

Gender Classification of Faces Using Adaboost*

Rodrigo Vershae^{1,2,3}, Javier Ruiz-del-Solar^{1,2}, and Mauricio Correa^{1,2}

¹ Department of Electrical Engineering, Universidad de Chile

² Center for Web Research, Department of Computer Science, Universidad de Chile

³ CMLA, ENS Cachan, France

Abstract. In this work it is described a framework for classifying face images using Adaboost and domain-partitioning based classifiers. The most interesting aspect of this framework is the capability of building classification systems with high accuracy in dynamical environments, which achieve, at the same time, high processing and training speed. We apply this framework to the specific problem of gender classification. We built several gender classification systems under the proposed framework using different features (LBP, wavelets, rectangular, etc.). These systems are analyzed and evaluated using standard face databases (FERET and BioID), and a new gender classification database of real-world images.

1 Introduction

The computational analysis of face images plays an important role in many computer vision applications. Among them we can mention visual human-computer interaction, biometry, video conferencing, human-robot interaction, surveillance, video summarizing, image/video indexing and retrieval, and drivers monitoring. So far, face detection systems that perform with high accuracy in real world applications have been developed [13][9][3][11]. However, face classification systems (gender classification, race classification, age classification, etc.) do not achieve similar performance, specifically when analyzing low-resolution faces obtained under uncontrolled conditions (uncontrolled illumination, non-uniform background, etc.).

We aim at reverting this situation by describing a framework for building robust face classification systems using Adaboost [7] and domain-partitioning [8] based classifiers. Within this framework we make use of diverse feature types: (i) rectangular features (similar to Haar wavelets) [12], (ii) LBP-based features [2], (iii) wavelet features [9], and (iv) LBP-based features applied over a wavelet decomposition. Some of these features have not been used before in face classification problems.

We apply this learning framework to the specific problem of gender classification. Gender classification is a relevant problem within many applications: (1) human-robot interaction and visual human-computer interaction: it allows building systems that behaves differently depending on the user's gender, (2) video summarizing, and image/video indexing and retrieval: it allows the use of gender information in the

* This research was funded by Millenium Nucleus Center for Web Research, Grant P04-067-F, Chile.

retrieving/indexing process, and (3) face recognition biometric systems: it allows improving the system performance by having specific models for each gender.

We built several robust gender classification systems using this learning framework and diverse features. Their main strengths are the ability of processing low-resolution faces (up to 24x24 pixels), and being illumination invariant (no pre-processing is needed for photometric normalization). These systems are analyzed and evaluated using standard face databases (FERET, BioID), and a new database of real-world images created with this purpose (UCHGender DB).

The article is structured as follows. In section 2 some related work is outlined. The learning framework is presented in section 3. In section 4 the employed features are described. In section 5 a comparative analysis of different gender classification systems is presented. Finally, in section 6, some conclusions of this work are given.

2 Related Work

Several methods have been proposed for solving the gender classification problem, among them systems based on neural networks (RBF, back propagation, etc.), PCA projections, decision trees, SVM classifiers, and Adaboost classifiers can be mentioned. Best reported results have been obtained using SVM and Adaboost. We will analyze some of these relevant works. In [5] it is proposed a gender classification system based on the use of SVM classifier. The employed features are the pixel elements themselves. The obtained results are very good, 3.38% overall error rate when using a RBF kernel, but the test set consists of only 259 faces. In [1] is proposed a gender classification system based on a SVM classifier and features obtained using PCA (Principal Component Analysis), CCA (Curvilinear Component Analysis) and SOM (Self Organizing Maps). Best results are obtained using 759 PCA components, 7.75% overall error rate, but the size of the faces is restricted to 128x128 pixels and the test set is composed by only 80 faces.

