International Journal of Natural and Engineering Sciences 7 (3): 63-70, 2013
ISSN: 1307-1149, E-ISSN: 2146-0086, www.nobel.gen.tr

A Comparative Study of Face Feature Metrics for a Dynamic and Self-Organised Multimedia Indexing Tool

Sinem ASLAN[1*] E. Turhan TUNALI[2]

[1]International Computer Institute, Ege University, Izmir, TURKEY

[2]Department of Computer Engineering, Izmir University of Economics, Izmir, TURKEY

*Corresponding author:
E-mail: sinem.aslan@ege.edu.tr

Received: October 21, 2013
Accepted: December 16, 2013

Abstract

In this study, Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) methods are integrated into MUVIS, a remarkable multimedia indexing and classification system with an effective indexing structure, for face recognition problem. Their classification performance is compared with that of Gabor Filter which already exists in MUVIS via proposed performance metric. It has been observed that, as far as the classification performance in MUVIS is concerned, LDA performs slightly better than Gabor filter whereas PCA is the worst among them

Keywords: Feature integration, MUVIS, Principal Component Analysis (PCA); Linear Discriminant Analysis (LDA); Hierarchical Cellular Tree (HCT); Classification of Face Photos

INTRODUCTION

Due to widespread use of the Internet, it has been a necessity to facilitate the usage of multimedia data and improve the indexing and storing capabilities of multimedia databases. In conjunction with these databases, face recognition and classification are widely used for variety of purposes including security applications. Face of a person is one of his/her unique characteristic that distinguishes him/her from the others and reaching crime records of a criminal by searching his/her photo in the stored criminal photo database at security centers is a typical example of application. When number of photos rise into vast quantities, search time increases and searching a subset of the collection rather than the complete database increases efficiency. Metric and Space Access Methods are developed for indexing and classification purposes. Extraction of some feature vectors that hold face characteristics are used to index photos. By this way, photos belonging to the same person are stored in the same location leading to easier and lower retrieval cost.

MUVIS is a multimedia retrieval system that uses metric-based, dynamic, cell based and hierarchical indexing method. Gabor Filter is the feature extraction method which can be used for face recognition and has already been implemented in MUVIS. In this study, we integrated two more facial feature extraction methods, namely PCA and LDA into MUVIS and compare the performance of all three methods on a passport photo database. Although LDA and PCA are well known methods used in the literature for face recognition purpose, their performance in MUVIS which performs an effective indexing scheme Hierarchical Cellular Tree (HCT) has not been examined for face recognition purpose before. Encountered difficulties in the integration process are discussed at the proper sections.

The paper is organized as follows: In Section 2, we give preliminary information on face recognition and MUVIS multimedia indexing platform. In Section 3, the adaptation of PCA and LDA methods to MUVIS are explained. In Section 4, we first introduce the performance metrics we developed, followed by the extensive comparison results for Gabor, PCA and LDA. Finally, in Section 5, concluding results are made.

RELATED WORKS

Indexing and Classification

Today, as a result of widespread usage requirement of huge-sized multimedia databases, improving database management structures to provide low memory cost and low item retrieval time has been a necessity for the efficient resource management. Spatial Access Methods (SAM) and Metric Access Methods (MAM) are developed to decrease query and retrieval times. SAMs are indexing structures storing geometric shapes that are the representations of various objects in multidimensional space [1]. They have some drawbacks especially for large scale multimedia databases. Multimedia databases have several feature types whereas SAMs work with only one feature space. Additionally, while these methods have better performance for low dimensional feature space, they exhibit insufficiency for multidimensional feature space due to the curse of dimensionality problem [2]. MAMs require a similarity distance function that satisfies symmetry, non-negativity and triangular inequality rules thereby allowing several features with different sub-features to be used for a multimedia database [2]. Characteristic problems of MAMs such as disk page access costs and static and unbalanced tree structure are attacked leading to M-trees [3]. In M-tree, similarity is computed between a new item and cell nuclei in a top-down manner on the tree and new item is inserted into the cell which has the most similar nucleus to the new item. It has been shown in [2] that M-Tree method, in fact, does not yield optimal subtrees.

Besides the indexing methods, some content-based indexing and retrieval systems have been developed such as MUVIS [2], QBIC [4], PicHunter [5], Virage [6], Photobook [7], VisualSeek [8], VideoQ [9]. All of these application systems provide a framework in which methods for indexing and retrieving still image and/or audio together with video data compression are included. Among these, MUVIS is designed for the purpose of providing a unified and global framework which implements capturing, recording, indexing and retrieval combined with browsing and various other visual and semantic capabilities. It also provides a user-friendly platform in which robust algorithms can be implemented, tested, configured and compared [10].

MUVIS uses a MAM based, dynamic and self-organized indexing method named as Hierarchical Cellular Tree (HCT), which is composed of one or more levels [11]. Each of these levels contains one or more cells that have similar items inside. HCT body is constructed by formation of levels after cell mitoses. Unlike M-Tree, item insertion in HCT is implemented by using preemptive cell search algorithm. This search technique takes into account not only the most similar cell nucleus to the new item, but also the most similar item of candidate cells [2]. In contrast to M-tree and its variations, the MUVIS system finds the optimal sub-tree by finding out all possible nucleus items on the upper level that are likely to provide the most similar items on the lower level. Moreover, by deleting the old nucleus item in the upper level and inserting the two new nucleus items not into the old cell but into the optimum cell on the upper level, MUVIS obtains optimal sub-tree.

