Online Activity Recognition through Kernel Methods

Evgeni Pavlidis

2014-12-05

Outline

1. Introduction
   - Motivation
   - Problem Statement
   - Gaussian Process - Latent Variable Model
2. Approaches
   - k-Means
   - Discriminate Sequence Back-Constrained GP-LVM
   - Gaussian Process - Latent Motion Field
3. Conclusions & Outlook
   - Lessons learned
   - Outlook

The Spencer Project

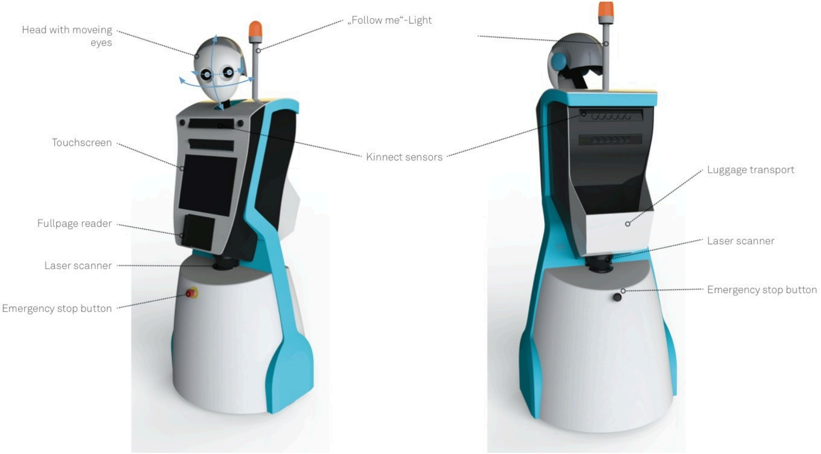


Figure: The final design of the SPENCER robot and its sensors.

Problem statement

Devise an algorithm which:

- is capable of online activity recognition
    -
- is not dependent on large amount of training data

Gaussian Process - Latent Variable Model

- A model for non-linear dimensionality reduction
- Has a Bayesian interpretation

Probabilistic Principal Components Analysis

Dual Probabilistic Principal Components Analysis

[2]

$$P(W) = \prod_{i=1}^{d} \mathcal{N}(\boldsymbol{w}_i | 0, \boldsymbol{I})$$

$$p(Y|X) = \prod_{i=1}^{d} \mathcal{N}(\boldsymbol{y}_i | 0, XX^T + \sigma^2 I)$$

We can interpret: $XX^T + \sigma^2 I = K$. as a linear kernel, leading to:

$$P(Y|X) = \prod_{i=1}^{d} \mathcal{N}(\boldsymbol{y}_i | 0, K)$$

Thus the *Dual Probabilistic PCA* can be interpreted as product of *Gaussian Processes* with a linear kernel.

$$log\ p(Y|X) \propto -\frac{p}{2} \log(|K|) - \frac{1}{2} tr(K^{-1} Y Y^T)$$

Back Constrained GP-LVM

[3]

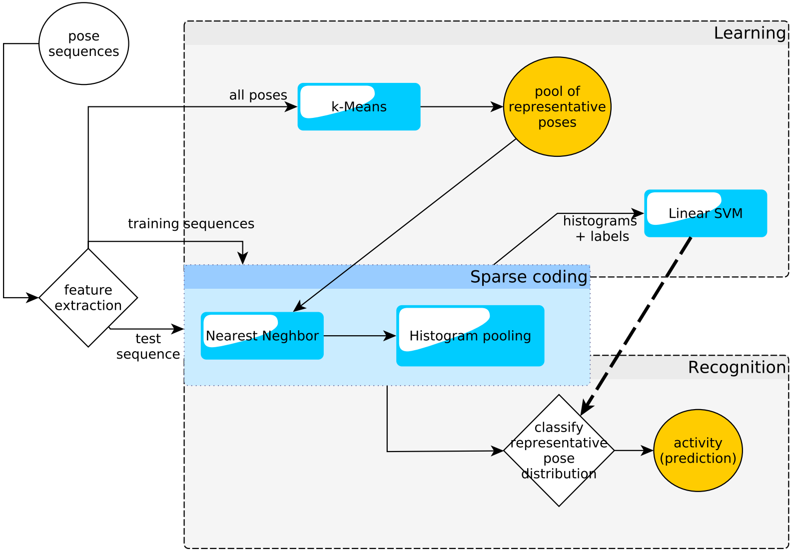
$$x_{i,j} = g_j(\boldsymbol{y}_i, \gamma)$$

$$g_j(\boldsymbol{y}_n, A, l, \sigma) = \sum_{i=1}^{n} A_{j,i} k(\boldsymbol{y}_n, \boldsymbol{y}_i)$$

