Efficient Wavelet Entropy-Based Face Recognition System

DOI:10.36909/jer.15971

Khaled Daqrouq, Emad Khalaf, and Ahmed Balamesh

Electrical and Computer Engineering Department, King Abdulaziz University

ABSTRACT

In the presented paper, the face images modeling in different circumstances such as occlusion, poses, and illumination was investigated. A novel research based on a new combination of the discrete wavelet transform (DWT) and wavelet packet (WP) methods in conjunction with log energy and sure entropies is proposed. Probabilistic neural network (PNN) was used for classification. The performance of the proposed method was investigated using a collection of experiments with different challenges ORL face database and extended Yale face database was conducted. For testing the proposed face recognition system on the ORL database, many classification scores were calculated for different training/testing sets. The comparative results obtained on the ORL database for different published methods based on the efficiency were performed and showed that our method was superior with rate reached 98.5% and relatively small, elapsed time about 4.4 sec was found. For the extended Yale face database B, the recognition rate was averaged at each time for 42 iterations. Furthermore, the recognition rate was calculated for different training/testing sets and excellent results were achieved. The contribution of the study was to use wavelet entropy for the face recognition task by finding a best combination of the method that leads to achieving a competent form guarantying better results. The major contribution of this study is that it is possible to use just a few real face images to model the illumination conditions in the extended Yale face database B.

Keywords: Entropy; Face image; Log energy; PNN; Sure entropy; Wavelet.

INTRODUCTION

Human face recognition is a significant field of image processing that is a very attractive aspect of research because of the practical implication in many systems. There are several challenging aspects of face recognition such as expression, illumination, and variation in head pose. By comparison, illumination possesses a much greater challenge for face recognition than pose and expression. Face features might change remarkably by illumination changing, which can affect the total recognition performance.

In literature, a wide range of methodologies has been suggested to tackle this challenging problem. Most of these approaches can be approximately divided into three groups: The first group is preprocessing and normalization: In this type, the images are tackled using the techniques of image processing for normalization for face image’s stability within illumination change effects. It is used for that Gamma correction and logarithmic transformation, block-based histogram equalization, region-based histogram equalization, and adaptive histogram equalization (Pizer et al., 1987, Savvides and Kumar 2003). The second group is 3D illumination models: In this category the idea is that the changes are represented in a subspace and guesstimate individual model parameters (Georghiades et al., 2001, Basri and Jacobs 2003). The algorithm is not easy to be conducted in real time due to very computational complexity. The third category is known as Extraction of illumination invariant features: where, the facial features are extracted, such as 2D Gabor-like filters methods, edge maps, and the gray-level's derivatives (Gao and Leung 2002, Chen et al., 2000, Liu and Wechsler 2002).

Recently, face recognition systems have been matured, however, challenging problems may have not expectable forms in the real world. Occlusion challenge that includes sunglasses, hats, eyeglasses, scarves, and also other objects that could be positioned in front of the face. Shadows, because of extreme illumination disturb, the model of low-dimensional linear illumination assumption (Zhou et al., 2009, Georghiades et al., 2001) are acting as kinds of occlusion. A successful face recognition system should be able to be robust to the occlusion that makes it suitable for the real world. Conventionally, the core method to tackling with occlusion is to examine the occluded face images locally by occlusion mask, or partly match occluded test images with un occluded trained images (Hong and Aleix 2009, Tan et al., 2009). For more reading, there are different other methods in literature (Zhang and Zhu, 2019, Lander et al., 2018, Tu, et., 2019, Wan**,** et al., 2019)

The contribution of the study: **first point**, the method investigates a new combination of the DWT and WP methods in conjunction with log energy and sure entropies for the face recognition task by finding a best combination of the method that leads to achieving a competent form guarantying best results. **Second point**, the investigation is conducted over different datasets with limitations. Where the purpose of the study is using the method for all kinds of challenging problems existing in the testing of three databases. The main contribution of this study is that it is possible to use just a few real face images to model the 64 illumination conditions in the extended Yale face database B.

