3D Human Motion Synthesis with Recurrent Gaussian Processes

Yeziwei Wang

University of Cambridge, Amazon

Introduction

- Human motion modelling is an active problem in fields such as people tracking in computer vision, and motion synthesis for computer graphics.
- Recurrent Gaussian Processes (RGP) are good at propagate long-term uncertainty, which is non-trivial in human motions generation.
- The aim of the project is to explore scalable RGPs to produce natural 3D human motions.

Recurrent Gaussian Processes(RGP)

A Gaussian Process (GP) is a distribution over functions and it can be specified by a mean function and covariance function,

$$f(x) \sim \mathcal{GP}\big(m(x), k(x, x')\big) \quad (1)$$

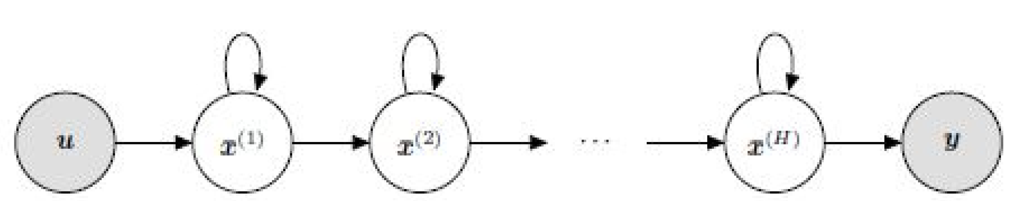


Fig. 1: RGP with H hidden layers

$$x_i = f(\bar{x}_{i-1}, \bar{u}_{i-1}) + \epsilon_i^x; \quad y_i = g(\bar{x}_i) + \epsilon_i^y \quad (2)$$

$$\bar{x}_i = [x_i, ...x_{i-L+1}]^\top, \quad \bar{u}_i = [u_i, ...u_{i-L_u+1}]^\top \quad (3)$$

where x is hidden layer, u is control signal and y is observation sequence. In a deep RGP, the mappings between each layer, as well as through recurrent layers, are GPs[1].

$$f \sim \mathcal{N}(0, K_f) \quad g \sim \mathcal{N}(0, K_g) \quad (4)$$

Considering a RGP with only one hidden layer x.

$$p(y, x|u) = p(y|x)p(x|u) \quad (5)$$

Apply variational inference

$$\log p(y) \geq \int_x Q \log\left[\frac{p(y|x)p(x|u)}{Q}\right] \quad (6)$$

$$Q = q(x) \quad (7)$$

M inducing points Z are introduced to represent each GP.

Implementation

Fully observed models are built to predict 3D walking and running motions.

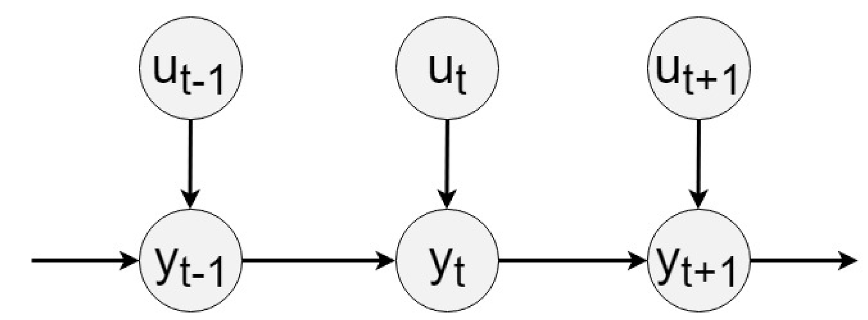


Fig. 2: Fully observed RGP

- The delta values are used as control signals.

$$\mathbf{u} = \Delta = [y_2 - y_1, ..., y_t - y_{t-1}, ...y_N - y_{N-1}, 0] \quad (8)$$

- The window size for both control signal and observation sequence is 20 frames.

$$\bar{x}_i = [x_i, ...x_{i-19}]^\top, \quad \bar{u}_i = [u_i, ...u_{i-19}]^\top \quad (9)$$

- 29 joints are treated as independent models,

$$p(\mathbf{y}) = p(y^1)p(y^2)...p(y^{29}) \quad (10)$$

- For prediction, the initial 20 frames of the test sequence are fed to the trained model for motion generation.