In this study, we will use MUVIS framework that implements HCT indexing scheme.

Face Recognition

There is a vast amount of literature on face recognition that includes holistic-based, feature-based and hybrid methods [12]. Although some color, texture and shape features are integrated in MUVIS in order to evaluate the retrieval performance of natural images [13-15], face recognition problem is not specifically studied. In [16], a new method, named as MAT, is proposed and compared with PCA for dimension reduction of HSV, YUV, and RGB color histogram features of still images. As far as face recognition is concerned, the most important procedure implemented in MUVIS is Gabor wavelets which is a feature-based technique leading to a considerably compact representation [18-19]. Main important components on face region such as eyes, mouth, nose edges and other points such as moles, dimples, scars etc. are enhanced via Gabor functions, hence a feature map corresponding to the face characteristics belonging to an individual can be obtained with these enhanced feature locations and used in recognition processes. Due to its robustness in face recognition problem, Gabor wavelet is a popular tool and it is going to form our reference point in our comparisons. Color histogram that is a holistic-based feature extraction method is also implemented in MUVIS, yet, due to its limited robustness in face recognition problem [19-20], we are not going to include it in our comparisons.

PCA is proposed by [21] to characterize face images in low-dimensional space. Then, [22] implemented a face recognition application by using eigenspace projection method, which is also known as Karhunen-Loéve or Principal Component Analysis in statistics. Instead of using the entire image space for the face representation, it is more efficient to use a subspace which has a much lower dimension with ability of defining faces successfully. This subspace is named as face space and the main vectors of this space are named as principal components. PCA aims to decrease the dimension of a space while it provides new axes that define the set model better. Set model is constructed by utilizing sample images in such a way that axes of face space or eigenvectors would be orthogonal and the variation between data points is maximized. Projection of a face image on face subspace creates a lower representation of the face image that holds the most important characteristic features of it. Projection of any face image in the database on this face space would give a weight vector and this vector which is unique to the specified face image is used as a feature vector for that image. Moreover, a face image can be reconstructed by using these weight vectors. There is a vast amount of literature on PCA that we are not going to cite. Although PCA is firstly proposed as a dimension reduction method, it is also reported to be robust as far as face recognition is concerned. PCA is also used by [16] in dimension reduction process of color histogram features and its computational complexity is compared with that of so called MAT method that is also developed in [16]; however, its classification performance as a facial feature on HCT is not studied yet. We are going to include two different versions of PCA with different number of eigenvectors in our comparison platform.

LDA was first developed in [23] and it has been applied to various problems in computer vision. It aims to find axes that bring similar data points closer to each other while setting dissimilar data points apart from each other through construction of a lower dimensional [24]. For face recognition, images belonging to same person that are taken at different illumination conditions or with different facial mimics constitute a class.

The difference between LDA and PCA is the way they calculate the subspace. Since PCA maximizes total scatter with the addition of unwanted variations due to lighting and facial expression, LDA seems to be more optimal from a discrimination standpoint as far as face recognition is concerned [25].

FACE RECOGNITION IN MUVIS

Notes on PCA Implementation

In the integration process of PCA into MUVIS, we assumed that training images of people are available previously. By extraction of feature vectors from these training images, some distinctive information is acquired before the actual classification process. In other words, in order to perform the classification processes in the future, we construct the initial projection subspace by these training images. Additionally, an important property of the PCA integration process is the projection subspace is updated by the contribution of each newly inserted image. The computational difficulty of large dimensional covariance matrix is solved by the method proposed in [26]. To determine the number of eigenvectors to be included, we have constructed the face space by calculating the first 10, 80, 120 and 240 eigenvectors of training sets which consisted of 10, 80, 120, 240 images respectively [27]. We repeated the same procedure by using 400 images in all cases. Then, we have reconstructed an input image by using its weighting vector which was obtained by projecting the input image onto the differently constructed face spaces. Finally, the similarity between the reconstructed and the input images were evaluated. Fig. 1, Fig. 2, and Fig. 3 show the reconstructed images with the specified conditions. Since the best reconstructed images were achieved with the face space constructed by 120 and 240 eigenvectors, tests are carried out with 120 and 240 eigenvectors where the size of the images is 112×92 yielding vectors of size 10304. The initial condition is constructed by the face space including 120 and 240 eigenfaces and a HCT tree is constructed according to this face space by the first 120 and 240 images in the database. Just after the construction of the initial condition, the face space updates itself along with the insertion of 121st and 241st images. This means that each image inserted after the 120th and 240th images also becomes a member of the training set.

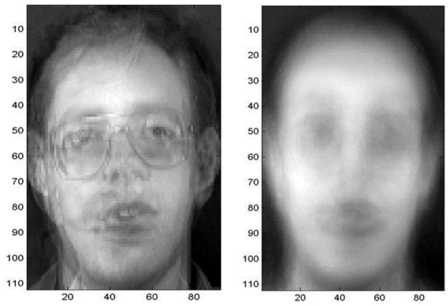


Figure 1. Reconstructed image where (a) number of images in the training set is 10 and number of eigenvectors used in face spaces is 10 (b) number of images in the training set is 400 and number of eigenvectors used in face spaces is 10.

