of biological motion?

- to ensure physically plausible pose estimates
- to reduce reliance on mocap data
- to understanding interactions

Modeling full-body dynamics is difficult



[Liu et al. `06]



[Kawada Industries HRP-2. `03]

Passive dynamics

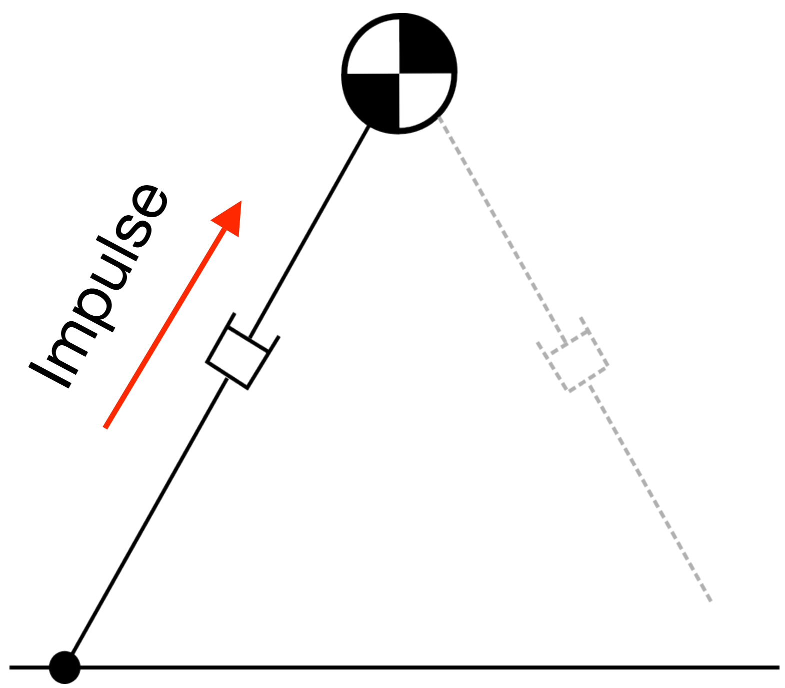
But much of walking is essentially passive.

[McGeer 1990]

[Collins & Ruina 2005]

Simplified planar biomechanical models

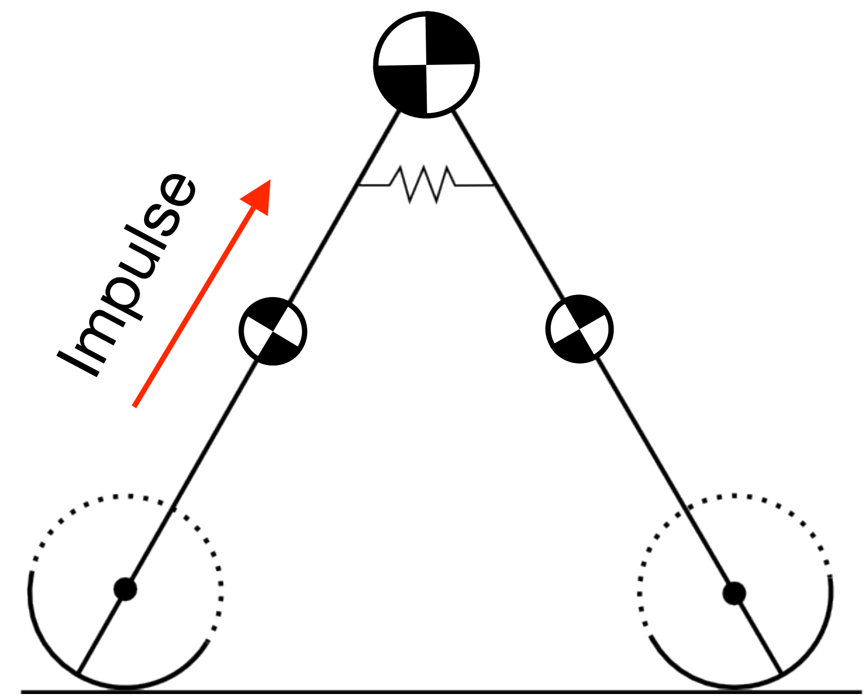
Monopode



[Blickhan & Full 1993; *Srinivasan & Ruina 2000]*

- point-mass at hip, massless legs with prismatic joints, and impulsive toe-off force
- inverted pendular motion

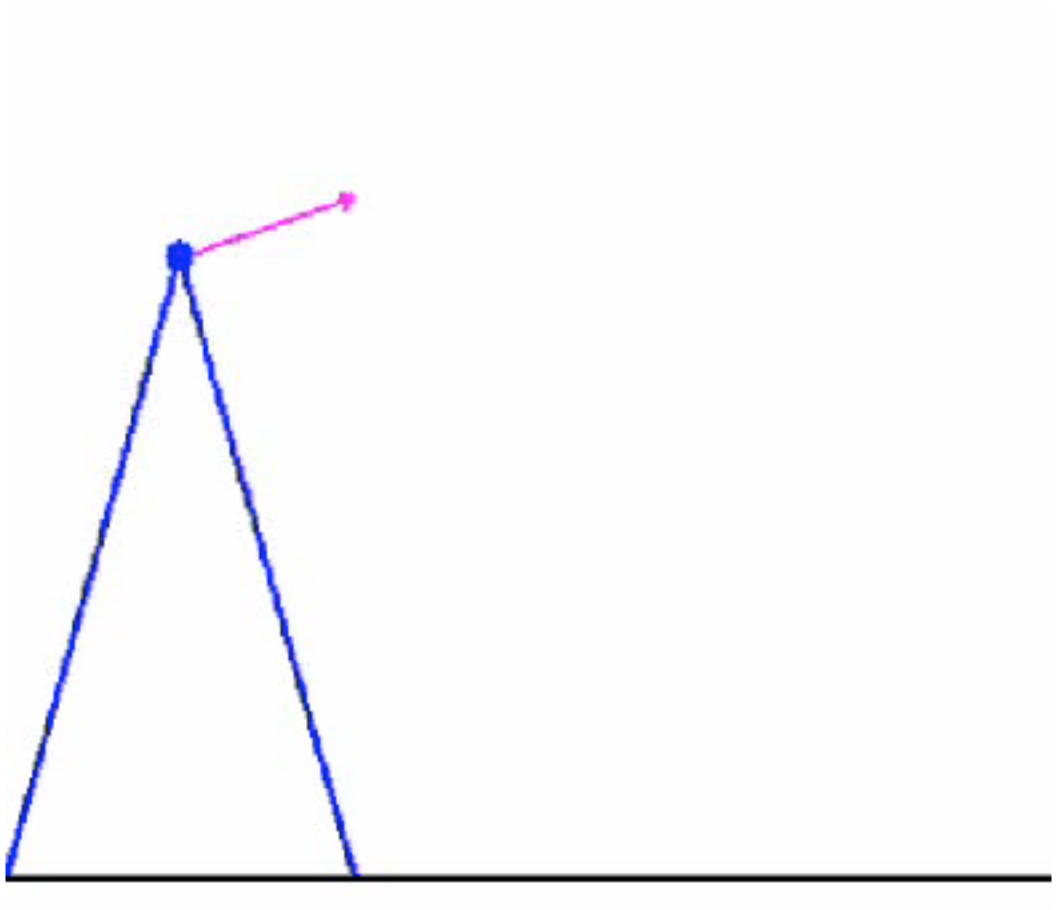
Anthropomorphic Walker



[McGeer 1990; Kuo 2001,2002]

- rigid bodies for torso and legs
- forces due to torsional spring between legs and an impulsive toe-off

Anthropomorphic walker gait



Different speeds and step-lengths

Speed: 6.7 km/hr; Step length: 0.875m

[Brubaker et al. `07]

