Time Efficient Face Recognition Using Stable Gram-Schmidt Orthonormalization

I. Sajid, M.M. Ahmed and I. Taj
Department of Electronic Engineering
Mohammad Ali Jinnah University (MAJU), Jinnah Avenue, Islamabad
{imtiazsajid, mansoor, imtiaztaj}@jinnah.edu.pk

Abstract

Most commonly used face recognition algorithms are based on extraction of global features using eigenvalue decomposition of some relational matrix of image intensity values. Real time face recognition applications require a computationally efficient algorithm for eigenvalues generation. Fast principal component analysis (FPCA) is an algorithm for efficient generation of eigenvalues which improves the computational efficiency to $O(n^2)$ as compared to normal decomposition method which gives the solution in $O(n^3)$ time. In FPCA however, non-convergence state can be resulted for high resolution images because in this case the number of Grams-Schmidt (GS) iterations for orthonormalization convergence may exceed the maximum limit. To overcome this problem we present a modified FPCA algorithm to generate eigenvalues for images including those at high resolution. An overall efficient face recognition scheme has also been proposed using the generated eigenvalues, which can work satisfactorily under varying image resolutions. The validity of the proposed system has been checked by varying the feature vectors and the training sets. The developed technique provides an efficient and a low error rate solution for high speed image recognition systems.

1. Introduction

Biometric based human verification and identification have found many applications in recent years ranging from security of personal money transactions to large scale security applications such as airport security. Centuries ago Babylonian kings used clay finger prints for authenticity [1]. Various biometrics have since been employed for personal verification and identification such as fingerprints, palmprints, iris patterns, face features, voice and human gait etc. Human face is one of the most famous and natural biometric, and face recognition systems are usually preferred over the other biometric systems due to their wider acceptability and ease of operation. However, contrary to the general perception, face recognition is not a straightforward problem and is vulnerable to many variations such as intensity variations, pose change, aging, occlusions etc. A variety of techniques have been proposed in the literature to model a human face in such a way that the final matching is immune to these variations [2-6]. These techniques can generally be classified into local feature extraction techniques and global features extraction techniques [3]. Local feature extraction techniques usually calculate auto-correlation matrix within each image in addition to the co-variance matrix between images [4], and they are computationally expensive. In global feature extraction techniques only co-variance is computed, and thus they are relatively more efficient.

In 1991, Turk and Pentland [2] proposed their famous approach to extract the most meaningful face details for matching, using principal component analysis (PCA). In their approach they mapped the face vector onto a smaller number of basis vectors called eigen faces corresponding the highest eigenvalues of the covariance matrix. They also presented a scheme to efficiently compute the eigenvalues by considering a matrix which is sufficiently smaller than the original covariance matrix. It is obvious that in their approach the timing efficiency of the recognition system directly depends on the number of highest eigenvalues but they did not suggest appropriate number of eigenvalues for a given face recognition application. Real time face recognition applications require an efficient system [3], thus appropriate selection of eigenvalues that minimizes the computational requirements and maximizes the system performance within constraints is vital.

The PCA which was followed by Turk *et al*. [2] is one of the holistic algorithms [3-5] which is based on global feature extraction. In PCA these features are extracted by computing the eigenvalues normally by using decomposition (EVD) methods. EVD method requires tridiagonalization by taking co-variance symmetric matrix and then decomposition of tridiagonal matrix is accomplished. Therefore, operational time complexity of PCA, because of tridiagonalization and decomposition, becomes $O(n^3)$ [7, 8]. Decomposition method cannot be stopped for desired number of leading eigenvalues. However usually 3-4 leading eigenvalues occupy more than 90% weight of the total variance [9,10], whereas recognition accuracy depends on variance in the eigenspace.

To improve the efficiency of PCA based systems, the concept of FPCA was introduced [7]. FPCA used Gram-Schmidt orthonormalization to calculate leading eigenvalues (LEVs) but with lesser time complexity i.e. $O(n^2)$ [7]. However, mean square error (MSE) in FPCA is higher [7] than PCA, for images of size greater than 4000 pixels. Moreover, face recognition using FPCA decreases the decision time of a system but could lead to a non-convergence state, especially when high resolution images are used [11]. Whereas images may vary from as few as 256 pixels to 452736 pixels [2,6,11], according to the resolution used. In other words, FPCA based system is affected by non-convergence state of the algorithm and high MSE.

For a deterministic state, the system should identify the parameters or characteristics which control the convergence condition. We recently established that a definite convergence can be attained, for high resolution images, with modified Grams-Schmidt (GS) process used in FPCA [11]. The technique, called adaptive FPCA (AFPCA) has the capability to adjust minimum possible tolerance value according to image resolution and make the real-time face recognition deterministic and time efficient.

Accuracy, decision and learning time of AFPCA increases by increasing the number of principal components (*h),* images in the training set (TS) and image resolutions in TS respectively whereas a face recognition system tries to achieve maximum possible accuracy within minimum given time. Thus appropriate selection of *h*, TS and images resolution provides efficient face recognition than conventional PCA based system. Selection of *h* is not a straight forward decision [12]. The proposed technique also minimizes error rate in face recognition, but at the expense of time. This paper, focuses only on the time efficient aspect of a face recognition system based on AFPCA.

