International Journal of Advanced Science and Technology
Vol. 29, No. 7s, (2020), pp. 2709-2716

Facial Face Recognition Based On High-Performance Topographical Features Selected Using Adaptive Feature Selection Method

Arif Sameh Arif

Lectruer, Middle Technical University, Inistitute of Admistration, Department of Computer Systems Techniques, Baghdad, Iraq

Abstract

In this paper, a new features selection method is represented to enhance a feature type that is already adopted in previous classification researches. By studying successive differences between image values of face image, a relatively huge set of topographical features (TGF) are extracted from assigning each value image pixel to the related feature. An Efficient Feature Selection for face recognition (EFS) method is proposed to study and analyze TGF features. Such method used standard measures used to determine feature performance adopted in standard methods, in addition to the major concept of Principal Component Analysis (PCA) in dimension reduction for determining efficient features. EFS analyzed the dispersion of yielded values for each feature within images of the dataset in order to determine the performance level of each feature. Eventually, the level of determined performance is assigned to the related features which are sorted in descending manner. Based on a controlled threshold, the highest set of performance values, which are the best expected set of features, are selected as candidate features for Support Vector Machine (SVM) classifier. Collecting proposed features and the selection method yielded encouraging results up to (94.12%) as classification accuracy with considerable differences comparing with the State of Art.

Keywords*: Facial Recognition, Feature Selection, Feature efficiency, Topographical features, SVM)*

1. Introduction

Face recognition has been widely studied in the recent decades. Many types of features with different accuracy levels in recognition were presented [1]. Each feature type provides an amount of extracted information from the studied image, whereas more significant information represented in the chosen feature provides higher performance in recognition results [2]. Spatial types of filtering were proposed in [3] to extract points of interest, which are gathered in order to build the potential features. A statistical model depends on probability has been built to select the features with most powerful performance for face recognition. While a distance-based classifier is used to classify extracted features from Principal Component Analysis (PCA) combined with Scale Invariant Feature Transform (SIFT) [4]. Jie Chen *et al.* [5] preferred to use the transformation feature by adopting a specific version of the Local Binary patterns (LBP) as a face descriptor. Such features provided encouraging classification accuracy, but with low level of performance in the classification [5]. Critical features of the human face were studied and analyzed using the standard Analysis of Variance (ANOVA) table to be fed in deep neural networks, which yielded unstable results over large dataset of face images [6]. Next section (2) discusses TGF features and how they provide more information than traditional edges by assigning all image pixels into corresponding features. In such approach, all image information is included, and accordingly section (3) discusses choosing high performance features and ignoring other ones.

Since TGF features are used to provide more information than traditional image edges, they are tested with the proposed method of feature selection in order to select the best expected features among them. In the rest section this of this research, a brief illustration of TGF and the suggested feature selection method, followed by results in the face recognition and their analysis. The last section contains the extracted conclusions from this work.

ISSN: 2005-4238 IJAST

2. Proposed Face Features.

(a) Original image (b) Image edges (c) TGH Features

Fig 1: Traditional image edges ignore significant information, while TGH study all available features.

Although Topographical Features TGF features were used in previous object recognition researches, classification accuracy of their yielded results was relatively low considering the other types of features. One of the expected reasons for the lack of their usage is the huge amount of features produced, since TGF assigns each image pixel to a corresponding feature [7, 8]. However, TGF provides larger amount of information than traditional image edges and lines. On the other hand adopting only image lines and edges may ignore significant information, as shown in Fig. 1.

Such problem is solved by selecting the most efficient set of features chosen from all produced features overall image, which handed a previously noticeable obstacle [10]. After processing the studied image with a set of masks to find the orthogonal polynomial of the image [9], image first derivative is used to provide gradient and Hessian matrix is computed from second derivatives to extract TGF features. Such features contain Peak, Pit, Saddle, Ridge, Ravine, Zero Crossing, Flat, Increasing Area and Decreasing Area [10]. Therefore, dealing with TGF provides significant amount of information, and proposed feature selector in the next section explains the suggested technique to overcome the high number of provided features. Since TGF is previously provided as a feature type, this research concerns with illustrating the method of feature selection, and aforementioned literature is recommended for discussion TGF in detail.