Daily Living Activity Recognition

[7]

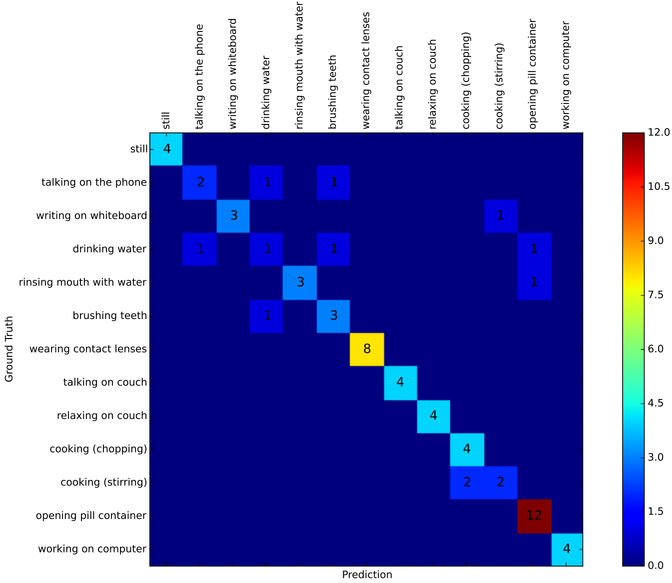
Illustration



Extensions:

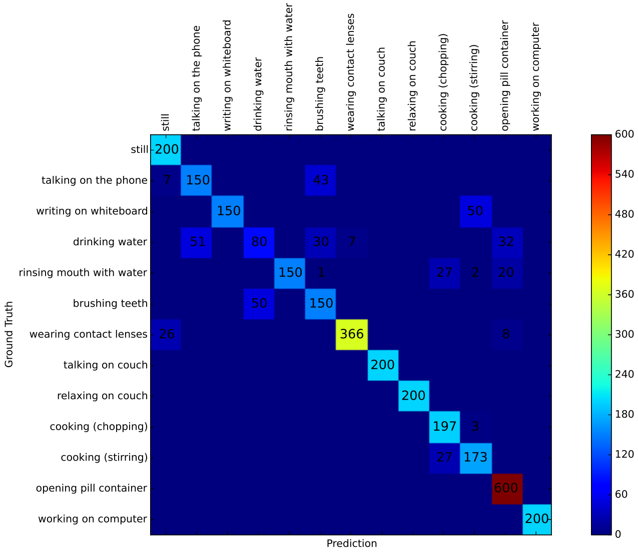
- Extract representative poses for each class
- Use sequence alignment functions for classification
    - Longest Common Subsequence
    - Dynamic Time Warping

Results

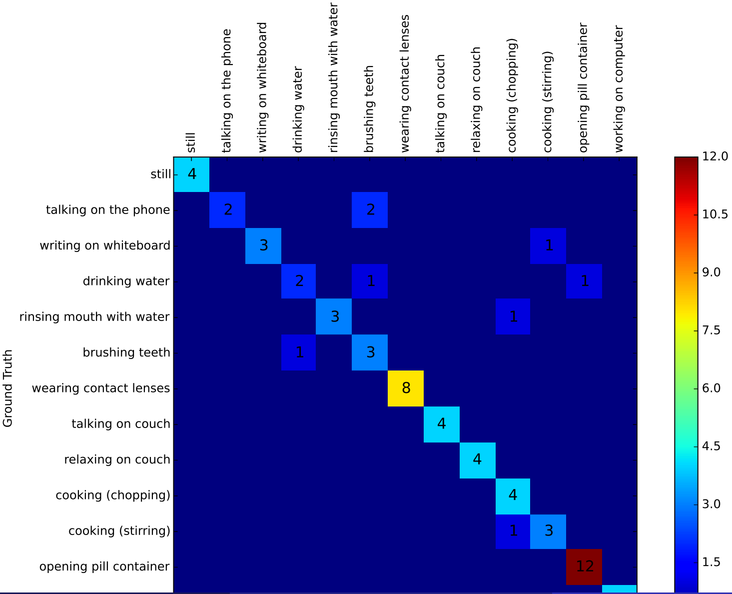


still
talking on the phone
writing on whiteboard
drinking water
rinsing mouth with water
brushing teeth
wearing contact lenses
talking on couch
relaxing on couch
cooking (chopping)
cooking (stirring)
opening pill container
working on computer
Ground Truth
Prediction
4
2
1
1
3
1
1
1
1
1
3
1
1
3
8
4
4
4
2
2
12
4
12.0
10.5
9.0
7.5
6.0
4.5
3.0
1.5
0.0

