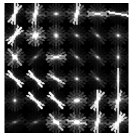
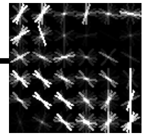
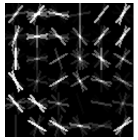
Bounding Part Scores for Rapid Detection with Deformable Part Models

$\epsilon = \langle$  $-$  $,$  $\rangle$

Iasonas Kokkinos

Center for Visual Computing
Ecole Centrale de Paris

Galen Group
INRIA-Saclay

17th October 2012
Parts and Attributes Workshop, ECCV 2012

Problem

Efficient detection with Deformable Part Models

Previous work: Dual-Tree Branch-and-Bound (DTBB)

Acceleration over Generalized Distance Transforms (GDT)

Problem: real bottleneck is part score computation (pre-GDT)

Current work: Incorporate part score computation in DTBB

Combine with Cascaded-DPM detection

P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, 'Object Detection with Discriminatively Trained Part Based Models', PAMI 2010

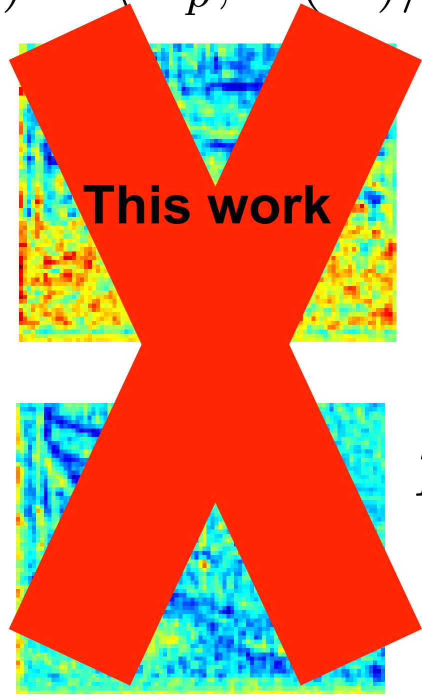
I. Kokkinos, Rapid DPM Detection using Dual-Tree Branch-and-Bound, NIPS 2011

P. Felzenszwalb R. Girshick and D. McAllester Cascade object detection with DPMs CVPR 2010

Object detection with Deformable Part Models (DPMs)

$$U_p(x') = \langle \mathbf{w}_p, \mathbf{H}(x') \rangle$$

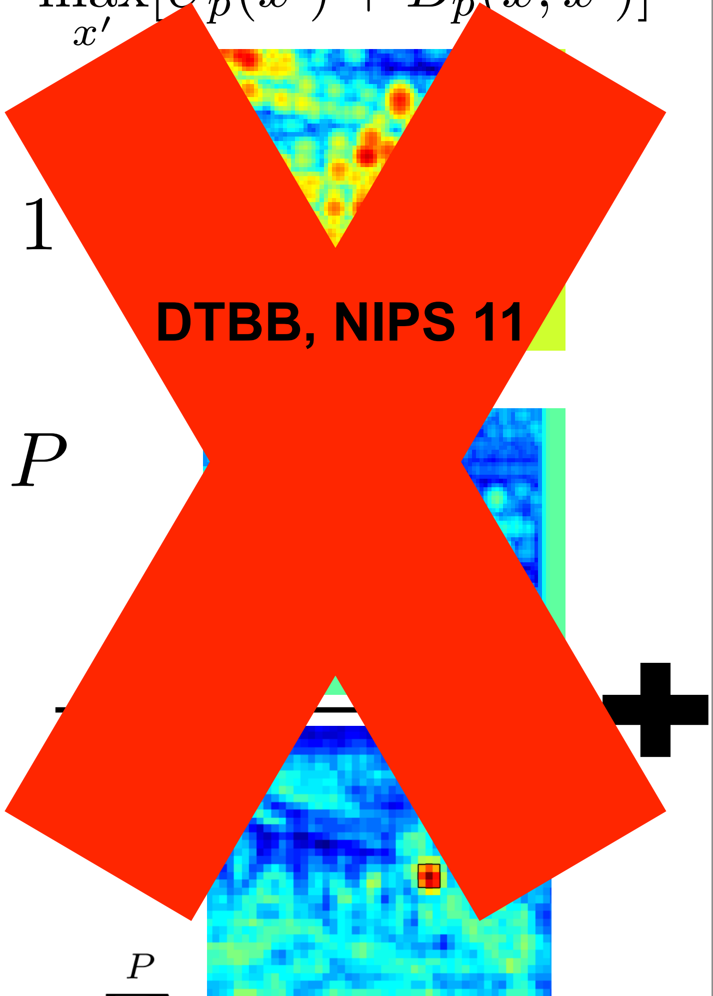
$$\max_{x'}[U_p(x') + B_p(x, x')]$$

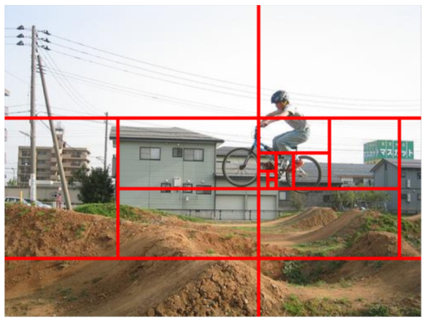


$p = 1$

$\vdots$

$p = P$





$$S(x) = \sum_{p=1}^{P} \max_{x'}[U_p(x') + B_p(x, x')]$$

Accelerating detection with DPMs

Efficient detection with DPMs

P F Felzenszwalb R B Girshick and D A McAllester Cascade object detection with DPMs CVPR 2010

B. Sapp, A. Toshev, and B. Taskar. Cascaded models for articulated pose estimation, ECCV, 2010

M. Pedersoli, A.Vedaldi, and J.Gonzalez. A coarse-to-fine approach for object detection, CVPR 2011

I. Kokkinos, Rapid DPM Detection using Dual-Tree Branch-and-Bound, NIPS 2011

Efficient part score computation

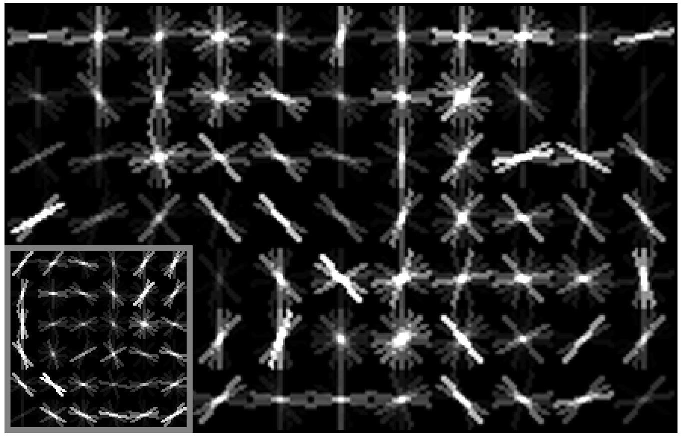
H. Pirsiavash and D. Ramanan, Steerable Part Models, CVPR 2012

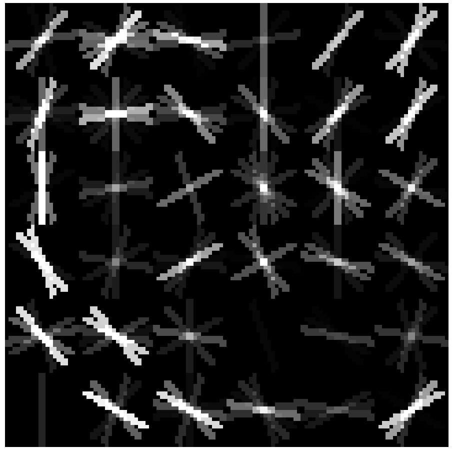
A. Vedaldi and A. Zisserman, Sparse Kernel Maps and Faster Product Quantization Learning, CVPR 2012

H.O. Song, S. Zickler, T. Althoff, R. Girschick, M. Fritz, C. Geyer, P. Felzenszwalb, T. Darrell, Sparselet Models for Efficient Multiclass Object Detection, ECCV 2012

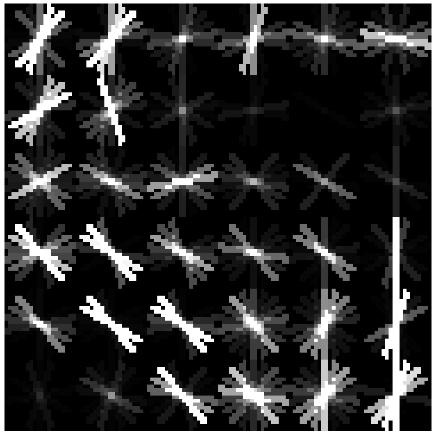
C. Dubout and F. Fleuret. Exact Acceleration of Linear Object Detectors, ECCV 2012

