Distributions of Diverse Features For Single- and Multiple-View Robot Vision

Justus Piater, University of Innsbruck

Table of Contents

1. Introduction

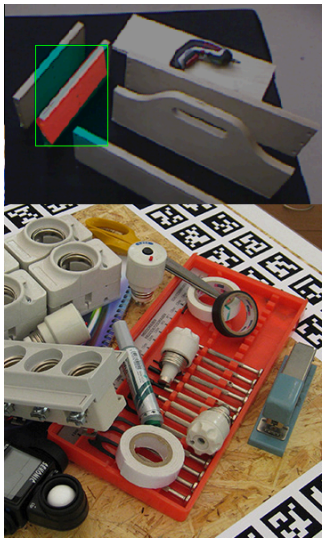
1.1. Pose — A Robotics Perspective

Grasping requires perception:

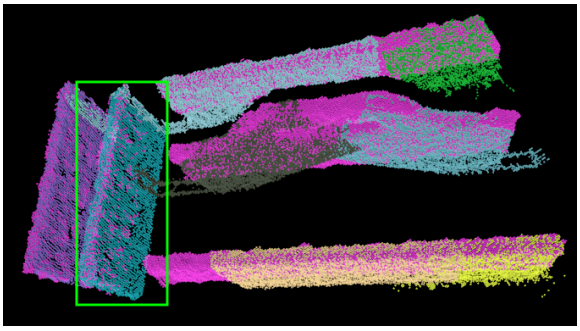
- Object location, and
  - nothing more (wrap grasp)
  - Or: shape
  - Or: identity + pose (the focus of this talk; applies also to parts)

1.2. Generative Methods

- Mutually-independent object representations
- Scenarios involving non-standard objects for which neither WWW images nor CAD models exist
- Generalization from very little training data
- Easy handling of missing perceptual modalities
- Robustness to clutter and partial occlusion



1.3. Part One: View-Based, Probabilistic Models



Both edges and surfaces are useful.

- RGB is good for edges.
- D is good for (some) surfaces.

1.4. Part Two: 3DPose 2015 Challenges

Imperial College Multi-Object Pose Estimation Challenge

University of Birmingham Highly Occluded Object Challenge (partial results)

Dresden Articulated Object and TU Graz & EPFL Tracking Challenges: not relevant to us

Dresden Occluded Object and Prague T-LESS Challenges: no time

2. Part One: View-Based, Probabilistic Models

2.1. Credits

Damien Teney

Özgür Erkent

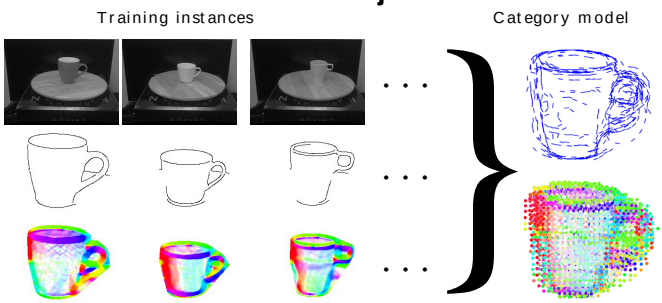
Dadhichi Shukla

University of Innsbruck

Intelligent and Interactive Systems

https://iis.uibk.ac.at/

2.2. View-based Probabilistic Object Models



Uses:

- Local pieces of edge
- Local image gradients at larger scale
- …

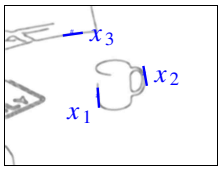
Can do:

- Detect objects at category level
- Estimate continuous pose
- Generate arbitrary views
- Generate a 3D model

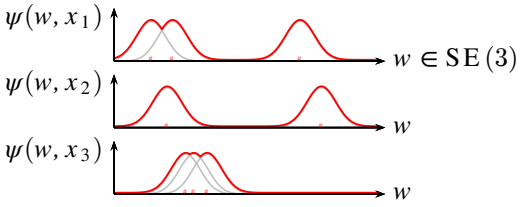
[Teney and Piater 2014]