This paper is constructed as following: First the introduction section that includes a background of the face recognition task and literature survey. The method section, where a full description of the method is included. The results and discussion section that includes all the experiments and its analysis and discussion. At the end we have the conclusion section and finally the references.

THE METHOD

In this paper, we investigate the face images in different circumstances such as occlusion, poses, and illumination. A novel research based on a new combination of the DWT and WP methods in conjunction with log energy and sure entropies is proposed. The wavelet transform is one of the most attractive methods for most of researchers of the signal processing, data mining, and pattern recognition field (Daqrouq and Azzawi 2012). The reason behind that is the trick of decomposing the image into many sub-band passes of frequency. For this study, the following steps are used:

- Reading the image as a numeric matrix.
- We decompose each row as well as column signal into the DWT of each level:

$F = [\, D_1, D_2, D_3, D_4, D_5, D_6, A_6]$

Where

$$
\begin{aligned}
D_1 &= d_{11}, d_{12}, \dots, d_{1N/2} \\
D_2 &= d_{21}, d_{22}, \dots, d_{2N/4} \\
D_3 &= d_{31}, d_{32}, \dots, d_{3N/8} \\
D_4 &= d_{41}, d_{42}, \dots, d_{4N/16} \\
D_5 &= d_{51}, d_{52}, \dots, d_{5N/32} \\
D_6 &= d_{51}, d_{52}, \dots, d_{6N/64} \\
A_6 &= d_{51}, d_{52}, \dots, d_{6N/64}
\end{aligned} \tag{1}
$$

where D_l represents the original signal part with high-frequency, which is known as detail sub-band; A_L represents the original signal part with low-frequency and denoted as the approximation DWT sub-band of the last level L, and N is the initial signal length (Daqrouq and Al Azzawi 2012).Framing the approximation sub signal at level 6 (A_6) into two frames, then Log Energy and sure entropies of WP at level 2 are calculated for each frame of the approximation sub signal and get the average vector of the two frames. The feature extraction vector (FV) is:

$$FV = [\mathrm{logenergy}(\mathrm{WP}_{ROWS}), \mathrm{logenergy}(\mathrm{WP}_{COLUMNS}), \mathrm{sure}(\mathrm{WP}_{ROWS}), \mathrm{sure}(\mathrm{WP}_{COLUMNS})] \tag{2}$$

The mother wavelet function type is very essential, and it varies for different applications. In our presented paper, we have tested various types of wavelet families by studying the system performance. Based on our study of the best performance, we have chosen to use the Daubechies ten (also known as db10) wavelet function type because it generates the best recognition rate. Thus, it will be utilised for our system analysis for the rest of the work. For classification, the PNN is used (see Figure 1).

Figure 1. Block diagram of the presented method showing the method stages.

Entropy as a common concept in many disciplines such as information theory, mechanical and energy studies and plays a crucial role in signal processing. The entropy that is the measured of the energy distribution is giving an idea about the randomness in the data that gives an amazing measure of the color concentration for the image. In the following equation list, different entropy types are presented. Many other types of entropies are testable and can be easily combined. In the following equations, "s" is presenting the original signal or image, whereas (s_i) presents the coefficients of "s" in the wavelet orthonormal basis. E denotes for entropy that must be an additive cost function such that E(0) = 0 and

$$E(S) = \sum_i E(S_i) \quad (3)$$

- Log energy entropy: $E(S_i) = \log(S_i{}^2)$, so

$$E(S) = \sum_i \log(S_i{}^2) \quad (4)$$

with the convention *log*(0) = 0.

- Sure entropy:

E(s) = n - # {i such that $|S_i| \leq p + \sum_i \min(S_i{}^2, p^2)$ (5)

RESULTS AND DISCUSSION

In the present section, we operate a collection of experiments to investigate the performance of the suggested procedure with different challenges such as 1) contiguous occlusion (glasses, beard, or moustaches, etc.), which are covered by ORL face database (figure 2), and 3) different poses and illumination conditions, which are covered by the extended Yale face database B (figure 3).

