Evolutionary Intelligence (2022) 15:585–591
https://doi.org/10.1007/s12065-020-00531-4

RESEARCH PAPER

Patch-based pose invariant features for single sample face recognition

Wasseem N. Ibrahem Al-Obaydy[1] · Zainab Mahmood Fadhil[2] · Basheer Husham Ali[1]

Received: 6 March 2020 / Revised: 27 October 2020 / Accepted: 16 November 2020 / Published online: 27 November 2020

Abstract
Pose variation is considered as one of the major challenges that degrade the performance of face recognition systems. Existing techniques address this problem from different attitudes. However, these methods may be inefficient or impractical in the case of single sample face recognition. This article presents an automatic patch-based pose invariant feature extraction method that can handle pose variations for the aforementioned case. The proposed method extracts Gabor and histograms of oriented gradients features from landmark-based patches. The features are then concatenated, dimensionally reduced using principal component analysis, fused using canonical correlation analysis, and normalized using min-max normalization. Experimental results carried out on the FERET database have shown the outstanding performance of the proposed method compared to that of the state-of-the-art approaches. The proposed approach achieved 100% and 96% and 94.5% recognition rates for moderate and wide pose variations, respectively.

Keywords Patch-based feature extraction · Single sample face recognition · Pose invariant face recognition · Gabor magnitudes · Histograms of oriented gradients

1 Introduction

Face recognition has been widely studied by the vision community over the past few decades. The importance of this technology comes from its use in various applications such as law enforcement, security and access control. Recently, the research of face recognition has been directed towards the single sample face recognition (SSFR) [1]. However, recognizing human faces in the SSFR scenario is extremely challenging due to the presence of limited single reference samples in the gallery and the large sensitivity of intra-person variations for instance pose, illumination, facial expression and partial occlusion in probe images. In particular, pose variation is considered as the most complex problem that changes the out-of-plane rotations of the face resulting in self-occluded faces [2]. Such modification alters the shape and appearance of faces in a way that some discriminated facial details are lost due to self-occlusion. This loss in information leads to severe performance degradation of the frontal face recognition systems. A vast amount of pose invariant face recognition (PIFR) approaches has been introduced to address the pose variation problem from different perspectives. For comprehensive details on PIFR literature, the reader is referred to the recent surveys [2, 3]. However, most of the current PIFR approaches may be impractical in the SSFR scenario due to the reasons that are described at the end of Sect. 2.

The main contribution of this paper is a patch-based pose invariant feature extraction method that is efficiently applicable in the SSFR framework. The proposed method extracts pose invariant facial features from the landmark-based patches located at the face organs, namely eyebrows, eyes, nose, and mouth, rather than the whole face image. Since the local patches may expose relatively small out-of-plane rotations compared to the global image, extracting discriminated details from these small regions may produce pose invariant facial features.

The privileges of the proposed method over the existing PIFR methods are fourfold. In the first place, the proposed

✉ Wasseem N. Ibrahem Al-Obaydy
wasseem.nahi@aliraqia.edu.iq

Zainab Mahmood Fadhil
120094@uotechnology.edu.iq

Basheer Husham Ali
basheer.husham@aliraqia.edu.iq

[1] Computer Engineering Department, College of Engineering, Al-Iraqia University, Baghdad, Iraq

[2] Computer Engineering Department, University of Technology, Baghdad, Iraq

method can efficiently handle moderate pose variations (± 45° yaw) and show a significant performance under large poses (± 60° yaw) compared to the state-of-the-art pose robust feature extraction methods. In the second place, it does not need the abundant multi-pose training face data compared to the learning-based PIFR approaches. Finally, it is fully automatic and does not need manual landmark annotation.

The remaining part of this paper is organized as follows. Section 2 reviews related research attempts of PIFR approaches. In Sect. 3, the proposed method is described in detail. The experimental results carried out on a benchmarking face database are reported in Sect. 5. Finally, the conclusion of this paper is outlined in Sect. 6.

