GP-LVMs for studying human grasping actions

Thomas Feix Javier Romero Hedvig Kjellström Danica Kragic

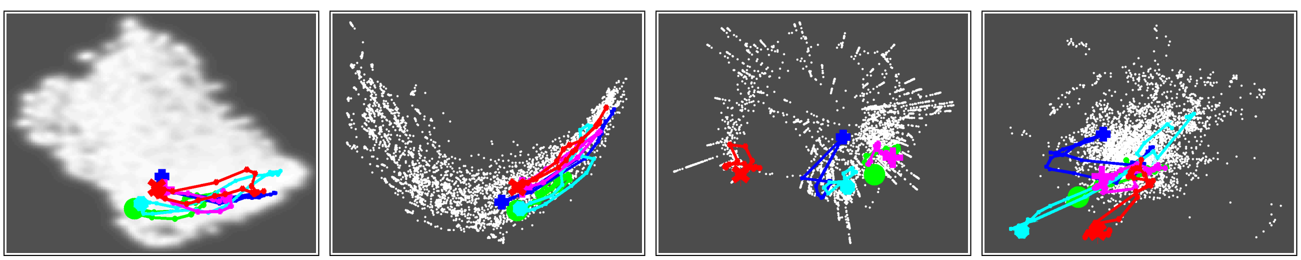
A multitude of grasping taxonomies have been proposed for grasping actions, [1]. Most of them have been designed heuristically and were never compared or evaluated with respect to how effectively they describe the space of natural grasping motions. Furthermore, they are based solely on the final grasp posture of the hand, without taking the approach and retreat phase into account.

In the work presented here we study the generation of a latent space which can help us to understand the underlying structure in human grasping actions. This will be done in a data driven manner, in contrast to the classical approach where it is based solely on the intuition of the author [1].

The data used in the system was extracted from human demonstrations (5 subjects) of 31 different grasp types recorded with a Polhemus magnetic tracker. The data space is spanned by the position (3 parameters) and orientation (4 parameters of the quaternion) of the 5 fingertips. This resulted in a space of dimensionality 35. From each trial we took 30 equally spaced samples, which lead to a total of 4650 datapoints.

We used Gaussian Process Latent Variable Models (GPLVMs) from Lawrence Matlab toolbox ([2]) to construct our latent space. GPLVMs have been used in the last years for modeling full-body motion, showing some advantages over other dimensionality reduction techniques [3]. To force the data to have a smooth mapping between data space and latent space we used back constraints which ensures that nearby points in data space stay near in the latent space [4].

The results of our 2D latent space were compared with other dimensionality reduction techniques, showing that the GPLVM space (Figure 1(a)) had a better inter-grasp separability and lower inter-subject variance, as well as better time continuity. PCA (Figure 1(b)) was not able to spread the different grasp types, all trajectories lie on a rather narrow arc. Isomap (Figure 1(c)) and LLE (Figure 1(d)) were not able to extract any meaningful structure. There is no common pattern between subjects as well as there is no smooth path of the movement. This is due to the fact that these two methods are much more sensible to error in the data.



(a) GPLVM (b) PCA (c) Isomap (d) LLE

Fig. 1. The figure shows the trajectories of grasp number 1 for all 5 subjects. The other grasp types show similar patterns.

A continuous temporal path for each grasp type was estimated on the latent space using Gaussian Mixture Regression (GMR) toolbox from Calinon [5], see Figure 2. As can be seen in [5], the GMR representation can be used for generating actions that resemble the demonstrated examples but are also conform to some constraints in the latent space. In our work we used the GMR representation to compute the similarity of different grasps, including the approach and retreat phase. This similarity measure was used for clustering the demonstrated grasps, creating in that way a taxonomy based on demonstrated grasps. Figure 3 shows the clusters found by the algorithm where there are clear trends visible within these clusters. E.g. in *Cluster 1* all grasps are power grasps with the four fingers in a similar position. The MCP joint is extended and PIP and DIP are flexed. In addition the thumb is in a very adducted and extended position. There are some conflicts with the classifications presented in [1], which seems to be due to the fact that the GPLVM algorithm gives each finger the same importance. In the taxonomies which are based on the authors intuition, the thumb is placed in a prominent position over the other fingers.

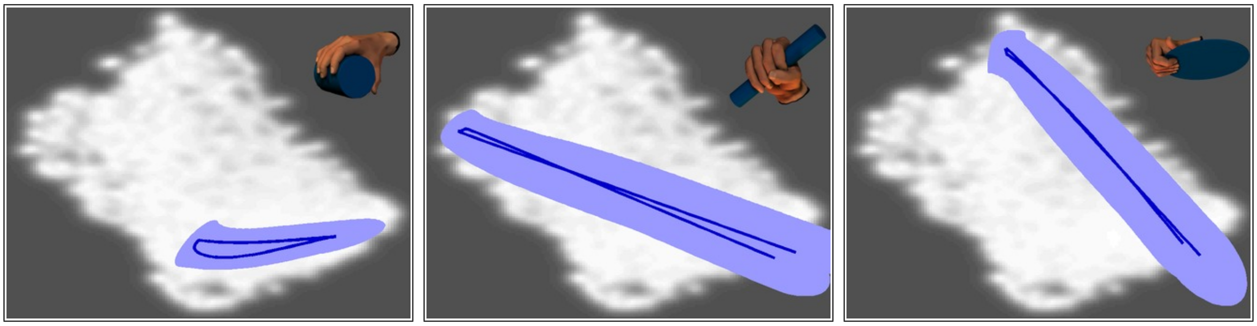


Fig. 2. GMR regression for 3 of the 31 grasp movements of all subjects. The dark line indicates the mean trajectory and the light area correspond to the variance