One of the drawbacks of the SVM based system is the fact that they are not real-time. Therefore systems based in Adaboost have been proposed in the last years. In [10] is presented a gender classification system that uses a threshold-weak-classifier based Adaboost algorithm and rectangular features. The system achieves a performance of 79% correct rate in a set of face images obtained from Internet that were manually annotated prior to classification (500 features were employed in this case). This system was favorably compared against the one proposed in [5] using the same dataset; it is 1,000 times faster and it has a higher classification rate (79% against 75.5%). In [14] is described a LUT-based Adaboost system for gender classification that uses rectangular features. Prior to classification faces are aligned. This is done using a face alignment method called SDAM that is a kind AAM (Active Appearance Model). After alignment, grey-level normalization (histogram equalization) is performed. The system achieves a classification rate of 88% in images downloaded from Internet (using 36x36 face windows), and using this database it is favorably compared against a SVM-based system and a threshold-Adaboost system.

Our Adaboost-based gender classification system employs the domain-partitioning approach, and the main improvements over previous works are: (1) the use of more suitable features for addressing this problem (mLBP features behave better than

rectangular features; see section 5), (2) the usage of smaller face windows (24x24) which allows analyzing smaller faces, and (3) a faster processing, because, besides the eye alignment, we do not perform any geometric or photometric normalization.

3 A Learning Framework for Building Robust Face Classification Systems

The key concepts used in the considered framework are boosting and domain partitioning classifiers. Adaboost [7] (a Boosting algorithm) is employed for finding highly accurate hypotheses (classification rules) by combining several weak hypotheses (classifiers). We use domain partitioning weak hypotheses [8], each one having a moderate accuracy, and giving self-rated confidence values that estimate the reliability of each prediction. These weak classifiers are linearly combined, obtaining a classifier of the form shown in (1). Each function $h_t(x)$ is a weak classifier, T is the number of weak classifiers, and b is a threshold value that defines the operation point of the classifier. The class assigned to the input corresponds to the sign of $H(x)$.

$$H(x) = \sum_{t=1}^T \alpha_t h_t(x) - b \quad (1)$$

The weak classifiers are applied over features computed in every pattern to be processed. Each weak classifier has associated a single feature. Following [8], domain-partitioning weak hypotheses make their predictions based on a partitioning of the domain X into disjoint blocks X_1, \dots, X_n , which cover all X , and for which $h(x) = h(x')$ for all $x, x' \in X_j$. Thus, the weak classifiers prediction depends only on which block X_j a given sample instance falls into. In our case the weak classifiers are applied over features, therefore each feature domain F is partitioned into disjoint blocks F_1, \dots, F_n , and a weak classifier h will have an output for each partition block of its associated feature f :

$$h(f(x)) = c_j \ni f(x) \in F_j \quad (2)$$

For each classifier, the value associated to each partition block (c_j), i.e. its output, is calculated for minimizing a bound of the training error. This value depends on the number of times that the corresponding feature, computed on the training samples (x_i), fall into this partition block (histograms), and on the class of these samples (y_i) and their weight $D(i)$. For minimizing the training error, c_j is set to [8]:

$$c_j = \frac{1}{2} \ln \left(\frac{W_{+1}^j + \varepsilon}{W_{-1}^j + \varepsilon} \right), \quad W_l^j = \sum_{i: f(x_i) \in F_j \wedge y_i = l} D(i), \text{ where } l = \pm 1 \quad (3)$$

were ε is a regularization parameter. The outputs, c_j , of each the weak classifier, obtained during training, are stored in a LUT for speeding up its evaluation. The pseudo code of the whole algorithm is shown in figure 1.

```

Given  $(x_1, y_1), \dots, (x_m, y_m)$  where  $x_i \in X, y_i \in Y = \{-1, +1\}$ 
 $D_1(i) \leftarrow \frac{1}{m}, i = 1, \dots, m$ 
for  $t = 1 : T$  do
    normalize  $D_t()$  to a p.d.f
    for each feature  $f_p \in F, h_p \in H, p = 1, \dots, P$  do
        
$$h_p(j) = \frac{1}{2} \ln \left( \frac{W_{+1}^j + \varepsilon}{W_{-1}^j + \varepsilon} \right), j = 1, \dots, J$$

        
$$Z_p = 2 \sum_j \sqrt{W_{+1}^j W_{-1}^j}$$

        select  $h_t$  for which  $Z$  is minimized:  $h_t = \underset{h_p \in H}{\operatorname{argmin}} Z_p$ 
        
$$D_{t+1}(i) \leftarrow D_t(i) \exp(-y_i h_t(x_i)), i = 1, \dots, m$$


$$H(x) = \sum_{t=1}^T h_t(x)$$


