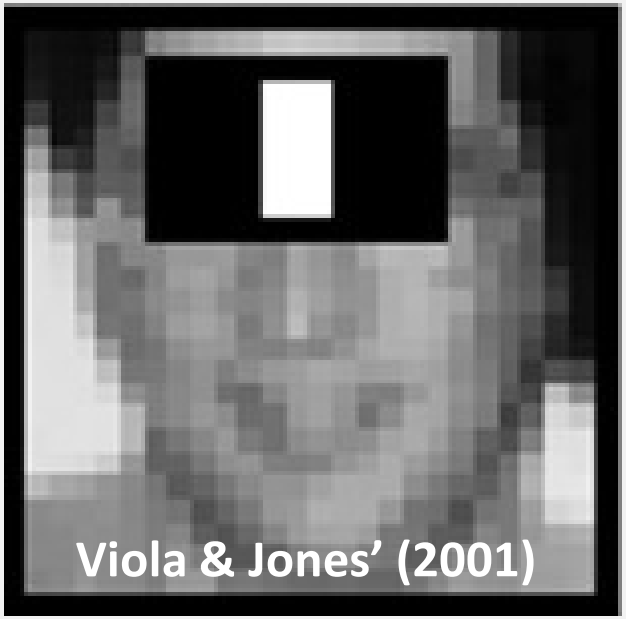
Department of Computer Science
University of Bristol