Part score computation





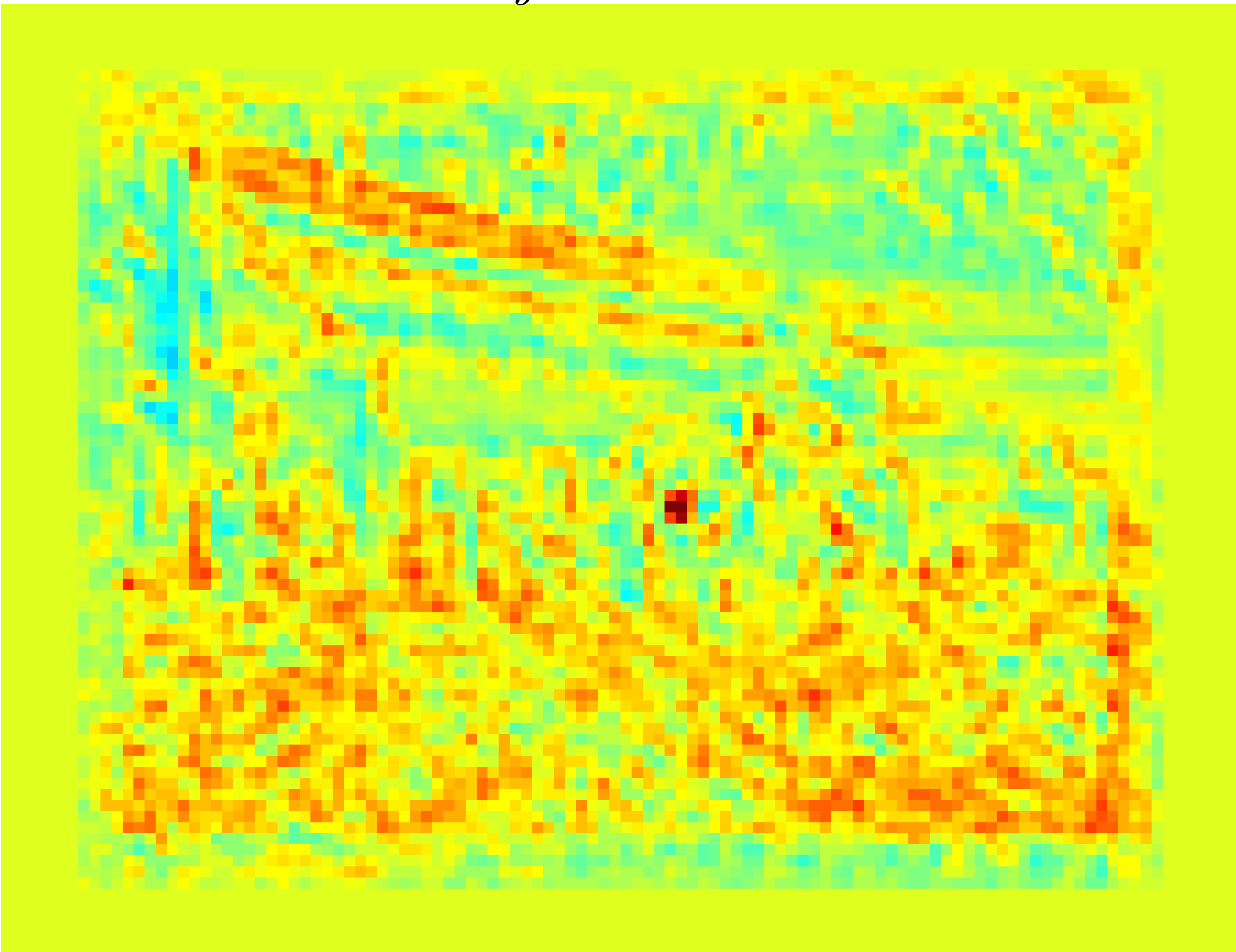
$\mathbf{w}[y]$



$\mathbf{h}[x+y]$

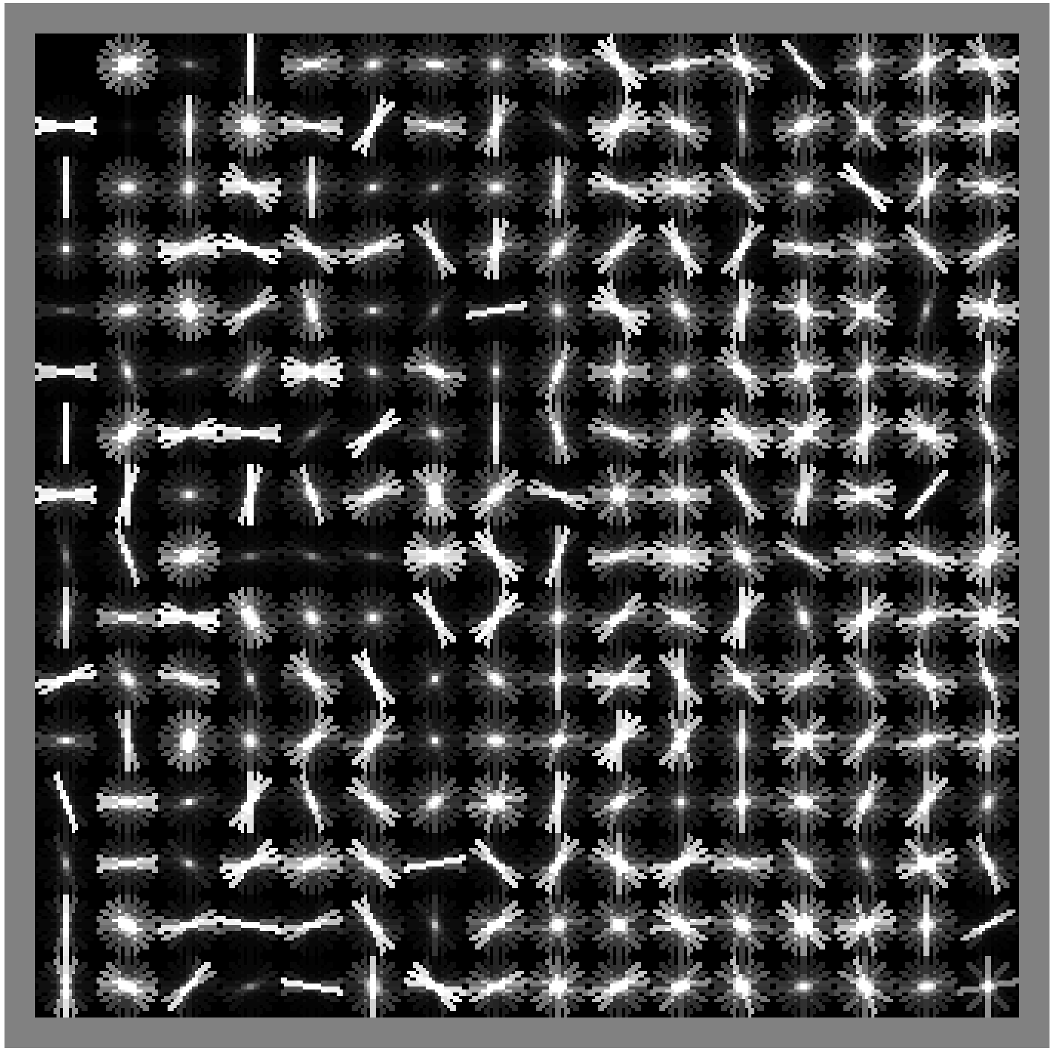
$$s[x] = \sum_{y} \langle \mathbf{h}[x+y], \mathbf{w}[y] \rangle$$

Part scores $s[x] = \sum_{y} \langle \mathbf{h}[x+y], \mathbf{w}[y] \rangle$



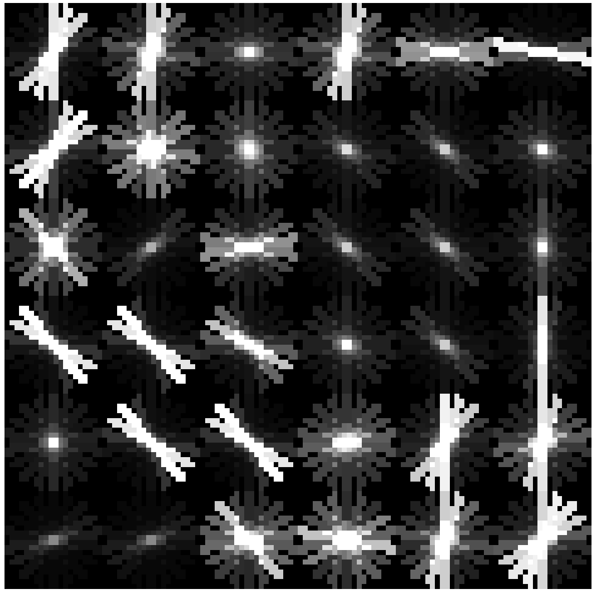
HOG cell quantization: visual 'letters'

$$\mathcal{C} = \{C_1, \ldots, C_{256}\}$$



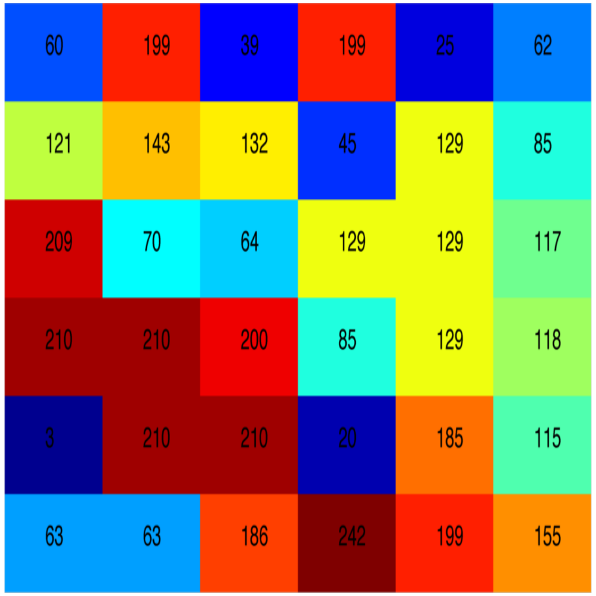
HOG feature quantization

HOG detail



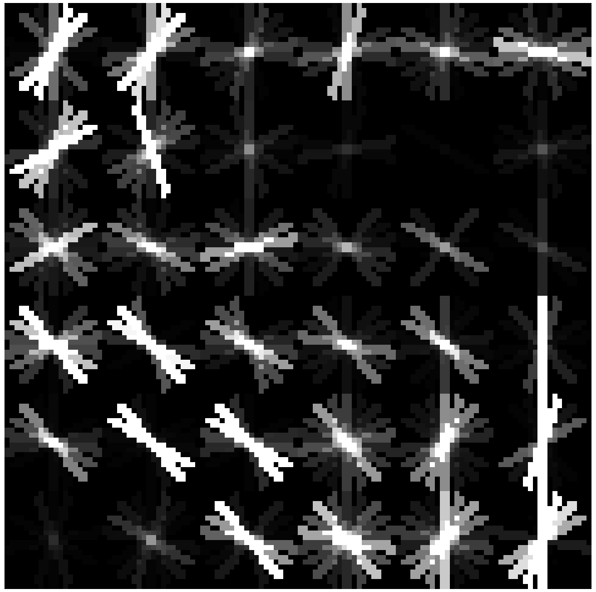
$$\mathbf{h}[x]$$

Codebook indices



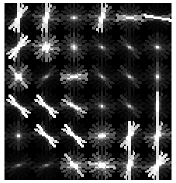
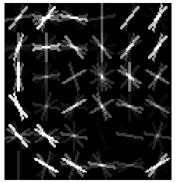
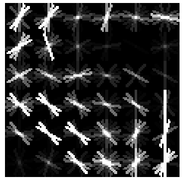
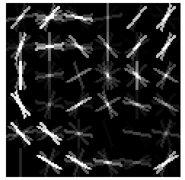
$$i[x] = \arg\min_k d(\mathbf{h}[x], C_k)$$

Quantized HOG



$$\hat{\mathbf{h}}[x] = C_{i[x]}$$

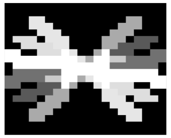
Efficient inner product approximation

$\Big\langle$  ,  $\Big\rangle \simeq \Big\langle$  ,  $\Big\rangle$

$$s[x] \simeq \hat{s}[x]$$

$$\sum_y \langle \mathbf{h}[x+y], \mathbf{w}[y] \rangle \simeq \sum_y \langle \hat{\mathbf{h}}[x+y], \mathbf{w}[y] \rangle$$