2.3. Edge Observations Vote For Poses

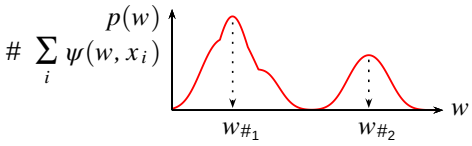
(a)



(b)

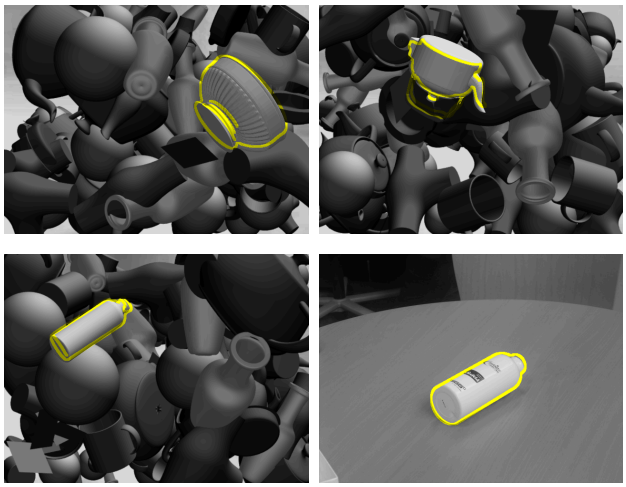


(c)

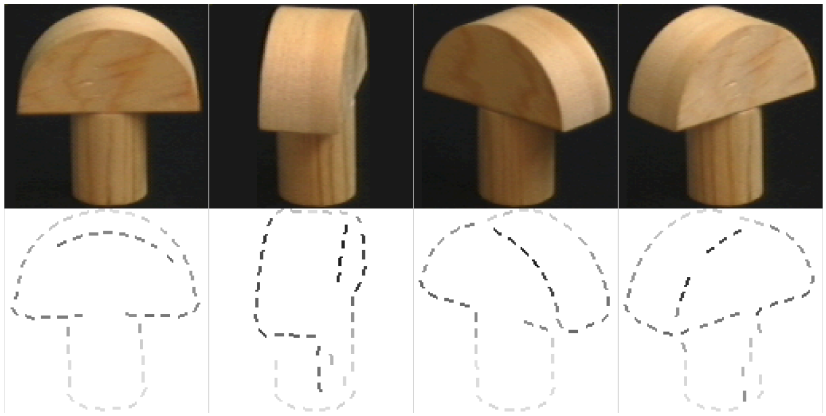


(d)

