LIE ALGEBRA-BASED KINEMATIC PRIOR FOR 3D HUMAN POSE TRACKING

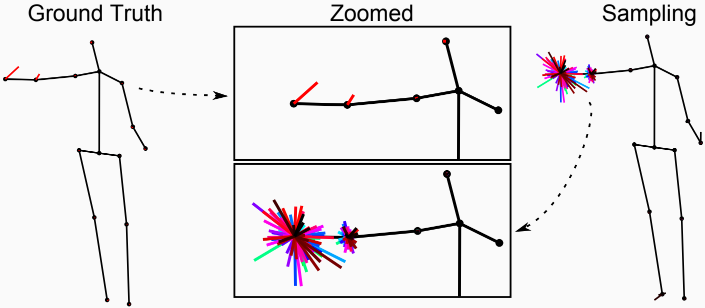
Edgar Simo-Serra, Carme Torras, Francesc Moreno-Noguer

Tokyo. 21st of May, 2015

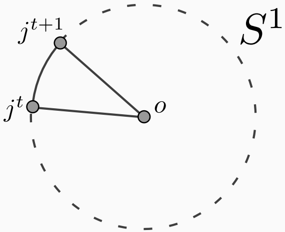
Institut de Robòtica i Informàtica Industrial
CSIC
UPC

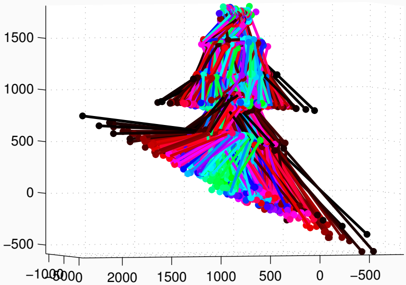
CONTENTS

- Predict change in position (velocity) in subsequent frames given the current position
- Application to 3D pose tracking as a prior



- Predict change in position (velocity) in subsequent frames given the current position
- Application to 3D pose tracking as a prior
- Only generate feasible configurations constrained by the manifold





Prior	Complexity	Scales	Consistent
Gaussian diffusion	Low	**Yes**	No
GPLVM [1]	Low	No	No
GPDM [2]	Medium	No	No
hGPLVM [3]	Medium	No	No
CRBM [4]	High	**Yes**	No
GCMFA [5]	High	No	No
GFMM (Ours)	Low	**Yes**	**Yes**

[1] N. D. Lawrence. Probabilistic Non-linear Principal Component Analysis with Gaussian Process Latent Variable Models. JMLR, 6:1783–1816, 2005.

[2] J. Wang, D. Fleet, and A. Hertzmann. Gaussian process dynamical models. In NIPS, 2005.

[3] M. Andriluka, S. Roth, and B. Schiele. Monocular 3D Pose Estimation and Tracking by Detection. In CVPR, 2010.

[4] G. Taylor, L. Sigal, D. Fleet, and G. Hinton. Dynamical binary latent variable models for 3d human pose tracking. In CVPR, 2010.

[5] R. Li, T.-P. Tian, S. Sclaroff, and M.-H. Yang. 3d human motion tracking with a coordinated mixture of factor analyzers. IJCV, 87(1-2):170–190, 2010.

1. Consider data to lay on a joint pose and kinematic manifold

$$(x, v) \in \mathrm{SO}(3)/\mathrm{SO}(2) \times \mathfrak{so}(3)/\mathfrak{so}(2)$$

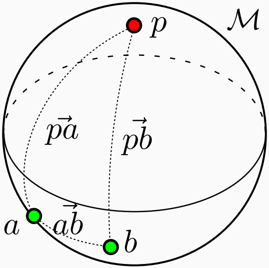
2. Learn joint probabilistic generative parametric model

$$p(x, v|\theta)$$

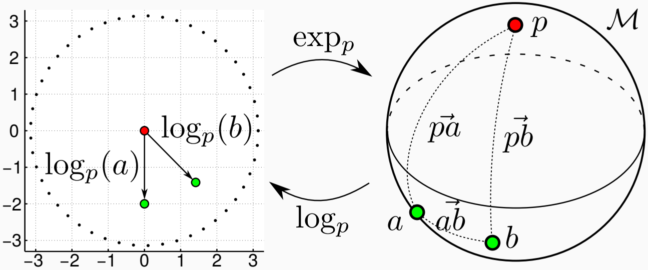
3. Infer kinematics from pose

$$p(v|x, \theta)$$

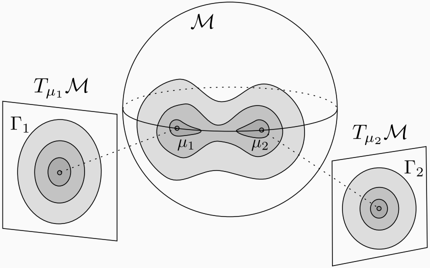
- **Geodesic distance** between two points on a manifold is the shortest distance along the manifold



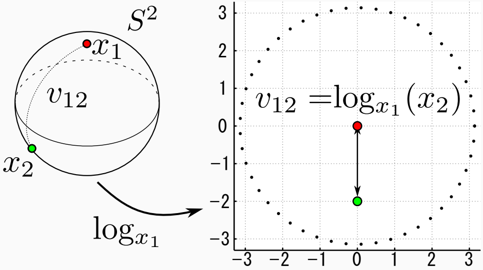
- **Geodesic distance** between two points on a manifold is the shortest distance along the manifold
- **Tangent space** is a local approximation of a manifold that is a **Euclidean space**
    - logarithm and exponential map project to and from a tangent space respectively