ORL Face Database

The first stage of our experiments is organized using one of the most common database called ORL face database, which was available by Cambridge University. All its images collection of faces were captured in the 1992,1993 and 1994 at Cambridge University, United Kingdom (The Olivetti & Oracle Research Laboratory). Since that time, it has been utilized as a testing benchmark face database in many faces identification and verification systems. It includes four hundred face images captured from forty different persons, for each one ten were taken. The images were taken against a dark homogeneous background with the subjects in an upright, with rotation tolerance up to about 20 degrees, frontal position. The images were taken with several facial expressions, poses, and different conditions of lighting. The images' format is bitmap of 92 x 112 pixels resolution, and gray of 256 levels. We can find some differences in the images of various persons based on glasses, beard, or mustaches, etc. We use five images for testing and five images for training per each person in ORL database. For testing of the proposed method with ORL database, several statistical parameters were calculated such as:

False rejection error (FRR) that depends on the false negative (FN), and true positive (TP) indicated in the following formula:

$$\mathrm{FRR} = \frac{\mathrm{FN}}{(\mathrm{TP}+\mathrm{FN})}$$

where FN is the number of testing images that are not correctly recognized for a person, wherein our system FN $\leq$ 5 (because the maximum number of testing images per a person is 5 images), and TP is the number of testing images that are correctly signified for one person, wherein our system TP $\leq$ 5.

False acceptance error (FAR) that depends on the false positive (FP) and true negative (TN) presented in the following formula:

$$\mathrm{FAR} = \frac{\mathrm{FP}}{(\mathrm{FP}+\mathrm{TN})}$$

where FP is the number of testing images that are correctly recognised as a person, whereas another person's image is testing, wherein our system FP $\leq$ 5, TN is the number of testing images that are correctly signified for the rest of the individuals, wherein our system TN $\leq$ 39x5. For more understanding of the way we calculated the above verification parameters, Table 1 is prepared. In table 1, part of the testing results is tabulated for ORL database.

Table 1. Part of testing results tabulated for ORL database for 5/5 training/testing set.

Person number	1	2	3	…	18	…	26	…	40
Results of testing by PNN	1,1,18, 1,1	2,2,2, 2,2	26,3,3, 3,26	…	18,18,18, 18,18	…	26,26,26, 26,26	…	40,40,40, 40,40
FN	1	0	2	…	0	…	0	…	0
TP	4	5	3	…	5	…	5	…	5
FP	0	0	0	…	1	…	2	…	0
TN	192	192	192	…	192	…	192	…	192

The verification parameters could be used in several ways in a purpose of calculating some objective evaluation rates, such as the receiver operating characteristics (ROC), equal error rate (EER), or others. Here, the proposed face recognition system is evaluated by the following statistical parameters:

Sensitivity (S): $S = \frac{TP}{(TP+FN)}$,

Specificity (P): $P = \frac{TN}{(TN+FP)}$

Positive Predictive (PP): $PP = \frac{TP}{(TP+FP)}$,

Accuracy (AC): $AC = \frac{TP+TN}{(TP+FN+FP+TN)}$,

Efficiency (EF): $EF = 100 * [1 - \frac{FN}{(TP+FN)}]$

F1-Score: $F1= 2 * [\frac{S*PP}{(S+PP)}]$

The scores of all above classification were calculated for ORL database for different training/testing sets are used for testing the proposed face recognition system. The scores FAR, FRR, S, P, PP, F1, and AC are calculated on a person basis, whereas the final score is obtained by the mean value of all databases. The results are tabulated in table 2. the proposed system achieved very high scores mainly for 5/5 training/testing.

Figure 2. Image samples from ORL database.

Figure 3. Image samples from extended Yale face database B database.

Table 2. The scores FAR, FRR, S, P, PP, and AC are calculated for ORL database for different training/testing sets.