2. Limitations of FPCA

It is observed that both PCA and FPCA generates similar MSE, if the chosen value of ε = 0.01 which has been suggested in [7], under the following conditions:

i. At least six samples are used in TS.

ii. Features vector *(d)* of orthogonal space should be equal to 100.

iii. The value of *h* should be 10.

iv. Image size should be less than 4000 pixels.

The maximum probability of convergence of FPCA is observed for multi pose images having resolution 118 x 122, *d*=100 and *h*=10. Variation of six faces with different facial expression is treated as multi pose. The concept of multi pose is used in different experiments as shown in Table-1. The number of iterations depends heavily on the image resolution. It has been demonstrated that the technique is valid even for the images which have been acquired with various lighting effects. Non-convergence in FPCA is addressed when iterations in GS process exceed 1000. These non-convergence conditions are highlighted with shaded cells in Table-1.

The convergence criterion in FPCA is based on the rule that dot product of orthogonal vectors, which are generated by GS process, should be unity [13]. This product value is calculated with current Φ_p^{+T} and previous Φ_p base vectors and its difference from unity is computed as

$$\left| \Phi_p^{+T} \cdot \Phi_p - 1 \right| \qquad (1)$$

Table 1. Limitation of FPCA when ε = 0.01, varying dimensionalities whereas shaded cells show non-convergence conditions.

Image Size	*h*	Faces / Objec	Single / Multi pose	Epsilon	Max. Iterations	Training Images
66 x 59	10	Faces	Multi	0.01	10	all
66 x 59	10	Faces	Multi	0.01	> 1000	10
66 x 59	10	Faces	Single	0.01	> 1000	10,20
66 x 59	10	Faces	Three	0.01	> 1000	10
66 x 59	10	Object	Single	0.01	11	10,20
118x 122	10	Faces	Muti	0.01	10	All
118x 122	10	Faces	Single	0.01	> 1000	10,20
118x 122	10	Faces	Three	0.01	> 1000	10,20
118x 122	10	Object	Single	0.01	> 1000	all

For convergence condition, the value of Equation (1) should be less than the chosen value ε. It is observed that when variation in poses are less than six in the TS having less than 100 images then Equation (1) approaches slowly to ε. This slow convergence is because of the stochastic and iterative nature of GS process. In case of non-convergence, GS multipies the co-variance with the modified vector Φ_p and then check the convergence condition. Whereas the initial values of Φ_p depend on the system random generation process. GS process tries to orthogonalize the current Φ_p^{+T} and previous Φ_p

vectors until it is less than ε which defines tolerance of the system. Ideally this tolerance should be zero, but practically the value of ε varies according to the randomness of Φ_p in FPCA and rounding effect in all sorts of other iterative processes [14].

It is suggested that an adaptive value of ε will facilitate a definite convergence in a GS process. For this purpose, a modification in the existing algorithm is proposed. The MSE of the modified algorithm (AFPCA) is studied. AFPCA ensures avoidance infinite looping by adjusting value of ε, or by deciding the maximum number of iterations.

3. Modification in FPCA algorithm

Transformation is of dimension *d* x *h*

$$\begin{aligned}
&\Phi : a \rightarrow b \\
&a \textit{ should be zero mean} \\
&\textit{where } a \in R^d \ (d-\dim\textit{ensional space}) \\
&\& \quad b \in R^h \ (\textit{reduced } h-\dim\textit{ensional space})
\end{aligned}$$

step i. *Select h, the number of leading principal components and compute co-variance* Σ_x *set* $p \leftarrow 1$

step ii. *Eigenvector* Φ_p *of size d x 1 is initialized randomly and initialize matrix vv of size d x h and set converge ← 1.*

   a. *Set ε ← 0.01 and initialize iteration counter (IC) with zero for post-condition*

step iii. *Assign* Φ_p *as* $\Phi_p \leftarrow \Sigma_x \Phi_p$

step iv. *Loop the Gram- Schmidt orthogonalization process*

step v. $\Phi_p^+ \leftarrow \Phi_p - \sum_{j=1}^{p-1}\left(\Phi_p^T.\Phi_j\right).\Phi_j$

step vi. *Normalize* Φ_p^+ *by dividing it by its norm* $\Phi_p^+ \leftarrow \dfrac{\Phi_p^+}{\left\|\Phi_p^+\right\|}$

step vii. *Compute the convergence condition as* $\left|\Phi_p^{+T}.\Phi_p - 1\right| < \varepsilon$ *and make new* Φ_p^+ *as old* Φ_p

step viii. *Update iteration counter and set ε incremental value*

step ix. *Increment ε value as per step vi until the convergence is occurred*

   a. *Record ε and iteration values for time complexity of algorithm*
   b. *Increase ε if IC has reached to threshold*