```

Fig. 1. Domain-Partitioning Adaboost training algorithm

4 Features

We analyze the use of different kinds of features applied to the problem of face classification: rectangular, mLBP, wavelet, and wavelet+mLBP features. In all cases the feature space is partitioned so it can be used directly with the domain-partitioning Adaboost classifier described in section 3. In all cases LUTs (Look-Up Tables) are used for a fast evaluation of the weak classifiers.

4.1 Rectangular Features

Rectangular features resemble Haar wavelets and can be evaluated very quickly, independently of their size and position, using the integral image [12]. They correspond to the difference between sums of pixels values in rectangular image regions. The output value defines a domain that is partitioned using intervals (or bins) of equal size [13].

4.2 Modified LBP

The LBP (Local Binary Pattern), also known as texture number or census transform, corresponds to an illumination invariant descriptor of the local structure in a given image neighborhood. We use their modified version [2] (mLPB = modified LBP), which overcomes some problems of the original LBP. The mLBP is computed as follows: for a given window of 3x3 pixels, the average of the pixels in the window is calculated. Then, each of the pixel values is compared against the obtained average. From these comparisons, 9 bits are generated, with 0 indicating that a pixel is smaller

than the average, and 1 otherwise. After that, the concatenated 9 bits corresponds to the mLBP feature. As it can be noticed, for the mLBP feature the domain partition is already defined, there are 512 bins (actually 511 [2]).

4.3 Wavelet Coefficients

The wavelet transformation allows analyzing images in the spatial-frequency domain. In the context of face detection, wavelets coefficients have been successfully employed as local features. As in [9], the features correspond to groups of 8 neighbor coefficients (in space, orientation or scale) that are used as a partition of the feature space. Each of the coefficients is quantized in 3 levels and then grouped together, defining a partition of 6,561 bins. Given that we use small training sets (see section 5), such a large number of bins may lead to overfitting. Therefore, we partition the wavelet coefficients directly in a small number of bins (16, 32 and 64).

4.4 Modified LBP of the Wavelet Coefficients

The mLBP is applied over the wavelet transform for avoiding the quantization of the wavelet coefficients, and for summarizing information from groups of coefficients. In this way we can reduce the number of partitions to 511 bins.

5 Evaluation

5.1 Evaluated Gender Classification Systems

We use three baseline systems for comparing the performance of our Adaboost based gender classification systems:

- SVM: SVM classifier and face pixels as features. Parameters: RBF kernel, $\sigma=1,125$, $C=25$, 3,082 support vectors.
- SVM+PCA: SVM classifier and PCA projection of the face pixels as features. Parameters: RBF kernel, $\sigma=4,950$, $C=50$, 1,981 support vectors, and 300 PCA components.
- PCA: Face gender is determined using the minimal reconstruction error after projection of the face pixels in a PCA model of men faces and a PCA model of women faces. Parameters: 100 PCA components in each model.

We build different gender classification systems using the described learning framework (domain partitioning Adaboost) and different features:

- Adaboost-Rect: Adaboost & 1,000 rectangular features, 16 bins.
- Adaboost-mLBP: Adaboost & 1,000 mLBP features.
- Adaboost-Wav: Adaboost & 426 wavelet-based features (2-level W), 64 bins.
- Adaboost-Wav-mLBP: Adaboost & 958 wavelet+mLBP features (3 levels WT).

For all systems the selection of the best parameters was done using a validation database (see next section).