Different speeds and step-lengths

Speed: 4.0 km/hr; Step length: 0.875m

Different speeds and step-lengths

Speed: 2.7 km/hr; Step length: 0.875m

Different speeds and step-lengths

Speed: 2.7 km/hr; Step length: 0.625m

Different speeds and step-lengths

Speed: 2.7 km/hr; Step length: 0.375m

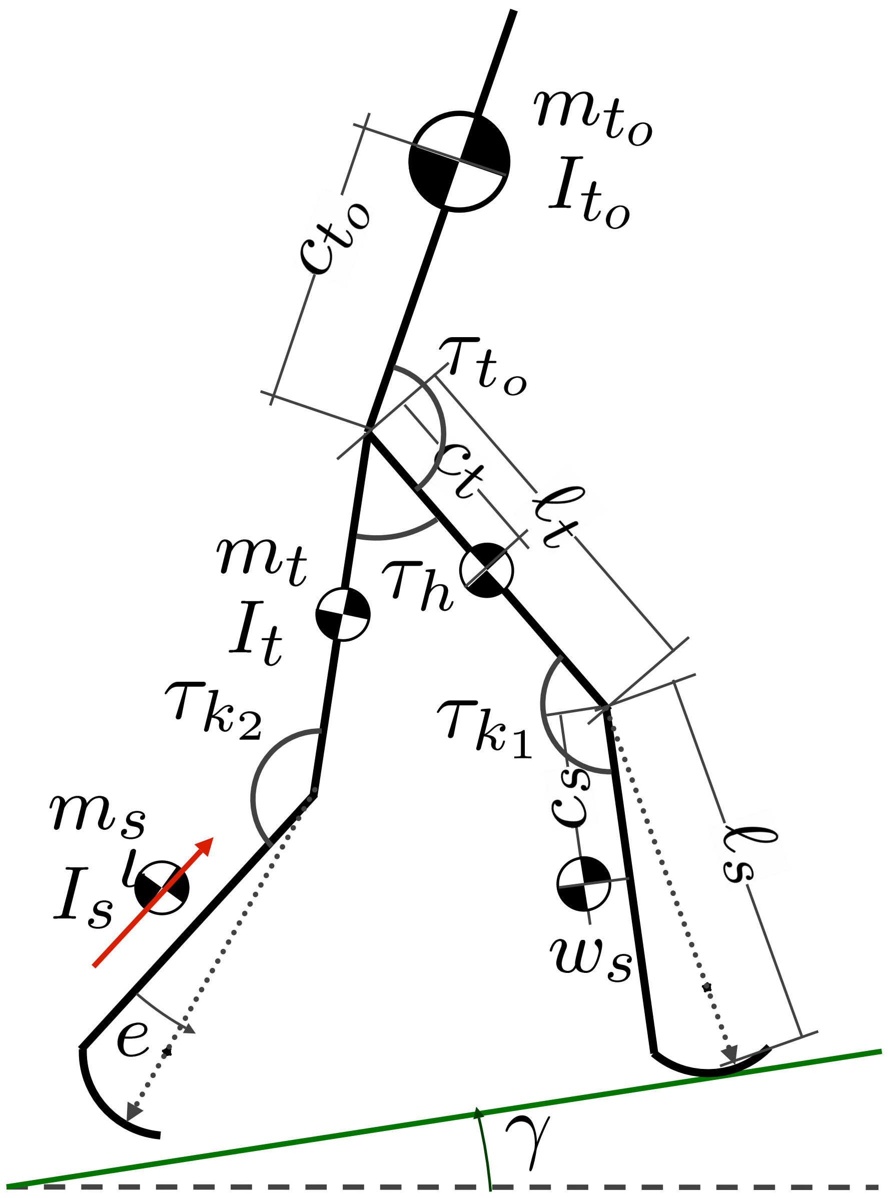
The *Kneed Walker*

Kneed planar walker comprises

- torso, legs with knees & feet
- inertial parameters from biomechanical data

Dynamics due to:

- joint torques $\tau_{t_o}, \tau_h, \tau_{k_1}, \tau_{k_2}$ (for torso, hip, & knees)
- impulse applied at toe-off (with magnitude ι)
- gravitational acceleration (w.r.t. ground slope γ)



[Brubaker and Fleet `08]

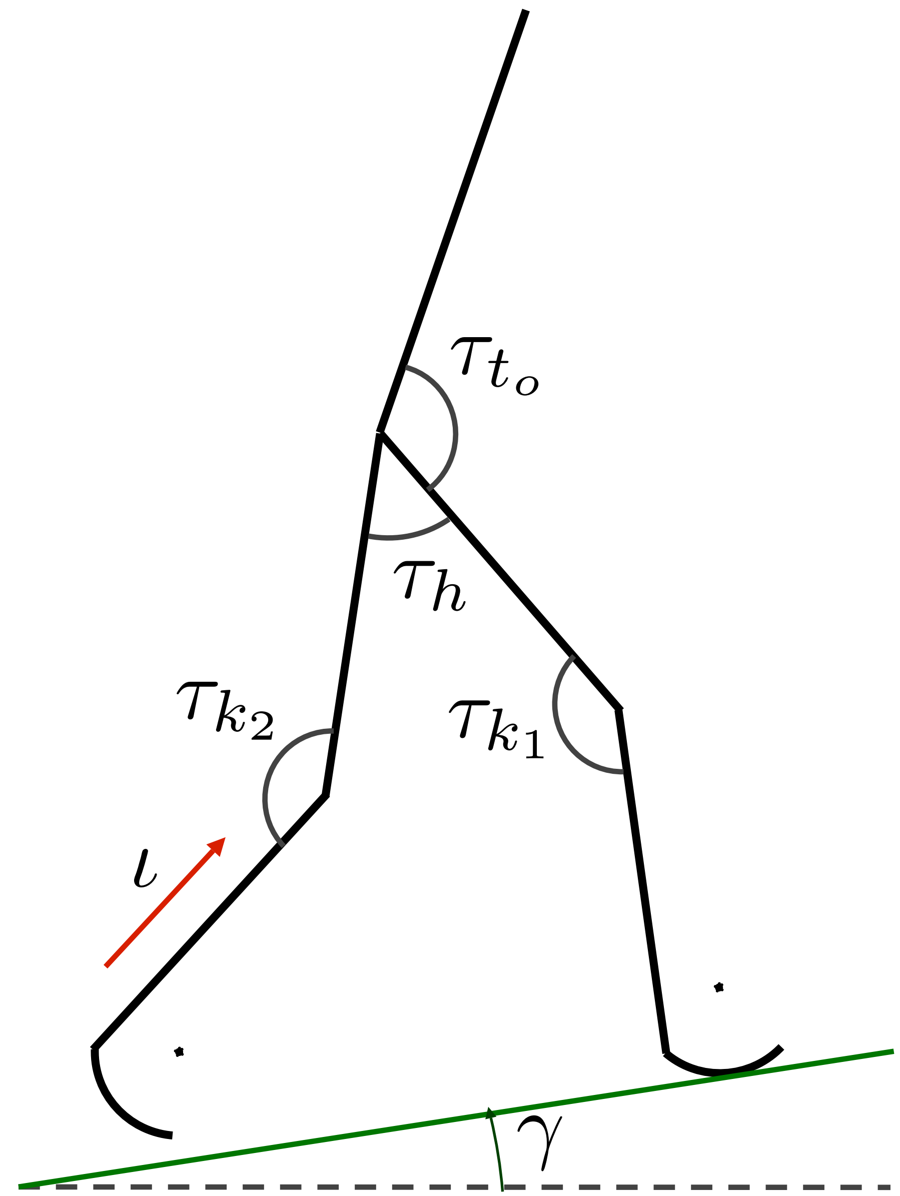
The *Kneed Walker*

Joint torques are parameterized as damped linear springs.

For hip torque

$$\begin{aligned}\tau_h &= \kappa_h\,(\phi_{t_2}+\phi_{t_1}-\phi_h)\\ &\quad - d_h(\dot\phi_{t_2}+\dot\phi_{t_1})\end{aligned}$$

with stiffness and damping coefficients, κ_h and d_h, and resting length ϕ_h



The *Kneed Walker*

Equations of motion

$$\mathcal{M}\ \ddot{\mathbf{q}} = f_s(\vec{\kappa},\, \vec{d},\, \vec{\phi}) + f_g + f_c$$