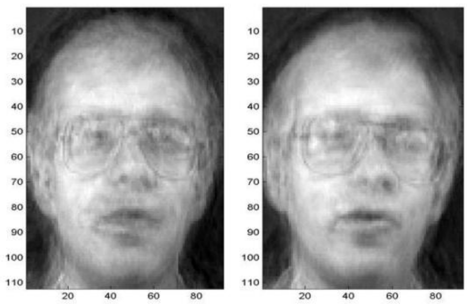


Figure 2. Reconstructed image where (a) number of images in the training set is 120 and number of eigenvectors used in face spaces is 120 (b) number of images in the training set is 400 and number of eigenvectors used in face spaces is 120.

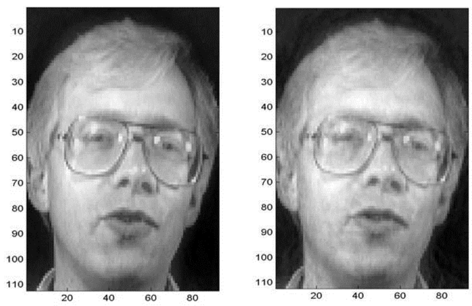


Figure 3. Reconstructed image where (a) number of images in the training set is 240 and number of eigenvectors used in face spaces is 240 (b) number of images in the training set is 400 and number of eigenvectors used in face spaces is 240.

Notes on LDA Implementation

The critical question in adaptation of LDA into MUVIS is how to define the class information in the training process. In this study, we used variety of training datasets including the one with first five images of each person. These five images of each person constitute a class and after using this class information in the construction of subspace, a feature vector for each person would be obtained by the projection process. To overcome the time and space complexity problems, the face vectors are downsized from 10304 to 2576 by averaging neighboring four pixels. After specifying the class information, the Total Scatter Matrix (TS) and Within-class Scatter Matrix (WS) are computed and Between-class Scatter Matrix (BS) is obtained as BS=TS-WS [28]. In order to determine the subspace holding the discrimination information, eigenvalues and eigenvectors that maximize the BS/WS ratio must be calculated. However, despite of the fact that BS and WS are real symmetric matrices, WS-1BS becomes a non-real and non-symmetric matrix. In order to overcome this problem, we used the diagonalization method of two symmetric matrices given in [28]. The discrimination subspace is then obtained by sorting the eigenvalues and their respective eigenvectors. The LDA features for each image is obtained by projecting each inserted image on this subspace and finally HCT is constructed by these feature vectors. The computations are carried out in MATLAB.

RESULTS

The ORL database of face images that is constructed at Olivetti AT&T Laboratories is used in the tests. There are 400 photos in this database belong to 40 different people each of which 112×92 pixels. Each person has 10 different photos that were taken from frontal view at different illumination conditions and different times. A sample of the images at the database can be seen in Fig. 4.

Table 1 contains the parameters used in HCT construction. These parameters control variety of factors in cell formation [2].

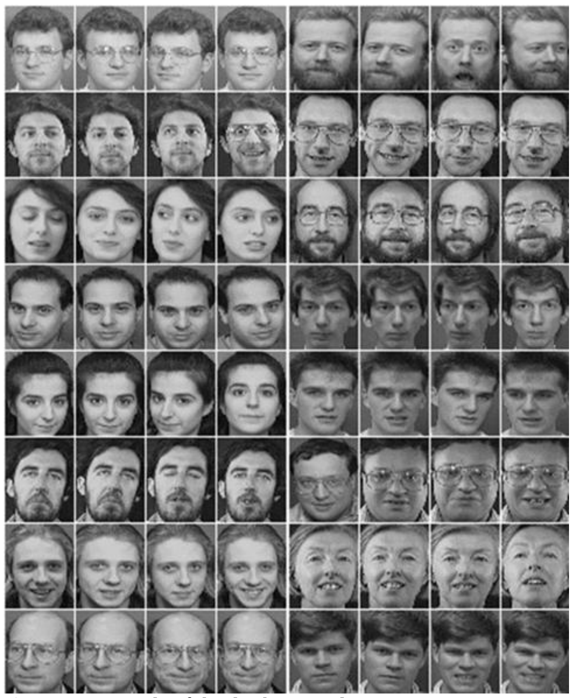


Figure 4. A sample of the database used in tests.

Table 1. Parameters used to construct the HCT.

Parameter	Value
Maturity level	11
Cell Size in the top level	40
Cell Compactness Trend Factor	0.1
Fitness Check Time Period	100
Cell Compactness Updating Time Period	100

Diversity of People in Cells

This is the number of cells containing different number of people. For example, in Fig. 5, there are two cells that contain images belonging to five different people for the Gabor filter case. In the ideal case, all images belonging to an individual should be located in only one cell.

As it can be seen from Fig. 5, the best distribution is obtained by LDA followed by Gabor filter, PCA 240 and PCA 120 respectively.

