Face Recognition in compressed domain by applying Canonical Correlation Analysis based feature vector optimization and Neural Network matching

Menila James† and Dr. Arockiaswamy††

† The Department of Computing, Muscat College, Sultanate of Oman
†† College of Information Systems, University of Nizwa, Sultanate of Oman

Summary
This research work presents an efficient approach for performing face recognition in compressed domain by applying Canonical Correlation Analysis based feature vector optimization. CCA is a dominant method for multivariate analysis and therefore a powerful method of feature projection based on CCA is proposed for compressed facial images. A major advantage of the proposed approach is the fact that face recognition systems can straightly work with JPEG2000 compressed images as it uses entropy points as input to the new face recognition system based on CCA. Cascade Forward back propagation neural network is used in the new method for matching of the images. Labeling of images is applied in the lateral part of the research for better recognition of images with varying expressions. The experimental results proved that the proposed method is very effective in achieving high Recognition Rate (RR) and Normalized Recognition Rate (NRR) with great reduction of computational time.

Key words:
Image compression, Discrete Wavelet Transform, Face Recognition, Canonical Correlation Analysis, Neural Networks.

1. Introduction

In the present networked world, the necessity of preserving the security of information or physical belongings is becoming both progressively more important and increasingly difficult. During recent years, there is lot of crimes regarding the credit card fraud, computer break-in by hackers, or security breaches in an organization or government offices. Biometric access controls are computerized techniques of authenticating or identifying the identity of a person according to their physiological characteristics, like fingerprints or facial features, or certain features of the person's activities, for instance his/her handwriting style or keystroke patterns [1]. In recent times, Biometry is one of the most actively researched areas traversing many sub-disciplines such as computer vision, pattern recognition and image processing [2]. The main advantage of face recognition as a biometric method is that it attains merits of both high precision and low intrusiveness [3] [4]. Users do not feel that they are under surveillance and that their privacy is being invaded and hence face recognition has good precision of a physiological approach without being intrusive [5].

Until recently uncompressed images were used in face recognition and image compression was a totally ignored area of face recognition research. There are various key reasons for using compressed images in face recognition systems. The main motivation was the introduction of e-passports which use face image as the main identifier [6]. Transmission requirements and storage requirements would be greatly reduced by using compressed images [7]. However, compressed image needs to be reconstructed (pixel) in order to carry out face recognition in classical setups. The reconstruction of each and every image is very computationally expensive and hence avoiding full decompression of image will be very helpful in face recognition. Therefore, the entropy coefficients are used by the face recognition system in compressed domain, without executing inverse quantization and inverse transform [8]. As a result, computational time and storage requirements will be extremely reduced since it allows the use of compressed images in face recognition.

The main objective of this research is to observe the potential of performing face recognition techniques directly into the compressed domain avoiding full decompression of facial images. JPEG2000 is the new standard for the compression of images which offers new features increasingly demanded by modern applications and superior compression rates at low bit-rates. As ANSI and ISO already decided to employ the wavelet transform based JPEG2000 as the de facto compression standard for biometric passports, the proposed research work mainly investigate the feasibility of achieving face recognition directly into JPEG2000 compressed domain and recommending a new algorithm for classification (i.e. recognition) of compressed images.

2. Related Work

In a hybrid approach to handle face recognition performed by Chien and Wu [9] multi resolution wavelet transform is applied to extract feature vector. To calculate the

Manuscript received May 5, 2012
Manuscript revised May 20, 2012

approximation band to be used in recognition algorithms, two wavelet decompositions are applied. The performance was marginally improved when compared to the traditional PCA method.

Ekenel and Sankur [10] in their studies examined multi resolution analysis to face recognition. This is done by applying wavelet decomposition using Daubechies wavelet 4 on standard face recognition algorithms PCA and ICA. The main purpose of the research was to identify the subbands which are least sensitive to expression and illumination changes. Good results are observed for images with varying illumination and no major changes for images with varying expressions when DWT coefficients were applied.

The feasibility of achieving face recognition in compressed domain is studies by K. Delac in [11] [12] and the authors conducted experiments on standard recognition systems like PCA , LDA and ICA. Image decompression was stopped before inverse transformation and the obtained DWT coefficients were used as input to these recognition systems. The results noticed no significant drops in recognition rates when compared with pixel domain. The authors also suggested finding more efficient methods for feature extraction and projection in compressed domain.