5.2 Databases

For the classifiers' training (Adaboost, PCA and SVM based) we built a training database of 4,245 face images, containing images from the CAS-PEAL dataset [4] and UCHGenderTrainDB (our own dataset). We use 2,009 images from CAS-PEAL (1,222 men and 787 women), and 2,165 images from UCHGenderTrainDB (1,150 men and 1,115 women). We use a validation database of 2,745 face images, containing images from CAS-PEAL and UCHGenderValDB (our own dataset). The validation CAS-PEAL dataset contains 1,312 images (768 men and 544 are women), while UCHGenderValDB contains 1,433 images (744 men and 689 women). The validation database is used for model selection during training. In the case of domain-partitioning Adaboost, it is used for selecting the number of weak classifiers and the number of bins in the LUTs. In the case of SVM, it is used for selecting the parameters of the kernel, while in the case of PCA, for selecting the number of components. In both databases, training and validation, for each face a second one was generated, which corresponds to a random variation in the position of cropping. It is important to note that CAS-PEAL includes only Asians, while UCHGenderTrainDB considers other races. With this combination we intend to make our gender classifier race independent to a large degree.

For evaluating the proposed system we use 3 databases: (1) the UCHGender database (real world images, see fig. 2), (2) the Feret database [6], and (3) the BioID database [15]. See table 1 for details on the number of images in each class for these datasets. No image in the evaluation database is also included in the training or in the validation databases.

Faces were cropped using the same procedure during training and evaluation. The cropping was done using the position of the eyes. In the case of the training and validation datasets, the cropping of the faces was done using ground truth data, while in the case of evaluation two cases are considered: cropping using ground truth data and cropping using automatic face and eyes detection (both systems are described in [11]). The obtained results using both alternatives are analyzed (see figs. 3-5).

Table 1. Summary of Databases used for evaluation

Test database	# images	# Faces	# Men	# Women	% Men	% Women
UCHGender	142	343	192	151	55.9	44.1
Feret	2,745	2,745	1,650	1,095	60.1	39.9
BioID	1,521	1,521	975	546	64.1	35.9

The images employed for the training and testing of the SVM and Adaboost based systems were not preprocessed at all. In the case of the PCA based systems (PCA and PCA+SVM) the standard preprocessing required by PCA analysis was employed (subtraction of the mean face image plus variance normalization).

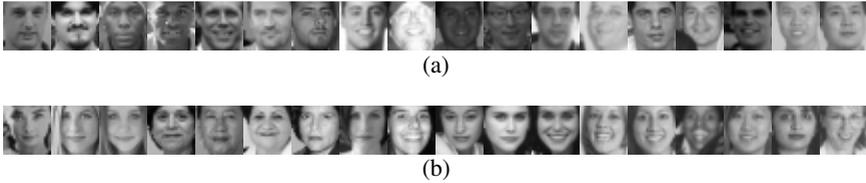


Fig. 2. UCHGender DB (examples) (24x24 pixels): (a) Men, (b) Women faces

5.3 Results

All the here presented results, with the exception of PCA and PCA+SVM, consider faces of 24x24 pixels. PCA based methods use faces of 100x185 pixels. The usage of larger face sizes (48x48) slightly improves the performance of some methods (for example Adaboost-Rect features and Adaboost-Wav), however the training time increases exponentially (from hours to days). Because of this and for not introducing important restrictions in the size of the faces to be analyzed we consider faces of 24x24 pixels.

Figures 3, 4 and 5 show the results of the evaluation of the different methods in the Feret, BioID and UCHGender databases, respectively. Figures 3(a), 4(a) and 5(a) show the results in the case when the eyes were annotated, while figures 3(b), 4(b) and 5(b) show results when the eyes (and faces) were automatically detected. In this last case only correct face detections were considered in the statistics (% of correct face detections: Feret: 99.49%, BioID: 98.22% and UCHGender: 96.79%). Table 2 shows some numerical results for the case of equal error rates in both classes.

Table 2. Correct Classification Rates at Operation points with equal error rates in both classes. Only best performing methods are shown. Faces were cropped using annotated eyes (left), and automatically detected eyes (right). Results are separated by a “/”. Best results are shown in bold.