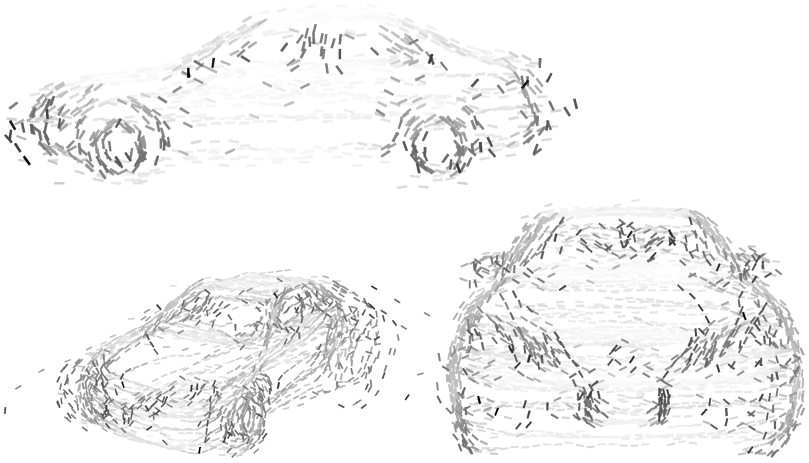
2.4. Some Illustrative Results



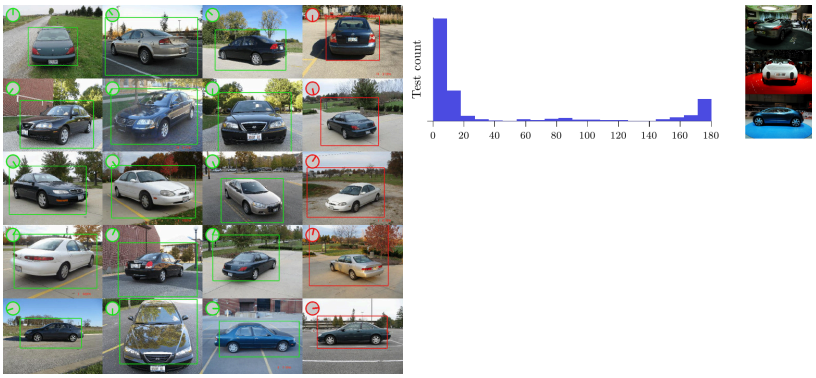
2.5. Feature Weighting According to Discriminative Power



2.6. Feature Weighting According to Discriminative Power

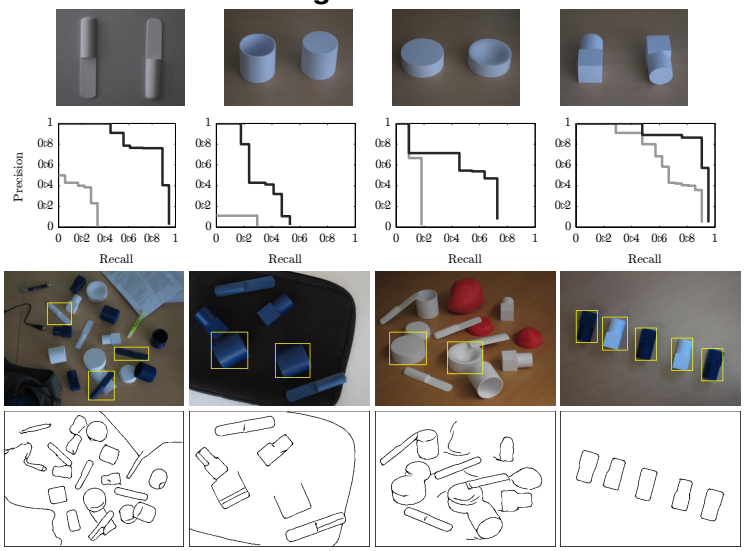


2.7. Results With Continuous Pose Interpolation

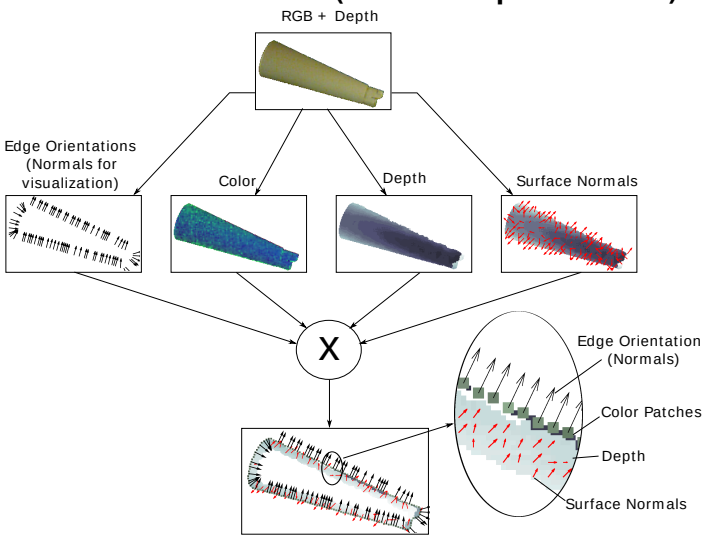


[Teney and Piater 2013]

