Proceedings of the 23rd IASTED International Multi-Conference
ARTIFICIAL INTELLIGENCE AND APPLICATIONS
February 14-16, 2005, Innsbruck, Austria

OPTIMIZATION OF A FACE VERIFICATION SYSTEM USING BAYESIAN SCREENING TECHNIQUES

Raquel Montes Diez, Cristina Conde, Ángel Serrano, Licesio J. Rodríguez-Aragón, Enrique Cabello
Universidad Rey Juan Carlos (URJC)
Escuela Superior de Ciencias Experimentales y Tecnología (ESCET)
Face Recognition & Artificial Vision Group (FRAV)
C/ Tulipán, s/n. Móstoles E-29833, Madrid, Spain
{r.montes, cristina.conde, aserranos, ljrodriguez, ecabello}@escet.urjc.es
http://frav.escet.urjc.es/

ABSTRACT

We present a face verification system. A 100 people face database has been created using a CCD video camera, with controlled illumination conditions and frontal upright face position. A Principal Component Analysis matrix has been computed with eight images per person, and only the 150 most important eigenvalues have been used. The results of PCA are fed into two classifiers (SVM and RBF), in order to perform a verification process in a control access.

The algorithm proposed here allows to compute automatically the optimal acceptance threshold to divide a population of candidates into genuine or registered and impostors or non-registered. A Bayesian approach based on screening techniques has been considered, so that the user provides the economical cost for false acceptances and false rejections within the system. According to the ratio between these two costs, the optimal acceptance threshold is computed as the value that minimizes the expected total cost for both acceptances and rejections.

Our experimental results show that our SVM classifier produces a lower false acceptance rate (FAR) for a given false rejection rate (FRR), and vice versa, than our RBF classifier. The FAR also appears to cancel for SVM for high security conditions.

KEY WORDS
Computer vision, pattern recognition, face verification.

1. Introduction

Biometrics-based applications have evolved largely in a few years, and have passed quickly from research labs to the first commercial implementations. There is still much expectation to this new technology, in special for face recognition, as it can be very powerful in cases such as airports, customs, official buildings, and so on, where security is important.

However, nowadays face recognition is not a panacea. It is still difficult to achieve high performances in real-time commercial applications without controlled conditions, as the false acceptance rates (FAR) and the false rejection rates (FRR) are still high enough to become a concern, letting impostors in or making the entrance of a considerable amount of registered users quite annoying due to their continuous rejection. In control access systems, the effect of the environment in the data acquisition is more controlled than in other applications, so better and more consistent results are obtained. Such systems could fulfill the performance criteria demanded by potential clients.

The experiment presented in this paper tests the performance of a control access system based on face verification technology. In such environments, it is possible to take advantage of a set of specific characteristics. Usually, the subject is in front of the camera, only one subject appears, the size of the face is more or less constant and the subject is usually collaborative. Our system uses these advantages and proposes a control access system designed to work in such situations.

In recent years, two main approaches to face recognition have appeared. On the one hand, Principal Components Analysis (PCA) and related methods such as Fisherfaces [1] [2] [3] [4] consider only the global information of the face. Likewise, methods based on Local Feature Analysis (LFA) [5] [6], similar to PCA, consider different kernel functions concentrated on local features (eyes, mouth and nose). In this case, selection of facial features and kernels is an open issue. On the other hand, Elastic Bunch Graph Matching (EBGM) [7] uses a wavelet transformation to obtain a local description of the face and a graph to obtain a global face description. As well, following the success of FERET tests [8] [9] [10], a recent and extensive survey of ten commercial products has been performed (FRVT 2002) [11].

Bearing in mind that the acceptance threshold is the value that determines whether a verification is acceptance or rejection, the decision of the optimum value that minimizes both FAR and FRR still remains an open problem (see Figure 1). These rates are tolerable depending on the level of security required for each situation.

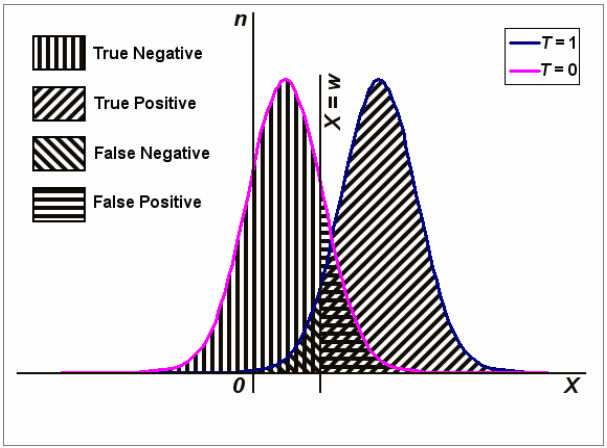

Figure 1 – Example of distribution of genuine cases ($T = 1$) and impostors ($T = 0$), as a function of the screening variable X, which corresponds to the ouput of a classifier such as SVM or RBF. The value $X = w$ is the threshold that best separates genuine cases from impostors. For clarity, both distributions have been drawn with the same height and standard deviation.