Byung-Joo Oh [13] introduced a method to improve the robustness of a face recognition system based on LDA using Radial Basis Function Network. Combination of two compensating classifiers was explored in this study. Using appearance based projection approaches like PCA and LDA, the images were preprocessed and the resultant feature projection was fed as input to Radial Basis Function Network which in turn used a feedforward multi-layer network classifier. The experiments were conducted using images from ORL database and 93.5% recognition rate was achieved.

Canonical Correlation Analysis (CCA) has been used by experts recently as a method which helps to relate sets of observations describing different aspects of appearance. For any two sets of multi-dimensional variables with a few pairs of canonical variables, CCA gives a high dimensional association [14]. Since CCA was intended to describe the association between two sets of data sequences, this method has been extensively used in the fields of pattern recognition and signal processing [15]. The main advantage of CCA is it can deal with two sets of data when compared to Principal Component Analysis (PCA) [16] and Linear Discriminant Analysis (LDA) [17]. Dominic and Robert in [18] implemented a two dimensional face recognition system based on CCA in pixel domain which was tested on 20 objects from FERET database. Recognition rate achieved was superior when compared with standard PCA and LDA based recognition systems.

A new methodology for efficient face recognition in compressed domain was introduced by Arockiasamy and Menila in [19] using 2-dimensional Canonical Correlation Analysis. The decompression was stopped after entropy decoding and the resultant entropy points were fed as input to the new CCA based algorithm. Mode based Matching method was applied by the authors for carrying out image data matching. The recognition rates were greatly improved for images with varying illumination and images which are partially occluded. For image data with varying expression, slight drop in Recognition Rate was observed. The Authors suggested to develop a better algorithm using neural networks for better recognition of images with varying expressions.

3. Proposed Methodology

A novel approach is presented in this research work to achieve face recognition in compressed domain , which uses Wavelet Transform (WT) based image compression/decompression, Canonical Correlation Analysis (CCA) based optimization of feature vector and Cascade Forward Neural Network based algorithm for matching of images.

3.1 Wavelet based image compression/decompression

Wavelet transform is a time- frequency scale transformation tool for data, function and operator and hence is a good method for image compression. An image is represented by a combination of wavelet functions of different locations and scales in wavelet transform. No blocky artifacts are introduced in the reconstructed image after decompression. Also higher compression ratios can be achieved by applying wavelet transform regardless of the amount of compression achieved.

In order to execute image compression, the original image is first transformed using wavelets into transform coefficients, transform coefficients are quantized and then entropy coded. Quantization is controlled by the rate distortion unit to minimize the distortion within the available bit rate. These compression stages have an impact on the quality of the image obtained for a given compression ratio. Hence it is concluded that a better compression performance is gained by a better transform [20].

In decompression process, the data must be entropy decoded, inverse quantized and inverse transformed in order to reconstruct as image. This resultant image will be of poor quality due to information discarded during compression. The figure given above shows the compression/decompression procedure in wavelet transform coding.

