World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:3, No:5, 2009

Evolutionary Eigenspace Learning using CCIPCA and IPCA for Face Recognition

Ghazy M.R. Assassa, Mona F. M. Mursi, and Hatim A. Aboalsamh

Abstract—Traditional principal components analysis (PCA) techniques for face recognition are based on batch-mode training using a pre-available image set. Real world applications require that the training set be dynamic of evolving nature where within the framework of continuous learning, new training images are continuously added to the original set; this would trigger a costly continuous re-computation of the eigen space representation via repeating an entire batch-based training that includes the old and new images. Incremental PCA methods allow adding new images and updating the PCA representation. In this paper, two incremental PCA approaches, CCIPCA and IPCA, are examined and compared. Besides, different learning and testing strategies are proposed and applied to the two algorithms. The results suggest that batch PCA is inferior to both incremental approaches, and that all CCIPCAs are practically equivalent.

Keywords—Candid covariance-free incremental principal components analysis (CCIPCA), face recognition, incremental principal components analysis (IPCA).

I. Introduction

FACE recognition is one-to-many process that compares an input test image against all face templates used in training; the output is the identity of the input test image. The problem of human face recognition is a complex and highly challenging one with spatial and temporal variations, e.g., illumination, pose orientation, expression, aging, head size, make-up, image obscuring (eye glass effect), disguise, and face background [1, 2]. The problem of automatic human face recognition can be stated as follows [1]:given an image of a human face, compare it with pre-stored models of a set of face images labeled with the person's identity (the training set) and report the matching result.

When training is carried out using appearance based modeling such as the Principal components analysis, all training dataset images must be available before the training process can be initialized; this is termed as batch training. On the contrary, an incremental approach allows adding new images and updating the PCA representation accordingly; thus offering the great benefit of discarding the new added images after model update.

Ghazy M.R. Assassa is with the Center of Excellence in Information Assurance (CoEIA) and the Department of Computer Science, King Saud University, Riyadh 11543, Saudi Arabia, (corresponding author phone: +966-1-4676595; fax: +966-1-4675630; e-mail: gassassa@coeia.edu.sa), on leave from the Faculty of Engineering at Shoubra, Benha University, Egypt (ghazyassassa@gmail.com).

Mona F.M. Mursi is with the Center of Excellence in Information Assurance (CoEIA) and the Department of Information Technology, King Saud University, Riyadh 11543, Saudi Arabia (e-mail: monmursi@coeia.edu.sa).

Hatim A. Aboalsamh is with the Center of Excellence in Information Assurance (CoEIA) and the Department of Computer Science, King Saud University, Riyadh 11543, Saudi Arabia (e-mail: hatim@ksu.edu.sa).

This paper is organized as follows: section II presents related work on Principal components analysis and section III presents the two incremental principal components approaches used in the current study; section IV discusses the proposed training and relearning strategies; numerical experiments are presented and discussed in section V. Finally, section VI discusses some concluding remarks.

II. Related Work —The PCA Approach

Face recognition approaches may be categorized under two general approaches: appearance-based (holistic) and feature-based (structural). Both approaches are designed to use previous knowledge obtained from feature extraction to recognize human faces [1, 2, 3, 4]. The most popular appearance-based holistic approaches includes: (1) the eigenfaces, known also as the Principal Components (PC) Analysis (PCA) and also as Kahunen-Loeve transformation (KL) [5, 6, 7], (2) the Fisherfaces known as the linear Discriminant analysis (LDA) [8] , and (3) Independent Component Analysis (ICA) [9]. PCA is unsupervised technique for dimensionality reduction; it searches for directions in the dataset that have the largest variance and define a projection matrix to project the data onto it. This leads to a lower dimensional presentation of the data, and therefore removes some of the noisy directions. Batch mode determination of principal axes for data with varying reliability and missing data was studied in [10, 11, 12, 13].

