Multi-Cue Onboard Pedestrian Detection

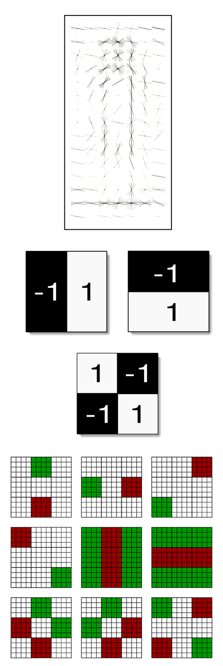
Christian Wojek, Stefan Walk, Bernt Schiele
Computer Science Department, TU Darmstadt, Germany
{wojek, walk, schiele}@cs.tu-darmstadt.de

Objectives

Detect pedestrians from a moving platform

- Exploit motion information
- Leverage complementarity of features
- Evaluate different classifiers
- New datasets with image pairs

Features



- HOG [1]
  - 8×8 pixel cells, 2×2 blocks
  - 9-bin histograms, unsigned gradients
- Haar wavelets [2]
  - 32 and 16 pixel masks
  - horizontal, vertical and diagonal responses
- IMHwd [3]
  - Regularized flow field [4]
  - 9-bin histograms, 8×8 pixel cells

Classifiers

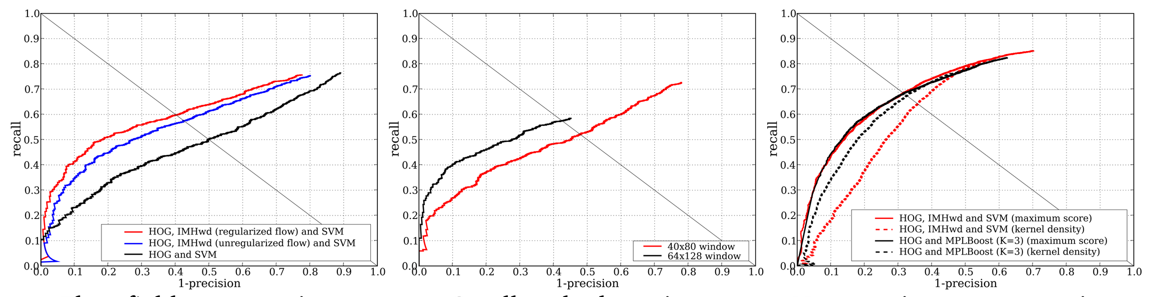
- Linear SVM
- Histogram intersection kernel SVM (HIKSVM) [5]
- AdaBoost
- MPLBoost [6, 7]

Input: $\{x_1, \dots, x_n\}, \{y_1, \dots, y_n\}, y_i \in \{-1, 1\}, K$
Output: K strong boosting classifiers
$H^k(x) = \sum_{t=1}^{T} \alpha_t^k h_t^k(x)$

```
for t = 1 to T do
  for k = 1 to K do
    Compute weights w_i^k = -∂L/∂H_i^k
    Train weak classifier h_t^k using weights |w_i^k|
        h_t^k = argmin_h Σ_i 1(h(x_i) ≠ y_i)|w_i^k|
    Find α_t^k via line search to minimize L(·, H^k, ·)
        α_t^k = argmin_α L(·, H^k + α h_t^k, ·)
    Update strong classifier H^k ← H^k + α_t^k h_t^k.
  end for
end for
```

Implementation details

- **Flow computation:** Regularized flow fields [4] outperform unregularized flow fields when computed on full images
- **Small scales:** Upscaling the tested image performs better than shrinking the detection window
- **Non-maximum suppression:** Maximum score in mode performs better than kernel density



Flow field computation | Small scale detection | Non-maximum suppression

TUD-MotionPairs and *TUD-Brussels* Datasets

TUD-MotionPairs for training with motion features:

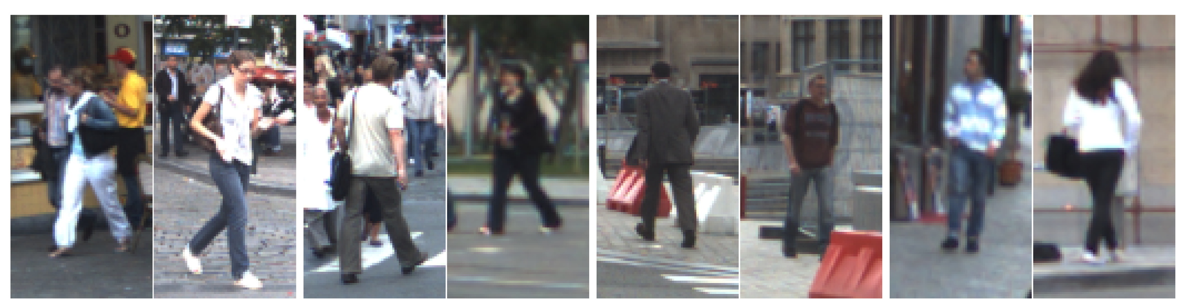
- 1092 image pairs with 1776 pedestrian annotations
- 192 image pairs in negative set
- Multi-view data recorded in pedestrian zones

