



Context-driven Clustering by Multi-class Classification in an Active Learning Framework



Martin Godec, Sabine Sternig,
Peter Roth and Horst Bischof

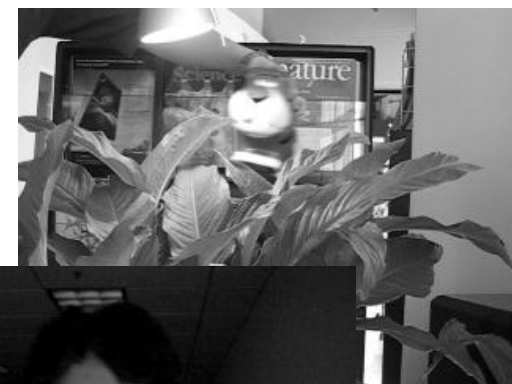


Institute for Computer Graphics and Vision
Graz University of Technology



Object Detection and Tracking

- Formulated as binary classification task
 - Train foreground vs. background
- Consider scene complexity
 - Clutter
 - Changing background
 - **Multi-modalities**
 - ...



Object Detection and Tracking

- Do we need complex binary classifiers?
 - More complicated/slower training
 - Slow evaluation
 - Overfitting
 - Differently complex classes
 - Complex task needs complex classifier?

GOAL: Adapt classifier complexity to current scene!

Split complex classes...

- Describe class by several clusters
 - Multi-pose learning [3]
 - Multi-classifier learning [11]
 - Cluster Boosted Tree [22]
 - ...
- But...
 - We need adaption over time
 - We do not want to specify number of clusters

Our Approach

- On-line clustering of classes
 - Depending on scene complexity
 - Without restricting number of clusters
 - Using multi-class classifications
- Using task-specific context knowledge
 - Perform active learning

Outline

- Motivation
- **Context-driven On-line Clustering**
 - Creating “virtual classes”
 - Active Learning Framework
 - On-line Multi-class GradientBoost
- Applications
- Conclusion

Context-driven On-line Clustering

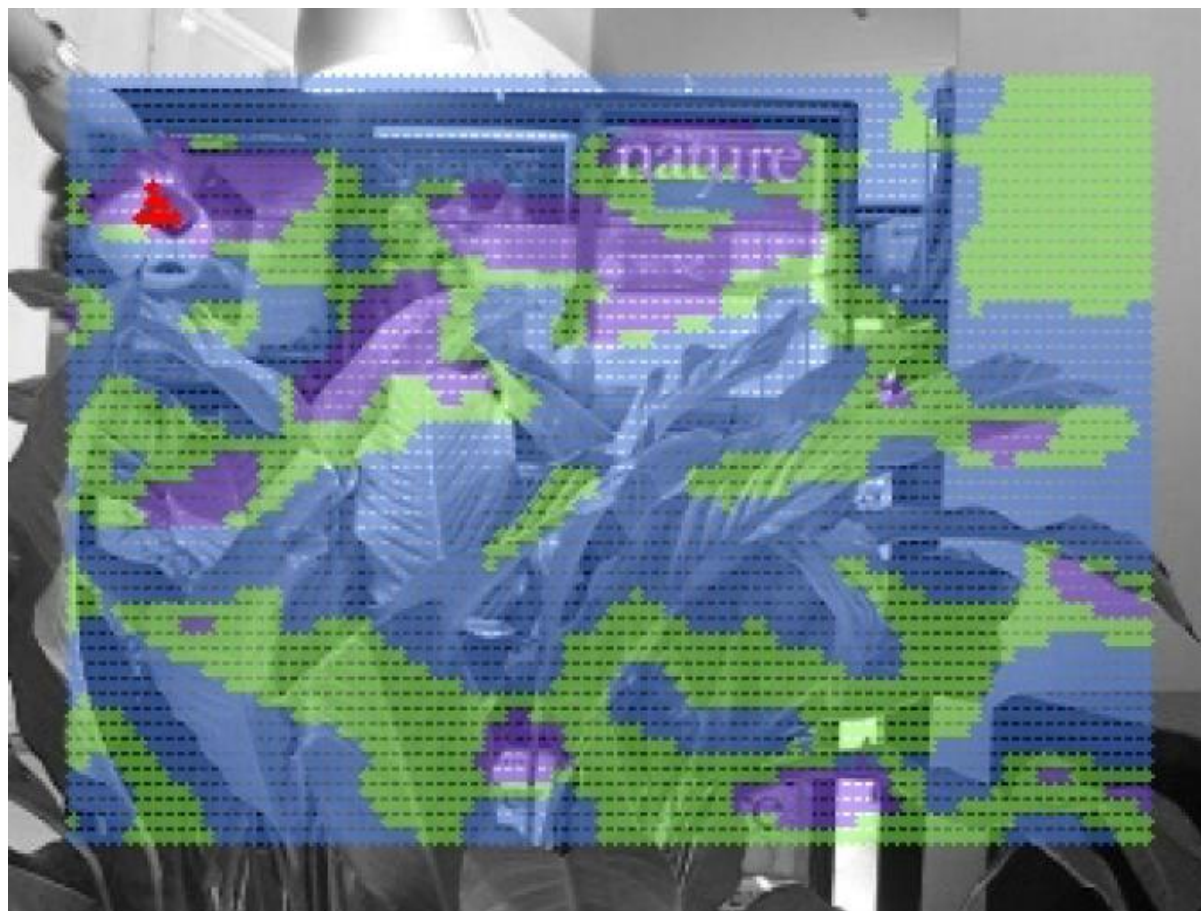
- Virtual classes
 - Describes cluster within a real class (FG/BG)
 - Several virtual classes Instead of one complex class
 - Clustering within feature space (\neq segmentation)
 - Embedded within multi-class classifier
 - Add new virtual classes if current model not sufficient