2 Related work

This section presents a review of the state-of-the-art PIFR approaches that addressed the pose problem from different points of view. According to Ding and Tao [2], the PIFR approaches can be classified into four categories namely, pose robust feature extraction, multi-view subspace learning, face synthesis, and hybrid approaches. The methods in the first category aim at designing face descriptors to extract discriminated facial features that are invariant against pose variations. For more details on face feature extraction approaches, the reader is referred to the recent survey [4]. The multi-view subspace learning-based methods use multi-pose face images to establish a shared latent subspace in which the features of different poses are projected and then matched. Face synthesis methods focus on transforming face images from one pose to another, so two faces originally in different poses can be matched in the same pose. Lastly, hybrid methods are simply a combination of two or more of the previous three groups. The next paragraphs explore the contemporary research endeavors in these categories and describe their limitations in the SSFR scenario.

Pose robust feature extraction methods extract facial features by either handcrafted or learning-based descriptors. The handcrafted approaches use a manually designed descriptor to extract features from either landmark-based or random facial keypoints-based patches. Zhou et al. [5] proposed a Huffman local binary pattern (LBP) to extract features from landmark-based patches. The authors applied a divide-and-rule strategy in both representation and classification to recognize faces across pose. Huang et al. [6] combined enhanced landmark-based multi-scale LBP (MSLBP) features with Gabor features by a proposed kernel-level fusion technique. Gao and Lee [7] presented a combined pose invariant scale invariant feature transform and personalized correspondence learning (PISIFT-PCL) method. The approach learns a generic correspondence between the poses to generate virtual patches from which the PISIFT features are extracted. The learning-based approaches extract facial features directly from the raw face images by machine learning techniques such as kernel-based and deep learning-based models. These methods learn to extract pose robust features by training on large-scale multi-pose face images. Duan and Tan [8] proposed a feature learning approach based on spatial self-similarity to extract the subject related information from a local feature by removing its pose related details. Shao et al. [9] proposed a pose invariant face representation learning approach based on sparse many-to-one encoders and a deep convolutional neural network. Ding and Tao [10] exploited convolutional neural networks (CNNs) to extract complementary facial features which are then compressed using a three-layer stacked auto-encoder (SAE).

The multi-view subspace learning-based PIFR approaches divide the nonlinear manifold of multi-pose face images into a separated set of pose spaces. Each pose space is considered as a single view from which pose specific projections to a shared latent subspace are learned. Guo et al. [11] utilized graph embedding to propose a multi-view linear discriminant analysis (MiLDA) for multi-pose face recognition. Cai et al. [12] proposed a regularized latent least square regression (RLLSR) method to map different poses of one person into a single point in the latent pose free space. Wang et al. [13] employed deep learning to design a deeply coupled autoencoder networks (DCAN) method to project samples from two poses into one common discriminating subspace.

Face synthesis methods generate a face image with the desired pose using a 2D or 3D model. In the 2D category, the process starts by fitting a 2D model to the face image and then 2D geometrical transformations (e.g., piecewise affine and thin-plate spline) are often used to warp face images to the desired pose. Sagonas et al. [14] introduced a robust statistical frontalization (RSF) technique based on the iterative procedure of both facial landmarks detection and warping to construct frontal face images. Haghighat et al. [15] proposed an improved version of active appearance models (AAM) by adopting automatic facial landmarks localization to enhance the AAM initialization part in the fitting procedure. Gao et al. [16] employed the discriminant appearance models (DAM) and partial least squares (PLS) to propose a view-based pose normalization method. The approaches in the 3D category accomplish face synthesis with the aid of a 3D model. Firstly, the 3D model is fitted to the face image based on facial landmarks. The face image texture is then mapped to the aligned 3D model. Lastly, the textured 3D model is rendered to the desired pose and a new synthesized face image is generated. Ding and Tao [17] invested a dense gird of 3D landmarks to design homography-based pose normalization (HPN) method. Deng et al. [18] developed a lighting-aware face frontalization approach depending on a five landmarks-based 3D generic model and quotient image

symmetry. Zhang et al. [19] adopted the use of a reference 3D face model, occlusion localization procedure, local face symmetry scheme, and Poisson image editing to design a face frontalization method. Recently, deep learning techniques have been incorporated into the 2D and 3D methods to achieve impressive face synthesis results. Huang et al. [20] proposed a deep architecture called a two-pathway generative adversarial network (TP-GAN) to synthesize a frontal face image exploiting the global structure and local texture of face image. Kan et al. [21] adopted a stacked progressive auto-encoders (SPAE) deep neural network to convert non-frontal face images to frontal views in a progressive manner.

