Dynamical Modeling of Clothing Articles using GP-LVM

Nishanth Koganti
Mathematical Informatics Laboratory

Nara Institute of Science and Technology

August 6, 2014

Contents

Table of Contents

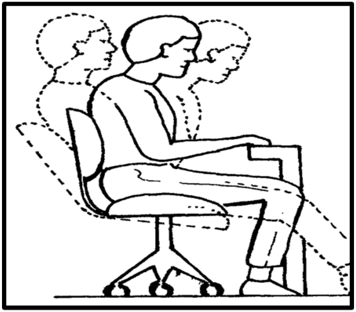
Robotic Clothing Assistance

- Clothing Assistance: Necessity in the daily life of the elderly and disabled people.

- Challenging problem involving close interaction of the robot *with non-rigid clothing material and assisted person with varying posture*.

Non-rigid clothing material

Varying posture of assisted person



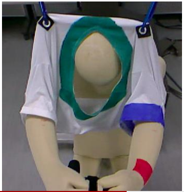
Estimation of Human-Cloth Relationship

Accurate estimation of human-cloth relationship is crucial to ensure efficient learning of motor skills.

We have previously shown that low-cost depth sensor can be used for real-time estimation with key features:

- Use of low-dimensional representation for human-cloth relationship (Topology Coordinates)
- Use of compact representation for clothing articles (Ellipse approximation)

Color Markers used on T-shirt for robust detection



Kinect Depth Sensor

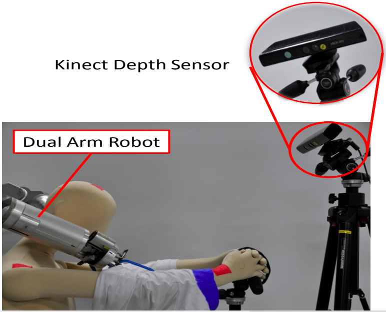


Table of Contents

Problem Statement

Problems with previous framework

Accuracy of human-cloth relationship estimation reduces under:

- Occlusion of T-shirt collar from mannequin
- Measurement noise in depth sensor

Proposed Solution

Hypothesis: Clothing material follows consistent dynamics during clothing tasks with minor variations depending on environment

Proposal: Extract low-dimensional dynamics model from ground-truth data and use model for constrained tracking of human-cloth relationship

Table of Contents

What is a Gaussian Process ($\mathcal{GP}$)?

Gaussian Process is a generalization of multivariate gaussian distribution to infinitely many variables.

A Gaussian distribution is fully specified by its mean vector ($\boldsymbol{\mu}$) and covariance matrix ($\boldsymbol{\Sigma}$):

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \ \mathbf{x} = [x_1 \ x_2 \ \cdots \ x_N] \tag{1}$$

A Gaussian process is fully specified by its mean function ($m(x)$) and covariance function ($k(x, x')$):

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \tag{2}$$

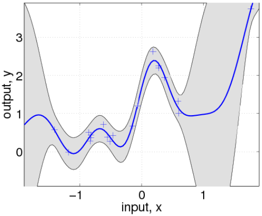
How is $\mathcal{GP}$ useful?

$\mathcal{GP}$ can be used as a prior over functions forming non-parametric model.

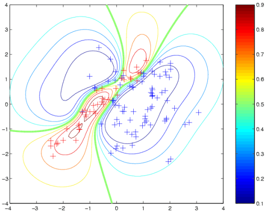
	Mapping	Prior
Parametric	$f(\mathbf{X}, \theta) = \mathbf{WX}$	On function parameters $p(\theta)$
Non-parametric	$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$	On function itself $p(f)$

Posterior/Predictive distribution will also be a $\mathcal{GP}$.

$\mathcal{GP}$ Regression



$\mathcal{GP}$ Classification



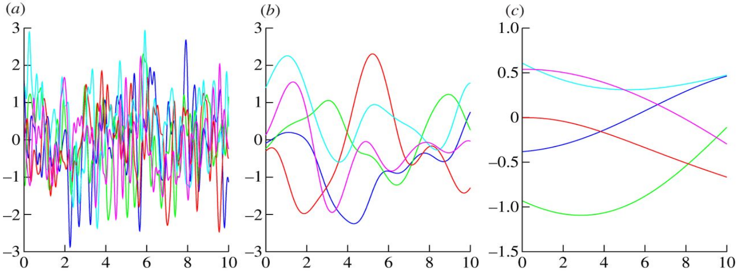
Covariance Functions

Mean Function $m(x)$: For most applications using zero mean function is sufficient.

