A Regularized Framework for Feature Selection in Face Detection and Authentication

Augusto Destrero, **Christine De Mol**, **Francesca Odone**, **Alessandro Verri**, **2008**

Present By

Mr. Apichon Witayangkurn, CSIM

ID: 106800

AT70.9011: Machine Vision for Robotics and HCI, Summer 2009

Content

- Introduction
- Feature Selection for Large Computer Vision Problems
- Face Detection
- Face Authentication
- Discussion

Introduction

- Proposes a general framework for selecting features.
- System requirement is a *Face Detection* module able to process the video in *real-time* and return a good localization of the frontal faces appearing in the video. Then, *Face Authentication* is performed on a video portion.
- *Rectangle features* for face detection and *Local Binary Pattern (LBP) features* for face authentication.

Feature Selection for Large Computer Vision Problem

- Thresholded Landweber
  - We consider the case of a linear dependence between input and output data, which means that the problem can be reformulated as the solution of the following linear system of equations:

$$g = Af$$

  - where $A = \{A_{ij}\}$, $i = 1, \ldots, n$ and $j = 1, \ldots, p$ is the n x p.
  - Feature matrix obtained representing the training set of n elements with a dictionary of p features

Feature Selection for Large Computer Vision Problem

- Thresholded Landweber
  - Lasso regression dealing with a feature selection problem.

$$f_L = \arg\min_f \{|g - Af|_2^2 + 2\tau |f|_1\}$$

  - where $|f|_1 = \sum_j |f_j|$ is the L1-norm of f.
  - τ is a regularization parameter regulating the balance between the data misfit and the penalty

Feature Selection for Large Computer Vision Problem

- Thresholded Landweber
  - Adopt a simple iterative strategy

$$f_L^{(t+1)} = S_\tau[f_L^{(t)} + A^\top(g - Af_L^{(t)})] \quad t = 0, 1, \ldots \tag{3}$$

with arbitrary initial vector $f_L^{(0)}$, where S_τ is the following "soft-thresholder"

$$(S_\tau h)_j = \begin{cases} h_j - \tau \operatorname{sign}(h_j) & \text{if } |h_j| \geq \tau, \\ 0 & \text{otherwise.} \end{cases}$$

Feature Selection for Large Computer Vision Problem

- Sampled Version of the Thresholded Landweber
  - The linear problems that we are about to build will be rather large.
  - Training set is made of *1000 positive and 1000 negative examples*, then a matrix A of *1 Gb* size will be easily obtained.
  - Based on resampling the features set and obtaining many smaller problems.
  - Subsets with about *10%* of the original feature set size. To choose the number of sub-problems S, we rely on the *binomial distribution* and estimate how many extractions are needed so that each feature is extracted at least 10 times with high probability.

Feature Selection for Large Computer Vision Problem

- Speeding up Feature Selection
  - The *speeding up heuristics* that suggested by two empirical considerations:
    - (a) for a given τ , a high percentage of the features is discarded in the first iterations;
    - (b) when a feature weight goes to zero it will not change in the next iterations.
  - The solution to the problem is reached from *2 to 6 times faster*, depending on the original size of the matrix.

Feature Selection for Large Computer Vision Problem

- The Choice of a Classifier
  - An *SVM classifier*, well known for its good generalization ability, on the selected features.
  - This method allows to achieve considerable space-time efficiency at run time, and this suits our requirement of real-time processing.