Efficient inner product approximation

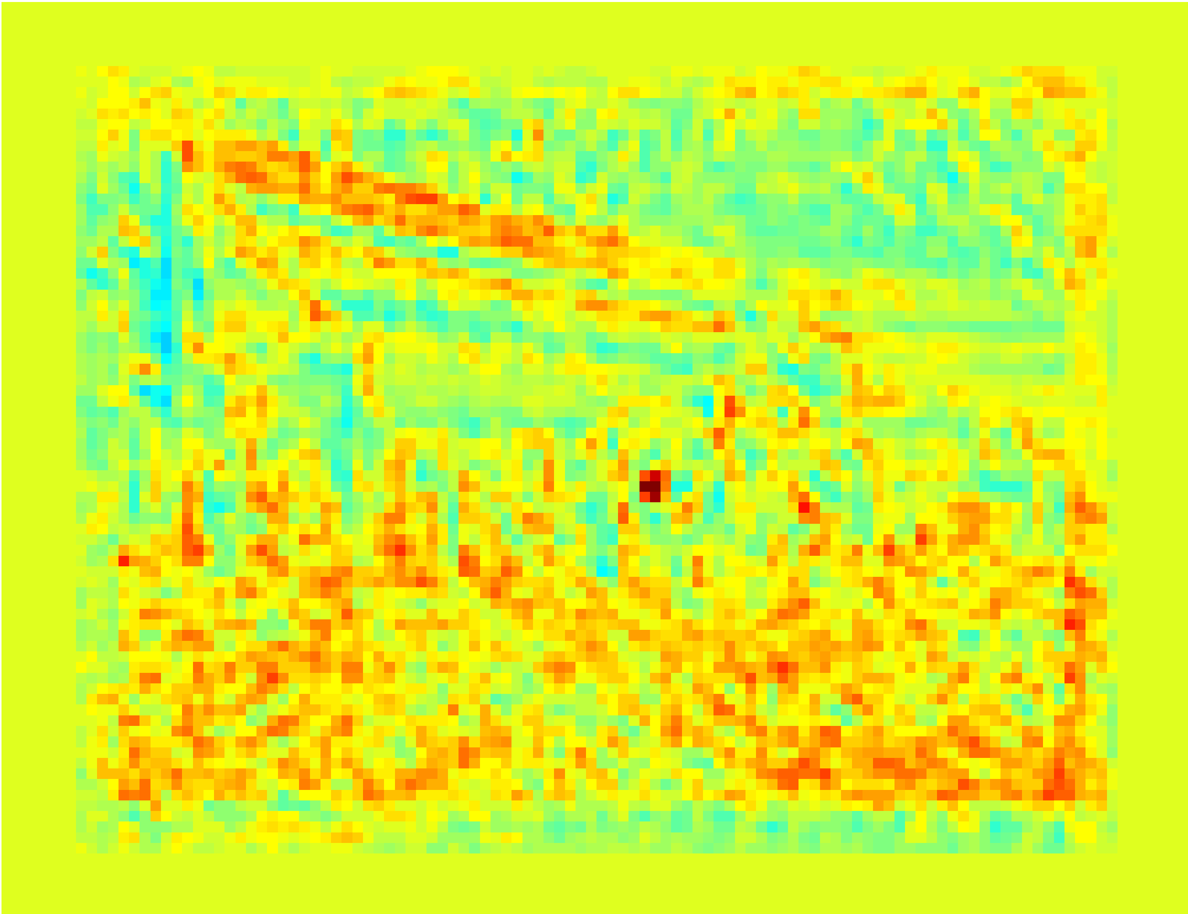
⟨  ,  ⟩ ≃ ⟨  ,  ⟩

$$\begin{aligned}
\langle \mathbf{h}[x+y], \mathbf{w}[y] \rangle \simeq \langle \hat{\mathbf{h}}[x+y], \mathbf{w}[y] \rangle &= \langle C_{I[x+y]}, \mathbf{w}[y] \rangle \\
&= \Pi[I[x+y], y]
\end{aligned}$$

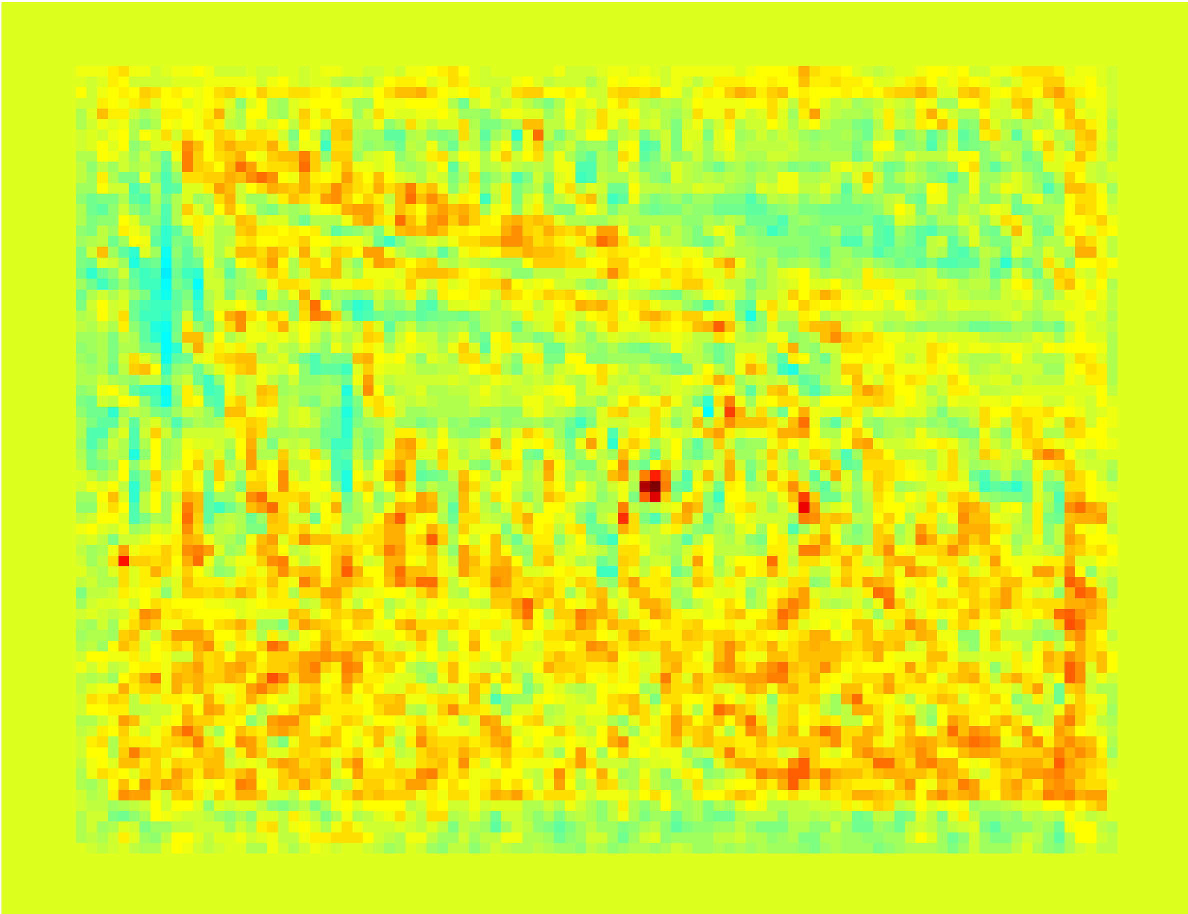
$$\Pi[k, y] = \langle C_k, \mathbf{w}_y \rangle$$

$$s[x] \simeq \hat{s}[x] = \sum_y \Pi[I[x+y], y]$$

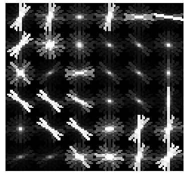
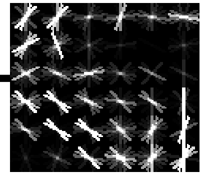
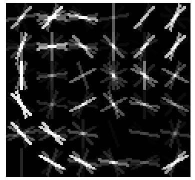
Lookup-based estimate demonstration: $s[x]$



Lookup-based estimate demonstration: $\hat{s}[x]$

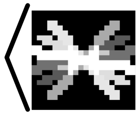


Part-level approximation error

$\epsilon \doteq s - \hat{s} = \langle$  $-$  ,  $\rangle$

$$= \sum_y \underbrace{\langle \mathbf{h}[y] - \hat{\mathbf{h}}[y], \mathbf{w}[y] \rangle}_{e[y]}$$

Cell-level approximation error

$$e[y] = \langle \mathbf{h}[y] - \hat{\mathbf{h}}[y], \mathbf{w}[y] \rangle = \langle$$  $$-$$  $$,$$  $$\rangle$$

$$= \langle \mathbf{e}[y], \mathbf{w}[y] \rangle$$

$$= \sum_{f=1}^{32} \mathbf{e}_y[f] \mathbf{w}_y[f]$$

Chebyshev inequality

For any zero-mean random variable, and any value of α:

$$P(|X| > \alpha) \leq \frac{E\{X^2\}}{\alpha^2}$$

Equivalently, with probability of error smaller than p_e:

$$X \in \left[-\sqrt{\frac{E\{X^2\}}{p_e}}, \sqrt{\frac{E\{X^2\}}{p_e}}\right]$$

Chebyshev inequality-II

For a weighted sum of i.i.d. zero-mean random variables:

$$X' = \sum_{k=1}^{K} w_k X_k$$

with probability of error smaller than p_e :

$$X' \in \left[-\sqrt{\frac{\left(\sum_k w_k^2\right) E\{X^2\}}{p_e}}, \sqrt{\frac{\left(\sum_k w_k^2\right) E\{X^2\}}{p_e}} \right]$$

Chebyshev inequality for cell-level error

$$e[y] = \sum_{f=1}^{F} \mathbf{e}_y[f]\mathbf{w}_y[f]$$

with probability of error smaller than p_e :

$$e_y \in \left[-\sqrt{\frac{\|\mathbf{w}[y]\|^2\|\mathbf{e}[y]\|^2}{p_e F}}, \sqrt{\frac{\|\mathbf{w}[y]\|^2\|\mathbf{e}[y]\|^2}{p_e F}}\right]$$

Chebyshev inequality for part-level error

$$\boxed{\epsilon = \hat{s} - s} = \sum_y e[y]$$

with probability of error smaller than p_e :

$$\epsilon \in \left[-\sqrt{\frac{\sum_y \|\mathbf{w}[y]\|^2 \|\mathbf{e}[y]\|^2}{p_e F}}, \sqrt{\frac{\sum_y \|\mathbf{w}[y]\|^2 \|\mathbf{e}[y]\|^2}{p_e F}} \right]$$

→ with probability of error smaller than p_e:

$$s \in \left[\hat{s} - \sqrt{\frac{\sum_y \|\mathbf{w}[y]\|^2 \|\mathbf{e}[y]\|^2}{p_e F}}, \hat{s} + \sqrt{\frac{\sum_y \|\mathbf{w}[y]\|^2 \|\mathbf{e}[y]\|^2}{p_e F}} \right]$$

