International Journal of Electrical and Computer Engineering (IJECE)
Vol.2, No.6, December 2012, pp. 766~773
ISSN: 2088-8708

The new method of Extraction and Analysis of Non-linear Features for face recognition

Ali Mahdavi Hormat*, Karim Faez, Zeynab Shokoohi*, Mohammad Zaher Karimi***
* Departement of Electrical and Computer Engineering, Qazvin Branch, Islamic Azad University
** Departement of Electrical Engineering, Amirkabir University of Technology

Article Info

Article history:

Received Jul 28, 2012
Revised Nov 19, 2012
Accepted Nov 28, 2012

Keyword:

Face recognition
Linear features
Nonlinear features
Image Processing

ABSTRACT

In this paper, we introduce the new method of Extraction and Analysis of Non-linear Features (EANF) for face recognition based on extraction and analysis of nonlinear features i.e. Locality Preserving Analysis. In our proposed algorithm, EANF removes disadvantages such as the length of search space, different sizes and qualities of imagees due to various conditions of imaging time that has led to problems in the previous algorithms and removes the disadvantages of ELPDA methods (local neighborhood separator analysis) using the Scatter matrix in the form of a between-class scatter that this matrix introduces and displayes the nearest neighbors to K of the outer class by the samples. In addition, another advantage of EANF is high-speed in the face recognition through miniaturizing the size of feature matrix by NLPCA (Non-Linear Locality Preserving Analysis). Finally, the results of tests on FERET Dataset show the impact of the proposed method on the face recognition.

Corresponding Author:

Ali mahdavi hormat,
Departement of Electrical and Computer Engineering,
Qazvin Branch, Islamic Azad University,
Nokhbegan Blvd, Qazvin, iran.
Email: ali.mahdavi.hormat@qiau.ac.ir

1. INTRODUCTION

The subspace learning methods have been considered according to their position in the pattern classification of the machine vision and learning in recent years. Whithin the past two decades, many subspace learning methods have been suggested for face recognition; These methods are generally divided into two categories: The first category is the learning methods with the supervisor and the second one is without the supervisor[2-11].

In this paper, we propose a new method for the analysis of extracting nonlinear features. This method reduces the dimensions of feature matrix to 3 * N (N is the number of image samples) and also to demonstrate the problem of size in the smaller samples, our objective function includs the separator matrix of within-class scatter feature. The results performed on the FERET Dataset shows the impact of EANF method.

2. LOCAL NEIGHBORHOOD SEPARATOR ANALYSIS

We consider the set of $X = [x1, x2, \cdots, xN]$ as the example of class C $\{\omega1, \omega2, \cdots, \omega C\}$ when $xi \in Rn$. Subspace learning methods try to find the transfer function Φ, so that the transformation from ndimensional space to the d-dimensional space (d «n) is possible through minimizing or maximizing the objective function when yi ∈ R^d.

the transformation function Φ is equal to yi = $\Phi^T x_i$. In the way that the dimensions of space reduces for between-class scatter matrix and increases for within-class scatter matrix.The structure of the between-class scatter matrix and within-class one is, respectively as following :

$$S_b^L = \sum_{ci=1}^{C}\sum_{cj=1}^{C}\left(\bar{x}c_i - \bar{x}\,c_j\right) W_{c_icj}\left(\bar{x}c_i - \bar{x}\,c_j\right)^T \quad (1)$$

$$S_w^L = \sum_{c=1}^{C}\sum_{xi,xj\in w0}^{1}\left(x_i - x_j\right) \mathrm{W}^{(c)}{}_{ij}\left(x_i - x_j\right)^T \quad (2)$$

$\bar{\mathbf{x}}\,\mathbf{c_i}$ shows the mean vector of ωci , and the weight of $w_{ij}^{©}$ for data pairs of $\boldsymbol{x_i}$ و $\boldsymbol{x_j}$ into c^{th} class is as following:

$$W_{c_icj} = \begin{cases}\exp\left(-\frac{\left|\left|\bar{x}c_i - \bar{x}\,c_j\right|\right|^2}{2\sigma^2}\right)\\ 0,\end{cases} \quad (3)$$

$$W_{ij}^{(c)} = \begin{cases}\exp\left(-\frac{\left|\left|xc_i - x\,c_j\right|\right|^2}{2\sigma^2}\right)\\ 0,\end{cases} \quad (4)$$

σ has also been experimentally determined.

