Topologically-Constrained Latent Variable Models

Raquel Urtasun, David Fleet, Trevor Darrell and Neil D. Lawrence

October 12, 2007

1 Introduction

Non-linear dimensionality reduction has become a popular approach to dealing with high dimensional data sets. Approaches such as locally linear embedding (LLE), Isomap and maximum variance unfolding (MVU) [3, 4, 7] all attempt to discover the topology of a low-dimensional representation, through interconnections between points in the data space. However, if the data are relatively sparse or noisy, these interconnections can stray beyond the 'true' local neighbourhood and the therefore produce a poor embedding.

Probabilistic formulations of latent variable models do not usually include explicit constraints on the embedding in terms of neighbouring points in the data space[1]. For these models, even with the correct topology and dimension of the latent space, if the initialization of the model is poor, the learning might get stuck in local minima. Moreover, sometimes the maximum likelihood solution does not reflect our intuitions about how the data should be embedded. To get better models in such cases more constraints need to be added to the model. In this abstract we describe how explicit topological constraints can be imposed within the context of probabilistic latent variable models. We describe two approaches, both within the context of the Gaussian process latent variable model (GP-LVM) [1]. The first uses prior distributions on the latent space that encourage a given topology. The second involves constraining the latent space and optimisation through constrained maximum likelihood [2].

Our approach is motivated by the problem of modeling human pose and motion for video-based people tracking and computer animation. Human motion is an interesting domain because, while there is an increasing amount of motion capture data available, the diversity of human motion means that we will necessarily have to incorporate a large amount of prior knowledge to learn probabilistic models that can accurately reconstruct a wide range of motions. Despite this, most existing methods for learning pose and motion models do not fully exploit useful prior information, and many are limited to modeling a single human activity (e.g. walking with a particular style).

This abstract describes how prior information can be used effectively to learn models with specific topologies that reflect the cyclic nature of human gaits. Importantly, with this information we can also model multiple activities, including transitions between them (e.g. from walking to running), even when such transitions are not present in the training data. As a consequence, we can now learn latent variable models with training motions comprising multiple subjects with stylistic diversity, as well as multiple activities, such running and walking.

2 Top Down Imposition of Topology

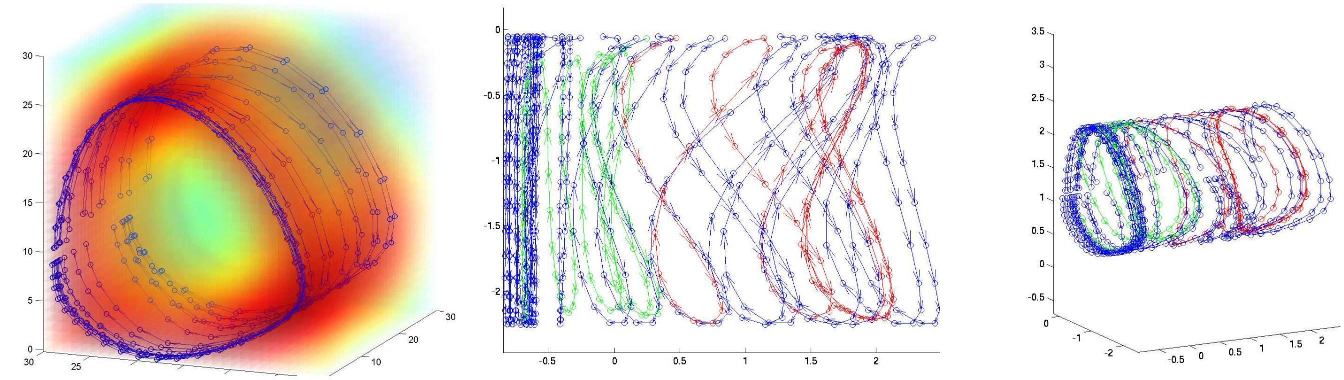
The smooth mapping in the GP-LVM ensures that distant points in data space remain distant in latent space. However, the mapping in the opposite direction is not required to be smooth. While the Gaussian process dynamics may be added to mitigate this effect (the GPDM [5, 6]), they often produces models that are neither smooth nor generalize well.

To help ensure smoother, well-behaved models, [2] suggest the use of *back-constraints*, where each point in the latent space is a smooth function of its corresponding point in data space. Nevertheless, when learning human motion data with large stylistic variations or different motions, neither GPDM nor back-constrained GP-LVM produce smooth models that generalize well.