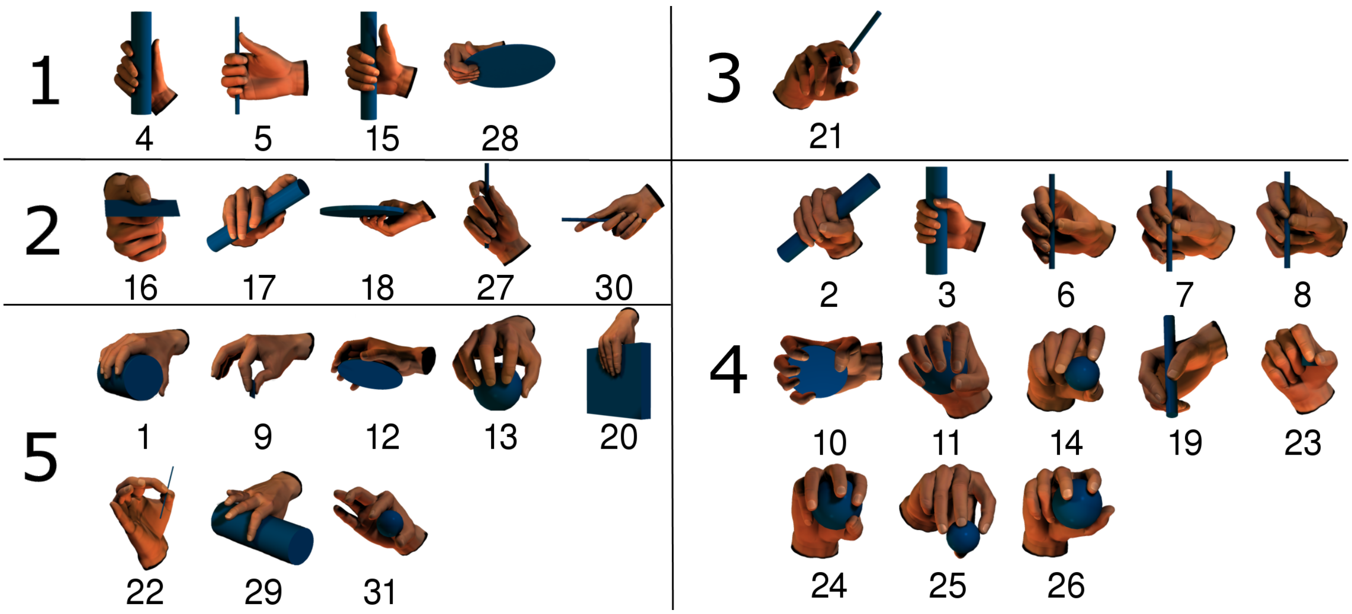


Fig. 3. Clusters based on the GMR regression of the demonstrated grasp types.

In conclusion, we showed that GPLVMs are a powerful tool for embedding grasp demonstration data into a low dimensional space which preserves time continuity and is robust to inter-subject variance. Such a latent space can be useful in multiple ways. First, it allows insights to human grasping actions which cannot be archived by looking at the high dimensional data. Second, it restricts the the data space to postures actually performed by humans during grasping actions which can simplify hand tracking. Third, the GMR representation of grasp types is flexible and therefore allows to synthesize previously unseen grasp types. Finally, it can be used to visualize the performance of a robotic hand in contrast to human hands when the movements of an anthropomorphic hand are projected to latent space.

References

[1] T. Feix, R. Pawlik, H. Schmiedmayer, J. Romero, and D. Kragic, "A comprehensive grasp taxonomy," in *Robotics, Science and Systems Conference: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation, Poster Presentation*, June 2009.

[2] N. Lawrence and A. Hyvärinen, "Probabilistic non-linear principal component analysis with gaussian process latent variable models," *Journal of Machine Learning Research*, vol. 6, pp. 1783–1816, 2005.

[3] R. Urtasun, D. J. Fleet, and P. Fua, "3D people tracking with gaussian process dynamical models," in *CVPR*, 2006, pp. I: 238–245.

[4] N. D. Lawrence and J. Quinonero-Candela, "Local distance preservation in the gp-lvm through back constraints," in *ICML06*. New York, NY, USA: ACM, 2006, pp. 513–520.

[5] S. Calinon, F. Guenter, and A. Billard, "On learning, representing and generalizing a task in a humanoid robot," *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 37, no. 2, pp. 286–298, 2007.