Using a sparse learning Relevance Vector Machine in Facial Expression Recognition

Dragoș Datcu

April 14, 2006

Introduction

The utility of facial expression recognition technology:

- Human computer interfaces
- Safety, surveillance
  - terorist identification
  - safe driving, somnolence detection
  - access control
- Psychology

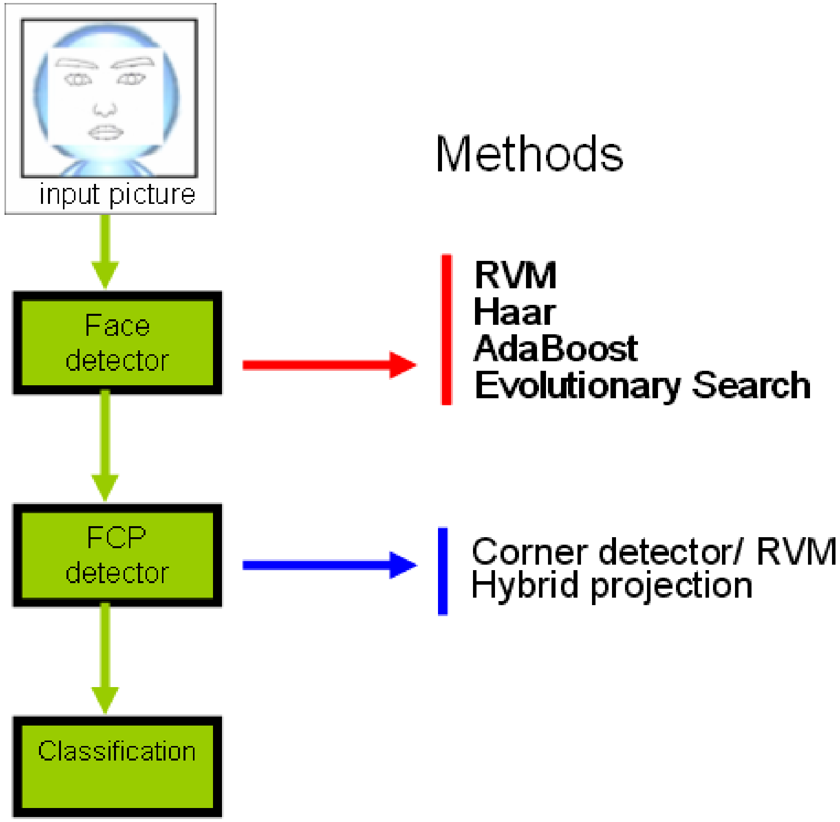
Problem definition

How to realize a fully automatic facial expression recognition system using a sparse learning Relevance Vector Machine?

The automatic facial expression recognition system includes:

- **face detector**
- **facial feature extractor** for *mouth*, *left* and *right eye*
- **Facial Characteristic Point - FCP extractor**
- **facial expression recognizer**

Facial expression recognition system



I. Face detection

The method makes use of:

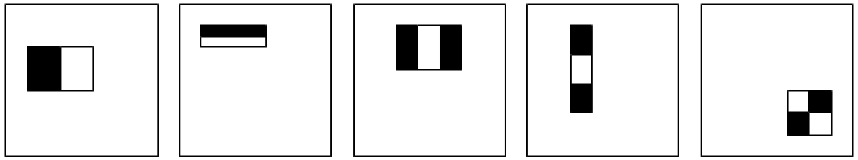
- Viola&Jones features (24x24 size samples, 162336 features/sample)
- Evolutionary AdaBoost (150 size population)
- Relevance Vector Machine (RVM) - weak classifier

- training data set includes: 4916 faces and 10000 non-faces.