Given the use of the linear PCA, ELPDA method has two problem: the high volume of feature vector and especial matrix error in the restoration of the original amount of matrix, which is solved using NLPCA. This function has selected all linear and non-linear parts and led to remove the additional parts and high precision in chosing the components, so we will have a high speed in face recognition.[23-24]

Linear transformation have poor performance in the separation of data for which their classes do not inherently become linearly separated. Like PCA, non-linear locality preserving component analysis (NLPCA) is used to determine and reduce the correlation of the data. PCA method defines the linear correlation between the features , but NLPCA defines the linear and non-linear correlation between the features, regardless of nonlinear nature of the data. In NLPCA method, a neural network has been taught for determining the non-linear mapping.[20,23-25]

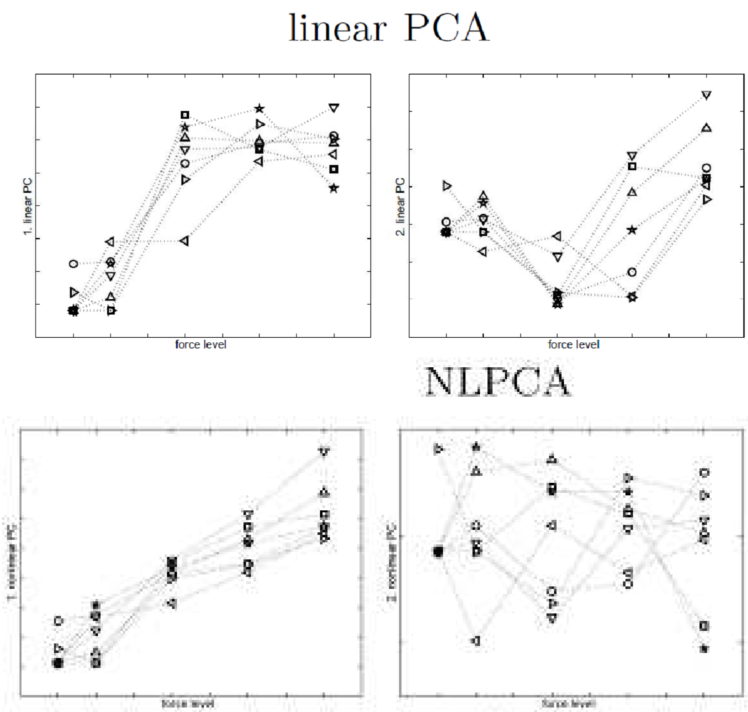

Figure 1. features of linear PCA and NLPCA in 5 level of force level

The analysis of original nonlinear components (NLPCA) is as a nonlinear generalization for the standard method of the analysis of original components (linear) (PCA). So far, many of these generalizations relied on one type of learning. Here, we suggest a algorithm for face recognision in which the extension from PCA to NLPCA is conducted through a deflationary type of learning [23-24].

When using any type of linear and non-linear analysis (PCA), it is important that their applications on the reduction of dimension and correct identification of the specific set of features based on the specific criteria be distinguishable. In the first set of applications, just subspace with the power of high description appears with emphasis on compressing and having no noise for the data.

It is no need for all features to be unique. The only necessary condition is that the mentioned subspace explains the mean square error (MSE) of information available in the data.

Implementation of a hierarchical algorithm PCA has two important properties i.e. the scalability and stability. The first one implies its ability to easily respond to increase in the amount of workload or shows the amount of system preparation for increasing in workload. For example, the scalability points out the system's ability to increase the overall performance while adding resources (e.g dimensions), distinguishable to the second one means that the i^{th} features from n features to have i solution for m features, in the way that (m ~ = n), then it reaches the equilibrium state and we are able to recover the original matrix, linear autoencode extracts and other features on the remaining error variance through training deflationary features. However, this method does not perform well enough in the non-linear one. The remaining variance can not be considered without regarding to non-linear vector mapping [15]. As the above figure shows, linear PCA loses some areas and selects a large interval for itself that will lead to an increase in the complexity.

