Available online at www.scholarsresearchlibrary.com

Scholars Research Library

European Journal of Applied Engineering and Scientific Research, 2015, 4 (3):24-30
(http://scholarsresearchlibrary.com/archive.html)

ISSN: 2278 – 0041
CODEN(USA): EJAEAJ

Face recognition using Eigen faces, PCA and support vector machines

[1]Parvaneh Shayghan Gharamaleki and [2]Hadi Seyedarabi[*]

[1]Department of Mecatronic Science, Ahar Branch, Islamic Azad University, Ahar, Iran
[2]Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

---

ABSTRACT

This paper is based on a combination of the principal component analysis (PCA), eigenface and support vector machines. Using N-fold method and with respect to the value of N, any person's face images are divided into two sections. As a result, vectors of training features and test features are obtained. Classification precision and accuracy was examined with three different types of kernel and appropriate number of face features was considered and the best function for system identification rate. Then, face features were fed into the support vector machine (SVM) with one vs. all classification. At first, 2-Fold method was examined for images of training and test system. The results indicated that the rotation of the sets in identical classifications had no impact on the efficiency of radial basis function (RBF). It was observed that the precision increased in the 5-Fold method. Then, 10-Fold method was examined which indicated that the average recognition rate further increased when compared with 2-Fold and 10-Fold methods. The results revealed that as the rotation number increases, the precision and efficiency of the proposed method for face recognition increases.

Key words: Face Recognition, Eigenfeatures, Eigenfaces, Multi-Layer Perceptron (MLP), Support Vector Machines(SVM), Principal Component Analysis(PCA), Radial Basis Function(RBF)

INTRODUCTION

Faces play outstanding roles in identifying and recognizing people and showing emotions at the level and scope of a society. The initial research studies in the realm of face recognition date back to the end of year 1980 [1].

Humans' ability in recognizing faces is remarkable since we can recall and recognize thousands of faces which we learned throughout our lives. That is to say, we can even recognize the faces of acquaintances many years later even though they have undergone changes in their facial features due to aging, growing beard, long hair, etc. Face recognition is considered to be a significant issue in security systems, identification of criminals, credit card control, etc. For instance, the ability to model a particular face and identify it among so many faces stored in a database can notably improve the identification of criminals [2].

In face recognition, based on the trained faces, a system can choose a face which is more similar to the respective face and consider it as the final response [3].

The most significant difference among the face recognition methods is related to the way they extract and display facial features and components. Until now, different methods have been proposed for extracting face features which can be divided into two general types, i.e. structure-based methods and feature-based methods [4]. Structure-based methods of face recognition are non-monitored methods which produce appropriate responses to the linear facial changes. As a case in point, principal component analysis (PCA) is regarded as a linear transformation which uses input data variance [5]. In general, face recognition methods consist of a feature extractor and a classifier [6]. The remaining sections of the paper are organized as follows. In section 2, PCA algorithm is described and discussed with respect to feature extraction. Section 3 introduces support vector machine (SVM) for classification. In section 4, we will introduce ORL picture data and in Section 5 and Section 6 concludes simulation is proposed.

2-Principal Component Analysis

In a collection of the information, it is considered as the main recognition patterns and methods being applied in the dimension reduction stating the similarities and differences based on their own main communicative elements. The main element analysis is a powerful tool for analyzing the data. By the reduction of the number of dimensions without losing the main volume, it will find the pattern saving the related data. The main purpose of the related algorithm is to summarize the data taking them in the process of saving and compacting issues [7].

2.1-Principal Component Analysis in the face recognition:

The main idea of the principal component analysis has been suggested by Trek and Pantland in 1991 [8] the main regular basis of the Pantland suggestion is to use the PCA for defining and describing the face features by Kirby and Sirovich [9]. This method has been also applied to reduce the dimension and extraction feature to be able to interact with the sub-spatial vectors. This makes a better display for the data distribution. This sub-space is called the face space when it appears on the face data. After specifying the vectors, the whole pictures will be transferred into the sub-space representing the same sub-space. By comparing the recent weighs similarities with the new picture weigh, it can specify the entered picture[10].

2.2-Eigenface methodology:

This method has been suggested by Pantland for continuing the studies in relation to recognize the face [11] making the data independent to the extent possible and this can be obtained by interacting the whole vertical vectors and PCA will be applied here. Indeed, the Eigenface uses the PCA Algorithm filteration for compacting its information; in other words, it finds firstly the covariance matrix feature vectors using these vectors for transferring the information and finally it achieves the reduced dimentional space for fulfilling the recognition process through its neighbor.