BoF Approach with subsequences



BoF with Longest Common Subsequence



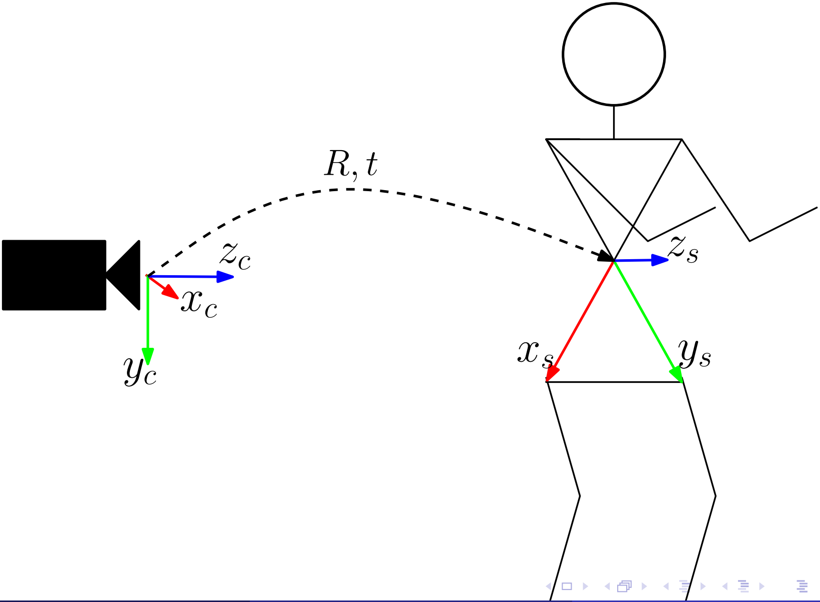
Issues

Very difficult to adjust the algorithm to perform online recognition.

Discriminate Sequence Back-Constrained GP-LVM

[5]

Feature selection



Sequence Back-Constraints

$$g_q(Y_s) = \sum_{m=1}^{S} a_{mq} k(Y_s, Y_m)$$

where the similarity measure is $k(Y_s, Y_m) = \gamma e^{\mathrm{DTW}(Y_s, Y_m)}$. This measure can be interpreted as a sequence alignment kernel. This measure is to be preserved in the latent spaces.

$$g_q(Y_s) = \mu_{sq} = \frac{1}{L_s} \sum_{n \in J_s} x_{nq}$$

Discriminative GP-LVM

[6]
Make the latent space more discriminative by minimizing inner-class variance and maximizing inter-class separability.

- The distance between the classes

$$S_b = \sum_{i=1}^{l} \frac{n_i}{n} (\boldsymbol{\mu_i} - \boldsymbol{\mu})(\boldsymbol{\mu_i} - \boldsymbol{\mu})^T$$

where n is the number of samples, n_i is the number of samples for class i and l is the number of classes. Furthermore $\boldsymbol{\mu_i}$ is the mean of the class and $\boldsymbol{\mu}$ is the mean across all classes.