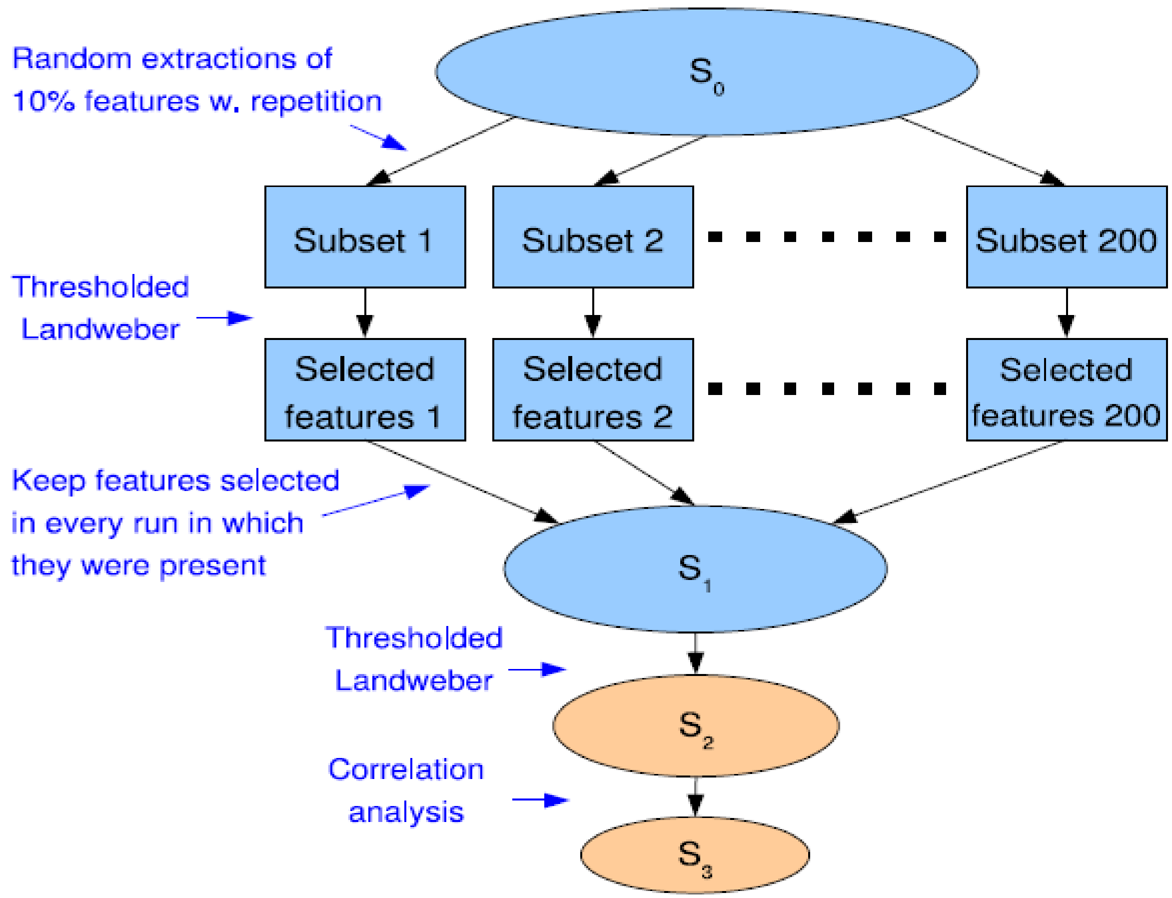
Face Detection

- Based on the *rectangle features* computed over different locations, sizes, and aspect ratios for each image.
- The size of training images: 19x19 pixels and compute about 64,000 features per image/patch.
- A dataset of *4000 training data*, evenly distributed between positive and negative, 2000 validation and *3400 test data*.
- Then use *iterative algorithm* to reduce number of redundancy feature points.
- Associated to an *eye detector* obtained from the same protocol and trained with eye-pairs from the FERET dataset.

Face Detection

- The structure of the 3 stages feature selection.

  - S_0 = 64,000 features.
  - S_1 = 345 features.
  - S_2 = 247 points.
  - S_3 = 42 points.



Face Detection

- Comparison of our second selection stage against others reduction method.