PCA eigensapce model may be determined using eigenvalue decomposition (EVD) (or singular value decomposition (SVD)) of the covariance matrix [16]. For a training set of images with covariance matrix C(n,n), ,the full eigen problem is to find the eigenvectors matrix U(n,n) and associated eigenvalues matrix λ(n,n) where C(n,n) U(n,n) = U(n,n) λ(n,n). For the sake of computational efficiency, some eigenvalues, according to a selected criterion, e.g., the smallest, are discarded and the eginvalue problem is reduced to solve the approximation C(n,n) U(n,k) = U(n,k) λ(k,k). For a set S of N training images $S=\{s_1, s_2, ,,,. s_N\}$, with each image $s_i \in \mathbf{R}^n$ where n is the dimension of the image space (n equals the number of pixels), the eigen space model yields three outputs: (1) a mean of the training set $\mu \in \mathbf{R}^n$, (2) a set of eigenvectors $\mathbf{U} \in \mathbf{R}^{nxk}$, that is a matrix of k n-dimensional principal axes, and (3) the eigenvalues $\lambda \in \mathbf{R}^{kxk}$ representing the spread of the training set over each eigenvector.

An eigensapce model Ω may be defined for a set $S=\{s_1, s_2, ..., s_N\}$ of N training images (observations):

$$\Omega = \Omega\ (\mu, \mathbf{U}, \lambda, S) \quad (1)$$

The eigensapce model can thus be viewed as a k-dimensional hyper-ellipse in the n-dimensional image space, where k, is the dimension of the reduced space in which the image is represented to a certain degree of accuracy depending on the retained value of k <= n; the higher value of k the more accurate is the presentation. The hyper-ellipse is viewed as centered at the mean of images used in training, its axes are the eigenvectors (columns of U(n,k), and the lengths of its axes are the square roots of the eigenvalues of the diagonal elements of the matrix λ(k,k).

PCA approach relies on modeling static datasets where training is performed in a batch mode on image set that is supposed to be available in advance of the training process. In PCA, a 2-dimensional face image with size of p rows and q columns can be viewed as a one dimensional vector of dimension pxq. The key idea of the PCA method is to find the vectors that best account for the distribution of face images within the entire pxq image space. These vectors define the subspace of face images. Because these vectors are the eigenvectors of the covariance matrix corresponding to the original face images, and because they are face-like in appearance, they are called eigefaces. A brief review of PCA is given hereafter [5, 14].

For a set $S=\{s_1, s_2, ..., s_N\}$ of N training images, the average vector image μ and the deviation matrix Φ of each image from the average image μ is given by:

$$\Phi = S - \mu \quad (2)$$

The covariance matrix C is given by:

$$C = (1/N) \sum_{i=1}^{N} \Phi_i \Phi^T_i = A\, A^T \quad (3)$$

where $A = [\Phi_1\ \Phi_2\ \Phi_3\ \dots \Phi_N]$, and Φ^T is the transpose matrix of Φ.

The eigenvectors of the product ($L = A^T A$) are obtained as:

$$L\, v_i = A^T A\, v_i = \lambda_i v_i, \quad L_{ij} = (\Phi_i)^T \Phi_j \quad (4)$$

Premultiplying both sides by matrix A,

$$A\, A^T A\, v_i = \lambda_i\, A\, v_i \quad (5)$$

where
vi and λi are respectively the N eigenvectors and N eigenvalues of matrix L, and
(A vi) and λi are respectively the eigenvectors and eigenvalues of the covariance matrix $C = A\, A^T$.

The looked for eigenvectors $u_i = A\, v_i$ of matrix C are the eigenfaces that are obtained by projection of the deviation matrix Φ on the eigenvectors v_i of L:

$$u_i = \sum_{j=1}^{N} v_{ij}\ \Phi_j \quad , i=1, 2, \dots, N \quad (6)$$

III. Incremental PCA

A shortcoming of the training process for PCA is that the entire training dataset images must be available beforehand in order to start the training process. This defect is elegantly handled by Incremental PCA (IPCA) methods which allow adding new images and updating the PCA representation accordingly; thus offering the great benefit of dispensing with the recently added images after model update. Incremental PCA methods have been studied by several researchers [15,16, 17, 18].