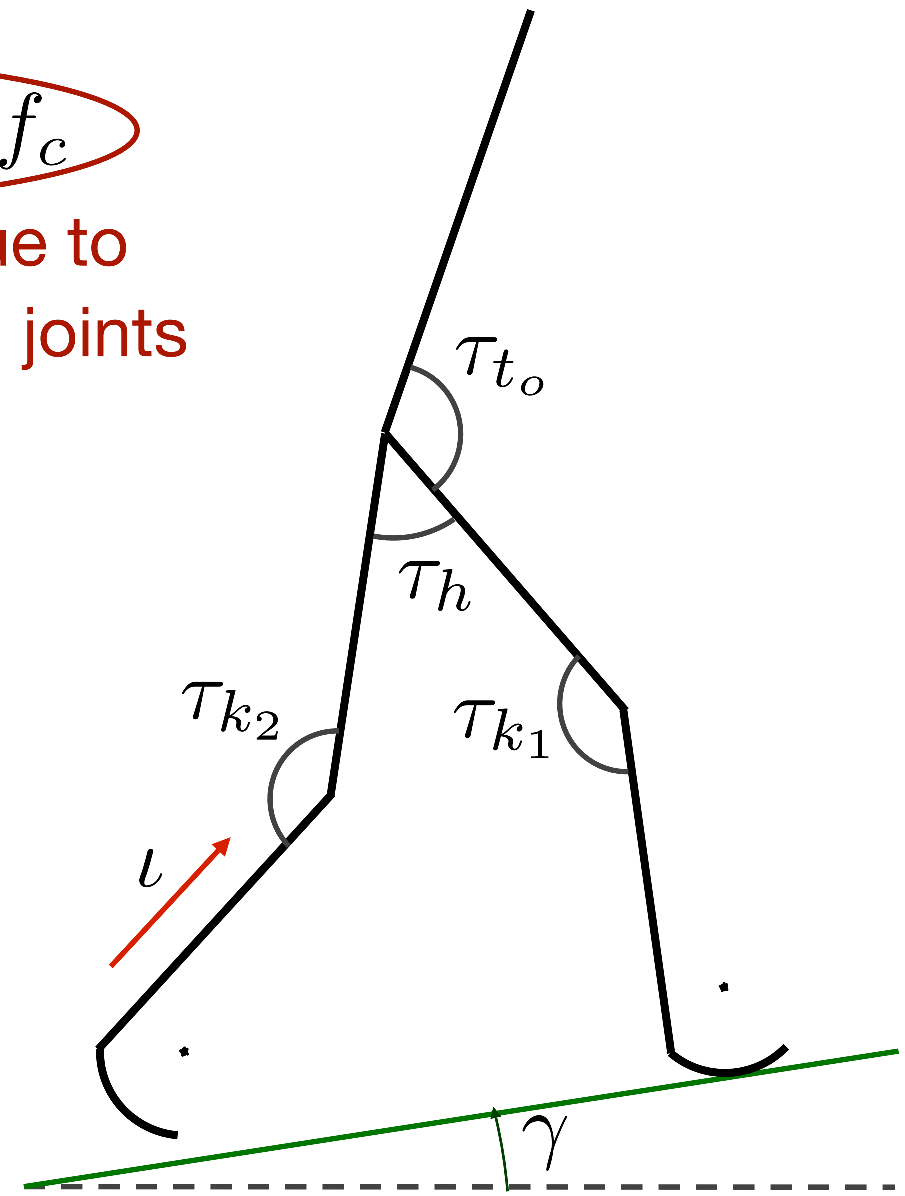
generalized mass matrix

acceleration

spring torques

forces due to gravity and joints

... plus ground collisions and joint limits (esp. knee)



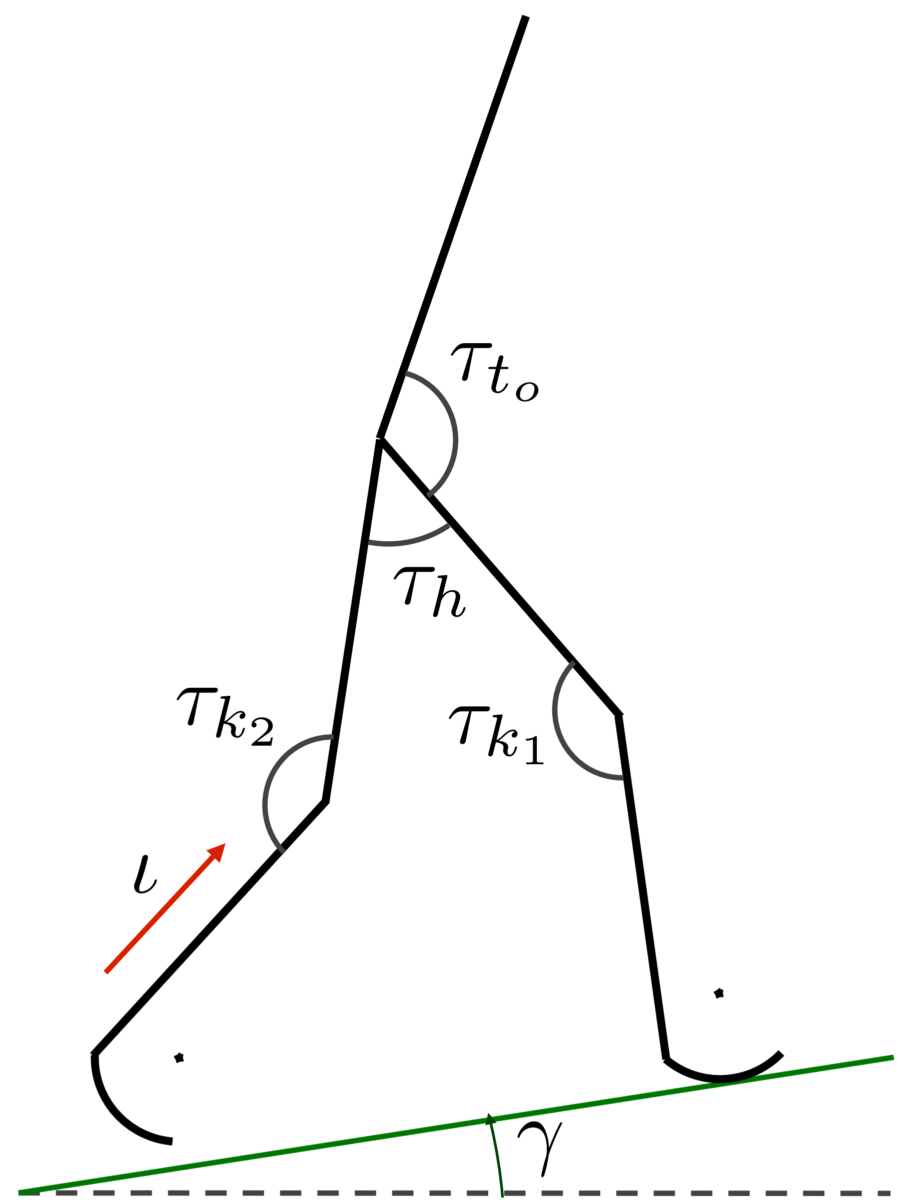
Kneed Walker: Ground collisions

Ground collisions modeled as instantaneous and inelastic.

Produces an instantaneous change in velocities

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

post-contact velocites
pre-contact velocities
impulse



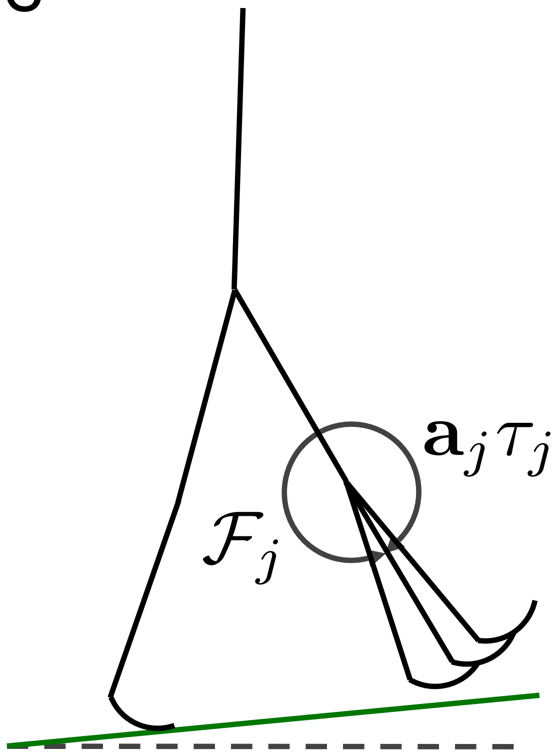
Kneed Walker: Joint limits

Joint limits easily expressed as constraints $\mathbf{a}^T\mathbf{q} \geq b$

When a joint limit violation is detected in simulation

- localize constraint boundary (i.e., the time at which joint limit reached), and treat as impulsive collision
- as long as constraint is then “active”, include a virtual reactive force to enforce joint limits
- augmented equations of motion

$$\begin{bmatrix} \mathcal{M} & -\mathbf{a} \\ \mathbf{a}^T & \mathbf{0} \end{bmatrix} \begin{pmatrix} \ddot{\mathbf{q}} \\ \tau \end{pmatrix} = \begin{pmatrix} \mathcal{F} \\ \mathbf{0} \end{pmatrix}$$



Prior for the Kneed Walker

How do we design a prior distribution over the dynamics parameters to encourage plausible human-like walking motions?

Assumption: Human walking motions are characterized by efficient, stable, cyclic gaits.

Approach:

- Find control parameters that, with minimal energy, produce optimal cyclic gaits over a wide range of natural human speeds and step lengths, for a range of surface slopes.
- Assume additive process noise in the control parameters to capture variations in style.

Efficient, cyclic gaits

Search for dynamics parameters $\vec{\theta} = (\vec{\kappa},\ \vec{d},\ \vec{\phi},\ \iota)$ and initial state $\mathbf{x} = (\mathbf{q},\ \dot{\mathbf{q}})$ that produce cyclic locomotion at speed s, step length ℓ, and slope γ with minimal "energy".