COMS30121 - Image Processing and Computer Vision

www.cs.bris.ac.uk/Teaching/Resources/COMS30121



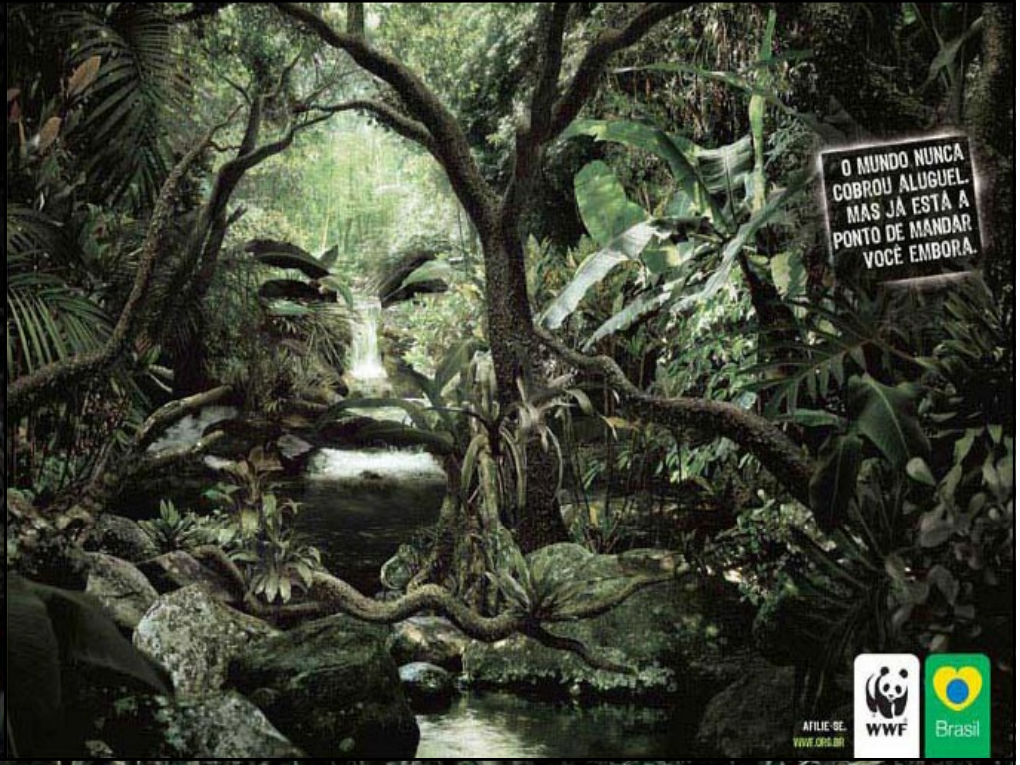
Lecture 06

Object Detection Basics I

Andrew Calway | andrew@cs.bris.ac.uk
Tilo Burghardt | tilo@cs.bris.ac.uk

14 Slides


A GENTE SÓ QUER
QUE O CÉU
CONTINUE AZUL.
A MATA CONTINUE
VERDE E SUA
CONSCIÊNCIA LIMPA.
AFILIE-SE.
WWF
Brasil


O MUNDO NUNCA
COBROU ALUGUEL.
MAS JÁ ESTÁ A
PONTO DE MANDAR
VOCÊ EMBORA.
AFILIE-SE.
WWF
Brasil


A SOLUÇÃO PARA
O AQUECIMENTO
GLOBAL NÃO
É COMPRAR UM
AR-CONDICIONADO.
AFILIE-SE.
WWF
Brasil


REFLORESTE-SE.
AFILIE-SE.
WWF
Brasil

What is ‘Object Recognition’?

- Object Recognition aims at bridging the ‘semantic gap‘ between...
  - given pixel values, *and*