Covariance Function $k(x, x')$: Defines prior properties of functions sampled from $\mathcal{GP}$ such as: Stationarity, Smoothness, Length-scales.

Squared Exponential Covariance Function:

$$k(x, x') = \alpha \exp\left(-\frac{\gamma}{2}(x - x')^T(x - x')\right), \ \alpha, \gamma : \text{hyperparameters} \qquad (3)$$



Motivation for Dimensionality Reduction

- For data with underlying “structure”, we expect:
  - Fewer distortions than dimensions.
  - Data to lie on a low-dimensonal manifold.
- Conclusion: Deal with high-dimensional data by looking for low-dimensional embedding.

Motivation for Dimensionality Reduction

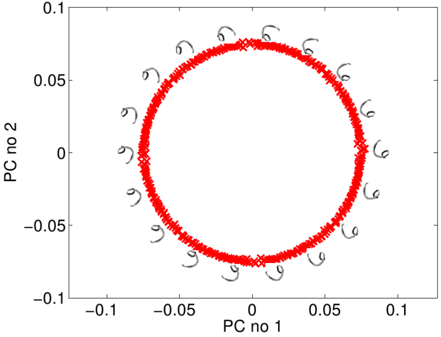
UPSC Handwritten Digit Dataset

3648 dimensional space

Digit 6 Image

Random Image

Low-dimensional manifold for digit rotation



Probabilistic Generative Model

- Observed (high-dimensional) data: $\mathbf{Y} = [y_1 \ y_2 \ \cdots \ y_N]^T \in \mathbb{R}^{N \times D}$
- Latent (low-dimensional) data: $\mathbf{X} = [x_1 \ x_2 \ \cdots \ x_N]^T \in \mathbb{R}^{N \times Q}, \ Q << D$
- Assume a relationship/mapping of the form:

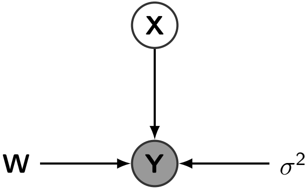
$$\begin{aligned} y_i &= \mathbf{W}x_i + \epsilon_i, \ \epsilon_i \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}) \\ y_i &= f(x_i) = \epsilon_i \end{aligned} \tag{4}$$

- Resultant likelihood on the data:

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{W}) = \prod_{i=1}^{N} \mathcal{N}(y_i|\mathbf{W}x_i, \sigma^2 \mathbf{I}) \tag{5}$$

Probabilistic Generative Model

Probabilistic PCA

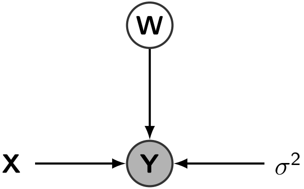


Places prior on latent space **X** and optimises linear mapping **W**

$$p(\mathbf{X}) = \prod_{i=1}^{N} \mathcal{N}(x_i|\mathbf{0}, \mathbf{I})$$

$$p(\mathbf{Y}|\mathbf{W}, \sigma^2) = \int p(\mathbf{Y}|\mathbf{W}, \mathbf{X}, \sigma^2)\ p(\mathbf{X}) \tag{6}$$

Dual Probabilistic PCA



Places prior on linear mapping **W** and optimises latent space **X**

$$p(\mathbf{W}) = \prod_{i=1}^{D} \mathcal{N}(w_i|\mathbf{0}, \mathbf{I})$$

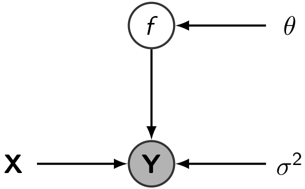
$$p(\mathbf{Y}|\mathbf{X}, \sigma^2) = \int p(\mathbf{Y}|\mathbf{W}, \mathbf{X}, \sigma^2)\ p(\mathbf{W}) \tag{7}$$

From Dual PPCA to GP-LVM

PPCA and Dual PPCA are equivalent eigenvalue problems with same Maximum Likelihood solution

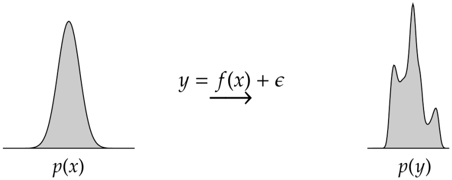
- GP-LVM: Instead of placing prior $p(\mathbf{W})$ on the function parameters in Dual PPCA, we can place a prior $p(f)$ directly on the mapping function i.e. $\mathcal{GP}$ Prior
- A $\mathcal{GP}$ Prior allows for non-linear mappings if the covariance function is non-linear. For example, the SE Covariance Function:

$$k(x, x') = \alpha \exp\left(-\frac{\gamma}{2}(x - x')^T(x - x')\right) \tag{8}$$