Solve

$$\min_{\vec{\theta},\mathbf{x}} E(\vec{\theta},\ \mathbf{x};\ s,\ \ell,\ \gamma) \quad \text{s.t.} \quad C(\vec{\theta},\ \mathbf{x};\ s,\ \ell,\ \gamma) < \epsilon$$

where

$$E(\vec{\theta}, \mathbf{x}; s, \ell, \gamma) \ = \ \alpha_\iota \iota^{1.5} + \sum_{j \in joints} \frac{\alpha_j}{T} \int_0^T \tau_j(t; \mathbf{x}_0, \vec{\theta}, \gamma)^2 \, dt$$

and $C(\vec{\theta}, \mathbf{x}; s, \ell, \gamma)$ measures the deviation from periodic motion with target speed and step-length

Efficient, cyclic gaits

Speed: 5.8 km/hr; Step length: 0.6 m; Slope: 0°

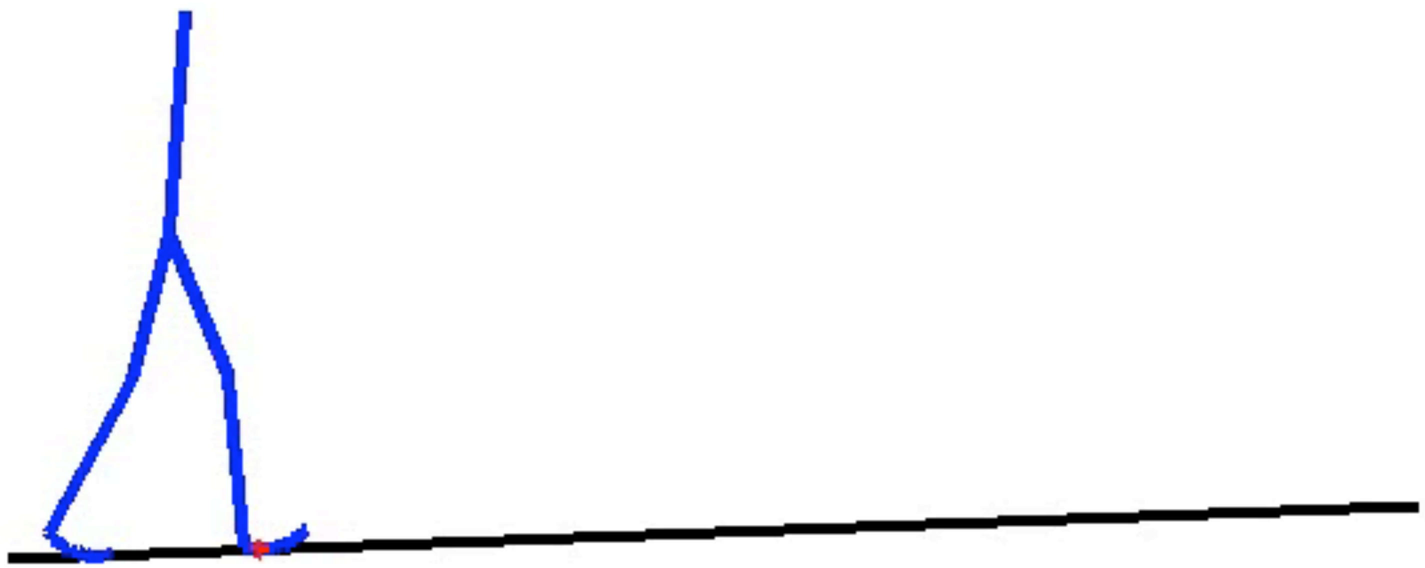
Efficient, cyclic gaits

Speed: 6.5 km/hr; Step length: 0.6 m; Slope: 4.3°

Efficient, cyclic gaits

Speed: 3.6 km/hr; Step length: 0.4 m; Slope: 4.3°

Efficient, cyclic gaits



Speed: 5.0 km/hr; Step length: 0.6 m; Slope: 2.1°

Efficient, cyclic gaits

Speed: 4.3 km/hr; Step length: 0.8 m; Slope: -2.1°

Efficient, cyclic gaits

Speed: 5.8 km/hr; Step length: 1.0 m; Slope: -4.3°

Stochastic dynamics

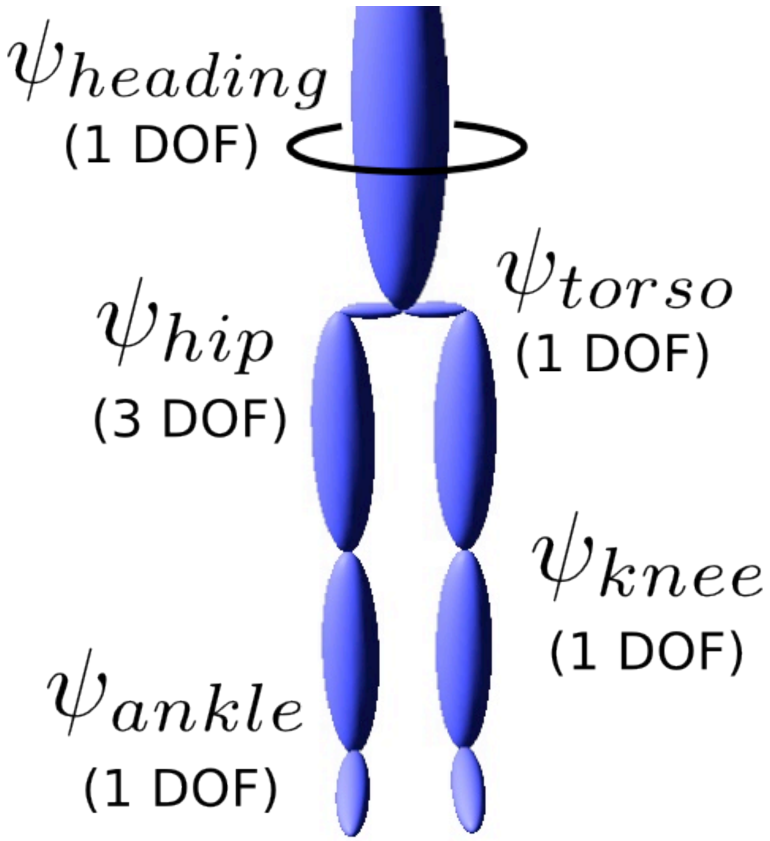
Our prior over human walking motions is derived from the manifold of optimal cyclic gaits:

- We assume additive noise on the control parameters (spring stiffness, resting lengths, and impulse magnitude).
- We also assume additive noise on the resulting torques.

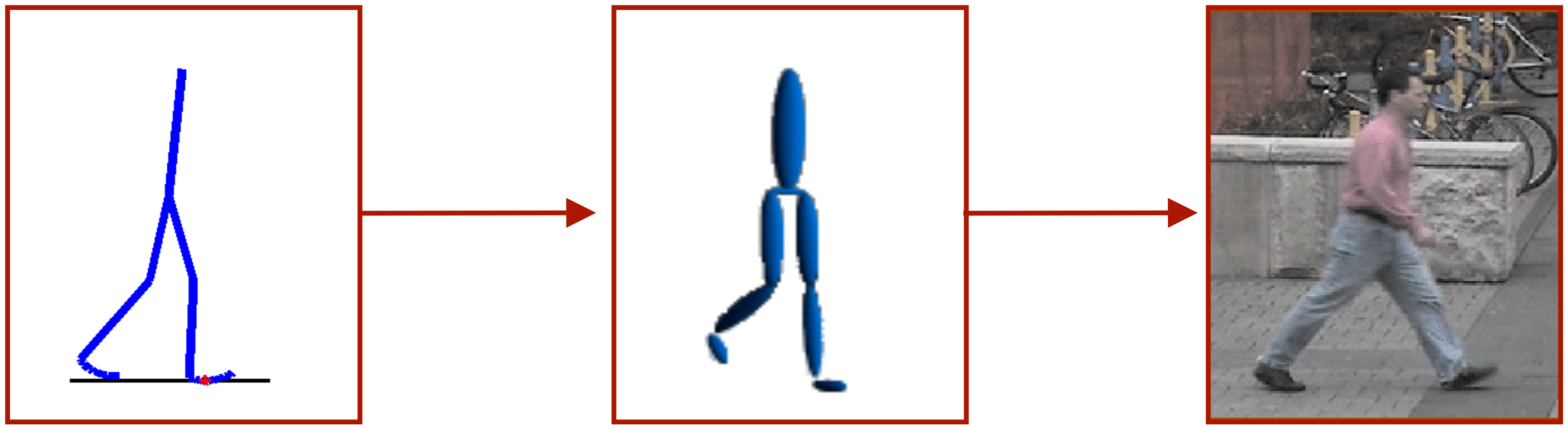
3D kinematic model

Kinematic parameters (15D) include global torso position and orientation, plus hips, knees and ankles.

