Meriem Fedias[1,2,*]. 2. Saigaa Djamel[2]
Laboratory of Energy Systems Modeling (LMSE), Dept. of Electrical Engineering, University of Biskra, Biskra, Algeria(1),
Department of Electronics, faculty of Technology, University of M'sila, Algeria(2),
ORCID: 1. 0009-0005-3951-7060; 2. 0000-0002-8781-8416

doi:10.15199/48.2024.09.12

Comparison between LDA and PCA with the use of the color to face authentication

***Abstract**. The face authentication systems generally use the representation in levels of gray of the face image like characteristic of entry; but in this work we studied the contribution of the color to the face authentication. for the extraction of face characteristics of the data base, we tested a different color spaces on two methods: the method of principal components Analysis PCA and the method of linear discriminant analysis LDA and once the characteristic vector is extracted, the next stage consists in comparing it with the vector characteristic of face witch authenticated, and with the use of each component color alone at the entry of this system, we calculates the error rates in the two sets of validation and test for the data base XM2VTS according to the protocol of Lausanne, and finally we compare between these two methods to see the improvements which can give us each method with the different color spaces.*

***Streszczenie.** Systemy uwierzytelniania twarzy zazwyczaj wykorzystują reprezentację obrazu twarzy w poziomach szarości, podobnie jak charakterystyczna cecha wejścia; ale w tej pracy badaliśmy wpływ koloru na uwierzytelnianie twarzy. w celu ekstrakcji cech twarzy z bazy danych przetestowaliśmy różne przestrzenie kolorów dwiema metodami: metodą analizy składowych głównych PCA oraz metodą liniowej analizy dyskryminacyjnej LDA i po wyodrębnieniu wektora charakterystycznego kolejnym etapem jest porównanie to z wektorem charakterystycznym twarzy, która została uwierzytelniona i przy użyciu każdego składowego koloru osobno na wejściu do tego systemu obliczamy współczynniki błędów w dwóch zestawach walidacyjnych i testowych dla bazy danych XM2VTS zgodnie z protokołem z Lozanny i na koniec porównujemy te dwie metody, aby zobaczyć ulepszenia, które mogą zapewnić nam każdą metodę z różnymi przestrzeniami kolorów (**Porównanie LDA i PCA z wykorzystaniem uwierzytelniania za pomocą koloru i twarzy**)*

Keywords: principal components Analysis PCA,, linear discriminant analysis LDA, face authentication, color spaces
Słowa kluczowe: główne składowe Analiza PCA, liniowa analiza dyskryminacyjna LDA, uwierzytelnianie twarzą, przestrzenie kolorów

Introduction

To prepare the purpose of a authentication system is to check the identity of an individual after this was identified. thus is not about a system of identification which him is given the responsability to discover the identity a priori unknown of an individual. In this context, we will develop or characterize an algorithm offering an expertise in a particular biometric field: face authentication.

The face is not rigid, it can undergo a large variety of changes due to the expression (joy, sorrow…), to the age, hair… etc, and research in this field is rather recent and the interest of the researchers for this last is significant.

In 1991 were revolving as regards theoretical research, with the publication of the article entitled "eigen faces for recognition" of Pentl and and Turk [1,3], of MIT (Massachusetts Institute of Technology). The article described a revolutionary algorithm, the "eigenfaces", While being based on a principal component analysis PCA and with the method of linear discriminant Analysis (LDA); it use the first clean vectors of the covariance matrix of the apprentissage data .

The spot is simple; the face image is collected by a camera. The subject can arise in front of this one and according to the technique used, the system extracts the characteristics from the face to make the comparison with the characteristics of the claimed person which are preserved in a data base.

This article is organized as follows: section 2 presents the problem of face authentication, section 3 explains the algorithm of the PCA and LDA used for the extraction of characteristics, in section 4 we present a comparison between the experimental results obtained, and finally in section 5 we give the conclusions and the prospects.

Authentification of face

A system of authentification must check the identity which is already known for more safety, if it is really the user or an impostor.