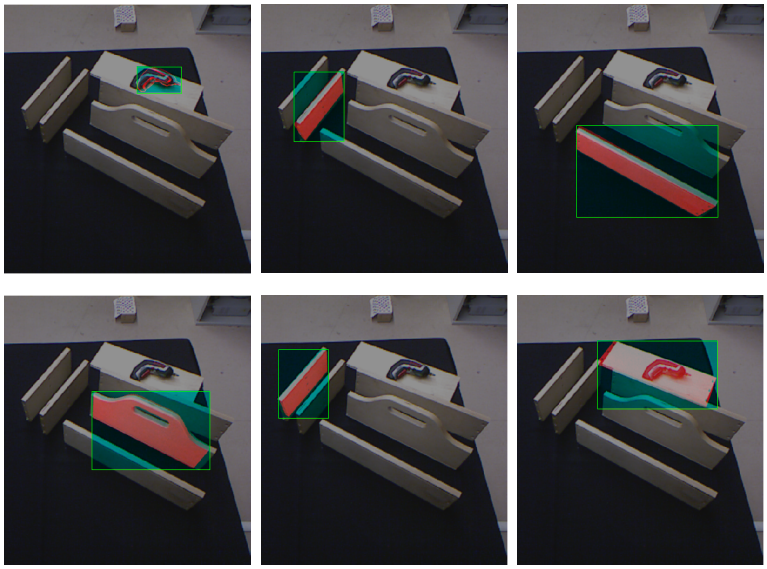
2.8. The Power of Shading



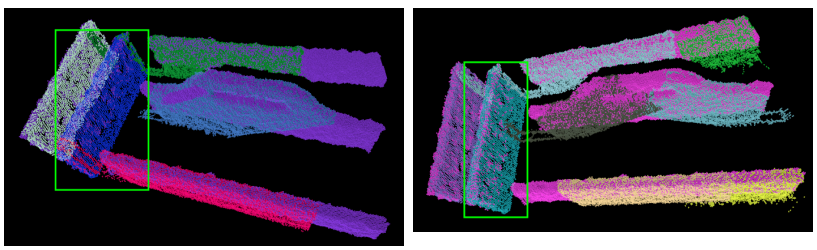
2.9. Features from RGB + D (at their respective best)



2.10. Detection + Pose Results



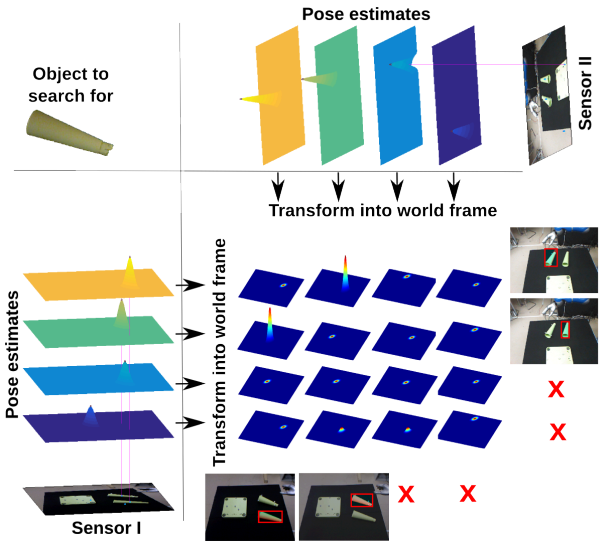
2.11. Baseline Detection + Pose Results



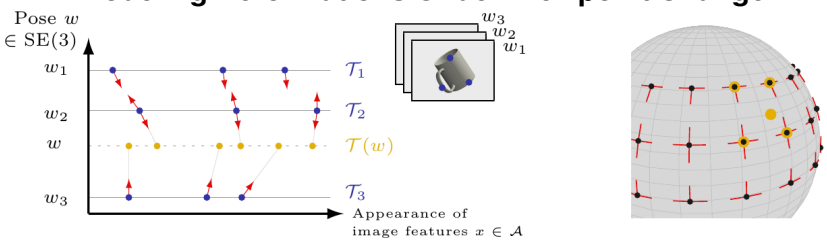
[Papazov and Burschka 2011] [Aldoma et al. 2012]

2.12.