- dynamics constrains contact of stance foot, hip angles (in sagittal plane), and knee/ankle angles
- other parameters modeled as smooth, second-order Markov processes.
- limb lengths assumed to be static



Graphical model



2D dynamics | 3D kinematics | image observations

Bayesian people tracking

Image observations: $\mathbf{z}_{1:t} \equiv (\mathbf{z_1}, ..., \mathbf{z}_t)$

State: $\mathbf{s}_t = [d_t, k_t]$

dynamics pose

Posterior distribution:

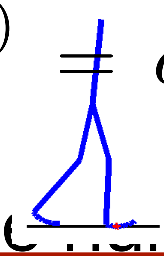
$$p(\mathbf{s}_{1:t} \,|\, \mathbf{z}_{1:t}) \;\propto\; p(\mathbf{z}_t \,|\, \mathbf{s}_t)\; p(\mathbf{s}_t \,|\, \mathbf{s}_{1:t-1})\; p(\mathbf{s}_{1:t-1} \,|\, \mathbf{z}_{1:t-1})$$

likelihood transition posterior

Sequential Monte Carlo inference:

- particle set $\mathcal{S}_t = \{ \mathbf{s}_{1:t}^{(j)}, w_t^{(j)} \}_{j=1}^{N}$ approximates $p(\mathbf{s}_{1:t} \,|\, \mathbf{z}_{1:t})$
- step 1. sample next state: $\mathbf{s}_t^{(j)} \sim p(\mathbf{s}_t \,|\, \mathbf{s}_{t-1}^{(j)})$
- step 2. update weight: $w_t^{(j)} = c\, w_{t-1}^{(j)}\, p(\mathbf{z}_t \,|\, \mathbf{s}_t^{(j)})$
- resample when the effective number of samples becomes small

$(a, \omega, \ldots)$

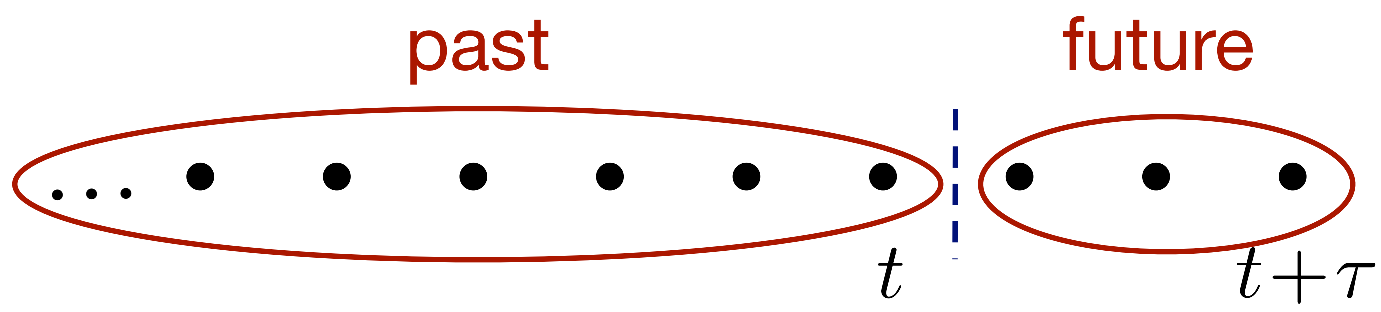


sample control parameters

simulate dynamics

sample kinematics

Bayesian people tracking



Proposals for re-sampling are given by Monte Carlo approximation, $\mathcal{Q}_t = \{\mathbf{s}_t^{(j)}, \hat{w}_t^{(j)}\}_{j=1}^N$, to the windowed smoothing distribution

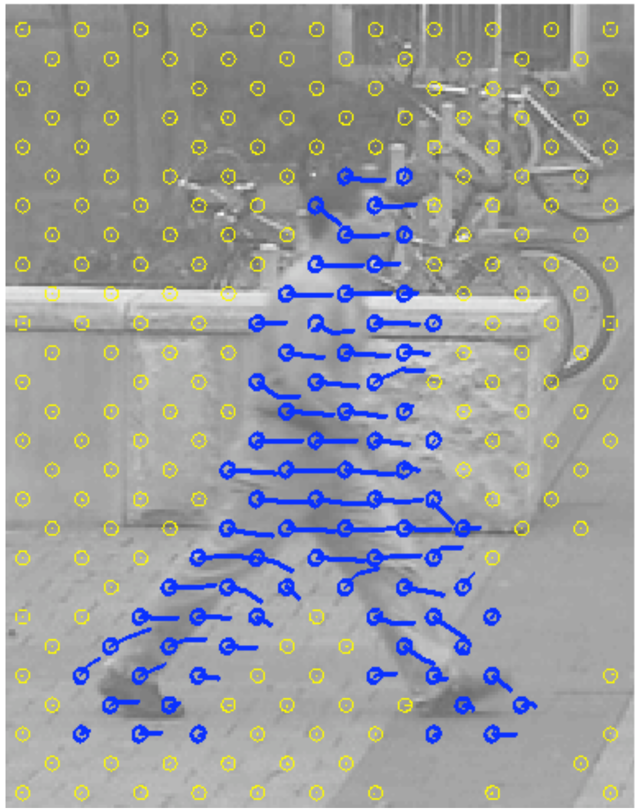
$$p(\mathbf{s}_t \,|\, \mathbf{z}_{1:t+\tau}) \;\propto\; \int_{\mathbf{s}_{t+1,t+\tau}} p(\mathbf{z}_{t:t+\tau} \,|\, \mathbf{s}_{t:t+\tau})\, p(\mathbf{s}_{t:t+\tau} \,|\, \mathbf{z}_{1:t-1})$$

Re-sample $\mathcal{S}_t$ when the effective sample size $[\sum_j (\hat{w}_j^{(j)})^2]^{-1}$ drops below threshold. Then,

- draw sample index $k(i) \sim multinomial\{\hat{w}_t^{(j)}\}_{j=1}^N$
- assign samples and perform importance re-weighting:

$$\mathbf{s}_t^{(k)} \leftarrow \mathbf{s}_t^{(i)} \qquad\qquad w_t^{(k)} \leftarrow w_t^{(i)} / \hat{w}_t^{(i)}$$

Image observations



Foreground model

Gaussian mixture model for color (RGB) of pixels in each part

Background model

mean color (RGB) and luminance gradient $E[\vec{I}(x,y),\ \nabla L(x,y)]$ with covariance matrix

Optical flow

robust regression for translation in local neighborhoods

Calibration and initialization

Camera calibrated with respect to ground plane.

Assume the ground plane orientation in known.

