Face Recognition under Partial Obstruction

Francinei Gomes de Morais
NCE/UFRJ – Universidade Federal do Rio de Janeiro,
Rio de Janeiro, RJ, Brazil
francinei@poli.ufrj.br

Antonio Carlos Gay Thomé
NCE/UFRJ – Universidade Federal do Rio de Janeiro,
Rio de Janeiro, RJ, Brazil
thome@nce.ufrj.br

***Abstract* — a comparative study and proposal of a new approach to the task of face recognition is done in this paper. The proposed approach addresses images with partial obstructions. Eigenfaces, Fisherfaces and Laplacianfaces methods are compared and taken as references. The experiments show that the suggested approach is more robust than the other methods.**

Keywords — Facial Recognition; Census Transform; Eigenfaces; Fischerfaces; Laplacianfaces.

I. INTRODUCTION

In recent years, researches on face recognition methods have received great attention from the scientific community. The first systems were proposed on the years 60, with the work of Bledsoe [1], who developed a method based on geometric distance between feature points of the face. Since then strong advances have been obtained, despite of the major issue, regarding to create systems that can operate robustly in real time, still remains as an important challenging. Since the beginning, various algorithms and variants have been proposed [18] and basically these algorithms can be classified into three major categories [19]: Template Matching; appearance (holistic) detection and local features correlation.

In the Template Matching approach a face or parts of it are represented by one or more templates inside a data base of reference. The recognition is achieved comparing the input face (template) with all faces (templates) stored in the reference data base and selecting the one that shows the largest matching [2, 20]. The major disadvantage of this method is related to the memory requirements.

In appearance based method, the image is transformed into a vector that is seen as a point in a high dimensional space [3]. The goal on this approach is to represent the vector in a reduced dimensional space, in which images from different people may form different groups, facilitating the task of classification. Principal Components Analysis (PCA) [4] and Linear Discriminant Analysis (LDA) [5] are two of the most popular techniques for the dimensional reduction purpose. Researches show that changes in lighting as well as in facial expression have great impact on the nonlinearity of the space domain of the faces and that different regions of the face have different importance for the recognition [21, 23]. In [6] is proposed an efficient method to treat with the nonlinearity this space that is called *Laplacianfaces* and is based on Local Preserving Projection technique (LPP).

The work of Wiskott et al [7], called Elastic Bunch Graph Matching (EBGM), had a great influence on local features-based approach. This study represents the face as a graph, with nodes positioned at some specific points (*fiducial points*), for example, eyes, mouth, nose. These points are extracted through Gabor wavelet [22], with different orientations and scales; they are called "jets". The edges of the graph are labeled with the distance between the nodes. Given a set of faces to be used for training the classifiers, a corresponding set of graphs are constructed, one for each face. The recognition process consists of seeking the graph that best fits the image to be recognized. The major disadvantage of this approach is the difficulty to automatically locate and select the feature points.

In general, the majority of the algorithms proposed in the literature put their focus only on problems considering frontal faces and changes in illumination; mustaches; eye glasses and facial expressions. Few or almost nothing can be seen addressing the problem of recognizing faces partially obstructed [24].

In this paper we describe a study realized with three of the most well-known holistic (appearance) approaches: *Eigenfaces*, *Fischerfaces* and *Laplacianfaces*, where they are applied over an artificially created problem simulating partial obstructions on faces extracted from Yale, ORL and PIE databases [8 – 10]. We also propose an approach to deal with this kind of problem and compare the performance of the four methods considering different degrees of obstruction (0%, 12.5%, 27.5% and 50%).