Recap

Lookup-based approximation:

$$s[x] \simeq \hat{s}[x] = \sum_y \Pi[I[x+y], y]$$

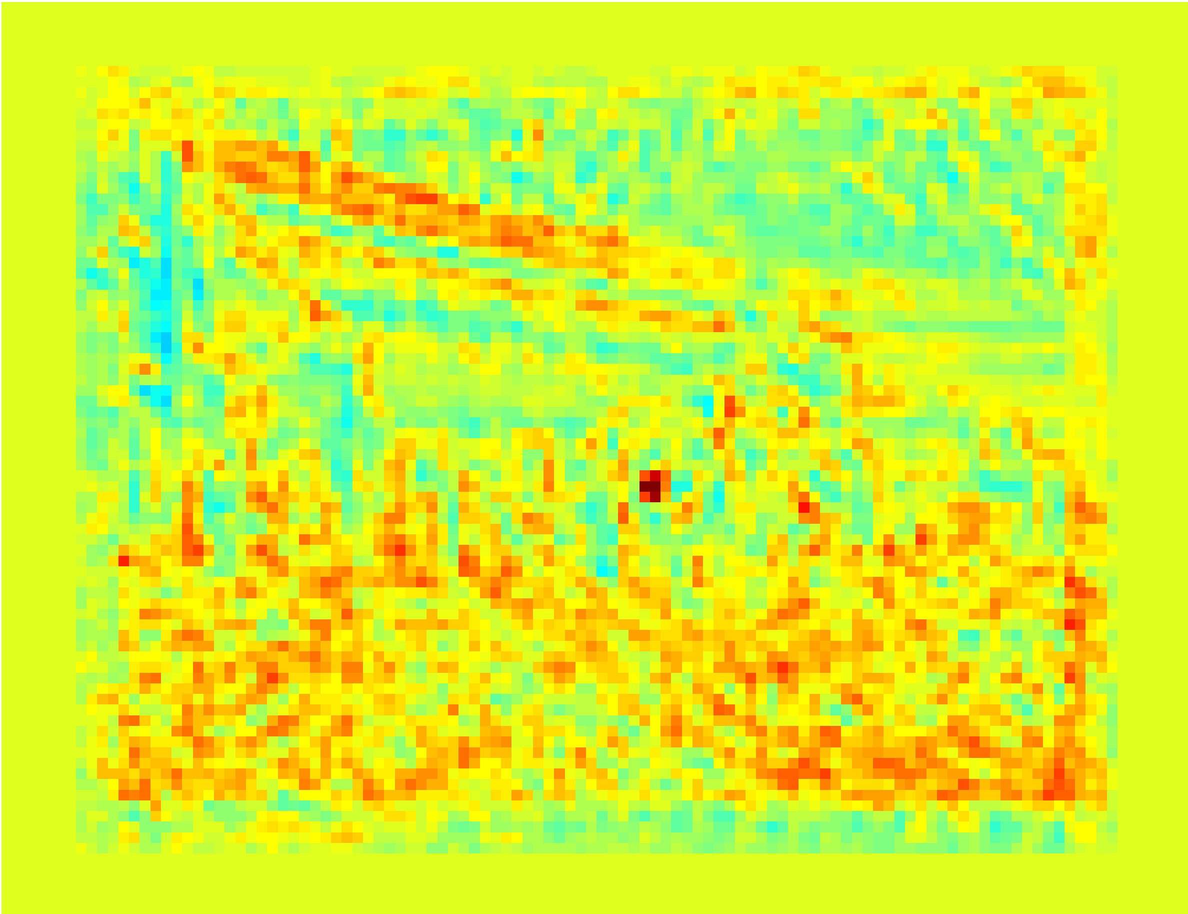
With probability of error at most p_e:

$$\underline{s}[x] \leq s[x] \leq \overline{s}[x]$$

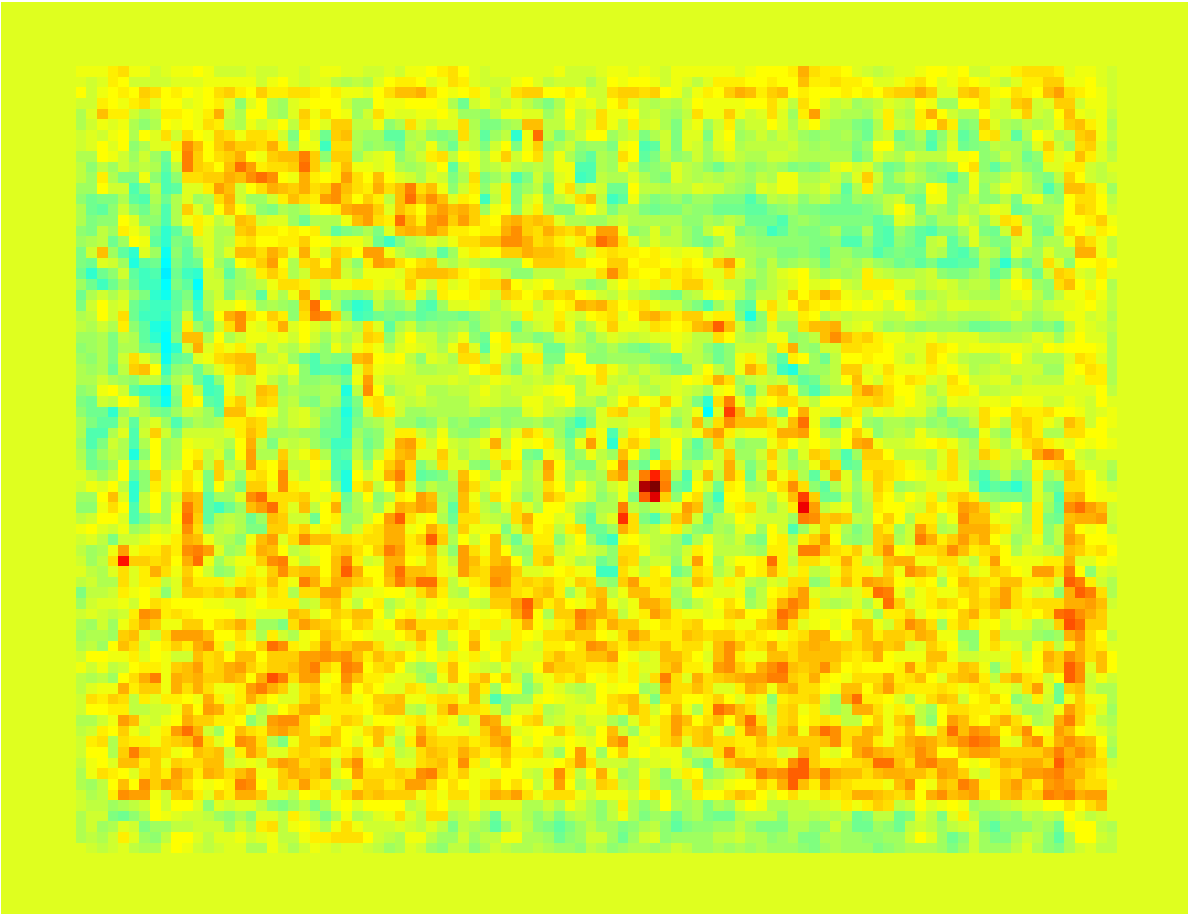
$$\underline{s}[x] = \hat{s}[x] - \sqrt{\frac{\sum_y \|\mathbf{w}[y]\|^2 \|\mathbf{e}[x+y]\|^2}{p_e F}}$$

$$\overline{s}[x] = \hat{s}[x] + \sqrt{\frac{\sum_y \|\mathbf{w}[y]\|^2 \|\mathbf{e}[x+y]\|^2}{p_e F}}$$