TUD-Brussels for testing with motion features:

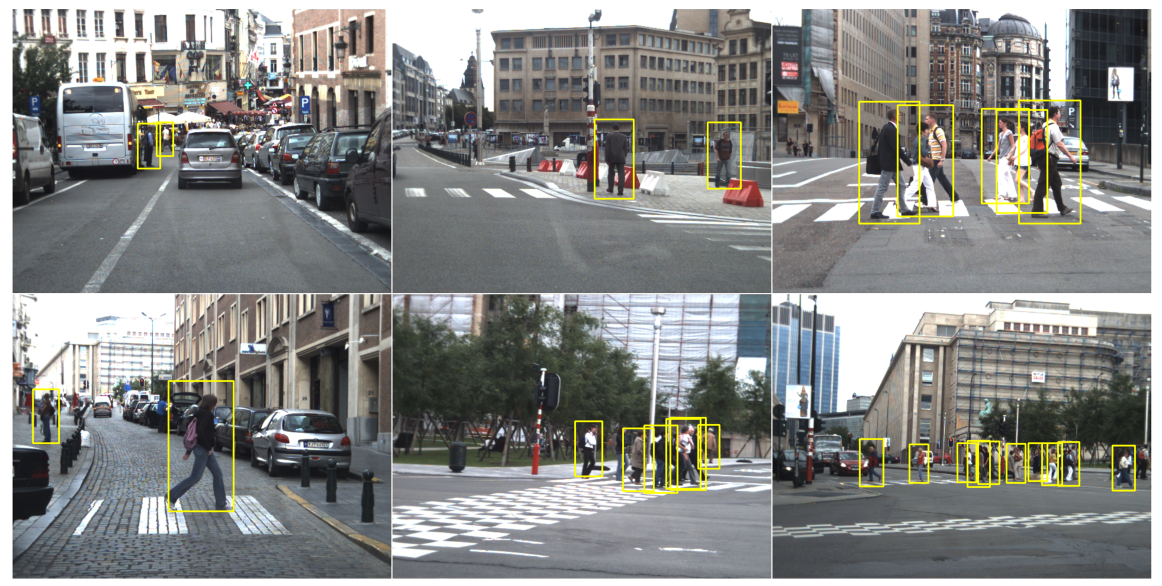
- Recorded in the center of Brussels from a driving car
- 508 image pairs with 640×480 pixel resolution
- 1326 pedestrian annotations

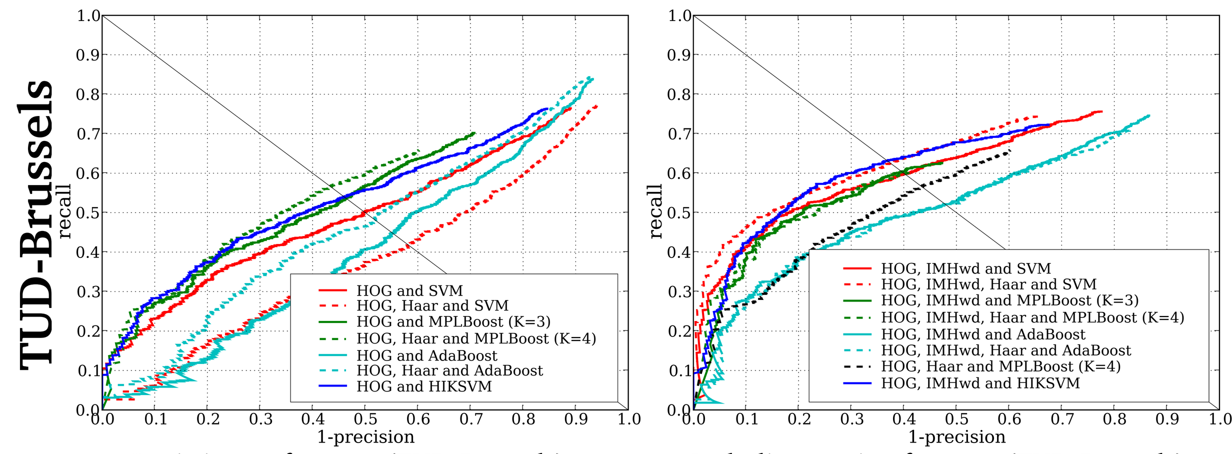
Datasets publicly available at
`http://www.mis.informatik.tu-darmstadt.de`.