The principle of the system of face authentication of an individual is the extraction of a vector X of characteristics of this last, in order to compare it with a vector Yi which contain the characteristics of this same individual extracted starting from his images which are stored in a data base ($l \le i \le p$, where p is the number of images of face of this person in the whole of training).To estimate the difference between two vectors, it is necessary to introduce a measurement of similarity. Several metric can be used such as for example the Euclidean distances L1 and L2, the distance from Mahalanobis the correlation… etc.

For example, if the Euclidean distance enter the vectors X and Yi is lower than a threshold, it is noted that the image of the face corresponds to the claimed person.

The problem which occupies us it contains two classes, namely on the one hand the customers and other share impostors. A system extremely strict indicates a TFA (Rate of False Acceptance) weak and a TFR (Rate of False Rejection) high. On the other hand a system laxist will be characterized by a high TFA and a rather low TFR. The medium locates some share between the two, and if the error rates are equal, it will be at the rate of equal error or TEE.

Extraction of caracteristic by the PCA

By looking at the images of the data base used here, one realizes directly that appear on the level of the neck of the characteristics not wished like the collars of shirt, the sports shirts, etc. In addition, the hair is also a characteristic changing in the course of time (change of cut, baldness…). the background appears on the images: it is not used for nothing, and inflates the size of the data unnecessarily. Finally the ears pose problem. So much so that two very similar images (with the human eye) could be extremely different if they are compared pixel by pixel. It is thus necessary to extract the characteristics suitable starting from the images [10].

The principal component analysis (PCA) is a linear mathematical method to analyze data , the rules is to seek the directions of the space (axes) which maximizes the variance of the data and minimizes the variation squared compared to the axes [8,12-13].

In the case of the face recognition we regard the set of the faces images of training as a set of random vectors (matrix of faces vectors), where each vector face is consisted the sequence of the lines or the columns of an image of face.

The PCA is applied to this matrix of the faces vectors. It primarily consists in carrying out a reduction of dimensionnality by coding the faces in a new base formed by the first clean vectors (EigenFaces) coming from the calculation of the PCA.

The method of Eigen Faces proceeds as follows:

That is to say $A=(X_1, X_2, X_i,....X_N)$ a matrix of data of dimension represents *nxl*, where each X_i is a face vector of dimension *n* Here *n* is the number of Pixel in the image of face represents , and l is the number of faces images in the set of training.

1. A face vector average $\overline{X}$ is calculated starting from the L vectors images of faces of the set of training.

(1) $$\overline{X} = \frac{1}{L}\sum_{1}^{l} Xi$$

2. The face vector average is withdrawn images of training (one thus eliminates the resemblances to concentrate on the differences), which generates the vectors of differences $\overline{X}i$ associated with each image:

(2) $$\overline{X}i = Xi - \overline{X}$$

3. The vectors $\overline{Xi}$ are combined, coast at coast, to create a data matrix of training $\widetilde{X}$ of size (*N X L*).

(3) $$\widetilde{X} = \left[\overline{X}1\,\overline{X}2.......\overline{X}l\right]$$

The matrix of data $\widetilde{X}$ is multiplied by its transposed to find the matrix of covariance Ω , given by [3]:

(4) $$\Omega = \widetilde{X}.\widetilde{X}^{T}$$

4. The linear transformation of a face vector is given by:

(5) $$Yi = W^{T}.\overline{Xi}$$

where Y_i is a vector characteristic of dimension *kxl* and which contains the coefficients of projection of the face vector $\overline{X}_i$ in the new space of transformation and *W* is a matrix formed by *K* first vectors clean of the matrix of covariance corresponding to *K* greater eigenvalues.

To note well that *K* is much lower than *n (k<<n)*.

Thus by the application of the PCA, the entry face vector of dimension *N* is reducing to a vector characteristic in under space of dimension *K*.

But it is necessary to note that the first vectors sorted according to the decreasing value of the eigenvalues are not inevitably most discriminating for these classes. Thus the *PCA* does not take into account the discrimination of the existing classes; one of the solutions on problem is the use of the linear discriminant analysis (*LDA*).

