**QUEEN'S
UNIVERSITY
BELFAST**

AdaBoost Multiple Feature Selection and Combination for Face Recognition

Martinez-Contreras, F., Orrite-Urunuela, C., & Martinez-del-Rincon, J. (2009). AdaBoost Multiple Feature Selection and Combination for Face Recognition. In H. Araujo, AM. Mendonca, AJ. Pinho, & MI. Torres (Eds.), *PATTERN RECOGNITION AND IMAGE ANALYSIS, PROCEEDINGS* (Vol. 5524 LNCS, pp. 338-345). Springer.

Published in:

PATTERN RECOGNITION AND IMAGE ANALYSIS, PROCEEDINGS

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

AdaBoost Multiple Feature Selection and Combination for Face Recognition

Francisco Martínez-Contreras, Carlos Orrite-Uruñuela, and Jesús Martínez-del-Rincón

CVLab, Aragon Institute for Engineering Research, University of Zaragoza, SPAIN
(franjmar,corrite,jesmar)@unizar.es

Abstract. Gabor features have been recognized as one of the most successful face representations. Encouraged by the results given by this approach, other kind of facial representations based on Steerable Gaussian first order kernels and Harris corner detector are proposed in this paper. In order to reduce the high dimensional feature space, PCA and LDA techniques are employed. Once the features have been extracted, AdaBoost learning algorithm is used to select and combine the most representative features. The experimental results on XM2VTS database show an encouraging recognition rate, showing an important improvement with respect to face descriptors only based on Gabor filters.

1 Introduction

Face recognition has become in a relevant biometrics modality due to the broad variety of potential applications in public security, law enforcement, access control and video surveillance. The main two factors that affect to the performance of most current systems are the large variability in facial appearance of a person, due to different poses, lighting and facial expressions, and the high dimensionality of the facial features.

In order to deal with facial expression and illumination, some authors [8, 9] propose to use a Gabor wavelet representation. The Gabor wavelet kernels allow to capture salient visual properties such as spatial localization, orientation selectivity, and spacial frequency characteristics.

Recently, others authors [21, 22] proposed the use of local oriented Gaussian derivative filters for face identification. The structure of an image can be represented by the outputs of a set of multi-scales Gaussian derivative filters (MGDF) applied to the image as proposed in [15], for the appearance representation in face recognition and image retrieval. More biologically motivated is the use of Steerable Gaussian first order derivative filters as proposed in [4] for detecting facial landmarks from neutral and expressive facial images.

The main drawback when using Gabor or other oriented filter approaches to obtain a new representation of the face is the great redundancy and high dimensionality they exhibit. Recent papers [19, 17] have proposed using the AdaBoost learning algorithm presented in [18] to select a small set of filters.

To cope with this dimensionality problem, some subspace projection techniques to reduce feature dimension are proposed. In addition to the “classical” techniques such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) or a combination of both (PCA+LDA)(see chapter 7 in [7] for a survey of face recognition in subspaces), recently two discriminant approaches have been proposed in junction with Gabor filters. In [9], the authors apply the Enhanced Fisher linear discriminant Model (EFM). In [16], Shen, et al. propose to transform the Gabor feature space into a General Discriminant Analysis (GDA) space. The mean nearest neighbor rule is used for classification in both methods, using different distance measures(i.e. Euclidean Distance, Mahalanobis Distance or Cosine Similarity).

2 Proposal Overview

Considering the good results given by the Gabor wavelet filters, as well as the multi-scales Gaussian derivative filters, this paper tries to fuse different face representations in order to find the best one among all.

In addition to the previous filters used in face recognition, this paper explores other feature extraction algorithms used in other computer vision applications. For example, Harris corner detector has become a popular interest-point detector due to its strong invariance to rotation, scale, illumination variation and image noise, just the kind of problems found in face recognition.

All these transformations constitute a rich representation of the face, invariant to lighting and facial appearance. However, the final representation of the face may exhibit a great redundancy and high dimensionality.