The incremental methods proposed in [19, 20] are tailored for temporally weighted learning allowing newer images to have a larger influence on the estimation of the current subspace than the older ones. Ref [21] studied incremental learning for online face recognition and proposed new approach to face recognition in which not only a classifier but also a feature space of input variables is learned incrementally to adapt to incoming training samples; as suggested, a benefit of this type of incremental learning is that the search for useful features and the learning of an optimal decision boundary are carried out in an online fashion. Incremental PCA algorithms that compute the principal components without computing the covariance matrix [22] are presented in [23, 24, 25].

A. The Candid Covariance-free IPCA Algorithm: CCIPCA

The candid covariance-free IPCA (CCIPCA) was introduced in [23] to compute the principal components of a sequence of samples incrementally without estimating the covariance matrix (thus covariance-free).

The algorithm keeps the scale of observations and computes the mean of observations incrementally. The method is suggested for real-time applications, and thus it does not allow iterations. It converges very fast for high dimensional image vectors. The CCIPCA algorithm generates "observations" in a complementary space for the computation of the higher order principal components. If we consider a sample vectors that are acquired sequentially, e.g., s(1), s(2), . . ., possibly infinite, the first k dominant principal components (PCs) u1(n), u2(n), . . . , uk(n) are obtained as follows [23].

For $n = 1, 2, \dots$, do the followings steps.
1) $s_1(n) = s(n)$.

2) For $i = 1, 2, \dots, \min(k, n)$, do:
a) if $i = n$, initialize the ith PC as $u_i(n) = s_i(n)$;
b) otherwise compute:

$$u_i(n) = (1/n)(n-1-l)\, u_i(n-1) + (1/n)(1+l)\, s_i(n)\, s_i^T(n)\, [u_i(n-1) / \| u_i(n-1)\|] \quad (5)$$

and

$$s_{i+1}(n) = s_i(n) - s_i^T(n)\,[u_i(n) / \| u_i(n)\|]\,[u_i(n) / \| u_i(n)\|] \quad (7)$$

where *l* is the amnesic parameter.
After normalization. the final eigen vector and eigen value are respectively given by:

$$u_i = u_i(n) / \| u_i(n)\| \quad (8)$$

and

$$\lambda_i = \| u_i(n)\| \quad (9)$$

It should be noticed that both batch PCA and CCIPCA are dimensionality reduction techniques searching for directions in the dataset that have the largest variance and define a projection matrix to project the data onto it. Both techniques avoid the costly computation of the covariance matrix; PCA computes the matrix L as an intermediate matrix leading to the eigen vectors, whereas CCIPCA is covariance-free approach.

B. The Incremental PC Subspace Learning Algorithm: IPCA

Ref [18] presented a weighted and robust incremental method for subspace learning based on incremental method. The approach sequentially updates the principal subspace represented by the egienspace model $\Omega = \Omega$ (μ, **U,** λ, S). In this method, the staring eigen space state, can be obtained by two approaches:

(1) A batch PCA may be applied on an initial set of images S^0 to obtain the average images μ^0, the eigenvectors $\mathbf{U^0}$, and the weight coefficients $\mathbf{A^0}$;

(2) The first training image vector $\mathbf{u}_1$ is used to set the initial eigen space Ω^0:

$$\boldsymbol{\mu}^0 = \mathbf{u}_1, \ \mathbf{U^0 = 0}, \text{ and } \mathbf{A^0 = 0} \quad (10)$$

For the latter case, the algorithm is considered as completely incremental from the start.