II. HOLISTIC ALGORITHMS

A. Eigenfaces

Eigenfaces [4] is the first method for face recognition proposed based on appearance (holistic). The method uses Principal Component Analysis - PCA for dimensionality reduction and Euclidean distance for classification. Below is a brief description of the main steps of the method:

Creation of the reference base:

- ✓ Step 1 - Selection of the training set.
  - Consists on the selection of a set of M images $A=[I_1, I_2 \cdots I_M]$, each one of the size NxN. The training set must contain images from all individuals under interest and ideally more than one image for each individual;
- ✓ Step 2 - Construction of the Image Vectors Matrix.
  - Each image from A is converted to grayscale and transformed into a column vector Γ_i with length of N^2 and form $B=[\Gamma_1 \quad \Gamma_2 \quad . \quad . \quad . \quad \Gamma_M]$;
- ✓ Step 3 – Creation of the Mean Vector.
  - $\Psi=[\Psi_1, \Psi_2, \cdots, \Psi_{N^2}]^T$. Where Ψ_i is the average value of the i-th line of the matrix B;
- ✓ Step 4 – Creation of the training set matrix.
  - Normalize B using Ψ, $x_i=\Gamma_i-\Psi$ where $X=[x_1 \quad x_2 \quad \cdots \quad x_M]$;
- ✓ Step 5 – Reduction of the dimensionality.
  - Calculate the eigenvectors transformation matrix W by applying PCA on the set X. The columns of the matrix W are called *Eigenfaces*. $Y=W^T X$;
- ✓ Step 6 – Construction of the reference base.
  - Select all vectors in Y that belongs to the same individual, compute the average of these vectors and include each average vector into the reference base matrix Ω_{ref}.

Recognition

- ✓ Step 1 – Formatting the current image.
  - The image I to be recognized is converted to grayscale and transformed into a column vector Γ_i;
- ✓ Step 2 – Formatting the input vector.
  - Construct the input $\vec{x}=\Gamma-\Psi$;
- ✓ Step 3 – Reducing the dimensionality.
  - Construct the final input vector $\vec{y}=W^T\vec{x}$;
- ✓ Step 4 – Computing the Euclidean distance.
  - Calculate the Euclidean distance between $\vec{y}$ and all vectors in Ω_{ref} (the reference base);
- ✓ Step 5 – Selecting the winner vector.
  - The input image is classified as belonging to the same group of individuals that present the lowest distance in step 4, i.e., the most similar.

B. *Fischerfaces*

Fischerfaces originally proposed by [5] is also a holistic method as *Eigenfaces*. The basic difference between *Fischer* and *Eigenface* methods is that the first applies a Linear Discriminant Analysis [11] over the reference base generated by the second. The main goal of Belhumeur was to create a fast algorithm with low computational complexity and capable to perform the task of recognition given a controlled environment, i.e., with images containing predominantly the face of a person, but with higher variations on illumination, expression and rotation than those addressed on the *Eigenfaces* approach.

The algorithm starts from the $X=[x_1 \quad x_2 \quad \cdots \quad x_M]$ matrix that is generated in the same way it is in the *Eigenfaces* algorithm. Considering now that in X there are more than one vector belonging to the same person, it is possible to cluster them in order to form different groups, one for each person, as represented by $X=[X_1 \quad X_2 \quad \cdots \quad X_C]$.

The next step consists on the generation of the transform matrix called W_{LDA} (equations 2 – 5) following the application of it in order to generate de reference base Ω_{ref} (equation 1).

$$\Omega_{ref}=W_{opt}^T \bullet X \tag{1}$$

$$W_{opt}^T=W_{LDA}^T \bullet W_{PCA}^T \tag{2}$$

$$W_{LDA}=\arg\max_W \frac{W^T S_B W}{W^T S_W W} \tag{3}$$

$$S_B=\sum_{i=1}^{C} N_i(\mu_i-\mu)\cdot(\mu_i-\mu)^T \tag{4}$$

$$S_w=\sum_{i=1}^{C}\sum_{x_k \in X_i}(x_k-\mu_i)\cdot(x_k-\mu_i)^T \tag{5}$$

Where:

S_B – is the dispersion matrix among all groups;

S_W – is the dispersion inner groups;

N_i – is the number of samples in the i-th group;

μ_i – is the average vector of samples of i-th group;

μ – is the average vector for all samples.