Let $\mathbf{F}$ be an arbitrary filtered image which is projected onto a PCA basis. This image (minus the mean) in the training set can be represented as a linear combination of the best K eigenvectors.

$$\mathbf{F} - \mathbf{mean} = \sum_{k=1}^K w_k a_k, (w_k = a_k^T \mathbf{F}) \quad (1)$$

where a_k are called Eigenfilters and w_k are the eigenvalues that can be used as features for the next fusion architecture.

Boosting is considered a powerful technique to combine multiple classifiers producing in this way a form of committee whose performance could be significantly better than that of any of the base classifiers [2]. Here we use a form of boosting called AdaBoost as proposed in [18].

There are several levels at which fusion can take place in a multi-biometric system [6]. In this paper, two different strategies based on AdaBoost algorithm are proposed: feature level fusion and score level fusion. Both approaches use the patterns obtained by applying PCA and LDA to the filtered facial representation after convolving the input image with a kernel filter.

Figure 1 (Left) shows the feature level fusion architecture. In this approach, the AdaBoost algorithm receives as input the eigenvalues w_k set corresponding

Fig. 1. Feature level fusion architecture (Left) and Score level fusion architecture (Right)

to an Eigenfilter (Fisherfilter) image, together with a group of thresholds. The similarity of two features is defined as a simple subtraction.

In the score level fusion architecture, see Figure 1 (Right), each filter formulates a weak classifier. The decision of acceptance or rejection is based on a measurement of similarity between the filtered face and the template set, with a global threshold. The similarity of two images is defined as the cosine distance. The threshold is selected at the equal error point, EER, at where the False Acceptance Ratio (FAR) is equal to the False Rejection Ratio (FRR) on the evaluation set.

The main contributions of this paper are summarized as follows:

- Different facial representations are considered to cope with different facial expressions and lighting.
- Two combination strategies, based on AdaBoost, are proposed in order to select a small set of filters, reducing redundancy and improving at the same time the rates of recognition.

3 Kernel Filters

3.1 Gabor Features for Face Recognition

In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The equations can be revised in previous experiments [9, 19, 17, 16], in which, the following parameters are used: five scales v and eight orientations u with $\sigma = 2\pi$, $k_{max} = \pi/2$ and $f = \sqrt{2}$. As pointed by [9] the convolution is speeded up by applying the Fast Fourier Transform (FFT). As a result, image $\mathbf{I}(\mathbf{z})$ can be represented by a set of Gabor complex coefficients. Like previous works, only magnitudes are used to form the final face representation.

3.2 Differential features

First order differential features The structure of an image can be represented by the outputs of a set of multi-scale Gaussian derivative filters. The simplest differential feature is a vector of spatial derivatives [20]. The initial image is incrementally convolved with Gaussians to produce images separated by a constant factor k in scale space. Each octave of scale space (doubling of σ) is divided into an integer number, s , of intervals, so $k = 2^{1/s}$. In the present approach, the values $\sigma = 1.6$ and $s = 4$ has been used due to empirical experiments. With these values, the number of used scales is seven. The magnitude and the orientation of all the pixels are calculated for all scales.

Oriented Filters Steerable filters refer to a class of filters in which a filter of an arbitrary orientation is synthesized as a linear combination of a set of "basis filters", see [3].

Following a similar approach as used for Gabor-based feature representation, a bank of filters is adopted, varying the orientations and the frequencies/scales. In this way, oriented edges are extracted by convolving a smoothed image with a set of different kernels. Each kernel is sensitive to a different orientation θ . This paper applies the following values for $\theta = k * 22.5$, $k = 0, \dots, 7$, the same ones used for Gabor filters, and for $p, q = \{-64, \dots, 64\}$, in accordance with the size of the image.

