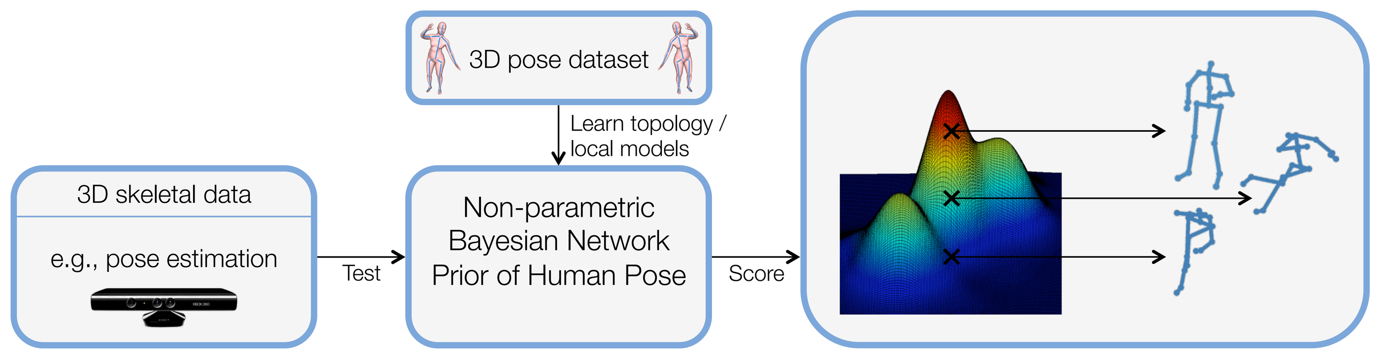
A Non-parametric Bayesian Network Prior of Human Pose

Andreas M. Lehrmann[1], Peter V. Gehler[1], Sebastian Nowozin[2]
MPI for Intelligent Systems[1], Microsoft Research Cambridge[2]

1 Overview

- We propose a general-purpose Bayesian network prior of human pose.



- **Fully non-parametric:** Estimation of both optimal information-theoretic topology and local conditional distributions from data.
- **Compositional:** Effective handling of the combinatorial explosion of articulated objects, thereby improving generalization.
- **Superior performance:** Better data representation than traditional global models and parametric networks on the large Human 3.6M dataset.
- **Real-time:** Fast and accurate computation of approximate likelihoods on datasets with up to 100k training poses.

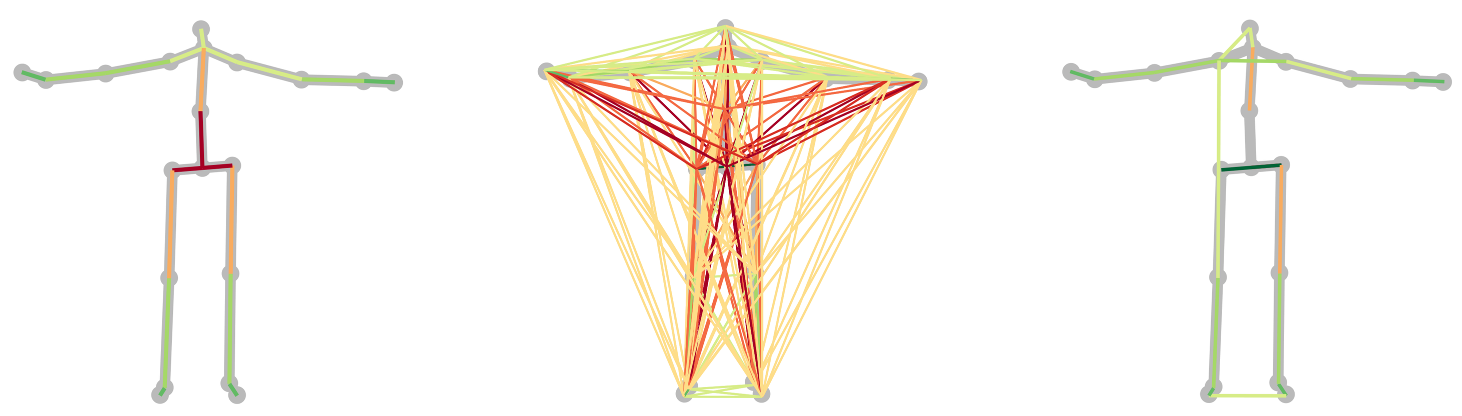
2 Non-parametric Networks

Learn a sparse and non-parametric Bayesian network $B = (p, \mathcal{G}(V, E))$.