Training/Testing sets per person	FAR	FRR	S	P	PP	AC	EF	F1-score
5/5	0.04	1.50	98.50	99.96	98.75	99.92	98.50	98.62
4/6	0.12	3.08	96.91	99.87	97.17	99.76	96.87	97.02
3/7	0.43	7.14	92.85	99.56	93.76	99.16	92.50	93.30
2/8	1.22	8.75	91.25	98.77	91.62	97.84	91.25	91.22

In table 3, we compare the efficiency for the ORL database of the suggested system method with other wavelet-based algorithms published in (Chao et al., 2007) such as LDC, LDB, MLDB and Wavelet Face. The results were produced for different training/testing sets. The execution of the suggested method, as an average of the results calculated for 5/5, 4/6, and 3/7 training/testing sets, is superior.

Table 3. The performance of the suggested method as a calculated average of the results for 5/4, 4/6, and 3/7 training/testing sets.

Training/Testing sets per person	LDC (Chao et al., 2007)	LDB	MLDB	Wavelet Face	Proposed method
5/5	96.95	96.65	96.43	94.20	98.50
4/6	95.81	95.60	95.18	94.56	96.87
3/7	93.43	93.30	92.49	92.92	92.50
Avr.	95.39	95.1833	94.70	93.89	95.95

To evaluate our system performance under different image conditions in the ORL database the ROC curve is used. The ROC graph uses true positive rate (TPR) on the y-axis vs. false positive rate (FPR) on the x-axis. ROC calculated for the proposed system is illustrated in figure 4. Maximum sensitivity links to a large y-axis value on the curve. Maximum specificity is related to a small x-axis value on the curve. Therefore, an excellent, test cutoff value is the value that shown at the upper left corner corresponds to a point on the curve. To review the performance of the ROC curve, the equal error rate (EER) that relates to the point where the line FPR = 1-sensitivity or 1-TPR intersects the ROC curve is utilized (Basri and Jacobs 2003).

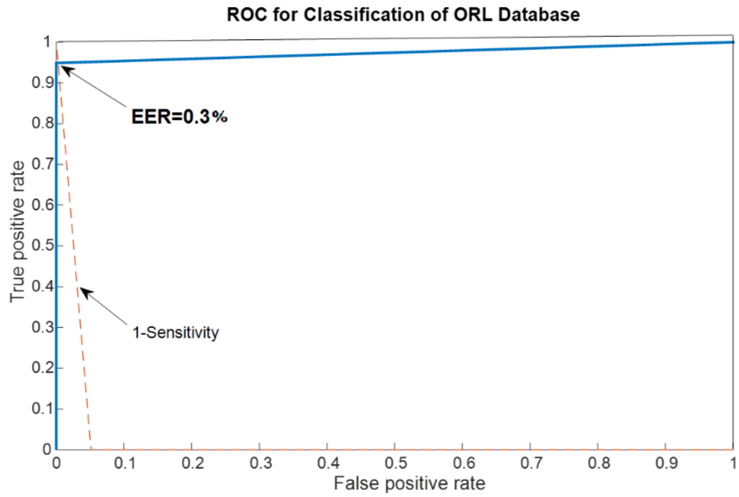

Figure 4. The ROC curve calculated for the proposed method for ORL database.

The quality, based on the recognition efficiency of the face recognition systems, was compared with some published systems for face recognition such as NeNMF (Guan et al., 2012), Block-based Steerable pyramid (El Aroussi et al., 2011). Moving average with neural network (MAFN) (Khalaf, 2019) and eigen value with moving average filter and neural network (EMAFN) (Khalaf, 2019) methods are also used for comparison. Table 4 shows the comparative results achieved for the ORL database. The experiments with these schemes declare the superiority of the proposed method.

Table 4. The comparative results obtained on ORL database for different published methods based on the efficiency.