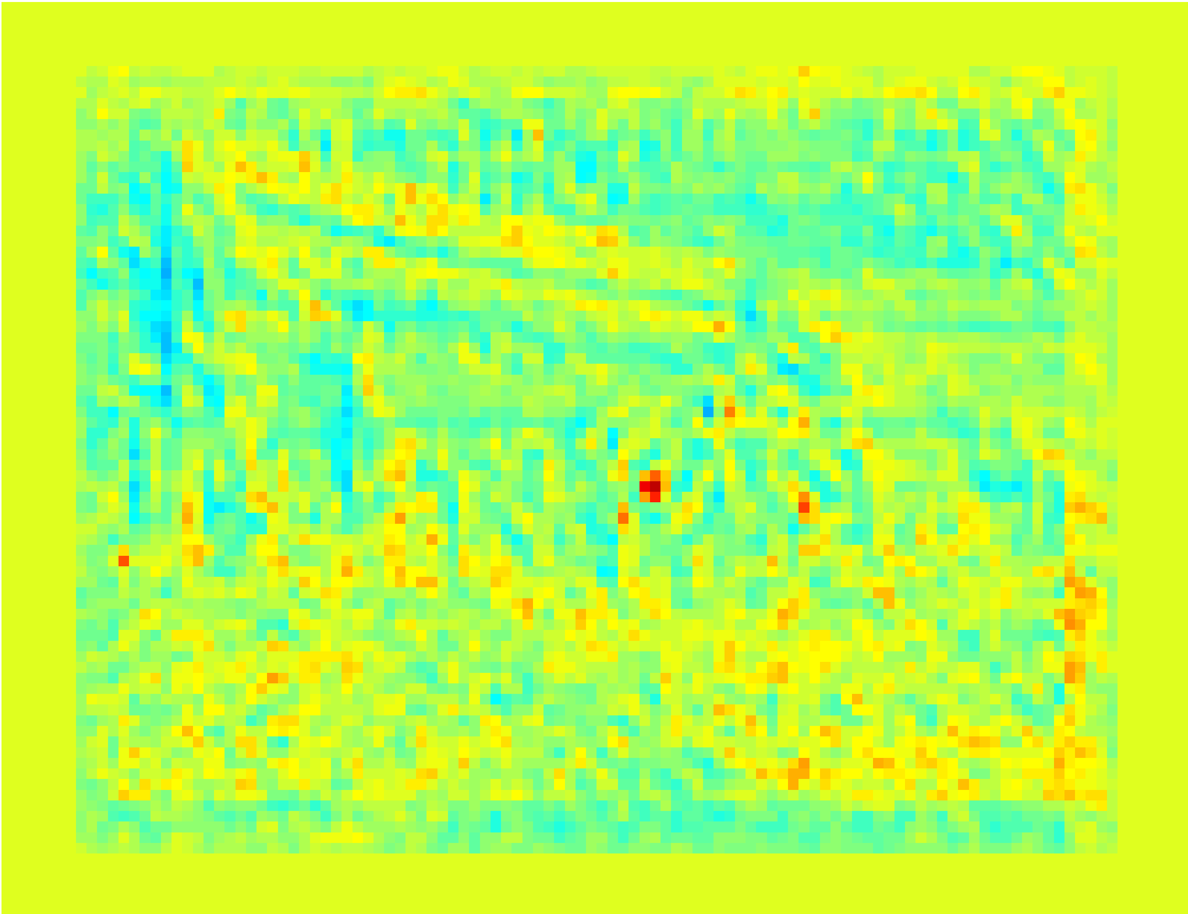
Bound demonstration: $s[x]$



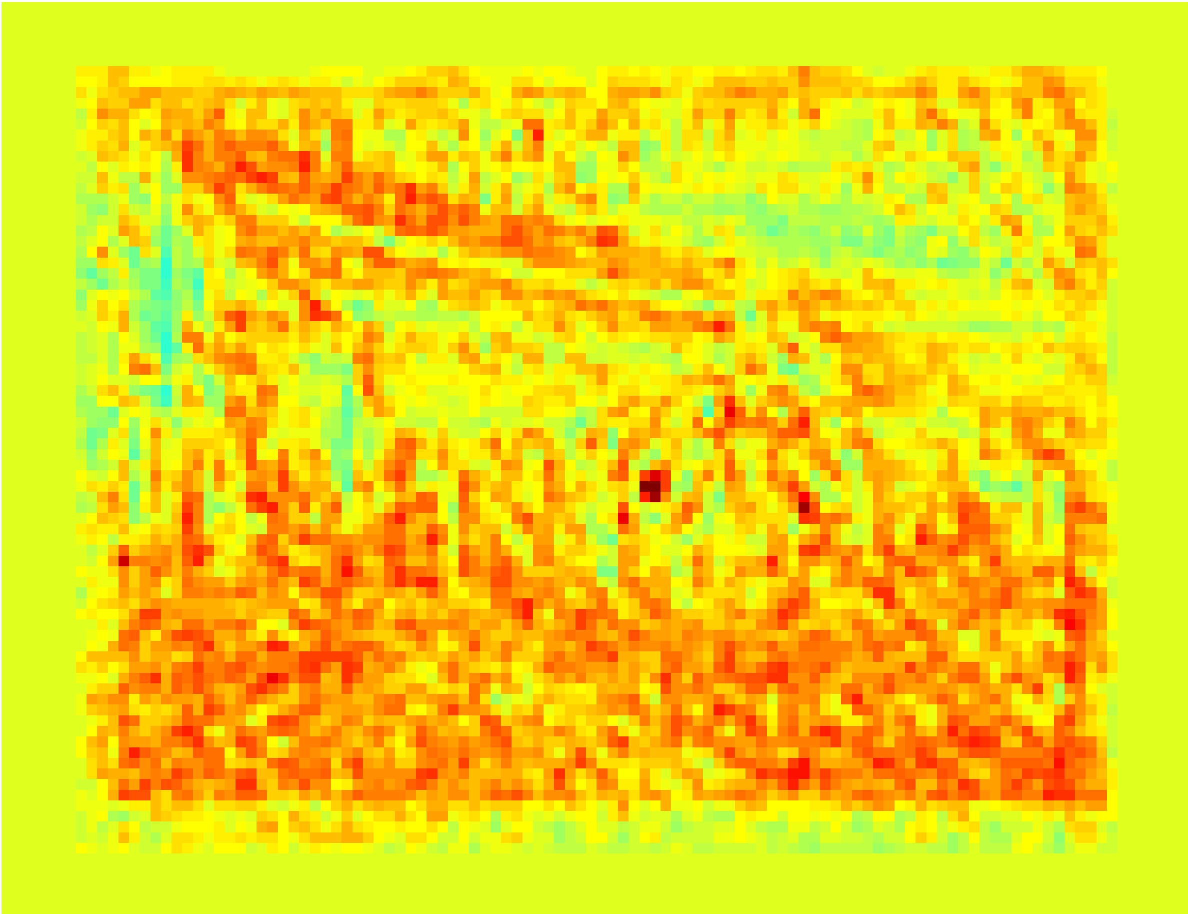
Bound demonstration: $\hat{s}[x]$



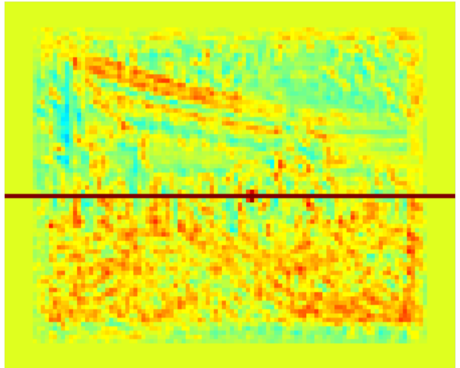
Bound demonstration: $\underline{s}[x],\ p_e = .05$

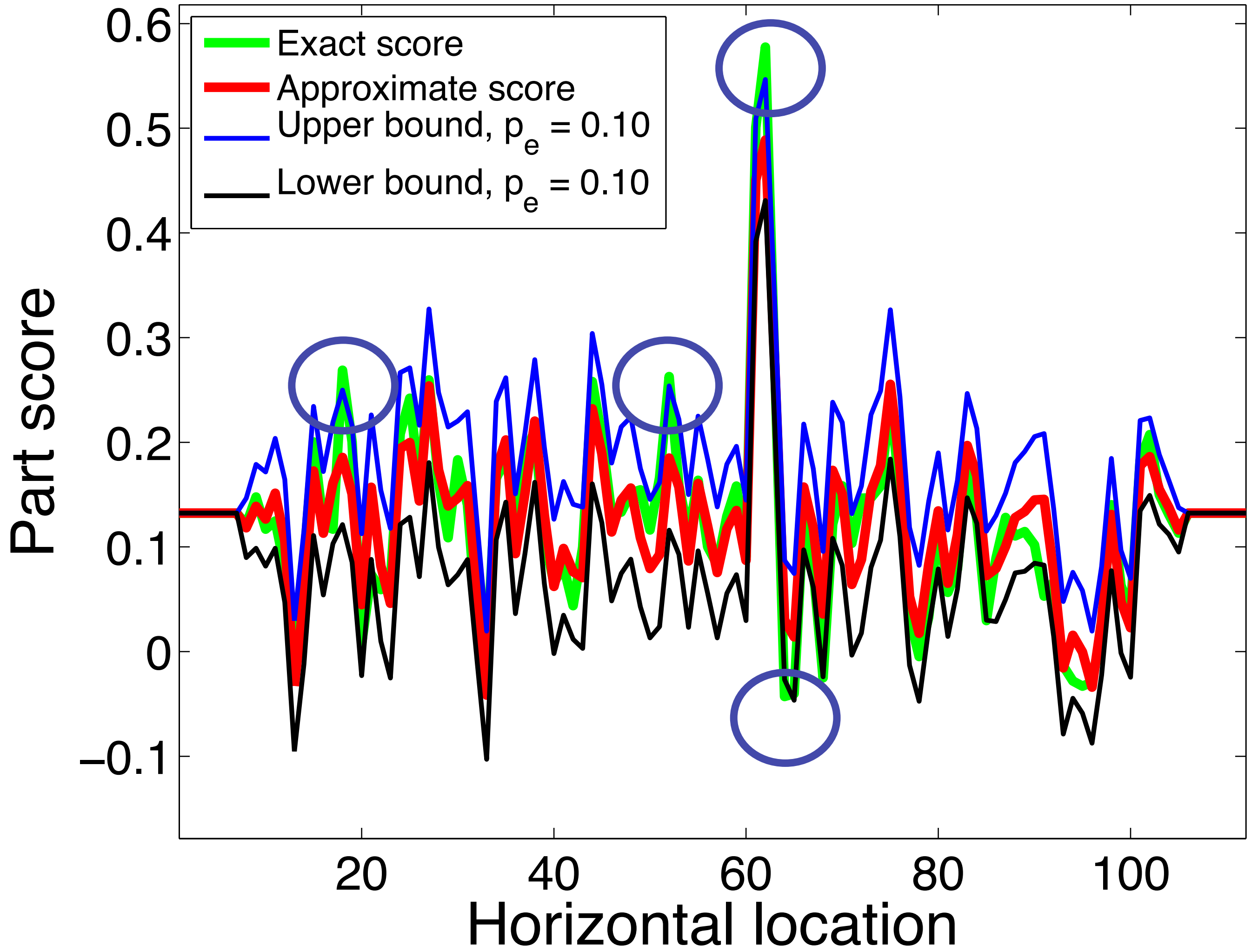


Bound demonstration: $\overline{s}[x],\ p_e = .05$

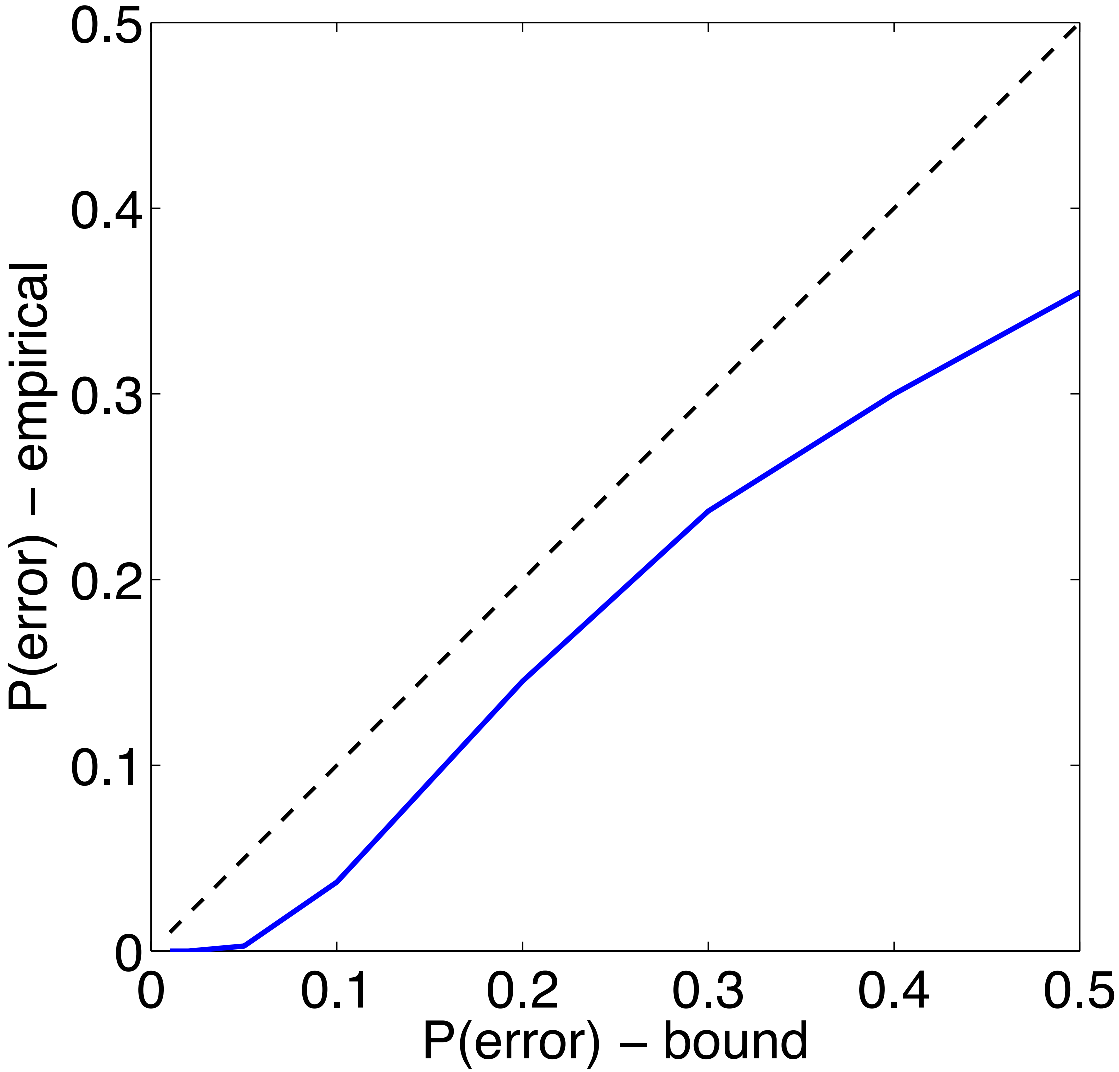


Bound demonstration for varying confidence





Bound tightness



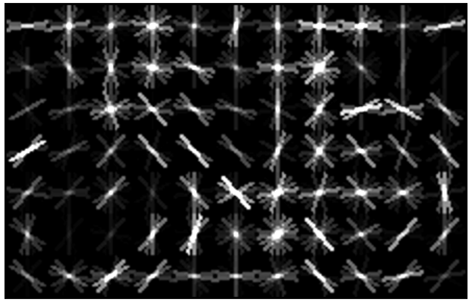
Integration with detection

Dual-Tree Branch-and-Bound

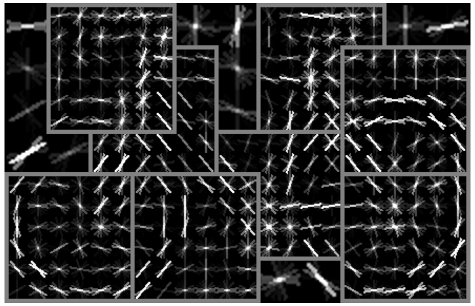
Cascaded DPMs (Felzenszwalb, Girschick et al, CVPR 2010)