- **Learning the graph structure:**
  Minimize KL-divergence between the high-dimensional pose distribution $q(\mathbf{X})$ and the tree-structured network $p(\mathbf{X}) = \prod_{j=1}^{|V|} p\left(X_j | X_{\mathrm{pa}(j)}\right)$,

$$\mathcal{G} := \operatorname*{argmin}_{pa} \mathrm{KL}\left(q(\mathbf{X}) \parallel p(\mathbf{X})\right) = \mathrm{MST}(\mathcal{G}'),$$

  where $\mathcal{G}'$ is the complete graph with edge weights $e_{jk} = \widehat{\mathrm{MI}}(X_j, X_k)$.



Kinematic chain — Mutual information — Chow-Liu tree

- **Learning the conditional distributions:**
  We use a conditional kernel density estimate (CKDE) to learn the local models of the inferred tree,

$$p\left(X_j | X_{\mathrm{pa}(j)}\right) = \frac{p\left(X_j, X_{\mathrm{pa}(j)}\right)}{\int_{X_j} p\left(X_j, X_{\mathrm{pa}(j)}\right) \mathrm{d}X_j} = \frac{\sum_i \mathcal{N}\left((X_j, X_{\mathrm{pa}(j)}) \,\middle|\, (X_j^{(i)}, X_{\mathrm{pa}(j)}^{(i)}), BB^\top\right)}{\sum_i \mathcal{N}\left(X_{\mathrm{pa}(j)} \,\middle|\, X_{\mathrm{pa}(j)}^{(i)}, (BB^\top)|_{X_{\mathrm{pa}(j)}}\right)},$$

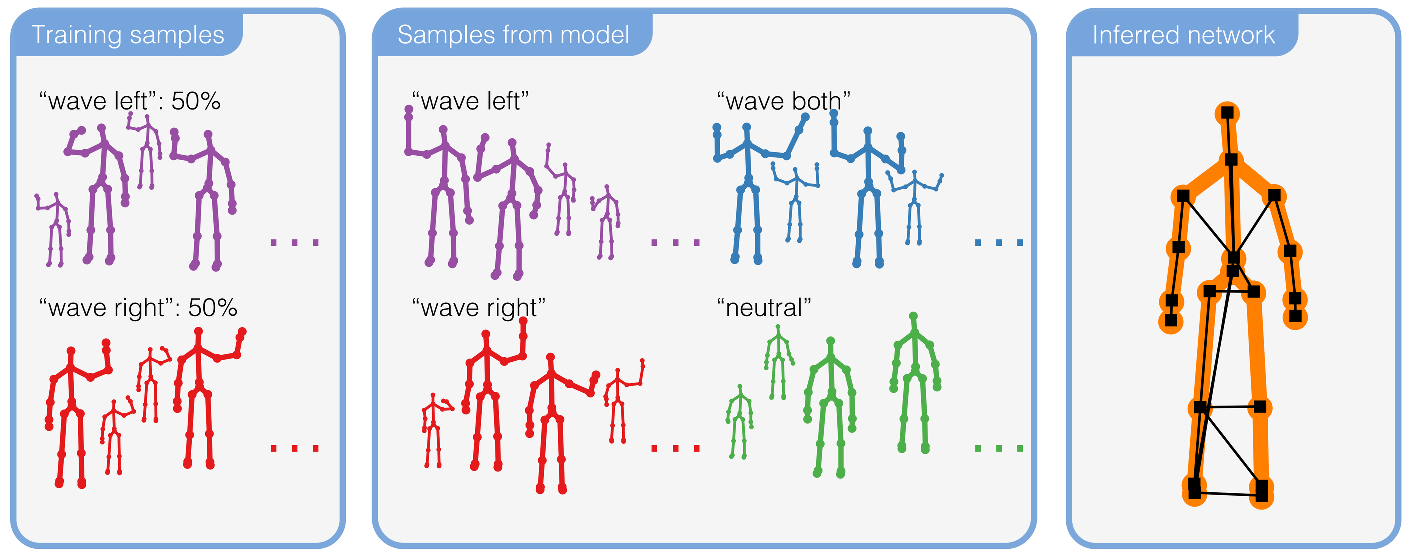
  where $p\left(X_j, X_{\mathrm{pa}(j)}\right)$ is an unconditional KDE with isotropic Gaussian kernel and bandwidth B proportional to the square root of the covariance.