Describe multi-modalities of a class

Creating Virtual Classes

```
for each sample  $a_i \langle x_i, y_i \rangle$   
   $y = \text{eval}(c; a_i)$   
  if  $y == y_i$  then  
    // Add new virtual class  
    update( $c; a_i; Y_{v++}$ )  
  else  
    // Update classifier  
    update( $c; a_i; y$ )  
  end if  
end for
```

On-line Clustering Result



Active Learning Framework

- Components of active learning (c; s; T; L; U)
 - c: classifier
 - s: sampling function (selects most valuable samples)
 - T: Teacher (assigns labels to selected samples)
 - L: set of labeled samples
 - U: set of unlabeled samples

Task-specific sampling (s) and teacher (T)

On-line Multi-class GradientBoost

- Based on On-line GradientBoosting [Leistner 09]
 - On-line Boosting for Feature Selection
 - Allows use of robust loss-function (e.g., Logit)
 - Using on-line histograms
- Multi-class extension
 - Symmetric multiple logistic transformation
 - Normalization for asymmetric datasets
 - Add new classes on-the-fly

Outline

- Motivation
- Context-driven On-line Clustering
 - Creating “virtual classes”
 - Active Learning Framework
 - On-line Multi-class GradientBoost
- **Applications**
- Conclusion

Applications

- Applied for two different tasks
 - Object Detection
 - Detecting persons or cars from a static camera
 - PETS 2006 and AVSS 2007 datasets
 - Single-target Tracking
 - Tracking-by-Detection with on-line learning
 - 8 publicly available sequences (300-1300 frames)
 - Representation
 - Simple Haar-like features