3. Efficient Feature Selector (EFS)

Proposed EFS method searches for efficient features using similar concepts adopted in Linear Discriminant Analysis (LDA) to estimate feature significance [11]. According to the non-parametric concept of evaluating feature performance in LDA, high-performance features must provide high differences over inter-class images against low differences over intra-class images. Such technique adopts examining each single feature in each round by measuring yielded differences over inter- and intra-class to record its level of efficiency. Then, the next feature is tested and so on till the last feature. Different yielded values of efficiency are filtered by supervised threshold that controls the selected set of features with highest efficiency. Determined locations of the candidate features are allocated to be chosen as recognition features over learning and test dataset in the SVM stages.

To explain the proposed technique of feature selection EFS, assume that the learning dataset of $n \times m$ images contain n persons and m images for each person. Each image, im_{ij} (where $i \in n, j \in m$), produces $h \times k$ TGF features, as below:

$$im_{ij} = \begin{bmatrix} f_{11} & f_{12} & f_{13} & \cdots & f_{1k} \\ f_{21} & f_{22} & f_{23} & \cdots & f_{2k} \\ f_{h1} & f_{h2} & f_{h3} & \cdots & f_{hk} \end{bmatrix} \qquad (1)$$

ISSN: 2005-4238 IJAST

From each image, the first feature in the learning dataset is transferred to the inspection matrix:

$$I_{11} = \begin{bmatrix} f11_{11} & \dots & f11_{1j} & \dots & f11_{1m} \\ f11_{i1} & \dots & f11_{ij} & \dots & f11_{im} \\ f11_{n1} & \dots & f11_{nj} & \dots & f11_{nm} \end{bmatrix} \quad (2)$$

where,

I_{11} = the inspection matrix of the feature f_{11} over the learning dataset, and

$f11_{ij}$ = the first feature $f11$ taken from the image in the row i and column j of the learning dataset.

In order to choose features that yield highest differences, two statistical laws of dispersion standard deviation (SD) and roughness coefficient (*rc*) are suggested to measure yielded changes within values over inter-class and intra-class[12]. Firstly, SD is computed over each row and used to evaluate the changes over different face images for the same person, which represents interior differences (intra-changes). Secondly, SD over each column is used to evaluate changes over different persons within specific age (inter-changes). The mean (μ) and SD of each row are computed as:

$$\mu_i = \frac{\sum_{j=1}^{m} f11_{ij}}{m} \quad (3)$$

$$SD_{ri} = \frac{\sum_{j=1}^{m} (f11_{ij} - \mu_i)^2}{m} \quad (4)$$

Accordingly, the μ and SD of each column are computed as:

$$\mu_j = \frac{\sum_{i=1}^{n} f11_{ij}}{n} \quad (5)$$

$$SD_{cj} = \frac{\sum_{i=1}^{n} (f11_{ij} - \mu_j)^2}{n} \quad (6)$$

This produces the next matrix:

$$I_{11} = \begin{bmatrix} f11_{11} & \dots & f11_{1j} & \dots & f11_{1m} & SD_{r1} \\ f11_{i1} & \dots & f11_{ij} & \dots & f11_{im} & SD_{ri} \\ f11_{n1} & \dots & f11_{nj} & \dots & f11_{nm} & SD_{rn} \\ SD_{c1} & \dots & SD_{cj} & \dots & SD_{cm} & \end{bmatrix} \quad (7)$$