- **Important operations are efficient:**
  - Computation of a log-likelihood requires $\mathcal{O}(|V|)$ KDE evaluations.
  - Ancestral sampling requires $\mathcal{O}(|V|)$ samples from the local models. [Gaussian mixture models with non-uniform weight distribution]

3 Compositionality & Generalization

- Our formulation allows to freely combine substructures, but only if they do not share a lot of information.

  $\Longrightarrow$ Compositionality exactly where needed and only where appropriate.



- We compute expected log-likelihoods for our Chow-Liu/CKDE model and several baselines on the Human 3.6M dataset.

Table 1: Expected log-likelihoods.

Method	Graph structure	Training	Testing
Gaussian	Global	−266.84	−271.15
KDE	Global	−239.61	−263.77
GPLVM*	Global	−327.85	−341.89
Gaussian linear network	Independent	−352.80	−345.94
	Kinematic chain (order 1)	−311.54	−310.98
	Kinematic chain (order 2)	−305.54	−307.88
	Chow-Liu tree	−283.82	−284.03
CKDE network	Independent	−322.64	−322.25
	Kinematic chain (order 1)	−260.04	−270.52
	Kinematic chain (order 2)	−247.35	−263.83
	Chow-Liu tree **(ours)**	−242.24	**−254.98**

*25% subsampling; FITC

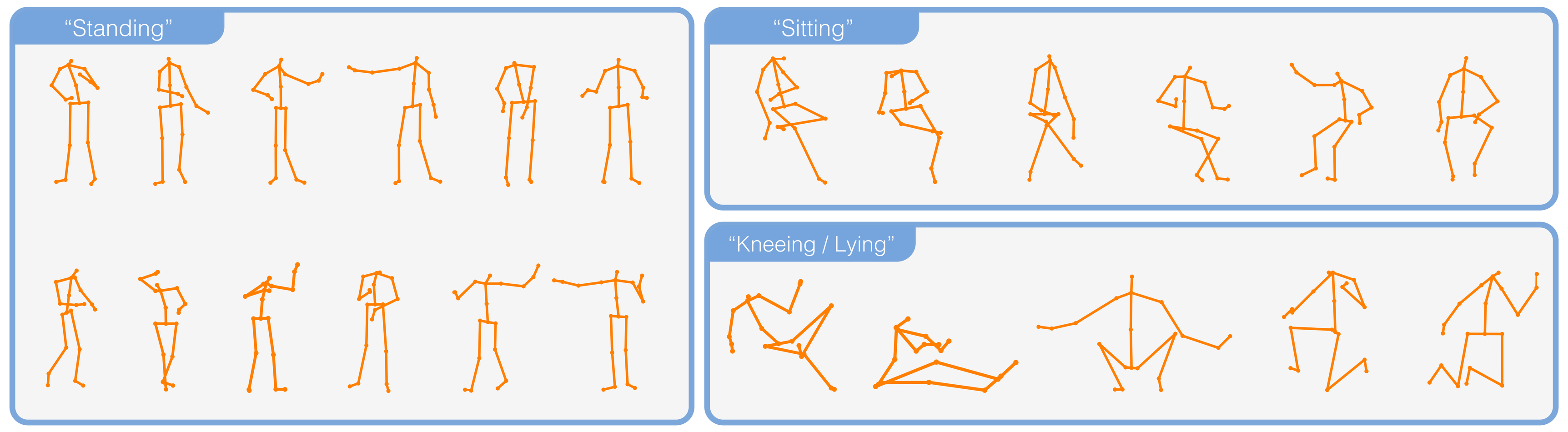


Figure 1: Samples drawn from a single Chow-Liu/CKDE model.