Accelerating detection with DPMs

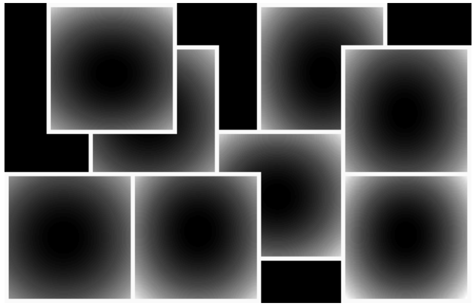
This work



w_1



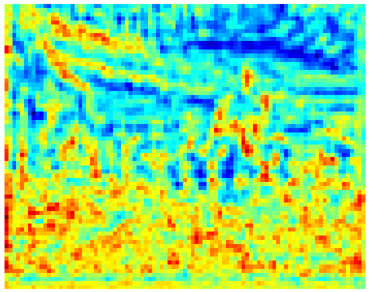
$w_2 \ldots w_P$



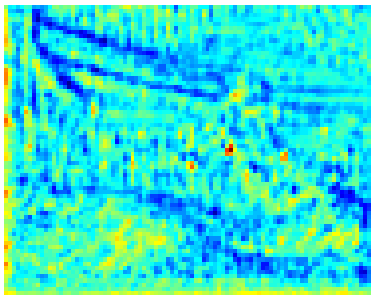
$B_2 \ldots B_P$

$$U_p(x) = \langle w_p, H(x) \rangle$$

$$\mu_p(x) = \max_{x'}[U_p(x') + B_p(x, x')]$$

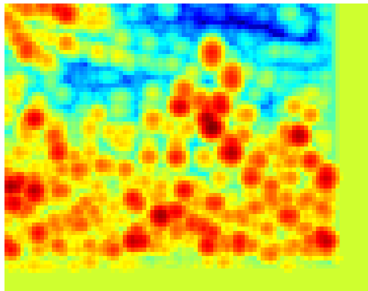


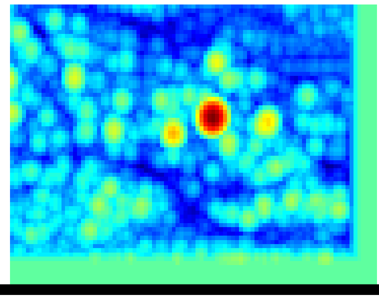
$p = 3$



$p = 5$

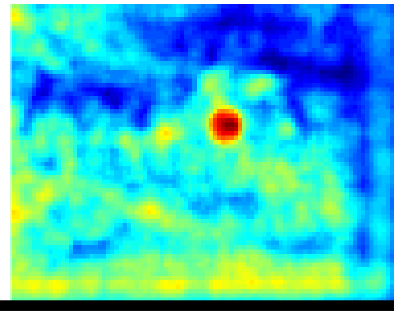
GDT





DTBB, NIPS 2011

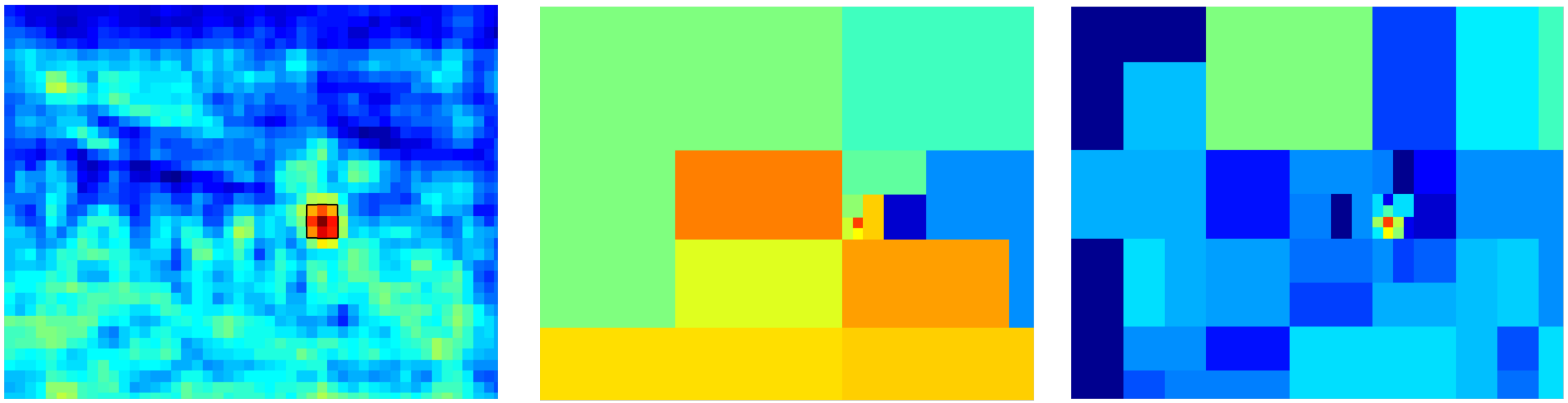
$$S(x) = \sum_{p=1}^{P} \mu_p(x)$$



Dual-Tree Brand-and-Bound

Input & Detection result

Detector score $S(x)$ BB for $\arg\max_x S(x)$ BB for $S(x) \geq -1$



I. Kokkinos. Rapid Deformable Object Detection using DTBB, NIPS 2011

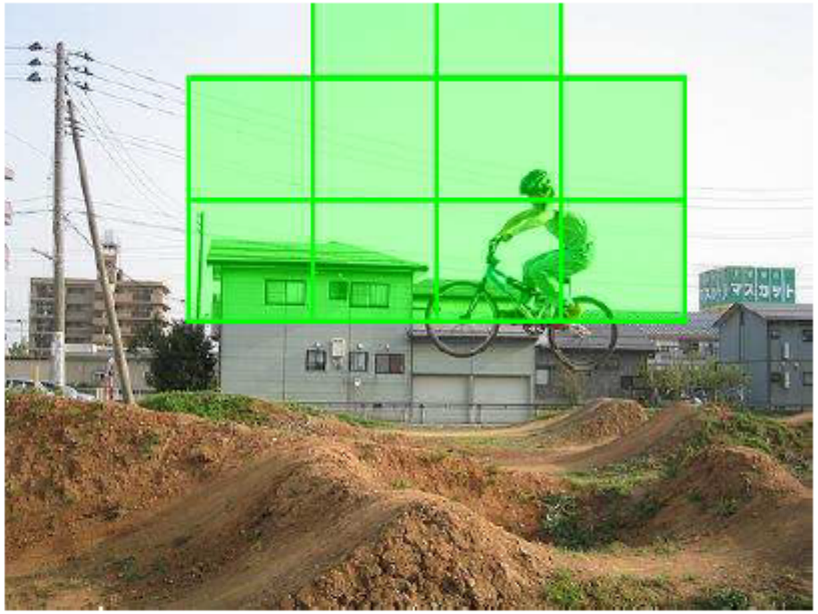
DTBB demonstration

DTBB demonstration

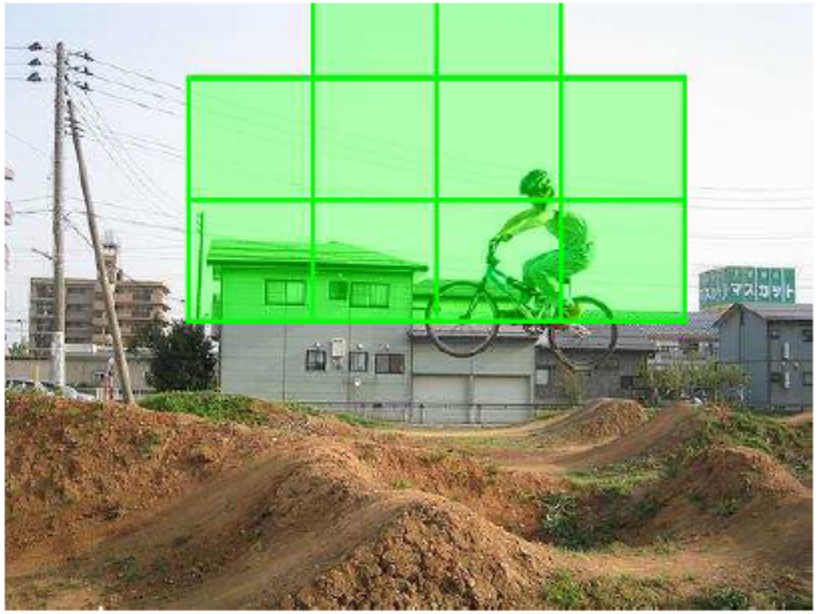
DTBB demonstration

DTBB demonstration