Detection using Virtual Classes

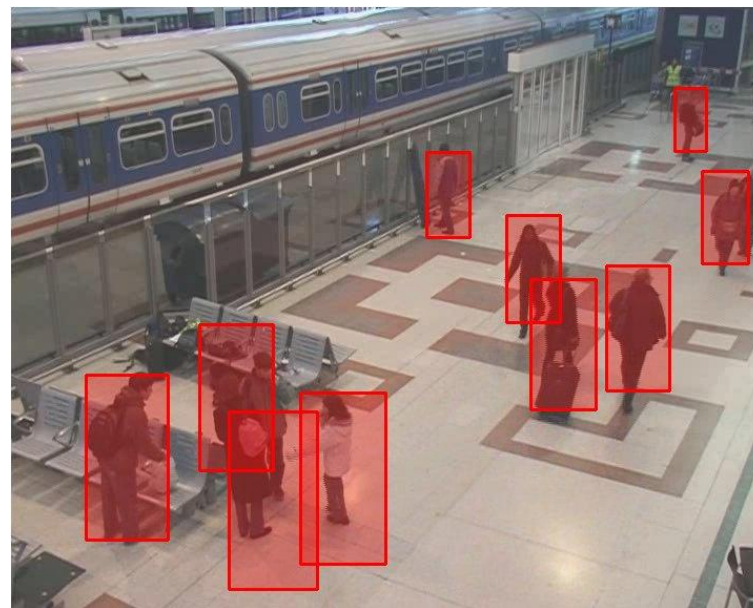
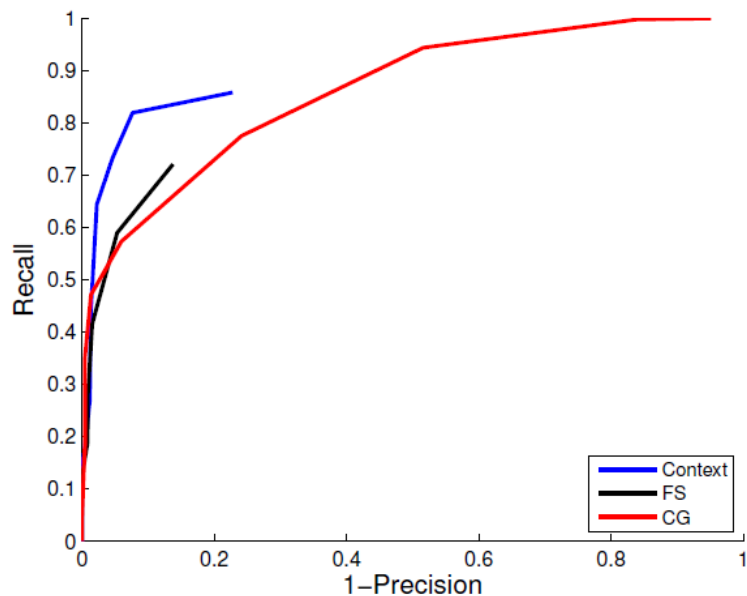
- Context Knowledge
 - Static camera
 - Background model is target-free
- Active Learning
 - s : sample from background model
 - T : background model belong to negative class
 - Create virtual classes during runtime

Detection using Virtual Classes

- Compared to
 - deformable part model (FS) [Felzenszwalb 08]
 - Static classifier
 - Part-based
 - Histogram of oriented gradients (HOG)
 - classifier grid (CG) [Roth 09]
 - Large number of simple, position-specific classifiers
 - On-line adaption
 - Haar-like features

