Contextual Modeling with Labeled Multi-LDA
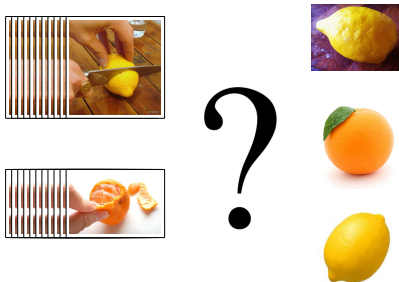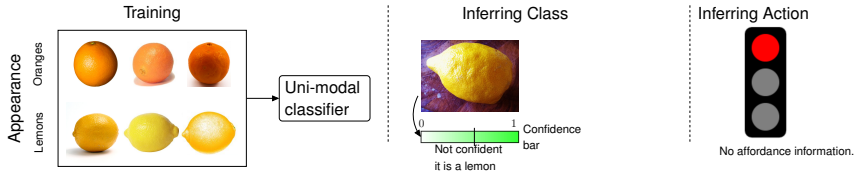
C. Zhang, D. Song and H. Kjellström

CVAP/CAS,KTH Royal Institute of Technology
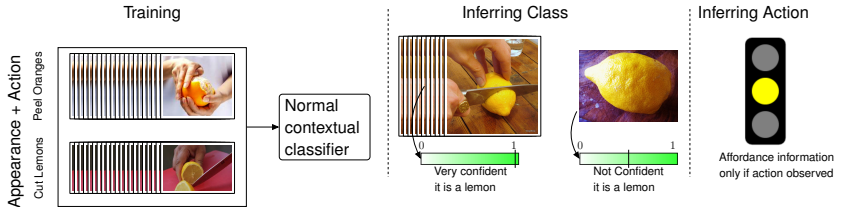
What object classes are observed?
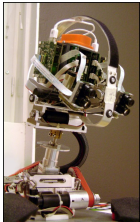
Training

Appearance
Oranges
Lemons

Uni-modal classifier

Inferring Class

0   1   Confidence bar

Not confident
it is a lemon

Inferring Action

No affordance information.

Training

Inferring Class

Inferring Action

Appearance + Action

Cut Lemons   Peel Oranges

Normal contextual classifier

0 — 1
Very confident
it is a lemon

0 — 1
Not Confident
it is a lemon

Affordance information
only if action observed

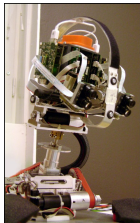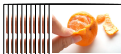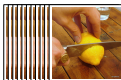**Robot observing objects**   **What object classes are observed?**   **What actions are afforded?**

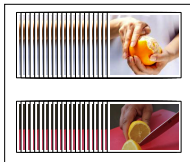# Contextual Modeling

Robot observing objects

What object classes are observed?

What actions are afforded?
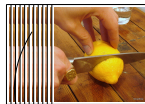
Training

Inferring Class

Inferring Action

Appearance + Action
Cut Lemons   Peel Oranges

LM-LDA

0 —————— 1
Very confident
it is a lemon

0 —————— 1
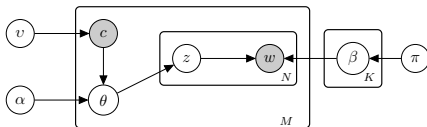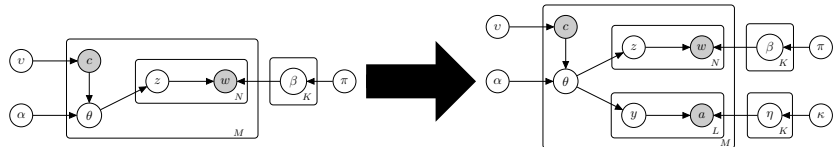Confident
it is a lemon

Affordance information
encoded in the model!

Labeled-LDA

1

---

[1] L. Fei-Fei and P. Perona, "A bayesian hierarchical model for learning natural scene categories," in CVPR, 2005.

