Facial Recognition using Modified Local Binary Pattern and Random Forest

Brian O'Connor and Kaushik Roy

Department of Computer Science,
North Carolina A&T State University,
Greensboro, NC 27411

Abstract

This paper presents an efficient algorithm for face recognition using the local binary pattern (LBP) and random forest (RF). The novelty of this research effort is that a modified local binary pattern (MLBP), which combines both the sign and magnitude features for the improvement of facial texture classification performance, is applied. Furthermore, RF is used to select the most important features from the extracted feature sequence. The performance of the proposed scheme is validated using a complex dataset, namely Craniofacial Longitudinal Morphological Face (MORPH) Album 1 dataset.

Keywords:

1. INTRODUCTION

Most current face recognition algorithms perform relatively well in a strictly constrained situation, where well aligned, well illuminated, and frontal pose face images are considered for performance evaluation to ensure the higher performance. Under controlled image acquiring constraints, it is possible to capture high quality images and attain an impressive accuracy with a very low error rate. However, recognition accuracies substantially decrease when the captured images do not have enough quality either due to subject's alignment problem to the camera, various facial expressions, gaze deviations or facial hair [1-5]. Several researchers proposed different facial recognition algorithms that perform well with unconstrained face images [1-3, 5-9]. Recently, the face recognition algorithms based on local appearance descriptors such as Gabor filters, SURF, SIFT, and histograms Local Binary Patterns (LBP) provide more robust performance against occlusions, different facial expressions, and pose variations than the holistic approaches [10, 11]. The LBP operator [10-12] has been regarded as one of the best descriptors for the appearance of local facial regions and it has been widely used in various applications. The LBP-based feature extractor has proven to be highly distinctive and its key advantages, including the robustness against illumination and pose variations, and the computational efficiency, make it suitable for high level image analysis tasks [13, 14]. In [10, 11], LBP operator was employed to extract the discriminative facial features. However, in [10, 11], only the basic LBP operator was used for textural feature extraction and this method may fail to detect the illumination variation

and facial expressions accurately. Addressing the above problem, in this paper, we propose to apply a modified LBP (MLBP), which fuses both the sign and magnitude features, to improve the facial texture classification performance [14]. Though the sign component of LBP operator preserves most of the information of local difference, the magnitude component provides additional discriminant information that enhances the overall recognition accuracy.

Previous work on face recognition has focused mainly on the issues of feature extraction and facial pattern classification [15-17]. However, less effort has been given to the critical issue of feature selection. In machine learning, selecting the best features from the higher-dimensional feature space has several potential benefits, including defying the curse of dimensionality to enhance the prediction performance, reducing the measurement and storage requirements and decreasing the training and prediction times [18]. The selection of the most representative facial features from the original feature set with a relative high dimension is an important issue in the field of face recognition [19-22].The MLBP operator used in this research effort utilizes both the sign and magnitude features, and therefore, lead to an extremely high number of features, most of which are redundant for the purpose of classification. To resolve this issue, we propose a variable selection algorithm to select the salient MLBP features based on Random Forest (RF) [23]. The proposed feature selection scheme using RF reduces the original feature space and speeds up the classification process. Fig. 1 shows the block diagram of the proposed face recognition system.
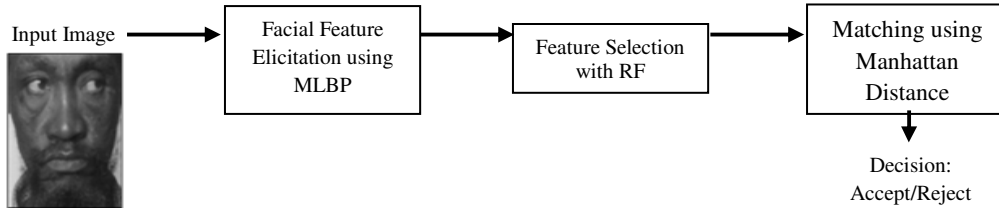

Fig. 1: Block diagram of the proposed face detection system.

The remainder of this paper is as follows. Section 2 provides a brief overview of the MLBP. Section 3 describes the feature selection technique using the RF. In Section 4, we present our experimental results. Section 5 provides the conclusions.