In order to take this into account, we propose a classification system based on costs for false acceptances and false rejections. The exact calculation of both costs (acceptance and rejection) could be difficult to find, but the rate between these costs is easier to fix. This is the input in the algorithm proposed. In this paper we present an innovative algorithm to calculate this optimal acceptance threshold by using economic screening techniques based on different costs for different error type.

2. Experimental Setup

As can be seen in Figure 2, the image acquisition set up consists on two diffuse light sources placed on both sides of a video camera.

In order to minimize distortions originated by changes in the lens focal length and the camera-subject distance, both have been fixed in every operation environment. These requirements are easily met in any exploitation site.

We used FRAV2D Database [12], obtained with 100 individuals in front of a dark uncluttered background and with controlled illumination conditions. Frontal and half-profile poses were considered, with more images for frontal views. Subjects were forced to change their pose between the acquisition of two consecutive images.

The image size is 320×240 pixels with face covering great part of the image (some examples are shown in Figure 3). Our face location system cropped the face to a window of 130×140 pixels. Eight frontal images per subject were used for computing a PCA matrix and training all classifiers (see below). For tests sets, four different frontal images per subject were considered.

Figure 2 – Experimental setup showing diffuse lighting and the CCD camera.

Figure 3 – Examples of FRAV2D Face Database.

3. Face verification system

The face verification process can be divided in four phases: Face location, PCA computation, classifier design and automatic optimal threshold calculation. The first three parts require a training or parameter computation phase and once all parameters have been adjusted and classifiers trained, a normal operation phase. This fourth process will be detailed in section 4.

3.1. Face location

We suppose a-priori that in all our images there is only an upright position face. As the distance from the video camera to the subject has been fixed, the size of faces comprises the natural range of human faces sizes. Therefore we do not expect huge variations in size from one image to another.

In order to improve the performance of this phase, all the images were converted from colour into a gray scale.

The first step in the face location is a background subtraction, which allows to obtain a rough initial estimate of the position. This is done with an image taken with no person in front of the camera.

Each subject in the database has a face template. This has been obtained by means of a convolution with a master face template, which is created only once the very first time our system is used. After the generation of the face template for each individual, its convolution with every single image obtained from the camera allows a reliable and accurate location of the face. According to our experience, we obtained better results when each subject in the database had his/her own template, which is part of the subject's face, instead of using the same master template for everyone.

	SVM			RBF		
	$\bar{x}_i$	s_i	n_i	$\bar{x}_i$	s_i	n_i
$T=1$	4.009	1.735	400	0.828	0.340	400
$T=0$	0.306	0.277	39600	-1.696	0.455	39600

Table 1 – Sufficient statistics for genuine and impostors for the SVM and RBF classifiers. See the text for more details.

When the convolution reaches the maximum over the images, a window containing the face is cropped. The final dimension was reduced to 130×140 pixels.

3.2. Principal Components Analysis Computation

Principal Components Analysis is the standard method in face verification systems. During the training phase, a PCA transformation matrix is computed, so that the eigenfaces retain almost 100% of the initial variance of the faces in the database. In our experiment eight frontal images per subject were considered in order to compute the PCA matrix, and in our tests 150 eigenvalues were considered.

3.3. Verification

We considered two classifiers: Artificial Neural Networks Radial Basis Functions (RBF) and Support Vector Machines (SVM). In all cases, training is performed with eight images per subject (as in PCA). Tests were carried out with four images per subject, without overlap with the training set. If the output value is greater than an acceptance threshold w, it is considered as a positive verification. This value has to be tuned in order to minimize false acceptance and false rejection rates. This threshold is different for each classifier.

RBF was used as an artificial neural network classifier for face verification, with a subject image and a personal identification number (PIN) code as inputs. The Gaussian functions considered were symmetric and centred in the middle of each face subject cluster.

