

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339324782>

Sorting Pixels based Face Recognition using Discrete Wavelet Transform and Statistical Features

Research · July 2019

DOI: 10.13140/RG.2.2.13374.64323

CITATIONS

0

READS

51

3 authors, including:



Raja K B

University Visvesvaraya College of Engineering

188 PUBLICATIONS 1,341 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Face recognition [View project](#)



2D and 3D Face recognition [View project](#)

Sorting Pixels based Face Recognition Using Discrete Wavelet Transform and Statistical Features

Divya A.

Dept. of ECE

University Visvesvaraya College of Engineering, Bangalore
University

Bangalore, India

e-mail: divyavsagar24@gmail.com

K. B. Raja

Dept. of ECE

University Visvesvaraya College of Engineering, Bangalore
University

Bangalore, India

e-mail: raja_kb@yahoo.com

Venugopal K. R.

Vice-Chancellor

Bangalore University

Bangalore, India

e-mail: Venugopalkr@gmail.com

Abstract—Face recognition is used to identify a person effectively and most effective physiological biometric trait. In this paper, we propose sorting pixels-based face recognition using Discrete Wavelet Transform (DWT) and statistical features. The novel concept of sorting pixel values in ascending order is introduced and segmented into two parts viz., Low Pixel Values (LPV) and High Pixel Values (HPV). The DWT is applied on LPV matrix to generate low and high frequency bands such as LL, LH, HL and HH. The low frequency LL band is considered for features as the coefficient values are enhanced compared to original image pixel values and also reduction in dimensionality. The statistical measure is applied on HPV to compute mean, median, mode, maximum and standard deviation features. The features of LL band and statistical features are concatenated to obtain final features. The Artificial Neural Network (ANN) is used as classifier to recognize human beings. It is perceived that the performance of the proposed method is enhanced compared with the existing methods.

Keywords—Biometrics; face recognition; discrete wavelet transform; statistical measures; artificial neural network

I. INTRODUCTION

Facial Recognition (FR) has expedited by decades and has gained more importance in the extensive research discipline viz., Image Processing, Pattern Recognition, Computer Vision and Neural Networks. FR is the most challenging yet interesting problem which continues to be a powerful tool due to its high reliability and security. FR has attracted many researchers in producing a variety of algorithms and shown promising results, but understanding the subtle changes in the facial geometry and extracting an appropriate facial feature is still a key task.

The FR system consists of two significant steps, namely, feature extraction and matching. The Feature extraction procedures are generally classified into local and heuristic

features, where local features like Local Binary Pattern, Local Ternary Pattern and Gabor descriptor are associated with some specific features of the face image such as eyes, nose etc., and heuristic approaches like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are associated by considering the whole image for feature extraction. In this paper, sorting based face recognition using DWT and statistical features is proposed. The new concept of sorting pixel values is introduced and DWT is applied on LPV to derive DWT features and statistical measures is applied on HPV to derive statistical features. The final features are derived by concatenating DWT and statistical features. ANN is used as a classifier to identify face images.

The paper is organized as following sections: In section II, the brief review of the feature descriptors, learning methodologies and classification techniques are discussed. Section III illustrates the detailed evaluation of the proposed methodology. The experiments on four standard face databases are demonstrated, the results are discussed and related with the existing methods in section IV. Section V concludes the paper.

II. RELATED WORK

In this section, the topics related to feature descriptor, feature learning and its classifications are conversed briefly.