  - meaningful objects (grouping of pixels + classification of groups)

→ **image regions need to be found and assigned with semantic labels from a space of object classes**

- Why do shape detection and segmentation rarely work for real-world object detection?
  - high intra-class, low inter-class variance
  - classes are rarely well defined
  - change of illumination, scale, pose + deformation, occlusion...

→ **object recognition is a difficult task**

Variable visual appearance

First Real-time Detection Method: Viola & Jones' (2001)
(base line standard for off-the-shelf method for almost a decade)

Selected Example Algorithm:
Viola & Jones' Real-time Method (2001)

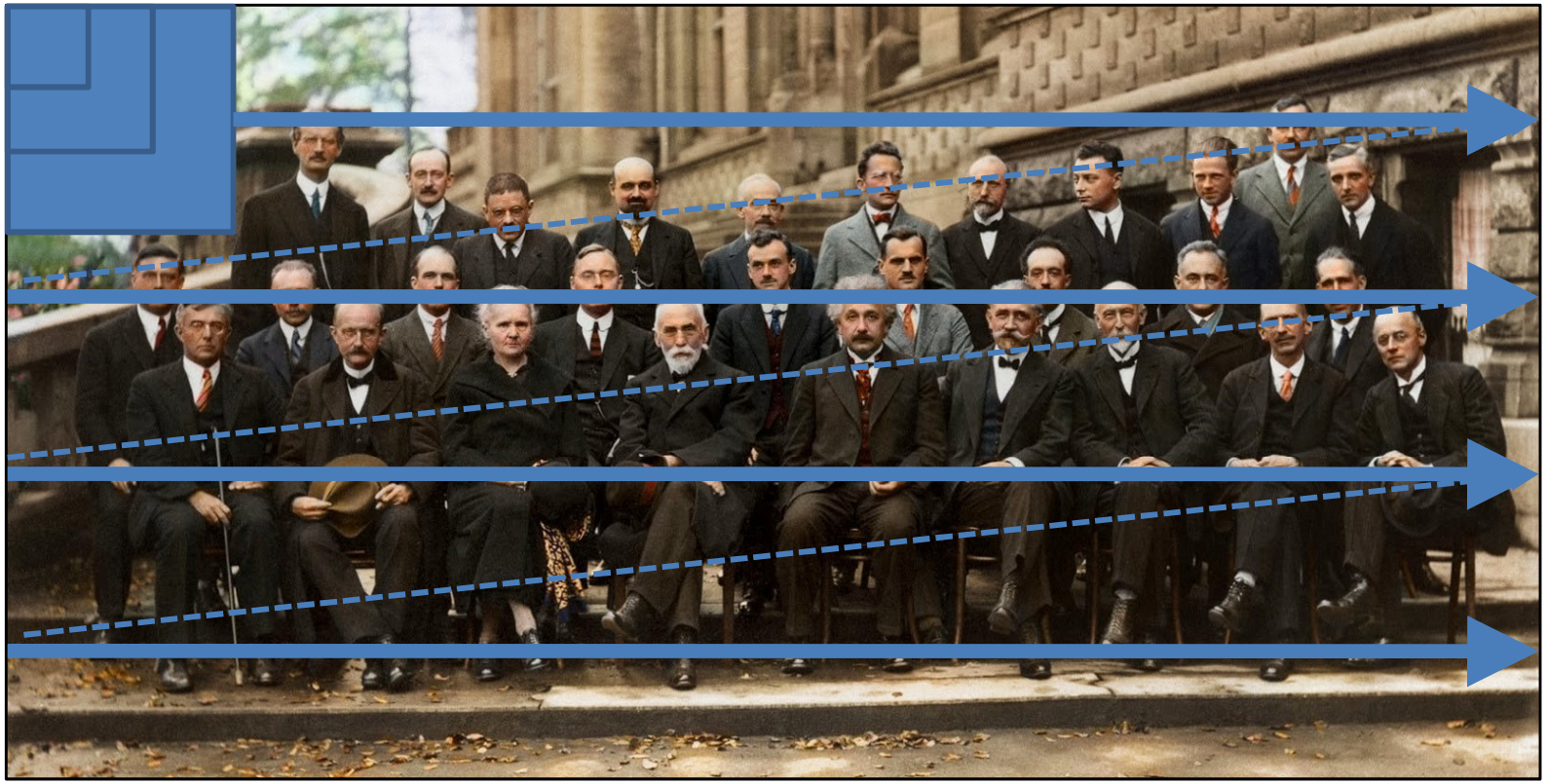
Our Agenda:

- Viola Jones technique overview
- Sliding Window Detectors
- Haar-like Features
- Feature Extraction and Integral Images
- Weak Classifiers
- Boosting and Classifier Evaluation
- Cascades of Boosted Classifiers

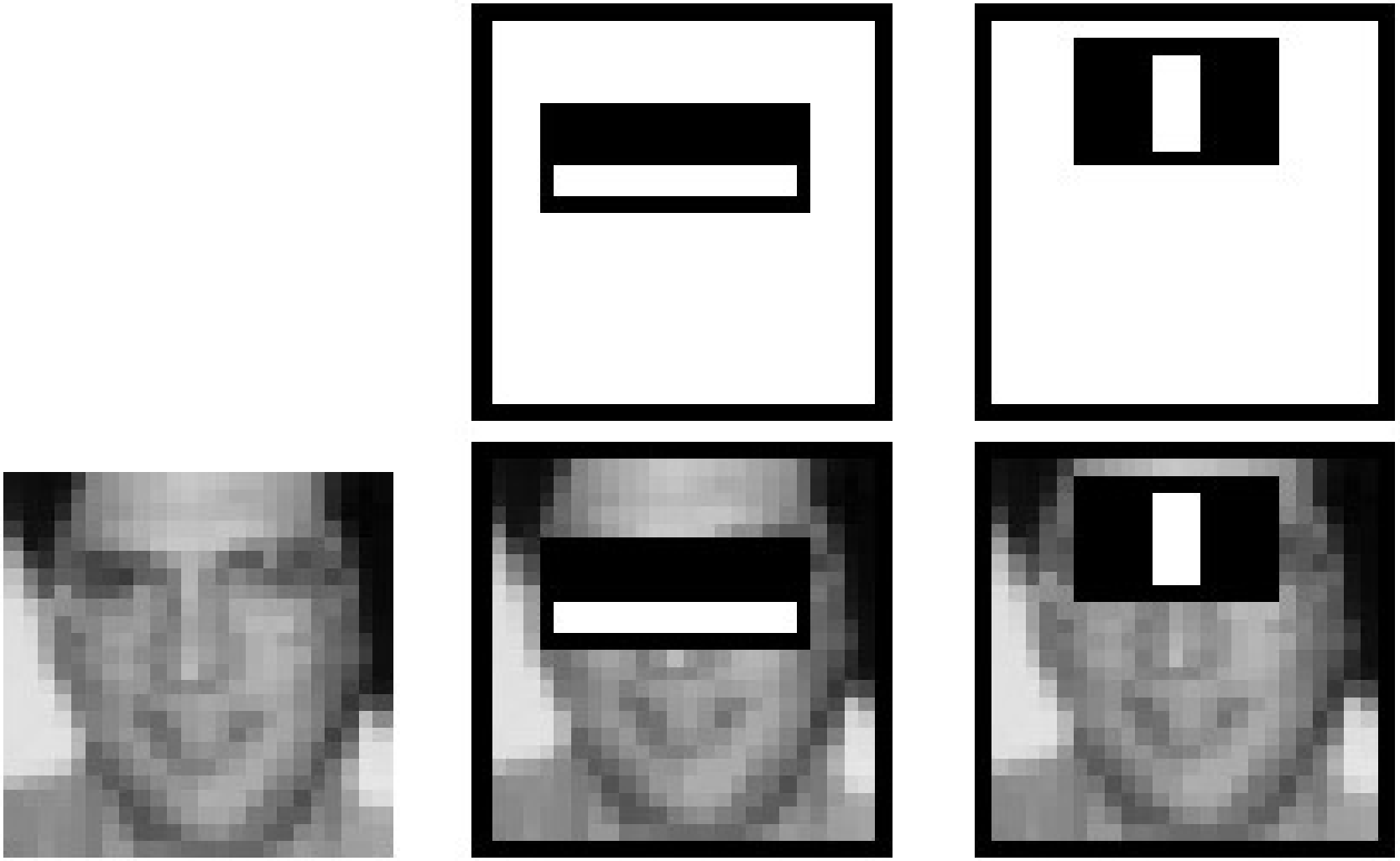
Best description of full details available in consolidated paper by Viola and Jones, International Journal of Computer Vision, 2004