Hybrid approaches consisting of two or more frameworks from the aforementioned categories have also been proposed for PIFR. Tran et al. [22] developed a disentangled representation learning-generative adversarial network (DR-GAN) approach to learn a generative and discriminative representation which can be used to synthesize frontal face image. Peng et al. [23] exploited a 3D model to synthesize multiple multi-view face images from which a rich feature embedding is learned by a deep neural network. Ding et al. [24] designed a hybrid approach based on combining a 3D-based frontal face synthesis, patch-based facial representation, and transformation dictionary-based subspace learning.

Despite the significant progress in PIFR research, most of the reported approaches may not be applicable in the strict SSFR condition for a number of reasons. Firstly, the performance of handcrafted pose robust feature extraction methods under large pose yaw $\pm 60°$ is still limited. Secondly, the learning-based pose robust feature extraction, multi-view subspace learning, and deep learning-based approaches require ample multi-pose training face images which are not available in the case of SSFR. Thirdly, face synthesis methods need to fit a 2D or 3D model to the face image in order to generate a face with the desired pose. The fitting process, however, may be time-consuming and computationally expensive. Moreover, some face synthesis approaches may require manual landmark annotation which is not feasible in real-world applications. They may also produce undesirable artifacts such as stretching due to inaccurate estimation of shape or pose parameters of the 2D or 3D model. These artifacts distort the face appearance and subsequently degrade the extracted features affecting the performance.

3 Proposed approach

In this section, the proposed patch-based pose invariant feature extraction method is described in detail. The main steps of the proposed approach include landmark detection, patch extraction, Gabor/HOG feature extraction, dimension reduction, feature fusion, and normalization. The diagram in Fig. 1 illustrates these steps. The details of each step are elaborated in the next subsections.

3.1 Landmark detection and patch extraction

Each face image is converted to the gray-scale. Then, 68 landmarks are detected using the constrained local neural field (CLNF) detection algorithm [25]. Only 51 landmarks located at the facial components, namely eyebrows, eyes, nose, and mouth, are used to determine patches. Each landmark represents the centroid of a patch $P(x, y)$ whose spatial size is predefined empirically for each dataset. Thus, 51 patches are segmented for patch-based local feature extraction. Due to the fact that local patches exhibit small pose variations compared to the global image, patch-based local feature extraction can produce features that are robust against pose variations.

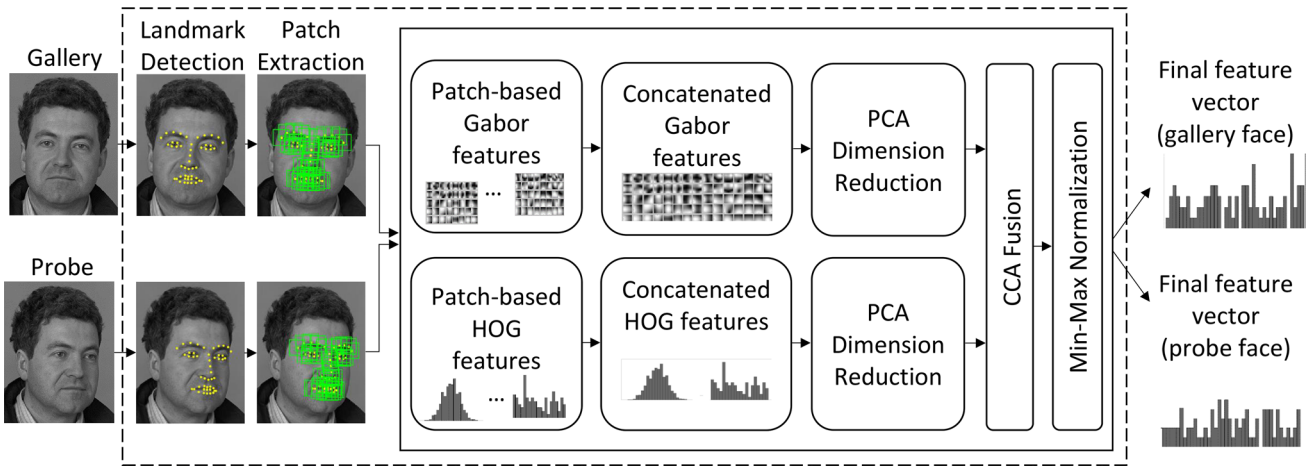

Fig. 1 The proposed patch-based pose invariant feature extraction method

3.2 Gabor/HOG feature extraction

Gabor filter [26, 27] and histograms of oriented gradients (HOG) [28] methods were adopted for local feature extraction. According to their original articles, Gabor and HOG features are robust against facial expression and illumination variations, respectively. To address the pose variation, Gabor magnitudes and HOG features were extracted from the 51 patches rather than the entire image.