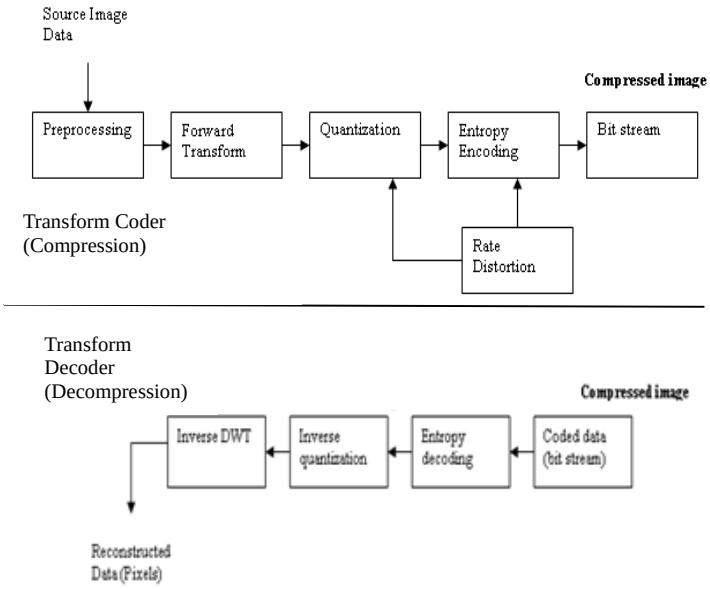

Figure 1: Block diagram of wavelet transform coder/decoder

For the better understanding of our research, it is essential to describe the compressed domain as it plays an important role in this research. Any point before inverse transform in the decoder scheme is considered as compressed domain. Hence compressed domain face recognition can be achieved by stopping the image decompression before inverse transformation stage and then subsequently use the resultant wavelet coefficients as input to recognition algorithms. The proposed methodology greatly saves computational time as the time consuming stages inverse quantization as well as the inverse transformation is avoided. The resultant entropy points were fed as input to the new face recognition system which uses CCA based feature vector optimization.

3.2 Canonical Correlation Analysis (CCA)

Canonical correlation analysis is a suitable and dominant technique which can be used for exploring the relationships among multiple dependent and independent variables. Therefore a powerful feature projection approach for facial images is proposed based on CCA. CCA recognize and measure the relationship between two sets of variables. Mainly CCA finds a pair of linear combinations which has the greatest correlation. For this purpose, CCA uses two sets of linear combination of variables. A pair of linear combinations with greatest correlation among all pairs is determined subsequently. Hence high dimensional association between two sets of variables with a few pairs of canonical variables is represented in by CCA [22].

CCA is extensively applied in various fields especially in pattern recognition and signal processing [23]. CCA as a novel method is applied to image processing and biometrics too. Compared to other projection approaches like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), CCA can concurrently deal with two sets of data.

For two sets of multivariate data, x and y, CCA finds two sets of basic vectors wx and wy so that the correlation is maximized between the projections of variables onto these vectors.

$x = w_x^T x$ and $y = w_y^T y$ where the projections x and y are identified as canonical variables [24]. CCA maximize the function:

$$p = \frac{E w^T{}_x C_{xy} w_x}{\sqrt{w^T{}_x C_{xx} w_x w^T{}_y C_{yy} w_y}}$$

whereby E represents empirical operation. The covariance matrix is a block matrix for two random variables X and Y, where Cxx and Cyy are the within sets covariance matrices of x and y respectively. Also Cxy=CxT is the between-sets covariance matrix [25]. By solving the following eigenvalue equations, the canonical correlations between x and y can be found.

$$p^2 w_x = C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx} w_x$$

$$p^2 w_y = C_{yy}^{-1} C_{yx} C_{xx}^{-1} C_{xy} w_y$$

where eigenvalues p^2 are the squared canonical correlations and eigenvectors w_x and w_y are the normalized canonical correlation vectors [25].

3.3 Cascade Feedforward Network

Neural Network is a computational system based on the structure and functional aspects of biological neural networks and also it simulates the activities of biological neural systems. The most interesting thing about a neural network is its ability to learn and can be implemented in any application without reprogramming. Learning means solving a task best by using a set of observations when a task is given to solve. Neural networks can be trained such a way that an application of set of inputs produces the desired set of outputs. Many training algorithms are available and choice of model depends on the data representation and the application.

Feedforward Backpropagation artificial intelligence model is made up of input layer, hidden layer and output layer. Backpropagation learning algorithm is sued to train a multilayer neural network. For a neural network with a certain set of inputs, this algorithm examines its output response to the sample input pattern given. The output

anticipated is then compared to the output response and the error value is calculated. The connection weights are adjusted based on the error value.
Feedforward Network will learn the relationship between the INPUTs and TARGETs quickly. But since we are using an application for face recognition, many faces seems to like. So Feedforward network finds difficult to train this situation. Compared to feed- forward networks, a weight connection is included in Cascade Feedforward models from input to each layer and from each layer to continual layers.
In the output layer the input is again added as weighted input to improve accuracy. That is the reason we have used Cascade Forward (CF) model in the proposed methodology.

Algorithm used in the proposed methodology (CCA using Neural Networks based matching):

There are two sessions in the process - Training session & Testing session

Training Session:

a. During the training session a set of INPUTS and their corresponding TARGETS are given and a trainable cascade-forward back propagation network is been created using it.

b. Apply CCA to all INPUT images and their feature vectors are taken.

c. Now the Network Should be trained in such way that for the respective CCA feature vector of an image the corresponding TARGET image is directed, since then it will respond the query image given. (In simple training is made in such a way that for the INPUTS images the output should be the corresponding images TARGET images, in our case CCA feature vectors are the INPUTS).