MPLBoost Clusters (K=4)

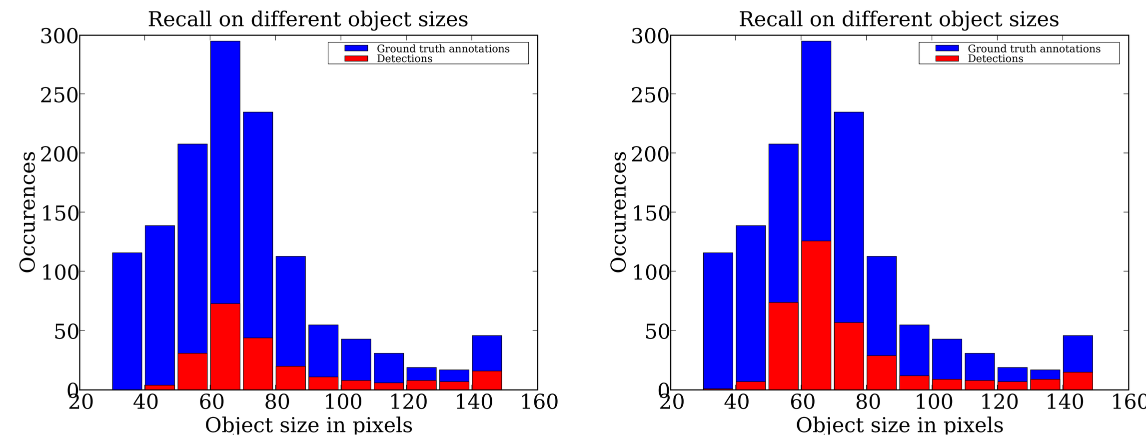


Experiments on *TUD-Brussels*





Static image features (TUD-Brussels) | Including motion features (TUD-Brussels)



Recall vs. Size at 90% precision (static image features) | Recall vs. Size at 90% precision (including motion features)