Extraction of characteristic by *LDA* [12-18]

The steps to follow to extract the discriminants for a set of images are [1-4]:

a) Calculate the within class scatter matrix

For the it h class, a scatter matrix (Si) is calculated as the sum of the covariance matrices of the centred images in that class.

(6) $$Si = \sum_{x \in xi}(x - m_i)(x - m_i)^T$$

m_i is the mean of the images in the class. The within class scatter matrix (S_W) is the sum of all the scatter matrices.

(7) $$S_W = \sum_{i=1}^{c} Si$$

C- is the number of classes.

b) Calculate the between class scatter matrix

The between class scatter matrix (S_B) measures the amount of scatter between classes.

(8) $$S_B = \sum n_i(m_i - m)(m_i - m)^T$$

ni is the number of images in the class, m is the mean of all the images.

3. Solve the generalized eigenvalue problem

Solve for the generalized eigenvectors (V) and eigen values (L) of the within class and between class scatter matrices.

(9) $$S_B V = \Lambda S_W V$$

Keep first *C-l* eigenvectors

Sort the eigenvectors by their associated eigenvalues from high to low and keep the first *C-1* eigenvectors. These eigenvectors form the Fisher basis vectors.

c) Project images onto basis vectors

Project all the original images onto basis vectors by calculating the dot product of the image with each of the basis vectors.

Experimental results

Base done

Our experiments were carried out on frontal face images of the data base *XM2VTS*. The principal choice of this data base is its big size, with 295 people and 2360 images in total and its popularity, since it became a standard in the audio and visual biometric community of multimodel checking of identity [4].

For each person eight catches were carried out in four sessions distributed for five months.

The protocol related to *XM2VTS* divides the base into two categories 200 customers and 95 impostors; the people are of the two sexes and various ages. The photographs are color of high quality and size (256x256).

The protocol of Lausanne shares the data base in three sets [6]:

1. The set of training (training): it contains information concerning the known people of the system (only customers)
2. The set of evaluation (validation): allows fixing the parameters of the face authentication system.
3. The set of test: allows to test the system by presenting images of people to him being completely unknown to him.

For the class of impostors, 95 impostors are divided in two sets: 25 for the set of evaluation and 75 for the set of test.

The sizes of the various sets are included in table 1.

Table 1. Distribution of the photographs in the various

set	customers	impostors
training	600 (3 by subject)	0
Evaluation	600 (3 by subject)	200(8 by subject)
Test	400(2 subject)	560 (8 by subject)

Figure 1 represents some examples of images of faces of the data base *XM2VTS*.

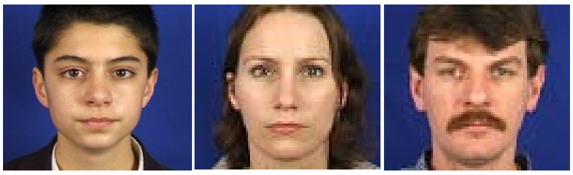

Fig.1. Examples of photographs of the data base XM2VTS.

Pre-treatment

Indeed, all information which is not used for nothing, but inflates the size of the data unnecessarily, for example: the hair, background, ears… etc. Require a reduction of image from which the operation is to extract only the essential parameters for the identifier and who more stable with time.

The decimation consists in taking a pixel on two. That reduces the resolution of the images of course. This operation is preceded by a filtering passes low, destroying the high frequencies, so as to observe the conditions of sampling. In our work we used a filter passes low uniform (2x2) in order to carry out a decimation of factor 2. That reduces by a factor 4 the size of the cut image. The image of face will pass thus from a dimension 256x256=65536 pixels towards a dimension of 66x60=3960 pixels (after cutting and decimation, as illustrated on figure 3).

Then we make the photo normalisation with the images it means: that for each image, we withdraw from each pixel the average value of those on the image, and that we divide those by their standard deviation. The photo normalisation for a double purpose: on the one hand it removes for any vector a possible shift compared to the origin, and then any effect of amplification.

Finally one applies the standardization which acts on a group of images (for each component, one withdraws the average of this component for all the images and one divides by the standard deviation) [6].