3. CONVERTING LINEAR TO NON-LINEAR PCA

The expansion of linear autoencoder includes non-linear mappings which accures through adding non-linear hidden layers. In Figure 2, you see this strategy being the analysis network of the original non-linear components analysis (NLPCA) in its heart[18,19].

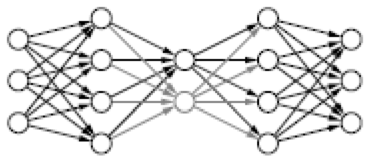

Figure 2. coding and decoding of nodes by mapping [3-4-2-4-3]

There are two general ways to introduce the extraction method of non-linear features in the feature space. First, the i^{th} featur is used to calculate the highest i^{th} variance like linear PCA. The second strategy is to search for the space of original data for the smallest mean squared error for the i^{th} feature. Finding a solution for the first strategy is much more difficult than that for the second one. Calculation of formula of MSE (mean squared error) is as following [21]:

$$E = \frac{1}{dN} \Sigma_n^N \Sigma_k^d (x_k^n - x^{\wedge}{}_k^n)^{2!} \qquad (5)$$

Where x and ˆx are the original and complimented data, respectively, n is the number of samples and d is dimention. All results can be generalized to any other dimension. E_1 and $E_{1,\,2}$ are the average errors when they are calculated only using one or two features. To performe h-NLPCA, we require the minimum value of E_1 and $E_{1,\,2}$ and we can create hierarchical method through minimizing the error:

$$E_H = E_1 + E_{1,2} \qquad (6)$$

However, we must obtain a balance between the value weights values for E_1 and $E_{1,\,2}$ with the α parameter:

$$E_H = \alpha E_1 + E_{1,2} \;; \qquad \alpha \in (0,\infty) \qquad (7)$$

Yet, for optimal selection of component α with the costs of computations, when α goes toward 1, it increases as it is shown in Figure 3.

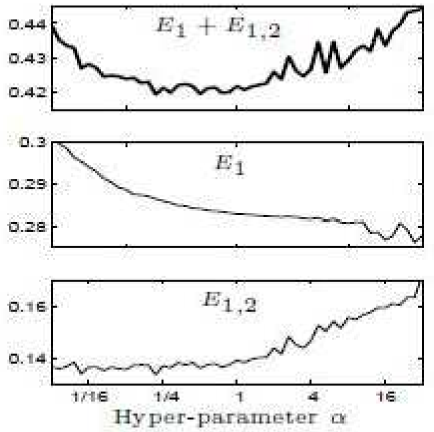

Figure 3. Dependent of errors by α component, for encoding ANN nodes with mapping [19-20-2-20-19] in 5 levels of s-NLPCA, while is in E_1 error α close to zero and in the E_1 error α close to ∞

To train the h-NLPCA, the error value of E1 and E1, 2 are separately calculated using a descending gradient, in each training iteration. This method in s-NLPCA is calculated by the network with one or two neurons in the features layer.

That respectively, the gradient of rEH is equal to the sum of rE1,2 + rE1 =rEH and the reduction of weight is calculated as the following sum:

$$E = E_H + \nu\, P_i\, w^2{}_i \qquad (8)$$

In most experiments, v = 0.001 is a good choice. In addition, the weight of non-linear layer is initialized to achieve more optimal results, so that the sigmoidal works nonlinearity in a linear system as non-linearity which is related to the starting point of h-NLPCA network with the simple solution of PCA.

4. RESULTS AND ANALYSIS

In this section, it is explained the results of research in two subchapter and given the comprehensive discussion.

4.1 Accuracy of Classification

The performed test includes the NLPCA features for the accuracy of classification. Here, we have a set of 50-dimensional samples that belongs to two classes with two levels of classification including image 1 and 2, and samples from A and B classe s[27].