2.13. Multi-View Integration



2.14. Modeling Deformations Under Viewpoint Change

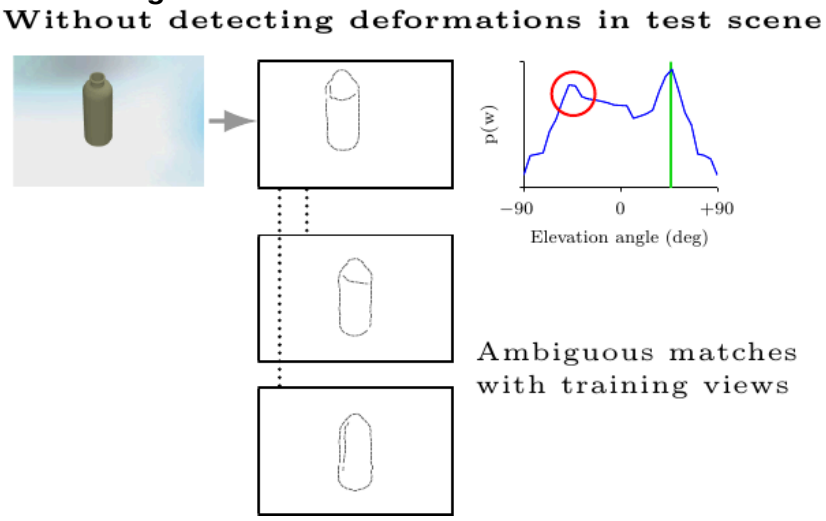


- Generate intermediate views for improved matching
- Disambiguate pose using motion parallax

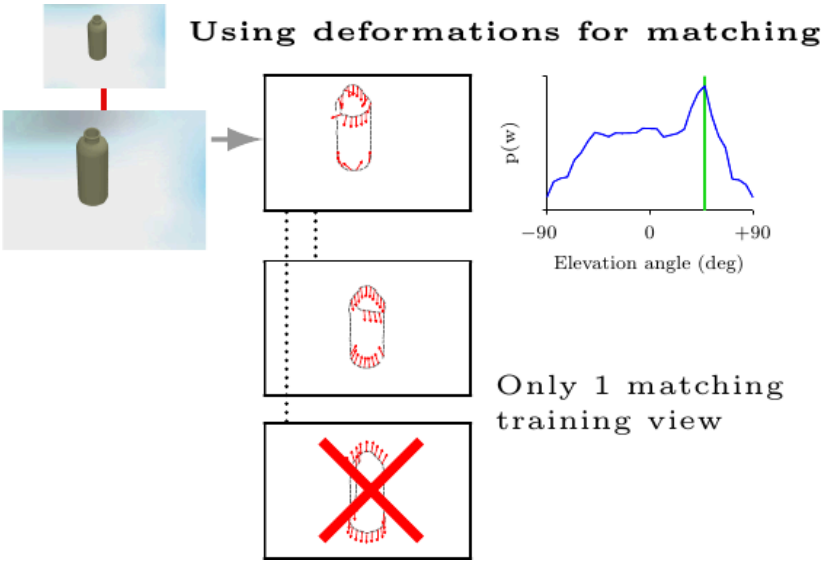
[Teney and Piater 2013]

2.15. View Interpolation

2.16. Ambiguous Pose

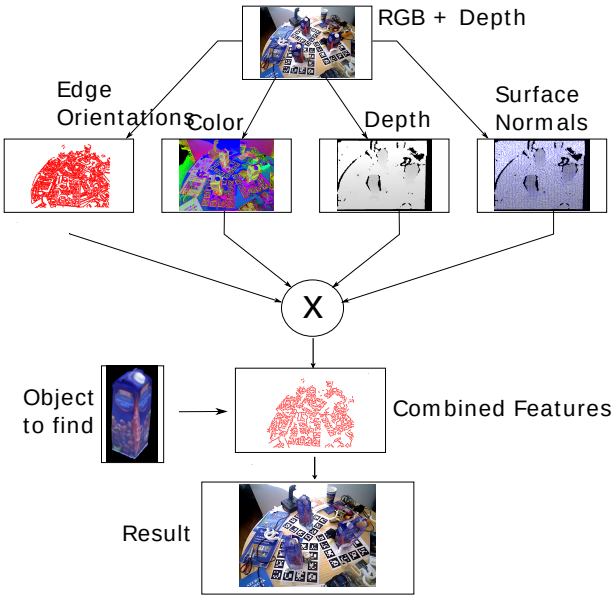


2.17. Disambiguated Pose



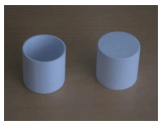
3. Part Two: 3DPose 2015 Challenges

3.1.