Fig. 1 illustrates a macro flow chart for the incremental PC Subspace Learning algorithm of [18]. For the sake of clarification, the following discussion is worth noting for any two successive eigenspace computations. For an image space (pixel space) of dimension n; all images $s_i \in \mathbf{R}^n$, the starting eigenspace is of dimension k, $\Omega(k)$: the mean $\mu \in \mathbf{R}^n$, the eigenvectors $\mathbf{U} \in \mathbf{R}^{nxk}$ and the eigenvalues $\lambda \in \mathbf{R}^{kxk}$. At the end of the eigenspace update, the dimension of the eigenspace is k+1, $\Omega(k+1)$, that is, $\mathbf{U} \in \mathbf{R}^{nxk+1}$ and $\lambda \in \mathbf{R}^{k+1xk+1}$.

This means that the dimension of the eigenspace is incremented by one after each update. This would lead to continuous increment the dimension of the eigenspace by one after each update. To overcome this problem of continuous growth in the eigenspace dimension, [18] and [2] suggest to keep the dimension of the eigenspace at its staring value (k), by discarding after each update the least significant principle vector

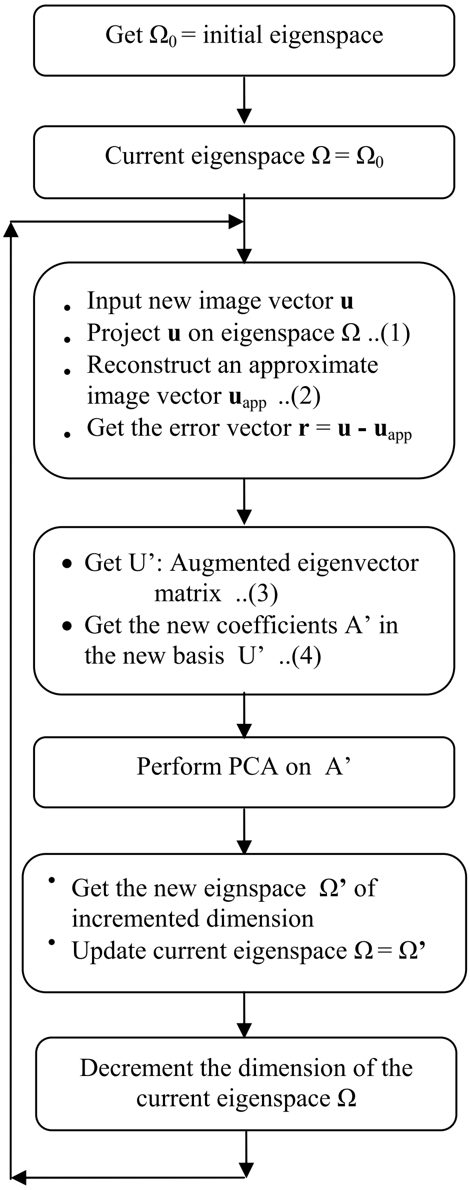

Fig. 1 Macro flow chart for the incremental PC subspace learning algorithm

Fig. 2 shows an example for the evolution in eigensapce dimension for the case of completely IPCA incremental learning algorithm. The figure illustrates the incremental continuous growth of the eigenspace up to a selected value of k=25 after which the eigenspace dimension is alternatively incremented and then decremented by unit value along the horizontal axis of incremental temporal steps; this leads to a constant eigenspace dimension of k=25.

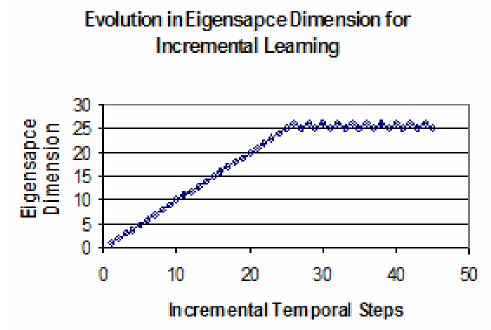

Fig. 2 Evolution in eigensapce dimension for completely incremental learning – IPCA algorithm

IV. CCIPCA and IPCA Training and Testing Strategies

ORL (Olivetti Research Labs) face database was used for CCIPCA and IPCA algorithms learning and testing processes. ORL includes 400 images (of size 112 x 92) for 40 subjects with 10 images per subject. Selection of the training and test images was done by two strategies:

Strategy A: all 400 face images were used in training, while testing was performed using100 test images that were randomly selected from within those used in training. For this case, the correct recognition rate CRR is given by:

CRR = (number of correctly recognized faces) / (number of randomly selected images) (11)

Strategy B: For each subject we used the first 80% images for training (first 8 images per subject) and the remaining 20 % for testing (the last two images per subject, the last 9th and 10th images).