- Mean estimated on the manifold using the **geodesic mean**
- Covariance estimated on the **tangent space** in closed form
- Expectation-Maximization algorithm
    - **Minimum Message Length** used to determine number of clusters



- Pose modelled using SO(3)/SO(2) joints
- Quotient of Lie algebras $\mathfrak{so}(3)/\mathfrak{so}(2)$ expresses velocity of a SO(3)/SO(2) joint
  - Equivalent to tangent space
  - Velocities are geodesic lines
- Joint pose and kinematic modelled as $SO(3)/SO(2) \times \mathfrak{so}(3)/\mathfrak{so}(2)$



- Learn joint distribution of poses and kinematics $p(x, v|\theta)$
  - θ are the mixture parameters
  - Number of clusters K determined automatically

$$p(x, v|\theta) = \sum_{k=1}^{K} \alpha_k p(x, v|\theta_k)$$

- Compute conditional distribution $p(v|x, \theta)$
  - Conditional distribution is a new mixture model
  - Cluster weights re-estimated given x

$$p(v|x, \theta) = \frac{p(x, v|\theta)}{p(x|\theta)} = \frac{\sum_{k=1}^{K} \alpha_k p(x|\theta_k) p(v|x, \theta_k)}{\sum_{k=1}^{K} \alpha_k p(x|\theta_k)} = \frac{1}{Z} \sum_{k=1}^{K} \pi_k p(v|x, \theta_k)$$

- Sampling is $\mathcal{O}(1)$, computing log-likelihood is $\mathcal{O}(K)$
  - 10^5 samples in under a second

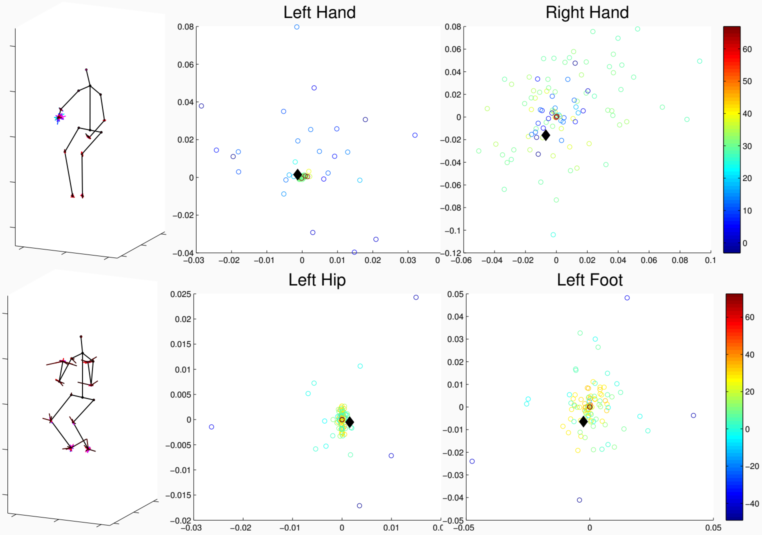
RESULTS

- Evaluation on Human3.6m dataset [1]
- 15 different actions with 2 subactions each
- 6 actors for training, 1 actor for testing
- Model body with 15 joints
    - 12 joints have 2 DoF, 2 joints have 1 DoF
    - Learn block-diagonal covariance matrices with 92 parameters each
- Scale pose and kinematic components to be relatively similar
- Subsample heavily correlated input data when learning

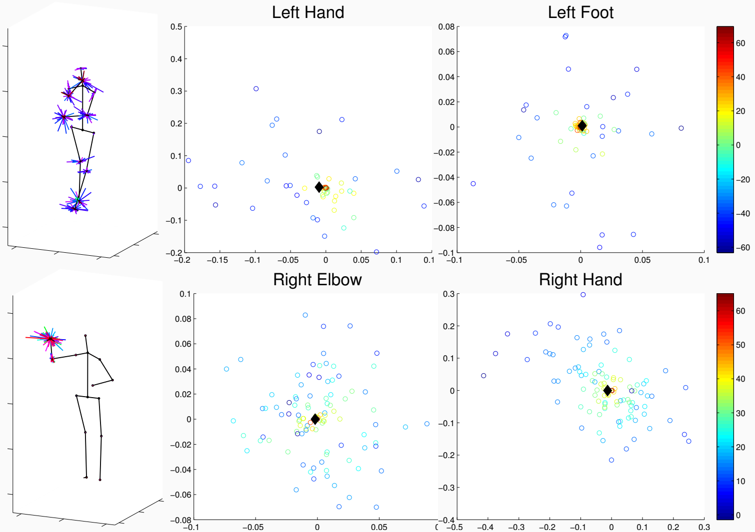
[1] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu. Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. PAMI, 36(7):1325-1339, 2014.

Method	Log-likelihood Train	Test
Samples	465,325	62,064
Gaussian diffusion	5.4325	5.4349
local Gaussian diffusion	6.4193	6.4206
Ours (30%, 211 clusters)	9.3382	11.7874
Ours (15%, 147 clusters)	8.9544	11.8714

RESULTS



RESULTS



- Robust kinematic prior
- Demonstrated performance over widely used gaussian priors
- Code for the GFMM framework is available [1]

[1] http://www.iri.upc.edu/people/esimo/

- Robust kinematic prior
- Demonstrated performance over widely used gaussian priors
- Code for the GFMM framework is available [1]
- Use in real world tracking framework
- Extend to more conditionals $p(v_t|x_t, x_{t-1}, \theta)$

[1] http://www.iri.upc.edu/people/esimo/

Questions?

http://www.iri.upc.edu/people/esimo/