Shift and Scale Invariance: Sliding Window Detectors

- image is tested for object presence window-by-window
- the window is `slided' and `scaled' throughout the image
- each resulting window is judged w.r.t. an object model giving a response indicating object presence or absence



Basic Object Model Idea:
Characteristic Set of Block Features



Viola & Jones' (2001)

Integral Images & Integration Rule

(INTEGRATION RULE OF CONVOLUTION)

$$(\mathbf{S_k} * \mathbf{I})^{[n]} = \mathbf{S_k}^{[q]} * \mathbf{I}^{[p]} \quad \text{given} \quad n = p + q$$

I

	0	1	2	3	4	5	6	7
0	1	1	1	2	3	1	2	1
1	1	2	0	0	0	3	1	1
2	1	1	1	1	1	2	3	1
3	1	1	1	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	1	1

$\sum$

II

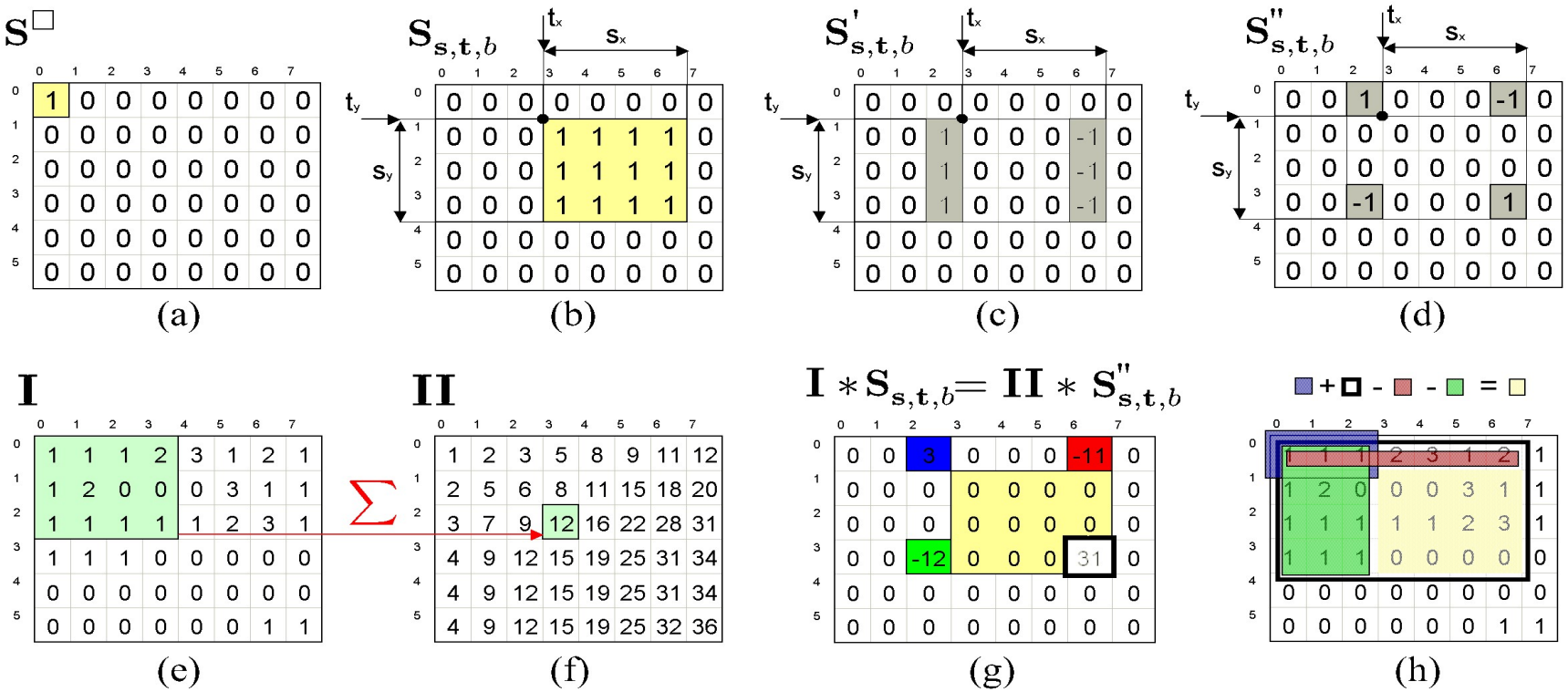
	0	1	2	3	4	5	6	7
0	1	2	3	5	8	9	11	12
1	2	5	6	8	11	15	18	20
2	3	7	9	12	16	22	28	31
3	4	9	12	15	19	25	31	34
4	4	9	12	15	19	25	31	34
5	4	9	12	15	19	25	32	36

(IMAGE INTEGRATION)

$$\mathbf{II}(-1, y) = 0; \qquad \mathbf{II}(x, y) = \mathbf{II}(x-1, y) + A(x, y);$$

$$A(x, -1) = 0; \qquad A(x, y) = A(x, y-1) + \mathbf{I}(x, y).$$

Fast 'BlockImage' Convolution



(FAST BLOCK IMAGE CONVOLUTION)

$$\mathbf{I} * \mathbf{S_k} = \mathbf{II}(\mathbf{t}_1 - 1, \mathbf{t}_2 - 1) + \mathbf{II}(\mathbf{s}_1 + \mathbf{t}_1 - 1, \mathbf{s}_2 + \mathbf{t}_2 - 1)$$
$$- \mathbf{II}(\mathbf{s}_1 + \mathbf{t}_1 - 1, \mathbf{t}_2 - 1) - \mathbf{II}(\mathbf{t}_1 - 1, \mathbf{s}_2 + \mathbf{t}_2 - 1)$$

where $\mathbf{k} = ((\mathbf{s}_1, \mathbf{s}_2), (\mathbf{t}_1, \mathbf{t}_2), b)$ holds the scale and translation parameters

Modelling Objects by Boosting

Positives Xp=((x1,y1),(x2,y2),(x3,y3),...)