Method	Efficiency
LR-SNN-T	90.25
Block based Steerable pyramid.	75.6
NeNMF	92.36
LSPBP	96
ORSA-NMF	94.5
PCA	87.5
LDA	75.6
wavelet-based generalized neural network	96.4
MAFN	91.45
EMAFN	91.75

Table 5. Other published methods with regard to testing time for ORL database are compared with proposed method.

Method	Testing time in seconds
PCA	8.9
MEN	8.6
LR-SNN-T	38.12
ORSA-NMF	6.88
NeNMF	8.2
wavelet-based generalized neural network	6.17
Proposed method	4.4

Proposed Method	98.50

Taking into consideration the results LDA, PCA, and LSPBP published in (El Aroussi et al., 2009), MEN (Zhou et al., 2011), low-resolution single neural network (LR-SNN-T) (Jahan et al., 2007), ORSA-NMF (*Guna et al., 2012), NeNMF (Guna et al., 2012), wavelet-based generalized neural network (Sharma et al., 2013), and our work results, the testing time of published face recognition systems is provided in table 5. Regardless the different processor speeds used in the calculations, and different quality of the machines, it is pretty safe to say that the proposed system is not worse than the other systems with regard to recognition rate and testing time for ORL database. That is a satisfactory result in terms of computational complexity, costs, and manufacturing ramifications.

The Extended Yale Face Database B

This database contains 16,128 images for 28 persons (9 poses and 64 illuminated images), then we have 576 images each of 640×480 pixels for each subject. Its format is similar to the Yale face database B (Georghiades et al., 2001). For the extended Yale face database B, we trained a part of the 576 images and tested the remaining part. The training images were chosen randomly 42 times, and recognition rates were averaged at each time. Furthermore, the confidence interval was calculated for the recognition rate of each iteration and illustrated in Fig.4. The confidence interval indicates that the obtained recognition rates of 95% for random combinations of training/testing tests contained in this interval. The Wider confidence interval is a sign of not suitable feature extraction method. For our method, the averaged interval calculated for the presented method is reasonable (99.50-99.60), see figure 5.

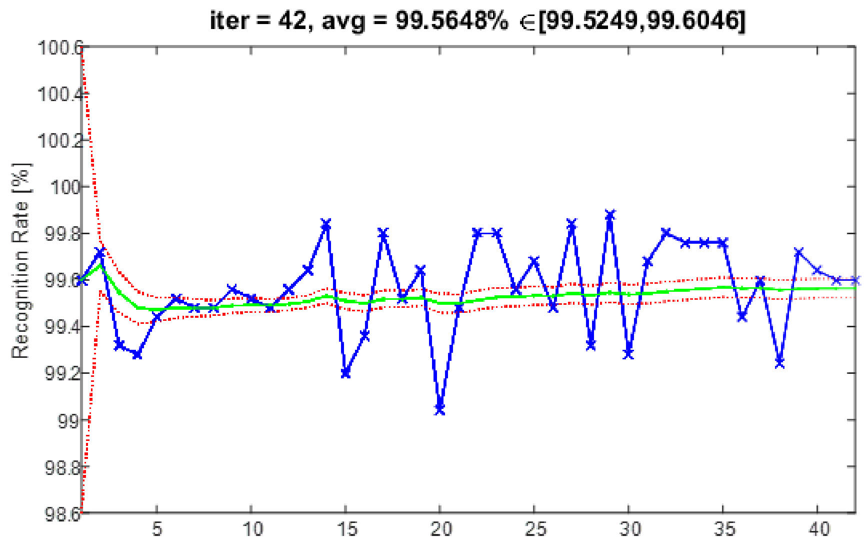

Figure 5. The extended Yale face database B results, where the half images of the human subject are used for training and the rest for testing that was chosen randomly for 42 times, and the recognition rates were averaged at each time. Furthermore, the confidence interval was calculated for the recognition rate of each iteration.

In table 6, we investigated the recognition rate of the suggested method for the extended Yale face database B. The results were conducted for different training/testing sets. The best result of the suggested method was for 250/326 training/testing set, which reached 99.57%. We can see that a good face recognition result can be attained for ten training images per person that were enough to achieve reasonable accurate face features to model the whole illumination cases. In conclusion, the main contribution of this study is that it is possible to use just a few real face images to model the 64 illumination conditions in the extended Yale face database B.