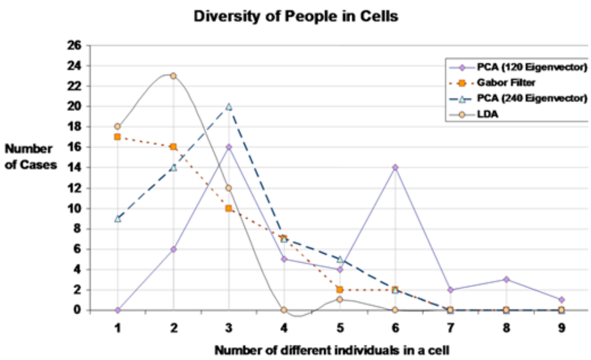


Figure 5. Diversity of People in Cells Graph

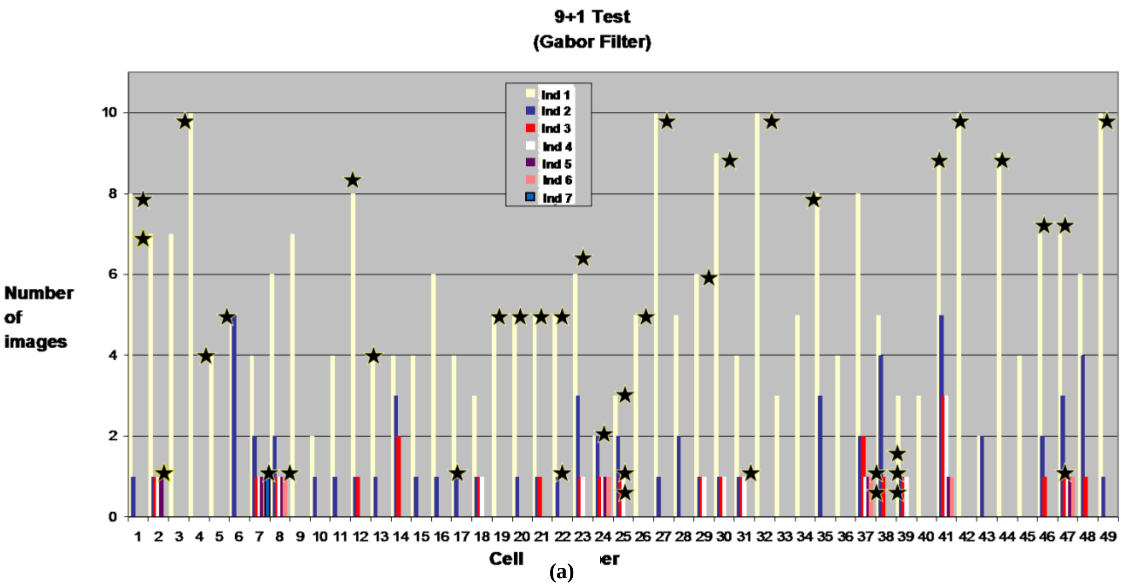
Classification Correctness

Features from nine photos of forty individuals are extracted and indexed by HCT. Feature vectors of the remaining one photo of each person are extracted afterwards and these vectors are inserted to the already constructed HCT. A destination cell is labeled as correct cell if the cell currently contains maximum number of photos belonging to that individual and this situation is named as ideal case. Consequently, the number of ideal cases is determined. We will call this as 9+1 test.

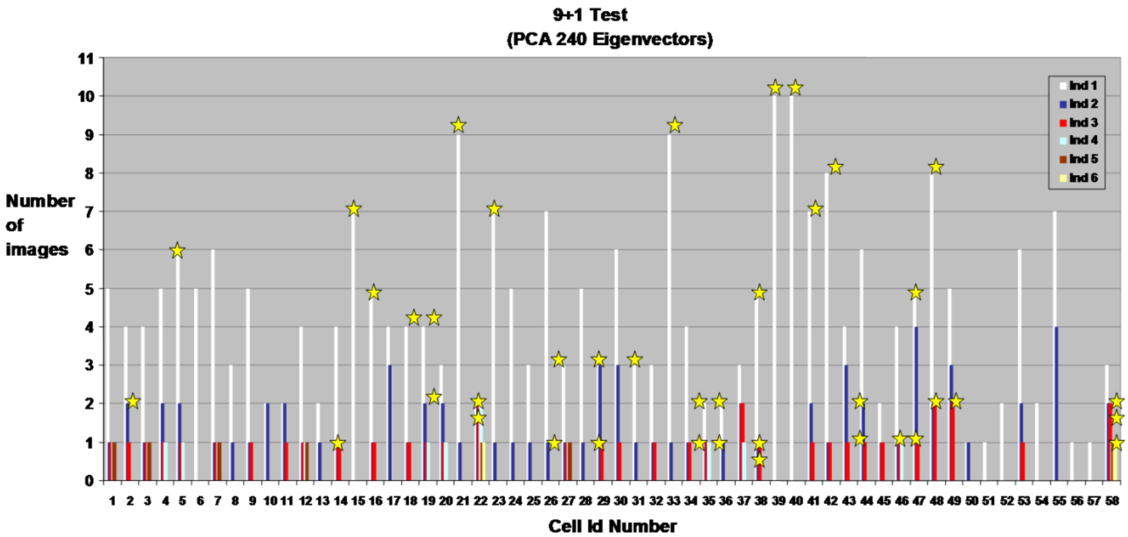
In Fig. 6, a star symbol is used to illustrate the last inserted image in the specified cell. It can be seen that there are more than one star signs in some of the cells which means that images belonging to more than one individual were increased by one in the specified cell. In the ideal case, we expect that each of the lastly inserted 40 images be located into the cell which contains maximum number of images of that person. It can be seen that LDA outperforms Gabor filter in this setting too.