step x. $vv \leftarrow \Phi_p$ *where vv is eigenvector space*

step xi. *Increment the p and go to step ii until p equal to h*

step xii. *Extract eigenvalue on the diagonal of* $dd^T * \Sigma_x * dd$

3.1. AFPCA Classifier Algorithm

AFPCA uses Grams-Schmidt orthonormalization process for the computation of eigenvalues and eigenvectors and then generates eigenfaces. Weight is assigned to each eigenface according to its corresponding eigenvalue [2]. For testing, each input face is converted into weight value by applying the same procedure as that of eigenfaces. The difference of given face magnitude value from each eigenface is treated as distances from eigenspace. These distances are recorded in a vector whose length is equal to the eigenfaces. The distance vector is treated as error vector (EV). This EV plays an important role in the calculation of threshold value which is based on the following steps:

i. Find minimum and maximum of EV for each input face, then difference of minimum and maximum value in EV is recorded as one value in difference error vector (DEV)
ii. The length of DEV becomes equal to test images because there is only one entry in DEV for each input face.
iii. Find the average of DEV which is denoted by ADEV, this becomes the threshold value for decision making.
iv. It is observed that value in ADEV is relatively high when input face belongs to the TS or a similar subject, otherwise this value is relatively low.

The threshold value facilitates to bifurcate test images into an accepted or a rejected class. On the basis of this classification, various experiments have been performed for the validity of the classifier.

4. Results and Discussion

The data set which was used for training and testing purposes contains male and female images of 270 x 300, 180 x 200 and 90 x 100 resolutions [15]. There were small changes in the face position because these images have been acquired in speech mode with no variation in hair style. Variation in lighting was also there. A total of 120 subjects with 20 images per subject having gray background with minor variation in head turn, tilt and slant were present. Out of 2400 images, 6 TSs 200, 160, 130, 100, 70, and 50 images were made.

It has been observed that system response regarding accuracy, learning and decision time varies by varying the number of training images, their resolution and images per subject. To reveal this complexity, the following TSs are made:

a) First group contained 200 images of 10 subjects where 20 images per subject are used. Three TSs are made having image resolutions 270 x 300, 180 x 200 and 90 x 100. For testing purpose, 10 images from training subjects and 10 from other subjects but having same characteristics are chosen randomly. Ten different combinations of leading eigenvalues are selected for the training of an AFPCA system while maximum eigenvalues are involved in case for a PCA base system.

b) Second group contained 160 images of 10 subjects where 16 images per subject are used. Three TSs are made image resolutions 270 x 300, 180 x 200 and 90 x 100 resolutions. For testing purpose, 10 images from training subjects and 10 from other subjects but having same characteristics are chosen randomly. Nine different combinations of leading eigenvalues are selected for training of an AFPCA system while maximum eigenvalues are involved for a PCA base system.

c) Third group contains 130 images of 10 subjects where 13 images per subject are used. Three TSs are made of said resolutions. For testing purpose, same testing set which was discussed in (a). Eight different combinations of leading eigenvalues are selected for training of an AFPCA system while maximum eigenvalues are involved for a PCA base system.

d) Fourth group contains 100 images of 10 subjects where 10 images per subject are used. Three TSs are made of said resolutions. For testing purpose, same testing set which was discussed in (a). Seven different combinations of leading eigenvalues are selected for training of an AFPCA system while maximum eigenvalues are involved for a PCA base system.

e) Fifth group contains 70 images of 10 subjects where 7 images per subject are used. Three TSs are made of said resolutions. For testing purpose, same testing set which was discussed in (a). Seven different combinations of leading eigenvalues are selected for training of an AFPCA system while maximum eigenvalues are involved for a PCA base system.

f) Sixth group contains 50 images of 10 subjects where 5 images per subject are used. Three TSs are made of said resolutions. For testing purpose, same testing set which was discussed in (a). Five different combinations of leading eigenvalues are selected for training of an AFPCA system while maximum eigenvalues are involved for a PCA base system.

In this research, results are presented for 138 experiments regarding face recognition using AFPCA and 18 by using PCA for comparison purposes. When the system does not recognize an image which belongs to the TS at testing phase, it is treated as a *true error*. And if the system gives positive decision for an image which does not relate to training subjects then it is noted as a *false error*. The time required by the system for the extraction of features vector is known as learning time while the time required for testing of 20 images is treated as decision time. Sum of these two time slots is considered as the total time required by the system.

Figure 1 (a), (b) &(c) illustrates error rate which is the sum of true and false errors, decision and learning time of a PCA based system for the data set having resolution of images 270 x 300, 180 x 200 and 90 x 100 respectively. It is observed that the decision and the learning time both increases by increasing with the number of leading eigenvalues while the magnitude of error is inversely proportional to it.