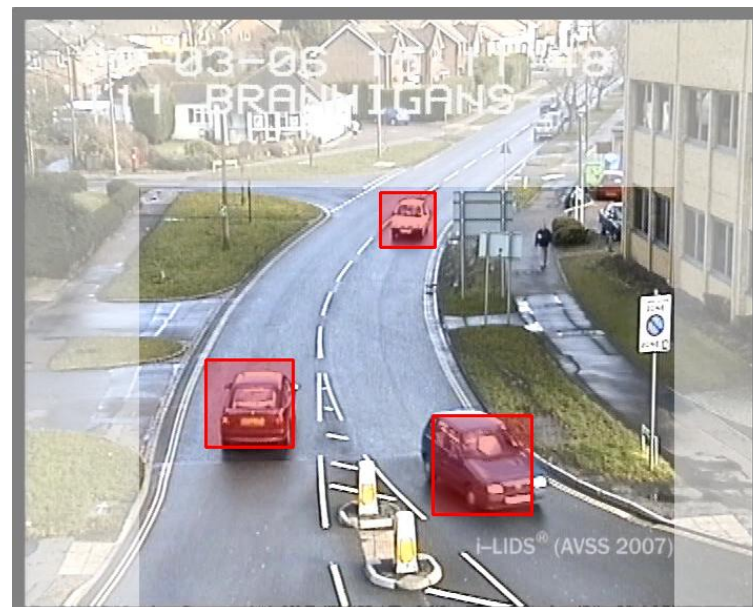
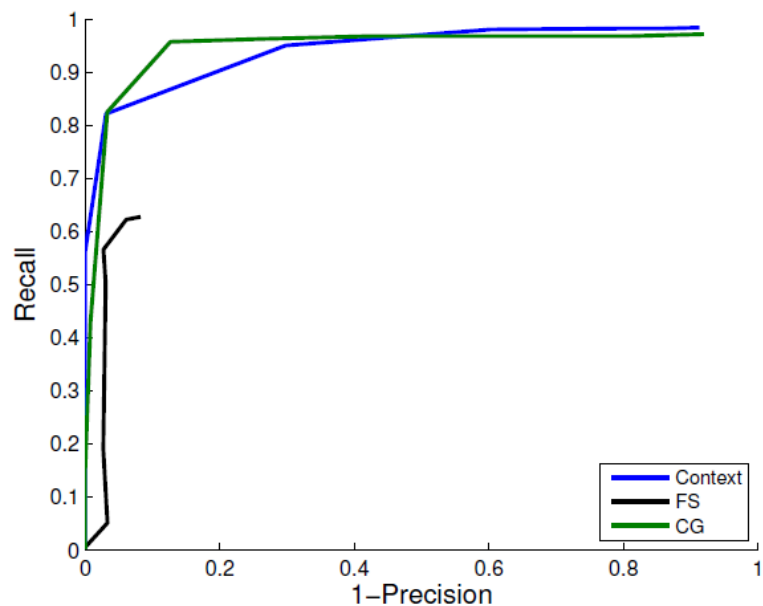
Detection using Virtual Classes

- PETS 2006
 - 308 frames, 720x576



Detection using Virtual Classes

- AVSS 2007 (AVSS PV Hard)
 - 500 frames, 720x576



Tracking using Virtual Classes

- Context Knowledge
 - Only one single target (position is known)
 - Other parts of the scene are target-free
- Active Learning
 - s: sparse random sampling from background
 - T: background models belong to negative class
 - Create virtual classes during runtime