Incremental Classification Correctness

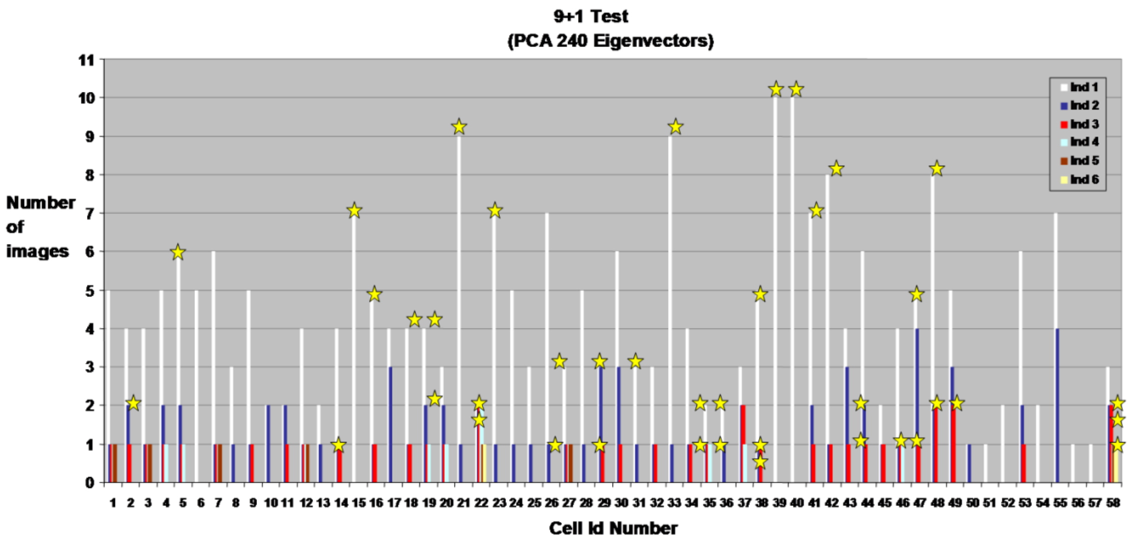
To evaluate the performance under limited number of sample images, a number of incremental results are obtained. In addition to 9+1 test of the previous case, 5+1, 6+1, 7+1 and 8+1 tests are performed. However, instead of reporting all of the cases, only those that have been located to full cells are reported. For example, in Fig. 7(a), only one of the forty images is located to a cell that contains five images of that person. When we examine Fig. 7, excluding 7+1 test which seems to be an exception, LDA is slightly better than Gabor filter. Note that Fig. 7 does not include cases where the last image is located into the cell that contains maximum number of images belonging to a person. To extract such information, figures similar to Fig. 6 are obtained. It has been observed that, on the average, LDA is slightly better that Gabor filter in all cases.



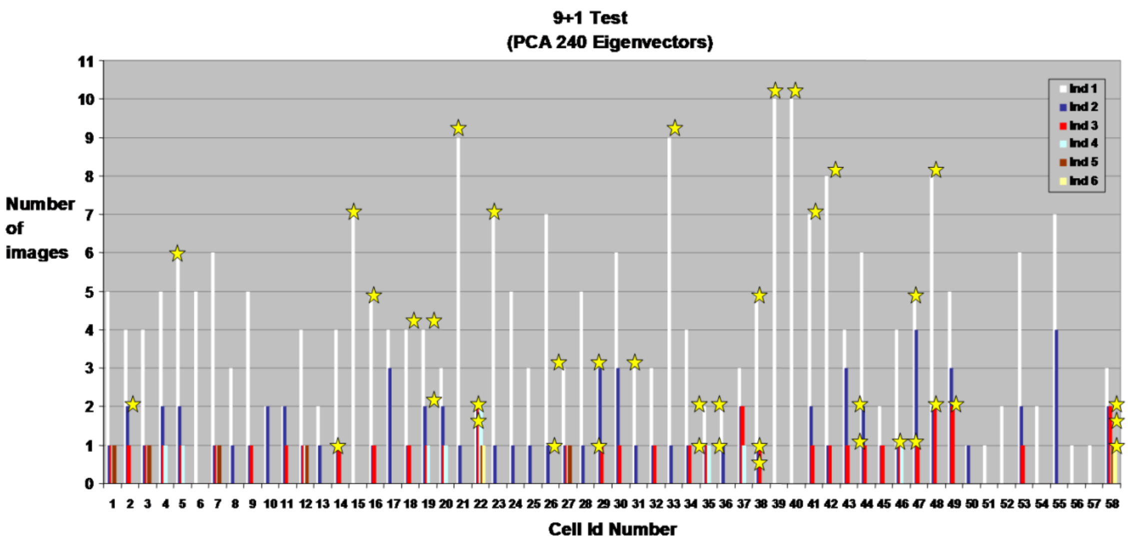
(a)



(b)



(c)



(d)

Figure 1: Classification correctness test (a) PCA using 120-eigenfaced space, (b) Gabor filter, (c) PCA using 240-eigenfaced space, (d) LDA.