2.3-The eigen extraction of the PCA:

The first phase of the eigen extraction is subjected to the person's face features in every face recognition system. The suggestive algorithm is subjected to PCA in this thesis. The extracted features of the pictures in one space with the higher dimensions are the most essential cases increasing the calculation complexities in every categorization phase. On the other hand, this can reduce the rate and performance of the categorizing system. The main reason is subjected to the distribution of the feature vectors in a space with higher dimension. To prevent the related problem, it is necessary to write a sub-space with lower dimension before categorizing along with a suitable method of feature with a sub–space higer dimentional vector or eigen vector. For achieving this, PCA linear separator will be utilized potentially[12].

2.4-Calculation of eigen face:

Every picture is considered as two-dimensional matrix T. In this method,in fact every T_i is a picture with $n\times m$ dimension which is transferred into $P = nm$; in other words, we consider a picture as a column vector with $P\times 1$ element *such as* $T^i = (T_1^i \ldots\ldots T_P^i)^T$ .The obtained vectors make the columns of matrix A respectively . The dimensions of this matrix is $M\times$P that M is related to the number of the pictures. In the next phase, the total average of faces will be obtained as following:

$$\Psi = \frac{1}{M}\sum_{n=1}^{M} T_n \qquad (1)$$

We will subtract the total average of and save these vectors into a matrix. Indeed, the data will be transferred into the zero centers. By doing this, the mean of the faces will also be zero getting ready for calculating the basic elements[4] .

$$\Phi_i = T_i - \Psi \qquad For \qquad i=1,2,\ldots\ldots, M \qquad (2)$$

The covariance matrix will be measured for the new matrix that it will be also obtained by the multiply of the transpose matrix A into itself.

$$C = A^T A \qquad (3)$$

Note that $(\Phi_i)s$ make the columns of the matrix A

$A=\{\Phi_1, \Phi_2, \Phi_3, \ldots\ldots, \Phi_i\}$ For $i=1,2,\ldots\ldots, M$ (4)

The covariance m atrix, eigen vectors and eigen values are obtained using the following equation:

$$C\,V' = \lambda' \, V' \quad (5)$$

V' refers to eigen vectors and λ' to eigen values.

The eigenvector matrix is the same vertical vectors that compose the sub-space feature transferring the data into these sub-spaces in order to be independent [13].

The eigen features can be applied to describe the human's face features [4] that can also be obtained by the following equation:

$$V = A.\,V' \quad (6)$$

In the next phase, the eigen values and eigen vectors should be arranged from the big towards small. We ignore vectors with small changes, because most of the changes are done to the special value, and gradually move on to the smaller eign values to reach the less scattering vectors . By innear product V^T into matrix A, the data matrix will be transferred into a new sub-space.

$$A_{project} = V^T.A \quad (7)$$

Pictures in the new spaces are defined by a matrix $A_{project}$. The new matrix is the basis for comparison.

3-Support Vector Machine (SVM):
The support vector machine is a training algorithm for learning the categorization and regulation of the regression from the related data. This algorithm has been suggested by Vapnik Vladimir in 1965s [14] as the most famous trainees' categorization ,a Russian researcher, [15] and has been also recovered by Vapnik and Corinna Cortes in 1995 for the nonlinear mood [16] coming from the Statistical Learning Theory being organized and arranged on the operational risk minimization process. This method is one of the most fairly newest approaches that has been innovated in the recent years in comparison to the traditional methods such as Perceptron Neural Nets [17]. The support vector machines take in to consideration the operational risk as the aim variable and calculate the optimized value. The main purpose of the support vector machine is to obtain the function F(x) as a determinant of the hyper-sheet. There have been many various hyper-sheets that are able to separate the data. But the main question is what hyper-sheet to choose? The training concept of the pictures being categorized into the higher dimensions in a one space is not unique. The main distinction of the related algorithm is subjected to the selection this hyper-sheet.

Figure 1 shows the best way of choosing the separalive.

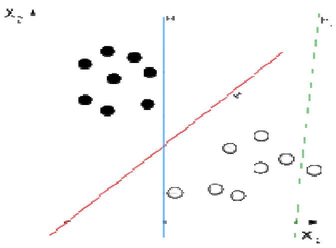

Figure 1: Way of the best separator [16]

As it is shown in figure 1, there have been established some separator to categorize the related data in this case. But due to the figure, the separators *H3* has not separated the two categories correctly. The separators *H1* and *H2* have achieved the best separation but if one case of data is measured again being established out of the separator *H1*, it will be specified that the separation has not been achieved efficiently, but the separator *H2* has categorized the task correctly. Indeed, the difference between *H1* and *H2* is correct in the operational risk or the lack of categorization risk. In the support vector machine the main purpose is subjected to maximize the margins of the two classes. Thus, a hyper-sheet should be selected whose distance from the nearest data in both sides of linear separators is maximum.

If we reach such a hyper-sheet, it will be introduced as the maximum margin separator [18]. Figure 2 shows the related explanation and description.