Two different Steerable filters are tested, one of them is developed in [4]. The value of σ in this case is $2\pi * i$, with $i = \{0.5, 1, 1.5, \dots, 4\}$. The second kind of Steerable filters are made likeness Gabor filters, with k_v used as scale (eq. 2). In this case, $\sigma=2\pi$, and $k_v = \frac{k_{max}}{f_v}$ is the same that in Gabor filters.

$$G_{\theta}^{-} = \frac{k_v^2}{2\pi\sigma^2} \cdot e^{-k_v^2 \frac{(p-\sigma \cos \theta)^2 + (q-\sigma \sin \theta)^2}{2\sigma^2}} \quad G_{\theta}^{+} = \frac{k_v^2}{2\pi\sigma^2} \cdot e^{-k_v^2 \frac{(p+\sigma \cos \theta)^2 + (q+\sigma \sin \theta)^2}{2\sigma^2}} \quad (2)$$

3.3 Harris Corner Detector

The use of local interest points in image matching can be traced back to the work of Moravec [14] on stereo matching using a corner detector. Later, Harris and Stephens [5] improved the Moravec detector making it more repeatable under small image variations and near edges. This work follows an approach similar to that given by Mikolajczyk [12]. The space of scale is a set of images, which are generated convolving the original image by a series of Gaussian kernels with different values of sigma. The sigma increase must be adequately fixed, since a high step would decrease the reliability of finding the most suitable correspondence. The value of an initial sigma is 1.6 and the sigma increase is $k = 2^{1/8}$, and the number of scales used is 5. To work with color images, some authors [13] propose a simple modification of Harris Detector, computing gradients in the RGB space. The parameters obtained by the Harris detector are the magnitude and the orientation and the two Eigenvalues λ_1 and λ_2 .

4 Feature Selection and Combination by AdaBoost

4.1 Fusion architectures

To use AdaBoost in the context of face recognition, the concepts of Intra-personal difference (called positive examples) and Inter-personal (called negative examples) are adopted in order to convert the multi-class problem into a two-class problem [17]. To accomplish this process, the original database is split into three ones: training, evaluation and test data set. The training data set is applied to obtain the new PCA (LDA) transformation parameters.

Figure 2 (Left) shows the flowchart for the feature level fusion architecture. Given an input image, different filters are applied to obtain a FilterFace representation at different scales and orientations. All the images of the evaluation data set are projected to the new PCA/LDA space. These new features w_k ((1)) are then used to generate positive and negative examples in relation to the individuals in the training data set. In this combined strategy, a simple subtraction between patterns is done, and the resulting features are introduced into the AdaBoost algorithm. For each feature, the weak learner determines the optimal threshold classification function, such that the minimum numbers of examples are misclassified.

Fig. 2. Flowchart for the face recognition system according to the feature level of the fusion architecture (Left) and according to the score level of the fusion architecture (Right)

The score level fusion architecture is depicted in Figure 2 (Right). The initial stages are the same than the feature level fusion, but once the K eigenvalues w_k are obtained, a classifier is built based on a measurement of similarity, the nearest neighbor decision rule, between the input and the average of client's training images with a global threshold. The threshold is selected at the equal error point, EER on the evaluation set. The similarity of two features $\mathbf{x}$, $\mathbf{y}$ is given by the cosine distance of its projected vector in PCA (LDA) subspace.

4.2 AdaBoost algorithm

AdaBoost is an algorithm which obtains a strong classifier from weak classifiers, combining each one of these individual classifiers with different weights to obtain a consensus decision which minimizes the error rate. The algorithm of AdaBoost can be seen at [18].

The weak classifier $h_j(x)$ is composed of one measurement $f_j(x)$ (Cosine Distance or subtraction, depending on the fusion architecture), and a threshold Θ_j . If the classifier returns 1, it means that the two patterns belong to the same person.

5 Experiments and Discussion

The proposed methods have been tested on the public XM2VTS face database [1], which contains 2360 images of 295 subjects, with 8 faces for each subject, divided into 4 sessions of 2 images in each one. The testing is performed following the Laussane protocol [11], which splits the database into 3 different sets for training, evaluation and test. Previous works [10] have reported that the configuration II gives lower error rates than configuration I. 200 “clients” and 95 “impostors” have been used.