Gabor features are extracted from each patch as follows. 40 Gabor filters of 5 scales and 8 orientations are defined as in (1)

$$\psi_{u,v}(x,y) = \frac{f_u^2}{\pi\kappa\eta} e^{-((f_u^2/\kappa^2)x'^2+(f_u^2/\eta^2)y'^2)} e^{j2\pi f_u x'} \tag{1}$$

where u and v define the scale and orientation, respectively, $x' = x\cos\theta_v + y\sin\theta_v$, $y' = -x\sin\theta_v + y\cos\theta_v$, $f_u = f_{max}/2^{(u/2)}$, and $\theta_v = v\pi/8$. The common parameters used for face recognition are $\kappa = \eta = \sqrt{2}$ and $f_{max} = 0.25$. Gabor features are extracted by filtering the patch $P(x, y)$ with the Gabor filter $\psi_{u,v}(x, y)$ as in (2)

$$G_{u,v}(x,y) = P(x,y) * \psi_{u,v}(x,y) \tag{2}$$

where $G_{u,v}(x, y)$ is the complex filtering output that is composed of real $R(G_{u,v}(x, y))$ and imaginary $I(G_{u,v}(x, y))$ parts. The magnitude $M_{u,v}(x, y)$ of filtering operation can be defined as in (3)

$$M_{u,v}(x,y) = \sqrt{R(G_{u,v}(x,y))^2 + I(G_{u,v}(x,y))^2} \tag{3}$$

All the 40 magnitude responses are downsampled by factor 5 and concatenated to form the Gabor feature vector of a single patch. All the Gabor vectors of 51 local patches are concatenated to construct the Gabor features of the global face image.

HOG features are extracted from each patch as follows. The gradient filter [−1, 0, 1] is used to compute the horizontal gradient $G_x(x, y)$ and vertical gradient $G_y(x, y)$ of the patch. The magnitude $|G(x, y)|$ and angle $\theta(x, y)$ of the gradient are defined as in (4) and (5), respectively

$$|G(x,y)| = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \tag{4}$$

$$\theta(x,y) = \arctan\left(\frac{G_y(x,y)}{G_x(x,y)}\right) \tag{5}$$

The patch is divided into cells, and each cell has 4×4 pixels. Then, a histogram of 10 evenly spaced orientation bins ranging from $0° - 180°$ is computed. Every bin is incremented by 1 when the magnitude ($|G(x, y)|$) whose angle ($\theta(x, y)$) belongs to the same bin. A block is formed by combining every four connected cells. The histograms of cells can be normalized in the block by $L2 - Hys$ (Lowe-style clipped $L2$ norm) normalization method. The combination of all histograms constructs the HOG feature vector of the patch. All HOG vectors from all local patches are concatenated to form the HOG feature vector of the global image.

3.3 Dimension reduction

Each of Gabor and HOG vectors has a high dimension which may slow down the performance. Many state-of-the-art approaches have been proposed for feature selection and dimension reduction [29–31]. However, these techniques have not been proved efficient for face recognition. A recent study [15] has shown that the use of principal component analysis (PCA) [32] with Gabor and HOG features yields significant face recognition performance. Hence, in the proposed method, the dimensional size of one-type vectors is reduced by PCA. Let the set of N feature vectors $\Gamma_1, \Gamma_2, \ldots, \Gamma_N$ be a training set. The average vector Γ_A of this set is defined as in (6)

$$\Gamma_A = \frac{1}{N}\sum_{n=1}^{N}\Gamma_n \tag{6}$$

The difference between each training vector and the average vector is defined as in (7)

$$\Phi_n = \Gamma_n - \Gamma_A \quad (n = 1, 2, \ldots, N) \tag{7}$$

The matrices $A = [\Phi_1\ \Phi_2\ \ldots\ \Phi_N]$ and $L = A^T A$ are constructed. Then, N eigenvectors e_i and eigenvalues of the matrix L are calculated. The eigenspace s_l can be defined as in (8)

$$s_l = \sum_{k=1}^{N} e_{lk}\Phi_k \quad (l = 1, \ldots, N) \tag{8}$$

The eigenspace represents a basis set from which the weights w_k can be obtained. The authors in [32] argued that a smaller N' dimensional eigenspace is sufficient to obtain weights. Only the N' eigenvectors with the highest eigenvalues are selected to generate the N' eigenspace. Each feature vector Γ is projected into the N' eigenspace to find its weight as in (9)

$$w_k = s_k^T(\Gamma - \Gamma_A) \tag{9}$$

where $k = 1, \ldots, N'$ and w_k is the contribution weight of kth eigenspace s_k. The resulted weights are grouped to construct the reduced feature vector $\Omega^T = [w_1, w_2, \ldots, w_{N'}]$.