After the training session the Neural Network will reconstruct the image, only if their feature vectors are given as input.

Testing Session:

a. During this Testing Session input image was first single plane separated and resized in to 128*128(i.e 128 rows and 128 Columns)

b. This is applied to Compression process which accomplishes wavelet transformation=>quantization=>Entropy coding and then decompressed

c. During the wavelet transformation only upper left corner block or Quadrant 2 is taken for operation so instead of 128*128=16384 coefficients, only 4096 coefficients are processed ,thus this quadrant contain enough in formation for face recognition operation.

d. Get the image data (as entropy points) from the compressed domain after entropy decoding

e. Then CCA is applied for the above processed Coefficients. this would be Query data

f. Then these coefficients are give to Neural network which we already trained.

g. The output of Neural Network will be the reconstruction of the query image.

h. Now this reconstructed query images is compared with each and every images in database.

i. During this comparison the error between this reconstructed image and the every images database were taken this is known as MSE (Mean Square Error)

j. Finally the Matched image is confirmed by finding image in database which has Minimum MSE from the above calculation.

4. Experimental setup and result analysis

During the Neural Network training, as by procedure some input's and its corresponding targets should be matched in this session. Normally Neural networks will be trained with already taken outputs with its corresponding inputs so that during the testing session that Neural Network will respond to the input given.
In our Project we train the network by the following process:
Step1: Normally set of Images was given as inputs but in our case we using Entropy points as input
Step2: From this Entropy points Plane separation was made (if this plane separation is made before the Entropy point extraction, then it is not needed in this stage)
Step3: For the Target Images the same procedures are made.
Step4: For the Inputs after plane separation, Canonical Correlation Analysis (CCA) was made and canonical coefficient is extracted this acts as Feature vector.
Step5: This Canonical Coefficient is will be in matrix form so we have to convert it into Vector format by reshaping the matrix

Step6: NN trainer needs two inputs to train the network so the Inputs from previous step is given and Target which already Plane separated is also given into the network. NN trainer will train itself by matching the respective inputs and outputs

Testing:

During the testing session the query image or query entropy points will be processed as like the **Inputs** in Training Session

Step1: Query data ie. Entropy points are first plane separated and **Canonical Correlation Analysis** is applied.

Step2: This feature vector will be in matrix form so it will be converted to vector form

Step3: Now the feature vector (Canonical Coefficient) is given in to trained network as input, since this Network is trained the output will be Image data corresponding the input feature vector.

Step4: Now the Output image is compared with the images in the database by checking the MSE (Mean square Error) between each image in the database. The Image which has lowest MSE in database is the matched one.

In the second part of the experiments, we propose labeling of the image in the preprocessing stage before compression. The input image is cropped in such a way to extract only the particular components of the image. By doing this, the background of the image and hair patters are been eliminated and hence the focus only will be in face. After this cropping process, a HIGH BOOST FILTERING is applied to improve the cropped image, so that it can improve the high frequency component in an image to much higher level.

