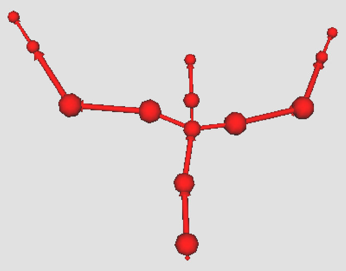
Three Dimensional Monocular Human Motion Analysis in End-Effector Space

S. Hauberg, J. Lapuyade, M. Engell-Nørregård, K. Erleben, and K.S. Pedersen

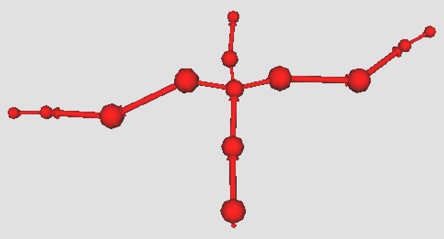
The Image Group, Dept. of Computer Science, University of Copenhagen

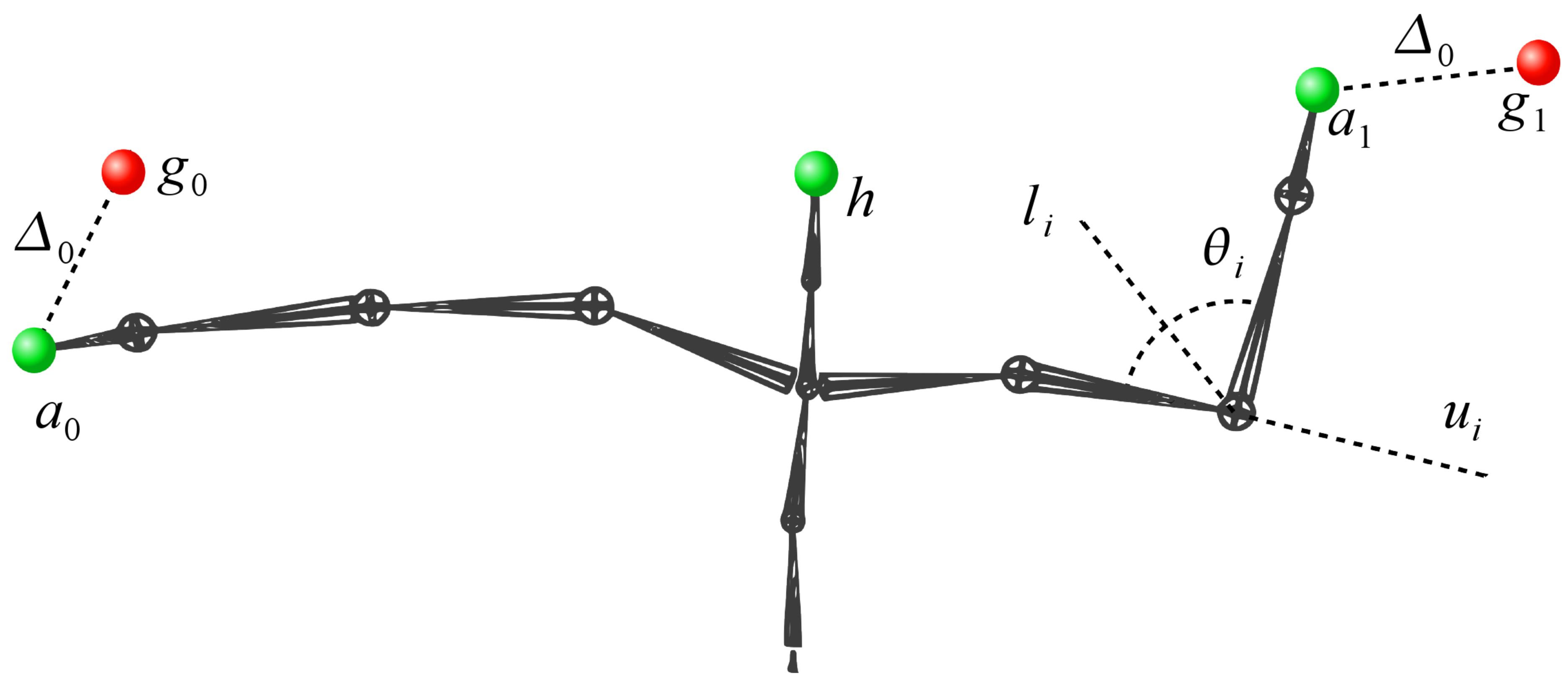
Difficulties of Motion Analysis



Visual human motion gives rise to **multi-modal state distributions.** This gives rise to solutions based on particle filters, which is troublesome as the obvious **state space is high dimensional.** A simple model can easily require more than 50 dof.

Our approach changes the state space to the end-effector space yielding a more efficient system.





Idea: change state space from angles to end-effectors

Joint Angles

A common representation of a skeleton is the set of joint angles. This assumes constant limb sizes.

Pros:

Simple representation

The skeleton is a collection of bones of constant size. Thus, the angles between these is a straight-forward representation.