Table 1.error rate for testing data by PCA and NLPCA

	Classification of '1' to '2'				
Features	**1**	**2**	**3**	**10**	**20**
PCA	40/6	30/3	30/3	-	-
NLPCA	48/5	38/3	-	-	-
	Classification of 'A' to 'B'				
Features	**1**	**2**	**3**	**10**	**20**
PCA	48/3	50/0	32/0	-	-
NLPCA	50/0	50/0	-	-	-

The non-linear quadratic equation is used to convert the non-linear data. According to Table 1, it is clear that each of which requires two features for each image and the other 3 rows are for the general specifications of the feature matrix. In both classification methods, NLPCA gives the best results for the smallest component. As it was said, the linear PCA has produced a weak classification [22, 26, 28].

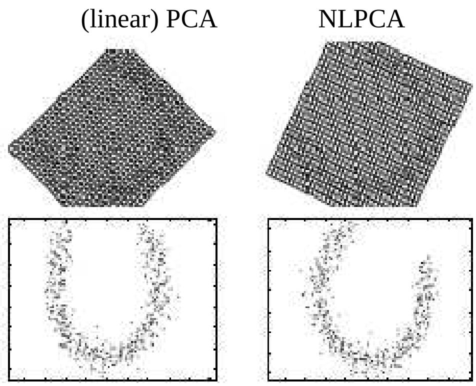

Figure 4.The result of linear PCA and NLPCA is noise of image

4.2 Testing EANF method for face classification

In this section, we compare the performance of the proposed EANF method for the FERET face Dataset. FERET face Dataset is the standard dataset in the evaluation of face recovery techniques including 14126 images out of 1199 persons. Also, our selected subset includes 1131 images out of 229 persons [14].

All images have been rotated and measured in two subsets, therefor; the center of eyes has been located on specific pixels and then changed into 32 * 32 pixels. The training stages includes 20 iteration and the simplest classifier is the nearest neighbor.

Figure 5.images of FERET dataset

4.2.1 Selecting EANF Parameter

In EANF, the number of neighbors close to the outer class K must be specified, to produce between-class scatter matrix, and any image should have different labels that this leads to high performance. To prove this technique, we study the effects of neighboring k on the base rate of recognition on the images in the FERET Dataset: In this experiment, 4 samples have been selected for each training class. it is clear in the figure, the curve of recognition rate can be obtained through changing k from 1 to 50 and it will be maximum when k = 1[13].

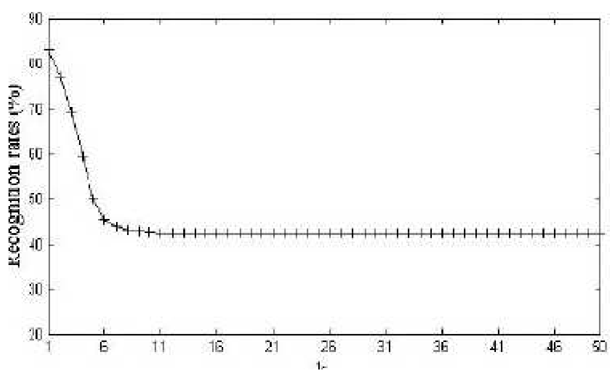

Figure 6. recognition rate and k parameter

In similar experiments with various training samples on FERE Database, when k = 1, EANF method always shows the most important recognition rate.

4.2.2 Tests on subsets of face recognition

In this section, we use the proposed EANF using five randomly subspace learning algorithm-LDA, PCA [1], NLDA, ELPDA (LDA empty space) and LPDA, with various number of images of each person for testing. In the following figure, the results of comparing the recognition rate in every 6 algorithms on FERET Dataset are shown. In table 2, the best recognition rate of six algorithms have been shown given that the number of dimensions are the same and the results considerably show the most optimal performance for EANF method.