Tracking using Virtual Classes

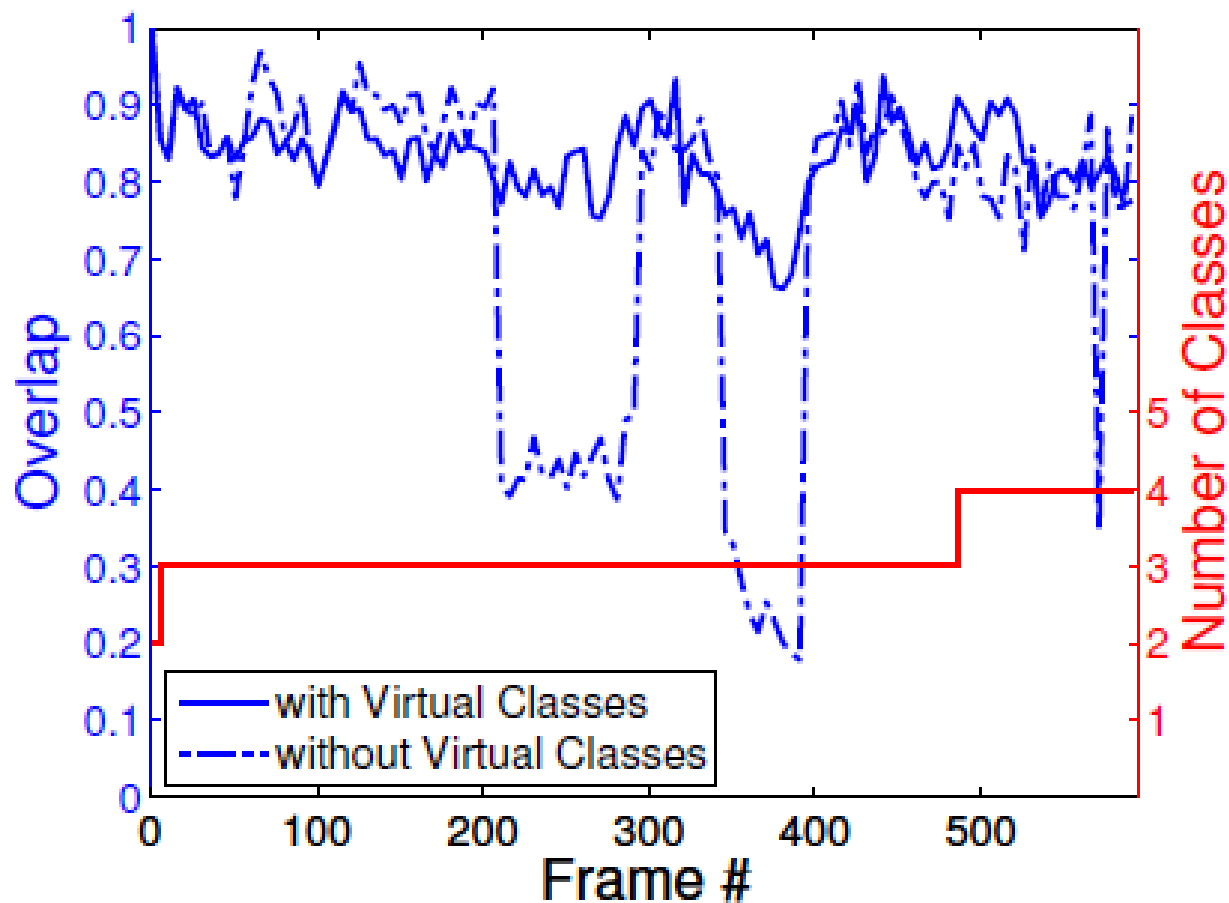
- Compared to
 - On-line multiple-instance boosting (MIL) [Babenko 09]
 - Multiple-instance learning
 - Using Haar-like features
 - Fragments-based tracking (Frag) [Adam 06]
 - Template based
 - Using intensity histograms of parts
 - On-line AdaBoost (OAB) [Grabner 06]
 - On-line boosting for feature selection
 - Using Haar-like features

Tracking using Virtual Classes

Sequence	Context	MIL	Frag	OAB
Sylvester	0.74	0.73	0.74	0.62
Face Occlusion 1	<i>0.93</i>	0.73	0.94	0.63
Face Occlusion 2	0.89	0.81	0.51	0.81
Girl	0.84	0.68	0.73	0.57
Tiger 1	0.65	0.65	0.26	0.33
Tiger2	<i>0.49</i>	0.69	0.22	0.41
David	<i>0.71</i>	0.73	0.52	0.39
Coke	<i>0.42</i>	0.47	0.10	0.25

Mean overlap over whole sequences (higher is better)

Tracking using Virtual Classes



Outline

- Motivation
- Context-driven On-line Clustering
 - Creating “virtual classes”
 - Active Learning Framework
 - On-line Multi-class GradientBoost
- Applications
- **Conclusion**

Conclusion

- Simple but efficient clustering algorithm
 - Using scene context
 - On-line adaptive
 - Concept for modeling multi-modalities
 - Automatic adaption to scene complexity
- Evaluation
 - Detection and tracking
 - Achieving state-of-the-art results

Thank you for your attention!