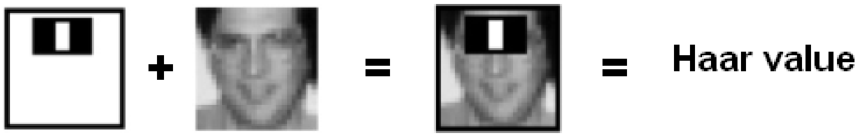
The detector consists in a 32 layer cascade of classifiers using 4297 features

Viola&Jones features

The basic types:



Applied on an image:



The value is: Haar value $= \sum Pixel_intensity_{dark_areas} - \sum Pixel_intensity_{light_areas}$

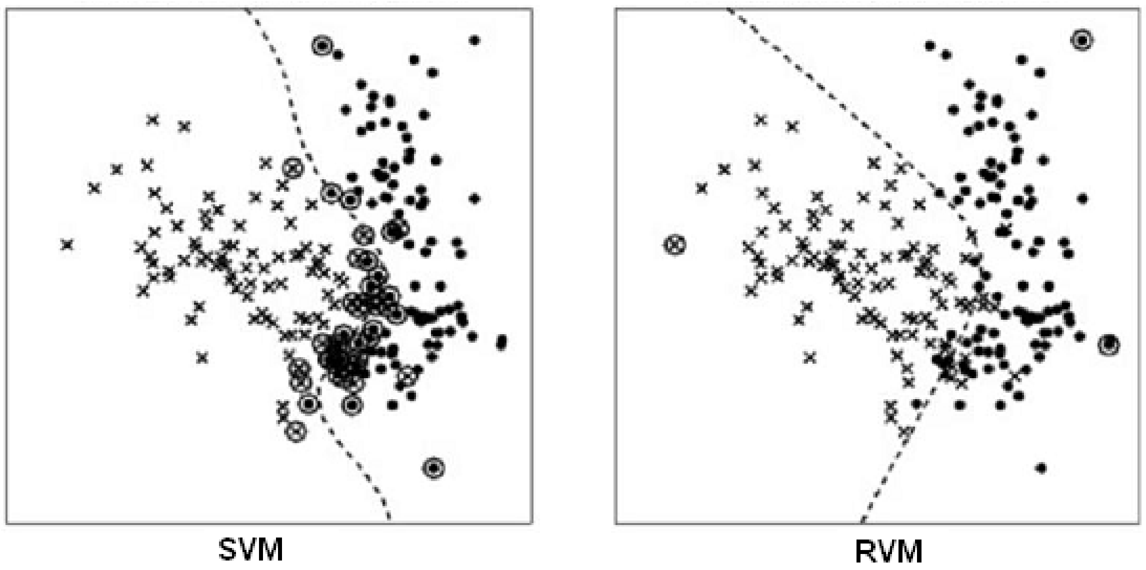
For a 24x24 image, there exist more than 160.000 such features.

AdaBoost algorithm

Discrete AdaBoost [Freund and Schapire (1996b)]