The SD for each row is denoted as SD_{ri} to be different from the SD for each column (SD_{ci}). Many of SD values are produced over intra-changes for each person, and o get a consolidated average SD value over all intra-class, SD_{intra} is computed using rule:

$$SD_{intra} = \frac{\sum_{i=1}^{n} SD_{ri}}{n} \quad (8)$$

As same as, SD_{inter} is computed as:

$$SD_{inter} = \frac{\sum_{j=1}^{m} SD_{cj}}{m} \quad (9)$$

The final value of feature efficiency, FP_{final} , for the feature f_{11} is computed by dividing SD_{inter} over SD_{intra} as given by:

$$FP_{final} = \frac{SD_{inter}}{SD_{intra}} \quad (10)$$

ISSN: 2005-4238 IJAST

Another measure of dispersion is the roughness coefficient which is a ratio between the differences over the successive values and their differences from the mean. Such a measure minimizes or deletes the effect of large differences between successive values [12]. This measure is calculated as:

$$rc = \frac{\sum_{i=2}^{n}(x_i - x_{i-1})^2}{\sum_{i=2}^{n}(x_i - \mu)^2} \tag{11}$$

The form of *rc* reduce the effects significant differences yielded in successive images due to different styles, facial hair, facial expression or other factors [13], which apparently has no significant differences in related ages. Accordingly, the new matrix that corresponds to Eq, (7) will be:

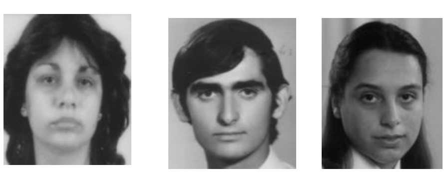
$$I_{11} = \begin{bmatrix} f11_{11} & \dots & f11_{1j} & \cdots & f11_{1m} & rc_{r1} \\ f11_{i1} & \dots & f11_{ij} & \cdots & f11_{im} & rc_{ri} \\ f11_{n1} & \cdots & f11_{nj} & \cdots & f11_{nm} & \\ rc_{c1} & \cdots & rc_{cj} & \cdots & rc_{cm} & rc_{rn} \end{bmatrix} \tag{12}$$

the bigger value of E_{final} the better since it refers to the SD_{inter} that has bigger value of SD_{intra} regarding SD_{inter}. E_{final} is transported to its related location in a matrix for calculation of the Efficiency of Features (*EFF*), e.g. E_{final} for the feature f_{11} is transferred to *EFF*(1,1). Proposed technique replies equations (2) to (10) to be applied on all next features in the image matrix im_{ij}, one-by-one.

Apparently, *EFF* matrix will contain a collection of performance values that differ over different features, and as a result, picking features with highest level of performance is expected to provide the best set of candidate features. In order to separate efficient features from weak ones, a specific threshold is proposed to choose features with bigger values of E_{final} than the threshold while values with performance lower than the threshold are considered as weak ones and ignored. For the purpose of classifying extracted features and estimate the related age, standard Support Vector Machine (SVM) standard classifier is adopted in this work [14, 15].

4. Results and Discussion

Experiments of this work used moderately distorted and high quality images picked from standard FG-NET database of face-images. Extremely distorted images are ignored in this paper to avoid undesirable effects on classification results. Hence 912 of the 1002 images from the studied dataset are viable for use in the experiments. For standardization, each image was converted to grayscale and a resized into 512×512 pixels. Fig. 2 shows some examples of test images used in this work.



(a) Fairly Distorted MAE=94.78 (b) Acceptable MAE=92.93 (c) High Quality MAE=91.12

Fig 2: Image quality affects the classification accuracy and causes error rate in recognition results over the whole dataset.

Building the mathematical model of 2D polynomial enhances image quality using a set of convolution masks, though image quality still has affect the classification accuracy (CA). Low image quality may hide or degrade original face features or may add non-real features [7].

ISSN: 2005-4238 IJAST

Chosen threshold in EFS measures affects the results of CA depending on two factors, performance level and number of selected features. Choosing higher threshold value increases performance level by focusing on higher performance scores. But such threshold selection may cause a loss of significant features. On the contrary, decreasing the value of chosen threshold ensures selecting most of high performance features, which may lead to including some features with low level-performance. Such selection can affect CA value. In addition, selecting wide number of features increases processing time. Conducted experiments showed that the best CA value is yielded with when the threshold value is set to select 26% of the features, as shown in Fig. 3.