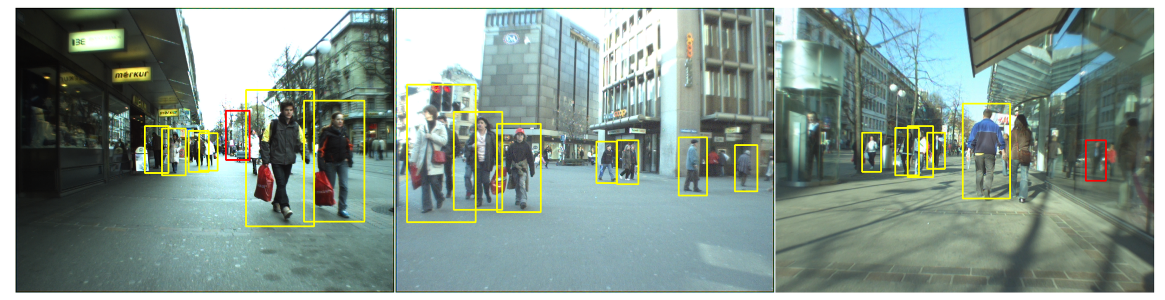
References

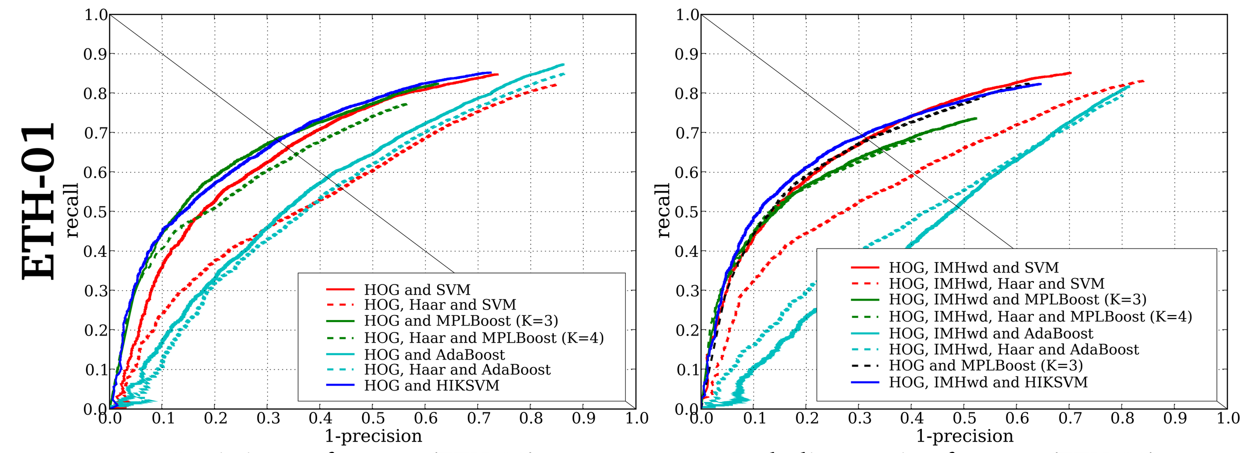
[1] N. Dalal and B. Triggs. Histogram of oriented gradient for human detection. In *CVPR*, 2005.
[2] C. Papageorgiou and T. Poggio. A trainable system for object detection. *IJCV*, 38(1):15–33, 2000.
[3] N. Dalal., B. Triggs., and C. Schmid. Human detection using oriented hist. of flow and appearance. In *ECCV*, 2006.
[4] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime $TV - L^1$ optical flow. In *DAGM*, 2007.
[5] S. Maji, A.C. Berg, and J. Malik. Classification using intersection kernel SVMs is efficient. In *CVPR*, 2008.
[6] B. Babenko, P. Dollár, Z. Tu, and S. Belongie. Simultaneous learning and alignment: Multi-instance and multi-pose learning. In *ECCV Faces in Real-Life Images*, 2008.
[7] T.-K. Kim and R. Cipolla. MCBoost: Multiple classifier boosting for perceptual co-clustering of images and visual features. In *NIPS*, 2008.
[8] A. Ess, B. Leibe, and L. Van Gool. Depth and appearance for mobile scene analysis. In *ICCV*, 2007.

Conclusion

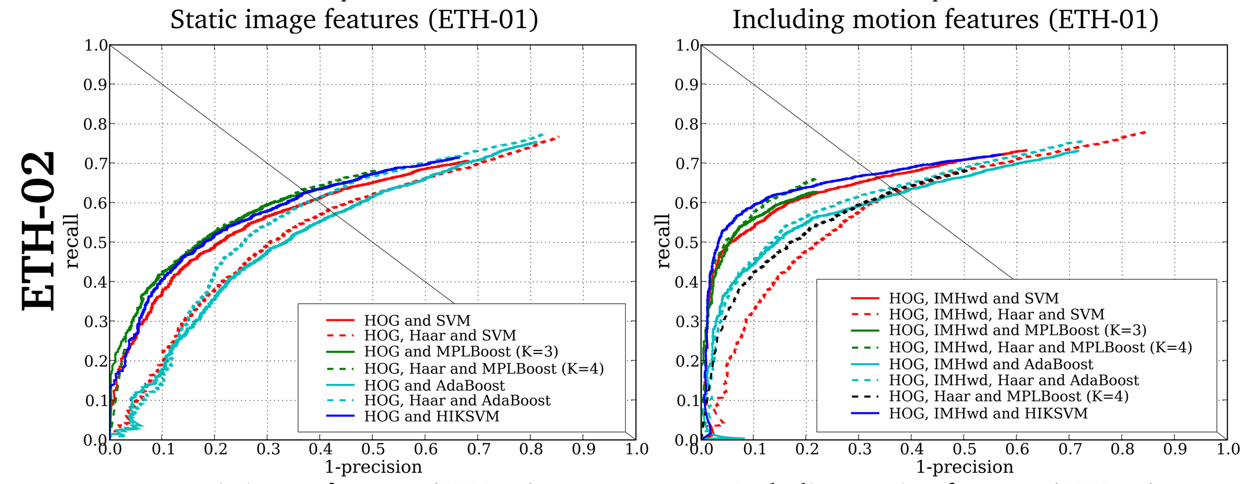
- Motion features improve onboard performance
- HIKSVM often performs best
- MPLBoost consistently outperforms AdaBoost and sometimes HIKSVM
- Haar features can allow for performance improvement