Database	SVM (RBF)	Adab.-Rect	Adab.-mLBP	Adab.-Wav	Adab.-Wav-mLBP
UCHGender	79.2 / 79.82	79.47 / 79.22	81.23 / 80.12	79.47 / 75.90	78.01 / 71.99
Feret	83.4 / 84.13	85.24 / 83.95	85.56 / 85.89	78.31 / 76.51	78.09 / 83.83
BioID	81.44 / 79.05	79.66 / 79.52	80.91 / 81.46	78.54 / 72.96	72.48 / 68.94

It can be noticed that in the Feret database (fig. 3) the best performing method is Adaboost-mLBP followed by Adaboost-Rect, SVM and SVM+PCA. It can also be notice that in this database the results of all methods are relatively independent of the way the eye positions were obtained. Main reason seems to be the fact that due to the characteristics of this database (homogeneous backgrounds and controlled illumination) the face detection rate is very high (99.49%) and the eyes detection very precise.

In the case of the BioID database (fig. 4), the performance of some of the methods increases when the eye detector is used. This happens in particular with the Adaboost-mLBP and Adaboost-Rect. When using annotated eyes (fig. 4(a)), best performing methods are Adaboost-mLBP and SVM depending on the operation point. This is the

only case where SVM works better than the other methods, for equal error rates in both classes (see table 2). When using automatic detected eyes (fig. 4(b)), the best performing method is Adaboost-mLBP followed by Adaboost-Rect. This shows that Adaboost-mLBP is more robust (an independent) of the eye detector being used. In the case of the UCHGender database (fig. 5), it can be noticed that again the Adaboost-mLBP outperforms other methodologies. In the case of UCHGender database we do not include results of the PCA and PCA+SVM methods because most faces are smaller than 100x185, and PCA using those images gives poor results.

From figs 3, 4 and 5 it can be noticed that best results are obtained in the Feret database, followed by BioID, and UCHGender. This is probably because the Feret database contains homogeneous backgrounds, controlled illumination, and only completely frontal faces. On the other hand, the UCHGender contains a large variation on backgrounds, races, illumination conditions, and faces are not necessarily frontal -- some of them present yaw (out-of-plane) rotation.

In table 3 it is shown the average time required by the different methods for the gender classification of a given face image. This time does not include the time required for the face detection/cropping, face scaling and eyes detection. It includes just the time required for the face analysis (feature extraction and classification). It can be seen that Adaboost-mLBP is about 10 times faster than SVM, while Adaboost-Rect is 6 times faster than Adaboost-mLBP, and 60 times faster than SVM. This evaluation was done in an Intel Pentium 4 CPU 1.80GHz with 2GB RAM, running Debian GNU/Linux.

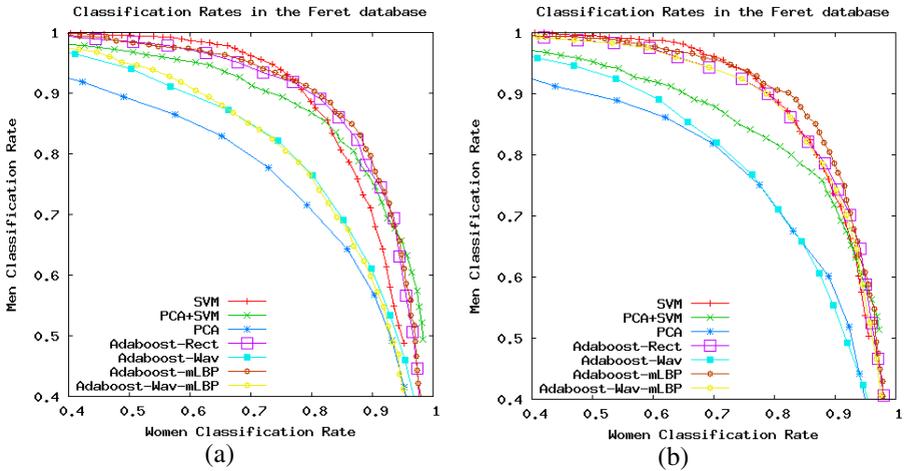


Fig. 3. Classification rates for the FERET dataset. Faces were aligned using: (a) annotated eyes, (b) detected eyes.

