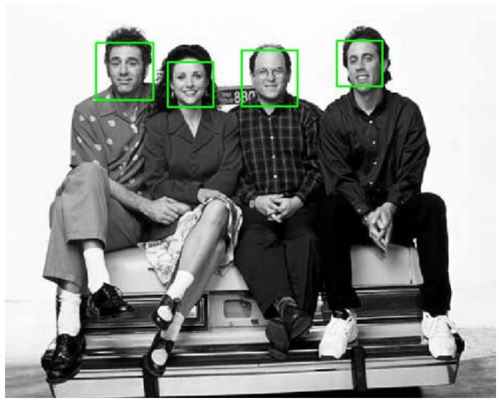
CS4670 / 5670: Computer Vision

Noah Snavely

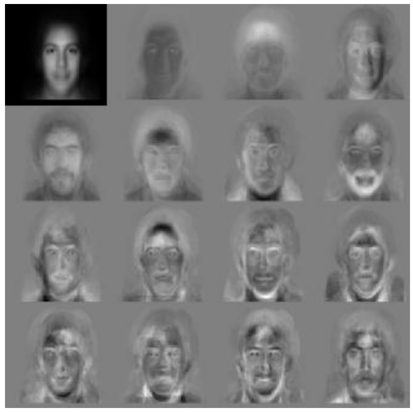
Lecture 29: Face Detection Revisited



Announcements

- Project 4 due next Friday by 11:59pm

Remember eigenfaces?



- They don't work very well for detection

Issues: speed, features

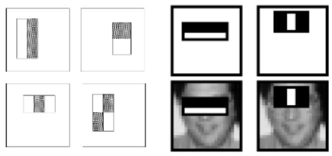
- Case study: Viola Jones face detector
- Exploits two key strategies:
  - simple, super-efficient, but useful features
  - pruning (cascaded classifiers)

- Next few slides adapted Grauman & Liebe's tutorial
  - http://www.vision.ee.ethz.ch/~bleibe/teaching/tutorial-aaai08/
- Also see Paul Viola's talk (video)
  - http://www.cs.washington.edu/education/courses/577/04sp/contents.html#DM

Visual Object Recognition Tutorial

Feature extraction

"Rectangular" filters

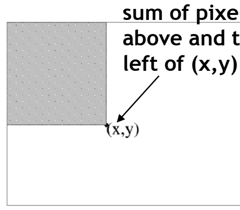


Feature output is difference between adjacent regions

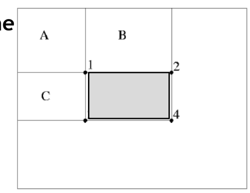
Efficiently computable with integral image: any sum can be computed in constant time

Avoid scaling images → scale features directly for same cost

Value at (x,y) is sum of pixels above and to the left of (x,y)



Integral image



$$D = 1 + 4 - (2 + 3)$$
$$= A + (A + B + C + D) - (A + C + A + B)$$
$$= D$$

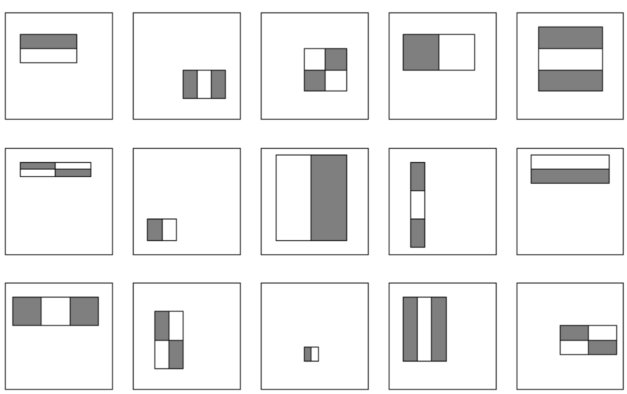
Viola & Jones, CVPR 2001

K. Grauman, B. Leibe

5

Visual Object Recognition Tutorial

Large library of filters



Considering all possible filter parameters: position, scale, and type:

180,000+ possible features associated with each 24 x 24 window

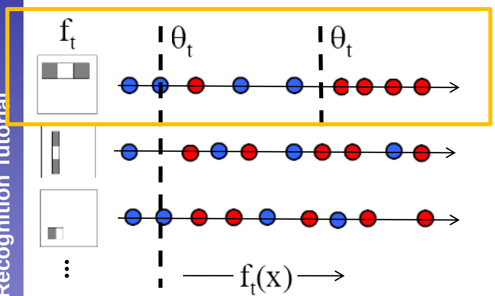
Use AdaBoost both to select the informative features and to form the classifier

Viola & Jones, CVPR 2001

K. Grauman, B. Leibe

AdaBoost for feature+classifier selection

- Want to select the single rectangle feature and threshold that best separates positive (faces) and negative (non-faces) training examples, in terms of *weighted* error.



Outputs of a possible rectangle feature on faces and non-faces.

Resulting weak classifier:

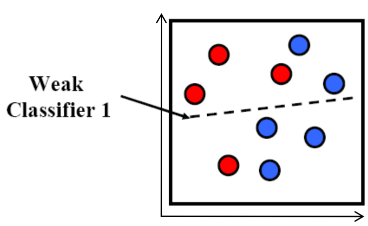
$$h_t(x) = \begin{cases} +1 & \text{if } f_t(x) > \theta_t \\ -1 & \text{otherwise} \end{cases}$$

For next round, reweight the examples according to errors, choose another filter/threshold combo.

Viola & Jones, CVPR 2001

K. Grauman, B. Leibe

AdaBoost: Intuition



Consider a 2-d feature space with positive and negative examples.

Each weak classifier splits the training examples with at least 50% accuracy.

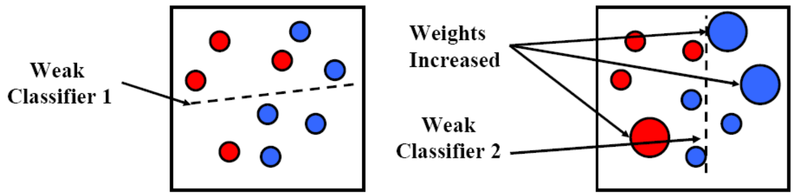
Examples misclassified by a previous weak learner are given more emphasis at future rounds.