Negatives Xn=((x4,y4),(x5,y5),(x6,y6),...)

t=0

Annotated Training Data

Greedy Learning Loop

Normalize Weights:

$$w_i = \frac{w_i}{\sum w_i}$$

Train Classifiers h_j on each single Haar Feature:

$$e_j = \sum_i w_i \left| h_j(x_i) - y_i \right|$$

lowest classification error → h_t

Update Weights:

$$w_i = w_i \cdot \left(\frac{e_t}{1 - e_t} \right)^{1-c_i}$$

Weak Classifier Output

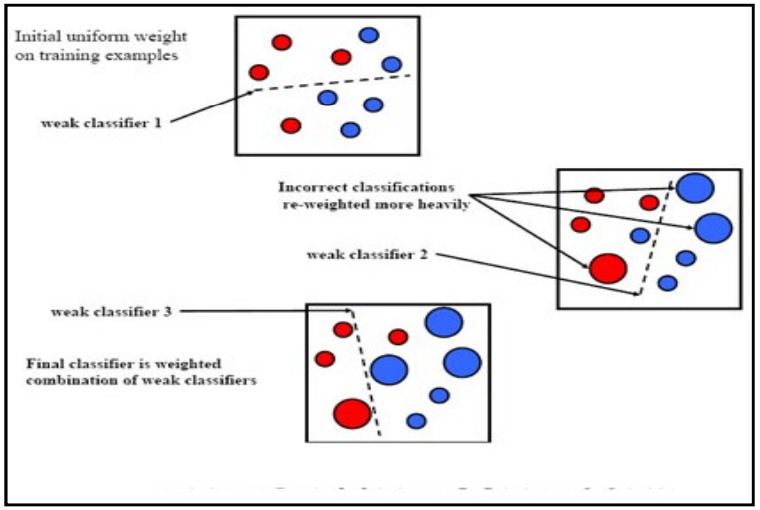
Strong Classifier Assembly

$$h(x) = \begin{cases} 1 & \sum_t \alpha_t h_t(x) \geq \frac{1}{2} \sum_t \alpha_t \\ 0 & otherwise \end{cases} \quad ; \alpha_t = \log \frac{1 - e_t}{e_t}$$