(a) (b) (c)

Fig. 2. a) image of entry, b) image after cutting and c) image after decimation.

Extraction of characteristic

The extraction of the characteristics is done by the method of Eigen Faces based on the PCA and LDA, we already detailed it with section 3.

Classification

The goal of an authentication system is knowing on the one hand the customers and other share impostors thus we have two classes. An authentication system extremely strict indicates a weak TFA and a high TFR. On the other hand a system laxist will be characterized by a high TFA and a rather low TFR. The medium locates some share between the two, and if the error rates are equal, it will be at the rate of equal error or TEE.

All these error rates are calculated in two sets: initially in a set of evaluation, which will make it possible to fix the TEE more or less vary some the parameters of acceptance and rejection of the system. Then in a set of test by using the parameters fixed previously. Thus, we can check the robustness of the system.

Comparison

The goal of our system is a binary decision with the rate minimum of equal error, in an authentication system we want to answer by yes if it is a customer or not if it is an impostor.

To do a comparison of results, we presented the latter with two basic methods the *PCA* and *LDA*, which have parameters:

- Pre-treatment with photo normalisation
- Coefficients: following coefficients of sorted projection decreasing eigenvalues.
- Measurement of score (similarity): correlation.
- Thresholding: Total.

We found that the use of the representation in levels of gray gives TS of about 89.16% with the method of PCA and TS about 93.03% with LDA method.

We present the various color errors in the two sets evaluation set and test set in table II with the *PCA* method, and table III with LDA method.

In table II we found that the *PCA* method achieves 3.82% equal error rate on face authentication system with the use of the component color U of the color space YUV as characteristic of the entry system and gives the best rate of succeed TS= 93.39 %.

This means that the use of color information by *PCA* method of the component *U* of color space *YUV*, as characteristic of entry in face authentication system, represents an improvement in the rate of succeed about 4.23% compared to the use of images represented in greyscale.

Table 2. Error rate with the PCA.

comp	evaluation set			Test set		
	TFR	TFA	TEE	TFR	TFA	TS
X	0,047	0,046	4,66%	0,050	0,060	89,04%
Y	0,045	0,045	4,52%	0,048	0,059	89,40%
Z	0,043	0,044	4,37%	0,045	0,056	89,88%
Y	0,048	0,047	4,79%	0,053	0,059	88,81%
Cr	0,035	0,035	3,52%	0,048	0,039	91,36%
Cb	0,037	0,037	3,66%	0,045	0,029	92,57%
R	0,052	0,051	5,14%	0,058	0,066	87,65%
G	0,043	0,043	4,34%	0,053	0,058	88,99%
B	0,045	0,044	4,46%	0,045	0,057	89,81%
Y	0,045	0,046	4,54%	0,050	0,059	89,12%
I	0,040	0,041	4,04%	0,045	0,034	92,09%
Q	0,030	0,031	3,04%	0,050	0,032	91,76%
Y	0,045	0,046	4,54%	0,050	0,059	89,12%
U	**0,038**	**0,038**	**3,82%**	**0,030**	**0,036**	**93,39%**
V	0,042	0,042	4,18%	0,035	0,037	92,77%
H	0,127	0,126	12,64%	0,143	0,148	70,91%
S	0,035	0,034	3,45%	0,043	0,047	91,09%
V	0,052	0,051	5,14%	0,058	0,066	87,63%
I1	0,047	0,047	4,71%	0,053	0,060	88,74%
I2	0,048	0,049	4,85%	0,048	0,063	88,92%
I3	0,037	0,037	3,69%	0,043	0,040	91,71%