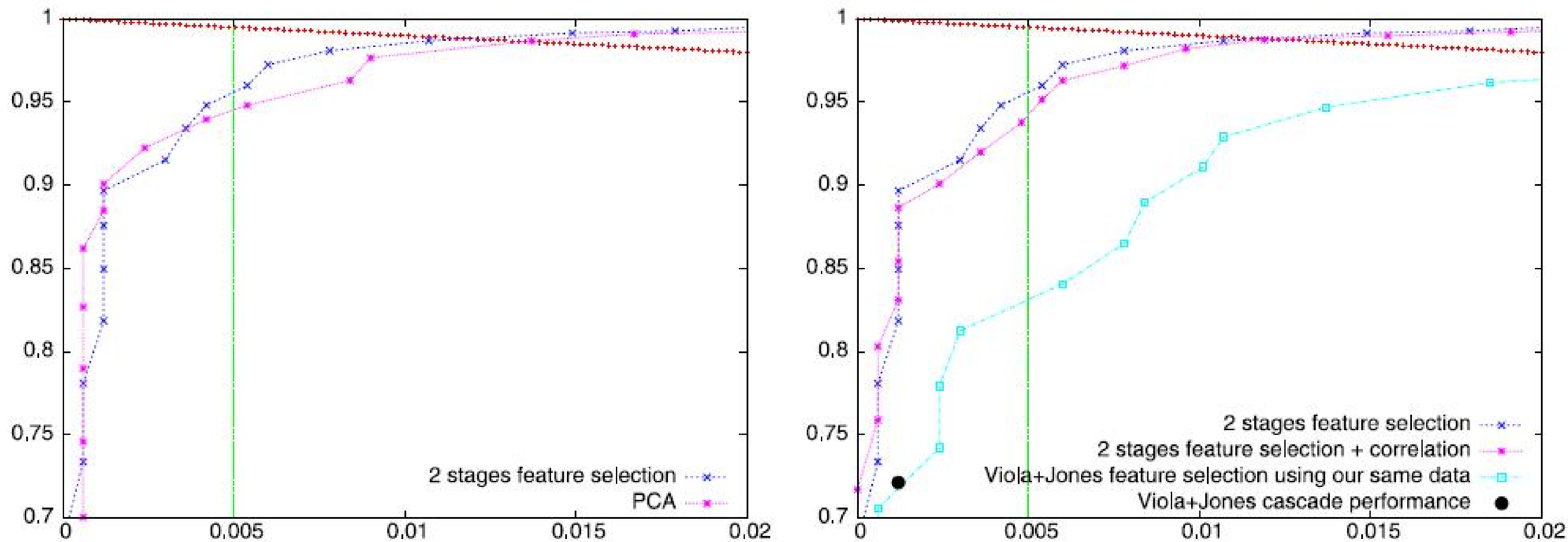
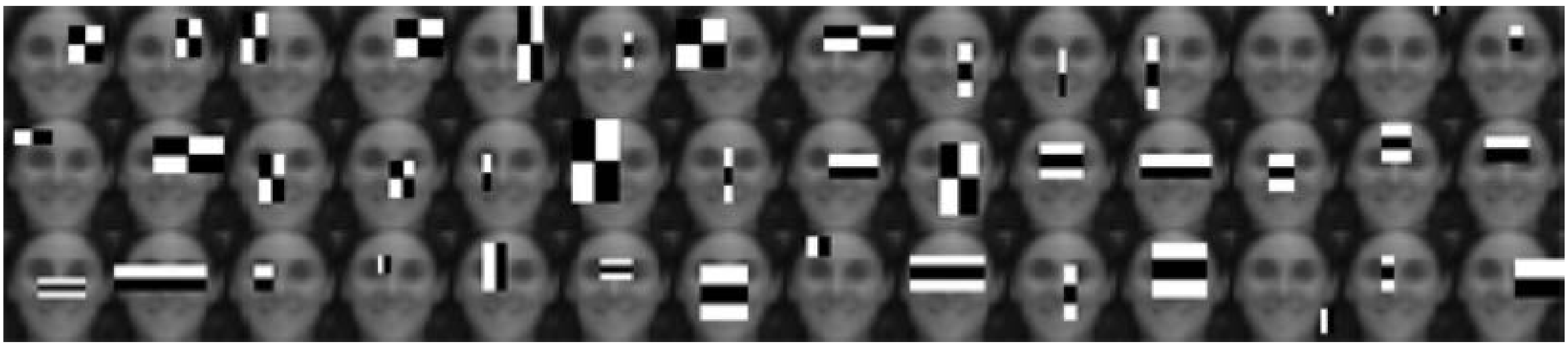