4 Live Scoring

- Applications in real-time environments require additional speed.
- **Training:** Cluster the training points into clusters $\{C^{(i)}\}_i$ using k-means and build a kd-tree for their centres.
- **Testing:** Given a test pose $\mathbf{x}$, use the kd-tree to compute a k-NN partitioning $\{C^{(i)}\}_i = C_e(\mathbf{x}) \uplus C_a(\mathbf{x})$ and approximate the likelihood as

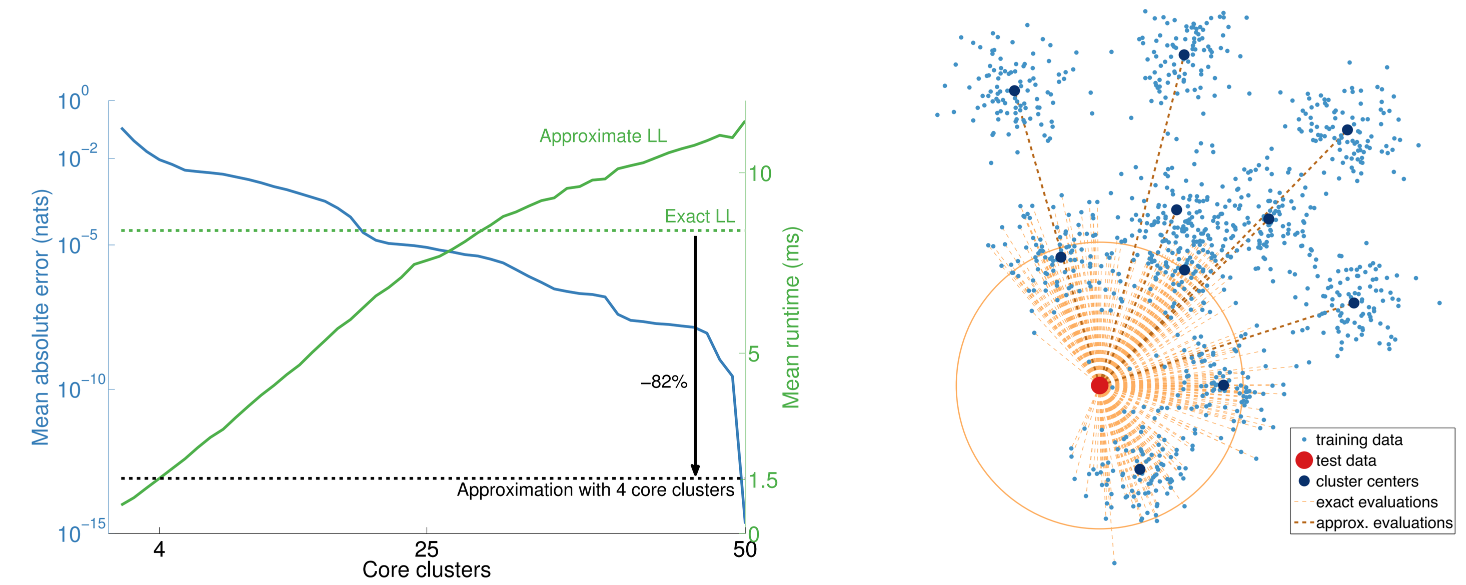
$$p(\mathbf{x}) \approx (S_e + S_a) \big/ (N \cdot \det(B)),$$

  with

$$S_e = \sum_{C \in C_e} \sum_{j \in C} \kappa\left(B^{-1}\left(\mathbf{x} - \mathbf{x}^{(j)}\right)\right), \qquad \text{[exact]}$$

$$S_a = \sum_{C \in C_a} |C| \kappa\left(B^{-1}\left(\mathbf{x} - \overline{C}\right)\right), \qquad \text{[approx.]}$$

  where $\overline{C}$ and $|C|$ denote the centre and size of cluster C, respectively.



References

[1] C. Chow and C. Liu. Approximating discrete probability distributions with dependence trees. IEEE Transactions on Information Theory, 1968.
[2] A. Gray and A. Moore. Nonparametric density estimation: Toward computational tractability. SIAM International Conference on Data Mining, 2003.
[3] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu. Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments. Technical report, University of Bonn, 2012.

Visit Us