Examination of Figure (1) clearly shows that maximum difference in decision and learning time is 174.68 seconds, 71.96 seconds and 15.46 seconds for resolution of 270 x 300, 180 x 200 and 90 x 100 respectively. The largest observed is for highest resolution having number of images in the TS. This difference decreases when the system uses less resolution with small TSs and it reaches to its minimum value when system employes 50 images in the TS having resolution of 90 x 100. Furthermore, Figure 1(a) shows that the maximum accuracy of PCA based system is 95%. This accuracy can be achieved by AFPCA with various combinations of leading eigenvalues as shown in Figure (2), although decision and learning time is small for 50 eigenvalues. Similarly for TS having images 180 x 200 resolution as shown in Figure (3) and for 70 eigenvalues in TS having images 90 x 100 resolution as shown in Figure (4).

On the other hand, the decision and the learning time shown in Figure (3) reflect almost a linear relationship when compared with the decision and the learning time of image 270 x 300 and 90 x 100 resolutions as depicted in Figure (2&4). The plot of the Figure clearly demonstrates that all TSs touches the same accuracy as that of 180 x 200 resolutions.

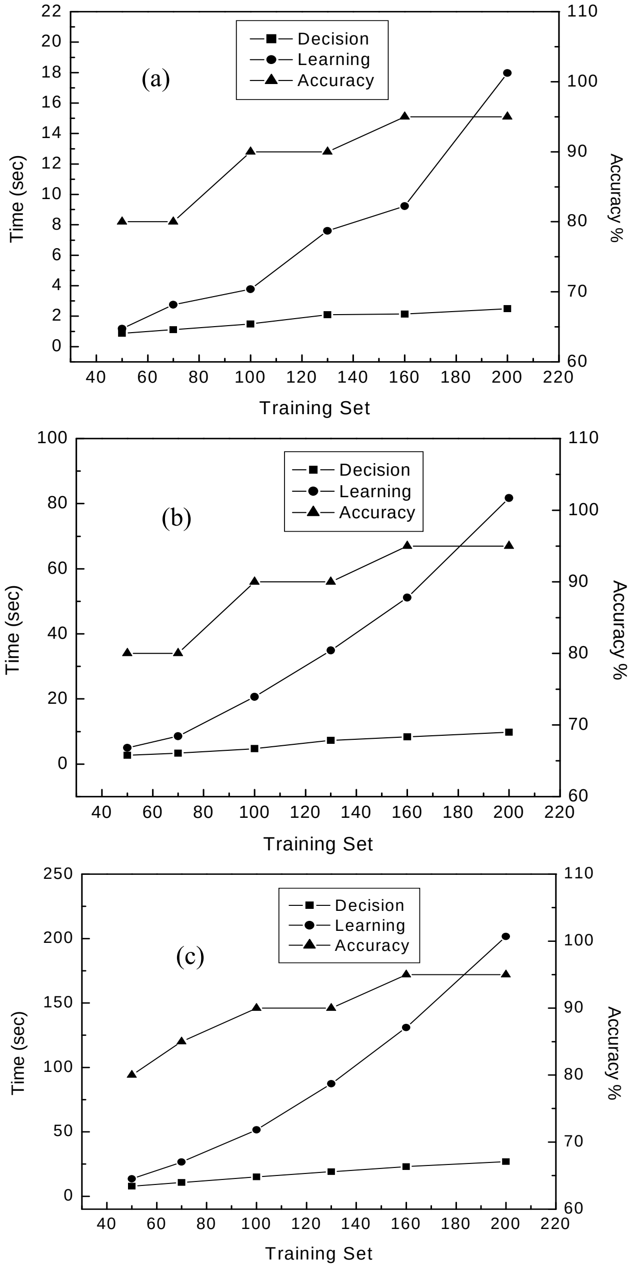


Figure 1. The resolution of training data sets are 90 x 100, 180 x and 200 270 x 300, of PCA in (a), (b) and (c) respectively.