2. MODIFIED LOCAL BINARY PATTERN

An LBP code is measured by comparing a pixel of an image with its neighboring pixels [12, 13]

$$LBP_{P,R} = \sum_{p=0}^{p-1} r(g_p - g_c)2^p, r(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} (1)$$

where g_c denotes the gray level value of the center pixel, g_p represents the value of the neighboring pixels of the center, P is the total number of neighboring pixels and R is the radius of the neighborhood. Let us consider an image of size *I*J*. LBP pattern is computed for each pixel of an image and a histogram is developed to represent the face texture [13]

$$Hist(k) = \sum_{i=1}^{I} \sum_{j=1}^{J} f(LBP_{P,R}(i,j), k), k \in [0, K], (2)$$

$$f(x,y) = \begin{cases} 1, & x = y \\ 0, otherwise \end{cases}$$

where K denotes the maximal LBP pattern value. Now, we calculate the local difference, $Dist_P$, between the center pixel g_c and the evenly spaced neighboring pixels, g_p, $p = 0, 1, 2, ..., P - 1$ as $Dist_P = g_p - g_c$. Thus, we obtain the image local structure at g_c with the local difference vector $[Dist_0, ... , Dist_{p-1}]$. Since the center intensity value, g_c is removed, the local difference vector provides robust performance against the illumination changes. We decompose the $Dist_P$ into two components [14]

$$Dist_P = s_p * m_p \text{and} \begin{cases} s_p = sign(Dist_P) \\ m_p = |Dist_P| \end{cases} (3)$$

where $s_p = \begin{cases} 1, Dist_P \geq 0 \\ -1, Dist_P < 0 \end{cases}$ and m_P are the sign and magnitude of $Dist_P$, respectively. In this paper, we utilize the complementary strengths of the sign and magnitude components of $Dist_P$ in order to improve the texture classification performance. The equation (3) represents the local difference Sign Magnitude Transform (SMT) [14]. To extract the facial textural features, first, we divide the facial image into several patches and apply the MLBP on each patch (see Fig. 2). Each facial patch is represented by 256 sign and 256 magnitude components. We concatenate the sign and magnitude components and present a single patch by 256×2=512 components. We apply MLBP on each patch of a cropped face image. Depending on the number of patches created, the overall features obtained can be dramatically large. For example, a subject's image split into 4×6 patches results in 24 patches. Each one of these patches has 512 features, leading to 12,288 features to represent the facial image. In the following section, we discuss the feature selection approach using the RF.

Fig. 2: Facial image with different patches.

3. RANDOM FOREST AS A FEATURE SELECTOR

RF is a well known and efficient algorithm for both the classification and regression problems [23]. RF is an ensemble learning process which generates many classifiers and aggregates their results. Two well known methods for ensemble learning are boosting and bagging of classification trees. In boosting method, extra weight is provided by the successive trees to the points which are incorrectly predicted by previous predictors. A weighted vote scheme is then considered for prediction. The trees are developed independently using a bootstrap sample of the data set in the bagging method and a simple majority vote is used for prediction purpose [23]. In RFs, an additional layer of randomness is included to the bagging process. RF constructs each tree using a

different bootstrap sample of the data and change the method of the classification or regression trees creation [19]. Unlike the standard trees in which each node is created using the best split among all the variables, RFs split each node using the best among a subset of predictors chosen randomly at the node. The specific size of the subset is a parameter of the forest. This strategy seems to be contradictory, however, performs relatively well compared to other classification techniques, including discriminant analysis, support vector machines and neural networks, and is robust to over fitting [23].

In RFs [19], the largest possible trees are grown without pruning. The root node of each tree contains a different bootstrap sample which is randomly selected from the original training data. The leaves of a tree provide the feature elements of the same class label. The class label of a new data is predicted based on the leaf in which that data lands. The RF algorithm can be summarized as follows [19]:

1. Extract n_t bootstrap samples from the original data.
2. For each n_t, develop an unpruned classification or regression tree. At each node, randomly sample m_t of the predictors and select the best split among those variables.
3. Predict new data by aggregating the predictions of the n_t. The majority votes and average are used for classification and regression, respectively.

In RF, an estimate of the error rate can be measured, based on the training data, according to following:

1. At each iteration of bootstrap process, predict the data which is not in the bootstrap sample, denoted as Out-Of-Bag (OOB) data, using the tree grown with the bootstrap sample.
2. Aggregate the OOB predictions. Measure the error rate, and define it as the OOB estimate of error rate.