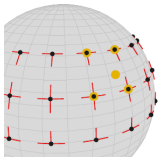


3.2. Not Used (not yet reimplemented)

Feature Weighting

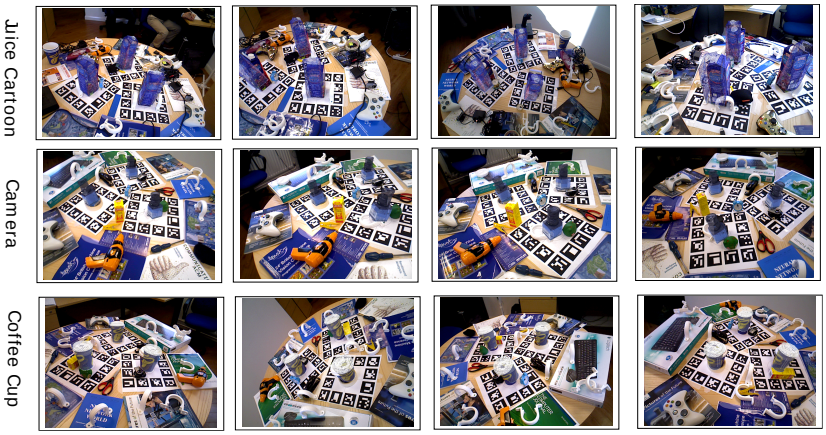


Large-scale gradient features to take advantage of shading

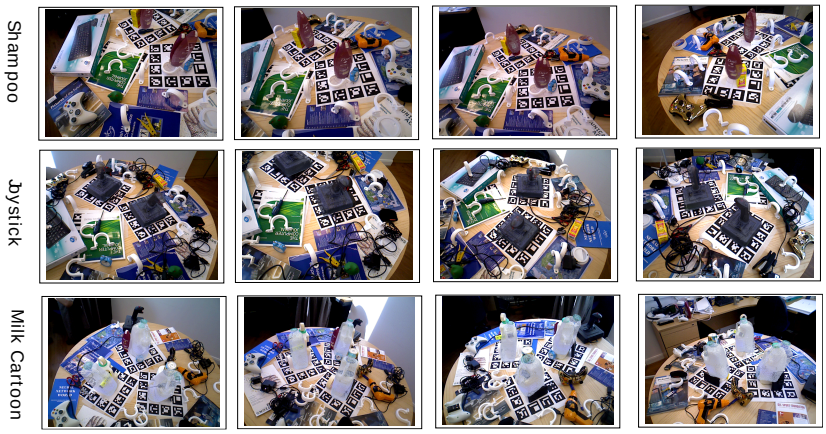


View/pose interpolation for pose refinement and disambiguation

3.3. ICVL Challenge: Results (A)



3.4. ICVL Challenge: Results (B)



3.5.

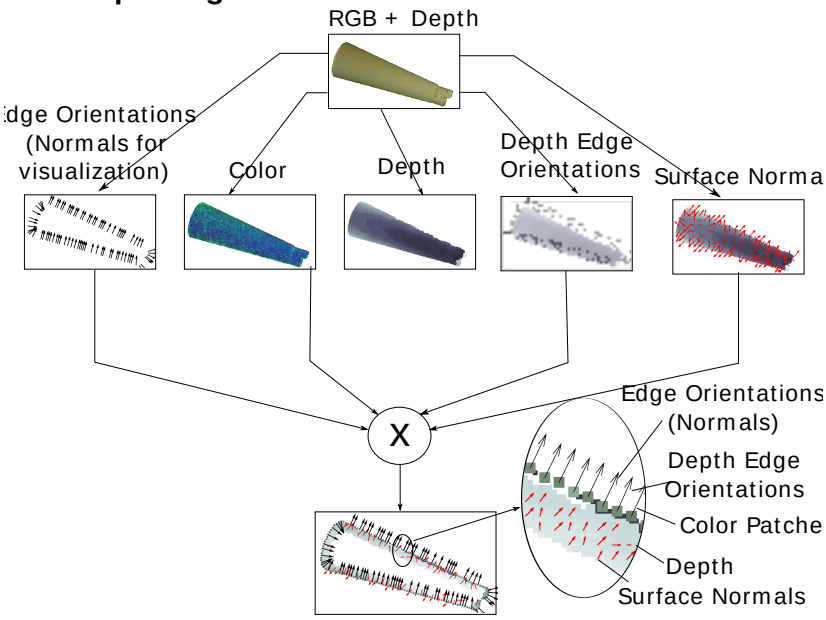
3.6.