Distribution of Images

The number of images of each person inside a cell is defined as a case. For example, for LDA, there are 6 cases at which 9 images of one person are located together in a cell in Fig. 8. When Fig. 8 and Fig. 9 are examined, we see that even though for cells containing more than two images, LDA and Gabor have approximately similar performance, for cells containing one and two images, LDA performance is considerably better.

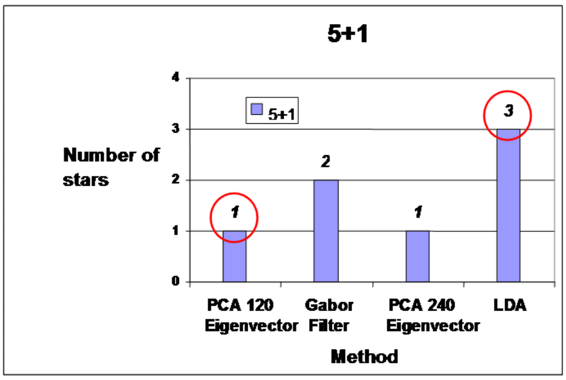
As it can be seen in Fig. 8, there are 145 cases where a cell contains one unique image of an individual when PCA with the face space of 120 eigenvectors is used. This would be evaluated as a comparatively good result if the remaining nine images belonging to each individual were collected in one distinct cell. However, we can see in Fig. 6 that there is only one case in which 5, 6, 7, 8 and 9 images belonging to one individual collected in one cell, thus the performance result of this method is quite poor. When the number of eigenvectors is increased to 240 in the face space construction for PCA, the number of cases where a cell contains one unique image of an individual decreases from 145 to 83, thus the performance rises nearly 100% and approaches to that of Gabor filter.

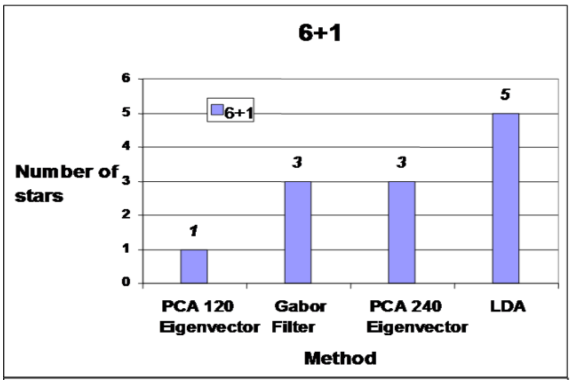
Incremental Distribution of Images

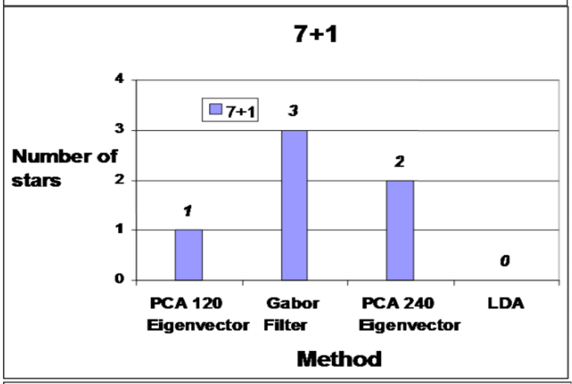
In this set of measurements, the purpose is to see if the feature extraction method leads to a learning capability for classification as new images of a person are inserted to the database. Starting with the first image of the first person, all the first images of forty people are inserted. Then, in the second round, second images of the forty people are inserted. All the images are inserted to the database in such a round robin fashion. At the end of each round, number of cases is plotted against number of images belonging to the same person. This measurement is similar to previous case with the only difference that images are fed to the HCT algorithm in a systematic way rather than totally random way. It has been observed that, in the first rounds, LDA exhibits the best learning ability with Gabor filter catching up later on.

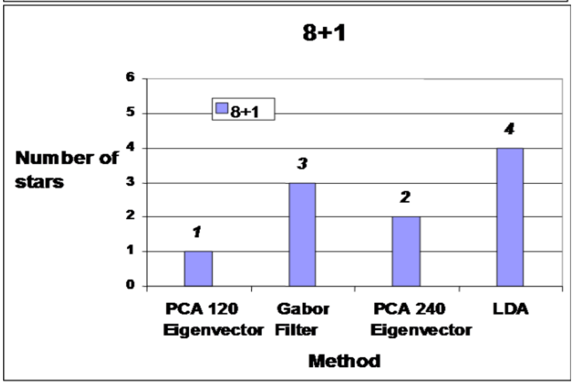
Classification Success Ratio

This is another test to measure accuracy of the classification. A collection of 200 images containing randomly chosen five images of each of the 40 individuals is inserted into MUVIS and after the feature extraction process they are indexed by HCT as usual. The rest of the collection of images is then inserted while observing the cases in which the image is located into a cell that contains at least five images of the individual. Classification success ratio is the ratio of these cases to the total of 200 images inserted in the second round. It has been observed that, as percentage, this ratio is 4.5, 19, 21.5 and 42 for PCA-120, PCA-240, Gabor filter and LDA respectively. We again observe that LDA outperforms Gabor filter in this test.