RF makes no distinction between the relevance of features during construction of the forest. Each node is created randomly with equal probability. Because of this, feature redundancy may arise and increase the generalization error. These redundant features can produce unreasonably large trees which will produce huge computational load. However, RF avoids this problem by producing a measure of importance of each variable, called variable importance. The algorithm estimates the importance of a variable by looking at how much prediction error increases when OOB data for that variable is modified while all other variables are left unchanged. We decide to use the variables with the highest variable importance to build our RF feature selector. This allows the vectors to be greatly reduced and potentially remove any harmful features that can cause erroneous predictions. The feature selection using RF can be summarized as follows [19]:

1. Select the bootstrap samples of the training data form a RF and estimate the error rate by using OOB predictions.

2. For each tree of the forest, exclude the OOB cases and count the number of votes for the correct class. Randomly permute the values of the first variable in the OOB cases and exclude these cases from the tree. Subtract the number of votes for the correct class in the first-variable-permuted OOB data from the number of votes for the correct class in the untouched OOB data.

3. The average of this number, over all trees in the forest, is regarded as the raw importance score for this variable.

4. Repeat the above step for all variables.

In this research effort, we have selected the LBP and MLBP features with the highest variable importance scores using the RF.

4. EXPERIMENTS AND RESULTS

In this research effort, the Craniofacial Longitudinal Morphological Face (MORPH) Album 1 dataset [24] is used for experiments. This dataset contains the facial images of subjects taken over a long period of time, starting from the 1960s. This dataset consists of males and females, with African, Asian, Hispanic and Caucasian ethnicities. MORPH-1 has 1683 facial images from 631 subjects, seven of those subjects only have one instance, so there is a total of 624 subjects with between 2 to 5 instances. In MORPH Album 1 dataset, most of the instances from each subject are affected by different nonideal factors, including subject's alignment to the camera, facial expression, gaze deviation, drastic age difference from one instance to the next, and facial hair styling. Therefore, these noise factors make it difficult to produce accurate results. Fig. 3 shows a few facial images of the same subject in the MORPH Album 1 dataset that has been preprocessed by cropping.As mentioned in Section II, we divide each subject's image into a multitude of different patch variations and generated both LBP and MLBP feature vectors for each patch. We process every combination of patches from 1x1 to 7x7 and then compare the results of such combination of patches using the Manhattan Distance. While each LBP patch provides 256 features (only sign components), each MLBP patch produces 512 features as we utilize both the sign and magnitude components for experiments. For each feature vector, we test each subject in one-to-many (1:N)and a many-to-one (N:1) setups.In (1:N) setup, one instance of a subject in the probeset is compared to multiple instances of that subject in the gallery set.

Fig. 3: Instances of a subject from MORPH Album 1 dataset.

In this setup, 1059 images are placed in the gallery set and 624 images are placed in the probe set. In the (N:1) setup, multiple instances of a subject of probe set are compared to one instance of that subject of the gallery set. Thus, in this setup, 1059 images are placed in the probe set and 624 images are placed in the gallery set. We analyze the feature extraction results using LBP and MLBP and select the patch configuration with the highest accuracy. The textural features extracted from the selected patch are provided as input to RF in order to reduce the dimensionally of the original feature vectors.

Tables1 and 2 report the top fifteen recognition accuracies for the LBP and MLBP methods with respect to (N:1) setup. It is observed from the tables 1 and 2 that facial feature extraction using

MLBP provides better results than the traditional LBP approach for every patch configuration, since the proposed MLBP uses both the sign and magnitude components. For the (N:1) setup, we can find that the 6×5 patch configuration, with 7680 features, exhibits the best accuracy of 32.96% using the MLBP operator. However, for the same setup of (N:1), we find that highest performance obtained by the LBP operator is 30.97% with the 7680 features. Similarly, with the (1:N) setup, we achieve a reasonable accuracy of 40.71% when the number of features is 21504

Table 1: Best LBP Patch Combinations on MORPH-1.