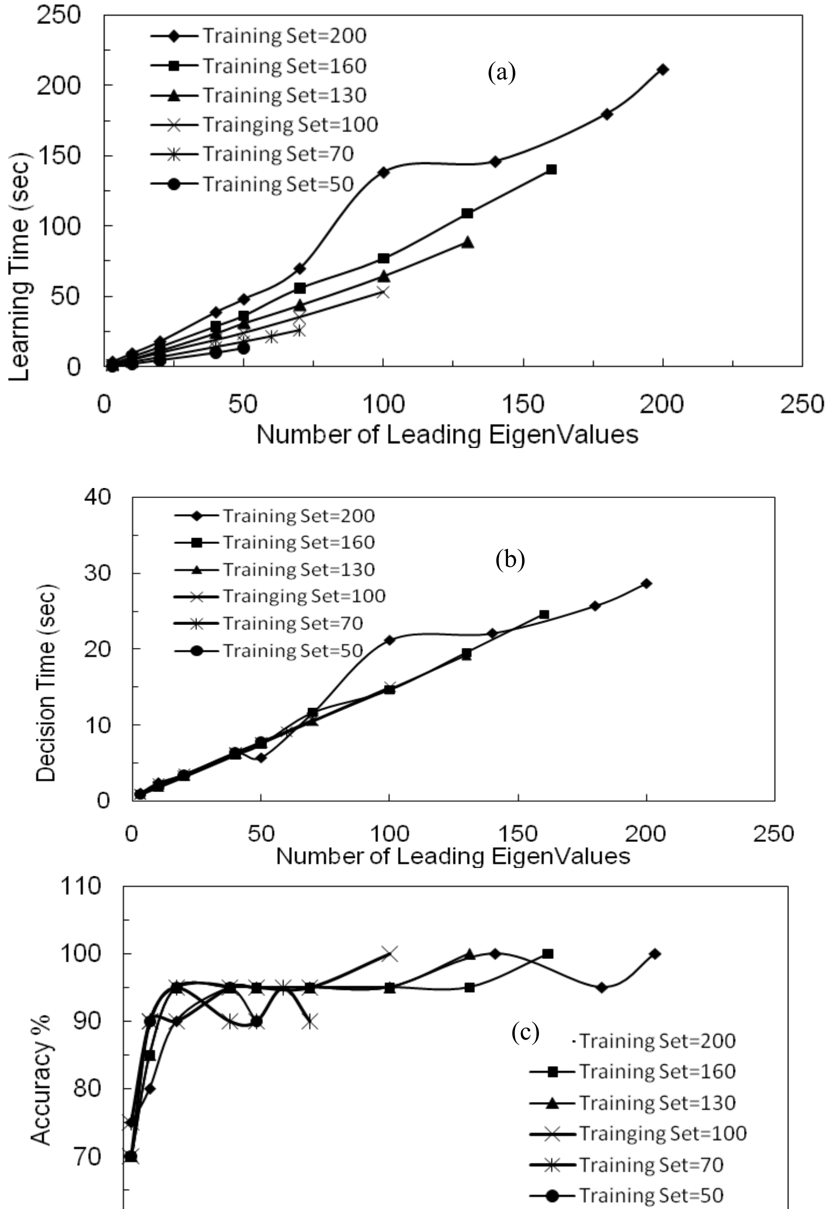


Figure 2. For the resolution 270 x 300, of AFP in (a),learning time (b) decision time (c) accuracy of the system against leading eigenvalues respectively.

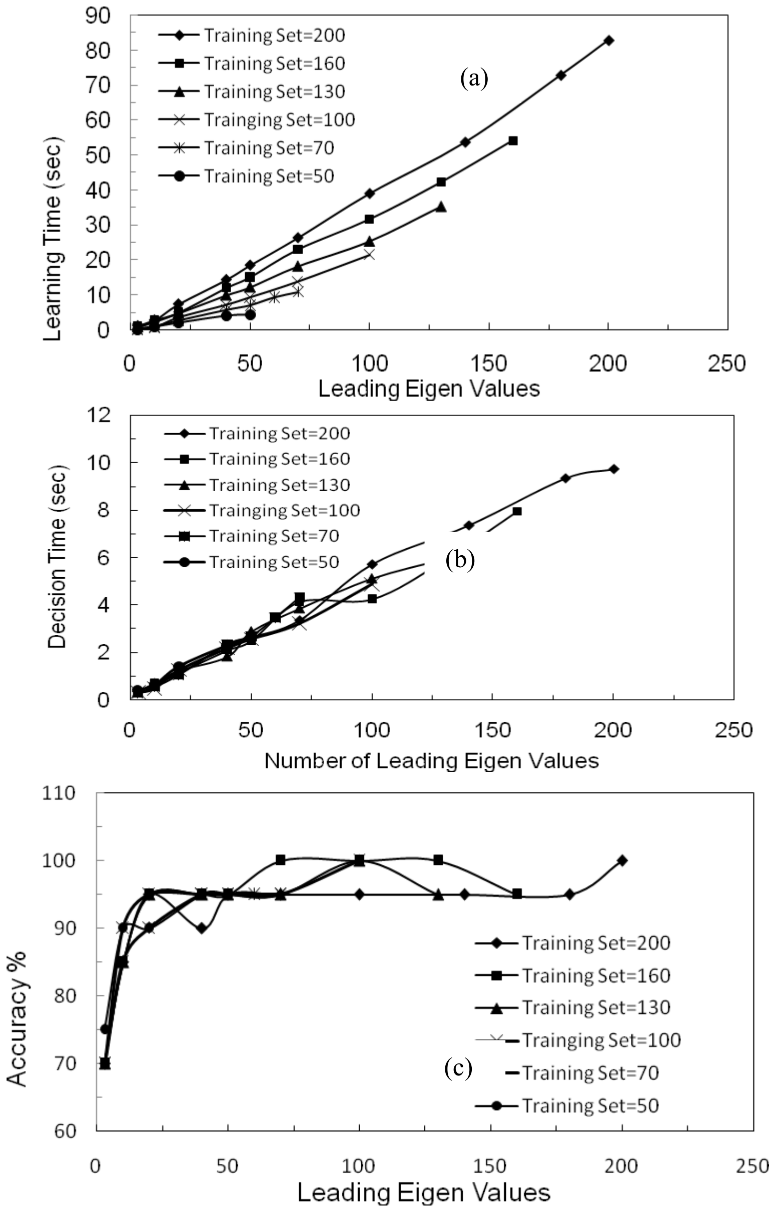


Figure 3. For the resolution 180 x 200, of AFP in (a),learning time (b) decision time (c) accuracy of the system against leading eigenvalues respectively.