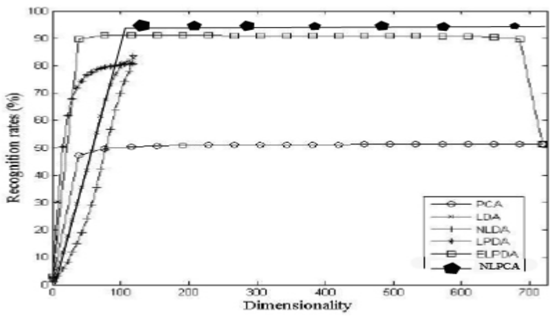

Figure 7. recognition rate of images of FERET dataset

Table 2. recognition rate of images of FERET dataset

Train	PCA	LDA
2	54/9±1/39(457)	61/8±1/95(109)
3	62/5±1/92(556)	68/9±2/44(228)
Train	**LPDA**	**ELPDA**
2	67/7±1/60(61)	69/0±1/62(289)
3	78/7±1/87(85)	82/5±1/56(186)
Train	**NLDA**	**NLPCAELPDA**
2	70/5±1/47(228)	67/7±1/60(61)
3	78/2±1/96(228)	89/5±1/56(186)

5. CONCOLUSION

In this paper, we presented a new method based on selecting a set of features of a set through the nonlinear PCA (NLPCA) (images with different dimensions) so that we will be able to improve the calculation speed of the features, decreasing dimension and high precision than the linear method, we can classify the available images with low error, higher speed and also any size without performing additional calculations than other algorithms. we showed, this algorithm will have better results than those by other methods.

REFERENCES

[1] P. Baldi and K. Hornik. Neural networks and principal component analysis:Learning from examples without local minima. Neural Networks, 2:53 – 58,1989.
[2] T. Hastie and W. Stuetzle. Principal curves. JASA, 84:502–516, 1989.

[3] M. Kramer. Nonlinear principal component analysis using auto-associative neural networks. AIChE Journal, 37(2):233–243, 1991.

[4] D. Mewett, K. Reynolds, and H. Nazeran. Principal components of recurrence quantification analysis of EMG. Proceedings of the 23rd Annual IEEE/EMBS Conference, Oct.25-28 2001.

[5] E. Oja. The nonlinear PCA learning rule in independent component analysis. Neurocomputing, 17(1):25 – 6,1997.

[6] B. Sch¨olkopf, A. Smola, and K.-R. M¨uller. Nonlinear component analysis as akernel eigenvalue problem. Neural Computation, 10:1299–1319, 1998.

[7] M. Scholz. Nonlinear PCA based on neural networks. Diploma Thesis, Dep. ofComputer Science, Humboldt-University Berlin, in preparation. In German.

[8] J. Stock and M. Stock. Quantitative stellar spectral classification. Revista Mexi-cana de Astronomia y Astrofisica, 34:143 – 156, 1999.

[9] V. Vapnik. The nature of statistical learning theory. Springer, New York, 1995.

[10] C. Webber Jr and J. Zbilut. Dynamical assessment of physiological systems andstates using recurrence plot strategies. J. of Appl. Physiology, 76:965 – 973, 1994.

[11] M. Turk and A. Pentland, "Eigenfaces for recognition," *J.Cognitive Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[12] X. He, S. Yan, Y. Hu, P. Niyogi, and Z. Hongjiang, "Facerecognition using Laplacianfaces," *IEEE Trans. Pattern Anal.Mach. Intell.*, vol. 27, no. 3, pp. 328–340, 2005.

[13] X. He, D. Cai, S. Yan, and H. Zhang, "Neighborhood preserving embedding," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2005, pp. 1208–1213.

[14] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, 2000.

[15] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A globalgeometric framework for nonlinear dimensionality reduction,"*Science*, vol. 290, pp. 2319–2323, 2000.

[16] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*,vol. 19, no. 7, pp. 711–720, 1997.

[17] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin, "Grapha embedding and extensions: A general framework for dimensionality reduction," *IEEE Trans. Pattern Anal. Mach Intell.*, vol. 29, no. 1, pp. 40–51, 2007.

[18] L. Yang, W. Gong, X. Gu, W. Li, and Y. Liu, "Bagging null space locality preserving discriminant classifiers for face recognition," *Pattern Recognit.*, vol. 42, no. 9, pp. 1853–1858,2009.