C. *Laplacianfaces*

Many authors advocate that algorithms based on LDA are stronger than those based on PCA, but in [12] they show that when the training set is small, PCA outperforms LDA and also that in these conditions PCA is less sensitive to variations in the set of training.

Following the reasoning that an image can be viewed as a point in a space of equivalent dimension to their size so, a set of images is nothing more than one set of points distributed on this high dimensional space. Works like [3, 13, 14] aimed to show that if it was possible to visualize the surface described by points in this high dimensional space, we would see that the points form a nonlinear surface, indicating that linear techniques of dimensionality reduction such as PCA and LDA would not work well in these cases.

Laplacianfaces is another holistic method. It is based on PCA (Principal Components Analysis) and LPP (Locally Preserving Projections). It is originally proposed by [15], with the goal of providing a method to reduce the dimensionality of images faces, modeling the possible nonlinearities of the space of faces and preserve its structure.

As in the case of *Fischerfaces*, the algorithm computes the reference database and the optimum transform matrix as in equations 6 and 7.

$$\Omega_{ref} = W_{opt}^T \bullet X \quad (6)$$

$$W_{opt}^T = W_{LPP}^T \bullet W_{PCA}^T \quad (7)$$

The LPP transform matrix is computed through three steps as following:

- ✓ Step 1 – Construction of the graph of "p" nearest neighbors of each image.
  - o The graph links the images to generate a mesh that models the space of faces and carry with them the location information. During the construction process of the graph, two points, or rather two images, are connected by an edge only if the two points x_i and x_j belong to the set of "p" neighbors of each other.
- ✓ Step 2 – Selection of weights for the edges.
  - o The weight of the edge that connects points x_i and x_j is given by equation 8 as follows:

$$S_{ij} = e^{-\frac{\left\| x_i - x_j \right\|^2}{t}} \quad (8)$$

Where:

S – is the matrix that models the space of faces;

t – is a constant value; and S = 0 whenever no edge connects x_i to x_j .

- ✓ Step 3 – Computation of W_{LPP} .
  - o This computation deals with the generalized eigenvectors **[**16**]** for solving the equations 9 and 10.

$$XLX^T w_i = \lambda XDX^T w_i \quad (9)$$

$$W_{opt} = \arg\max_W \frac{W^T XLX^T W}{W^T XDX^T W} \quad (10)$$

Where:

$D_{ii} = \sum_j S_{ji}$ a diagonal matrix;

$L = D - S$ the Laplacian matrix;

W_{LPP} is then obtained ordering W_{opt} by the increasing values of λ ($0 \leq \lambda_0 \leq \cdots \leq \lambda_{k-1}$).

III. Proposed Approach

The algorithm proposed here, named Eigenblocks, makes use of the Census Transform [17], split the original image into four sub-images and applies the Eigenfaces method over each of these sub-images. As a final step, there is a decision module based on the results obtained on each of the sub-images. Figure 1 shows the Eigenblocks scheme.

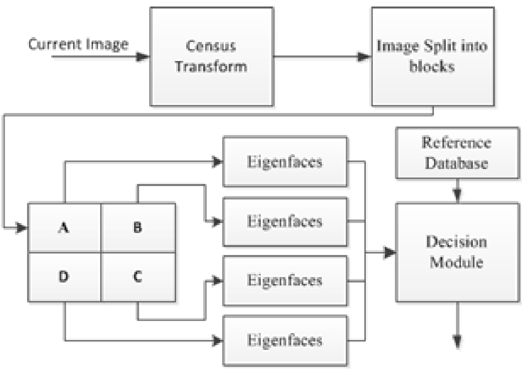

Fig. 1. Eigenblocks Diagram.

The decision module computes the recognition for each block of the input image and uses a majority like decision rule to provide the final response.