Support Vector Machines offer excellent results in 2-class problems and in verification problems. In our experiment, a linear kernel was considered.

4. Optimal acceptance threshold computation

A Bayesian screening approach is performed to optimize the acceptance threshold [13], considering two variables, namely:

- A binary performance variable T, which identifies whether an image of a given person has been taken ($T = 1$) or not ($T = 0$).

c_a / c_r rate	k	SVM			RBF		
		w	FRR (%)	FAR (%)	w	FRR (%)	FAR (%)
0.1	0.091	–0.717	1	0.17	1.081	2.00	2.45
1	0.500	–0.366	1	0.01	1.527	7.21	0.32
10	0.909	0.001	3.50	0	1.888	11.72	0.31

Table 2 – Optimum acceptance threshold variation with FAR and FRR for the three security levels discussed in the text.

- A screening variable X, which defines the output of a known classifier, for instance, SVM or RFB.

Since the screening variable X is not perfectly correlated with the performance variable, decisions made by using the screen are prone to error (false acceptance and false rejection).

4.1. Design of the Screen

Let our screening variable X be continuous and of the type "the larger the better", that is, a large value of X tends to indicate a matching image or genuine ($T = 1$), whereas a small value of X is a sign of an impostor ($T = 0$).

Under such an assumption, we expect a cut-off point or threshold w to exist, so that it divides our population of candidates between genuine or impostors. This distinction can be done if the variable X is greater than w (genuine) or if it is lower instead (impostors). In the precise situation that X equals exactly w, an arbitrary choice and a compromise has to be taken between acceptance or rejection. In our system, we have considered subjects with $X = w$ as genuine. Therefore, our screen has been precisely defined as:

- If $X \geq w$, the person is accepted.
- If $X < w$, the person is rejected.

4.2 Optimal Acceptance Threshold

The novelty of our method is the automatic computation of the value of the threshold w which minimizes the expected total cost of the procedure.

Let c_a and c_r be the cost paid for a false acceptation and a false rejection by the system, respectively.

The expected total cost of the classification of a subject according to the variable X (considered as the output of a classifier such as SVM or RBF) may be expressed as a function of w as the sum of the total cost due to badly rejected images ("false negative") plus the total cost due to badly accepted images ("false positive"). In other words:

$$ETC(w) = c_r P(T = 1, X < w) + c_a P(T = 0, x \geq w)$$

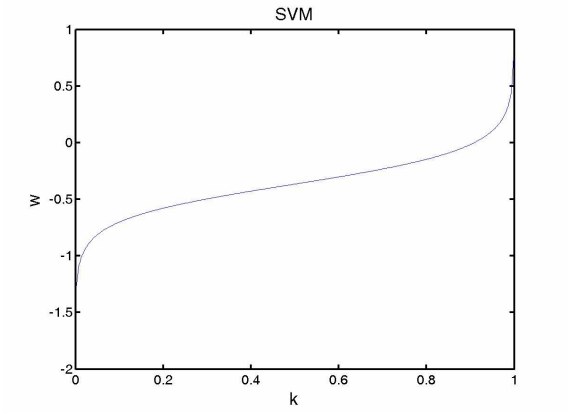

Figure 4 – SVM optimal threshold variation with *k* parameter.

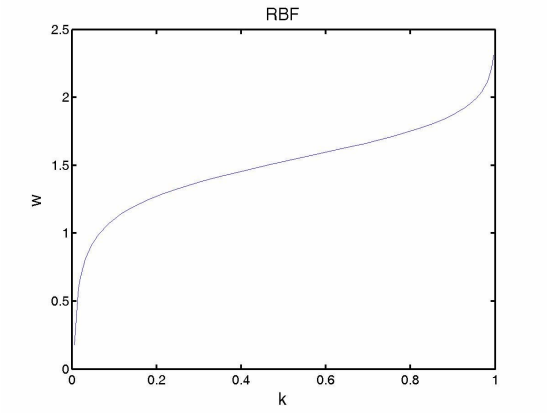

Figure 5 – RBF optimal threshold variation with *k* parameter.

Considering that X is continuous and ETC(w) can be minimized with respect to w, it can be obtained:

$$P(T=1|X=w)=k$$

This equation gives the optimal value w for the acceptance threshold, where $k=c_a/(c_a+c_r)$. This parameter can be more easily calculated than the costs c_a and c_r themselves.