Patches (W x L)	Number of Features	Accuracy (N:1)	Accuracy (1:N)
6 ×5	7680	30.97%	36.70%
6 ×7	10752	30.69%	38.46%
5 × 6	7680	30.59%	37.66%
5 × 7	8960	30.22%	38.62%
7 × 5	8960	30.22%	37.98%
6 × 6	9216	30.03%	36.70%
7 × 7	12544	29.84%	37.34%
5 × 5	6400	29.65%	37.18%
6 × 4	6144	29.56%	35.90%
7 × 6	10752	28.90%	36.22%
4 × 5	5120	28.61%	37.02%
4 × 6	6144	28.33%	36.06%
4 ×4	4096	28.14%	36.70%
7 × 4	7168	28.14%	36.06%
4 × 7	7168	28.05%	36.22%

Table 2: Best MLBP Patch Combinations on MORPH-1

Patches (W x L)	Number of Features	Accuracy (N:1)	Accuracy (1:N)
6 × 5	**15360**	**32.96%**	40.22%
5 × 6	15360	32.11%	38.78%
6 × 7	**21504**	32.01%	**40.71%**
7 × 6	21504	31.73%	39.10%
7 × 7	25088	31.73%	39.10%
6 × 6	18432	31.54%	39.90%
5 × 5	12800	31.44%	38.94%
6 × 4	12288	31.26%	39.10%
4 × 6	12288	31.16%	37.02%
5 × 7	17920	31.07%	38.78%
7 × 5	17920	30.97%	38.78%
7 × 4	14336	30.59%	37.98%
4 × 5	10240	30.31%	37.98%
4 × 7	14336	29.84%	36.86%
3 × 5	7680	29.65%	36.38%

(6×7 patch configuration) using the MLBP operator. For the same setup, LBP operator achieves a highest accuracy of 38.46% when the facialimage is divided into 5×7 patches with 8960 features.

We select the patch configuration with the highest accuracy using LBP and MLBP and provide the extracted features as input to the RF for the selection ofsalient facial features. Therefore, we input 15360 features (6×5 patch configuration) to the RF with N:1 setup as shown in Fig.4(a). We can see form Fig.4(a)that thefeature selection curve starts to level off when the number of features is 1150 with an accuracy of 32.30% for MLBP. Therefore, it is observed that the proposed RF feature selection technique is able to remove 14210 features from the original feature vector of size 15360 without affecting the recognition performance much. In Fig 4(b), we provide7680 features (6×5 patch configuration) to the RF with the N:1configuration.It is found that feature selection curve starts to flattenwhen the number of features is 972 with an accuracy of 27.10% with the traditional LBP approach.Similarly, we provide 21504 features (6×7 patch configuration) to the RF for 1:N setup and obtain 40.10% accuracy when the number of features is around 6000 as shown in Fig.5(a). Thus, we find that the RF-based feature selection reduces 15504 features from the original feature vector of size 21504 without losing the recognition in a great extent. From Fig. 5(b), we can find that LBL with RF achieves an accuracy of 36.05% at the 2304 features. Therefore, we find from the experiments that the proposed approach obtains a reasonable accuracy for both of the setups.
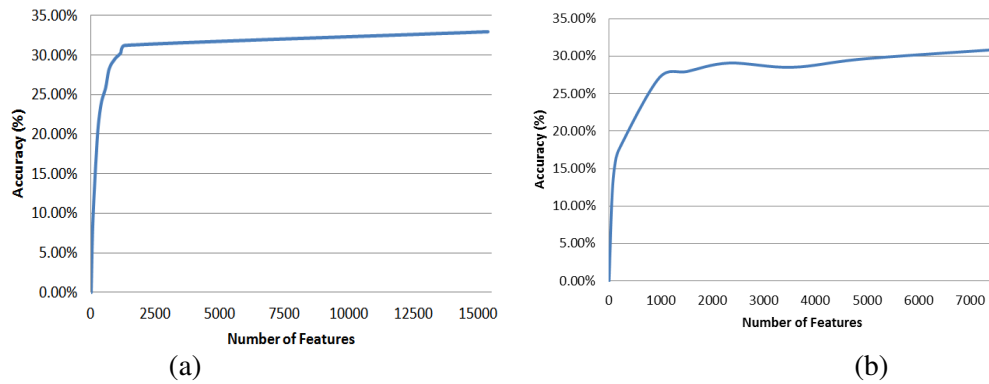

(a)   (b)

Fig. 4: Feature selection using RF (N:1 approach). (a) MLBP,and (b) LBP.