Questions?

Download software at <http://irs.icg.tugraz.at>

More on Multi-class Boosting:

Saffari, Godec, Pock, Leistner, Bischof,

On-line Multi-class LPBoost, CVPR 2010

This work was supported by the FFG project EVis under the FIT-IT program, the FFG project HIMONI under the COMET programme in cooperation with FTW, the FFG project SECRET under the Austrian Security Research Programme KIRAS, and the Austrian Science Fund (FWF) under the doctoral program Confluence of Vision and Graphics W1209.

References

- [1] A. Adam, E. Rivlin, and I. Shimshoni. Robust fragments-based tracking using the integral histogram. In Proc. CVPR, 2006
- [2] S. Agarwal, A. Awan, and D. Roth. Learning to detect objects in images via a sparse, part-based representation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004.
- [3] B. Babenko, P. Dollár, Z. Tu, and S. Belongie. Simultaneous learning and alignment: Multi-instance and multi-pose learning. In Faces in Real-Life Images, 2008
- [4] B. Babenko, M.-H. Yang, and S. Belongie. Visual tracking with online multiple instance learning. In Proc. CVPR, 2009.
- [5] D. A. Cohn, Z. Ghahramani, and M. I. Jordan. Active learning with statistical models. Journal of Artificial Intelligence Research, 1996
- [6] R. Collins, Y. Liu, and M. Leordeanu. On-line selection of discriminative tracking features. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005
- [7] P. Felzenszwalb, D. McAllester, and D. Ramanan. A discriminatively trained, multiscale, deformable part model. In Proc. CVPR, 2008
- [8] J. Friedman. Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 2001

References

- [9] H. Grabner, M. Grabner, and H. Bischof. Real-time tracking via on-line boosting. In Proc. BMVC, 2006
- [10] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton. Adaptive mixtures of local experts. Neural Computation, 1991
- [11] T.-K. Kim and R. Cipolla. Mcboost: Multiple classifier boosting for perceptual co-clustering of images and visual features. In NIPS, 2009
- [12] P. Laskov, C. Gehl, S. Krüger, and K.-R. Müller. Incremental support vector learning: Analysis, implementation and applications. Journal of Machine Learning Research, 2006.
- [13] C. Leistner, A. Saffari A. A., P. M. Roth, and H. Bischof. On robustness of on-line boosting - a competitive study. In Proc. On-line Learning for Computer Vision Workshop, 2009
- [14] M. Li and I. K. Sethi. Confidence-based active learning. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2006
- [15] N. J. B. McFarlane and C. P. Schofield. Segmentation and tracking of piglets. Machine Vision and Applications, 1995
- [16] J.-H. Park and Y.-K. Choi. On-line learning for active pattern recognition. IEEE Signal Processing Letters, 1996

References

- [17] D. A. Ross, J. Lim, R.-S. Lin, and M.-H. Yang. Incremental learning for robust visual tracking. *Intern. Journal of Computer Vision*, 2008. 5
- [18] P. M. Roth, S. Sternig, H. Grabner, and H. Bischof. Classifier grids for robust adaptive object detection. In *Proc. CVPR*, 2009
- [19] A. Saffari, M. Godec, T. Pock, C. Leistner, and H. Bischof. Online multi-class l_pboost. In *Proc. CVPR*, 2010
- [20] A. Saffari, C. Leistner, J. Santner, M. Godec, and H. Bischof. On-line random forests. In *Proc. On-line Learning for Computer Vision Workshop*, 2009. 2
- [21] A. Torralba, K. P. Murphy, and W. T. Freeman. Sharing visual features for multiclass and multiview object detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2007
- [22] B. Wu and R. Nevatia. Cluster boosted tree classifier for multi-view, multi-pose object detection. In *Proc. ICCV*, 2007
- [23] R. Yan, J. Yang, and A. Hauptmann. Automatically labeling video data using multi-class active learning. In *Proc. ICCV*, 2003