Table 3. Error rate with LDA

comp	Evaluation set			Test set		
	TFR	TFA	TEE	TFR	TFA	TS
X	0,030	0,030	3,01%	0,030	0,031	93,91%
Y	0,032	0,033	3,22%	0,025	0,035	93,97%
Z	0,027	0,027	2,69%	0,018	0,030	95,23%
Y	0,030	0,030	2,99%	0,033	0,031	93,62%
Cr	**0,017**	**0,016**	**1,65%**	**0,023**	**0,017**	**96,10%**
Cb	0,032	0,032	3,20%	0,025	0,029	94,62%
R	0,037	0,036	3,65%	0,048	0,033	91,94%
G	0,030	0,031	3,04%	0,023	0,032	94,56%
B	0,027	0,026	2,65%	0,015	0,030	95,53%
Y	0,032	0,033	3,22%	0,030	0,032	93,78%
I	0,028	0,028	2,83%	0,028	0,022	95,03%
Q	0,020	0,020	1,99%	0,030	0,020	94,99%
Y	0,032	0,033	3,22%	0,030	0,032	93,78%
U	0,030	0,031	3,05%	0,020	0,027	95,33%
V	0,022	0,021	2,13%	0,025	0,022	95,27%
H	0,067	0,066	6,65%	0,055	0,083	86,22%
S	0,023	0,024	2,38%	0,030	0,026	94,38%
V	0,033	0,034	3,35%	0,040	0,031	92,89%
I1	0,030	0,031	3,03%	0,030	0,033	93,73%
I2	0,033	0,032	3,21%	0,038	0,032	93,06%
I3	0,017	0,017	1,67%	0,030	0,015	95,50%

Table 4. Comparison between PCA and LDA

comp	TS (LDA)	TS (PCA)	TS_{LDA}-TS_{PCA}
X	93,91%	89,04%	4,87%
Y	93,97%	89,40%	4,57%
Z	95,23%	89,88%	5,35%
Y	93,62%	88,81%	4,81%
Cr	96,10%	91,36%	4,74%
Cb	94,62%	92,57%	2,05%
R	91,94%	87,65%	4,29%
G	94,56%	88,99%	5,57%
B	95,53%	89,81%	5,72%
Y	93,78%	89,12%	4,66%
I	95,03%	92,09%	2,94%
Q	94,99%	91,76%	3,23%
Y	93,78%	89,12%	4,66%
U	95,33%	93,39%	1,94%
V	95,27%	92,77%	2,50%
H	86,22%	70,91%	15,31%
S	94,38%	91,09%	3,29%
V	92,89%	87,63%	5,26%
I1	93,73%	88,74%	4,99%
I2	93,06%	88,92%	4,14%
I3	95,50%	91,71%	3,79%
Level of greys	93,03%	89,16%	3,87%

in table III we observed that the LDA method achieves 1.65% equal error rate on face authentication system with the use of the component color Cr of the color space YCrCb as characteristic of the entry system and gives the best rate of succeed TS= 96.10 %.

This means that the use of color information by LDA method of the component Cr of color space YCrCb, as characteristic of entry in face authentication system, represents an improvement in the rate of succeed about 3.07% compared to the use of images represented in Level of greys.

Conclusion

The face authentication system uses usually the representation of the face image in levels of gray like characteristic of entry. And since our principal goal is the improvement of the performances of our authentication system, we have to introduce thus color information by the use of many color spaces.

we found that the PCA method achieves 3.82% equal error rate on face authentication system with the use of the component color U of the color space YUV as characteristic of the entry system and gives the best rate of succeed TS= 93.39 %. the use of a single component color with the linear discriminate analysis LDA achieves 1.65% equal error rate so 96.10% rate of succeed using only 100 features apply the color component of color space *YCrCb*.

So the color information achieves 6.94% in rate of succeed; if we compare the best rate of succeed in LDA method of the color component Cr of color space *YCrCb* with the rate of succeed in level gray in PCA method.

In future work we propose the nonlinear fusion of the various components or spaces color with other methods like: analysis in independent components (ICA)...