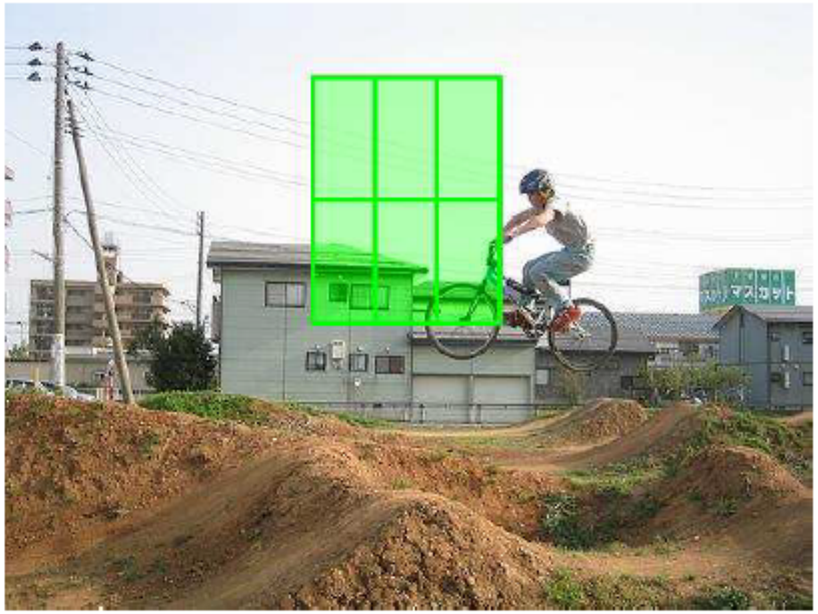
DTBB demonstration



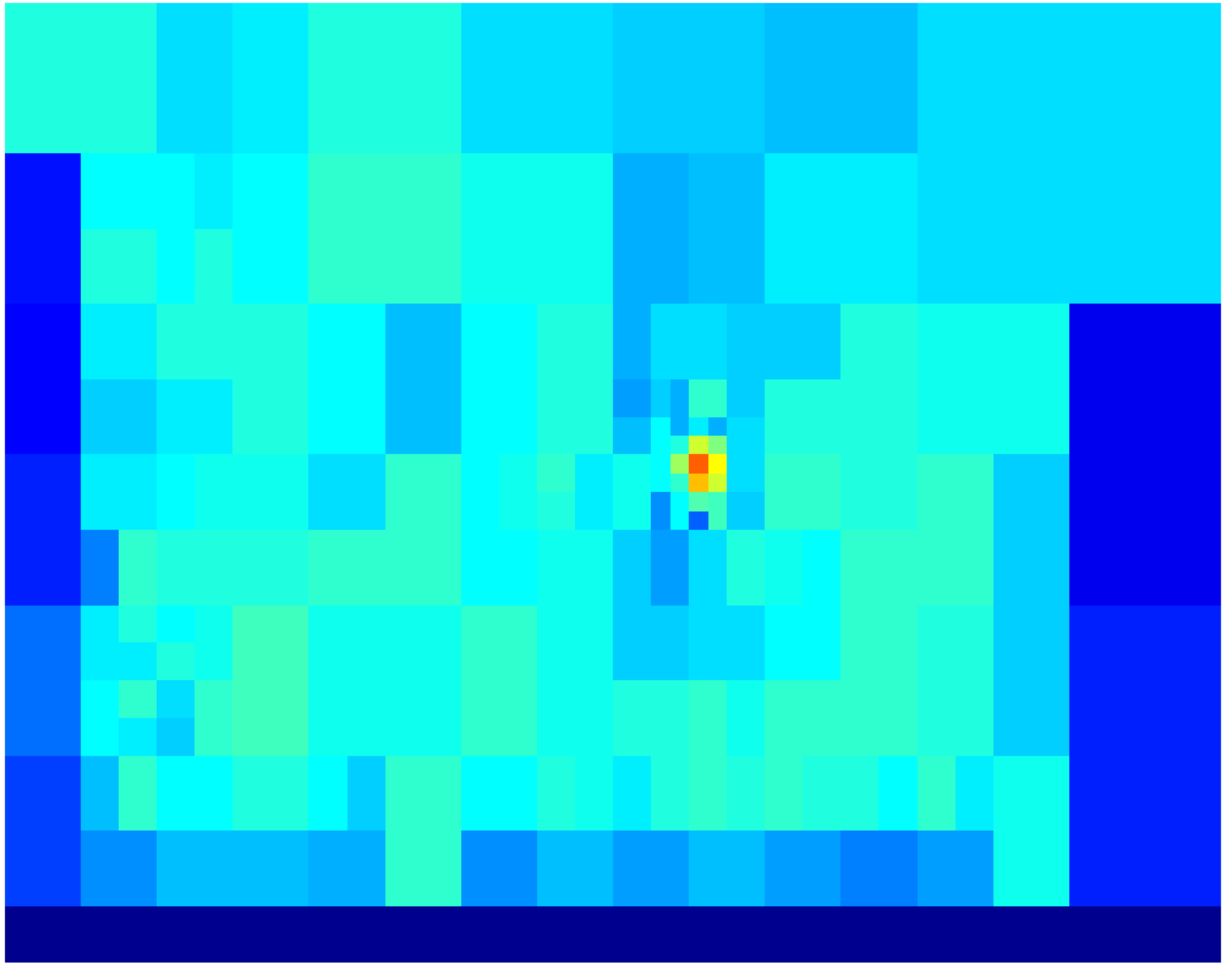
DTBB demonstration



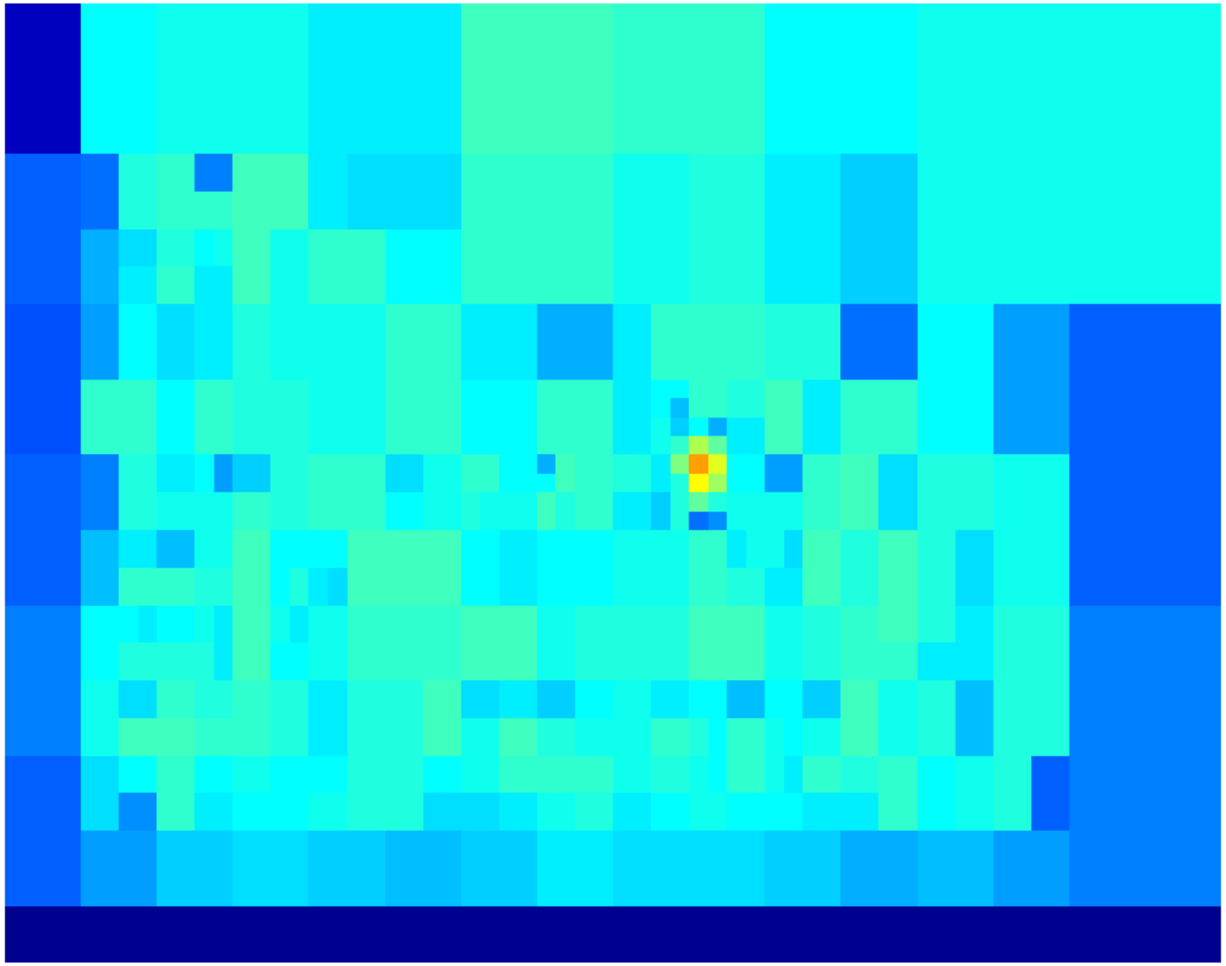
DTBB demonstration



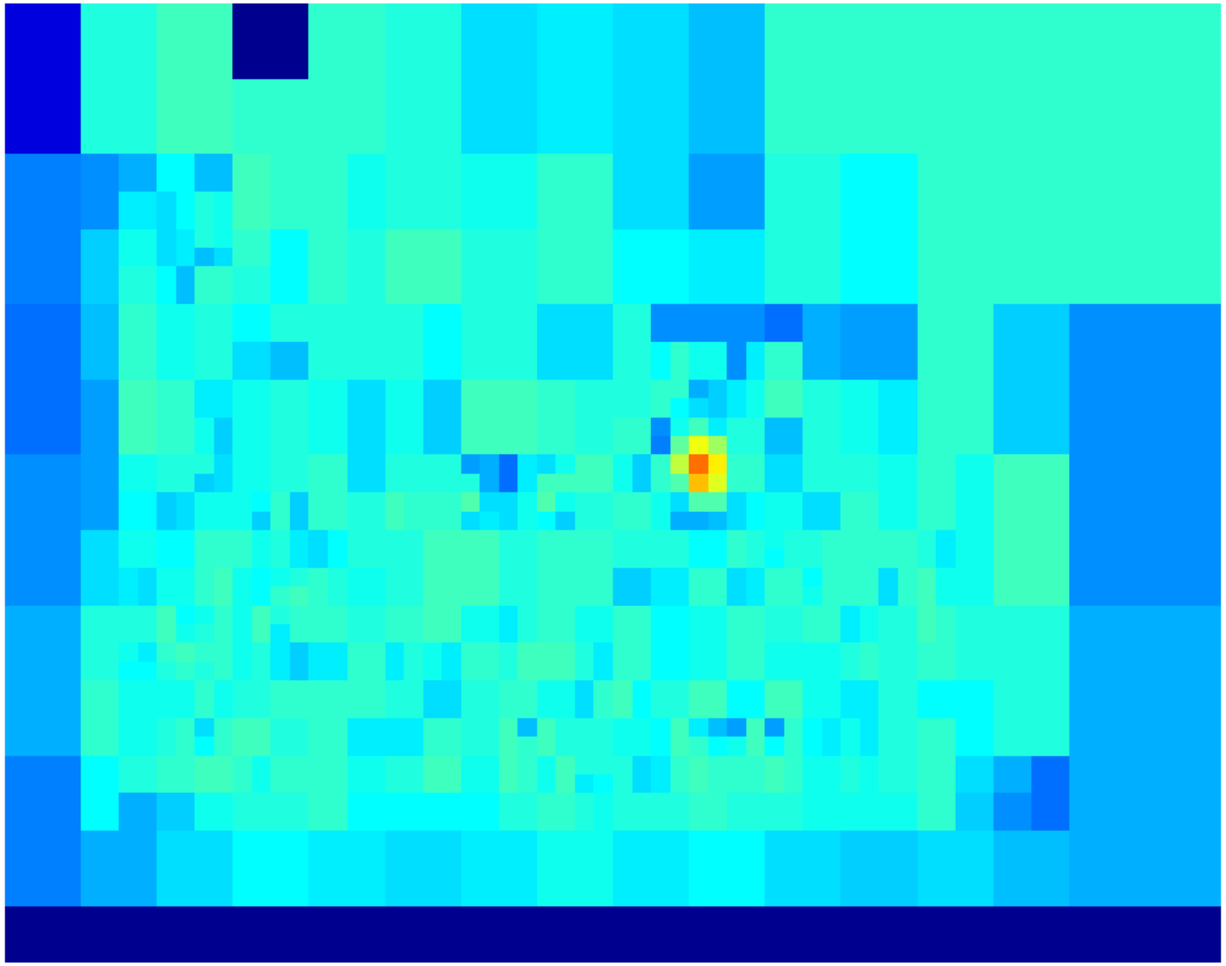
DTBB results: exact part scores



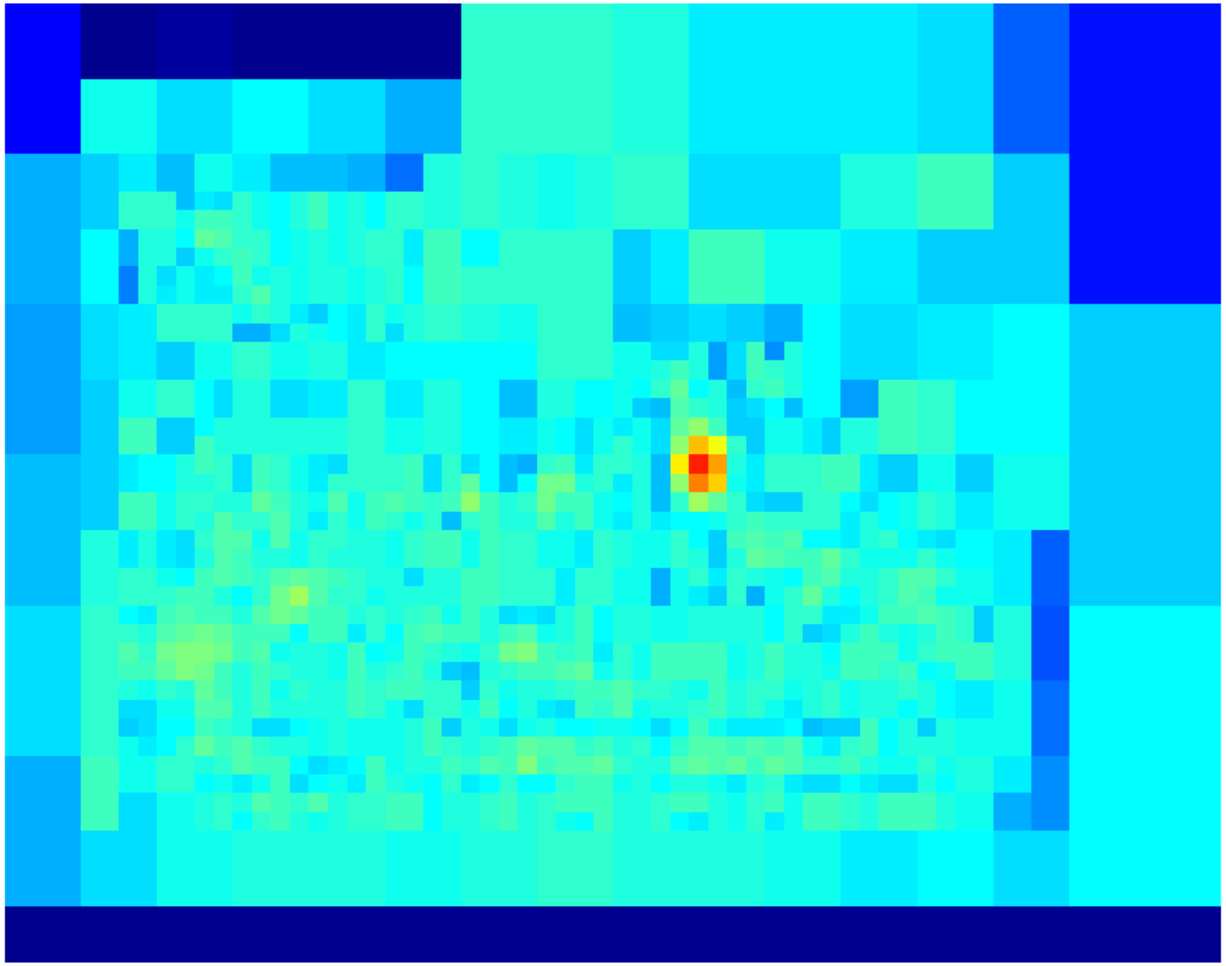
DTBB results, part score bounds @ $p_e = 0.2$



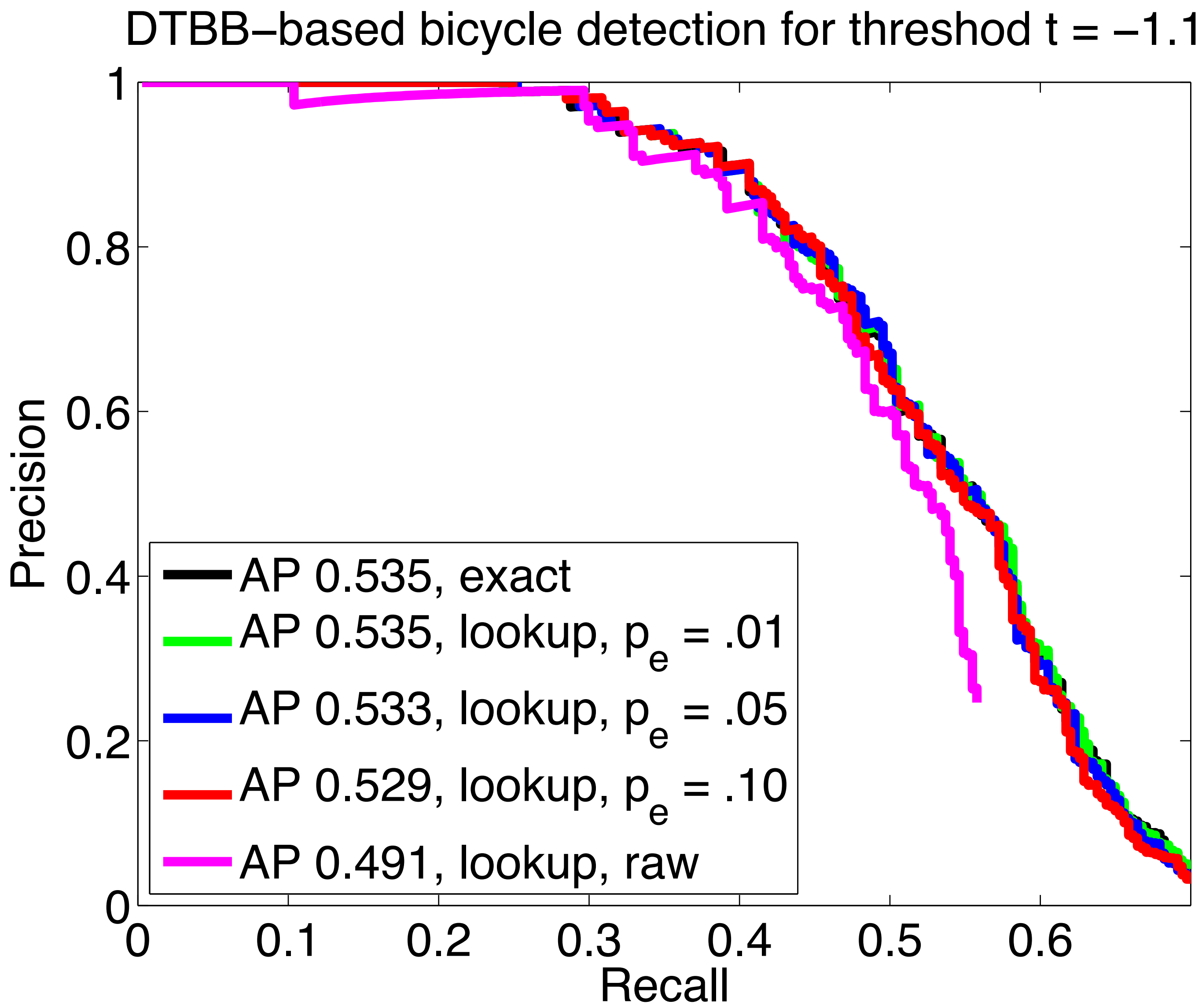
DTBB results, part score bounds @ $p_e = 0.1$



DTBB results, part score bounds @ $p_e = 0.05$



Impact on performance



Speedup results

	GDTs	DTBB	$p_e = 0.05$	$p_e = 0.01$
Part terms	8.35 ± 0.77	1.69 ± 0.18	0.69 ± 0.03	0.69 ± 0.06
$\theta = -0.5$	0.60 ± 0.05	0.21 ± 0.06	0.47 ± 0.11	1.04 ± 0.25
Sum	8.95 ± 0.82	1.90 ± 0.23	1.17 ± 0.12	1.74 ± 0.32
$\theta = -0.7$	0.60 ± 0.05	0.42 ± 0.10	1.00 ± 0.23	2.33 ± 0.65
Sum	8.95 ± 0.82	2.10 ± 0.24	1.70 ± 0.27	3.00 ± 0.71
$\theta = -1.0$	0.60 ± 0.05	1.31 ± 0.31	3.80 ± 0.90	9.40 ± 2.70