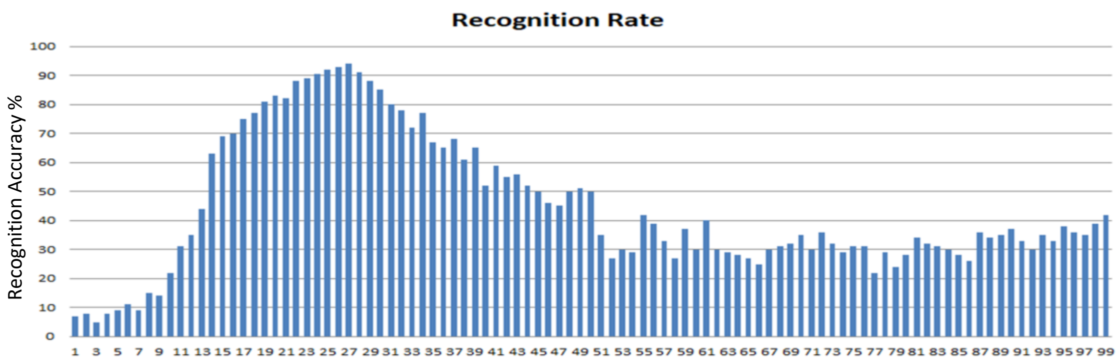


Figure 3: Recognition rate due to the value of chosen threshold.

Standard FG-Net face dataset is constructed of real-life face images, which are not specialized for scientific purposes. Therefore, it contains a large set of images with facial expression, glasses, facial hair and hair styles (see Fig. 4). Some faces in FG-Net images were rotated on multiple axes, the first as rotation around the vertical axis when facing to the left of right, while some others rotated around the horizontal axis by tilting their head (see Fig. 5). In spite of such degradation factors that affect CA value, the experiments still yielded encouraging results, as shown in Table 1.

Table 1: Recognition rate under the effects of glasses, hair style and face hair

Condition	RR
Hair Style	93.41
Face Hair	93.03
Glasses	92.81
Rotation	93.25

Figure 4: Glasses, hair style and facial hair can affect the recognition by hiding some features or creating new ones.

ISSN: 2005-4238 IJAST

Figure 5: Some faces are rotated around the horizontal or vertical axes

Yielded results illustrated that the TGF features are more robust against rotation around the vertical axis than around the horizontal axis. Firstly, the justification of such case can be explained by two points. As TGF features are produced using differences between values without depending on the own values in specific, they are more robust against rotation and illumination effects [7]. Secondly, rotations around the vertical axis changes the features positions with fewer changes on the value [15], in other words, image has the same set of values but in different positions. However, rotation around the horizontal axis can hide or create features due to the illumination and the view angle, i. e. side view can give significantly different view from the frontal one.

The achievement of the proposed EFS using TGF features were compared against the results reported by other recent studies on facial recognition employing different approaches, as given in Table 2. It is found the EFS yielded encouraging results. The highest yielded CA by the benchmark techniques was 93.89%, whereas the proposed technique achieved an average recognition accuracy of 94.05%. When applying the proposed technique on images with best conditions, the proposed technique managed to achieve recognition rates of up to 94.42%.

Table 2: Benchmarking results with other techniques

Techniques	**CA (%)**
Affine invariant feature extraction [16]	92.50
Local and Holistic features with Neural Networks [17]	93.46
Signal Reconstruction with Neural Networks [18]	93.89
Proposed EFS Technique (Average)	**94.05**
Proposed EFS Technique (Best)	**94.42**

5. Conclusions

In this paper, extracting topographical features is accomplished based on the constructed 2D polynomial rather than the image itself. The extraction is by computing first and second order derivatives of the image. Since that TGF produces wide set of features, the proposed EFS technique is proposed to select efficient features and ignore weak ones based on standard feature evaluation concepts. Proposed type of feature provided more robustness against rotation and other variations in the images including spectacles, hair styles and facial hair. The proposed technique yielded encouraging results in the comparison with variety of other techniques in field. For future work, more in-depth study of the effect of facial expressions on TGF features extracted by EFS could be undertaken.