Difficulty with Non-linear Mapping

- Normalization of probability distribution after passing through non-linear mapping becomes difficult:

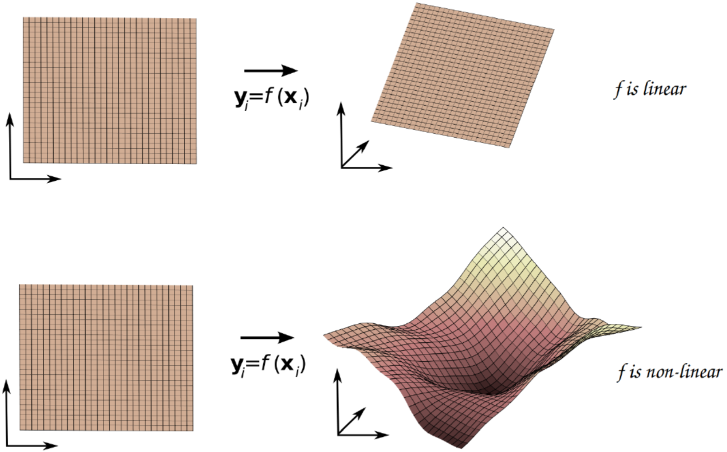


- No longer possible to optimize wrt $\mathbf{X}$ as an eigen value problem

$$\mathbf{X}, \theta = \text{argmax}_{\mathbf{X},\theta} p(\mathbf{Y}|\mathbf{X}, \theta) \tag{9}$$

- Instead we need to use iterative approach and find gradients wrt $\mathbf{X}, \alpha, \gamma, \sigma^2$

Linear vs. Non-linear Dimensionality Reduction



Extensions of GP-LVM

Back Constrained GP-LVM: Ensures points close in the observation space (Y) will be close in latent space by constraining back mapping $f' : Y \to X$

GP-LVM with Dynamics Model: Computes latent space assuming that the latent positions ($\mathbf{X}$) are sequential:

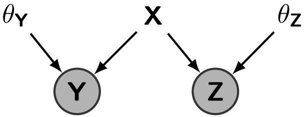
$$x_t = h(x_{t-1}) + \epsilon_{dyn}, \epsilon_{dyn} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{dyn}\mathbf{I}) \tag{10}$$

A $\mathcal{GP}$ Prior is placed on the function $h(x)$. The resultant optimization becomes:

$$\mathbf{X}, \theta, \theta_{dyn} = \text{argmax}_{\mathbf{X},\theta,\theta_{dyn}}\ p(\mathbf{Y}|\mathbf{X}, \theta)\ p(\mathbf{X}|\theta_{dyn}) \tag{11}$$

Shared GP-LVM

- Objective: Learn a shared latent structure ($\mathbf{X}$) for two observation spaces ($\mathbf{Y}, \mathbf{Z}$) of the same phenomenon

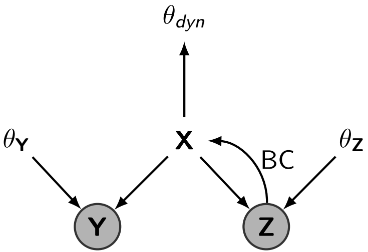


- This model indicates the following:

$$
\begin{aligned}
X, \theta_{\mathbf{Y}}, \theta_{\mathbf{Z}} &= \mathrm{argmax}_{X,\theta_{\mathbf{Y}},\theta_{\mathbf{Z}}}\ p(\mathbf{Y}, \mathbf{Z} | X, \theta_{\mathbf{Y}}, \theta_{\mathbf{Z}}) \\
&= \mathrm{argmax}_{X,\theta_{\mathbf{Y}},\theta_{\mathbf{Z}}}\ p(\mathbf{Y} | X, \theta_{\mathbf{Y}})\ p(\mathbf{Z} | X, \theta_{\mathbf{Z}})
\end{aligned} \tag{12}
$$

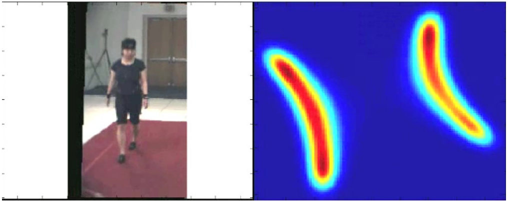
Human Motion Modeling using GP-LVM

- Ek et al. (2007) applied shared GP-LVM to infer human pose from ambiguous silhoutte information.
- Extension to Shared GP-LVM:
  - Back constraint from pose space (**Z**)to latent space (**X**) to enforce *one-to-one* mapping.
  - Dynamic model in the latent space to enforce representation to follow data's dynamics.