We are therefore interested in the conditional probability of an item with screening value $X = x$ being successful, which, by using a Bayesian approach, results:

$$P(T=1\mid X=x,data)=\frac{f(x|T=1,data)}{\sum_{i=0,1} f(x|T=i,data)},$$

where we are also assuming that $P(T=1)=1/2$, so that both groups are considered as equally probable.

By assuming normal distributions for the variable X in each group, and using non-informative prior distribution for the unknown parameters, the conditional posterior predictive densities $f(x|T=i,data)$ for i = 0, 1 are Student-t density functions,

$$f(x\mid T=i,data)\propto\frac{1}{\sqrt{p_i}}\left\{1+\frac{(x-\bar{x}_i)^2}{(n_i-2)p_i}\right\}^{-\frac{1}{2}(n-1)},$$

where $p_i=(1+n_i^{-1})s_i^2$ and where $\bar{x}_i$, s_i and n_i are the sample mean, sample standard deviation and sample size for each one of the two different groups.

Optimal values of the acceptance threshold w are then calculated by employing numerical techniques.

5. Results and Discussion

We shall present the results in two stages. First the variation of the optimal acceptance threshold w as a function of different acceptance and rejection costs rates shall be shown for both clasifiers, SVM and RBF. Then the variation of FRR and FAR in each cost case will be discussed.

First of all, our results show that the screening variable X is continuous and of the type "the larger the better", as we had previously assumed in our screening setup. Table 1 gives the statistics of our experiment. $\bar{x}_i$ represents the sample mean of our screening variable X, s_i is the sample standard deviation and n_i is the sample size for each case.

A 100-face database with 4 images per subject produces 400 genuine cases (T = 1). For each of the 100 individuals, 99 impostors with 4 images can be considered, yielding a total of 39600 impostors (T = 0).

We have computed the changes of the acceptance threshold w for different c_a and c_r, with k in the range $0 \le k \le 1$, for the two considered classifiers (SVM and RBF). In Figures 4 and 5 we present these results.

Bear in mind that different values of the parameter k can be interpreted as representative of different security levels of an access control. For example:

- In a low security level system, the acceptance cost would be much smaller than the rejection cost. If we consider a standard ratio of $c_a = 0.1c_r$, for instance, then k = 0.091. This is the case of a not very restrictive system, where the FRR is forced to be very low. This security level could be applied, for instance, in a supervised parking access control, where it is more important to let the registered users in, than avoiding intruders.

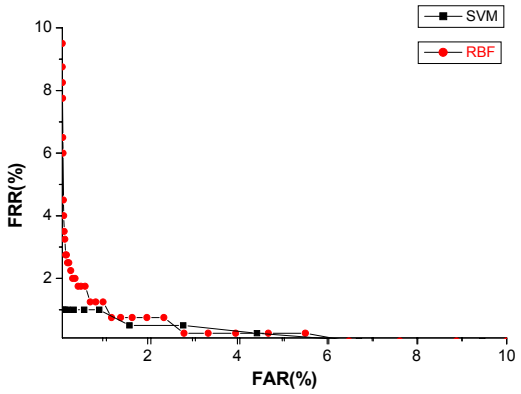

Figure 6 – ROC curves for our system, which plot FRR as a function of FAR. Our SVM classifier (squares) gives more reliable results than RBF (circles), as its curve is closer to the coordinate origin. Bear in mind that the data are expressed in percentages within a range 0 – 10 %, using a linear scale.

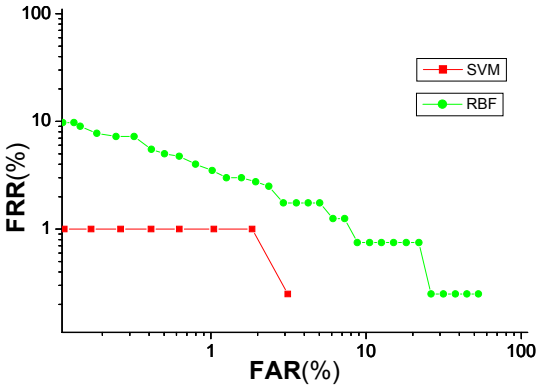

Figure 7 – DET curves for our system. Our SVM classifier (squares) gives better results than RBF (circles), as it yields a lower FAR for a given FRR, and vice versa. Considering EER as a measure of the system performance, the superiority of SVM is clear: EER(SVM)=0.99 and EER(RBF)=2.43.

- A medium security level system could be represented with equivalent rejection and acceptance costs, i.e. $k = 1/2$. In this case false negatives are considered as annoying as false positives.