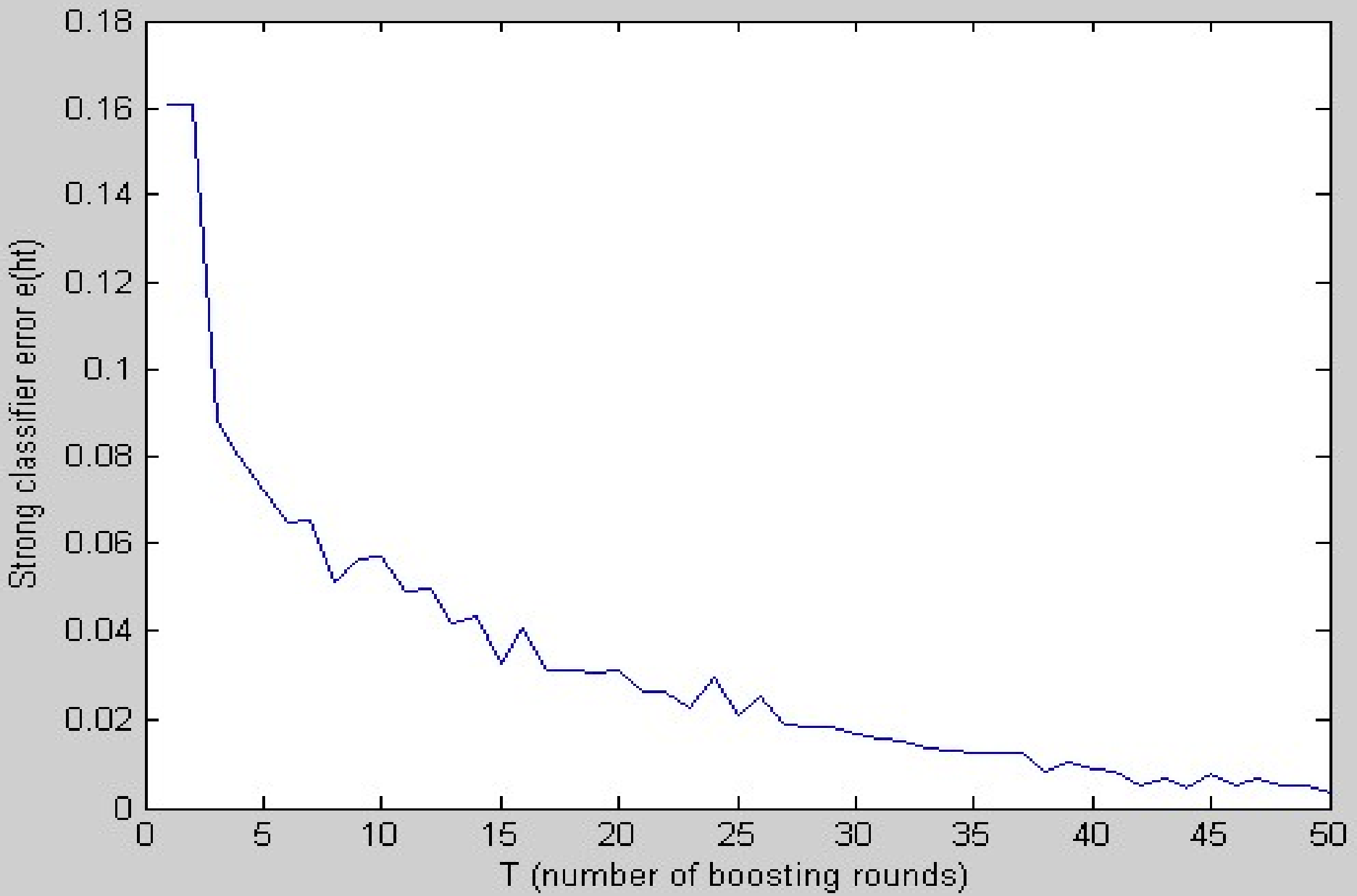
Strong Classifier Output

AdaBoost Algorithm

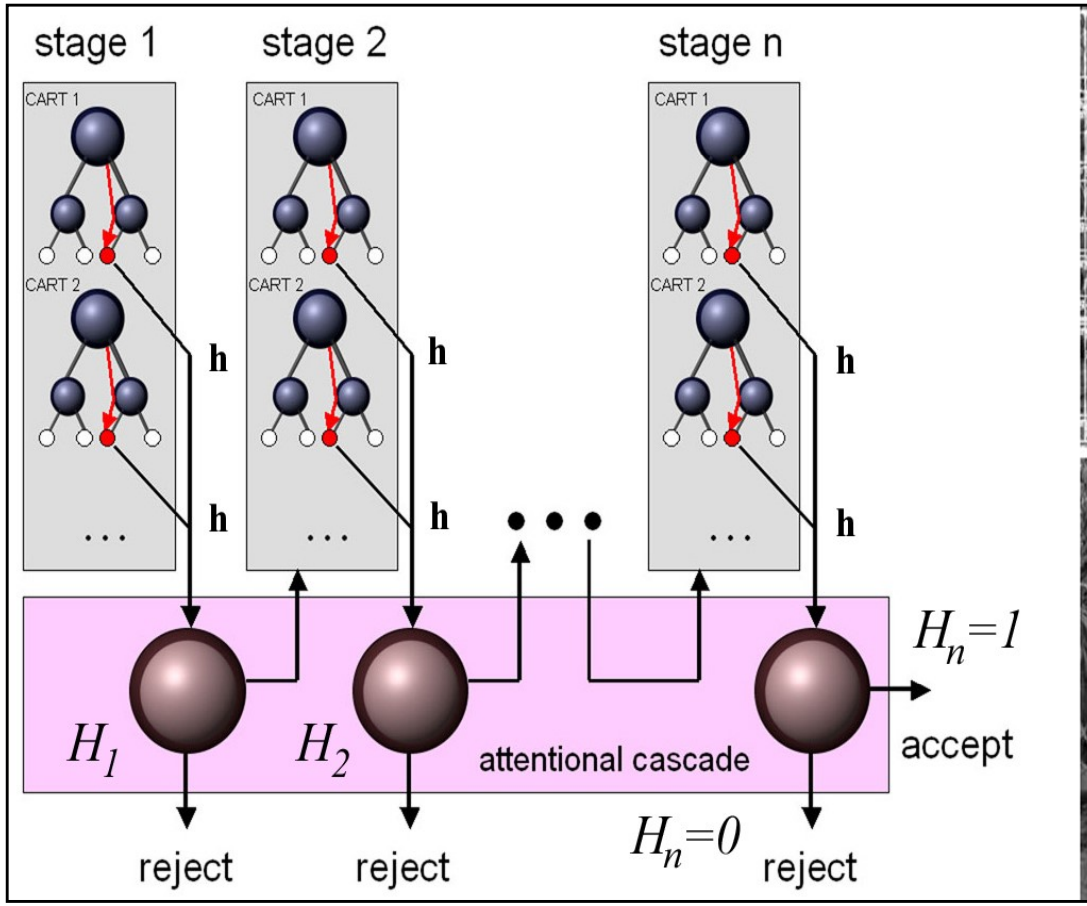
(see paper by Viola and Jones 2004)