Labeled-LDA

Labeled Multi-LDA

Data: Activities of Daily Living (ADL) dataset.[2]

SVM+bag AO: 38.31%   SVM+ pyramid AO: 38.58%

Train STIP
Test STIP
30.49%

Train both
Test STIP
38.23%

Train both
Test both
46.04%

[2] H. Pirsiavash and D. Ramanan, "Detecting activities of daily living in first-person camera views", in CVPR, 2012.
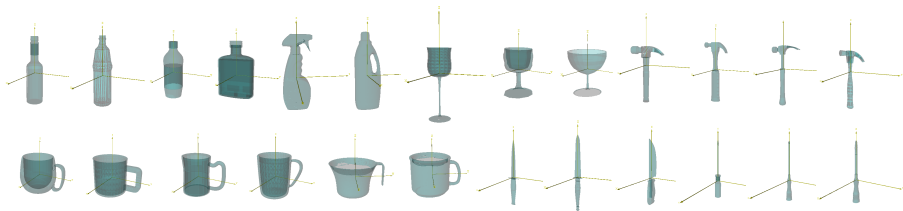
All the instances used in the Appearance and Task-Orientated Grasp experiment: 6 Bottles, 3 Glasses, 4 Hammers, 6 Mugs, 3 knives and 3 Screwdrivers.[3]

---

[3] D. Song, K. Huebner, V. Kyriki, D. Kragic, "Learning task constraints for robot grasping using graphical models", in IROS, 2010

	Btl	Gls	Hm	Knf	Mug	SD
Bottle	1	0	0	0	0	0
Glass	0	0.33	0.33	0	0.33	0
Hammer	0	0	1	0	0	0
Knife	0	0	0	0.67	0	0.33
Mug	0	0	0	0	1	0
ScrewDriver	0	0	0	0.33	0	0.67

Train SIFT,Test SIFT: 77.78%

	Btl	Gls	Hm	Knf	Mug	SD
Bottle	1	0	0	0	0	0
Glass	0	0.33	0.33	0	0.33	0
Hammer	0	0	1	0	0	0
Knife	0	0	0	0.67	0	0.33
Mug	0	0	0	0	1	0
ScrewDriver	0	0	0	0	0	1

Train both,Test SIFT: 83.33%

	Btl	Gls	Hm	Knf	Mug	SD
Bottle	1	0	0	0	0	0
Glass	0	0.67	0.33	0	0	0
Hammer	0	0	1	0	0	0
Knife	0	0	0	0.83	0	0.17
Mug	0	0	0	0	1	0
ScrewDriver	0	0	0	0	0	1

Train both,Test both: 91.67%
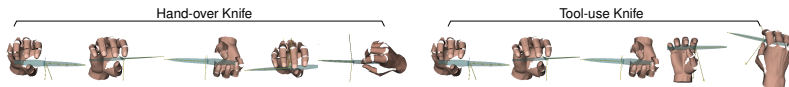
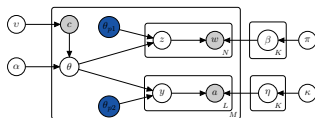The most probable grasps on an unseen knife with task-oriented grasp Hand-over and Tool-use.
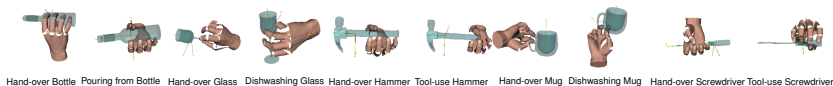
Grasp examples on other objects.

Future Work:

1. Factorized - LMLDA[4]

2. Efficient Learning Algorithm

`http://www.csc.kth.se/~chengz/TopicModelCode.html`[5]

---

[4] C. Zhang, C. H. Ek, A. Damianou, and H. Kjellström, "Factorized topic models," in International Conference on Learning Representations, 2013.

[5] C. Zhang, C.H. Ek, X. Gratal, F.T. Pokorny and H. Kjellström, "Supervised Hierarchical Dirichlet Processes with Variational Inference", In ICCV inferPGM, 2013 .

Contact:

`{chengz, hedvig}@kth.se`

Thanks & Questions