[19] L.-F. Chen, H.-Y. Mark Liao, M.-T. Ko, J.-C. Lin, and G.-J. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Pattern Recognit.*,vol. 33, pp. 1713–1726, 2000.

[20] D. Tao, X. Li, X. Wu, and S. J. Maybank, "Geometric mean for subspace selection," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 31, no. 2, pp. 260–274, 2009.

[21] Z. Li, D. Lin, and X. Tang, "Nonparametric iscriminant analysis for face recognition," *IEEE Trans. Pattern Anal.*

[22] *Mach. Intell.*, vol. 31, no. 4, pp. 755–761, 2009.

[23] T. Zhang, K. Huang, X. Li, J. Yang, and D. Tao, Discriminative orthogonal neighborhood-preserving projections for classification," *IEEE Trans. Syst. Man Cy. B*, vol. 40, no. 1,pp. 253–263, 2010.

[24] A. M. Martinez and R. Benavente, "The AR face database,"CVC Technical Report #24, Tech. Rep., 1998.

[25] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, 2000.

BIIOGRAPHIES OF AUTHORS

Ali Mahdavi Hormat was born in Hamedan, Iran. He received his degrees in Computer Engineering from Hamedan Branch, Islamic Azad University in September 2011, and is now studying student in artificial intelligence Engineering at Azad University of Qazvin.
Emails: ali.mahdavi.hormat@qiau.ac.ir, ali.mahdavi.hormat@gmail.com

Karim Faez was born in Semnan, Iran. He received his BSc. degree in Electrical Engineering from Tehran Polytechnic University as the first rank in June 1973, and his MSc. and Ph.D. degrees in Computer Science from University of California at Los Angeles (UCLA) in 1977 and 1980 respectively. Professor Faez was with Iran Telecommunication Research Center (1981-1983) before Joining Amirkabir University of Technology (Tehran Polytechnic) in Iran in March 1983, where he holds the rank of Professor in the Electrical Engineering Department.
He was the founder of the Computer Engineering Department of Amirkabir University in 1989 and he has served as the first chairman during April 1989-Sept. 1992.
Professor Faez was the chairman of planning committee for Computer Engineering and Computer Science of Ministry of Science, Research and Technology (during 1988-1996). His research interests are in Biometrics Recognition and authentication, Pattern Recognition, Image Processing, Neural Networks, Signal Processing, Farsi Handwritten Processing, Earthquake Signal Processing, Fault Tolerance System Design, Computer Networks, and Hardware Design. Dr. Faez coauthored a book in Logic Circuits published by Amirkabir University Press. He also coauthored a chapter in the book: *Recent Advances in Simulated Evolution and Learning*, Advances in Natural Computation, Vol. 2, Aug. 2004, World Scientific. He published about 300 articles in the above area. He is a member of IEEE, IEICE, and ACM, a member of Editorial Committee of Journal of Iranian Association of Electrical and Electronics Engineers,and International Journal of Communication Engineering.
Emails: kfaez@aut.ac.ir, kfaez@ieee.org, kfaez@m.ieice.org.

Zeynab Shokoohi was born in Shiraz, Iran. He received his Technician degree in Computer Software from Shiraz Branch, Higher Education Center for Girls Kazeroun University in September 2004, and his BSc. degrees in Computer Engineering from Ahvaz Branch, Shahid Chamran University in September 2007, and is now studying student in artificial intelligence Engineering at Azad University of Qazvin.
Emails: z.shokoohi @qiau.ac.ir

Mohammad Zaher Karimi was born in West Azerbaijan, Iran. He received his Technician degree in Computer Software from Kermanshah Branch, Technical College's number one Kermanshah in September 2005, and his BSc. degrees in Computer Engineering from Tabriz Branch, Daneshvaran University in September 2008, and is now studying student in Mechatronics Engineering at Azad University of Qazvin.
Emails: zaherkarimi78@gmail.com, zaherkarimi78@qiau.ac.ir