The decision rule used in this work selects the winner label based on a majority scheme as follow:

- ✓ Each block presents its own answer;
- ✓ If there is a label with the majority of votes than it will be the answer of the system;
- ✓ In case of two draw results the answer of the system will be the label with smallest join Euclidean distance (summation of all blocks with the same label);
- ✓ Otherwise, the system rejects the classification.

A. *Census Transform*

In the Eigenfaces method the percentage of accuracy tends to decrease with the increasing nonlinearity of the space of faces. In view of this, it is common to apply techniques such as histogram equalization and Discrete Cosine Transform (DCT) [25] to minimize the effects of variation in brightness, once this variable exerts a strong influence on the nonlinearity of the space of faces.

The Census Transform (CT) is a non-parametric local transformation that tries to reduce the brightness changing influence through a different approach. It was initially proposed by [17]. The operation is performed by comparing the luminance of the pixels of a given region with the luminance of the central element of this region. Formally, the Census Transform defines square regions, where the most common sizes are of the order of 3x3 and 5x5. The output of the CT is a bit string where each bit is calculated by equation

11. Finally the central pixel new value is given by the conversion of the bit string in decimal.

$$C(x) = \bigotimes_{y \in N} \zeta(I(x), I(y)) \tag{11}$$

Where:

$I(x)$ – is the central pixel of the squared neighborhood;

$I(y)$ – is any pixel of the squared neighborhood;

$\zeta(I(x), I(y))$ – is the comparison function, $\zeta = 1$ if $I(y) > I(x)$ and $\zeta = 0$ otherwise;

$\otimes$ – denotes the concatenation operation;

$C(x)$ – is the kernel function, with null center, defined over a region $N(x)$.

Figure 2 shows a subset of all local features that can be represented by the Census Transform from 3x3 regions. Such matrices are called Kernels of Census Transform.

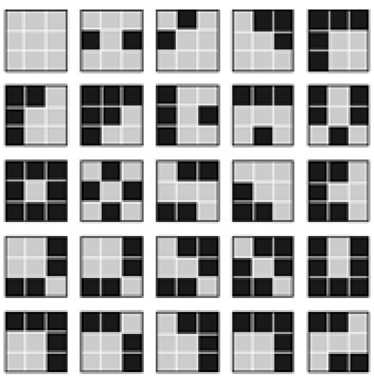

Fig. 2. Possible 3x3 Census Transform Kernels.

Figure 3 shows the obtained results when applying Census Transform on two images of the same person that were taken in different lighting conditions. Notice that although the illumination varies significantly in the original images, the images resulting from the Census Transform remain quite similar.

Fig. 3. Census Transform – example of invariance to illumination.

IV. Experiments and Results

To validate the proposed approach, we compared its performance against the previously described holistic methods: *Eigenfaces*, *Fischerfaces* and *Laplacianfaces*. Several experiments were performed with the use of three publicly available face databases and widely used: Yale [8]; ORL [9] and CMU PIE [10].

A. Face Databases

The Yale database was constructed at the Center for Computational Vision and Control of Yale University. It contains 165 images of 15 persons, in other words, 11 images per each person. The face images are frontal and present variations in illumination, facial expression (happy, sad, winking, sleepy and surprise) and persons wearing or not glasses.

The training set used in these experiments was formed with seven images per person, chosen at random. The test set was formed with the four remaining images per person. The images, in grayscale, were normalized in terms of scale and orientation, so that the eyes were horizontally aligned and then, they were cropped to contain only the internal structures of the face. Finally they were resized to the dimensions 128x128, 64x64, 32x32 and 16x16. Figure 4 shows a set of images in its original format and figure 5 shows the same images after these preprocessing steps.

Fig. 4. Examples of Yale original images.

The ORL database contains images of 40 individuals, 10 images per person, taken in different situations. The variations of facial expression are (eyes open, eyes closed, smiling and normal); the facial details are (wearing or not glasses). All images have a black background, their faces are frontal and with small variations in pose and orientation, a maximum of 20 degrees. The original resolution is 92x112 with 256 grayscale levels. Some examples are shown in figure 6.