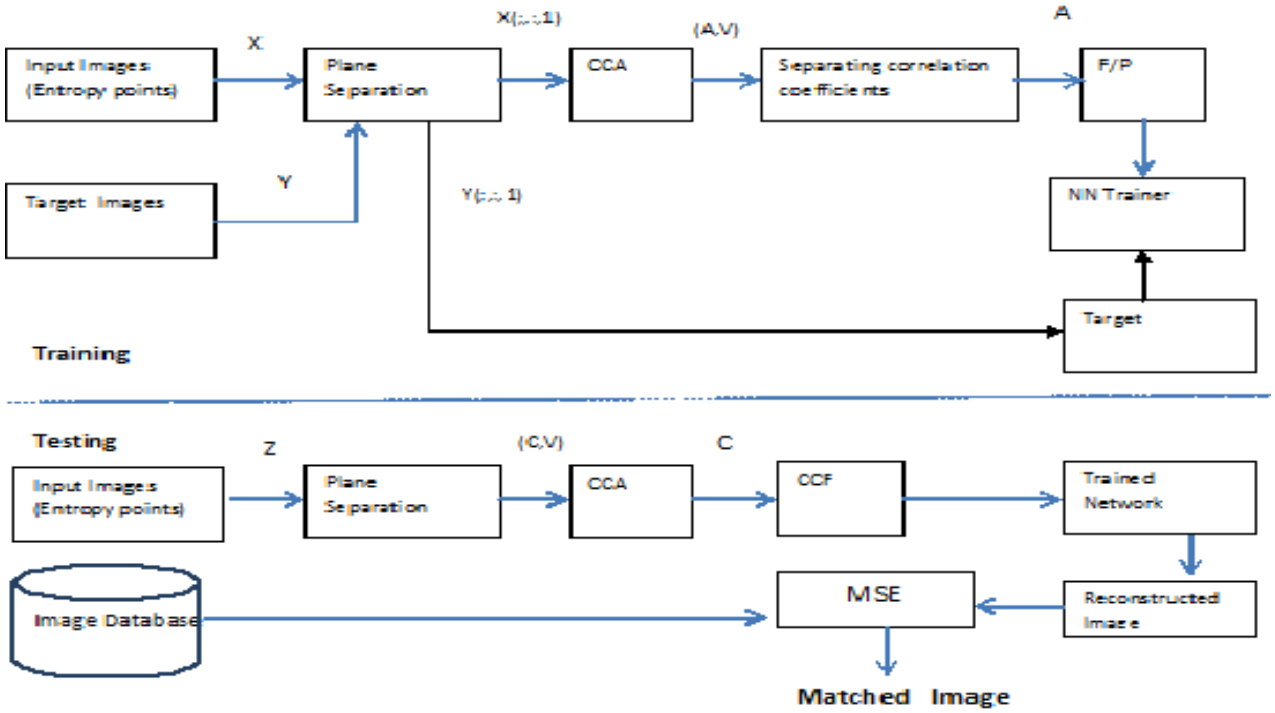

Figure 2: Experimental setup of the proposed method

In the second part of the experiments, we propose labeling of the image in the preprocessing stage before compression. The input image is cropped in such a way to extract only the particular components of the image. By doing this, the background of the image and hair patters are been eliminated and hence the focus only will be in face. After this cropping process, a HIGH BOOST FILTERING is applied to improve the cropped image, so that it can improve the high frequency component in an image to much higher level.

The algorithm for high boost filtering is as follows:

1. Get the Input image, let it be '**L**'

2. Apply Gaussian filtering for the input image(Gaussian filtering is a type Low Pass Filter which eliminates all high frequency components in an image, so image look like blurred one since only low frequency components are present) and let it be '**b**'

3. Now subtract this Blurred image from Input image, so the output will be only high frequency components 'M= **L-b**'

4. Now again add this high frequency image with Input image **H=L+(R*M)** where
R==> high frequency improvement factor and
H was High Boost filtered image
Because of high boost filtering, the weak high frequency components are boosted up and extraction of high frequency components of that cropped image is made.

5. Only this high frequency components of a cropped image is applied for Compression on the next step.

Numerous experiments were conducted using different image data sets and a comparison of results of the proposed CCA and Neural Network based classification method with different recognition systems is presented in Table 1. Comparatively better results are observed for the new method for images with varying illumination and partial occlusion in the first set of tests. The Recognition Rate of images with varying expressions is also slightly improved by applying labeling of images in the second part of the experiments.

Table 1: Comparison of Recognition Rates in compressed domain for various image data sets

Feature Projection Technique	Original image (Pixel Domain)			1bpp			0.5bpp			0.3bpp			0.2bpp		
	fb	fc	fd	fb	fc	fd	fb	fc	fd	fb	fc	fd	fb	fc	fd
PCA (Euclidean matching)	79.4	47.9	38.5	77.8	49.0	37.1	79.0	50.0	38.2	79.5	52.6	38.1	80.6	53.6	36.6
ICA (Euclidean matching)	83.0	68.6	44.3	83.0	68.0	42.9	82.8	67.5	43.5	83.8	65.5	43.5	83.6	57.5	38.9
CCA with Mode Based Matching	78.6	83.0	59.0	78.0	81.0	55.0	77.4	76.0	51.0	74.2	71.3	49.7	77.0	74.1	48.4
CCA with NN	64.91	70.18	50.88	79.95	81.46	59.65	78.72	82.70	61.40	78.19	85.96	56.1	80.70	86.72	57.89
CCA with NN and labeling	71.93	78.95	52.76	80.44	82.70	61.32	82.46	84.21	63.16	79.95	86.21	57.86	81.13	87.24	58.73

Evaluation of Normalized Recognition Rate (NRR) of the proposed method is also observed for various compression rates and the results are compared in Figure 2. It is clearly evident from the figure that the new CCA based method is superior when compared to the recognition rate in spatial domain.