Human Pose Inference using Shared GP-LVM

Ambiguity in heading direction



Ambiguity in positon of legs

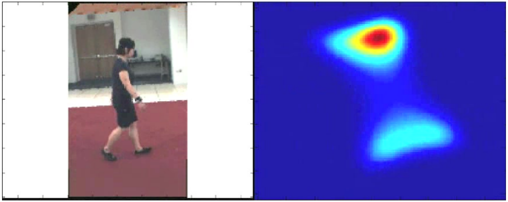


Table of Contents

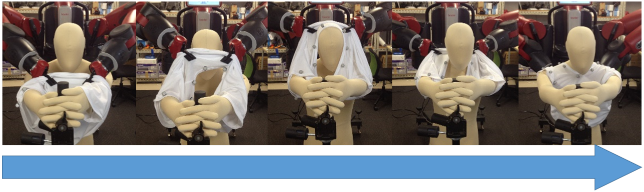
Dynamical Modelling of Clothing Articles

Objective: Verify effectiveness of GP-LVM for clothing assistance task

Measured T-shirt collar and sleeve shape using motion capture system

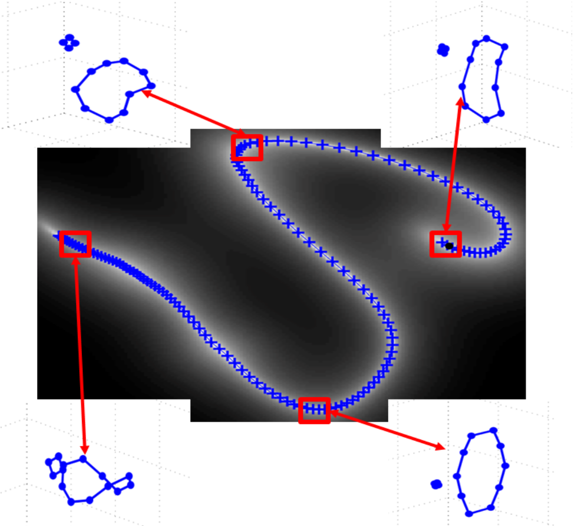
Observed Data, $\mathbf{Y} \in \mathbb{R}^{45}$, Location of 15 Markers

Latent Positions, $\mathbf{X} \in \mathbf{R}^2$, 2-Dimensional space



Motion Model from Ground-truth Data

GP-LVM Model learnt for single clothing trial ($N = 141$)



Motion Model for Multiple Clothing Trials

GP-LVM Model learnt for 4 different clothing trials ($N = 620$)