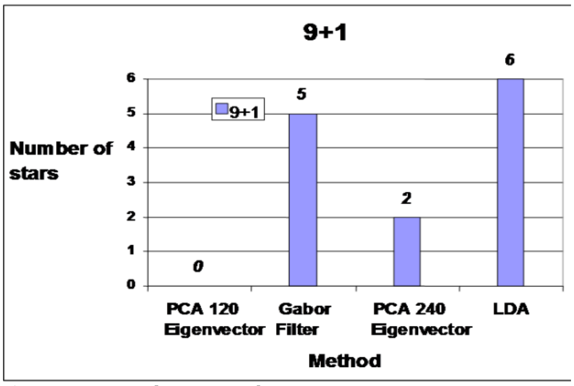


Figure 7. 9+1 evaluation graphs.

CONCLUSION

We integrated the PCA and LDA feature extraction methods into MUVIS to classify passport photographs and aimed to compare their performance with that of Gabor filter that has already been integrated into MUVIS. To be able to do this comparison, we have introduced extensive classification performance metrics and carried out tests to compare the above mentioned methods.

Our observations revealed that LDA outperforms Gabor filter in all cases whereas PCA lags behind both of the methods. When the number of eigen vectors increase, the performance of PCA approaches Gabor filter. We have also observed that, the higher the number of images belonging to a person, the better the classification performance.

Several problems still affect the experiments. The first one is the importance of the order in which images are inserted. That is, the formation of HCT is dependent on the incoming order of the images to be classified. This drawback is overcome by repeating the randomized ordering process variety of times. Another drawback is that while PCA face space updates itself by the insertion of each new image, the discrimination space of LDA remains constant. An improvement in LDA that can update this space is expected to yield even better results.

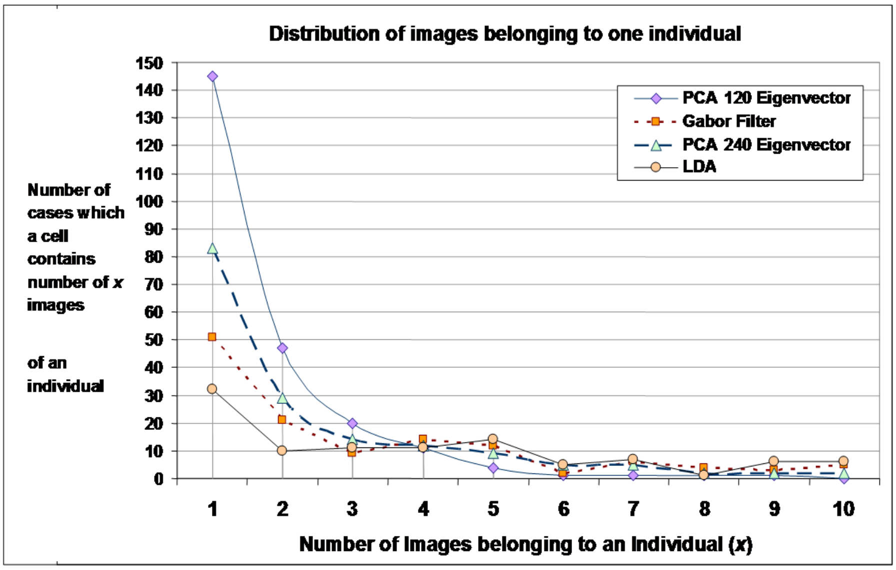


Figure 8. Diversity of People in Cells Graph

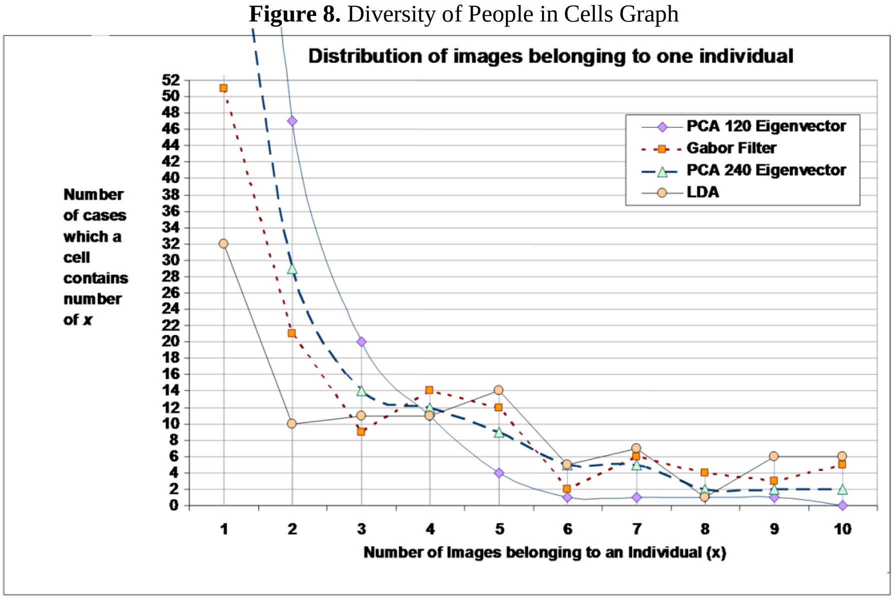


Figure 9. Diversity of People in Cells Graph (zoomed)

Acknowledgment

We would like to thank Prof. Dr. Serkan KIRANYAZ of Tampere University of Technology for his support about MUVIS and HCT software.

