Multivariate Discretization for Bayesian Network Structure Learning in Robot Grasping

Dan Song, CH. Ek, K. Huebner, and D. Kragic

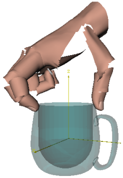
dsong@kth.se

KTH – Royal Institute of Technology, Stockholm, Sweden
CVAP – Computer Vision & Active Perception Lab

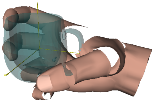
International Conference on Robotics and Automation
ICRA-2011, Shanghai, China

Multivariate Discretization for Bayesian Network Structure Learning in Robot Grasping

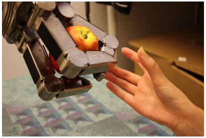
Bad for Pouring



Good for Pouring



Bad for Hand-over



Good for Hand-over

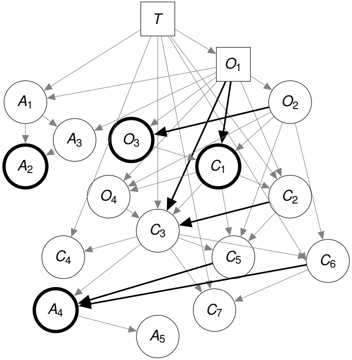
We want a MODEL OF GRASPING TASKS that allows:

- **selecting objects** that afford an assigned task,
- **planning grasps** that satisfy the task constraints.

We need to model the joint distribution:

$$p(\mathbf{Y}),\ \ where\ \mathbf{Y} = \{Y_1, Y_2, \ldots, Y_N\} \leftarrow \{O, A, C, T\} \qquad (1)$$

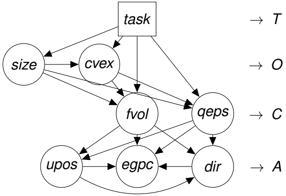
Bayesian Network (BN)



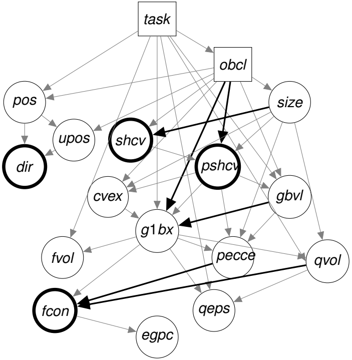
$$p(\mathbf{Y}) = p(\mathbf{Y}|\boldsymbol{\theta}, S) = \prod_{i=1}^{N} p(Y_i|\mathbf{pa}_i, \boldsymbol{\theta}_i, S) \ , \qquad (2)$$

- Factorization of joint distribution through structure S.

Structure Learning



IROS'10 Song *et al.* (2010) specified by experts



ICRA'11, learned from data

To learn structure of a Bayesian network (BN)

- Need to discretize the continuous, high-dimensional data
- Existing methods for multi-variate discretization cluster in original high-dimensional space.

Discretization

Direct discretization suffers from “curse of dimensionality”

Discretization

Direct discretization suffers from "curse of dimensionality"

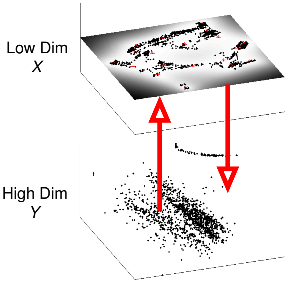
- IDEA: exploit intrinsic, low-dimensional representation for efficient discretization

Discretization

Direct discretization suffers from "curse of dimensionality"

- IDEA: exploit intrinsic, low-dimensional representation for efficient discretization
- How to find the intrinsic representation?
  1. Mapping between the observed and the intrinsic spaces.
  2. Generative mapping –> Gaussian Process Latent Variable Models (GP-LVM)

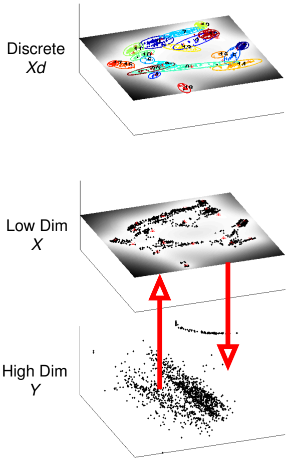
Discretization



Direct discretization suffers from "curse of dimensionality"

- IDEA: exploit intrinsic, low-dimensional representation for efficient discretization
- How to find the intrinsic representation?
  1. Mapping between the observed and the intrinsic spaces.
  2. Generative mapping –> Gaussian Process Latent Variable Models (GP-LVM)

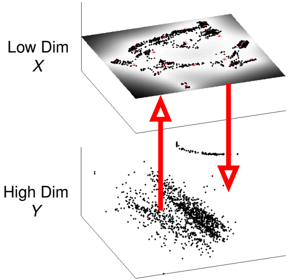
Discretization

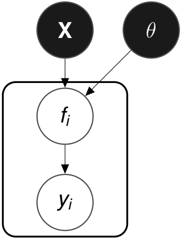


Direct discretization suffers from "curse of dimensionality"

- IDEA: exploit intrinsic, low-dimensional representation for efficient discretization
- How to find the intrinsic representation?
    1. Mapping between the observed and the intrinsic spaces.
    2. Generative mapping –> Gaussian Process Latent Variable Models (GP-LVM)
- Then a Gaussian mixture model (GMM) can be used to discretize the data in this low-dimensional latent space.