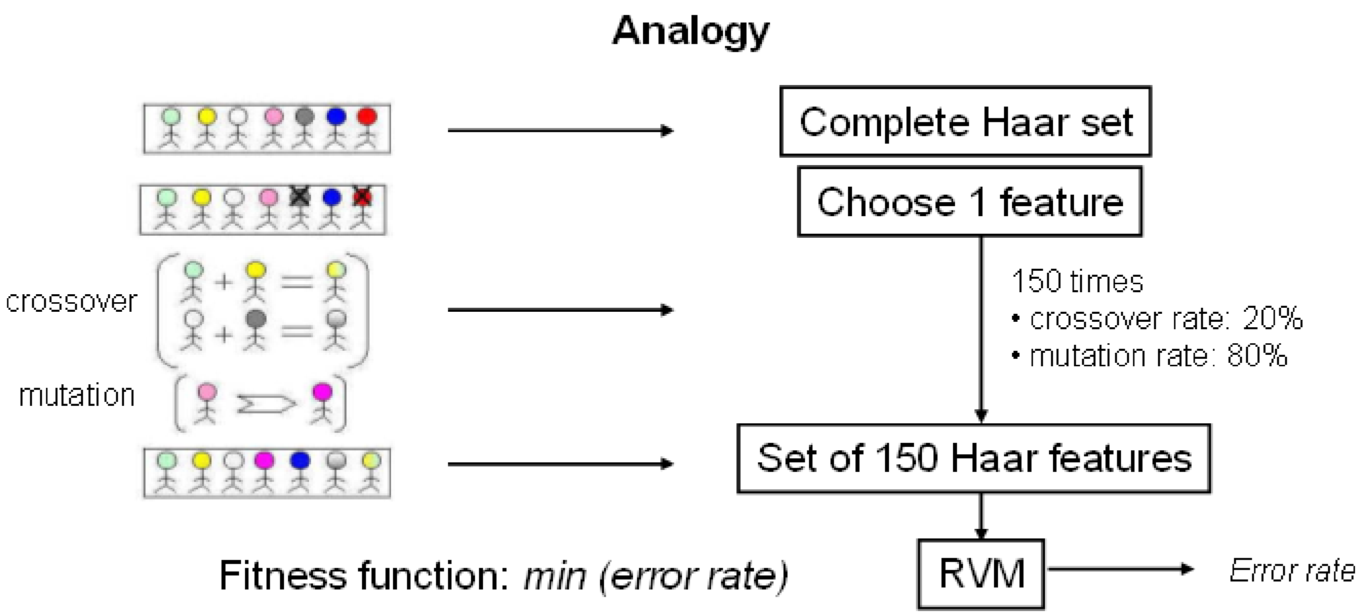
1. Start with weights $w_i = 1/N$, $i = 1, ..., N$.
2. Repeat for $m = 1, 2, ..., M$:
   (a) Fit the classifier $f_m(c) \in \{-1, 1\}$ using weights w_i on the training data.
   (b) Compute $err_m = E_w[1_{(y \neq f_m(x))}]$, $c_m = log((1 - err_m)/err_m)$.
   (c) Set $w_i \leftarrow w_i exp[c_m 1_{(y_i \neq f_m(x_i))}]$, $i = 1, 2, ..., N$, and renormalize so that $\sum_i w_i = 1$.
3. Output the classifier $sign[\sum_{m=1}^{M} c_m f_m(x)]$.