Based on the recognition rate of the face recognition system, we compare the quality of the proposed method with some published systems for face recognition such as SRC (Wright et al., 2009), SSRC1 (Ou et al., 2014), and RSC (Yang et al., 2011) on extended Yale face database B. The comparison was conducted based on the results published in (Ou et al., 2014) as an average of the recognition rates of the four sets 1,2,3 and A, with the fifth number of the data for training and the rest for testing. Table 7 tabulates the comparative results. The investigations of these methods show that the proposed method is the best.

Table 6. The recognition rate for the extended Yale face database B. The results were conducted for different training/testing sets.

Training/Testing sets per person	Avr. [%]	Highest [%]
10/566	83.30	95.50
50/526	94.80	98.50
150/426	99.00	99.55
250/326	99.57	99.95

Table 7. The comparison of the proposed method with other published face recognition systems such as SRC, SSRC,1 and RSC on extended Yale face database B.

Method	Efficiency
SRC	83.35
SSRC1	97.61
RSC	90.54
Proposed method	98.79

CONCLUSION

Wavelet transform and entropy have been used for face recognition task. An investigational study for testing the quality of the presented method with ORL face database and extended Yale face database has been conducted. For the sake of testing the proposed face recognition system, statistical scores such as FAR, FRR, S, P, PP, F1-Score, and AC were computed. Different training/testing sets over ORL database have been investigated. The comparative results conducted on the ORL database for different published methods based on the efficiency (for comparison), showed that our method was superior. For the extended Yale face database B, the training images were chosen randomly 42 times, and recognition rates were averaged at each time. Furthermore, the confidence interval was illustrated. The recognition rate of our method for the extended Yale face database B was calculated for different training/testing sets, and excellent results were achieved. The main contribution of this study is that it is possible to use just a few real face images to model the 64 illumination conditions in the extended Yale face database B. At the end of this study, we conclude that the presented method is very promising with excellent results, especially for 5/5 training/testing set. The face recognition phone application is intended to be created as a future work.

ACKNOWLEDGEMENTS

This project was funded by the Deanship of Scientific Research (DSR). King Abdulaziz University, Jeddah, under grant no. 1434/135/463. The authors, therefore, acknowledge with thanks to DSR technical and financial support.

REFERENCES

Aroussi, M., El Hassouni, M., Ghouzali, S., Rziza, M., Aboutajdine, D. 2011. Local appearance-based face recognition method using block based steerable pyramid transform. Signal Processing 91,38–50.

Basri, R., Jacobs, D. 2003. Lambertian reflectance and linear subspaces. IEEE Transaction on Pattern Analysis and Machine Intelligence 25, 218–233.

Chao-Chun, L., Qing-Dai, D., Yan, H. 2007. Local Discriminant Wavelet Packet Coordinates for Face Recognition. Journal of Machine Learning Research 8,1165-1195.

Chen, H., Belhumeur, P., Jacobs, D. 2000. In search of illumination invariants. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 254–261.

Daqrouq K, Al Azzawi K. Y. 2012. Average framing linear prediction coding with wavelet transform for text independent speaker identification system. Computers and Electrical Engineering, Volume 38(6), Nov., 1467–1479.

El Aroussi, M., El Hassouni, M., Ghouzali, S., Rziza, M., Aboutajdine, D. 2009. Local steerable pyramid binary pattern sequence LSPBPS for face recognition method. Signal Processing 5 (4), 281–284.

Gao, Y., Leung, K. 2002. Face recognition using line edge map. IEEE Transaction on Pattern Analysis and Machine Intelligence 24, 764–779.

Georghiades, A., Belhumeur, P., Jacobs., D. 2001. From few to many: illumination cone models for face recognition under variable illumination and pose. IEEE Transaction on Pattern Analysis and Machine Intelligence 23, 643–660.