REFERENCES

[1] V. Gaede and O. Günther, Multidimensional access methods, *ACM Computing Surveys*, 30(2) (1998), pp. 170-23.

[2] S. Kiranyaz, Advanced Techniques for Content-Based Management of Multimedia Databases, *Ph.D. Dissertation*, Tampere University of Technology, Tampere, Finland (2005)

[3] P. Ciaccia, M. Patella and P. Zezula, M-tree: An Efficient Access Method for Similarity Search in Metric Spaces, *Proceedings of the 23rd International Conference on Very Large Data Bases* (1997), pp. 426-435.

[4] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, Query by image and video content: the QBIC system, *Computer*, 28(9) (1995), 23-32.

[5] I.J. Cox, M.L. Miller, S. M. Omohundro, and P.N. Yianilos, PicHunter: Bayesian relevance feedback for image retrieval, *Proceedings of the 13th International Conference on Pattern Recognition*, 3 (1996), pp. 361-369.

[6] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C.F. Shu, The Virage image search engine: An open framework for image management, *Proceedings of Storage and Retrieval for Image and Video Databases*, doi:10.1117/12.234785 (1996).

[7] A. R. Pentland, W. Picard, and S. Sclaroff, Photobook: Content-based manipulation of image databases, *International Journal of Computer Vision*, 18(3) (1996), pp. 233-254.

[8] J.R. Smith and S.F. Chang, VisualSEEK: a fully automated content-based image query system, *Proceedings of ACM International Conference on Multimedia*, (1996) pp. 87–98.

[9] SF. Chang, W. Chen, and H. Sundaram, VideoQ: a fully automated video retrieval system using motion sketches, *Proceedings of 4th IEEE Workshop on Applications of Computer Vision*, (1998) pp. 270-271.

[10] M. Gabbouj, E. Guldogan, M. Partio, O. Guldogan, M. Birinci, A. Iftikhar, S. Uhlmann, and S. Kiranyaz, An Extended Framework Structure in MUVIS for Content-based Multimedia Indexing and Retrieval, *Proceedings of 4th International Conference on Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisia (2007).

[11] S. Kiranyaz, and M. Gabbouj, Hierarchical Cellular Tree: An Efficient Indexing Scheme for Content-based Retrieval on Multimedia Databases, *IEEE Transactions on Multimedia*, 9(1) (2007), pp. 102-119.

[12] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, Face recognition: A literature survey, *ACM Computing Surveys*, 35(4) (2003), 399-458.

[13] E. Guldogan, S. Kiranyaz, K. Caglar, and M. Gabbouj, Dynamic Integration of Explicit Feature Extraction Algorithms into MUVIS Framework, *Proceedings of Finnish Signal Processing Symposium*, Tampere, Finland (2003).

[14] M. Partio, E. Guldogan, O. Guldogan, and M. Gabbouj, Applying texture and color features to natural image retrieval, *Proceedings of Finnish Signal Processing Symposium*, Tampere, Finland, (2003) pp. 199-203.

[15] S. Kiranyaz, M. Ferreira, and M. Gabbouj, A Generic Shape/Texture Descriptor Over Multiscale Edge Field: 2-D Walking Ant Histogram, *IEEE Transactions on Image Processing*, 17(3) (2008), pp. 377-391.

[16] E. Guldogan, and M. Gabbouj, Mapping by adaptive threshold method for dimension reduction of content-based indexing and retrieval features, *Proceedings of European Signal Processing Conference*, Antalya, Turkey, (2005), pp. 4.

[17] J.G. Daugman, Two-dimensional spectral analysis of cortical receptive field profiles, *Vision Research*, 20(10) (1980), pp. 847-856.

[18] C. Liu, and H. Wechsler, A Gabor Feature Classifier for Face Recognition, *Proceedings of 8th IEEE International Conference on Computer Vision*, (2001) pp. 270.

[19] S. Marcel, and S. Bengio, Improving Face Verification Using Skin Color Information, *Proceedings of the 16th International Conference on Pattern Recognition*, 2 (2002), pp 11-15.

[20] C. Nastar, and M. Mitschke, Real-time face recognition using feature combination, *Proceedings of 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, (1998) pp. 312-317.

[21] I. Sirovich, and M. Kirby, Low-dimensional procedures for the characterization of human faces, *Journal of the Optical Society of America*, 4(3) (1987), pp. 519-524.

[22] M. Turk, and A. Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, 3(1) (1991), pp. 71-86.

[23] R.A. Fisher, The Use of Multiple Measures in Taxonomic Problems, *Annals of Human Genetics*, 7(2) (1936), pp. 179-188.

[24] R. Gutierrez-Osuna, Lecture Notes on Fisher Linear Discriminants, http://research.cs.tamu.edu/prism/lectures/pr/pr_l10.pdf, *Resource document* (2002), Accessed 15 July 2013.

[25] 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*, 19(7) (1997), pp. 711-720.

[26] M.A. Turk, Random Walk through Eigenspace, *IEICE Transactions on Information and Systems*, E84-D(12) (2001), pp. 1586-1595.

[27] S. Aslan, A Comparative Study of Feature Metrics For Classification Of Human Passport Photos, *MSc. Dissertation*, Ege University, International Computer Institute (2007).

[28] W.W. Cooley, and P. R. Lohnes, Multivariate data analysis, *John Wiley & Sons* (1971).