A problem for modelling human motion data is the sparsity of the data relative to the diversity of naturally plausible motions. For example, while we might have a data set comprising different motions, the data may not contain transitions between motions. In practice however, we know that these motions will be broadly cyclic and that transitions can only physically occur at specific points in the cycle. How can we encourage our model to respect such topological constraints which arise from prior knowledge?

In this abstract we first consider an alternative approach to the 'hard' constraints on the latent space suggested by [2]. We introduce topological constraints through a prior distribution in the latent space, based on a neighborhood structure learn through a generalized local linear embedding (LLE) [3]. The distance metric in LLE can be adjusted to better reflect different types of prior knowledge (such as the location of possible transition points). We then can define similarity measures for use with the functionally constrained GP-LVM. Both these approaches force the latent space to construct a representation that reflects our prior knowledge and they can be combined. This allows us to develop motion models with specific topologies that incorporate different activities and transitions between them.

[1] An exception is the back-constrained GP-LVM [2] where a constrained maximum likelihood algorithm is used to enforce these constraints.



(a) (b) (c)

Figure 1: Hybrid model learned using local linearities for the style and backconstraints for the content. The training data is composed of 9 walks and 10 runs performed by different subjects and speeds. (a) Likelihood for the reconstruction of the latent points (b) First two components and (c) 3D view of the latent trajectories for the training data in blue. Sampled motions are shown in green and red.

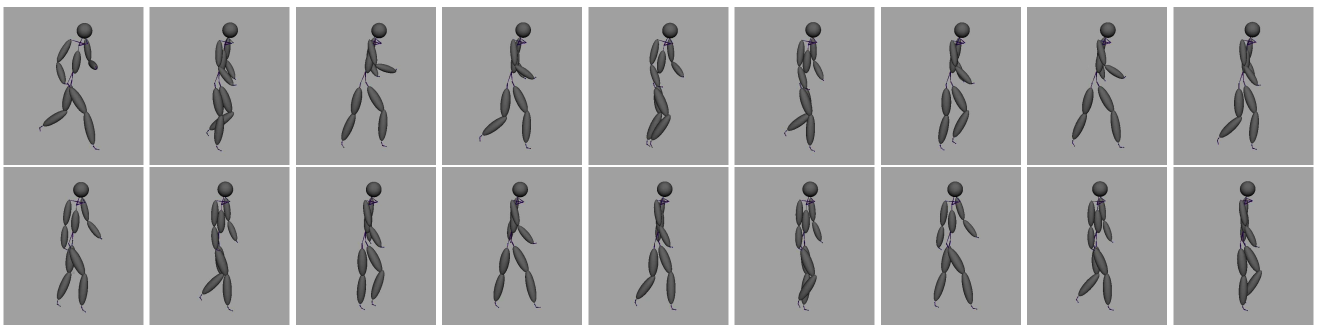


Figure 2: **Transition from running to walking**: The system transitions from running to walking in a smooth and realistic way. The transition is encouraged by incorporating prior knowledge in the model. The latent trajectories are shown in green in Fig. 1 (b,c).

Figure 1 (a) shows a hybrid model learned using priors for smoothness, and back-constraints for topology. The larger training set comprises approximately one gait cycle from each of 9 walking and 10 running motions performed by different subjects at different speeds (3 km/h for walking, 6–12 km/h for running). Colors in Fig. 1 (a) represent the variance of the GP as a function of latent position. Only points close to the surface of the cylinder produce poses with high certainty. Under different initial conditions the model is able to simulate different walking and running styles and transitions between then as depicted by Fig. 2.

3 Conclusions

In this abstract we have proposed a general framework of probabilistic models that learn smooth latent variable models of different activities within a shared latent space. We have also described a principled way to include prior knowledge, that allow us to learn specific topologies and transitions between the different motions. We focussed on models composed of walking and running, but our framework is general, being applicable in any data sets where there is a large degree of prior knowledge for the problem domain, but the data availability is relatively sparse compared to its complexity.

References

[1] N. Lawrence. JMLR 6:1783–1816, 2005.

[2] N. Lawrence and J. Quinonero-Candela. In *ICML*, 2006

[3] S. Roweis and L. Saul. *Science*, 290(5500):2323–2326, 2000.

[4] J. Tenenbaum, V. de Silva, and J. Langford. *Science*, 290(5500):2319–2323, 2000.

[5] R. Urtasun, D. J. Fleet, and P. Fua. In *CVPR*, 2006.

[6] J. Wang, D. J. Fleet, and A. Hertzman. In *NIPS*, 2005.

[7] Kilian Q. Weinberger and Fei Sha and Lawrence K. Saul. In *ICML*, 2004