- For a high level security system, the acceptance cost would be much greater than the rejection cost. For instance, if we assume a standard ratio of $c_a = 10c_r$, then it turns to be $k = 0.909$. In this case impostors have to be repelled at all costs, to the detriment of a possible rise in false rejections.

These three situations are summarized in Table 2. As can be seen, the FAR decreases as the acceptance cost increases. We have observed that, for a RBF classifier, the FAR drops down to 0.31%, while it becomes null for a SVM classifier.

A standard Receiver Operating Characteristic curve (ROC) has been derived for a wide range of optimal acceptance thresholds w (Figure 6). Our SVM classifier (squares) yields a lower FAR for the same FRR, and vice versa, unlike our RBF classifier (circles). Bear in mind that the scale in this graph has been zoomed, in order to distinguish both curves.

In Figure 7, we have used a conventional DET curve [14], which allows a better separation between curves, as both axes are in a logarithmic scale. The point corresponding to FAR = FRR is called Equal Error Rate (EER). This value, which should be as low as possible, is an estimate of performance of our system. As shown in the figure, our SVM classifier produces better results, as EER(SVM) = 0.99, while EER(RBF) = 2.43.

6. Conclusions

In this paper a reliable face verification system with an innovate module has been presented; automatic evaluation of the optimal acceptance threshold using Bayesian screening techniques. This assures that the security level is under control while keeping a minimum error levels.

Using the algorithm proposed, the user is allowed to provide the cost that is assumed to pay for false acceptance or false rejection. This allows the tailoring of our system to user security requirements. Furthermore, the user may indicate the value of the level of security required in an intuitive way, and parameter computation is hidden to the user. The system proposed can work under several security conditions that can be changed by the user.

The method proposed here is valid for all face verification systems, independently of the classifier. Its integration in an existing system has been performed and results show that integration of the algorithm is not expensive.

In this way, any face verification system could be adapted to the environment and the specific conditions of the future application to obtain satisfactory results.

This work is being improved considering faces with non-frontal views, occlusions and different illumination conditions.

7. Acknowledgements

This paper has been supported by grants from the Universidad Rey Juan Carlos (PPR-2003-41, GCO-2003-15) and Comunidad de Madrid (Spain). The authors would also like to thank Javier Arjona for his work.

8. References

[1] M. Turk, A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*. *3* (1), 1991, 71-86.

[2] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman. Eigenfaces vs Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions in Pattern Analysis and Machine Intelligence*, *19* (7), 1997, 711-720.

[3] L. Wiskott, J-M. Fellous, N. Krüger, C. von der Malsburg. Face Recognition by Elastic Bunch Graph Matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. *19* (7), 1997, 775-789.

[4] P. S. Penev, J. J. Atick. Local feature analysis: a general statistical theory for object representation. *Network: Computation in Neural Systems*. *7* (3), 1996, 477-500.

[5] J. J. Atick, P. A. Griffin, A. N. Redlich. Statistical approach to shape from shading: reconstruction of 3D face surfaces from single 2D images. *Neural Computation*. *8* (6), 1996, 1321-1340.

[6] R. L. Hsu. Face detection and modelling for recognition. PhD. Thesis. Michigan State University. Dpt. Computer Science and Engineering. 2002.

[7] P. S. Penev, L. Sirovich. The global dimensionality of face space. *Proc. Fourth IEEE International Conference on Face and Gesture Recognition*. 2000, 264-270.

[8] P. J. Phillips, H. Moon, S. Rizvi, and P. Rauss. The FERET Evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, *22* (10), 2000.

[9] P. J. Phillips, H. Wechsler, J. Huang, and P. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, *16* (5), 1998, 295-306.

[10] P. J. Phillips, A. Martin, C. L. Wilson, and M. Przybocki. An introduction to evaluating biometric systems, *Computer*, *33*, 2000, 56-63.

[11] P.J. Phillips, P. Grother, R.J Micheals, D.M. Blackburn, E. Tabassi, and J.M. Bone. FRVT 2002: Evaluation Report. 2003. http://www.frvt.com/

[12] http://frav.escet.urjc.es/databases/FRAV2D/

[13] R. Montes Diez. Optimal Design of Two-Stage Screens: A Bayesian Approach. PhD. Thesis. University of Nottingham. Maths and Science Dpt. 2000.

[14] A. Martin et al. The DET curve in assessment of detection task performance. *Eurospeech 97*, *4*, 1997, 1895-1898.