

Crosstalk Cascades for Frame-Rate Pedestrian Detection

Piotr Dollár
Microsoft Research Redmond
pdollar@microsoft.com

Ron Appel
California Institute of Technology
appel@caltech.edu

Wolf Kienzle
Microsoft Research Redmond
wkienzle@microsoft.com

Microsoft
Research

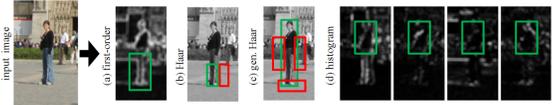


Overview

PROBLEM: detection is computationally demanding
OBSERVATION: adjacent windows are evaluated independently, no information is shared
IDEA: exploit neighborhood correlations during cascade: combine *excitation* and *inhibition*
GAINS: **5-30 x** speedup over standard cascades
SPEED: **30-60 fps** on 640x480px images (1 core, no GPU)

Previous State-of-the-Art

Integral Channel Features using FPDW for fast scale pyramid

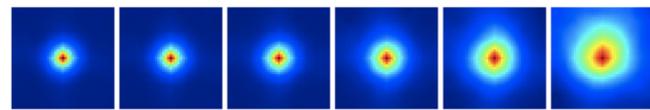


Classify using dense sliding window (4px steps)

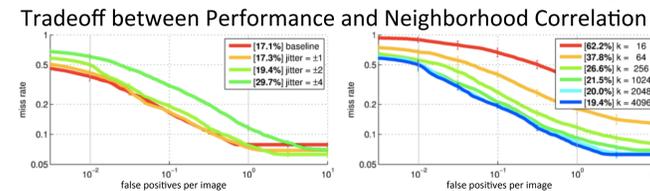
Classifier is Boosted depth-2 trees $H_K(x) = \sum_{i=1}^K \alpha_i h_i(x)$

Don't Ignore thy Neighbors

Neighboring windows are correlated (average classifier responses around true positives)



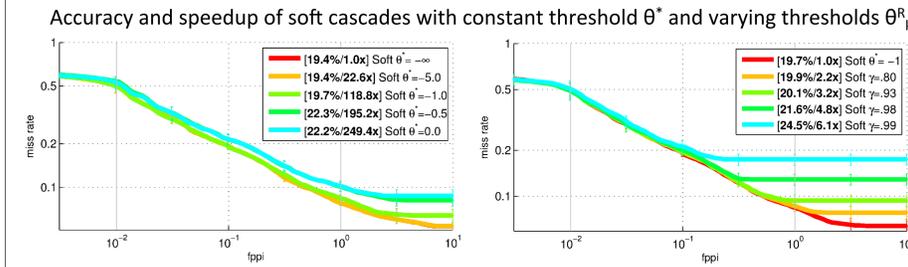
Less jitter / Smaller K vs More jitter / Larger K



Use: jitter = ±2, K = 4096, Neighborhood of [7x7x3] (12px steps)

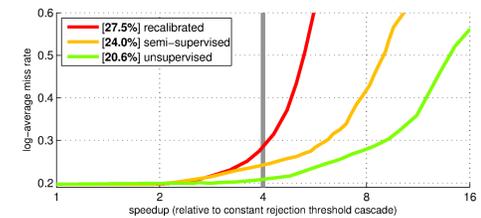
Crosstalk Cascades

Baseline Soft Cascades

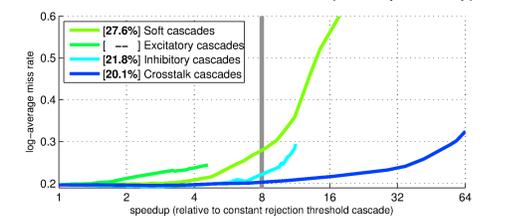


Parameter (γ) Controls Speed-Accuracy Tradeoff

Effectiveness of soft cascades (sweep over γ)

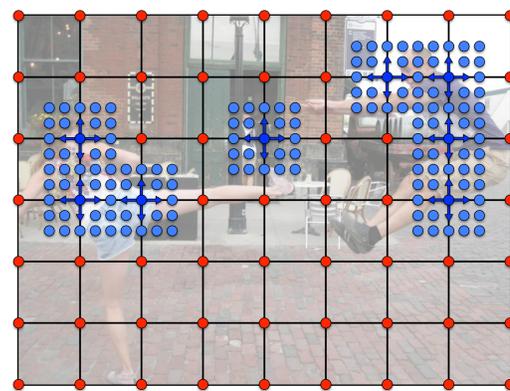


Crosstalk versus soft cascades (sweep over γ)

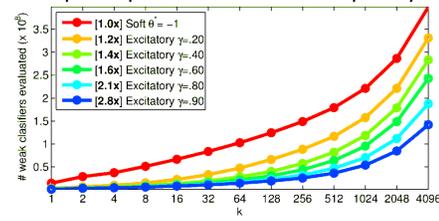


Excitatory Cascades

After k stages in the Boosting classifier, neighbors are **excited** if: $H_k > \theta_k^*$