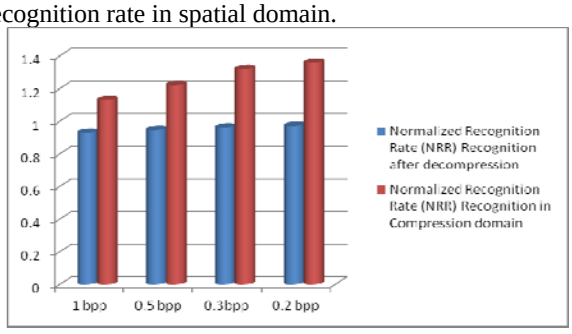

Figure 2: NRR evaluation of the proposed method (CCA +ANN + Labeling)

Comparison of computational time of the proposed method with former recognition algorithms which uses different matching techniques is shown in Figure 3. A significant reduction in computational time is observed in the proposed method with a comparable recognition rate.

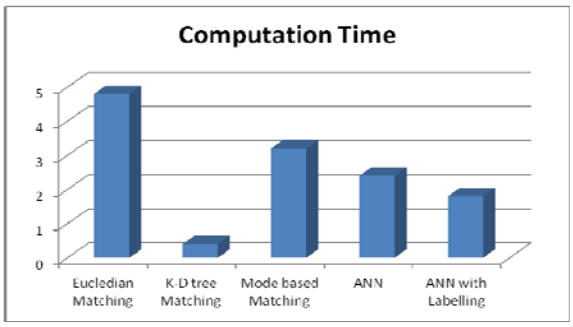

Figure: 3 Comparison of computational time

5. Conclusion

In this research work, a new methodology for performing face recognition in JPEG2000 compressed domain is presented. We have introduced a technique for extracting feature vector from JPEG2000 compressed images and hence the major computationally expensive part of decompression scheme is avoided. Also a new algorithm based on the CCA approach is proposed for feature vector optimization. Cascade Forward back propagation Neural Network is used in the new method for matching of the images. Labeling of images is applied in the lateral part of the tests for better recognition of images with varying expressions.

Numerous experiments were conducted and the obtained results confirmed that the new approach achieves high recognition rate and considerable reduction in computational time. Significant improvement in recognition rate is observed for images with varying illumination and also for partially occluded images. For image data set of varying expression also a slight improvement in recognition rate is observed when compared with results obtained for the approach of CCA with mode based matching.

Future research will include finding a method for extracting feature vectors from compressed image therefore completely excluding decompression of images completely in face recognition. Also a better algorithm which will further enhance the recognition rate of images with varying expression is another interested area of further research.

Acknowledgment

The authors would like to acknowledge the suggestions received from all experts which in turn lead us in the right pathway in this research.

References

[1] S.Z. Li and A.K. Jain (Eds.), Handbook of Face Recognition, Springer, New York, USA, 2005.

[2] Dominique Ginhac, Fan Yang, Xiaojuan Liu, Jianwu Dang and Michel Paindavoine, “Robust Face Recognition System Based on a Multi-Views Face Database”

[3] W. Zhao, R. Chellappa, A. Rosenfeld and P.J. Phillips, “Face Recognition: A Literature Survey”, ACM Computing Surveys 35 (4) 399–458, 2003.

[4] K. Delac and M. Grgic (Eds.), “Face Recognition”, I-Tech Education and Publishing,Vienna, 2007.

[5] Ekenel, H.K.; Stallkamp, J. Gao, H. Fischer and M. Stiefelhagen, "Face Recognition for Smart Interactions", IEEE International Conference on Multimedia and Expo, Pp: 1007-1010, 2007

[6] D.P. McGarry, C.M. Arndt, S.A. McCabe and D.P. D’Amato, “Effects of Compression and Individual Variability on Face Recognition Performance”, Proc. of SPIE 5404, 362–372, 2004.

[7] S. Rakshit and D.M. Monro, “An Evaluation of Image Sampling and Compression for Human Iris Recognition”, IEEE Trans. on Information Forensics and Security 2 (3) ,605–612, 2007.