3.7.

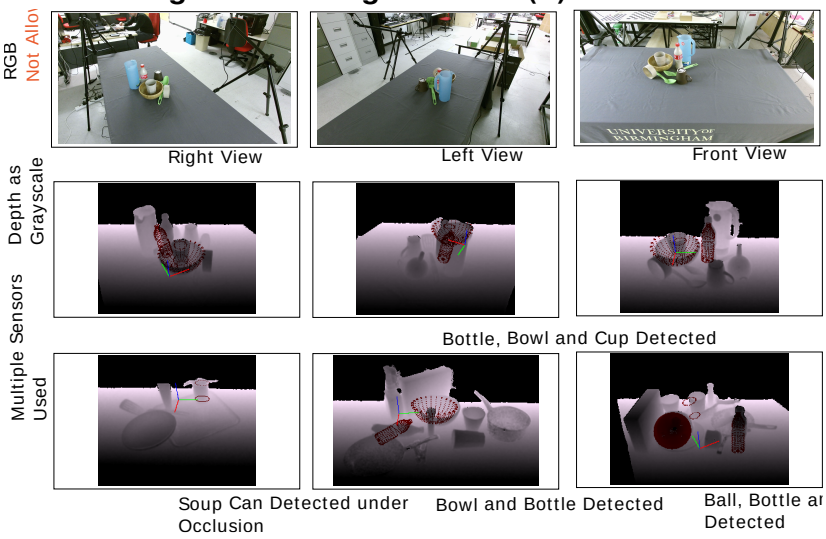
3.8.

3.9.

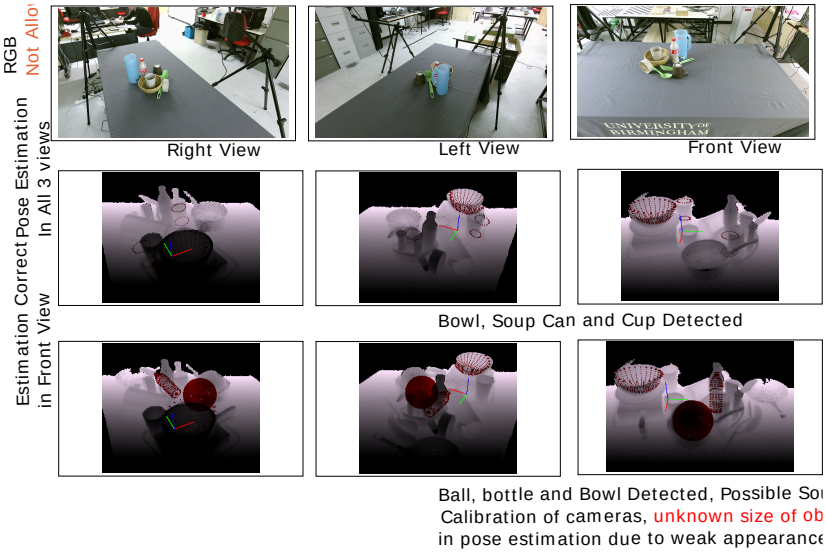
3.10. Depth Edge Features



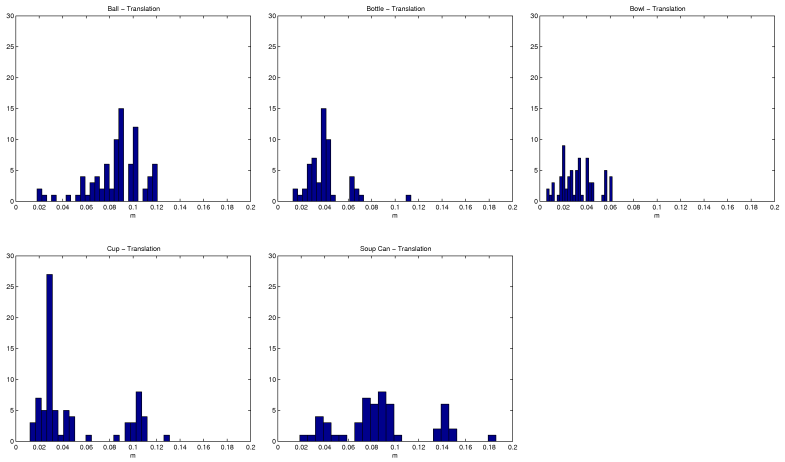
3.11. Birmingham Challenge: Results (A)