Xiang Wu et al., [1] proposed Max-Feature-Map process to acquire a compressed and low dimensional face images. They developed a Light Convolutional Neural Network (CNN) structure to learn for representation of faces. Smaller kernel sizes of convolution layers and Residual Blocks have been applied to condense the parameters space and improve performance. Kaihao Zhang et al., [2] proposed a feature extraction technique for FR using spatial and temporal methods. They presented Part-based Hierarchical Bidirectional Recurrent Neural Networks to evaluate the information of face expression of temporal features. Multi Signal CNN was developed to extract the features from

spatial domain. Xi Yin and Xiaming Liu, [3] proposed a multitask learning system called multitask CNN for face identification and classification. The dynamic weighting system was developed to allocate the loss weights. They developed a pose focussed multitask CNN for learning pose-specific features. Se-In Jang et al., [4] proposed Recursive learning design for heterogeneous FR. Incremental PCA is combined with the Gaussian Mixture Model and are used for feature extraction. Extreme learning machine is used as learning algorithm with a single hidden layer for supervised learning. Jiwen Lu et al., [5] proposed a simultaneous feature and dictionary learning method used to learn the image set based FR. Learning feature projection and the structural dictionary together result in the final features from the image sets. Jiaojiao Zhao et al., [6] proposed a set-to-set distance matching used to find the similarity between the two sets. PCA is used for feature extraction and K-NN average pooling is applied on the extracted features to get the similarity scores. Then by comparing the scores between the sets the efficiency of the model is evaluated.

III. PROPOSED METHOD

The flow of the proposed scheme is illustrated using the block diagram shown in Fig. 1, which demonstrates the ideology behind the model.

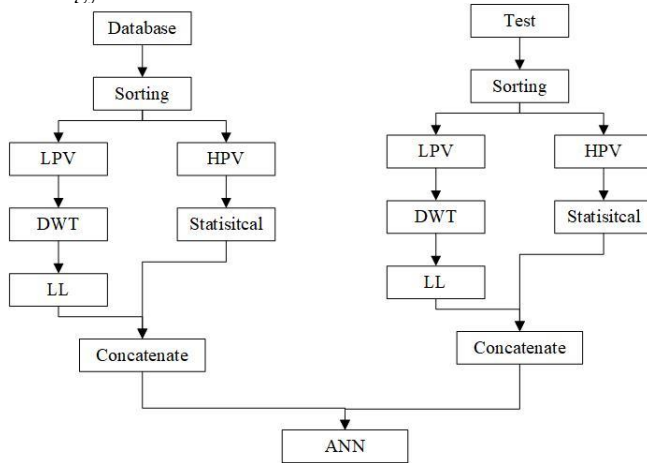


Figure 1. Block diagram of the proposed model for face recognition.

A. Database

The proposed model is tested with standard publicly available face databases such as ORL, Yale, JAFFE and FERET. The number of persons, images per person and size of each image vary with face database.

B. Sorting pixel values of face images

The quick sort is considered for sorting the elements and it is a recursive algorithm which uses divide and conquer strategy. It is an in-place sorting algorithm where constant amount of additional memory is used for sorting, which makes sorting fast and effective in many practical circumstances. The pixel values of all the images are rearranged in ascending order and divided into two equal parts namely Low Pixel Values (LPV) / (low spatial

resolution) and High Pixel Values (HPV). The LPV and HPV are used to generate frequency and spatial domain features.

C. Discrete Wavelet Transform (DWT)

It is used to transform spatial domain images into frequency domain. The 2D-DWT is implemented using a set of filter banks, consisting of cascaded structures of low-pass and high-pass filters.

The single level 2D DWT is a segmentation of the image into four non-overlapping multi-resolution sub-bands, namely, LL, LH, HL and HH represents approximation version of an image, the horizontal edges of the image, the vertical edges of the image and the diagonal edges of the image respectively. The DWT is applied on LPV; the LL sub band is considered for further process.

A sample image of size 384X256 is considered, all the pixel values of the image are sorted in ascending order and divided into LPV and HPV with size of each 49152X1. The single level 2D DWT is applied on the LPV, resulting in four sub-bands namely LL, LH, HL and HH. The LL sub-band is considered for the further processing as it enhances the low pixel values and contributes to low dimensional features for face recognition.

D. Statistical Parameters

A number of statistical measures [7] mean, median, mode, standard deviations, mac etc., are used in the field of digital image processing for image enhancement. The statistical model involves in analyzing the input image $I(x, y)$ using five statistical measures like mean, mode, median, max and standard deviation to get an optimized features of an image.