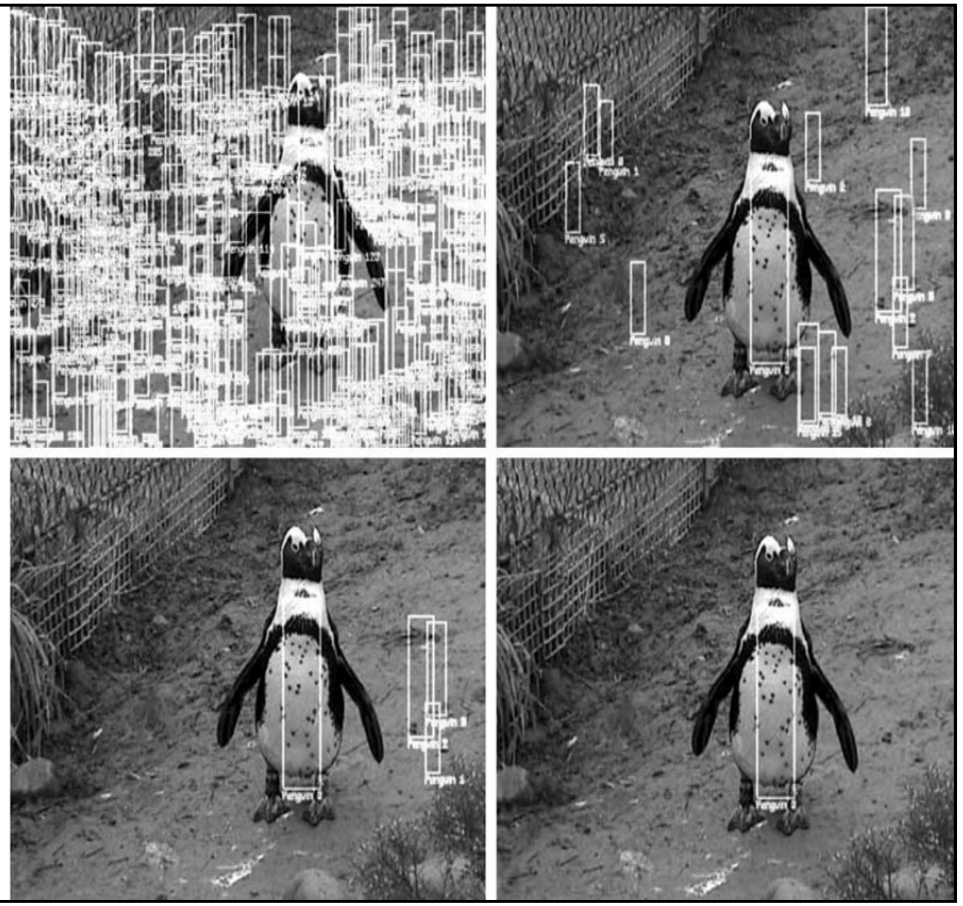


Error Reduction as Boosting adds Classifiers



Concept of Attentional Cascading





On Window Resolution