Strategies A and B were applied for learning and testing both CCIPCA and IPCA algorithms using ORL face database.

CCIPCA Training and Testing Strategies

The effect of the number of increments and size of the eigen vectors on the correct recognition rate (CRR) were studied for various CCIPCA training strategies. The eigen space dimension (# of eigenfaces) was varied as 15, 20, 25, 30, 35, and 40. Five training and relearning strategies were investigated:

(1) Batch PCA: The entire training set of images must be available in advance of training, there is no relearning process, contrary to the next four IPCAs;
(2) CCIPCA1: Incremental PCA using CCIPCA algorithm;
(3) CCIPCA2: Image level relearning using CCIPCA, relearning after each added image (learning the same image twice, the second learning being immediately after the first);
(4) CCIPCA3: Increment level relearning using CCIPCA, relearning after each added increment, that is learning the same increment twice the second being immediately after the first;
(5) CCIPCA4: Set level relearning using CCIPCA, relearning after adding all increments, which is learning the entire set of images twice, the second being immediately after the first.

For strategy B and ORL, the total number of training images is 8 x 40 = 320 images. Eight learning increments (each including 40 images) are defined; the first two increments are shown hereafter for illustration:

- Increment # 1 (40 images): images # 1, 11, 21, …, 371, 381, 391;
- Increment # 2 (40 images): images # 2, 12, 22, …, 372, 382, 392.

Strategy B assumes 8 training images and 2 test images per subject, thus there are two test sets, the last 9^{th} and 10^{th} images per subject totaling 40+40 = 80 images:

- Test set 1: the 9^{th} image per subject for 40 subjects: images # 9, 19, 29, 39, … , 379, 389, 399;
- Test set 2: the 10^{th} image per subject for 40 subjects: images # 10, 20, 30, 40, … , , 380, 390, 400.

V. Numerical Experiments and Discussion

Numerical experiments were carried out according to the above introduced learning and testing strategies. The features vector length (FVL) was varied as 15, 20, 25, 30, 35, and 40; for each FVL the correct recognition rate CRR was determined using batch and the four strategies CCIPCA1-4 [26]; batch training is the well known PCA, for which the entire image dataset must be available before the start of training. From the obtained CRR results, the average CRR was determined. The average CCR for increment # i is defined as:

$$CCR(i)_{avg} = \Sigma\, CCR(i, CCIPCA(,j)) / 4 \qquad (12)$$

where i is the increment #, i = 1, 2, 3…, 8, and CCIPCA(j) is the training strategy, j = 1, 2, 3, 4.

Fig. 3 illustrates the variation of the average CRR for batch, CCIPCA1, CCIPCA2, CCIPCA3, CCIPCA4; more details can be found in [26]. The figure confirms the general findings that batch PCA training is inferior to the four CCIPCA1-4 training strategies and that CCIPCA3 learning yields slightly better CRR. This statement should be verified against other face databases.

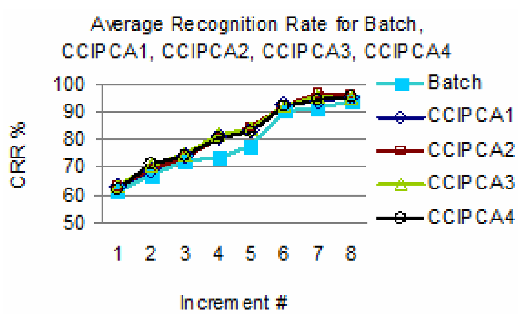

Fig. 3 Variation of the average correct recognition rate with the increment number for batch, CCIPCA1, CCIPCA2, CCIPCA3, CCIPCA4 approaches

Using the above proposed testing strategy B, the variation of the correct recognition rate with the increment number is shown in Fig. 4 for the two incremental algorithms CCIPCA1 and IPCA for the case of feature vector length of 40. For all considered increments from 1 to 8 for testing strategy B on ORL face database, the figure suggests that the IPCA algorithm yields higher correct recognition rates than the CCIPCA1 algorithm.