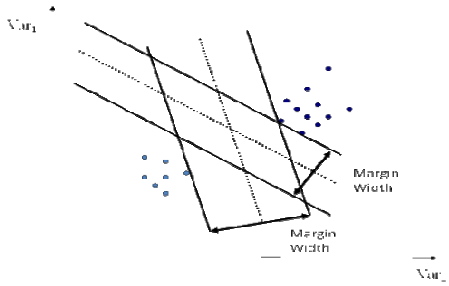

Figure 2: Shows the selection of the maximum hyper-sheet [16]

In order to separate the data, two territorial sheets which are in parallel with separation sheet, must be drawn and then made away from each other to the extent that they contact the related data; this makes the appearance of the maximum margin. The best separator is a sheet that has the highest distance from the territorial sheets. The nearest training samples to the separator hyper-sheet are called the support vectors so that they compose the classes' territories.

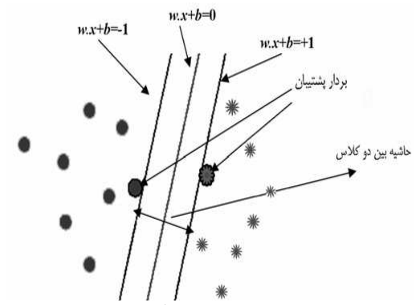

Figure 3: Optimized hyper-sheet for showing two complete separate cases [19]

The decision-making function for categorizing the data is determined with the support vectors and the use of these support vectors instead of the whole data can lead this algorithm to compact the related data. The linear support vector machine is also used as a rapid and accurate categorizing machin for seeking the face in a two-dimensional space [20].

4-ORL Face database:

This database has been gathered from 1992 to 1994 in AT & T laboratory [21] and it can be stated that it is one of the most applied face databases for the face recognition algorithms. This database includes 400 different pictures of 40 persons from each 10 pictures have been taken. Photographing process has been carried out in different lightening conditions in different time periods. There have been established different moods for the pictures such as open and close eyes with other details like having beard or being beardless, laugh or without laugh.

The forehead and hair of people can be observable in the related pictures. The face situation towards the camera angle is variable from top to bottom and left to right side. All the pictures are black and white with 92×112 pixels. The most common approach for evaluating the face recognition system is the application of ORL database in which everybody's face pictures have been categorized into two sections according to the system requirement; the first section of the pictures has been applied for training and the second section is subjected to the test pictures issues. For example, five pictures out of ten sections can be considered for training and others left for the test pictures issues.

In the training section of the investigated algorithm, the pictures in the training section are used in order to produce the models of 40 persons in the same 40 classes. In the next phase, the pictures in the test pictures section can be applied to determine the rate of the accuracy of the system. In order to obtain the value of the recognition accurately, the above mentioned evaluation is usually repeated for several times along with training and test pictures collections and general recognition will be represented by measuring the mean of the results.

Figure 4: Related pictures of four persons from ORL database in four lines

5-Simulation:

When the results were obtained, one vs. all classification was conducted using three kernel functions. N-fold category was used to obtain new results. That is to say, among many different poses for each person (10 images), a few images are selected as training samples and the rest are regarded as test samples. According to the one vs. all method, the samples were examined based on the number of features which had the best efficiency in the kernel function and the recognition precision was obtained in the first stage. In the next stages, the images of the training and test samples were periodically replaced. Then, the recognition rate is calculated again for each stage. Finally, after a complete rotation in the images, an average of the total obtained results is calculated which is referred to as average system recognition rate. The most appropriate number of face features and the best function in system recognition rate, obtained in the previous stages, are used for the next stage and the tests are conducted again so as to obtain the new results.

In this section, the impact of classification on the efficiency of the RBF function is examined. The 2-fold method was used in the classification in which the number of training and trial images was the same (five images for each person). The location of training images and trial images will be replaced. The obtained results are given in table (1).

Table (1): The results of the conducted tests in the 2-fold method

The conducted stages	The obtained results (%)
The first stage (2-fold method)	96
The second stage (2-fold method)	96
Average recognition precision	96

As shown in table above, in systems where the training and test sets have been equally classified, the replacement of these sets have no impact on the efficiency of the RBF function. In this stage, the training images are selected with the proportion of 8 to 2 in relation to test images. The average recognition rate is obtained after five sages. That is, in the first stage, 10 images (for one person) are selected so that the first and second images are considered for test and the others are used for training. In the second stage, the same number of classification is used. The third and fourth images are used for test and the remaining images are selected for training. In this way, the same procedure continues until the fifth stage. Then, an average is calculated from the obtained results and the recognition rate of this classification is obtained.