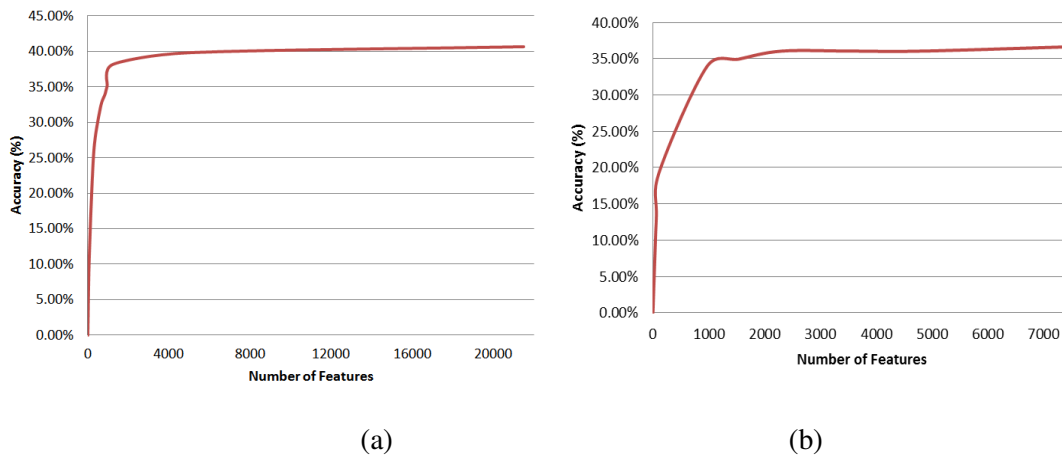(a)   (b)

Fig. 5: Feature selection using RF (1:N approach).(a) MLBP, and (b)LBP.
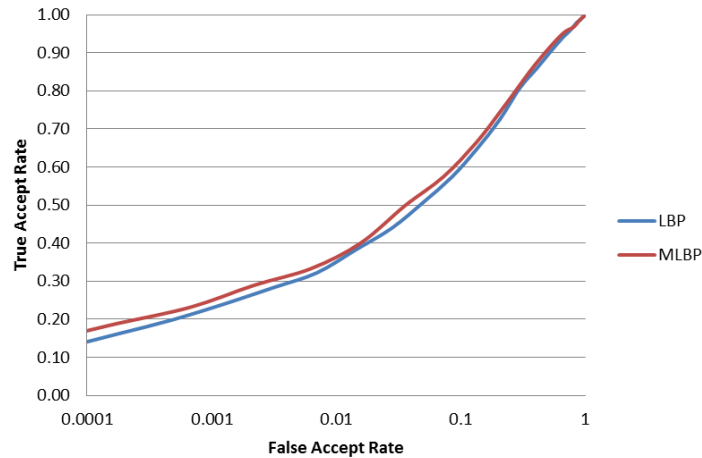
Fig. 6: ROC Curve shows the verification performance of the proposed scheme.

We also measure the performance in terms of *True Accept Rate (TAR)* and *False Accept Rate(FAR)* with the assumption that a test sample comes from a specific subject. From Receiver Operating Characteristic (ROC) curve of Fig.6, we find that the proposed approach achieves relatively a verification rate of 33.10% at the FAR of 0.01. LBP obtains a verification rate of 32.10% at the FAR=0.01.

5. CONCLUSIONS

In this research effort, the MLBP, which combines the sign and magnitude features for the improvement of facial texture classification performance, is applied. We examine the performance of LBP and MLBP for multiple patch variations on a challenging facial dataset. Whereas the traditional approach using LBP for facial recognition can provide accurate results, the proposed approach with MLBP has proven to be consistently more accurate than the traditional LBP. Furthermore, RF-based feature selection method can reduce feature vectors drastically without affecting the recognition accuracy. We validated the proposed iris recognition scheme on the MORPH Album 1 dataset with an encouraging performance.