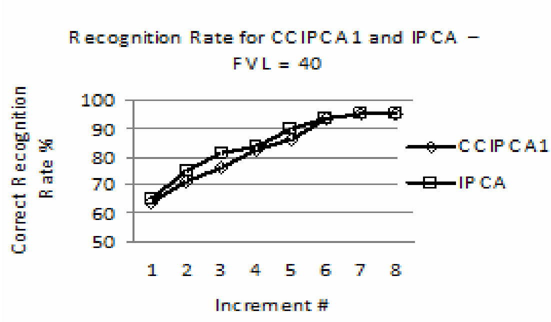

Fig. 4 Variation of the correct recognition rate with the increment # for the two incremental algorithms CCIPCA1 and IPCA for a feature vector length (FVL) = 40 and using testing strategy B on ORL face database

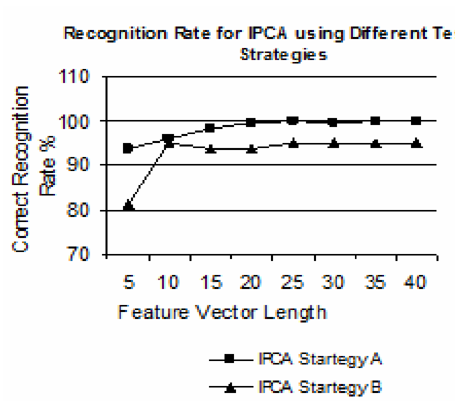

Fig. 5 Variation of the correct recognition rate with the length of feature vector for the testing strategies A and B using the incremental IPCA algorithm on ORL face database

Numerical experiments were also carried out to compare the two testing strategies A and B on the incremental algorithm IPCA. Fig. 5 displays the variation of the correct recognition rate with the length of feature vector for both testing strategies A and B using ORL face database. The figure suggests that in general, the correct recognition rate for testing strategy A outperforms that for strategy B roughly by 5%.

VI. Conclusion

This paper considers the batch and incremental PCA appearance-based (holistic) approach for face recognition. The paper examines two incremental approaches: the candid covariance-free IPCA "CCIPCA" algorithm, and the IPCA algorithm. The paper proposed various training and testing strategies and applied them to ORL face database. The effect of the number of increments and size of the eigen vectors on the correct rate of recognition are numerically investigated. Analysis of the obtained results suggests that batch PCA is inferior to all CCIPCAs training strategies, IPCA1, IPCA2, IPCA3, and IPCA4. Comparison of training strategies A a nd B suggests that in general, the correct recognition rate for testing strategy A outperforms that for strategy B roughly by 5%. Future work will include study the processing time for the various training strategies. Other popular batch appearance-based holistic approaches will also be considered for incremental study.

Acknowledgement

The authors would like to thank engineer Yasser Ali for his preliminary work towards a face recognition system.