The RVM-based weak classifier

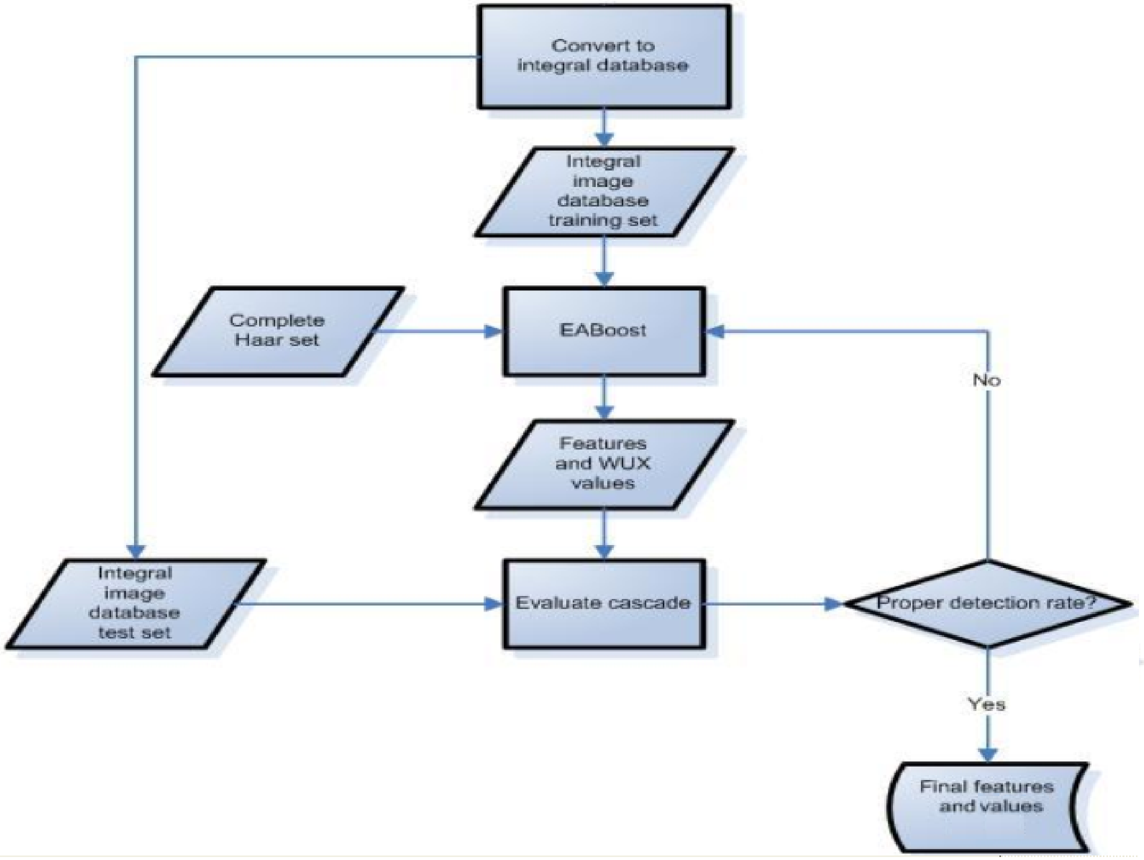


Evolutionary AdaBoost

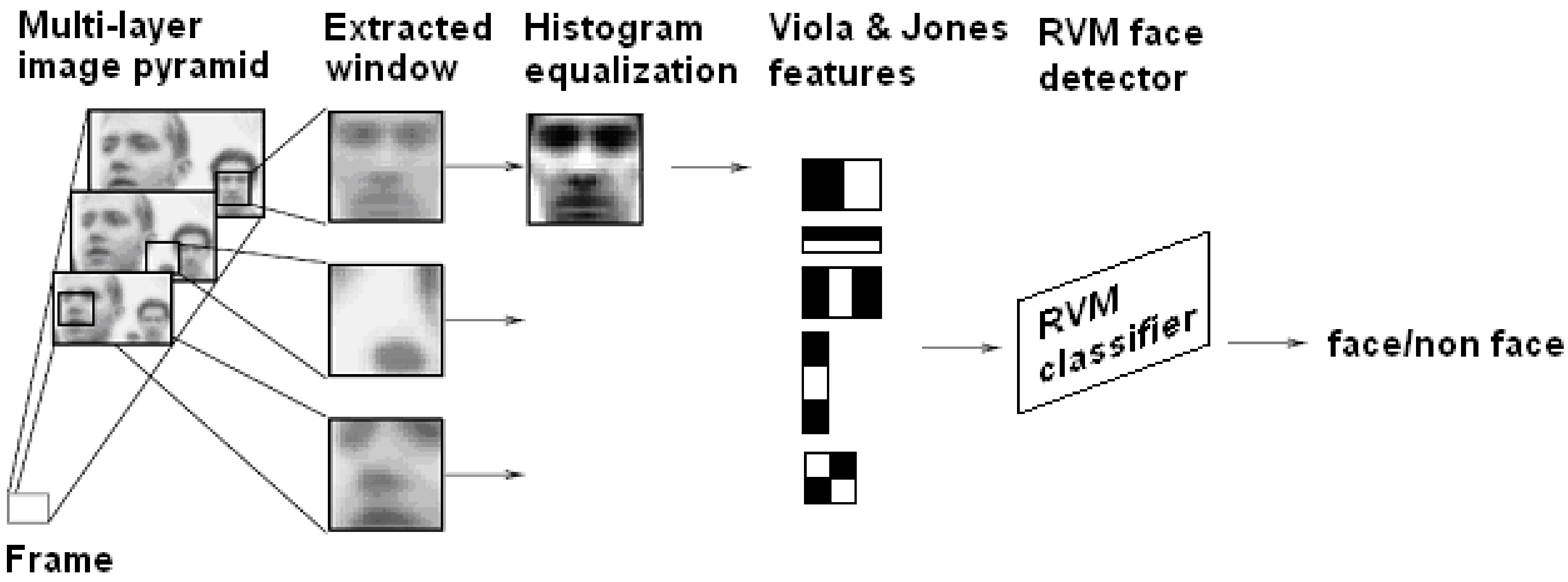
E.A. performs an efficient search for the representative Viola&Jones features for classification