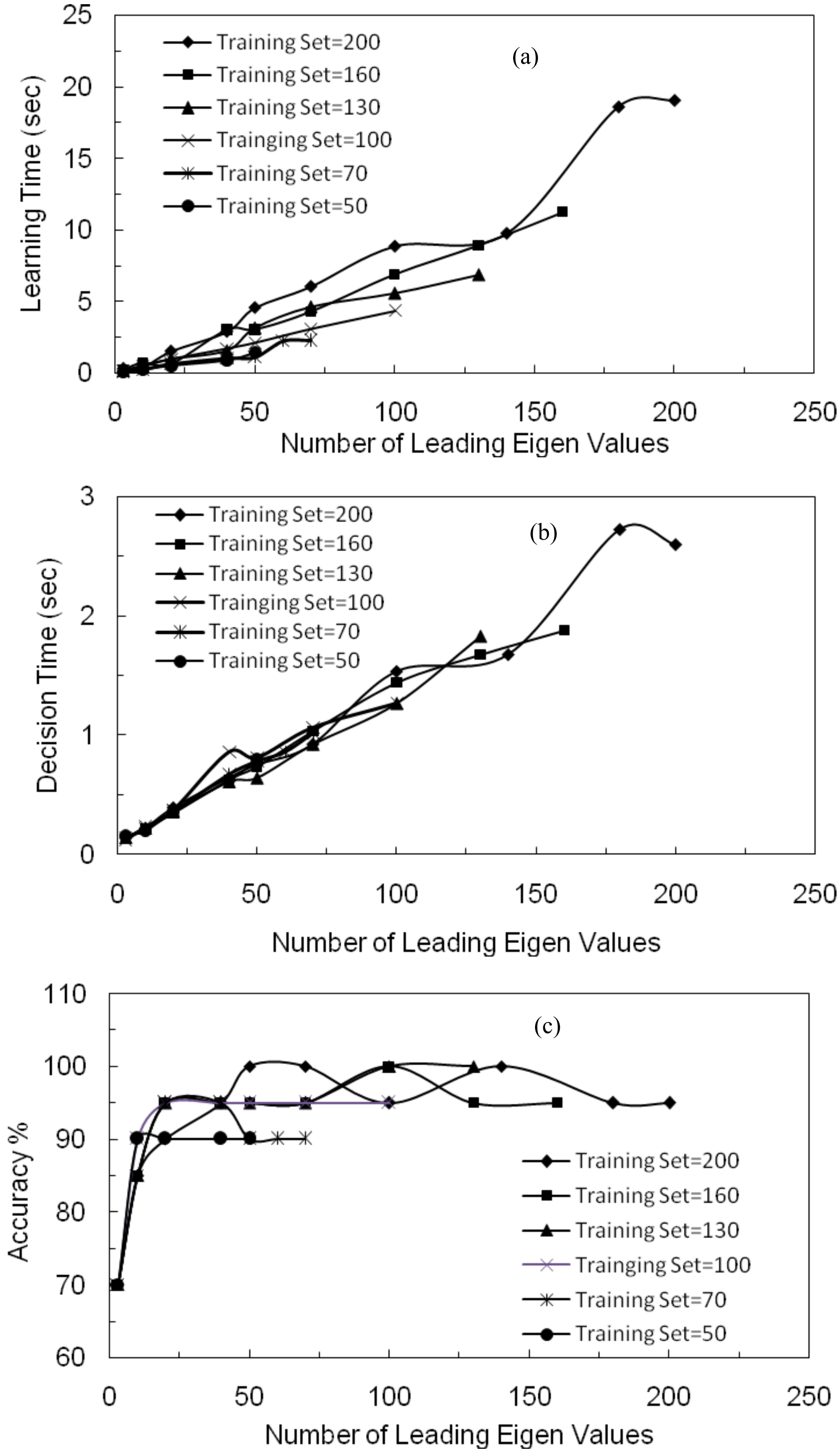


Figure 4. For the resolution 90 x 100, of AFPCA in (a),learning time (b) decision time (c) accuracy of the system against leading eigenvalues respectively.

A comprehensive view of best possible combination of decision and learning time against accuracy of PCA and AFPCA is shown in Table-2. The data of the table shows that for PCA, the accuracy remains the same for different TSs, whereas for AFPCA number of leading eigenvalues (LEVs) plays a crucial role. Entries in Table-2 show best possible combination of the system variable. Decision and learning time increases by increasing the accuracy for both the systems whereas the rate of increase in decision and learning time for AFPCA is sufficiently less than a conventional PCA system. The difference of decision time in PCA and AFPCA is about 19 seconds for 95% accuracy while 100% cannot be achieved for PCA based system as shown in Figure 5 (a).