Table 3. Processing times of some of the different methods

Method	SVM	PCA	SVM+PCA	Ada-Rect	Ada-mLBP	Ada-Wav	Ada-Wav-mLBP
Time [mseg]	10.48	625	205	0.244	1.465	1.8554	2.7592

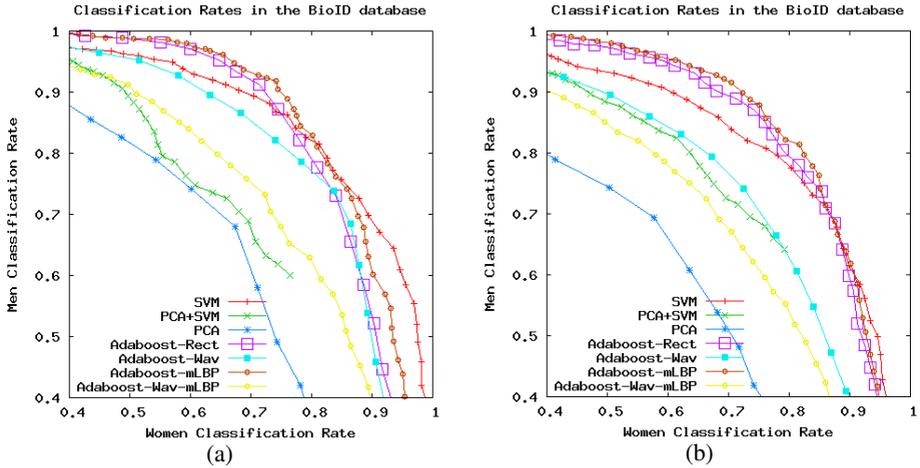


Fig. 4. Classification rates for the BioID dataset. Faces were aligned using: (a) annotated eyes, (b) detected eyes.

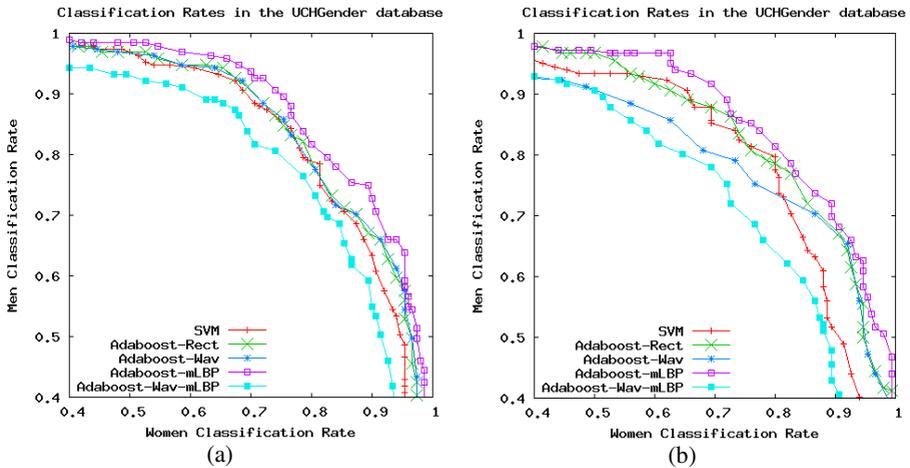


Fig. 5. Classification rates for the UCHGender dataset. Faces were aligned using: (a) annotated eyes, (b) detected eyes.

6 Conclusions

In this article it was presented a framework for classifying face images using Adaboost and domain-partitioning based classifiers. We built several gender classification systems using the proposed framework and different features (LBP, wavelets, rectangular, etc.). These systems are analyzed and evaluated using three databases (Feret, BioID and UCHGender). The obtained results indicate that Adaboost-mLBP

outperforms all other Adaboost-based methods, as well as baseline methods (SVM, PCA and PCA+SVM), in terms for classification rate. The Adaboost-mLBP behavior is robust to changes in the way eyes positions are obtained for performing the face alignment.