Face detection

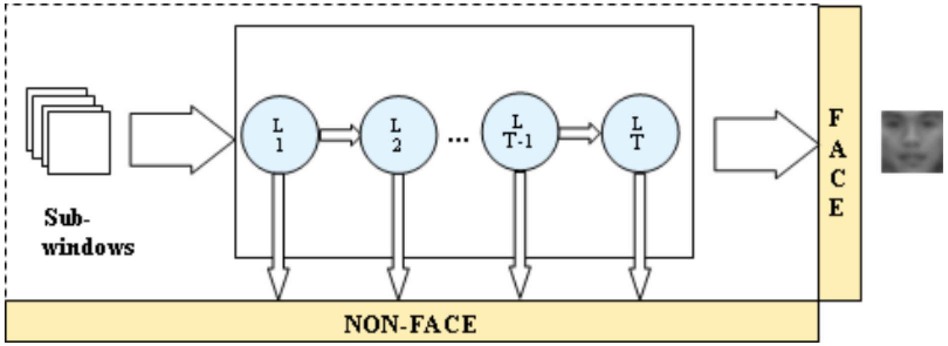


Face detection



Face detection

Cascaded classifier with T layers



Face detection

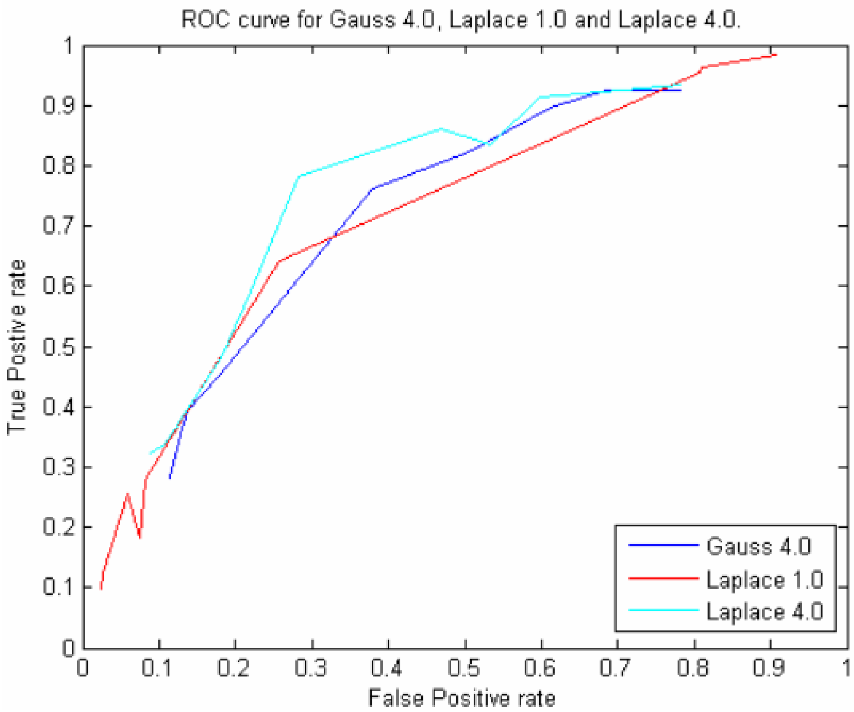
Example: choosing the propper weak classifier for three different V&J features.

