On the Complementarity of Face Parts for Gender Recognition

Yasmina Andreu Ramón A. Mollineda
University Jaume I, Castellón, Spain

Abstract

The expected complementarity between the most prominent parts of the face for the gender recognition task is evaluated.
Given the image of a face, five important parts (right and left eyes, nose, mouth and chin) are extracted and represented as appearance-based data vectors. Several mixtures of classifiers based on these five parts were designed using simple voting, weighted voting and other learner as combiners.
Experiments using the FERET database prove that ensembles perform significantly better than plain classifiers based on single parts.

Methodology

Given a grey image of a frontal pose of a face and the coordinates of the two eyes, seven subimages containing the two eyes, nose, mouth, chin, the rectangular area of the internal face and the full face are defined from expected proportions of an aesthetic face.
The process has the following steps:

1. Grid estimation 2. Histogram equalization 3. Select regions of interest.

Experiments

The FERET database

The experiments are based on a well-known database of face images called FERET, which contains human faces acquired in a controlled environment with no restrictions as to age, race, or facial expression. Only faces in frontal pose without glasses were used, because their presence could strongly distort the effectiveness of eyes for gender recognition. The experiments involved 2147 medium-sized 256x384 pixel images from 834 different subjects separated into 842 female faces and 1305 male faces.
From this image database, seven related vector datasets were inferred containing descriptions of the left and right eyes, nose, mouth, chin, internal face and the full face. These datasets were transformed by PCA, where new attributes were selected to explain 99% of the total variance.

The Classification Models

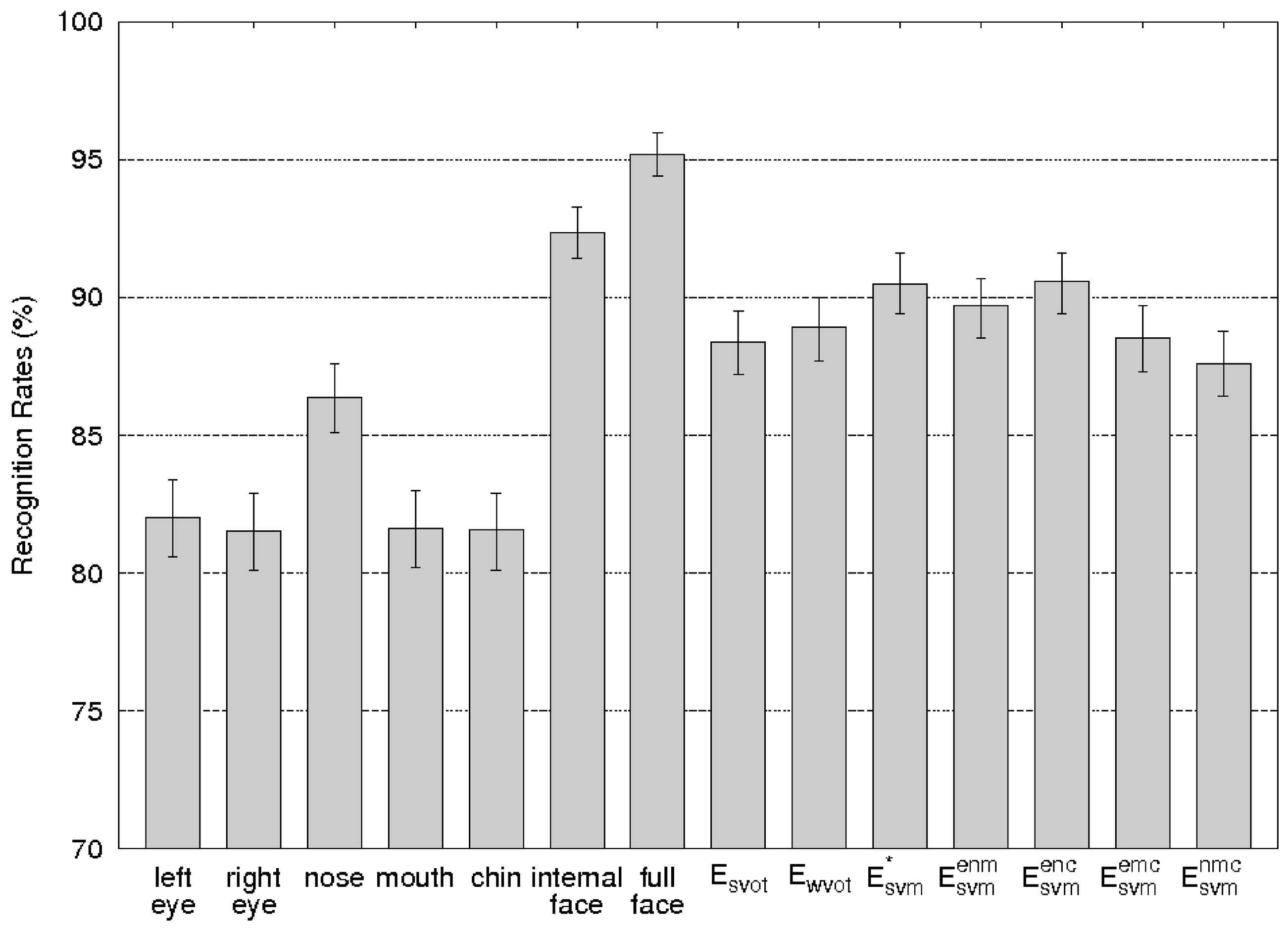
Several ensembles of parallel SVMs were defined, whose decisions were combined by simple and weighted voting and by other SVM, respectively. The SVM with a linear polynomial kernel was chosen as the base classifier of the ensembles because of its proved effectiveness in this task.
The ensembles of classifiers studied are:

- $\mathbf{E_{svot}}$ counts the number of base binary decisions for each class and chooses the most-voted one.
- $\mathbf{E_{wvot}}$ chooses the greater of the two sums of the *a posteriori* probabilities of both classes over all the base classifiers.
- $\mathbf{E^{*}_{svm}}$ combines the two *a posteriori* probabilities of all the base classifiers based in the five individual parts in a new vector space. New training subsets were created from the projection in this probability space of the original training partitions.
- $\mathbf{E^{enm}_{svm}}$ uses the same combination that $\mathbf{E^{*}_{svm}}$, but with only the left eye, the nose and the mouth probabilities.
- $\mathbf{E^{enc}_{svm}}$ combines only the probabilities of the left eye, the nose and the chin.
- $\mathbf{E^{emc}_{svm}}$ uses the probabilities of the left eye, the mouth and the chin.
- $\mathbf{E^{nmc}_{svm}}$ combines the probabilities of the nose, mouth and chin.

The plain gender classification of individual parts was also performed by the SVM with a linear polynomial kernel. Results were computed by averaging five independent runs of a 5-fold cross-validation.

Results

The performances (and their 95% confidence intervals) of the ensembles and the plain classifiers are shown in the following graph and tables.



Plain classifiers

left eye	right eye	nose	mouth	chin	internal face	full face
82.0	81.5	86.4	81.6	81.6	92.4	95.2
[80.6, 83.4]	[80.1, 82.9]	[85.1, 87.64]	[80.2, 83.0]	[80.1, 82.9]	[91.4, 93.3]	[94.4, 96.0]

Ensembles

E_{svot}	E_{wvot}	E^{*}_{svm}	E^{enm}_{svm}	E^{enc}_{svm}	E^{emc}_{svm}	E^{nmc}_{svm}
88.4	88.9	90.5	89.7	90.6	88.5	87.6
[87.2, 89.5]	[87.7, 90.0]	[89.4, 91.6]	[88.5, 90.7]	[89.4, 91.6]	[87.3, 89.7]	[86.4, 88.8]

Conclusions

- The experiments proved that the joint contribution of separate parts is more effective for gender recognition than isolated parts. So, there is a complementarity relation between the face parts.
- The ensembles based on 3 parts appear to be as discriminant as the combinations of 5 parts. Particularly, those in which the eye and the nose coincide perform better. This would be useful to recognize the gender under partial occlusions of the face.
- The plain classifiers based on the holistic descriptions were more accurate than the ensembles. There are two major causes that explain these results: the holistic representation provides information about the configuration of the face, and the full faces contain more face parts than these studied.