Fig. 3 Comparison of our second selection stage against other dimensionality reduction methods applied to set S_1. *Left*: PCA; *Right*: Adaboost features obtained from our same pool of data (see text)

Face Detection

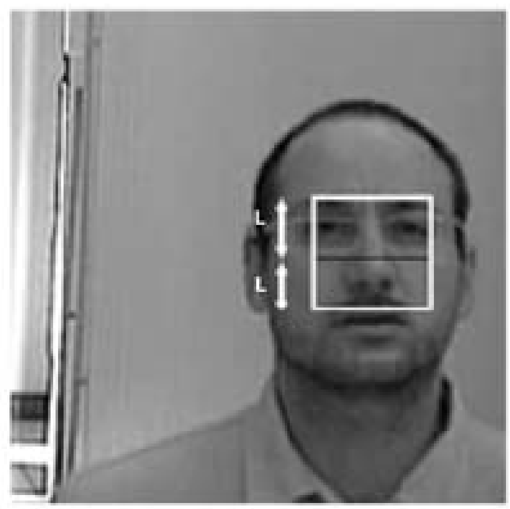
- The 42 features that are left after a third stage of correlation analysis (S3).



Face Detection

- The *eye detector* is applied to the face detector hits to discard false positives and non frontal faces. A second important is to automatically register faces to the purpose of authentication.
- The final set of faces is resized to 40x40 pixels.


face and eye detection


face normalization


cropped face

Face Authentication

- Closely related to face recognition.
- The probe (test) is composed by a test image (or test sequence) and an associated identity.
- Verify whether he/she is a genuine or an impostor.
- Analyze *intra-personal* and *extra-personal* variations.

Face Authentication

- Feature Selection for Authentication
  - Dataset is given by face detection module.
  - Manually labeled the stored videos and built models for all the individuals. In total, *15 individuals* were included in the training phase, while a total of *64 individuals* were gathered for testing.
  - For each image, we compute *Local Binary Pattern (LBP) histograms* on rectangular regions of at least 3x3 pixels, on all locations and aspect ratio.
  - Then compare corresponding LBP histograms between images using the *$\chi 2$ distance*.
  - Select features following the same *3-stages protocol* same as face detection.

Face Authentication

- Evaluation of the feature obtained from the 3 stages.
  - Three results are comparable.
  - A slight performance loss when adding the third stage

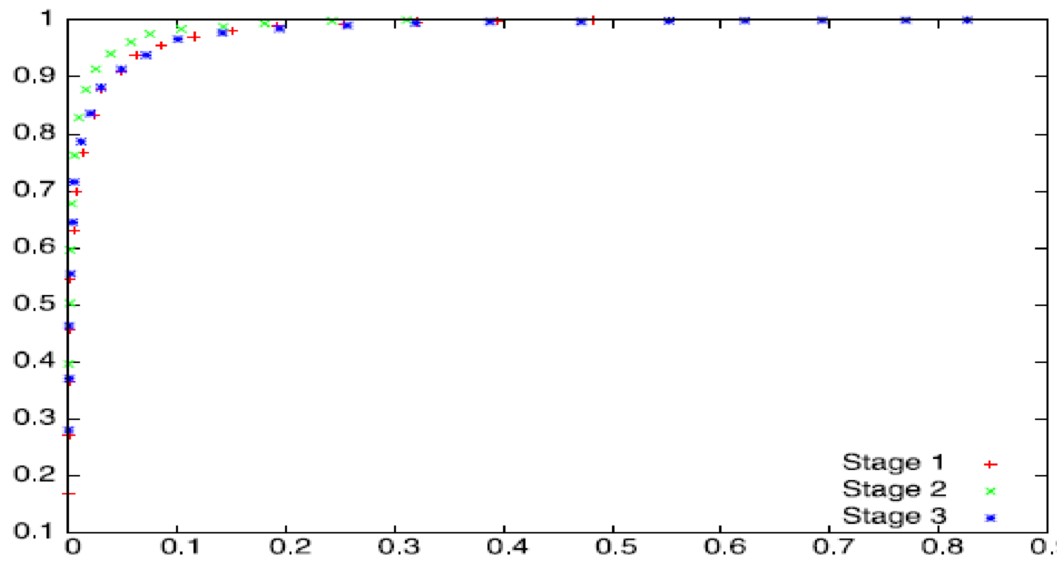


Fig. 7 Evaluation of the features obtained from the three stages, for an individual arbitrarily chosen