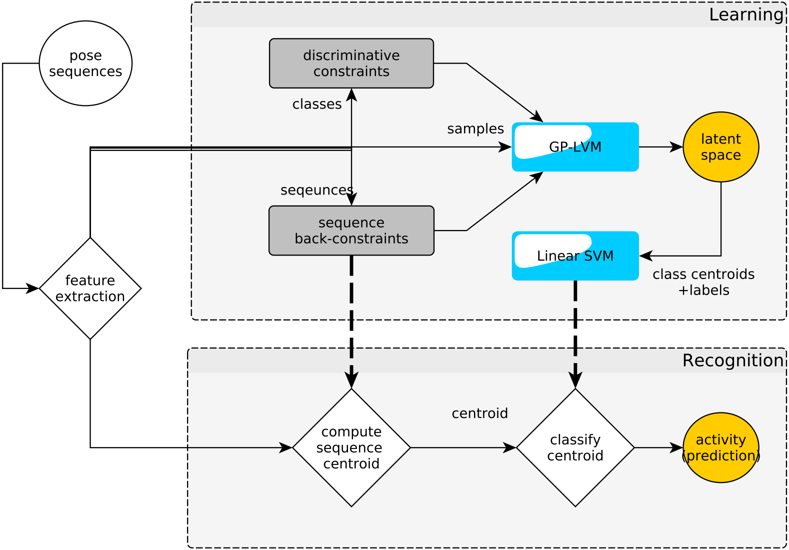
- The variance within each class

$$S_w = \frac{1}{n} \sum_{i=1}^{l} \sum_{j=1}^{n_i} \frac{n_i}{n} (\boldsymbol{x_{i,j}} - \boldsymbol{\mu_i})(\boldsymbol{x_{i,j}} - \boldsymbol{\mu_i})^T$$

where $\boldsymbol{x_{i,j}}$ is the j-th sample from class i.

$$J(X) = tr(S_w^{-1} S_b)$$

Illustration



Issues

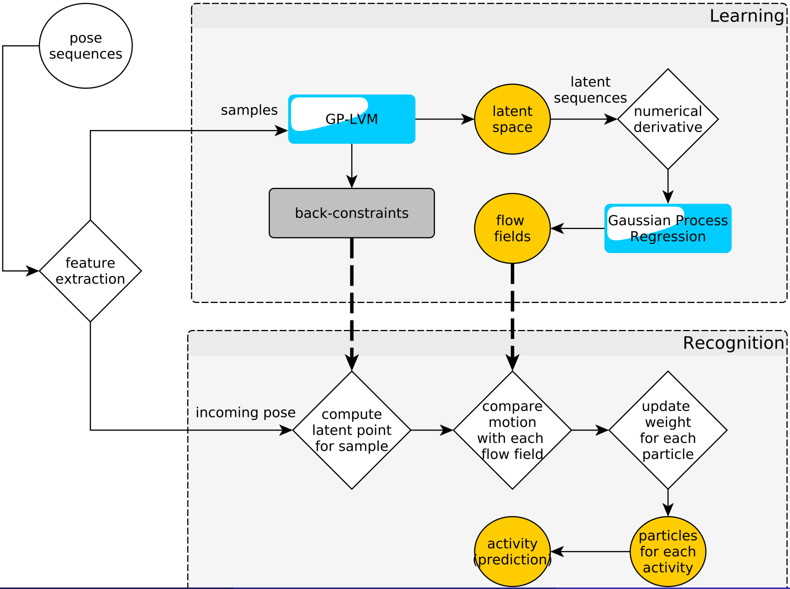
Due to the nature of more complex activities and the huge search space, the optimization of the GP-LVM did not perform well.

Gaussian Process - Latent Variable Model

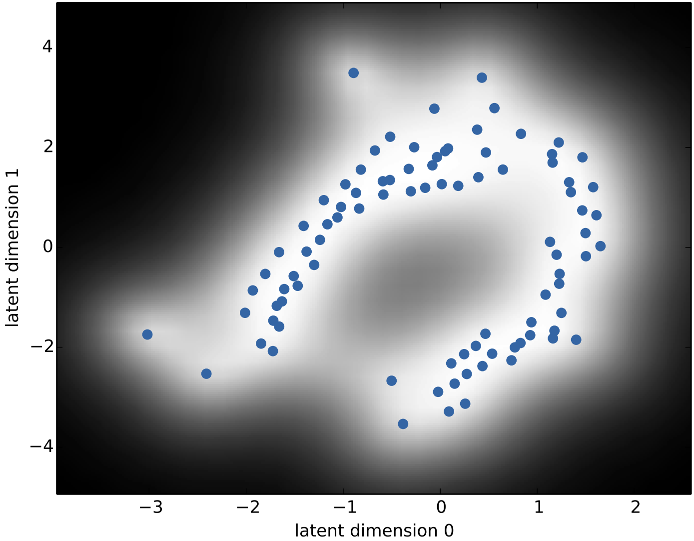
Inspired by:[1]

- Perform a separate dimensionality reduction for each activity class
- Learn a motion flow field by GP regression on the velocity function
- Online activity by comparing the incoming motion with each flow field and updating the belief