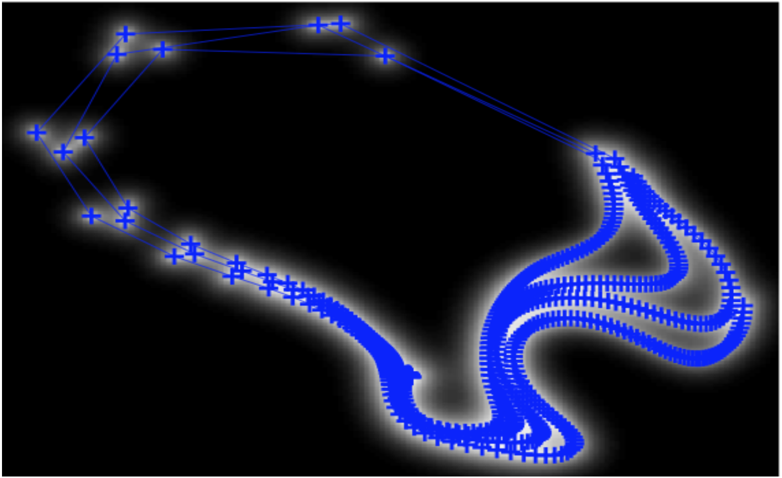


Table of Contents

Cloth State Prediction from Multiple Observations

Objective: Infer cloth state from noisy observation and shared GP-LVM model

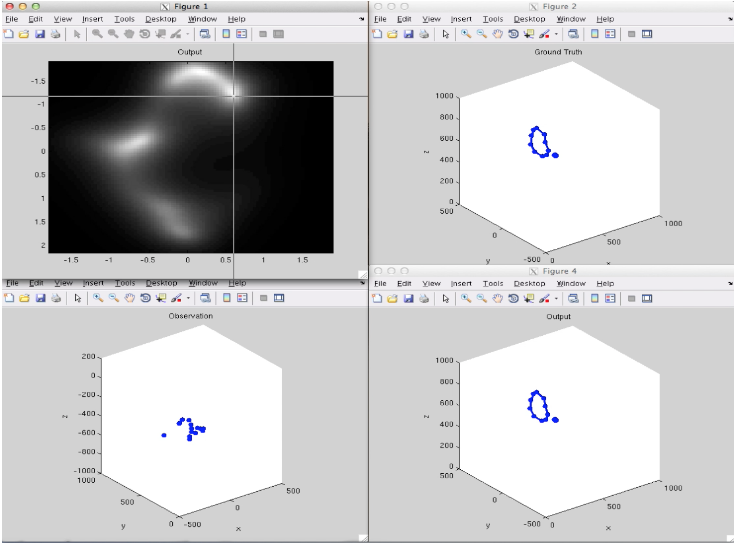
Measured T-shirt collar and sleeve shape using both motion capture system and kinect depth sensor

Pose Space, $\mathbf{Z}$: Motion Capture Data

Feature Space, $\mathbf{Y}$: Kinect Depth Data

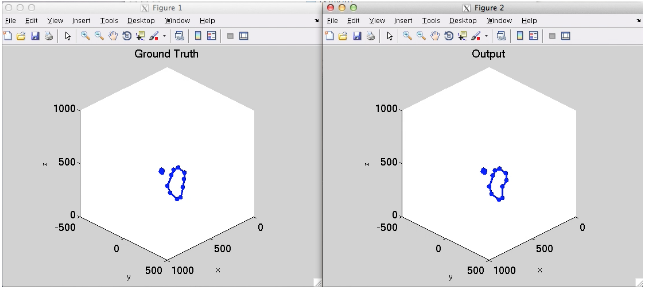
Latent Space, $\mathbf{X}$, 5-dimensional representation

Likelihood over Latent Space



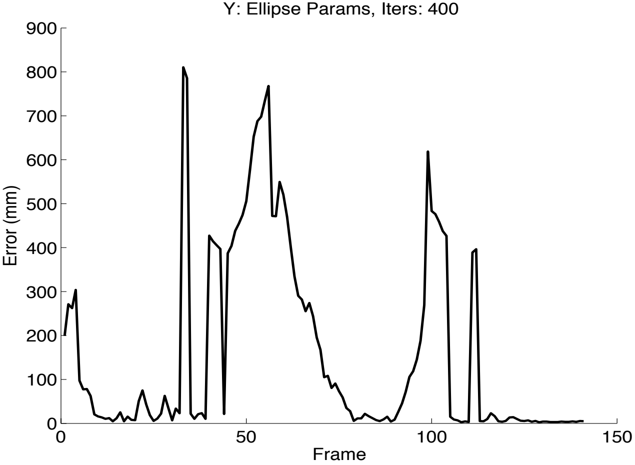
State Prediction using Shared GP-LVM

T-shirt state predicted from noisy observation corresponding to latent space point that maximizes likelihood



State Prediction using Shared GP-LVM

Euclidean distance error between ground-truth and predicted T-shirt state



Effect of Feature Space Representation

Feature space representation plays important role in performance of model

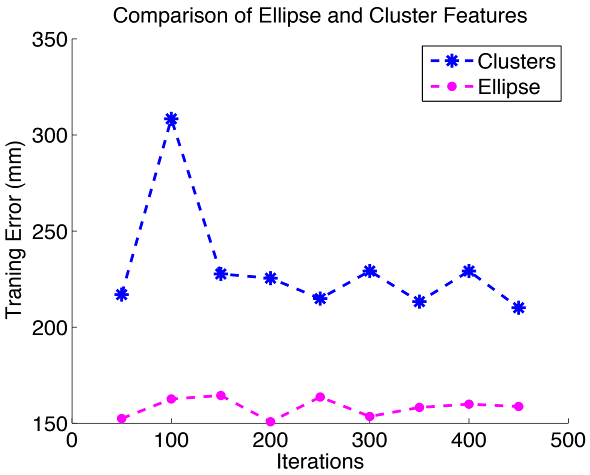


Table of Contents

Conclusion

We demonstrated that

- GP-LVM can model dynamics of clothing articles for clothing assistance tasks
- Shared GP-LVM can be used to infer true cloth state from noisy depth sensor readings.

We plan to

- Improve accuracy of Shared GP-LVM by using better representations.
- Extend dynamics model to include human posture, robot's proprioceptive information.

Thank you for your kind attention

Questions and Comments