3.4 Feature fusion and normalization

The reduced Gabor and HOG vectors are then fused by canonical correlation analysis (CCA) [33] to yield a more

discriminated and robust feature vector. Let $X \in \mathbb{R}^{p\times n}$ and $Y \in \mathbb{R}^{q\times n}$ are matrices of two feature sets from two different modalities. Let $S_{xx} \in \mathbb{R}^{p\times p}$ and $S_{yy} \in \mathbb{R}^{q\times q}$ are the within-sets covariance matrices of X and Y. Let $S_{xy} \in \mathbb{R}^{p\times q}$ is the between-set covariance matrix (note that $S_{yx} = S_{xy}^T$). CCA aims to find the linear combinations, $X^* = W_x^T X$ *and* $Y^* = W_y^T Y$, that maximize the pair-wise correlations across the two feature sets. The transformation matrices, W_x and W_y, are found by solving the eigenvalue equations defined in (10) and (11), respectively

$$S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} \hat{W}_x = \Lambda^2 \hat{W}_x \tag{10}$$

$$S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy} \hat{W}_y = \Lambda^2 \hat{W}_y \tag{11}$$

where $\hat{W}_x$ and $\hat{W}_y$ are the eigenvectors and Λ^2 is the diagonal matrix of eigenvalues or squares of the canonical correlations. The non-zero eigenvalues in each equation are then sorted in descending order. The transformation matrices, W_x and W_y, consist of the sorted eigenvectors corresponding to the non-zero eigenvalues. Thus, X^* and Y^* are calculated and known as canonical variates. Feature-level fusion is performed either by concatenating or adding the transformed feature vectors as defined in (12) and (13), respectively

$$Z_1 = \begin{pmatrix} X^* \\ Y^* \end{pmatrix} = \begin{pmatrix} W_x^T X \\ W_y^T Y \end{pmatrix} = \begin{pmatrix} W_x & 0 \\ 0 & W_y \end{pmatrix}^T \begin{pmatrix} X \\ Y \end{pmatrix} \tag{12}$$

$$Z_2 = X^* + Y^* = W_x^T X + W_y^T Y = \begin{pmatrix} W_x \\ W_y \end{pmatrix}^T \begin{pmatrix} X \\ Y \end{pmatrix} \tag{13}$$

where Z_1 and Z_2 are called the canonical correlation discriminant features. In this paper, the summation method defined in (13) is used. Finally, the fused vectors are normalized using min-max normalization.

4 Parameters settings

For the proposed method, there are two main parameters (i.e. the patch size and the number of eigenfaces) and other auxiliary parameters (i.e. Gabor filter's scale and orientation and HOG's cell size, block size, and bins number) need to be tuned for better recognition performance. The value of patch size is empirically set depending on the dataset. The eigenfaces' number is set to the number of training classes for optimal performance. In the experiments, the patch size was set to 40×40 pixels and the eigenfaces' number was set to 200 for the FERET b-series [34] dataset. Gabor filter's scale and orientation were set to 5 and 8, respectively. HOG's cell size, block size, and bins number were set to 4×4 pixels, 2×2 cells, and 10, respectively.