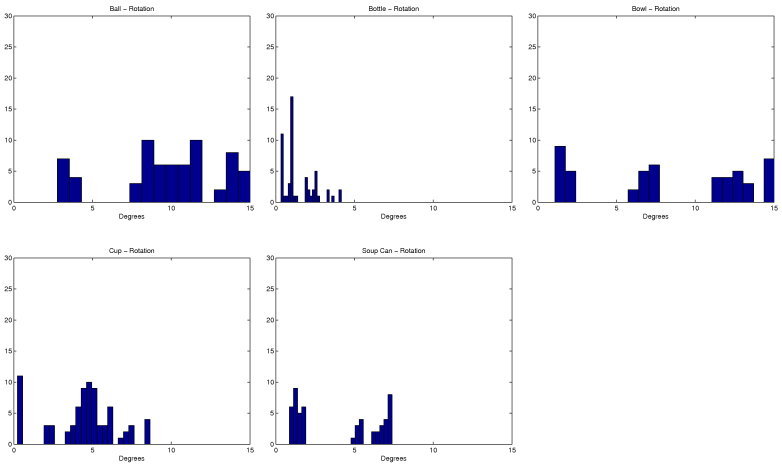
3.12. Birmingham Challenge: Results (B)



3.13. Birmingham Challenge: Translation Errors



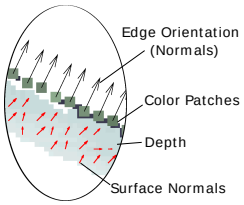
3.14. Birmingham Challenge: Rotation Errors



3.15. Conclusions

- Significantly better results expected from full method (pose interpolation, feature weighting, large-scale gradients)
- Category-level pose estimation of imprecisely known objects using generative models
- Deployed and further developed under EU-FP7-ICT 3rdHand for pose estimation of diverse objects and human hands
- Visible edges are a high-fidelity cue to pose

  but not visually distinctive ⇒ learn features
- Extension to tracking under investigation



3.16. The IIS Group (Approximation)

4. References

4.1. References

Bibliography

A. Aldoma, F. Tombari, R. Rusu, M. Vincze, "OUR-CVFH – Oriented, Unique and Repeatable Clustered Viewpoint Feature Histogram for Object Recognition and 6DOF Pose Estimation[1]". *Joint 34th DAGM and 36th OAGM Symposium*, pp. 113–122, 2012.

C. Papazov, D. Burschka, "An Efficient RANSAC for 3D Object Recognition in Noisy and Occluded Scenes". *Proceedings of the 10th Asian Conference on Computer Vision*, pp. 135–148, 2011.

D. Teney, J. Piater, "Continuous Pose Estimation in 2D Images at Instance and Category Levels[2]". *Tenth Conference on Computer and Robot Vision*, pp. 121–127, 2013.

D. Teney, J. Piater, "Modeling Pose/Appearance Relations for Improved Object Localization and Pose Estimation in 2D images[3]". *6th Iberian Conference on Pattern Recognition and Image Analysis*, pp. 59–68, 2013.

D. Teney, J. Piater, "Multiview feature distributions for object detection and continuous pose estimation[4]". *Computer Vision and Image Understanding* 125, pp. 265–282, 2014.

[1] http://dx.doi.org/10.1007/978-3-642-32717-9_12
[2] https://iis.uibk.ac.at/public/papers/Teney-2013-CRV.pdf
[3] https://iis.uibk.ac.at/public/papers/Teney-2013-IbPRIA.pdf
[4] https://iis.uibk.ac.at/public/papers/Teney-2014-CVIU.pdf