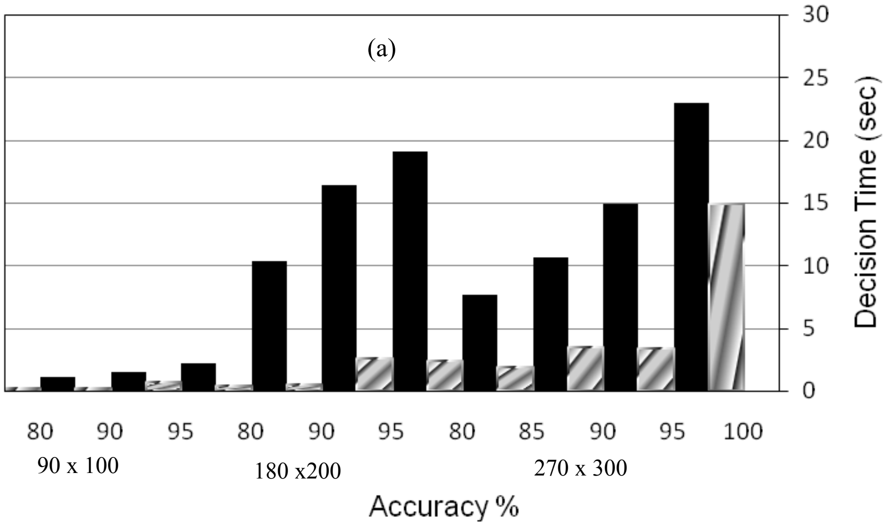
Table 2. Decision and learning time varies with error rate associated with different image resolutions.

IR	Accuracy%	PCA_DT	PCA_LT	min. TS	AFP_DT	AFP_LT	min. TS	LEVs
270 x 300	100	NA	NA	NA	14.938	52.984	100	100
270 x 300	95	22.96	130.92	160	3.375	4.9531	50	20
270 x 300	90	14.937	51.438	100	3.45	9.9531	100	20
270 x 300	85	10.61	26.531	70	1.875	5.8281	130	10
270 x 300	80	7.672	13.469	50	2.4066	9.3594	200	10
180 x 200	95	19.094	51.125	160	2.5625	4.3281	50	50
180 x 200	90	16.432	20.641	100	0.5	0.8125	100	10
180 x 200	80	10.375	5.0313	50	0.4375	0.17188	50	3
90 x 100	95	2.14	9.2344	160	0.6719	1.0469	70	40
90 x 100	90	1.484	3.7813	100	0.234	0.48438	100	10
90 x 100	80	1.1094	2.75	70	0.156	0.078125	50	3

PCA provides 95% accuracy for the three said resolutions involved. Whereas, in this discussion it requires a minimum 2.14 second for decision and 9.23 second for learning when it is trained by using 90 x 100 resolution. On the other hand, AFPCA requires 1.04 second for decision and 0.6719 for learning, when AFPCA is trained with 70 images and uses at 40 LEVs. Thus, the optimum selection of principal components and TS, AFPCA gives efficient face recognition with PCA. In general, it has been demonstrated that AFPCA provides time efficient decision for various image resolutions. Further, the technique provides a definite decision for a real-time face recognition system even for high resolution images.

5. Conclusion

FPCA can generate eigenvalues however, it has limitations in convergence especially when the images are of high resolution. Further, its mean square error is relatively high. The conditions for non-convergence of FPCA have been investigated and a modification in FPCA has been proposed to achieve a definite convergence. It has been demonstrated that the proposed technique offers an efficient solution to calculate eigenvalues even for high resolution images. Furthermore, it has been shown that the developed technique provides time efficient face recognition for a given error rate. The observed error rate in AFPCA is lower than PCA subject to the appropriate selection of TS and leading eigenvalues.



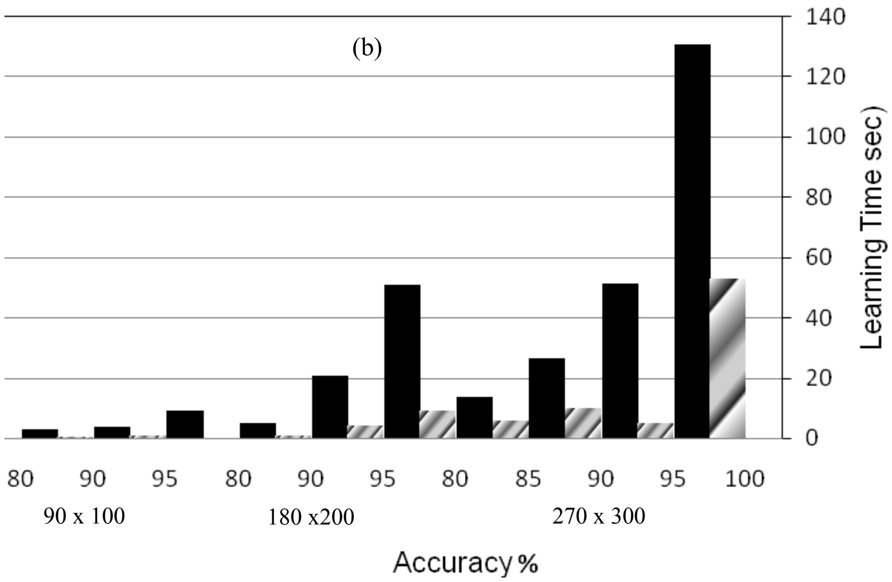


Figure 5. Shows possible accuracy for 270 x 300, 180 x 200 and 90 x 100 resolutions of AFP in (a),decision time (b) learning time

6. References

[1] J. Ashbourn. Biometrics: Advanced Identity Verification. Springer-Verlag London, UK, 2000.