Figure adapted from Freund and Schapire

K. Grauman, B. Leibe

Visual Object Recognition Tutorial

AdaBoost: Intuition

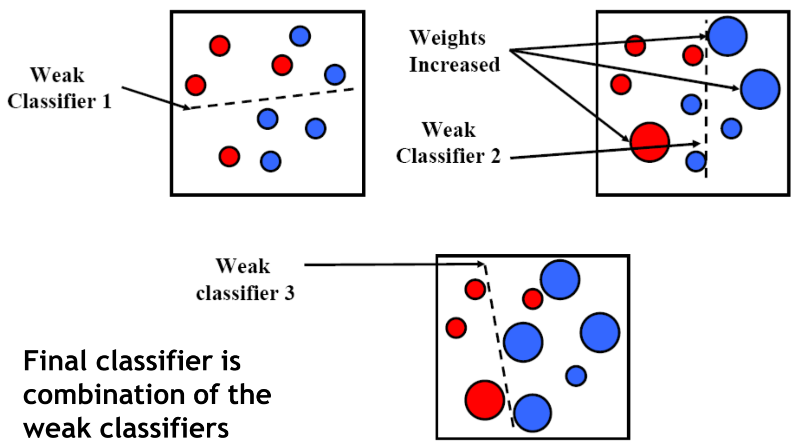


K. Grauman, B. Leibe

9

Visual Object Recognition Tutorial

AdaBoost: Intuition



Final classifier is combination of the weak classifiers

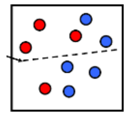
K. Grauman, B. Leibe

10

- Given example images $(x_1, y_1), \ldots, (x_n, y_n)$ where $y_i = 0, 1$ for negative and positive examples respectively.
- Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0, 1$ respectively, where m and l are the number of negatives and positives respectively.
- For $t = 1, \ldots, T$:
  1. Normalize the weights,
     $$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$$
     so that w_t is a probability distribution.
  2. For each feature, j, train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to w_t, $\epsilon_j = \sum_i w_i \left| h_j(x_i) - y_i \right|$.
  3. Choose the classifier, h_t, with the lowest error ϵ_t.
  4. Update the weights:
     $$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$
     where $e_i = 0$ if example x_i is classified correctly, $e_i = 1$ otherwise, and $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$.
- The final strong classifier is:
$$h(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases}$$
where $\alpha_t = \log \frac{1}{\beta_t}$

AdaBoost Algorithm

Start with uniform weights on training examples



For T rounds

Evaluate *weighted* error for each feature, pick best.

Re-weight the examples:
Incorrectly classified -> more weight
Correctly classified -> less weight

Final classifier is combination of the weak ones, weighted according to error they had.

an, B. Leibe

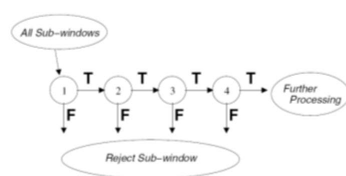
Freund & Schapire 1995

Visual Object Recognition Tutorial

Cascading classifiers for detection

For efficiency, apply less accurate but faster classifiers first to immediately discard windows that clearly appear to be negative; e.g.,

- Filter for promising regions with an initial inexpensive classifier
- Build a chain of classifiers, choosing cheap ones with low false negative rates early in the chain



Fleuret & Geman, IJCV 2001
Rowley et al., PAMI 1998
Viola & Jones, CVPR 2001

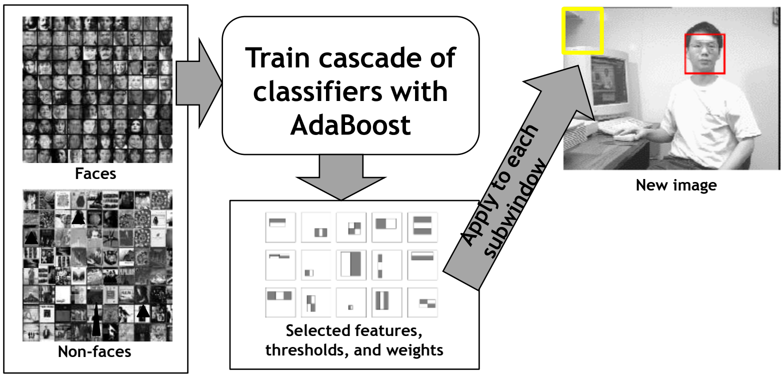
K. Grauman, B. Leibe

Figure from Viola & Jones CVPR 2001

12

Visual Object Recognition Tutorial

Viola-Jones Face Detector: Summary



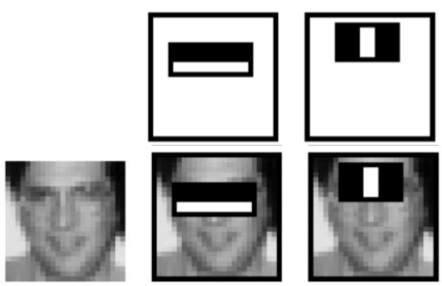
- Train with 5K positives, 350M negatives
- Real-time detector using 38 layer cascade
- 6061 features in final layer
- [Implementation available in OpenCV: http://www.intel.com/technology/computing/opencv/]