Face Authentication

- For person 6, for instance, *the beard is spotted*.
- For person 8 seems to be the top part, *the eyes region*.
- Person 14 is mainly characterized by *vertical features*.

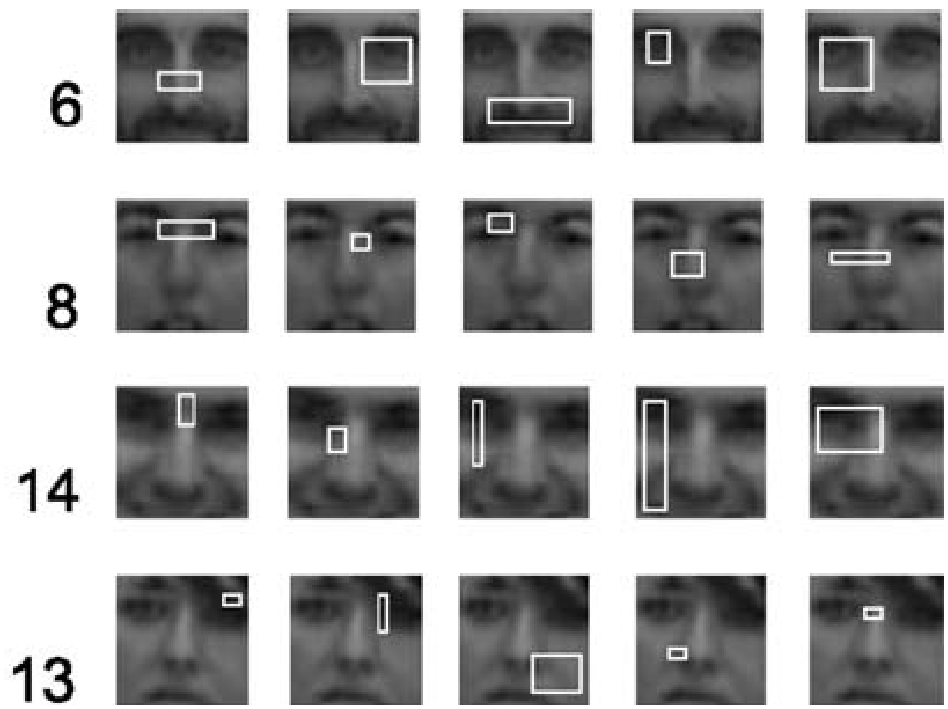


Fig. 8 Top 5 features for some individuals: they often capture distinctive face regions

Face Authentication

- Compare the performance of our method with PCA

Individual	# of features	LBP		PCA	
		Precision	Recall	Precision	Recall
1	17	0.92	1.00	0.91	0.98
2	12	1.00	0.86	0.65	0.82
3	19	1.00	0.96	0.78	0.95
4	19	0.99	0.86	0.54	0.88
5	16	0.96	0.90	0.81	0.64
6	13	1.00	1.00	0.78	1.00
7	14	0.95	0.94	0.84	0.92
8	13	0.98	1.00	0.65	0.91
9	25	0.94	0.87	0.67	0.91
10	14	0.97	0.78	0.65	0.61
11	19	0.99	0.98	0.65	0.98
12	20	0.96	0.96	0.52	0.79
13	7	0.91	0.88	0.50	0.28
14	16	1.00	0.99	0.63	0.99
15	23	0.98	0.99	0.73	0.99

Face Authentication

- Test the effectiveness of the authentication by run each classifier on the test set.