Experiments on *ETH-Pedestrians*

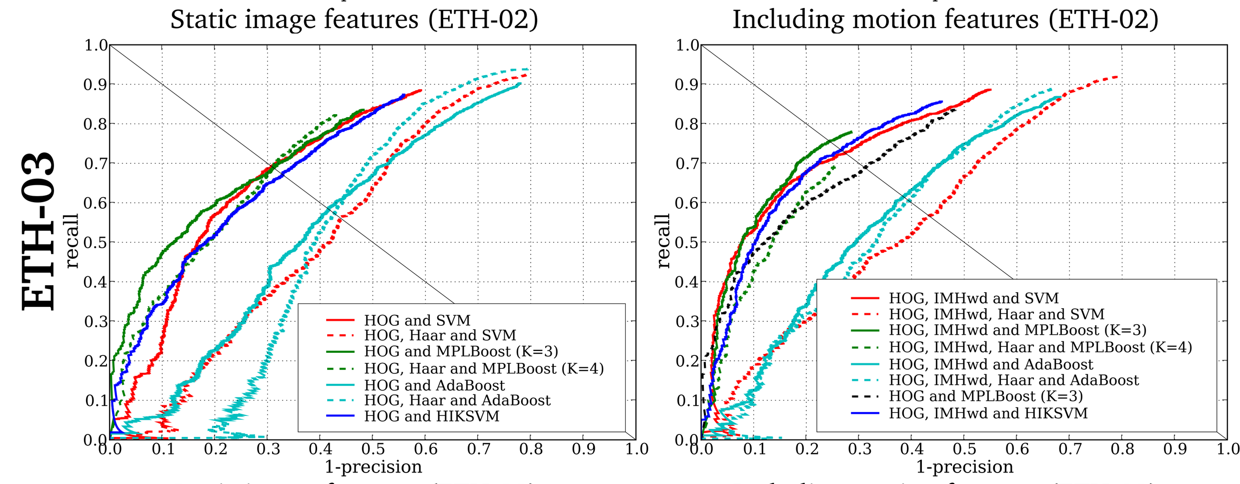




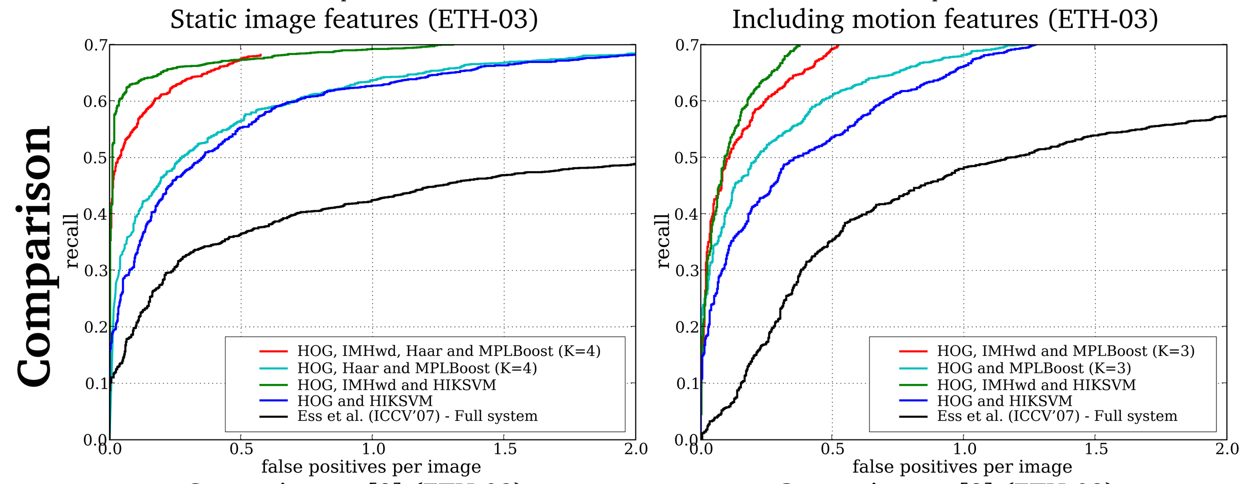
Static image features (ETH-01) | Including motion features (ETH-01)



Static image features (ETH-02) | Including motion features (ETH-02)



Static image features (ETH-03) | Including motion features (ETH-03)



Comparison to [8] (ETH-02) | Comparison to [8] (ETH-03)

Acknowledgements: This work has been funded, in part, by Toyota Motor Europe. Further we thank Christoph Zach and Thomas Pock for publicly releasing OFLib, Andreas Ess for his dataset and results and Piotr Dollár for helpful discussion.