[8] Harguess and J. Changbo Hu Aggarwal, J.K. Dept. of ECE, Univ. of Texas at Austin, Austin, TX, USA, "Fusing face recognition from multiple cameras", Workshop on Applications of Computer Vision (WACV), Pp: 1-7, 2009.

[9] Chien J.T. and Wu C.C., “Discriminant Waveletfaces and Nearest Feature Classifiers for Face Recognition”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.24, No.12, December 2002, pp. 1644-1649

[10] H.K. Ekenel and B. Sankur, “Multiresolution Face Recognition”, Image and Vision Computing 23 (5) 469-477, 2005.

[11] Delac. K. , Grig. M. and Grgic. S., “ Towards Face Recognition in JPEG2000 Compressed Domain”, Proc. Of the 14th International Workshop on Systems, Signals and Image Processing (IWSSIP), pp. 155-159, 2007.

[12] K. Delac, M. Grgic and S Grgic, “ Face Recognition in JPEG and JPEG2000 compressed domain”, Image and Vision Computing 27, 1108-1120, 2009.

[13] Byung-Joo Oh , “Face Recognition using Radial Basis Function Network based on LDA”, PROCEEDINGS OF WORLD ACADEMY OF SCIENCE, ENGINEERING AND TECHNOLOGY VOLUME 7, ISSN 1307-6884, AUGUST 2005.

[14] Sun Ho Lee and Seungjin Choi, “Two-Dimensional Canonical Correlation Analysis”, IEEE Signal Processing Letters, VOL. 14, NO. 10, 2007.

[15] H. Hotelling, “ Relations between two sets of variates”, Biometrika 28, 321, 1936.

[16] M. Turk, and A. Pentland, “Eigenfaces for Recognition”, Journal of Cognitive Neurosicence 3 (1) 71–86, 1991.

[17] M.S. Bartlett, J.R. Movellan and T.J. Sejnowski, “Face Recognition by Independent Component Analysis”, IEEE Trans. on Neural Networks, Vol 13 (6), Pages 1450– 1464, 2002.

[18] Dominik Jelsovka, Robert Hudec, and Martin Breznan, “ Face Recognition on FERET Face Database Using LDA and CCA Methods”, 978-1-4577-1411-5/11 , IEEE 2011.

[19] S. Arockiasamy, Menila James, “Human Face Recognition in wavelet compressed domain using Canonical Correlation Analysis”, IJCA, MENILA

[20] Uni Grassman and Risto Miikkulainen, “ Effective Image Compression using Evolved Wavelets”, GECCO’05, ACM 1-59593-010-8/05/0006, 2005.

[21] K. Delac, M.Grgic and Grig M, “ Towards Face Recognition in JPEG2000 Compressed Domain, Proc . Of the 14th International Workshop on Systems, Signals and Imag Processing (IWSSIP), pp. 155-159, 2007

[22] Sun Ho Lee and Seungjin Choi, “ Two-Dimensional Canonical Correlation Analysis”, IEEE Signal Processing Letters, VOL. 14, N0.10, 2007.

[23] M. Borga. “Canonical correlation a tutorial”, http://www.imt.liu.se/~magnus/cca/tutorial/ tutorial.pdf, January 12, 2001

[24] W. Yang, Z. Lei and J.Sang, “2D-3D face modeling using CCA method, National Laboratory of Pattern Recognition Institute of Automation , Chinese Academy of Sciences, Beijng, 2008

[25] G. Kukarev, E.Kamenskaya, "Application of two-dimensional canonical correlation analysis for face image processing and recognition", Pattern Recognition and Image Analysis, 2010.

Dr.S.Arockiasamy, Head, College of Information Systems, University of Nizwa. He received MSc and MPhil in 1987 and 1994 respectively. He received PhD. degree from Bharathiar University in 2007. Dr. Arockiasamy is specialized in Applications of Image processing. He has published a considerable number of Research papers and articles in various leading International journals and International Conferences.

Menila James received the Master of Computer Applications degree from Bharathiar University , India in 1994 and Mphil in Computer Science from Bharathidasan University, India in 2005 and currently a Ph.D. candidate in Computer Science, Karunya University, Coimbatore, India. She is now senior lecturer at Muscat College, Sultanate of Oman. She has published papers in International journals and International conferences.