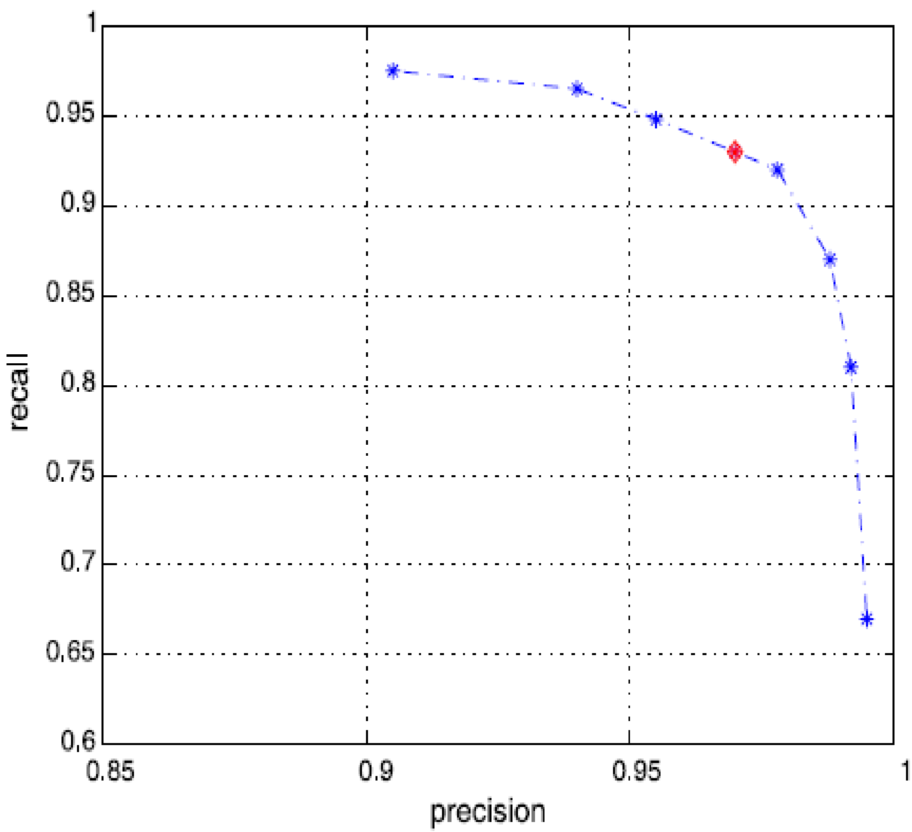


Fig. 9 The trend of precision-recall obtained by varying from 20 to 90 the percentage of positive test data obtained from a single probe image. The *darker spot* indicates the precision-recall at 50% currently used by our system

Table 3 Comparison between average authentication results obtained with automatic feature selection of the LBP features most appropriate for each individual (averages of the results reported in Table 2) and manually set LBP features according to Ahonen et al. (2006)

Automatic LBP selection		Manual selection	
Precision	Recall	Precision	Recall
0.97	0.93	0.93	0.86

Face Authentication

- A more challenging comparison can be obtained by replacing in the presented pipeline our automatic feature selection with the manual selection of LBP features.
- This approach selects the *most meaningful features per each individual*, while the manual selection is based on *extracting meaningful areas for an average face* (eyes, nose, and mouth regions).

Table 3 Comparison between average authentication results obtained with automatic feature selection of the LBP features most appropriate for each individual (averages of the results reported in Table 2) and manually set LBP features according to Ahonen et al. (2006)

Automatic LBP selection		Manual selection	
Precision	Recall	Precision	Recall
0.97	0.93	0.93	0.86

Discussion

- Apply another image classification task.
  - A pedestrian detection system.
  - Three different datasets: MIT, USC, and Daimler Chrysler dataset.
  - Dataset size is of 1404 images.