References

[1] J. Haddadnia, K. Faez, and A. Majid, "N-Feature Neural Network Human Face Recognition," Image Vision Computing 22, pp. 1071-1082, 2004.
[2] B.L. Zahng, H. Zhang, and S.S. Ge, "Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory," IEEE Trans. Neural Networks, vol. 15, no. 1, pp. 166-177, Jan. 2004.
[3] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D Face Recognition: A Survey," Pattern Recognition Letters, 28, pp. 1885-1906, 2007.
[4] W. Zhao and R. Chellappa, ed, Face Processing: Advanced Modeling and Methods, Academic Press, Elsevier Inc, 2006.
[5] A. Eleyan and H. Demirel, "PCA and LDA based Neural Networks for Human Face Recognition," Face Recognition, edited by Kresimir Delac and Mislav Grgic, p. 558, I-Tech, Vienna, Austria, June 2007.
[6] K-C. Kwak and W. Pedryez, "Face Recognition: A Study in Information Fusion using Fuzzy Intrgral," Pattern recognition Letters, vol. 26, pp. 719-733, 2005.
[7] M. Turk, and A. Pentland, "Ejgenfaces for Recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.
[8] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: RecognitionUsing Class Specific Linear Projection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.
[9] M.S. Bartlett, J.R. Movellan, and T.J. Sejnowski, "Face Recognition by Independent Components Analysis," IEEE trans. Neural Networks, vol. 13, no. 6, pp.1450-1464, 2002.
[10] H. Shum, K. Ikeuchi, and R. Reddy, "Principal component analysis with missing data and its application to polyhedral object modeling," PAMI, 17(9):854–867, 1995.
[11] H. Sidenbladh, F. de la Torre, and M. J. Black, "A framework for modeling the appearance of 3D articulated figures," in AFGR00, pages 368–375, 2000.
[12] F. De la Torre and M. J. Black, "Robust principal component analysis for computer vision," in ICCV'01, I: 362–369, 2001.
[13] K. Gabriel and S. Zamir, "Lower rank approximation of matrices by least squares with any choice of weights," Technometrics, 21(21):489–498, 1979.
[14] H. A. Aboalsamh, "Human Face Recognition using Eigen and Fisher Faces," Egyptian Computer Science Journal, vol. 31, no. 1, (accepted), January 2009.
[15] S. Chandrasekaran, B. S. Manjunath, Y. F.Wang, J.Winkeler, and H. Zhang, "An eigenspace update algorithm for image analysis," Graphical Models and Image Processing, 59(5):321–332, September 1997.
[16] P. Hall, D. Marshall, and R. Martin, "Incremental eigenanalysis for classification," in British Machine Vision Conference, volume 1, pages 286–295, September 1998.
[17] P. Hall, D. Marshall, and R. Martin, "Merging and splitting eigenspace models," PAMI, 22(9):1042–1048, 2000.

[18] D. Skoˇcaj and A. Leonardis, “Weighted and Robust Incremental Method for Subspace Learning,” Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV’03), 2003.
[19] A. Levy and M. Lindenbaum, “Sequential karhunen-loeve basis extraction and its application to images,” IEEE Trans. On Image Processing, 9:1371–1374, June 2000.
[20] X. Liu and T. Chen, “Shot boundary detection using temporal statistics modeling,” in ICASSP 2002, May 2002.
[21] S. Ozawa, S.L Toh, S. Abe, Shaoning Pang, and N. Kasabov, “Incremental learning for online face recognition,” Proceedings of the IEEE International Joint Conference on Neural Networks, IJCNN, 2005.
[22] H. Zhao, P.C. Yuen, and J.T. Kwok, “A Novel Incremental Principal Component Analysis and Its Application for Face Recognition,” IEEE Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics, vol. 36, no. 4, pp. 873-886 August 2006.
[23] J. Weng, Y. Zhang, and W. S. Hwang, “Candid covariance-free incremental principal component analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 8, pp. 1034–1040, Aug. 2003.
[24] T. D. Sanger, “Optimal unsupervised learning in a single-layer linear feed forward neural network,” Neural Netw., vol. 2, pp. 459–473, 1989.
[25] E. Oja and J. Karhunen, “On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix,” J. Math. Anal. Appl., vol. 106, no. 1, pp. 69–84, 1985.
[26] H. A. Aboalsamh, H. Mathkour, G. M.R. Assassa. And M F. Mursi, “Face Recognition using Incremental Principal Components Analysis,” International Conference on Computing in Engineering Science and Information, ICC 2009, University of California, Fullerton, April 2009, (accepted).