[2] M. Turk, A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neuroscience. vol. 3, no. 1. 1991, pp. 71-86.

[3] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey", ACM Computing Surveys, vol. 35, no. 4, 2003, pp. 399-458.

[4] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face Recognition Using Laplacian faces", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 3, 2005, pp. 328-340.

[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, 1997, pp. 711-720.

[6] R. Zhang, h. Chang, " A literature survey of face recognition and reconstruction technique", Technical report, University of Texas, December, 2005.

[7] A. Sharma, K. K. Paliwal, "Fast Principal component analysis using fixed-point algorithm", Pattern Recognition Letters, Elsevier, vol. 28, no. 10, 2007, pp. 1151-1155.

[8] W. H. Press, S. A. Teukolsky and W. T. Vetterling, Numerical recipes in C: the art of scientific computing, 2nd ed., Cambridge University Press, 1992.

[9] I. Sajid, M. M. Ahmed, I. Taj, M. Humayun, F. Hameed, "Design of High Performance FPGA based Face Recognition System", PIERS 2008 in Cambridge, USA, 2-6 July, 2008.

[10] J. Siwabessy, Penrose, J., Kloser, R., Fox, D., Seabed habitat classification, 1999, International Conference on High Resolution Surveys in Shallow Water, Sydney, Australia.

[11] I. Sajid, M. M., Ahmed, I., Taj, Design and Implementation of a Face Recognition System Using Fast PCA, 2008, IEEE proceeding at the International Symposium on Computer Science and its Applications, Australia.

[12] Rzempoluch, E. J., Neural Network Data Analysis Using Simulnet, 1998, Spring, 195-201.

[13] Anton, H., C. Rorres, Elementary Linear Algebra, ninth edition, Oxford University Press, pp. 138.

[14] Burden, R. L., J. D. Faires, Numerical analysis, seventh edition, Thomson, ISBN:981-243-106-3.

[15] L. Spacek, "Collection of Facial Images: Faces94",
http://cswww.essex.ac.uk/mv/allfaces/ (May 31, 2008).

Authors

Imtiaz Sajid received the M.Sc. Physics degree from Department of Physics , University of Lahore Pakistan, M.S in Computer Sciences from National University of computer and emerging sciences Islamabad Pakistan. He is Ph.D. scholar at Mohammad Ali Jinnah University Islamabad Pakistan. His research interests include pattern recognition, machine vision and reconfigurable computing.

Imtiaz Ahmad Taj received the B.Sc Electronics Engineering degree from UET, Peshawar, The MS and the Ph.D degrees both from Hokkaido University, Japan, Department of Electronics & Information Engineering. His research interests include machine vision and digital communication.

Prof. M. M Ahmed received the MS degree in electronics from Quaid-e-Azam University, Islamabad, Pakistan, in 1987, the M. Phil. degree in Microelectronics from Punjab University, Lahore, Pakistan, in 1990, and the PhD degree in Microelectronics from Cambridge University, U.K., in 1995. From 1987 to 1988, he worked with the Pakistan Atomic Energy Commission, Islamabad as a Scientific Officer where he was involved in instrumentation and design. From 1991 to 1992, he worked as a Research Assistant in Cavendish Laboratory, Cambridge on MMICs.

In 1996, Dr. Ahmed joined Ghulam Ishaq Khan (GIK) Institute of Engineering Sciences and Technology, Topi, Pakistan, as an Assistant Professor in the Faculty of Electronic Engineering, and remained involved actively in teaching and research. From 1996 to 2004, he was rated consistently as an outstanding professor of GIK Institute. In 1999, he was promoted to the rank of Associate Professor in the Faculty of Electronic Engineering and was also assigned the additional responsibility of Dean of the Faculty. In 2004 he joined Mohammad Ali Jinnah University as Dean Faculty of Engineering and Applied Sciences where he is currently working as Executive Vice President of the University.

Dr. Ahmed was a member of organizing committees for the 2nd and 4th IEEE International Multitopic Conferences (INMICs) held in Islamabad in 1997 and 2000, respectively. He was conference chair for INMIC 2003 and 2006. He was chair technical committee for International Conference on Machine Vision, 2007, and International Conference on Emerging trend in Engineering Education, 2007. He has been an invited speaker and session chair in many conferences and symposia.

Dr. Ahmed is a chartered engineer from UK and fellow of Institution of Electrical Engineer, UK. He was awarded title of European Engineer (Euro Ing) in 2002 by European Federation. He is life member of IEEE electron devices, USA and IEEE microwave theory and technique society, USA. He is member of IEEE Communication and Computer Societies.