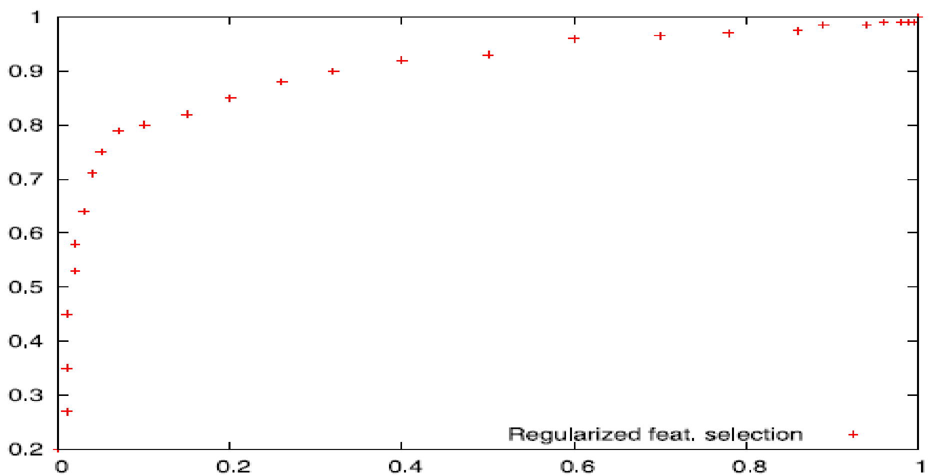


Fig. 12 Performance achieved with a pedestrian detection system automatically built with our data-driven procedure

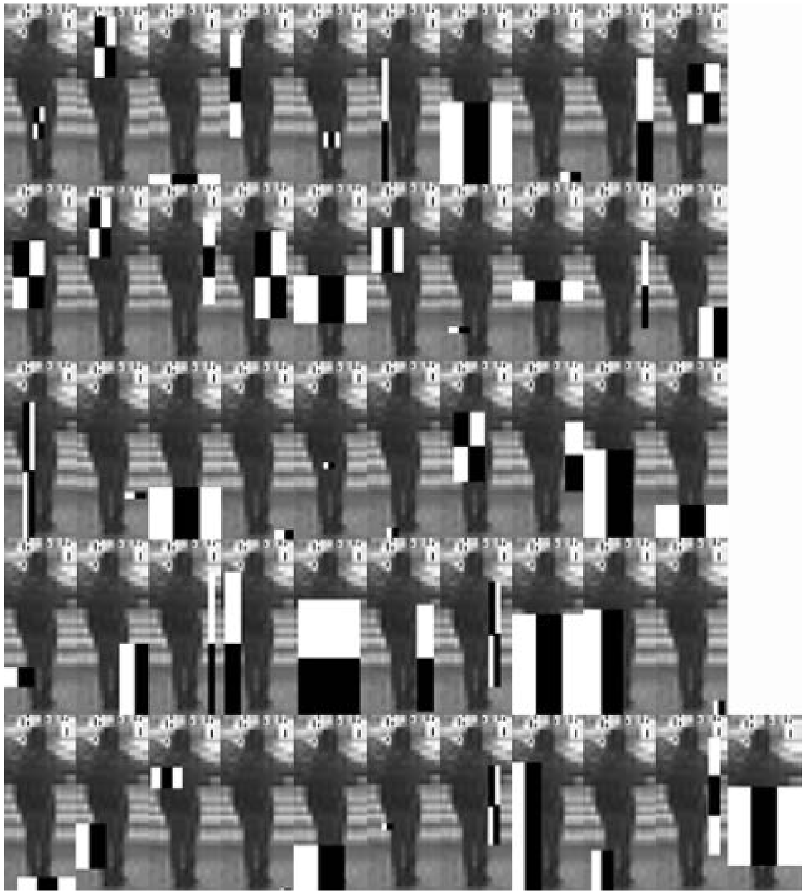


Fig. 11 The 51 features most representative of the pedestrian class automatically extracted by our 3-stages feature selection method

Discussion (con't)

- This work is built around an *iterative algorithm*, the *Thresholded Landweber* algorithm, implementing the so-called *Lasso* scheme.
- Propose effective solutions to the crucial problems
  - The need for solving large systems
  - The request for limiting the computational cost of the training phase coming from real world applications (up to 6 times faster using heuristics).
  - The processing speed at run time is about 8 fps with a PAL frame format.
- A future work will be an extension of the face authentication that uses *EVLBP features* to construct new feature matrix.

Thank you