Cons:

High dimensional

Even for simple skeletons, the degrees of freedom can easily exceed 50. This makes particle filtering and/or motion learning intractable.

Non-standard topology

If constraints on the angles are enforced, this representation is topological equivalent to the unit cube. If constraints are not enforced, the space is circular. In either case, most standard motion models cannot be easily applied as they are designed for real vector spaces.

VS

End-Effectors

Humans tend to plan motion in terms of end-effectors rather than joint angles. Thus, it seems reasonable to represent a pose in terms of end-effectors.

Pros:

Low dimensional

Each end-effector is represented as a 3D spatial coordinate. This makes the state space 3N-dimensional. This can be further reduced by assuming independence between the end-effectors.

Real vector space

The spatial representation makes the real vector space a suitable state space. This allows for straight-forward application of standard motion models.

Cons:

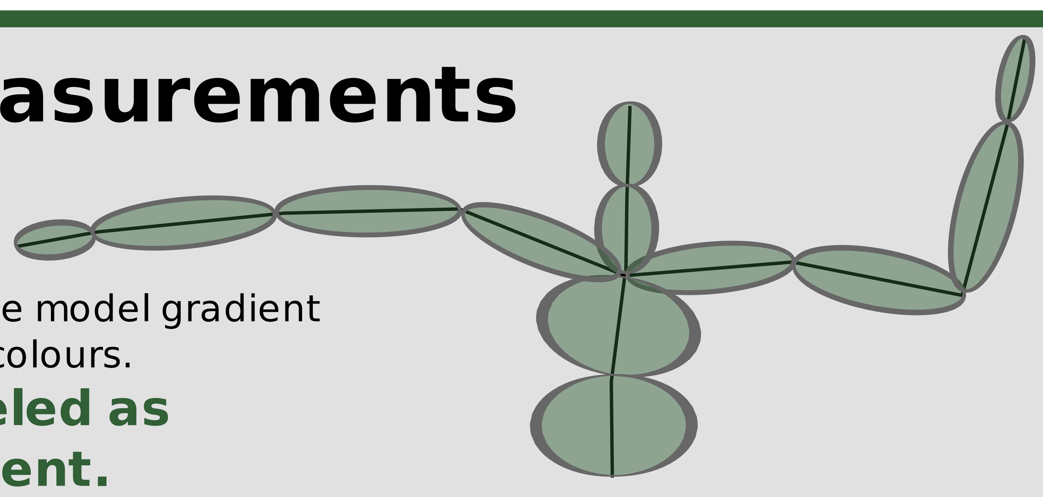
Less exact

For a given configuration of the end-effectors, several angle configurations are possible. This mismatch makes the end-effector representation inherently less exact.

Technical stuff: measurements and inverse kinematics

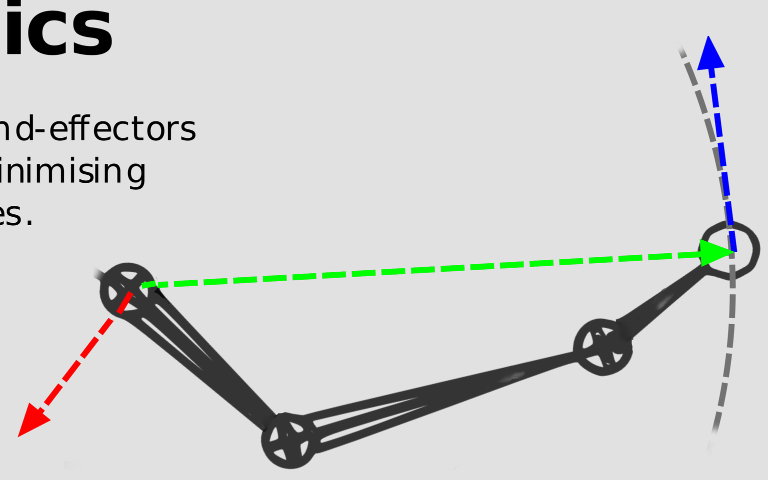
Visual Measurements

We use a simple **MRF texture model** for each limb, where we model gradient orientations and hue colours. **Limbs are modeled as being independent.**



Inverse Kinematics

We compute joint angles, such that the end-effectors attains given positions. This is done by minimising the distance using non-linear least squares. This **allows us to convert end-effectors to angles** . We can, however, not expect to find a unique solution.



Experiments show: we need fewer particles

Experiments

We implemented and compared the state spaces. In both state spaces a **simple linear predictive motion model** was used.

We first determined a suitable number of particles in end-effector space. Then the number of particles in angle space given rise to the same run-time was determined. Finally, we found the number of particles in angle space needed to give similar results to the end-effector tracker.

The conclusion was that the end-effector state space yielded a much more efficient tracker with a slight loss in accuracy.

Angles
100 particles
Run-time: 5 min.

Angles
5000 particles
Run-time: 10 h.

End-effectors
25 particles
Run-time: 5 min.