Preliminary Results

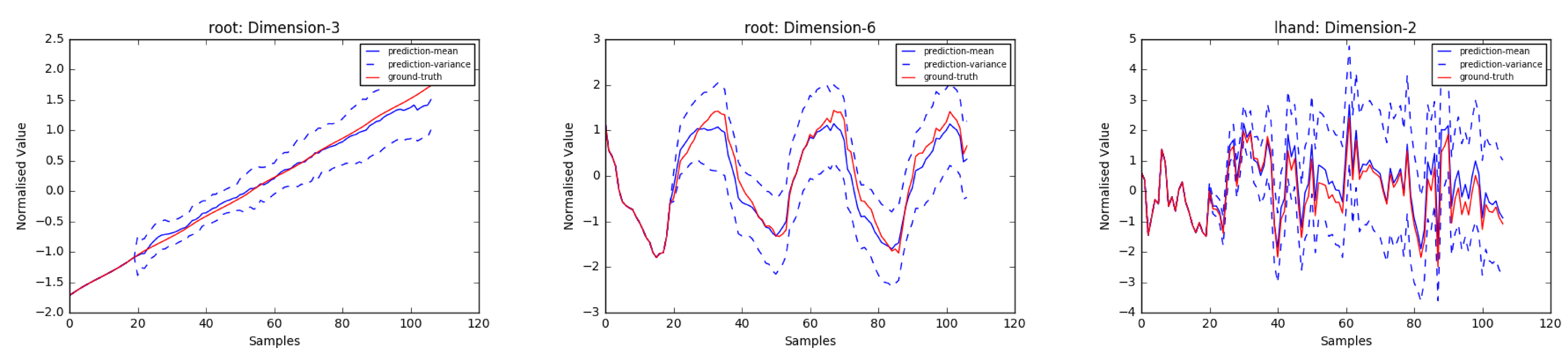


Fig. 3: Prediction of single dimension

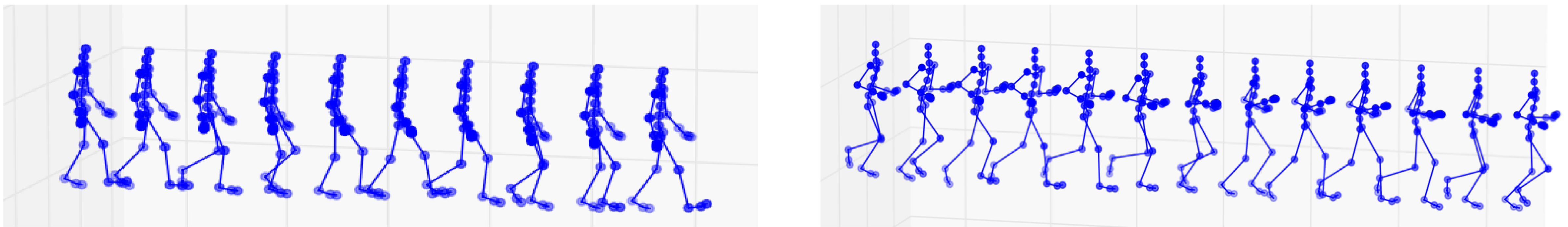


Fig. 4: Original motions

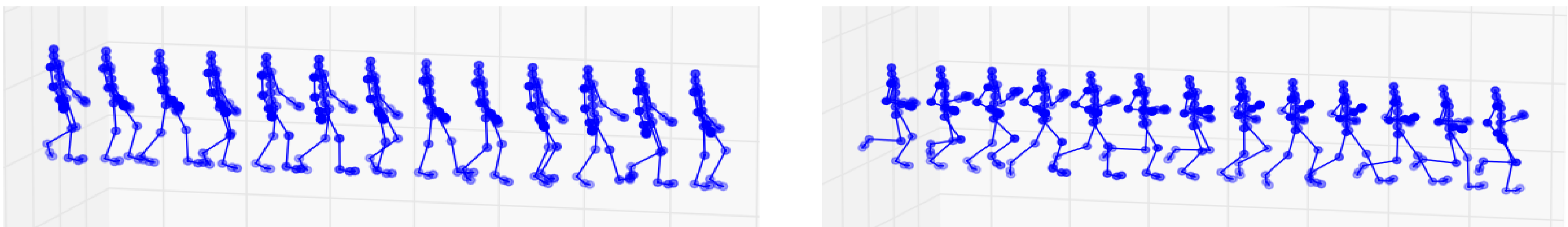
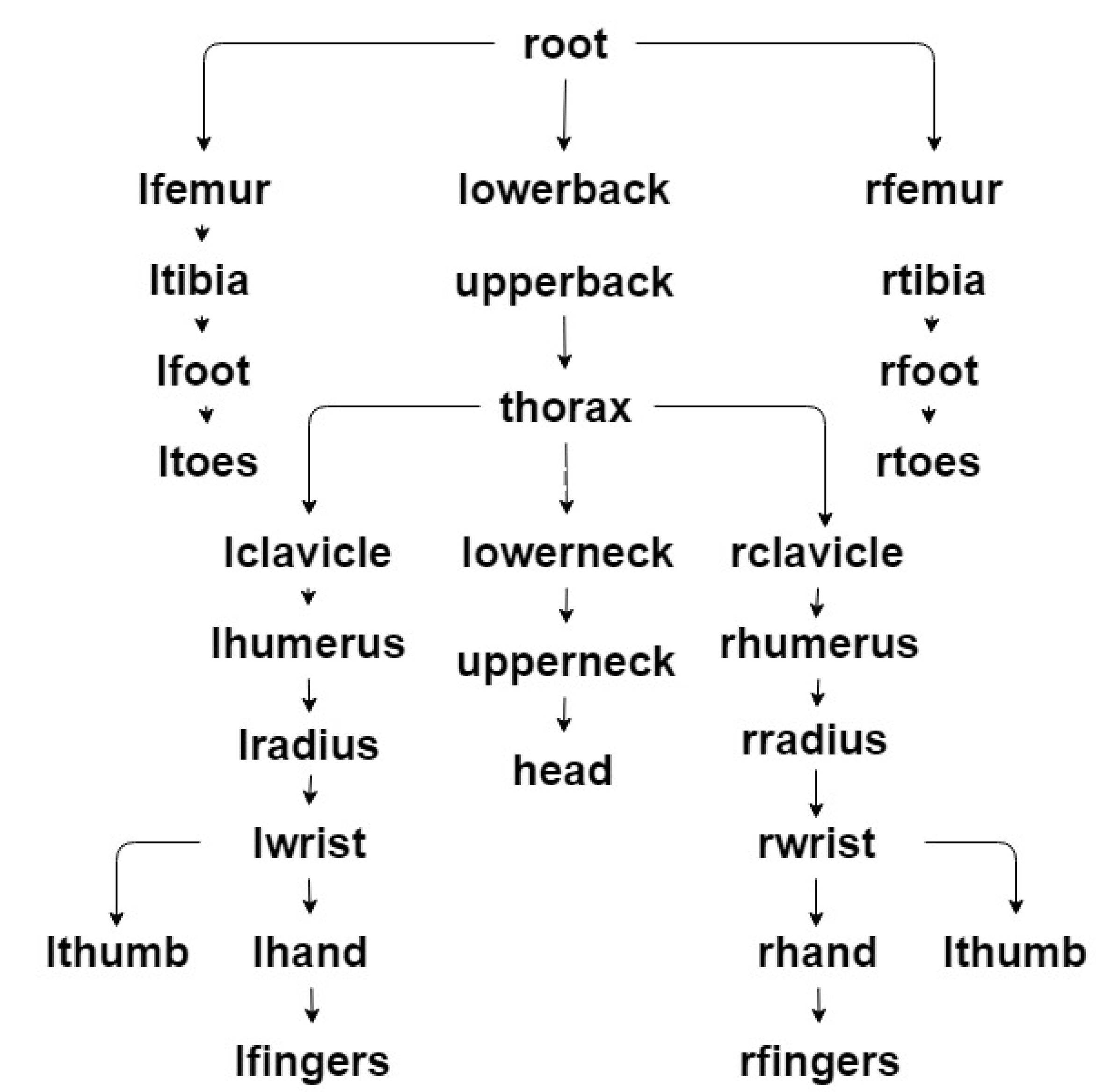


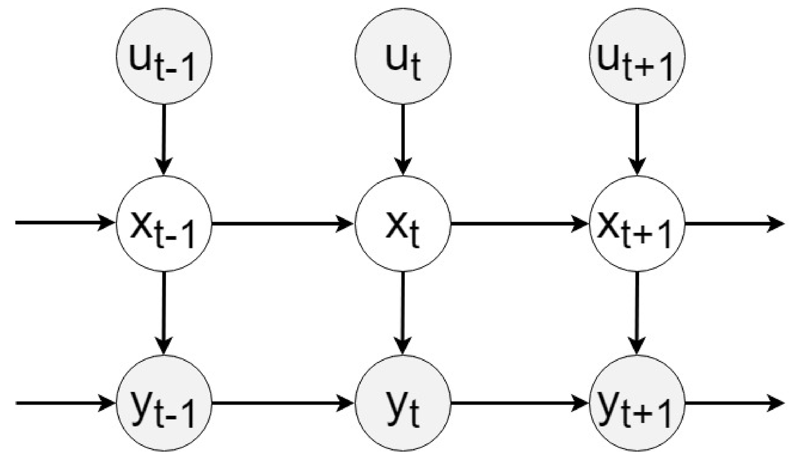
Fig. 5: Synthesised motions

Future Experiments

1. Introduce dependencies between joints. In mocap dataset, each joint is dependent on its parent, for example **lowerback** is dependent on **root**.



2. Introduce hidden layers to the model.



3. Different types of control signals will be further investigated for improving model performance.
4. Different motions will be explored to test how well the model generalises.

References

[1] César Lincoln C. Mattos, Zhenwen Dai, Andreas Damianou, Jeremy Forth, Guilherme A. Barreto, and Neil D. Lawrence. Recurrent Gaussian processes. In Hugo Larochelle, Brian Kingsbury, and Samy Bengio, editors, *Proceedings of the International Conference on Learning Representations*, volume 3, 2016.