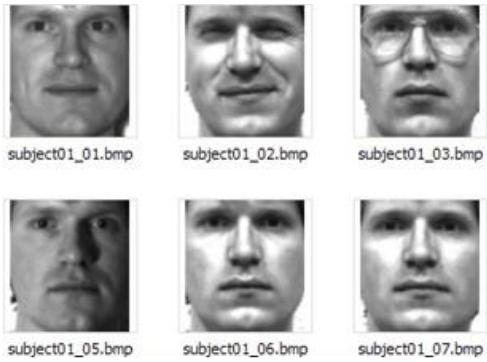

Fig. 5. Same images after normalization steps.

Fig. 6. Example of ORL images, (a) original ones and (b) after preprocessing steps.

The CMU PIE database (Pose, Illumination, and Expression) is composed of 68 individuals, totaling 41,368 images that were recorded by 13 cameras in 21 different conditions of flashes. The subset used for the experiment was formed by 20 individuals, 24 images per person, all frontal and normal facial expression. Each one of the 24 images represents a different illumination condition. The purpose on this set of images was to be able to evaluate the behavior of the algorithms when addressing only changes in the illumination and keeping fixed all the others variables, such as facial expression, rotation and orientation.

The training and test pairs were (12,12), (10,14), (8,16), (6,18) and (4,20), where the first coordinate represents the number of images per individual used in the training set and the second coordinate shows the number used in the test set. Figures 7 and 8, show examples of the PIE database images in their original and preprocessed formats.

Fig. 7. PIE database, original images.

Fig. 8. PIE database, preprocessed images.

B. Obstruction Experiments

We also performed a set of tests considering two types of partial obstructions of the face. The partial obstructions were artificially generated over the original images as shown in figures 9 and 10.

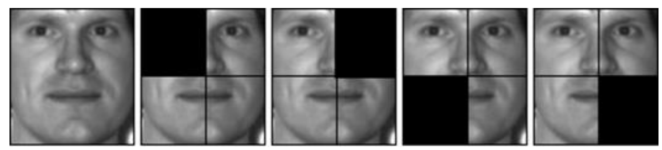

Fig. 9. Partial obstruction of an entire quadrant.

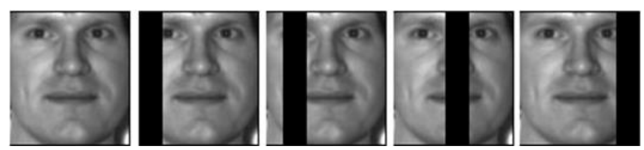

Fig. 10. Partial obstruction, simulated by a black stripe dislocating from the left to the right. We used four different sizes for the black stripe: 12.5%, 25%, 37.5% and 50% of the image's wide.

C. Experiments and results

The following three tables show the results using different quadrant obstructions applied over the three face's databases: Yale, ORL and PIE.

TABLE I. RESULTS USING YALE DATABASE IMAGES OF THE SIZE OF 32X32. ORIGINAL MEANS THE IMAGE WITHOUT OBSTRUCTION AND Q1 TO Q4 MEANS THE OBSTRUCTED QUADRANT

Yale 32x32	**Original**	**Q1**	**Q2**	**Q3**	**Q4**
Eigenfaces	78.3	36.7	43.3	26.7	53.3
Fisherfaces	68.3	33.3	28.3	30.0	60.0
Laplacianfaces	**96.7**	51.7	58.3	36.7	61.7
Eigenbloks	76.7	**73.3**	**73.3**	**68.3**	**66.7**

TABLE II. RESULTS USING ORL DATABASE IMAGES (20 INDIVIDUALS IN THE REFERENCE BASE)

ORL-32x32/20P	**Original**	**Q1**	**Q2**	**Q3**	**Q4**
Eigenfaces	86.0	21.0	27.0	22.0	30.0
Fisherfaces	68.0	15.0	28.0	18.0	44.0
Laplacianfaces	**99.0**	31.0	46.0	45.0	33.0
Eigenbloks	92.0	**81.0**	**83.0**	**79.0**	**78.0**