2-fold cross validation results on three week classifiers for face detection based on Haar-like features

Kernel	Error rate		
	Feature 1	Feature 2	Feature 3
Gauss 2.0	26.30% ± 0.85	35.45% ± 7.71	38.60% ± 1.84
Gauss 5.0	25.35% ± 5.30	32.40% ± 2.83	36.55% ± 3.61
Laplace 0.5	35.20% ± 14.42	26.70% ± 0.99	42.70% ± 9.62
Laplace 2.0	29.25% ± 9.83	32.00% ± 7.50	41.60% ± 7.78
Laplace 5.0	26.20% ± 2.12	25.90% ± 1.84	37.45% ± 7.57

Face detection

ROC curves of three kernels, obtained by adjusting each classifier's threshold



Face detection results

RVM test results, both training and testing are performed on MIT CBCL database

Kernel	Nr. of test samples	Detection rate %	Nr. of false negatives	Nr. of false positives
Gauss 5.0	2500	93.92	103	49
Gauss 7.0		95.08	58	49
Laplace 2.0		83.88	339	64
Laplace 5.0		95.04	60	64

Face detection results

RVM test results, the training is done using MIT CBCL, the testing is done on CMU database

Kernel	CMU database consisting of faces only		CMU database consisting of non-faces only	
	Number of test samples	Detection rate %	Number of test samples	Detection rate %
Laplace 2.0	472	22.03	5036	100
Laplace 5.0		51.91		97.34
Gauss 5.0		38.77		96.68
Gauss 7.0		30.30		97.86

II. Facial feature extraction

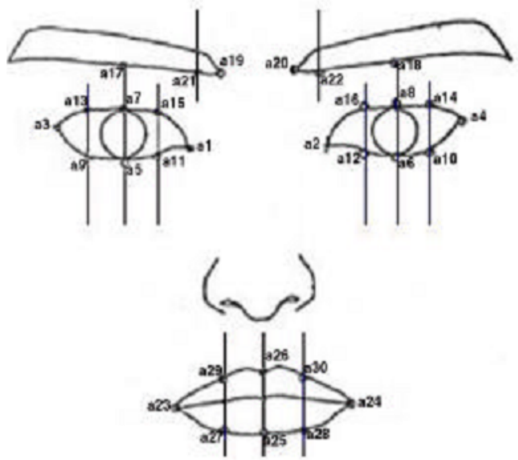
The facial features to be extracted are: left/right eye and mouth areas.

Dataset parameters			
Database source	BioID	BioID	BioID
Sample size (h x w)	15x25 pixels	15x25 pixels	20x40 pixels
Number of classes	2	2	2
Class 0	Non-left eye	Non-right eye	Non-mouth
Class 1	Left eye	Right eye	Mouth
Number of samples (0/1)	500/500	500/500	500/500

TPR	0.906	0.912	0.852
FPR	0.05	0.049	0.024

III. FCP model

Kobayashi and Hara model the face through 30 FCPs



There are three steps involved in the FCP detection:

1. FCP detection using *corner detectors*

2.a. FCP detection using *RVM classifier*

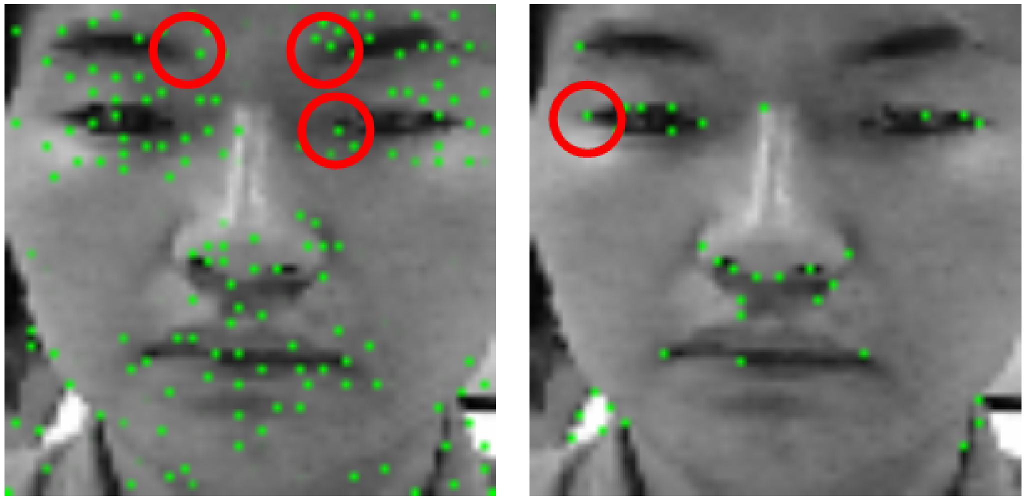
2.b. FCP detection using *integral projection method*