GP-LVM



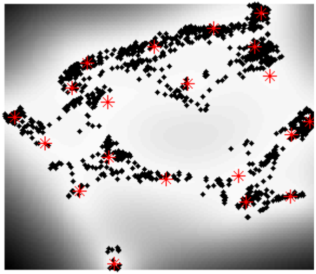


Learning

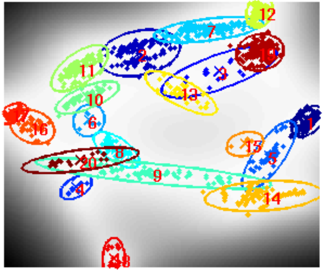
- GP-LVM is a generative dimensionality reduction approach which builds on Gaussian processes (GPs) (Lawrence, 2005).
- $\log p(\mathbf{Y}|\mathbf{X},\theta) = -\frac{1}{2}\mathbf{y}^{\mathrm{T}}\underbrace{(K(\mathbf{X},\mathbf{X})+\sigma^2\mathbf{I})^{-1}}_{\mathcal{O}(n^3)}\mathbf{y} - \frac{1}{2}\log|K(\mathbf{X},\mathbf{X})+\sigma^2\mathbf{I}| + \text{const.}$

Sparse GP-LVM

GP-LVM $\rightarrow$ 2D Latent Space



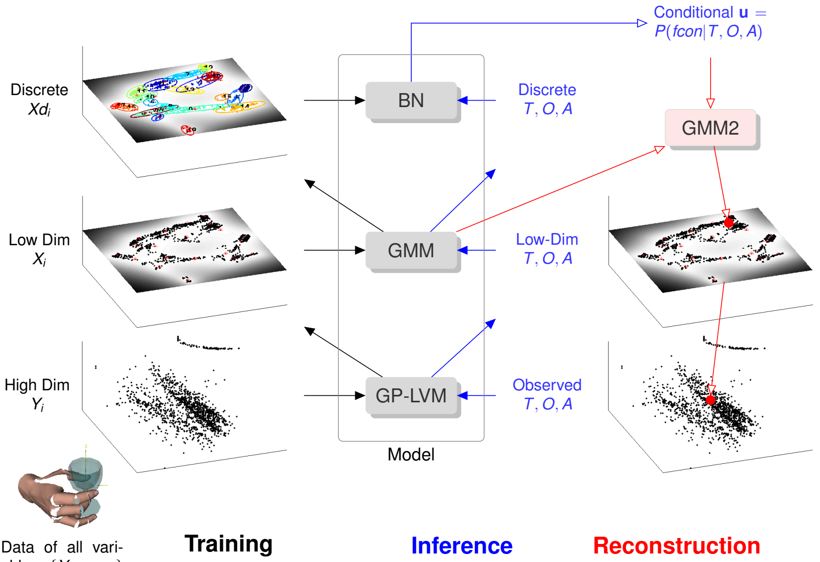
GMM $\rightarrow$ Discretization



- Introducing m inducing points $\mathbf{X}_u$, where $m \lll n$,
  - Efficiently learning the low-dimensional representation
  - Learning these "representative" inducing points $\mathbf{X}_u$
- *Discretize in terms of:* $\mathbf{X}_u$

Sparse GP-LVM provides a coherent way for both dimensionality reduction and subsequent discretization!

Model Overview



Data of all variables $\{Y_{i=1,\ldots,N}\}$ that describe a set of grasps

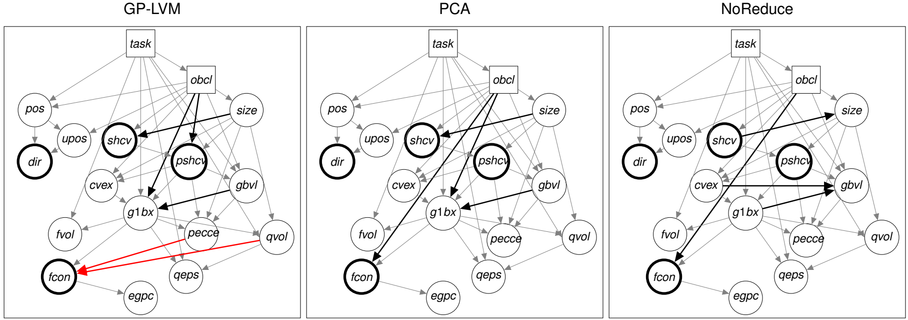
Training Data $\mathbf{X} = [\mathbf{O}; \mathbf{A}; \mathbf{C}; \mathbf{T}]$

- Object approximation – box decomposition
- Grasp hypotheses – approaching box facades
- Feature extraction – $[\mathbf{O}; \mathbf{A}; \mathbf{C}]$
- Manual task labeling – $[\mathbf{T}]$ by human | All in BADGr (Huebner, 2010)

Three Experiments:

1) Structure learning.
2) Task classification, T.
3) Prediction of grasp final configuration, *fcon*.