TABLE II. RESULTS USIING PIE DATABASE IMAGES (12 IMAGES PER INDIVIDUAL FOR CONSTRUCTING THE REFERENCE BASE AND 12 FOR TESTING)

PIE 32x32/12_12	Original	Q1	Q2	Q3	Q4
Eigenfaces	95.8	58.8	77.9	67.9	65.8
Fisherfaces	98.3	32.9	74.6	29.6	21.3
Laplacianfaces	**100.0**	62.9	90.8	64.6	80.8
Eigenbloks	98.8	**92.1**	**92.1**	**97.5**	**95.8**

Next tables show the results obtained when using the black stripe of different sizes (12.5%, 25%, 37.5% and 50% of the image's wide) dislocating pixel by pixel from the left to the right.

TABLE IV. YALE DATABASE

Yale 32x32	Original	12.5	25.0	37.5	50.0
Eigenfaces	78.3	64.7	40.7	23.7	18.4
Fisherfaces	68.3	20.0	15.0	16.7	8.3
Laplacianfaces	**96.7**	**88.3**	63.3	48.3	30.0
Eigenbloks	76.7	70.2	**64.1**	**57.9**	**44.7**

TABLE V. ORL DATABASE

ORL-32x32/20P	Original	12.5	25.0	37.5	50.0
Eigenfaces	86.0	46.1	24.0	13.5	9.6
Fisherfaces	68.0	17.0	3.0	0.0	5.0
Laplacianfaces	**99.0**	51.0	24.0	11.0	5.0
Eigenbloks	92.0	**80.4**	**72.5**	**61.1**	**43.9**

TABLE VI. PIE DATABASE

PIE 32x32/12_12	Original	12.5	25.0	37.5	50.0
Eigenfaces	95.8	88.9	73.0	46.6	24.3
Fisherfaces	98.3	46.7	33.8	17.9	15.4
Laplacianfaces	**100.0**	95.0	**93.8**	**90.8**	**77.5**
Eigenbloks	98.8	**98.4**	83.9	67.0	41.5

V. CONCLUSIONS

As can be seen, Eigenblocks improved the performance of the original *Eigenfaces* in almost all cases, despite the face be or not partially obstructed. The proposed approach also provided better results than *Fischerfaces* and *Laplacianfaces* considering the cases where the face is under partial obstruction. Eigenblocks showed to be less sensitive to the types of obstruction tested in these experiments.

As future work, the authors intend: 1) replace the Census' Transform by other illumination normalization technique like DCT or homomorphic filter [26]; 2) replace the *Eigenface* by another holistic method such as the *Laplacianface*; 3) test real obstruction patterns (such as trees and poles) and finally; 4) build a system that check if the face image is obstructed or not, and then try to select the most adequate algorithm.

REFERENCES

[1] W. Bledsoe. “Man-Machine Facial Recognition”. Technical Report PRI 22, Panoramic Research Inc, Palo Alto, CA, 1966.

[2] R. Brunelli et al., “Face recognition: Features versus templates”, PAMI, 1993, 15(10): 1042-1052.

[3] H.S. Seung and D.D. Lee, “The Manifold Ways of Perception”. Science, vol. 290, Dec. 2000.

[4] M. Turk and A.P. Pentland, “Face Recognition Using Eigenfaces”. IEEE Conf. Computer Vision and Pattern Recognition, 1991.

[5] P.N. Belhumeur, J.P. Hespanha, and D.J.Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection”. IEEE Trans. Pattern Analysis and Machine Intelligence, 19(1997)7.

[6] Xiaofei He, Shuicheng Yan, Yuxiao Hu, Partha Niyogi, and Hong-Jiang Zhang, “Face Recognition Using Laplacianfaces”. IEEE Trans. Pattern Analysis and Machine Intelligence, 27(2005)3.