Body position, pose and dynamics coarsely set manually

Speed change

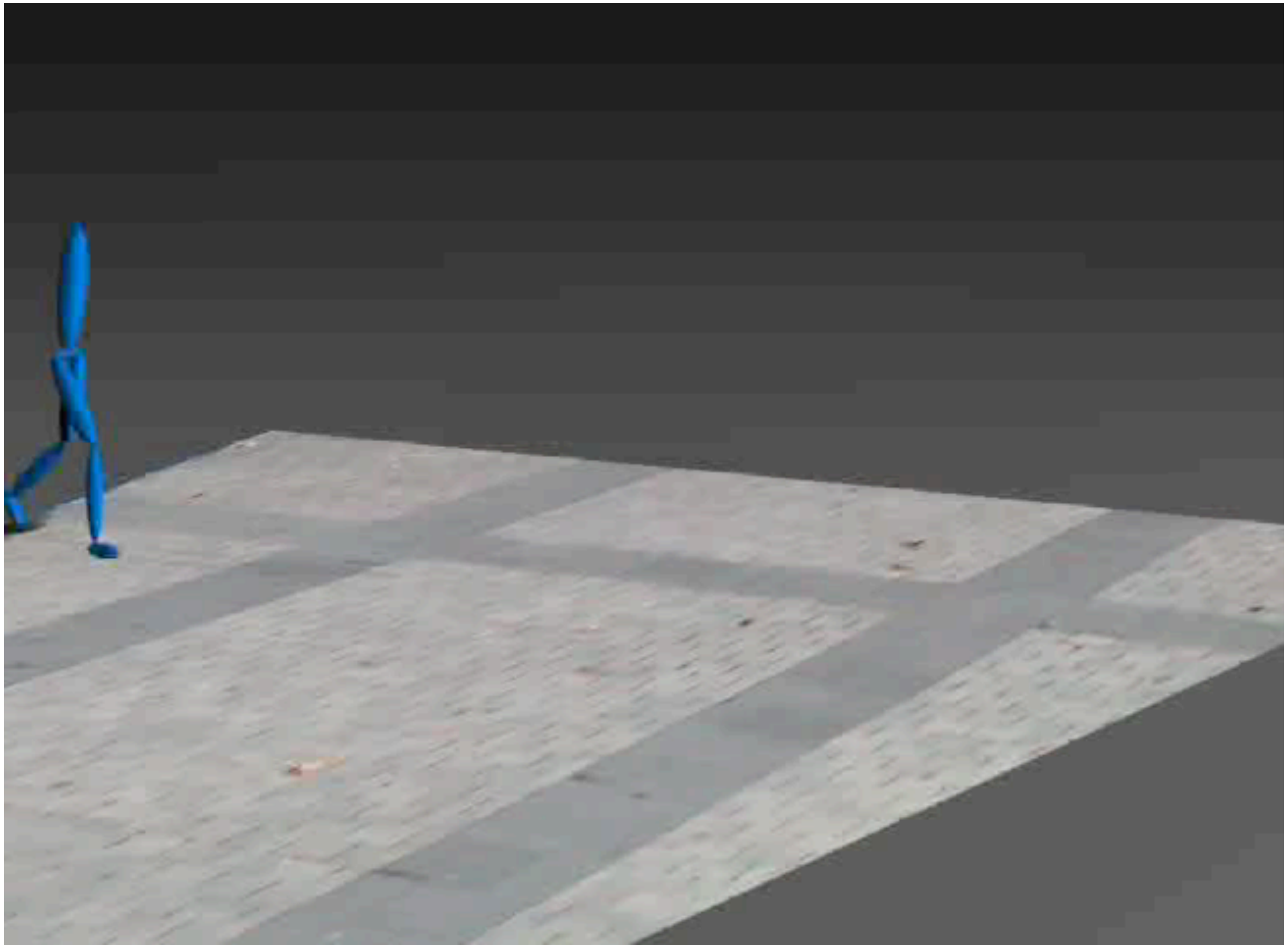
Image observations

negative log background likelihood

Speed change

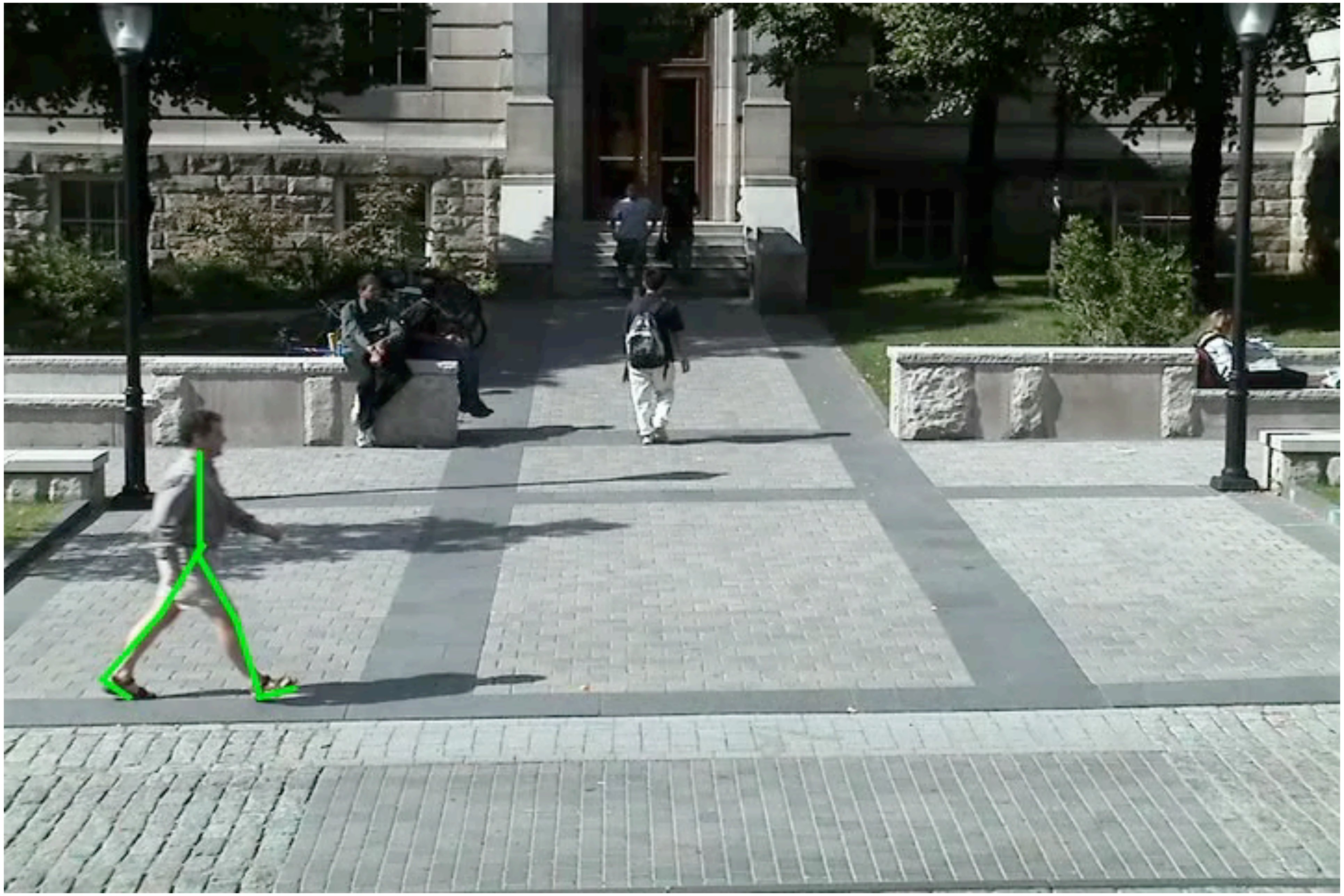
MAP Pose Trajectory (half speed)

Speed change



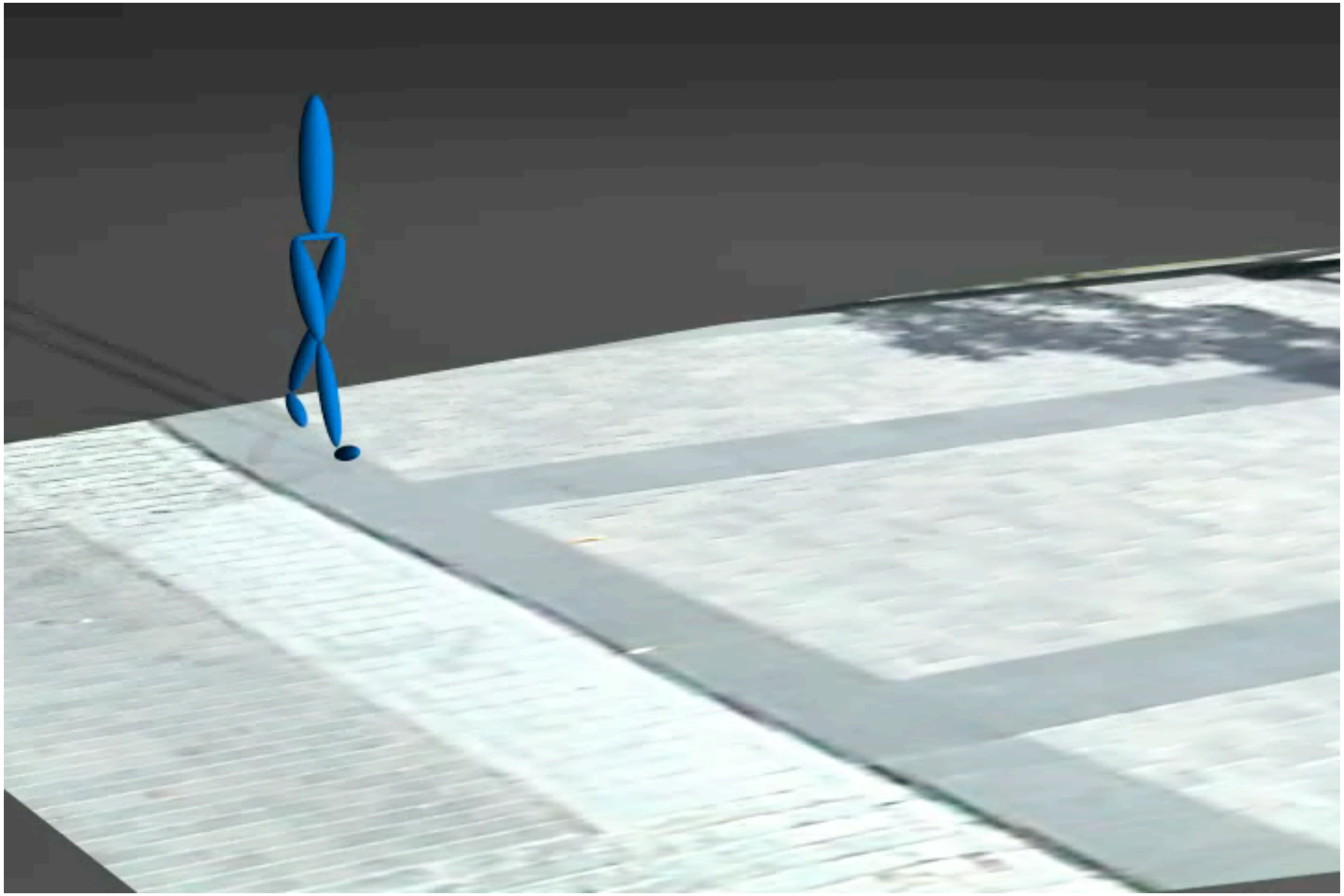
Synthetic rendering of MAP Pose Trajectory (half speed)

Occlusion



MAP Pose Trajectory (half speed)