1) Structure learning



- Structure learning successfully reveals the conditional dependencies among a large pool of variables.
- Main differences in structures lie in the connections to *fcon*:
  - GP-LVM reveals dependencies such as *pecce* $\rightarrow$ *fcon*, while others do not.

2) Task classification, $P(T|*)$

$\downarrow \rightarrow$: *Hand-over, Pouring, Tool-use*

	O			O, A			O, A, C		
GP-LVM	0.70	0.12	0.17	0.74	0.09	0.17	0.86	0.04	0.10
	0.12	0.88	0.00	0.05	0.95	0.00	0.15	0.85	0.00
	0.16	0.00	0.84	0.03	0.00	0.97	0.04	0.01	0.95

- Classification performances improve as more features are observed.

2) Task classification, $P(T|*)$

$\downarrow\rightarrow$: *Hand-over, Pouring, Tool-use*

GP-LVM (O, A, C)

0.86	0.04	0.10
0.15	0.85	0.00
0.04	0.01	0.95

PCA (O, A, C)

0.87	0.06	0.07
0.26	0.73	0.01
0.13	0.00	0.87

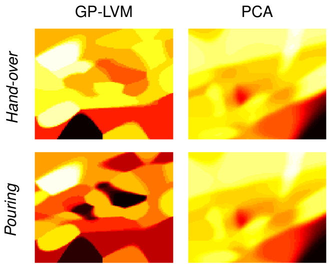
NoReduce (O, A, C)

0.79	0.11	0.10
0.32	0.51	0.17
0.29	0.14	0.57

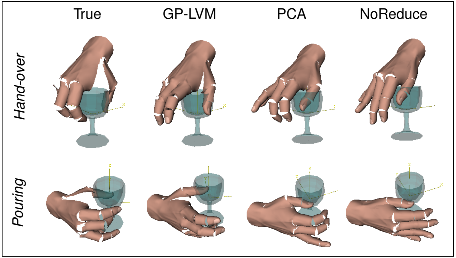
- Classification performances improve as more features are observed.
- GP-LVM discretization scheme results in the best classification performances in most observation conditions.

3) Prediction of grasp final configuration, $P(fcon|T, O)$

Likelihood Maps of *fcon* in 2D Latent Space
$P(fcon|T, O = glass)$



Most Probable Prediction of *fcon* in 20D Space
$P(fcon|T, O = glass)$



- GP-LVM scheme results in clearly different likelihood maps between the two tasks, whereas PCA does not.
- This indicates the GP-LVM scheme has captured the potential constraining effects of the tasks on *fcon*.
- GP-LVM scheme enables the most natural, intuitive data reconstruction in the original observation space *fcon*.

Conclusions

- Sparse GP-LVM-based discretization method provides a compact, efficient representation of high-dimensional data.
- This method allows fast and effective structure learning for Bayesian networks.
- The resulting composite modeling system is fully generative, and allows better task classification and data reconstruction in original observation spaces.

Future Work

- Embodiment-Specific Representation of Robot Grasping using Graphical Models and Latent-Space Discretization – IROS 2011 submission.
- Integration with real vision systems: combining object categorization with task constrained grasping – IROS 2011 submission.
- Integration with real robot platforms: task-constrained grasp online adaptation based on stability measure using haptic sensory feedback (Bekiroglu *et al.*, 2011), which will be presented in Session ThA111.
- Introducing prior on inducing points in sparse GP-LVM, which will be presented in Workshop on Friday: 'Manipulation Under Uncertainty'.

Thanks

Thanks for your attention!

EU project IST-FP7-IP-215821

Swedish Foundation for Strategic Research

Literature

Bekiroglu, Y., Huebner, K., and Kragic, D. (2011). Integrating Grasp Planning with Online Stability Assessment using Tactile Sensing. In *Proceedings of the IEEE International Conference on Robotics and Automation*. To appear.

Huebner, K. (2010). BADGr - A Toolbox for Box-based Approximation, Decomposition and GRasping. In *IEEE International Conference on Intelligent Robots and Systems, Workshop on Grasp Planning and Task Learning by Imitation*.

Lawrence, N. D. (2005). Probabilistic Non-linear Principal Component Analysis with Gaussian Process Latent Variable Models. *Journal of Machine Learning Research*, **6**, 1783–1816.

Song, D., Huebner, K., Kyrki, V., and Kragic, D. (2010). Learning Task Constraints for Robot Grasping using Graphical Models. In *IEEE International Conference on Intelligent Robots and Systems*. To appear.