III.1. FCP detection using *corner detectors*

There are two corner detectors that are used as a first stage for FCP detection. The hybrid corner detector stands for a combination of two corner detectors:

- Harris
- Sojka

Test result of the corner detectors on face images (64x64 pixels)



Harris Sojka

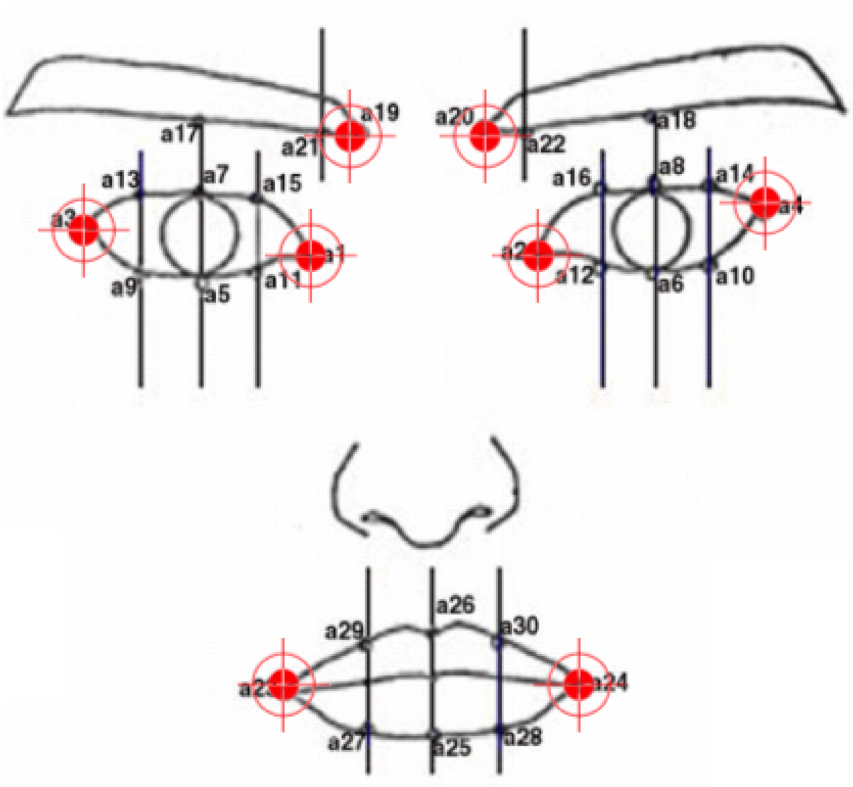
III.2.a. FCP detection using *RVM classifier*

The method makes use of:

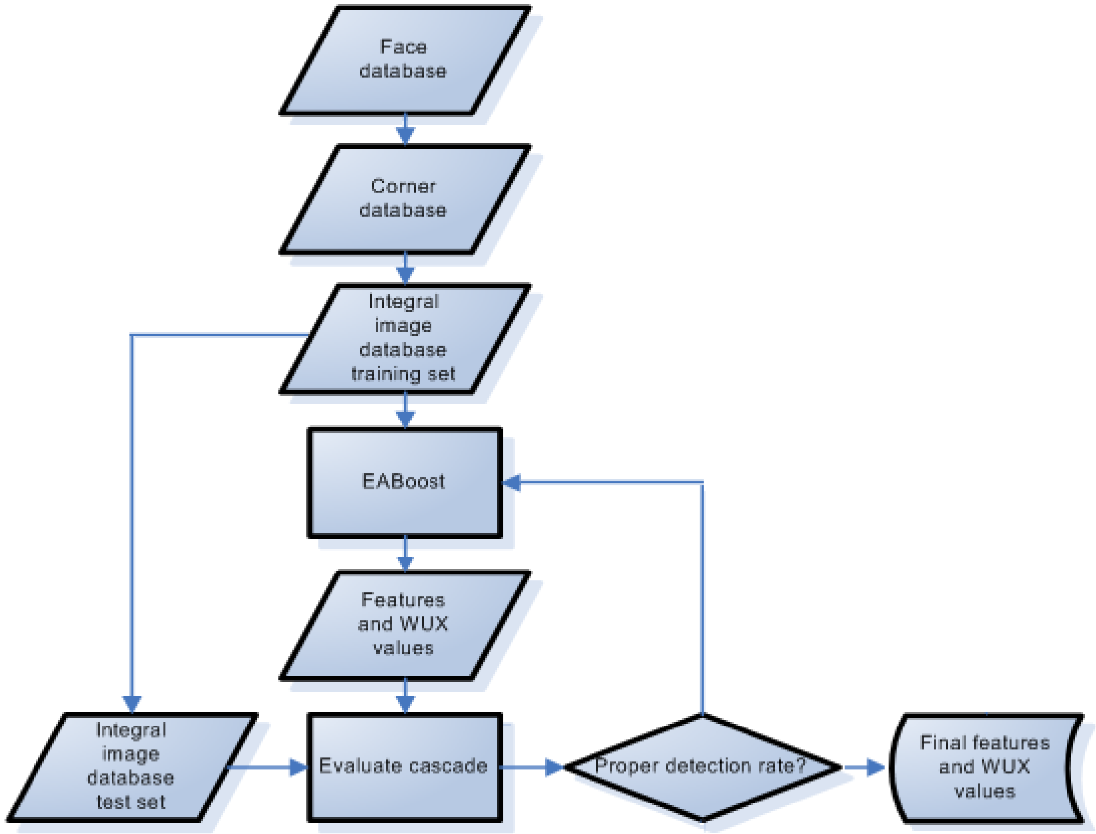
- Viola&Jones features (13x13 size samples, 14140 features/sample)
- Evolutionary AdaBoost
- Relevance Vector Machine (RVM) - weak classifier

Note:
DCT - Discrete Cosine Transform method has been found to be too sensitive to illumination

III.2.a. The FCPs to be extracted with RVM based E.A. classifier



III.2.a. The stages of training the FCP detector



III.2.a. FCP data set

BioID dataset and the Carnegie Melon dataset

Database description:		
Database source:	BioID/Carnegie Melon	
Database sample size:	64x64 pixels	
Extracted sample size:	13x13 pixels	
Training:	2-class (0/1)	
Type	**Nr. of samples (0/1)**	
	Training set	**Testing set**
Left eye inner corner	500/500	500/500
Right eye inner corner	500/500	500/500
Left eye outer corner	500/500	500/500
Right eye outer corner	500/500	500/500
Nose left corner	250/250	259/250
Nose right corner	250/250	326/350
Mouth left corner	500/500	500/495
Mouth right corner	500/500	500/495

III.2.a. EA characteristics

EABoost parameters	Values and description
Population size	250
Crossover/Mutation rate	0.20/0.80
Classifier/ kernel	RVM/ Laplace 4.0
Feature false positive rate	0.3
Feature true positive rate	0.75
Target false positive rate	< 0.05
Target true positive rate	> 0.90
Number of features needed	variable

III.2.a. FCP detection results

Type	Acronym	TPR	FPR
Left eye inner corner	LEIC	0.922	0.02
Right eye inner corner	REIC	0.89	0.106
Left eye outer corner	LEOC	0.916	0.064
Right eye outer corner	REOC	0.892	0.034
Left nostril corner	NLC	0.935	0.06
Right nostril corner	NRC	0.893	0.049
Mouth left corner	MLC	0.955	0.049
Mouth right corner	MRC	0.916	0.049

III.2.b. FCP detection, *integral projection method*

It is used to extract the rest of the points.

- projects the image into the vertical and horizontal axes
- obtain the boundaries
- the boundaries of the features have relatively high contrast
- the image is presented by two 1D orthogonal projection functions: IPF_v, IPF_h, $MIPF_v$, $MIPF_h$, $VIPF_v$, $VIPF_h$, GPF_v, GPF_h