Authors. *1. Meriem Fedias was born in Biskra Algeria. After succeeding in the baccalau-reat in 1999, she graduated from the University of Mohammed khider Biskra, Al-geria, with a state Engineering degree in Electronics Majoring in Control, in June 2004.*
She received the Magister degree in electronic from the University of Biskra in 2007, and since then she working with the MSE Laboratory. She received her PhD in Automatic from Biskra University, Algeria in 2013. She is currently an Associate Professor at the Genie Electric Department of Biskra University, Algeria.
Automatic Department, University Mohamed khider B.P 45
BP Biskra (07000), Algeria, LMSE laboratory.
E-mail: m_fedias@yahoo.fr.
2.Djamel Saigaa received his Ingenious and his 'Magister' degree from Sétif University, Algeria in 1990 and 1993 respectively. He received his PhD in Automatic and signal processing from Biskra University, Algeria in 2006. He is currently an Associate Professor at the Electronics Department of M'Sila University,Algeria. His research interests include Digital signal processing and Biometric recognition Systems. E-mail: dj-saigaa@yahoo.fr.

REFERENCES

[1] Turk M.A., Pentland A.P., Face recognition using eigenfaces. IEEE Comput. Sco. Press, (1991), 586-591.
[2]Chellappa R., Wilson C.L., Sirohey S., Human and machine recognition of faces: A survey, Proceedings of the IEEE, (1995), 705-740.
[3] Matthew T., Alex P., Eigenfaces for recognition. Journal of cognitive neuroscience, 3 (1991., nr 1, 71–86
[4] Messer K., Matas J., Kittler J., Jonsson K., Xm2vtsdb : The extended m2vts database. Audio- and Video-based Biometric Person Authentication (AVBPA), (1999), 72–77.
[5] Luettin J., Maitre G., Evaluation protocol for the extended M2VTS database. (1998).
[6] Saigaa D., Benoudjit N., Benmahamed K., Leland-ais S., Authentification d'individus par reconnaissance de visage, courier de savoir Biskra, Algerie, (2005), nr6.
[7] Hardenberg J.Y., Acquisition et reproduction d'images couleur : approches colorimétrique et multi spectrale, doctorat thesis, Ecole Nationale Supérieure des Télécommunications Specialité : Signal et Images 1999.

[8] Havran C., Hupet L., Czyz J., Lee J., Vandendorpe L., Verleysen M., Independent Component Analysis for face authentication" KES'2002 proceedings – Knowledge- Based Intelligent Information and Engineering Systems, Crema (Italy), 16- 18 September (2002).
[9] Saigaa D., Benoudjit N., Bemahammed K., Lelandaius S., Face Authentication using Enhanced Fisher linear discriminant Model (EFM), WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics (CIMMACS'05), in Miami Florida (USA), Nov. 17-19 (2005), 155-160.
[10] Belhumeur P., Hespanha J.P., Kriegman D.J., Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. on Pattern Analysis and Machine Intelligence, (1997), 711-720.
[11] Fedias M., Saigaa D., L'apport de la couleur à l'authentification de visage, Conférence internationale à constantine JIG'2007 proceedings, (2007), 120-126.
[12] Wendy S. Yambor, Analysis of PCA-based and fisher discriminantbased image recognition algorithms, Technical Report CS-00-103, july (2000).
[13] Belhumeur P.N., Hespanha J.P., Kriegman D.J.. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Trans. on Pattern Recognition and Machine Intelligence, 19 (1997), nr7, 711– 720.
[14] Zhao W., Chellappa R., Krishnaswamy A., Discriminant analysis of principal components for face recognition. In Proc. of Int. Conf. on Automatic Face and Gesture Recognition,(998), 336– 341.
[15] Liu C., Wechsler H., Enhanced fisher linear discriminant models for face recognition, In Proceeding of the International Conference on Pattern Recognition (ICPR), 1998.
[16] M. Fedias, D. Saigaa " Nonlinear fusion of colors to face authentication ", Proc. Conf. ICEEDT'08, International conference on electrical engineering design and technologies, Hammamet, Tunisia, Nov. 2008.
[17] M. Fedias, D. Saigaa "Non linear fusion of colors to face authentication using EFM method, Journal of Applied Computer Science & Mathematics Romania (JACSM), (2010), nr9, 42-50.
[18] Saigaa D., Fedias M., Harrag A., Bouchelaghem A., Drif D., Color space MS based feature extraction method for face verification, the 11th International conference on hybrid intelligent systems HIS' 2011, (2011), 328-333.