The most interesting advantage of the Adaboost-based methods is its high accuracy in dynamical environments, achieved with high processing speed. In terms of processing time, the faster method is Adaboost-Rect, being at least 6 times faster than the other methods (60 times faster than SVM-based methods). It is followed by Adaboost-mLBP, which is 10 times faster than SVM-based methods.

Another interesting characteristic of the developed Adaboost-based methods is their relatively high training speed, about one hour in the case of Adaboost-mLBP and about 48 hours for Adaboost-Rect, for a training database of 4,245 face images and a validation database of 2,745 face images.

Future work can be done in extending this framework to multi-class problems (age and race classification), and finding out a way of using (selecting) different kinds of features at the same time.

Acknowledgements

Portions of the research in this paper use the FERET database of facial images collected under the FERET program, and the CAS-PEAL face database collected under the sponsor of the Chinese National Hi-Tech Program and ISVISION Tech. Co. Ltd.

References

1. S. Buchala, N. Davey, R. J. Frank, T.M. Gale, M. Loomes, W. Kanargard, "Gender Classification of Face Images: The Role of Global and Feature-Based Information", *ICONIP 2004*, Calcutta, India, *Lecture Notes in Computer Science* 3316, pp. 763-768.
2. B. Fröba and A. Ernst, "Face detection with the modified census transform", *6th Int. Conf. on Face and Gesture Recognition - FG 2004*, pp. 91-96, Seoul, Korea, May 2004.
3. M. Delakis and C. Garcia, "Convolutional face finder: A neural architecture for fast and robust face detection", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 26, No. 11, pp. 1408-1423, 2004.
4. W. Gao, B. Cao, S. Shan, D. Zhou, X. Zhang, D. Zhao, "The CAS-PEAL Large-Scale Chinese Face Database and Evaluation Protocols", *Technical Report No. JDL_TR_04_FR_001*, Joint Research & Development Laboratory, CAS, 2004.
5. B. Moghaddam, M.-H. Yang, "Learning Gender with Support Faces", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 24, No. 5, pp. 707 - 711, 2002.
6. P. J. Phillips, H. Wechsler, J. Huang and P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms", *Image and Vision Computing J.*, Vol. 16, no. 5, pp. 295-306, 1998.
7. R.E. Schapire, "A brief introduction to boosting", *In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, 1999.
8. R.E. Schapire and Y. Singer, "Improved Boosting Algorithms using Confidence-rated Predictions", *Machine Learning*, 37(3):297-336, 1999.

9. H. Schneidermann and T. Kanade, "A statistical model for 3D object detection applied to faces and cars", *IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 746 – 751, 2000.
10. G. Shakhnarovich, P. Viola, and B. Moghaddam, "A Unified Learning Framework for Real Time Face Detection & Classification", *Int Conf. on Automatic Face & Gesture Recognition – FG 2002*, pp. 16 – 26, May 2002.
11. R. Verschae, J. Ruiz-del-Solar, M. Correa, and P. Vallejos, "A Unified Learning Framework for Face, Eyes and Gender Detection using Nested Cascades of Boosted Classifiers", Technical Report UCH-DIE-VISION-2006-02, Dept. of E. Eng., U. de Chile, 2006.
12. P. Viola and M. Jones, "Fast and robust classification using asymmetric adaboost and a detector cascade", *Advances in Neural Inform. Processing System 14*, MIT Press, 2002.
13. B. Wu, H. Ai, C. Huang, and S. Lao, "Fast rotation invariant multi-view face detection based on real Adaboost", *6th Int. Conf. on Face and Gesture Recognition - FG 2004*, pp. 79–84, Seoul, Korea, May 2004.
14. B. Wu, H. Ai, and C. HUANG. "LUT-based Adaboost for Gender Classification", *4th Int. Conf. on Audio and Video-based Biometric Person Authentication*, June 10-11, 2003, Guildford, United Kingdom.
15. BioID Face Database. Available on april 2006 in: <http://www.humanscan.de/support/downloads/facedb.php>