[7] L. Wiskott, J.M. Fellous, N. Kruger, and C.v.d. Malsburg, “Face Recognition by Elastic Bunch Graph Matching”. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, pp. 775-779, 1997.

[8] Yale Database: http://cvc.yale.edu/projects/yalefaces/yalefaces.html

[9] ORL Face Database : http://www.codeforge.com/article/109886

[10] PIE Face Database: http://vasc.ri.cmu.edu/idb/html/face/index.html

[11] Duda, R. O.; Hart, P. E.; Stork, D. G. “Pattern Classification”. New York: Wiley Interscience, 2001.

[12] MARTINEZ, A. M.; KAK, A. “PCA versus LDA”. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2001.

[13] Chang, Y.; Hu, C.; Turk, M. “Manifold of Facial Expression”. In: Proc. IEEE Int’l Workshop Analysis and Modeling of Faces and Gestures, October 2003.

[14] Roweis, S. T.; Saul, L. K. “Nonlinear Dimensionality Reduction by Locally Linear Embedding”. Science, 2000, v.290, pp. 2323-2326.

[15] Xiaofei, H.; Shuicheng, Y.; Yuxiao, H.; Niyogi, P. et al. “Face Recognition Using Laplacianfaces”. IEEE Trans. Pattern Anal. and Mach. Intell., v.27, n.3, pp. 328-340, Mar.2005.

[16] He, X.; Niyogi, P. “Locality Preserving Projections”. Advances in Neural Information Processing Systems, 2003.

[17] Ramin Zabih and John Woodfill. “A non-parametric approach to visual correspondence”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996.

[18] Dandpat, S.K.; Meher, S., "Performance improvement for face recognition using PCA and two-dimensional PCA," Computer Communication and Informatics (ICCCI), 2013 International Conference on , vol., no., pp.1,5, 4-6 Jan. 2013

[19] Teja, G.P.; Ravi, S., "Face recognition using subspaces techniques," Recent Trends In Information Technology (ICRTIT), 2012 International Conference on , vol., no., pp.103,107, 19-21 April 2012.

[20] Lone, M.A.; Zakariya, S. M.; Ali, R., "Automatic Face Recognition System by Combining Four Individual Algorithms," Computational Intelligence and Communication Networks (CICN), 2011 International Conference on , vol., no., pp.222,226, 7-9 Oct. 2011.

[21] Kavita. R. Singh, Mukesh. A. Zaveri and Mukesh. M.Raghuwanshi, “Illumination and Pose Invariant Face Recognition: A Technical Review”, International Journal of Computer Information Systems and Industrial Management Applications, Vol.2 (2010), pp.029-038.

[22] Chao-Chun Liu and Dao-Qing Dai, "Face recognition using Dual Tree complex Wavelet features," IEEE Transactions On Image Processing, vol. 18, no. 11, pp. 2593-2599 Nov 2009.

[23] De Marsico, M.; Nappi, M.; Riccio, D.; Wechsler, H., "Robust Face Recognition for Uncontrolled Pose and Illumination Changes," Systems, Man, and Cybernetics: Systems, IEEE Transactions on , vol.43, no.1, pp.149,163, Jan. 2013.

[24] Shengcai Liao; Jain, A.K.; Li, S.Z., "Partial Face Recognition: Alignment-Free Approach," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.35, no.5, pp.1193,1205, May 2013.

[25] Shermina, J., "Illumination invariant face recognition using Discrete Cosine Transform and Principal Component Analysis," Emerging Trends in Electrical and Computer Technology (ICETECT), 2011 International Conference on , vol., no., pp.826,830, 23-24 March 2011.

[26] Thamizharasi, A.M.E., "Performance analysis of face recognition by combining multiscale techniques and homomorphic filter using fuzzy K nearest neighbour classifier," Communication Control and Computing Technologies (ICCCCT), 2010 IEEE International Conference on , vol., no., pp.643,650, 7-9 Oct. 2010.