Occlusion

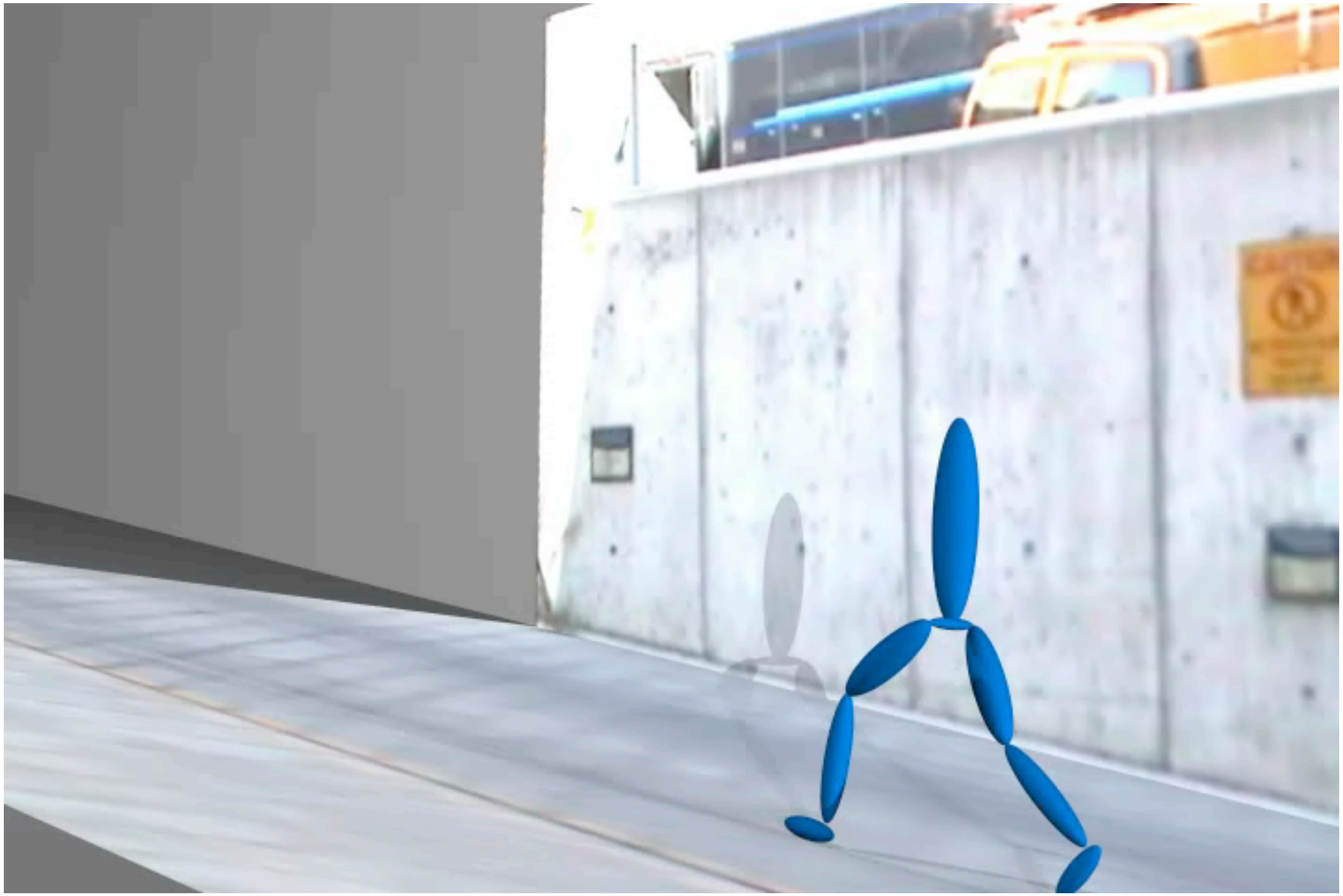


Synthetic rendering of MAP Pose Trajectory (half speed)

Sloped surface (~10º)

MAP Pose Trajectory (half speed)

Sloped surface (~10º)



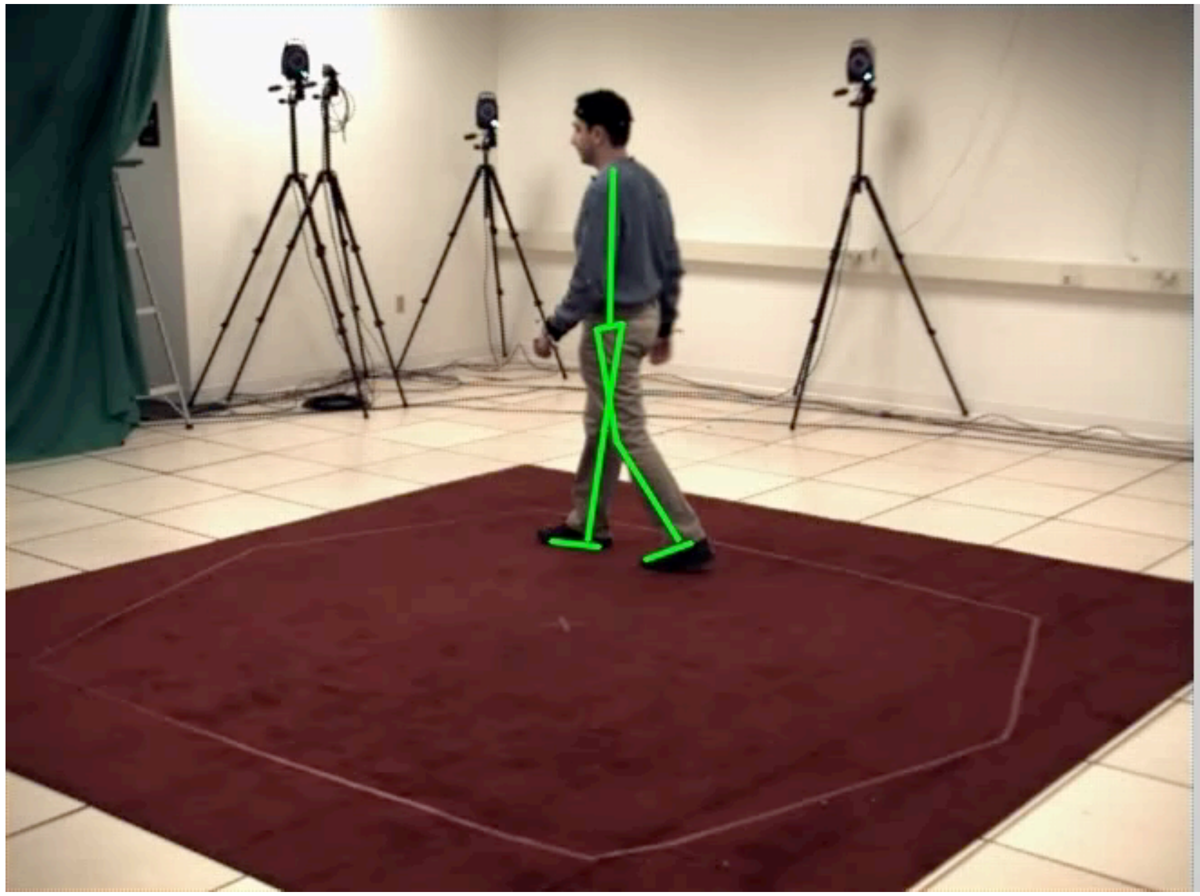
Synthetic rendering of MAP Pose Trajectory (half speed)

Expt 4: HumanEva Data

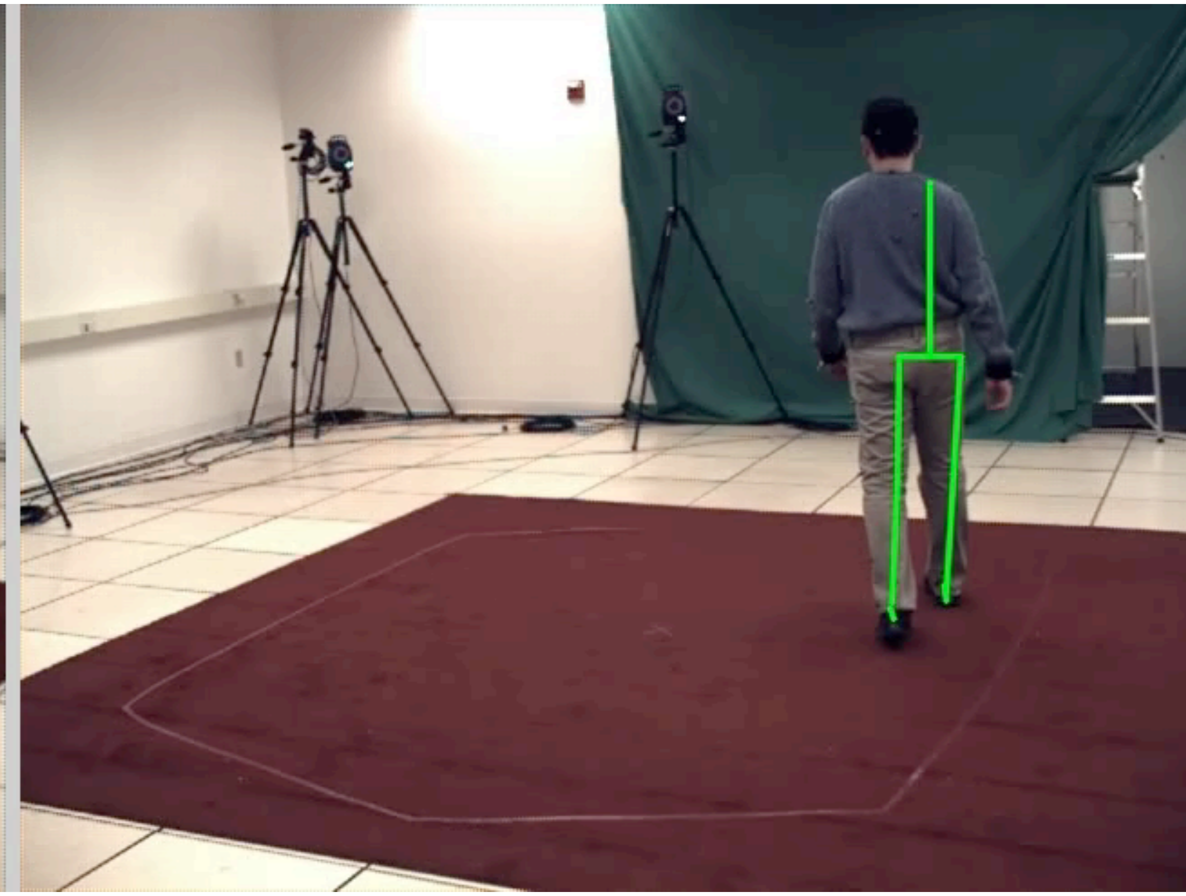
Synchronized motion capture and video

- mocap provides ground truth and data for learned models
- four cameras (for monocular and multiview tracking)
- diverse range of motions
- blind benchmark error reporting