Sum	8.95 ± 0.82	3.00 ± 0.42	4.50 ± 1.02	10.01 ± 2.82

Detection with Cascade DPMs (C-DPMs)

$$S_0(x) = 0,\ \mathcal{I}_0 = [1, N] \times [1, M]$$

$$S_k(x) = S_{k-1}(x) + \max_{x'} \left(U_p(x') + B_p(x', x)\right)$$

$$\mathcal{I}_k = \{x \in \mathcal{I}_{k-1} : S_{k-1}(x) \geq \theta_k\}$$

Felzenszwalb, Girschick, et al: use PCA-projection of $\mathbf{h}, \mathbf{w}$

Our work: use quick upper bounds, thresholds for full HOG

	GDTs	C-DPM	$p_e = 0.05$	$p_e = 0.01$
$\theta = -0.5$	8.95 ± 0.82	0.56 ± 0.07	0.19 ± 0.03	0.23 ± 0.04
$\theta = -0.7$	8.95 ± 0.82	0.72 ± 0.09	0.29 ± 0.04	0.36 ± 0.06
$\theta = -1.0$	8.95 ± 0.82	1.04 ± 0.16	0.51 ± 0.10	1.07 ± 0.29

Conclusions

Rapid upper and lower bounds

Blend of optimization and low-level processing

On-going work

- Part sharing
- Tighter bounds, cascades

http://vision.mas.ecp.fr/Personnel/iasonas/code.html

Agence Nationale de la Recherche ANR