- Mean (μ): It is a value of pixel intensity values is given in Equation (1).

$$\mu = \frac{\sum_{i=0}^{N-1} x_i}{N} \quad (1)$$

where i is the index that runs through the 0 to N-1 pixel values and x_i is pixel values

- Median: It is a measure of intensity level of a pixel that separate the high intensity and lower intensity value pixels.
- Mode: It is the pixel intensity value that occurs more frequently in the dataset.
- Max: It is a vector that contains the maximum value of each row or column of a matrix A. The maximum and minimum are the only useful basic operators that are used in mathematical morphology applied on images.
- Standard Deviation (σ): It is used to measure the variation or dispersion in the intensity levels at the edges of an image. The standard deviation for an image of A of the brightness within a region (R) with Λ pixels is given by Equation (2).

$$\sigma = \sqrt{\frac{1}{\Lambda-1} \sum_{a,b \in R} (A[a, b] - \mu)^2} \quad (2)$$

where μ is the mean.

The statistical measures are applied on HPV of size 384X128 to analyze the image in spatial domain which results in lesser number of features with significant information of an image. The statistical measures of HPV reduce dimensions of HPV drastically to 5X128.

The DWT on LPV and Statistical measures on HPV, resultant matrices are merged using concatenation to acquire the final feature vector.

E. Classification

Artificial Neural Network (ANN) is used for classification and comprises of 3 layers namely, the input, hidden and the output layers. Each layer consists of a set of neurons and each neuron of the input layer are associated with weights and these succeeding to hidden layers.

The training data is obtained from the input images after feature extraction using fusion of DWT and statistical model. In the training level, supervised learning is used in which the data is labelled with the help of target data and predicts the output to match with the output. The output nodes are assigned with the values '1' for the node conforming true class, and '0' for the incorrect class. The network compares the values calculated with the output nodes with correct values to compute an error term for each node. The errors generated are then fed back to hidden layers again to modify the weights to facilitate the output results to the target.

IV. EXPERIMENTS, RESULTS AND DISCUSSION

In this segment, the performance evaluation parameters and face database settings for the proposed model using four well-known face databases are illustrated. The comparison of the proposed method with the prior-arts of face recognition is discussed. The efficiency of the fused statistical and structural approach for FR are tested with the ORL database [8], Yale database [9], JAFFE face database [10] and FERET databases [11].

A. Performance Evaluation Parameters

The performance evaluation of proposed model is calculated by means of Recognition Rate (RR), Testing Rate (TR) and Total Error Rate (TER) using the following Equations (3), (4) and (5).

$$RR = \frac{\text{Number of Correct match}}{\text{Number of images in Training set}} \times 100 \quad (3)$$

$$TR = \frac{\text{Number of Correct match}}{\text{Number of images in Test set}} \times 100 \quad (4)$$

$$TER = \frac{\text{Number of Incorrect match}}{\text{Number of images in dataset}} \times 100 \quad (5)$$

B. ORL Dataset

The ORL dataset consisting of 400 face images of 40 subjects. The face images are diverse among the face expression and angle with illumination invariant. The grayscale images of the ORL dataset has size of 112X92.

1) Results on ORL dataset

In this experiment, the Back Propagation Neural Network (BPNN) is created and consists of an input layer (1518 inputs), fully-connected hidden layer (20 nodes) and the output layer (40 classes). The training images per person is varied to assess the performance of the model. Table I shows the recognition rate and error of this experiment. The results show that, as the training images per person is increased the recognition rate increases with decrease in total error.

TABLE I. RECOGNITION ACCURACY FOR ORL DATASET

Training IPP	Test IPP	Iteration	Recognition Rate (%)	Testing Rate (%)	Total Error Rate (%)
2	8	109	75.25	88.8662	8.91
4	6	230	92.25	97.1190	1.73
5	5	149	95.25	96.2134	1.89
6	4	211	97.25	96.9992	1.20
7	3	48	98.50	99.0048	0.72
8	2	179	99.50	99.7242	0.055186

2) Comparison

Table II tabulates the recognition rate of the proposed model and related with the prior arts face descriptors under supervised settings. It is observed that the proposed method obtained enhanced recognition rate than the existing methods.