HumanEva Results: Monocular

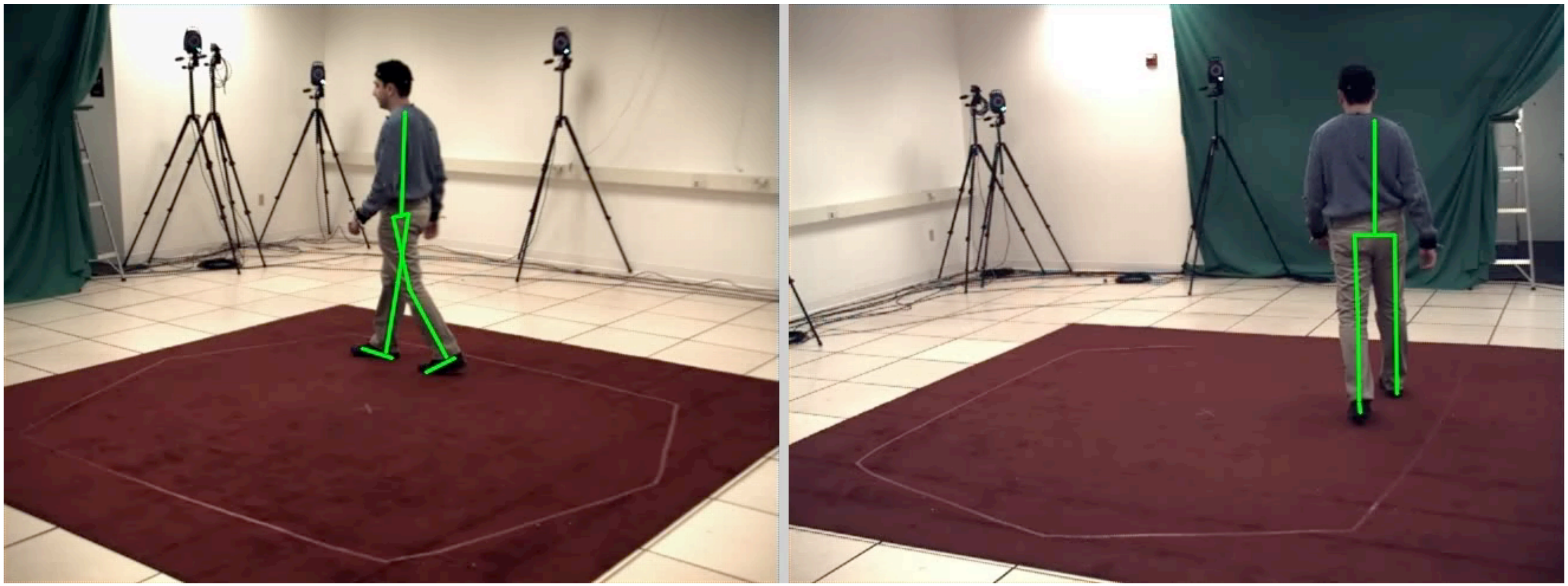


camera 2: Tracking input



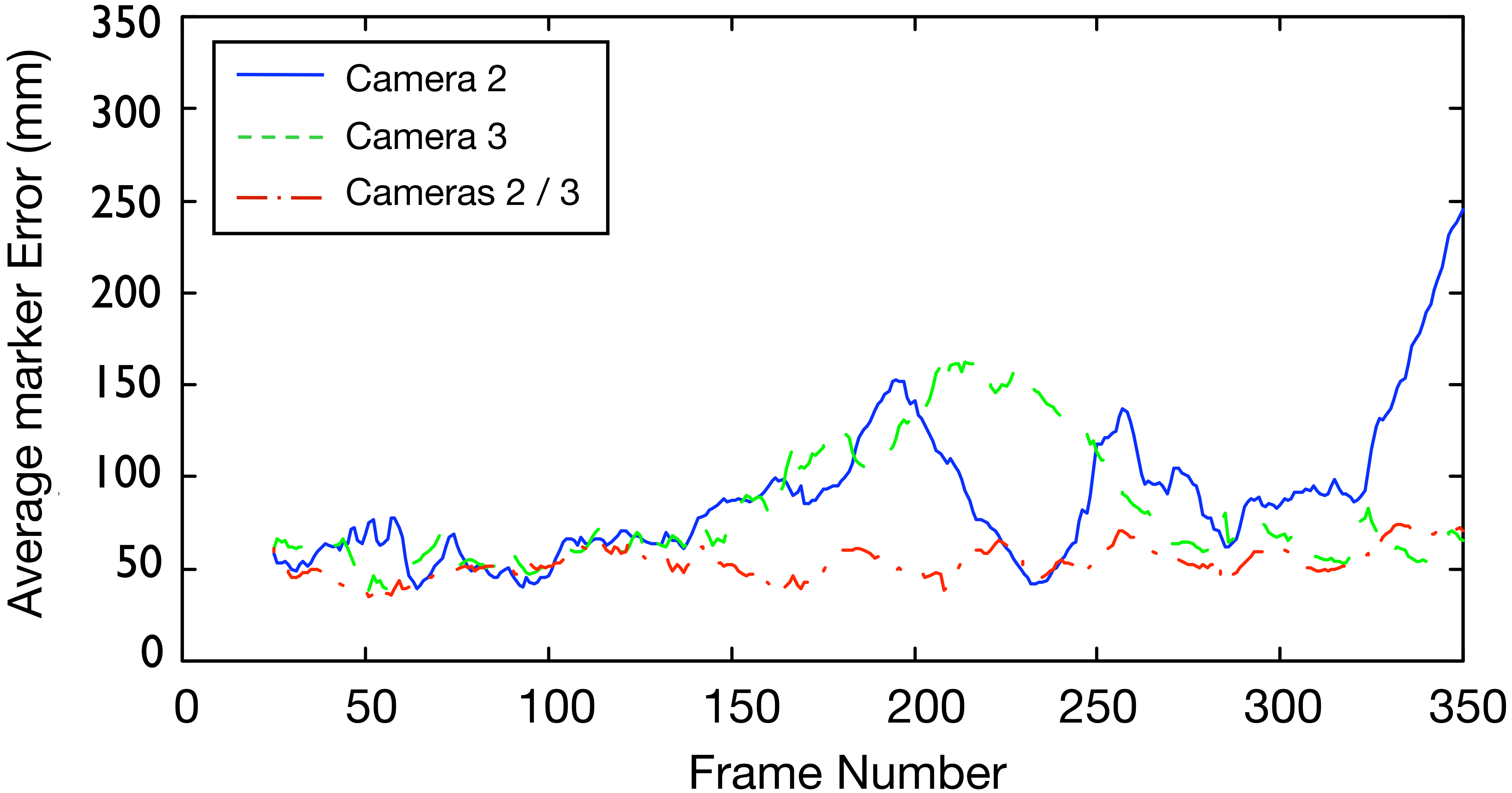
camera 3: Evaluation

HumanEva Results: Binocular



cameras 2 and 3: Binocular tracking input

HumanEva Results



Conclusions

Low-dimensional dynamics capture key physical properties of motion and ground contact. The models

- are stable and simple to control, and
- generalize to a wide range of walking motions

Combined with kinematic models, they provide useful walking models for human pose tracking.

Limitations & Future Work

Our work has just scratched the surface

- Extend locomotion dynamics to capture standing (two-foot contact) and running (no contact during flight phase)
- Learning
  - parameters of physics-based models from mocap
  - conditional kinematics.
- 3D physics-based models of locomotion
- …

Selected references

Brubaker M et al., Physics-based person tracking using the Anthropomorphic Walker. *IJCV*, 2009 (to appear)

Brubaker M and Fleet D. The Kneed Walker for human pose tracking. *Proc IEEE CVPR* 2008.

Brubaker et al., Estimating contact dynamics. *Proc IEEE ICCV,* 2009

Collins S et al. Efficient bipedal robots based on passive-dynamic walkers. *Science* 207(5712), 2005

Collins S and Ruina A. A bipedal walking robot with efficient human-like gait. *Proc ICRA*, 2005

Kuo A. Energetics of actively powered locomotion using the simplest walking model. *J Biomech. Eng*. 124, 2002

Liu K et al. Learning physics-based motion style with nonlinear inverse optimization. *ACM Trans Graphics* 24(3), 2005

McGeer T. Dynamics and control of bipedal locomotion. *J. Theoretical Biol.* 163, 1993