ISSN: 2005-4238 IJAST

References

[1] Rahimeh Rouhi, Rouhi, R., Amiri, M., and Irannejad, B., 2012. A Review on Feature Extraction Techniques In Face Recognition. Signal and Image Processing: An International Journal (SIPIJ), vol. 3, no. 6, pp 97-103.

[2] Guyon, I., Gunn, S., Nikravesh, M., Zadeh, L.A. 2006. Feature Extraction, Foundations and Applications. Series Studies in Fuzziness and Soft Computing, Physica-Verlag, Springer.

[3] King Choong and Roberto Cipolla, 1997. Feature-based Human Face Detection. Journal of Image and Vision Computing15 (1997) 713-735. Elsevier.

[4] Priyanka, D., 2015. A Study on Facial Feature Extraction and Facial Recognition Approaches. International Journal of Computer Science and Mobile Computing, vol. 4 no. 5, May 2015, pp. 166-174.

[5] Chen, J., Patel, V. M., Liu, L., Kellokumpu, V., Zhao, G., Pietikäinen, M., and Chellappa, R., 2017. Robust local features for remote face recognition. Journal of Image and Vision Computing, vol. 64, pp. 34-46. Elsevier.

[6] Abudarham, N., Shkiller, L., and Yovel, G., 2019. Critical features for face recognition. Journal of Cognition, vol. 182, pp. 73-83. Elsevier.

[7] Wang, L. and Pavlidis, T. (1993). Direct Gray-Scale Extraction of Features for Character recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1053-1067.

[8] Lee, S.-W. and Kim, Y. J (1995). Direct Extraction of Topographic Features for Gray Scale Character Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 7, pp. 724-728.

[9] Moody, E. B. (1994). Discrete Orthogonal Polynomial Restoration of Images Degraded by Spatially Varying Point Spread Functions. Image Processing, 1994. IEEE International Conference on Image Processing (ICIP-94), pp. 680-684.

[10] Haralick, R. M. (1984). Digital Step Edges from Zero Crossing of Second Directional Derivatives. IEEE Transactions on Pattern Analysis and Machine Intelligence, no. 1, pp. 58-68.

[11] Chen, W., Er, M. J. and Wu, S. (2005). PCA and LDA in DCT Domain. Pattern Recognition Letters, vol. 26, no. 15, pp. 2474-2482.

[12] Hogg, R. V., McKean, J., and Craig, A. T. (2005). Introduction to mathematical statistics. Pearson Education.

[13] Xie, S., and Hu, H. (2019). Facial Expression Recognition Using Hierarchical Features With Deep Comprehensive Multipatches Aggregation Convolutional Neural Networks. IEEE Transactions on Multimedia, vol. 21, no. 1, pp. 211-220.

[14] Suykens, J. A., and Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural Processing Letters, vol. 9, no. 3, pp. 293-300.

[15] Yu, L., Xia, X., Zhou, K., and Zhao, L. (2018). Affine invariant fusion feature extraction based on geometry descriptor and BIT for object recognition. IET Image Processing, vol. 13, no. 1, pp. 57-72.

[16] Yu, L., Xia, X., Zhou, K., and Zhao, L. (2018). Affine invariant fusion feature extraction based on geometry descriptor and BIT for object recognition. IET Image Processing, vol. 13, no. 1, pp. 57-72.

[17] Xie, S., and Hu, H. (2019). Facial Expression Recognition Using Hierarchical Features With Deep Comprehensive Multipatches Aggregation Convolutional Neural Networks. IEEE Transactions on Multimedia, vol. 21, no. 1, pp. 211-220.

ISSN: 2005-4238 IJAST

[18] Xu, B., Liu, Q., & Huang, T. (2018). A discrete-time projection neural network for sparse signal reconstruction with application to face recognition. IEEE transactions on neural networks and learning systems, (99), 1-12.

ISSN: 2005-4238 IJAST