TABLE II. RECOGNITION RATE COMPARISON FOR ORL DATASET

Method	Recognition Rate (%)
Mehmet and Nihat, [12]	96.00
Thangairulappan and Jeyasingh, [13]	97.00
Y Xu et al., [14]	97.80
Robert et al, [15]	97.89
Proposed method	99.50

C. Yale dataset

The Yale dataset is a publicly used dataset which consists of one hundred and fifty face images for fifteen subjects. All the images are expressions and light illumination variant but pose invariant. The grayscale images of the YALE dataset with the original size 243X320 i.e., total pixel values are 77760 is considered for the experiments.

1) Results on yale dataset

TABLE III. RECOGNITION ACCURACY FOR YALE DATASET

Training IPP	Test IPP	Iteration	Recognition Rate (%)	Testing Rate (%)	Total Error Rate (%)
2	8	363	83.33	85.0737	0.1194
4	6	161	94.67	93.6868	0.0379
5	5	164	93.33	91.2996	0.0435
6	4	203	98.00	95.9307	0.0163
7	3	97	98.67	98.7193	0.0024
8	2	200	99.33	97.9569	0.0041

The Back Propagation Neural Network (BPNN) is created and consists of 3 layers namely, the input layer

(10800 inputs), a hidden layer (12 nodes) and the output layer (15 classes). The training images per person is varied to assess the performance of the model and trained for number of iterations. Table III illustrates the variations in the recognition rate and error with respect to the training images per person. It is observed that the % Recognition Rate increase with increase in Training IPP.

2) Comparison

The recognition rate of the proposed model is related with the prior arts face descriptors under supervised settings as shown in Table IV. From the observation, the proposed method achieved higher recognition rate compared to the existing state-of-the-arts methods presented by Robert et al., [15] and C.Hu. et al., [16].

TABLE IV. RECOGNITION RATE COMPARISON FOR YALE DATASET

Method	Recognition Rate (%)
Robert et al.,[15]	96.47
C. Hu et al., [16]	96.67
Proposed method	99.33

D. JAFFE Dataset

The JAFFE dataset is a famous publicly used dataset consisting of total two hundred images of ten subjects. The images are in grayscale with the original size of 256X256 which are expression variant but invariant to the pose and illuminations.

1) Results on JAFFE dataset

The BPNN is created and consists of an input layer with 10752 inputs, a hidden layer with 25 nodes and the output layer of 10 classes. The network is trained under the supervised settings and the performance evaluation of the proposed model is performed by varying the training images per person for number of iterations. Table V shows the variations in the recognition rate and error with respect to the number of training images per person. It is observed that the % Recognition Rate is high for higher values of training IPP.

TABLE V. RECOGNITION ACCURACY FOR JAFFE DATASET

Training IPP	Test IPP	Iteration	Recognition Rate (%)	Testing Rate (%)	Total Error Rate (%)
4	11	92	92.00	93.1541	0.0548
8	12	125	97.00	94.8585	0.0308
10	10	169	98.50	98.1241	0.0094
12	8	139	99.00	98.3833	0.0065
14	6	37	98.00	99.0995	0.0042
16	4	151	100	99.9900	0.000020109

2) Comparison

The recognition rate for the proposed model is related with the prior arts face descriptors under supervised settings as shown in Table VI. From the results it is observed that the proposed model has achieved maximum recognition rate compared to the existing state-of-the-arts methods such as [13], [17], [18] and [19].