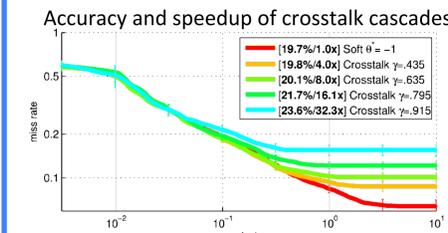
Speedup versus classifier complexity



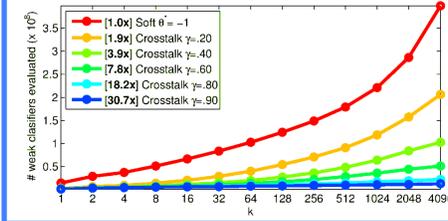
Reduces Computation for All Windows

Crosstalk Cascades
Combination of **Soft**, **Excitatory**, and **Inhibitory** cascades, reducing computation across the board

State-of-the-art detection accuracy at frame-rate speeds! Wow!



Speedup versus classifier complexity

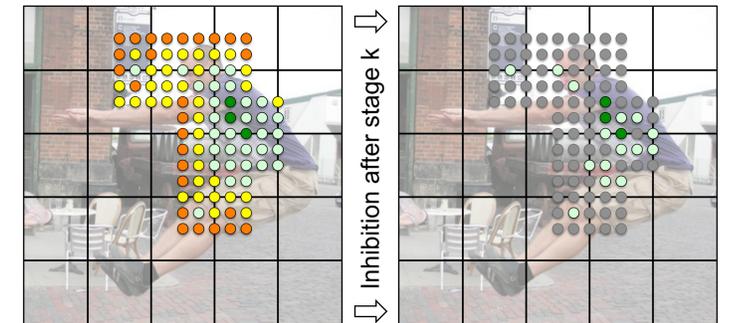


Reduces Overall Computation for Negatively-Classified Windows

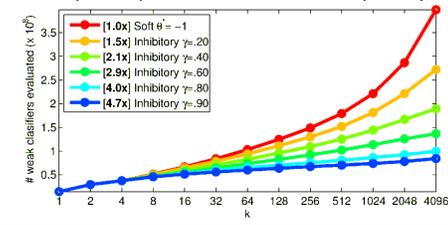
Reduces Overall Computation for Positively-Classified Windows

Inhibitory Cascades

After k stages in the Boosting classifier, an evaluating window is **inhibited** if it has a neighboring window \mathcal{N} such that: $H_k/H_{\mathcal{N}_k} < \theta_k^*$



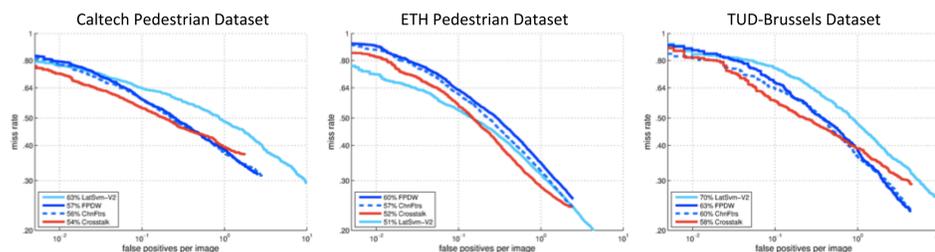
Speedup versus classifier complexity



Additional Pedestrian Datasets

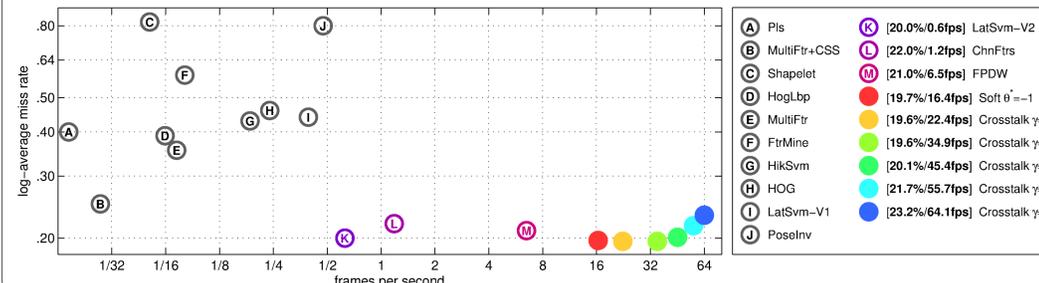
Performance on other datasets

http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/



Fast and Accurate Results

Log-average miss rate (MR) versus speed for various detectors on INRIA pedestrians



Summary

- FAST:** 30-60 fps (5-30x speedup)
- ACCURATE:** state-of-the-art detection
- ROBUST:** consistent performance across datasets
- GENERAL:** applicable to any multi-stage detector
- TRAINABLE:** no extra supervised data required
- CODE:** feature computation code available at: <http://vision.ucsd.edu/~pdollar/toolbox/doc/index.html>

see paper for more details