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]
  - 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 (≠ segmentation)
  – Embedded within multi-class classifier
  – Add new virtual classes if current model not sufficient

Describe multi-modalities of a class
Creating Virtual Classes

\[
\text{for each sample } a_i <x_i, y_i> \\
y = \text{eval}(c; a_i) \\
\text{if } y == y_i \text{ then} \\
  // Add new virtual class \\
  \text{update}(c; a_i; y_{v++}) \\
\text{else} \\
  // Update classifier \\
  \text{update}(c; a_i; y) \\
\text{end if} \\
\text{end for}
\]
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

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

Mean overlap over whole sequences (higher is better)
Tracking using Virtual Classes

![Graph showing overlap and number of classes over frames with and without virtual classes.](image)
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://lrs.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


References


References