TABLE VI. RECOGNITION RATE COMPARISON FOR JAFFE DATASET

Method	Recognition Rate (%)
Ke Shan et al., [17]	76.7442
Tumpa Dey and Tamojay Deb [18]	85.00
Bingrong et al.,[19]	94.80
Thangairulappan and Jeyasingh [13]	97.00
Proposed method	100

E. FERET Dataset

The dataset is large color dataset consisting of more than 13539 images of 1565 subjects. The face images are diverse with expressions, angles and illuminations. The original image is of size 384X256 i.e.; 98304 total pixel values are used for experimentation of the proposed method. In the experiment, the variations in Pose, Illumination Expression (PIE) are used to analyses and considered 280 subjects with eleven images per subject i.e., total 3080 images are used. The final features values of 13248 are obtained from the proposed model by reducing 86.52% of the original size of the image.

1) Results on FERET dataset

The BPNN network is created with an input layer 13248 inputs nodes, a hidden layer with 25 nodes and the output layer with 280 classes. The dataset is divided into training and test dataset and training set is the input feature values of 13248. Table VII illustrates the variations in the recognition rate and error with respect to the training images per person. It is observed that the % Recognition Rate increase with increase in Training IPP.

TABLE VII. RECOGNITION ACCURACY FOR FERET DATASET

Training IPP	Test IPP	Iteration	Recognition Rate (%)	Testing Rate (%)	Total Error Rate (%)
2	9	1000	74.2857	97.8338	01.77
4	7	1000	88.5714	99.3405	00.42
5	6	1000	91.5259	99.5772	00.23
6	5	1000	93.5389	99.7738	00.11
8	3	1000	96.3961	99.9568	00.09

2) Comparison

The recognition rate for the proposed method is compared with the prior arts face descriptors under supervised settings as shown in Table VIII. It is observed that the proposed method achieved better recognition rate compared to the existing state-of-the-arts methods such as [20] and [31].

TABLE VIII. RECOGNITION RATE COMPARISON FOR FERET DATASET

Method	Recognition Rate (%)
Tao Liu et al., [20]	65.40
Junying Zeng et al., [21]	93.90
Proposed method	96.3961

The proposed model is very effective for the following reasons.

- The novel concept of original image pixel values is sorted and converted into two equal parts LPV and HPV is introduced for appropriate feature extraction techniques.
- The DWT is applied on LPV to derive LL sub band with effective enhanced coefficient values compared to original LPV and also compressed dimensionality to obtain effective first set of initial features.
- The statistical features are extracted from HPV to derive effective second set of initial features.
- First and second set of initial features are concatenated to obtain final effective features for better classification of face images.

V. CONCLUSION

Biometrics are used in security systems to identify human beings for various applications. In this paper, sorting pixel values-based FR using DWT and statistical features is proposed. The pixel values of face images are sorted in ascending order and segmented into two equal parts. The DWT is applied on LPV pixel values matrix to generate frequency bands. The low frequency sub band is considered for initial features as it has substantial information of original face image and also reduced dimensionality. The statistical features are computed from HPV. The final features are derived by concatenating DWT and statistical features. The face images are classified based on final features using ANN. The performance of the proposed method is better compared to existing methods. In future, the proposed model can be implemented in real time using FPGA to build effective face recognition model since the number of final features are less.