K. Grauman, B. Leibe

13

Visual Object Recognition Tutorial

Viola-Jones Face Detector: Results



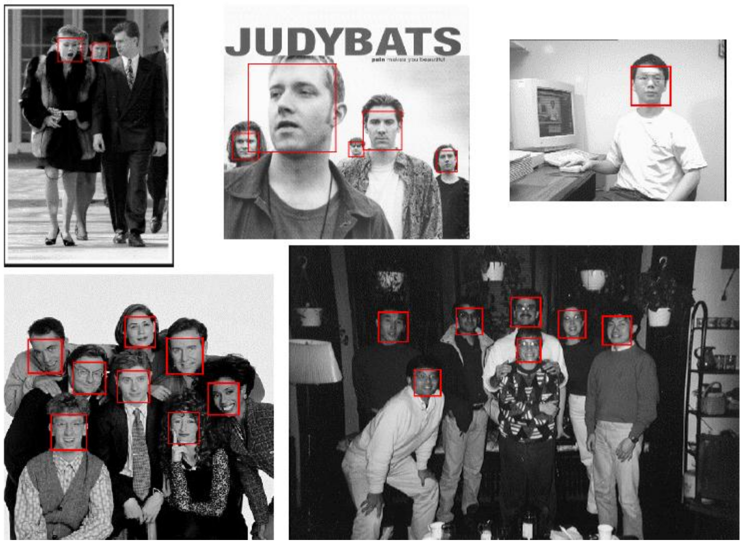
First two features selected

K. Grauman, B. Leibe

14

Viola-Jones Face Detector: Results

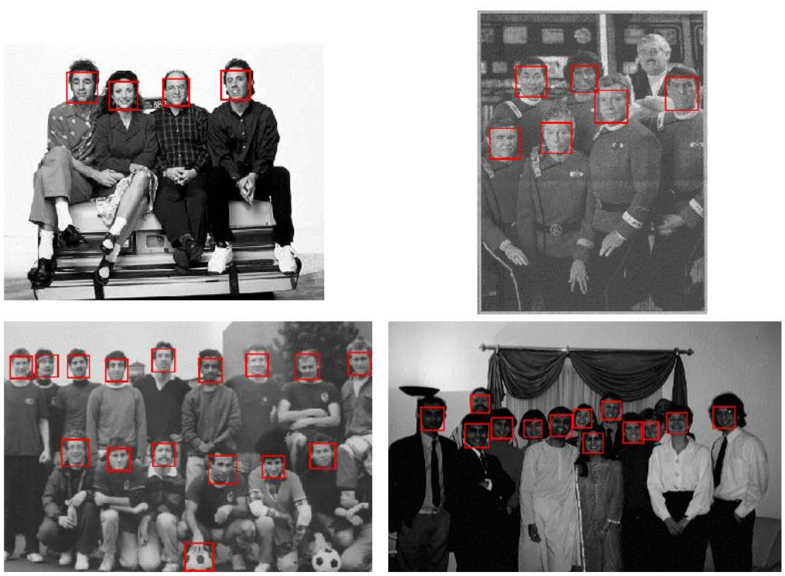
Visual Object Recognition Tutorial



K. Grauman, B. Leibe

Viola-Jones Face Detector: Results

Visual Object Recognition Tutorial



K. Grauman, B. Leibe

Visual Object Recognition Tutorial

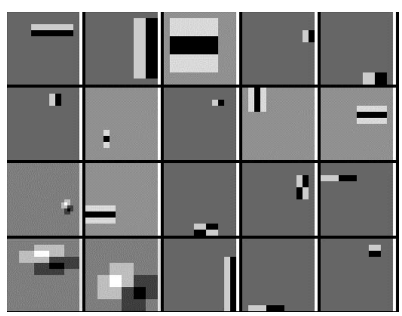
Viola-Jones Face Detector: Results

K. Grauman, B. Leibe

Visual Object Recognition Tutorial

Detecting profile faces?

Detecting profile faces requires training separate detector with profile examples.



K. Grauman, B. Leibe

Viola-Jones Face Detector: Results

Visual Object Recognition Tutorial

Paul

K. Grauman, B. Leibe

Questions?