Table (2): The results of the conducted tests in the 5-fold method

The obtained results (%)	The conducted stages
100	The first stage (5-fold method)
98.75	The second stage (5-fold method)
95.0	The third stage(5-fold method)
96.25	The fourth stage(5-fold method)
98.75	The fifth stage(5-fold method)
97.75	Average recognition precision

Finally, in this stage, the training images are selected with the proportion of 9 to 1 in relation to test images. The average recognition rate is obtained after ten stages; that is, ten images (for each person) are used in the first stage. The first image for test and the remaining images are selected for training.

Table (3): The results of the conducted tests in the 10-fold method

The obtained results (%)	The conducted stages
97.50	The first stage (10-fold method)
100	The second stage (10-fold method)
100	The third stage(10-fold method)
100	The fourth stage(10-fold method)
100	The fifth stage(10-fold method)
100	The sixth stage(10-fold method)
100	The seventh stage(10-fold method)
100	The eighth stage(10-fold method)
95.0	The ninth stage(10-fold method)
95.0	The tenth stage(10-fold method)
98.75	Average recognition precision

CONCLUSION

As a result of using N–fold method, each person's face images was divided into two sections. The first section of the images was used for training and the second section was used for test. Then, features were extracted from the images by means of the linear algorithm of the principal component analysis (PCA). The results of these stages are the attainment of the training and trial feature vectors. The vector of features (10 to 90) for the total set was examined by three kernel functions. The obtained results were as follows: 96% for the RBF kernel function with 20 feature vectors (20 eigenfaces), 97.5% for the kernel MLP function with 50 feature vectors and 93.5% for the polynomial kernel function with ten feature vectors. With respect to the most appropriate number of special faces (20) and the best kernel function in the system recognition rate, they were fed into the support vector machine classifier. The one vs. all method was used for classification. Firstly, 2-fold method was tested for the system. The recognition rate of the system in both stages was 96%. The obtained results indicated that, in systems in which training and test sets are classified equally, rotation of such sets have no impact on the efficiency of the RBF kernel function.

In 5-fold method, the 5-stage test revealed that the recognition rate is 97.75% which had increased in comparison with the 2-fold method. In the 10-fold method, the 10-stage test indicated that the obtained average recognition rate was 98.75%. The results indicated that as the number of rotations increase, the recognition rate and efficiency of the proposed method for recognizing face increases. However, it should be noted that the speed of recognition decreases which is due to an increase in the number of rotations.

REFERENCES

[1] S. K Roy, S Bandyopadhyay, *International Journal of Computer and Technology*, **2013**, 10(8), 1943-1953.
[2] I Atalay, Face Recognition Using Eigenfaces. M. Sc. Thesis, Istanbul Technical University Institute of Science and Technology,**1996**.
[3] L Xianwei, C Guolong, *Photonics and Optoelectronics(SOPO)*, **2012**, 1-4.
[4] M Elahi, J Shirzi, *The Eighth Symposium of Science and Technology Advances***, 2013**, Khavaran Higher Education Institute.
[5] H Ghajar, M seryani, A Kochari, *Biannual Journal of Signal and Data rocessing*, **2010**, 7(1).
[6] E Gumus, N Kilic, A Sertbas, O.N Ucan, The *Online Journal on Electronics and Electrical Engineering*, 2(4), 308-310.
[7] A tutorial on Principal Component Analysis, Lindsey Smith (**2002**): http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf
[8] M Turk, A Pentland, *J COGNITIVE NEVROSCI*, **1991**, 3(1).
[9] M Kirby, L Sirovich, *IEEE T PATTERN ANAL*, **1990**, 12(1), 103-108.
[10] http://www.boute.ir/iust-ai-92/face-recognition
[11] A Pentland, B Moghaddam,T Starner, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **1994**, 84-91.
[12] S Meshkini, Automatic Face Recognition Using Support Vector Machines, Ph.D. Thesis, University of Tabriz, **2012**.
[13]http://www.irmatlab.irm/metod -eigenface
[14] R Raye, S Fallahpur, *Review of Accounting and Auditing*, **2008**,15(53), 17 - 34.
[15] C J. C. Burges, *DATA MIN KNOWL DISC*, **1998,** 2, 121-167.
[16] http://fa.wikipedia.org/
[17] A Nikbakht Shahbazi, B Zahraie, M Naseri, *Water and WasteWater*, **2012**, 2, 72-84.
[18] SM KolahyHeris, SVM Support Vector Machines or in MATLAB Tutorial Comprehensive Fradars. http://faradars.org/product/mvrnn9102fh-support-vector-machine-in-matlab-video-tutorial/

[19] A Keshavarz, H GhasemianYazdi, *Computer and Electricity Magazine*, **2005**, 3(1).
[20]http://www.researchgate.net/profile/Bahman_Deldadeh_Barani/publication/270824704_Support_Vector.
[21] F.S Samaria, A.C. Harter, IEEE *Workshop on Applications of Computer Vision*, **1994**, 138 - 142.