REFERENCES

[1] J. Wright and G. Hua, "Implicit elastic matching with random projections for Pose-Variant face recognition," Proc. in Computer Vision and Pattern Recognition, pp. 1502-1509, 2009.
[2] P. Dreuw, P. Steingrube, H. Hanselmann, and H. Ney, "SURF Face: face recognition under viewpoint consistency constraints," Proc. in British Machine Vision Conference, pp. 7.1-7.11, 2009.
[3] J. Ruiz-del-Solar, R. Verschae, and M. Correa, "Recognition of faces in unconstrained environments: A comparative study," EURASIP Journal on Advances in Signal Processing, vol. pp. 1–20, 2009.
[4] G. Zhao and M. Pietikaèinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 6, 2007.
[5] X. Lu and A. K. Jain, "Integrating range and texture information for 3D face recognition," Proc. in IEEE Workshop on Applications of Computer Vision., pp.156-163, 2005.

[6]   Y. Wang, C. Chua, and Y. Ho, "Facial feature detection and face recognition from 2D and 3D images," Pattern Recognition Letters, vol.23, pp.1191-1202, 2002.

[7]   C. Benabdelkader and P. A. Griffin, "Comparing and combining depth and texture cues for face recognition," Image and Vision Computing, vol.23, pp.339-352, 2005.

[8]   K. W. Bowyer, K. Chang, and P. J. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," Computer Vision and Image Understanding, vol.101, pp.1-15, 2006.

[9]   X. Lu and A. K. Jain, "Integrating range and texture information for 3D face recognition," Proc. in IEEE Workshop on Applications of Computer Vision., pp.156-163, 2005.

[10] D. Maturana, D. Mery, and Á. Soto, "Face Recognition with Local Binary Patterns, Spatial Pyramid Histograms and Naive Bayes Nearest Neighbor Classification," Proc. in SCCC International Conference of the Chilean Computer Science Society, pp. 125- 132, Chile, 2009.

[11] T. Ahonen, A. Hadid, and M. Pietikinen, "Face Description with Local Binary Patterns: Application to Face Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037-2041, Dec. 2006.

[12] G. Heusch, Y. Rodriguez, and S. Marcel, "Local binary patterns as an image preprocessing for face authentication," Proc. in International Conference on Automatic Face and Gesture Recognition, 6 pp. - 14, Southampton 2006.

[13] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with Local Binary Pattern," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, 2002.

[14] Z. Guo, L. Zhang and D. Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification," IEEE Trans. on Image Processing, vol. 19, no. 6, pp. 1657-1663, June 2010.

[15] J. Yang, D. Zhang, and A. F. Frangi, and J.Y. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 131-137, 2004.

[16] M. Ma, M. Shao, X. Zhao, and Y. Fu, "Prototype based feature learning for face image set classification," Proc. in International Conference on Automatic Face and Gesture Recognition, pp. 1- 6, 2013.

[17] W. Chen, and Y. Gao, "Face recognition using ensemble string matching," IEEE Transaction on Image Processing, vol. 22, no. 12, pp. 4798 - 4808, 2013.

[18] K. Roy, P. Bhattacharya, and C. Y. Suen, "Towards nonideal iris recognition based on level set method, genetic algorithms and adaptive asymmetrical SVMs," Engineering Applications of Artificial Intelligence, vol. 24, no. 3, pp. 458–475, 2011.

[19] V. Ghosal, P. Tikmani, and P. Gupta, "Face classification using Gabor wavelets and random forest", Proc. in Canadian Conference on Computer and Robot Vision, pp. 68-73, 2009.

[20] S. H. Ki, J. H. Lee, B. Ko, and J. Y. Nam, "X-ray image classification using random forests with local binary patterns", Proc. in International Conference on Machine Learning and Cybernetics, Qingdao, pp. 11-14, 2010.

[21] K. C. Kwak and W. Pedrycz, "Face recognition using an enhanced independent component analysis approach," IEEE Transactions on Neural Networks, vol.18, pp.530-541, 2007.

[22] Veerabhadrappaand L. Rangarajan, "Multi-Level Dimensionality Reduction Methods Using Feature Selection and Feature Extraction," International Journal of Artificial Intelligence and Applications, vol. 1, no. 4, pp.54-68, 2010.

[23] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[24] K. Ricanek and T. Tesafaye, "MORPH: A longitudinal image database of normal adult age-progression", Proc. of the International Conference on Automatic Face and Gesture Recognition, pp.341-345, 2006.