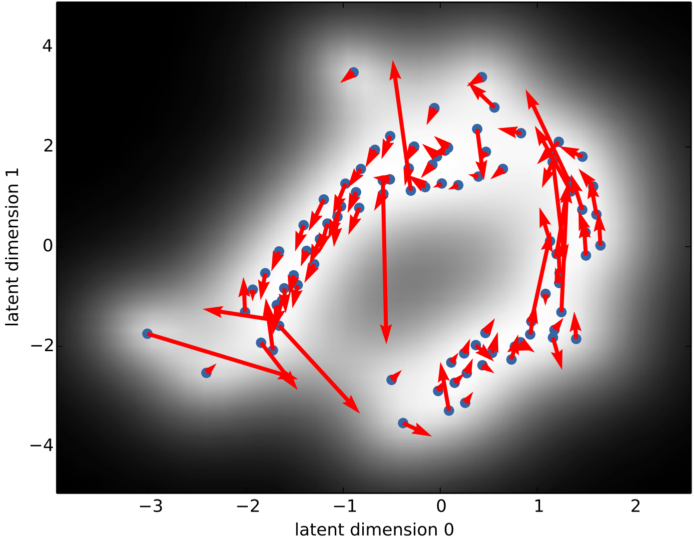
Illustration



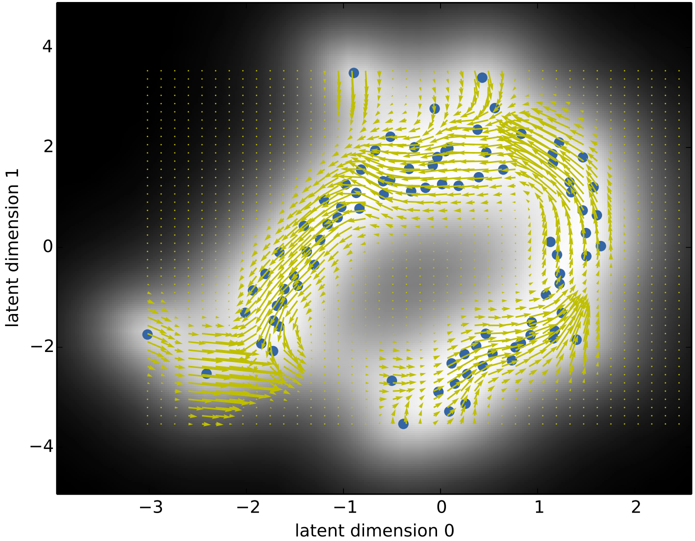
Latent Motion



Latent Motion - velocity



Latent Motion Flow Field



Issues

- No smooth mapping from observed space to latent space
- This leads to discontinuities when learning the regression
- Possible solution:
  - Use tempo-spatial constraints in the optimization ([4])

Lessons learned

- Common dimensionality reduction for a large number of activities is intractable
- Optimization of the *GP-LVM* is very difficult and strongly depends on the initialization
- Local motion tendencies are more discriminative for complex activities then the overall dynamics

Contributions

- Implementation and extensions of a k-Means based approach in Python
- Implementation of a *ROS* module capable of activity recognition in real-time
- Implementation of the Discriminative GP-LVM in Python
- Implementation of the Sequence Back-constraints in Python
- A novel approach for activity recognition using latent motion flow fields
- Advantages and disadvantages of dimensionality reduction with GP-LVM in the context of activity recognition

Outlook

Implementation of the GP-Latent Motion Field using spatio-temporal GP-LVM

As described earlier the GP-LMF approach failed, due to the fact that the optimization of the GP-LVM with back-constrained did not result in a smooth backward mapping. One possibility to solve this problem is by implementing the *Spatio-Temporal GP-LVM* as described in [4]. Doing so the latent space will also take the temporal order of the poses into account, resulting in a smoother backward mapping (observed space to latent space).

Adaptive GP Regression of the flow field

One issue of the proposed *GP-LMF* method, besides the difficulty to obtain a smooth mapping, is that the *RBF* kernel is stationary. This means that the lengthscale is defined over the whole latent space. We would like this lengthscale to adapt to the curvature of the flow field. This way small and large motion tendencies can be learned. An approach that could be tested to capture this, is using a non-stationary kernel.

Semi-supervised activity learning by automatic segmentation of activities

References

Kihwan Kim, Dongryeol Lee, and Irfan Essa.
Gaussian process regression flow for analysis of motion trajectories.
In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1164–1171. IEEE, 2011.

Neil D. Lawrence.
Gaussian process latent variable models for visualisation of high dimensional data.
In *Nips*, volume 2, page 5, 2003.

Neil D. Lawrence and Joaquin Quiñonero-Candela.
Local distance preservation in the GP-LVM through back constraints.
In *Proceedings of the 23rd international conference on Machine learning*, pages 513–520. ACM, 2006.