Georghiades, A., Belhumeur, P. Kriegman's, D. 2001. From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 643–660.

Guan, N., Tao, D., Luo, Z., Yuan, B. 2012. NeNMF: An optimal gradient method for non-negative matrix factorization. IEEE Transactions on Signal Processing 60 (6),2882–2898.

***Guan, N., Tao, D., Luo, Z., Yuan, B. 2012.** Online nonnegative matrix factorization with robust stochastic approximation. IEEE Transactions on Neural Networks and Learning Systems 23 (7), 1087–1099.

Hong, J., Aleix, M. 2009. Support vector machines in face recognition with occlusions. Conference on Computer Vision and Pattern Recognition, IEEE,136–141.

Jahan, Z., Javed, M., Usman, Q. 2007. Low resolution single neural network based face recognition. Proceedings of the Fourth International Conference on Computer Vision, Image and Signal Processing, vol. 22, 189–193.

Khalaf, E. (2019). An Investigation of the Use of Eigen Values in Human Face Modeling for Recognition Tasks, Journal of Advances in Mathematics and Computer Science, Page 1-9
DOI: 10.9734/jamcs/2019/v32i130135,

Lander, K., Bruce, V., and Bindemann, M. 2018. Use-inspired basic research on individual differences in face identification: Implications for criminal investigation and security, Cognit. Res., Princ. Implications, vol. 3, no. 1, pp. 1–13.

Liu, C., Wechsler, H. 2002. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Transactions on Image Processing 11, 467–476.

Ou, W., You, X. Tao, D. et al, 2014. Robust face recognition via occlusion dictionary learning. Pattern Recognition 47, 1559–1572.

Pizer, S., Amburn, E. et al. 1987. Adaptive histogram equalization and its variations. Computer Vision Graphics and Image Processing 39, 355–368.

Savvides, M., Kumar, V. 2003. Illumination normalization using logarithm transforms for face authentication. Proceedings of the IAPR AVBPA, pp. 549–556.

Sharma, P., Arya, K., Yadav, R. 2013. Efficient face recognition using wavelet-based generalized neural network. Signal Processing 93,1557–1565.

Tan, X., Chen, S., Zhou, Z., Liu, J. 2009. Face recognition under occlusions and variant expressions with partial similarity. IEEE Trans. Inf. Forensics Secur.4 (2), 217–230.

The Olivetti & Oracle Research Laboratory Face Database of Faces (Online). Available: ⟨http://www.cam-orl.co.uk/facedatabase.html⟩.

Tu, H., Li, K., and Zhao, Q. 2019. Robust face recognition with assistance of pose and expression normalized albedo images, in Proc. 5th Int. Conf. Comput. Artif. Intell. (ICCAI. pp. 93–99.

Wan, M.-H., and Lai, Z.-H. 2019. Generalized discriminant local median preserving projections (GDLMPP) for face recognition, Neural Process. Lett., vol. 49, no. 3, pp. 951–963, Jun. 2019.

Wright, J., Yang, A., Ganesh, A., Sastry, S., Ma, Y. 2009. Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. 31(2), 210–227.

Yang, M., Zhang, L., Yang, J., Zhang, D. 2011. Robust sparse coding for face recognition. Conference on Computer Vision and Pattern Recognition, IEEE, 625–632.

Zhang, D., and Zhu, S. 2019. Face recognition based on collaborative representation discriminant projections, in Proc. Int. Conf. Intell. Transp., Big Data Smart City (ICITBS), Jan. 2019, pp. 264–266.

Zhou, T., Tao, D., Wu, X. 2011. Manifold elastic net: a unified framework for sparse dimension reduction. Data Mining and Knowledge Discovery 22 (3),340–371.

Zhou, Z., Wagner, A., Mobahi, H., Wright, J. & Ma. Y. 2009. Face recognition with contiguous occlusion using Markov random fields. International Conference on Computer Vision, IEEE, 1050–1057.