All images are cropped and rectified according to the manually located eye positions supplied with the XM2VTS data. The normalized images are 128 x 128 pixels. By filtering and normalizing the facial image, a feature vector in a very high dimensionality space is obtained. The use of PCA allows to reduce each vector to N-1 dimensions keeping at least 90% of the variance in most of the cases, where N represents the number of classes. Although the reduction is enough, LDA is applied, which gets a better separation among feature vectors of different classes. The dimensionality is maintained after LDA. After that, the size of each pattern, which was initially of $128 \times 128 = 16384$ features, is composed of 199 features (number of clients - 1).

The two fusion architectures explained before are tested using Gabor filters. To measure the performance of the verification system, FAR, FRR and the Total Error Rate (TER), defined as the sum of the FAR and FRR are used.

Table 1 (Left) shows the rates given by both approaches. Clearly, the score level fusion outperforms the feature level fusion in FAR as well as FRR. In addition, the computational cost of the feature level architecture is quite higher than the score level. In this point we took the decision of using the score level fusion architecture for the rest of the facial representation.

Table 1 (Right) shows the best results obtained for every facial representation, together with the number of weak classifiers given by the AdaBoost algorithm.

The next step is to combine these filters in order to obtain the best error rates. In total, all combinations of 1-filter, 2-filters, 3-filters, and so on..., have been tested, $C_1^5 + C_2^5 + C_3^5 + C_4^5 + C_5^5 = 31$, where C_i^5 denotes the number of combinations of i filters chosen amongst the possible total number of 5 filters.

The results are shown in Table 2. The best combination in order to obtain a lower TER is given by Harris, Gaussian filters and the first Steerable filter proposal, with FAR = 1.47, FRR = 0.25 and TER = 1.72.

Architecture	FAR	FRR	TER
Score level	1.57	1.25	2.82
Feature level	2.05	1.75	3.80

Label	Filter	In	Out	FAR	FRR	TER
1	Gabor	56	14	1.57	1.25	2.82
2	Harris	20	6	2.28	1.5	3.78
3	Gaussian	14	6	2.65	1.5	4.15
4	Steerable 1	64	12	2.02	1	3.53
5	Steerable 2	64	15	1.98	1.5	3.48

Table 1. (Left) FAR and FRR for both fusion architectures using Gabor filters. (Right) Best rates of FAR and FRR obtained with the different filters

Combination	FAR(%)	FRR(%)	TER
1+2	2.03	1.25	3.28
1+3	0.93	1.25	2.18
1+4	1.50	0.75	2.25
1+5	1.53	1.5	3.03
2+3	1.15	1.25	2.40
2+4	1.69	0.5	2.19
2+5	1.65	1.75	3.40
3+4	1.78	0.5	2.28
3+5	1.64	1.25	2.89
4+5	1.98	1.5	3.48
1+2+3	0.81	1.25	2.06
1+2+4	1.49	0.5	1.99
1+2+5	1.51	1.5	3.01

Combination	FAR(%)	FRR(%)	TER
1+3+4	1.38	0.5	1.88
1+3+5	1.37	1	2.37
1+4+5	1.60	1.5	3.10
2+3+4	1.47	0.25	1.72
2+3+5	1.57	1.25	2.82
2+4+5	1.67	1.5	3.17
3+4+5	1.60	1.25	2.85
1+2+3+4	1.31	0.5	1.81
1+2+3+5	1.30	1	2.30
1+2+4+5	1.56	1.5	3.06
1+3+4+5	1.60	1.5	3.10
2+3+4+5	1.56	1.25	2.81
1+2+3+4+5	1.26	0.75	2.01

Table 2. Best rates of FAR and FRR obtained by means of the combination of different filters

6 Conclusions

In this paper, we have presented a novel technique of dimensionality reduction of features for face recognition using the AdaBoost algorithm as a selector of facial representations.