REFERENCES

- [1] Xiang Wu, Ran He, Zhenan Sun and Tieniu Tan, "A Light CNN for Deep Face Representation with Noisy Labels", *IEEE Transactions on Information Forensics and Security*, Vol. 13, Issue 11, pp. 2884-2896, November 2018.
- [2] Kaihao Zhang, Yongzhen Huang, Yong Du and Liang Wang, "Facial Expression Recognition Based on Deep Evolutional Spatial-Temporal Networks", *IEEE Transactions on Image Processing*, Vol. 26, No. 9, pp.4193-4203, September 2017.
- [3] Xi Yin and Xiaming Liu, "Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition", *IEEE Transactions on Image Processing*, Vol. 27, Issue 2, pp.946-975, February 2018.
- [4] Se-In Jang, Geok-Choo Tan, Kar-Ann Toh and Andrew Beng Jin Teoh, "Online Heterogeneous Face Recognition Based on Total-Error-Rate Minimization", *IEEE Transactions on Systems, Man, And Cybernetics: Systems*, pp.1-14, 2017.
- [5] Jiwen Lu, Gang Wang and Jie Zhou, "Simultaneous Feature and Dictionary Learning for Image Set Based Face Recognition", *IEEE Transactions on Image Processing*, Vol. 26, No. 8, pp.4042-4054, August 2017.
- [6] Jiaojiao Zhao, Jungong Han, and Ling Shao, "Unconstrained Face Recognition Using a Set-to-Set Distance Measure on Deep Learned Features", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 28, Issue 10, pp.2679-2689, October 2018.
- [7] Vijay Kumar and Priyanka Gupta, "Importance of Statistical Measures in Digital Image Processing", *International Journal of Emerging Technology and Advanced Engineering*, Vol. 2, Issue 8, pp.56-62, August 2012.
- [8] The ORL Database of Faces, AT & T (Olivetti) Research Laboratories, Cambridge, U.K.
- [9] Yale University (1997) 'The Yale Face Database'.
- [10] M. J. Lyons, M. Kamachi and J. Gyoba, "Japanese Female Facial Expressions (JAFFE)," Database of Digital Images, 1997.
- [11] P J Phillips, H Moon, S A Rizvi and P J Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp.1090-1104, October 2000.
- [12] Mehmet, K. and Nihat, Y, "Face Recognition by Using Back Propagation Artificial Neural Network and Windowing Method", *Journal of Image and Graphics*, Vol. 4, No.1, pp.15-19, 2016.
- [13] Thangairulappan, K. and Jeyasingh, J. B. V, "Face Recognition Based on Wavelet Packet Coefficients and Radial Basis Function Neural Networks", *Journal of Intelligent Learning Systems and Applications*, Vol. 5 No. 2, pp.115-122, 2013.
- [14] Y Xu, X Zhang, and H Gai, "Quantum Neural Networks for Face Recognition Classifier", *Procedia Engineering*, Vol. 15, pp.1319-1323, January 2011.
- [15] Robert Yao Aaronson, Wu Chen and Ben-Bright Benuwa, "Robust Face Detection using Convolutional Neural Network," *International Journal of Computer Applications*, vol. 170, no. 6, pp. 14-20, July 2017.
- [16] C. Hu, M. Ye, Y. Du, and X. Lu, "Vector Projection for Face Recognition", *Computer Electrical Engineering*, Vol. 40, No. 8, pp.51-65, 2014.
- [17] Ke Shan, Junqi Guo, Wenwan You, Di Lu and Rongfang Bie, "Automatic Facial Expression Recognition Based on a Deep Convolutional-Neural-Network Structure", *International Conference on Software Engineering Research, Management and Applications*, pp.123-128, 2017.
- [18] Tumpa Dey and Tamojay Deb, "Extended Study of K-Means Clustering Technique for Human Face Classification and Recognition", *IEEE International Conference on Electrical, Computer and Communication Technologies*, pp.1-4, 2015.
- [19] Bingrong Xu, Qingshan Liu and Tingwen Huang, "A Discrete-Time Projection Neural Network for Sparse Signal Reconstruction with Application to Face Recognition," *IEEE Transaction on Neural Networks and Learning Systems*, Vol. PP, Issue 99, pp. 1-12, June 2018.
- [20] Tao Liu, Jian-Xun Mi, Ying Liu and Chao Li, "Robust Face Recognition via Sparse Boosting Representation", *Neurocomputing*, Vol. 214, pp. 944-957, November 2016.
- [21] Junying Zeng, Xiaoxiao Zhao, Junying Gan, Chaoyun Mai, Yikui Zhai, and Fan Wang, "Deep Convolutional Neural Network Used in Single Sample per Person Face Recognition," *Hindawi Computational Intelligence and Neuroscience*, pp.1-11, 2018.