We have proposed two different fusion architecture based on score level or feature level fusion. The application of both approaches to the selection of Gabor filters showed that the score level fusion outperforms the feature level one with a lower computational cost.

Besides Gabor features, other kind of facial representations based on Steerable Gaussian first order kernels and Harris corner detector have been tested.

The experimental results on XM2VTS database show that individually, the Gabor filters are the best method for face authentication. The AdaBoost algorithm chooses 14 Gabor representation to give a TER=2.82.

The combination of different filters improves substantially the recognition rate, reaching a TER=1.72 using 16 filters. Surprisingly, Gabor filters are not included in the best combination, unlike other filters with less computational cost.

References

1. The xm2vtsdb. <http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb>.
2. C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
3. W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(9):891–906, 1991.
4. Gizatdinova and Surakka. Feature-based detection of facial landmarks from neutral and expressive facial images. *Pattern Anal. Mach. Intell.*, 28(1):135–139, 2006.
5. C. Harris and M. Stephens. A combined corner and edge detector. In *4th ALVEY Vision Conference*, pages 147–151, 1988.
6. ISO/IEC. Technical Report 24722, 2007.
7. A. K. Jain and S. Z. Li. *Handbook of Face Recognition*. 2005.
8. V. Krüger and G. Sommer. Gabor wavelet networks for object representation. In *Proceedings of the 10th International Workshop on Theoretical Foundations of Computer Vision*, pages 115–128, London, UK, 2001. Springer-Verlag.
9. C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing*, 11(4):467–476, 2002.
10. K. Messer, J. Kittler, J. Short, G. Heusch, F. Cardinaux, S. Marcel, Y. Rodriguez, S. Shan, Y. Su, W. Gao, and X. Chen. Performance characterisation of face recognition algorithms and their sensitivity to severe illumination changes. In *Advances in Biometrics, International Conference, ICB 2006*, pages 1–11, 2006.
11. K. Messer, J. Matas, J. Kittler, J. Luetttin, and G. Maitre. Xm2vtsdb: The extended m2vts database. In *Second International Conference on Audio and Video-based Biometric Person Authentication*, March 1999.
12. K. Mikolajczyk and C. Schmid. Indexing Based on Scale Invariant Interest Points. In *8th International Conference on Computer Vision (ICCV 2001)*, pages 525–531.
13. P. Montesinos, V. Gouet, and R. Deriche. Differential invariants for color images. In *ICPR '98: International Conference on Pattern Recognition-Vol 1*, page 838, 1998.
14. H. Moravec. Rover visual obstacle avoidance. In *International Joint Conference on Artificial Intelligence (IJCAI81)*, pages 785–790, 1981.
15. S. Ravela and A. Hanson. On multiscale differential features for face recognition. *Vision Interface, Ottawa*, 2001.
16. L. Shen, L. Bai, and M. Fairhurst. Gabor wavelets and general discriminant analysis for face identification and verification. *Image Vision Comput.*, 25(5):553–563, 2007.
17. Y. Su, S. Shan, W. Gao, and X. Chen. A multiple fisher classifiers combination for face recognition based on grouping adaboosted gabor features. In *Proceeding of British Machine Vision Conference, Oxford, UK*, volume 2, pages 569–578, 2005.
18. P. A. Viola and M. J. Jones. Rapid object detection using a boosted cascade of simple features. In *CVPR 2001*, pages 511–518, 2001.
19. P. Yang, S. Shan, W. Gao, S. Z. Li, and D. Zhang. Face recognition using ada-boosted gabor features. In *Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2004)*, pages 356–361, 2004.

20. J. J. Yokono and T. Poggio. Oriented filters for object recognition: an empirical study. In *Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR 2004)*, pages 755–760, 2004.
21. J. J. Yokono and T. Poggio. A multiview face identification model with no geometric constraints. In *FGR '06: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, pages 493–498, 2006.
22. X. Zhang and Y. Jia. Face recognition with local steerable phase feature. *Pattern Recogn. Lett.*, 27(16):1927–1933, 2006.