5 Experiments and results

In this section, an experimental evaluation of the proposed method is presented. The experiments were conducted on the publicly available database namely, FERET [34]. Two statistical classifiers, namely k nearest neighbor (kNN) with city-block distance function and support vector machine (SVM) were used to classify the proposed patch-based pose invariant features using a single sample per person. The experiments were implemented using MATLAB R2016a on a Windows 10 Professional laptop with Intel Core i7-3630QM CPU 2.4 GHz and 16 GB RAM. The b-series images of the FERET face database were used to evaluate the performance of the proposed method in comparison with several state-of-the-art approaches [5, 7, 8, 14–16, 21] that solved the pose variation problem from different perspectives. The b-series images in the FERET database are of size 256×384 pixels and were collected for 200 subjects. A subset of nine samples consisting of one frontal and eight pose varied images per subject were selected in the experiments. The frontal image is labeled as *ba* and holds a face in neutral conditions. The remaining eight non-frontal images include faces in different poses with $+60°$, $+45°$, $+30°$, $+15°$, $-15°$, $-30°$, $-45°$ and $-60°$ in yaw. These pose varied images are labeled as *bb*, *bc*, *bd*, *be*, *bf*, *bg*, *bh* and *bi*, respectively. Figure 2 shows face images of a sample subject. The size of patch was set to 40×40 pixels. Each subject in the gallery was represented by a single frontal face image *ba*, whereas the eight pose varied images *bb-bi* were presented in the testing operation. Table 1 reports the recognition rates of the proposed method and the peer approaches where the results of competed methods were transferred directly from their original papers. As can be seen in the table, the proposed method showed a comparable or better performance compared to the peer methods. The proposed scheme achieved 100% accuracy for pose variations with yaw between $+45°$ and $-30°$ as shown by the percentages highlighted in bold in the table. In the wide pose yaw $\pm 60°$, the proposed method achieved 96% and 94.5% rates outperforming the peer methods. This outstanding performance is due to the extraction of discriminated facial features from landmark-based patches located at the face organs rather than the entire face image. The experimental results under this constrained condition suggest that the proposed method is effective for a wide range of pose variations within $\pm$ 60° yaw.

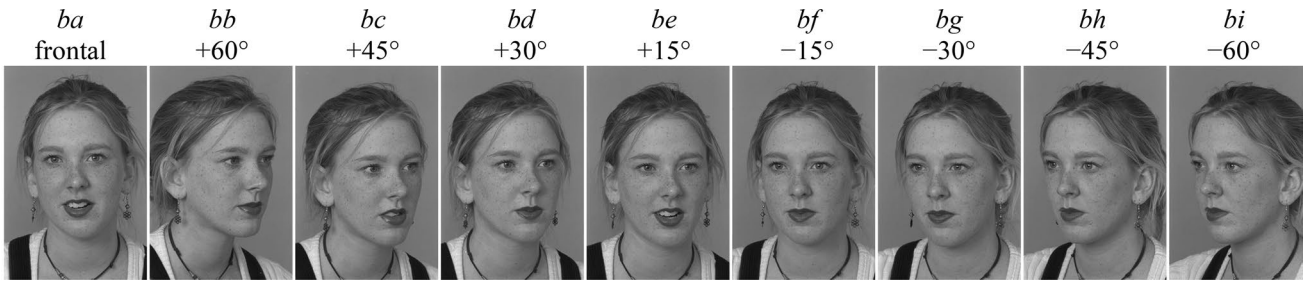

Fig. 2 Face images of a subject from FERET b-series database

Table 1 Recognition rates on FERET b-series database

Approach	Gallery subjects	+60°	+45°	+30°	+15°	−15°	−30°	−45°	−60°
SPAE [21]	200	77	95	99	99	99	98	96	77
DAM-PLS [16]	200	–	92	96	99	97	96	95	–
Haghighat et al. [15]	200	91.5	96	**100**	**100**	**100**	**100**	**99**	93
RSF [14]	200	–	96	**100**	**100**	**100**	99	96.5	–
PISIFT-PCL [7]	100	50	85	99	**100**	**100**	99	86	41
Duan and Tan [8]	100	73	92	97	98	99	98	94	83
Zhou et al. [5]	200	–	95	**100**	**100**	**100**	**100**	98	–
Proposed-*k*NN	200	95.5	**100**	**100**	**100**	**100**	**100**	97.5	93
Proposed-SVM	200	**96**	**100**	**100**	**100**	**100**	**100**	97.5	**94.5**

The bold values indicate the highest recognition rates

6 Conclusion

In this paper, a patch-based pose invariant feature extraction method is presented for single sample face recognition. This technique can be adopted to develop a face recognition system for large-scale identification applications, such as driver license, national ID card, or passport identification system in which only one training sample per person is enrolled in the database. The proposed approach consists of CLNF-based landmark detection, patch extraction, Gabor and HOG-based feature extraction, PCA-based dimension reduction, CCA-based feature fusion, and min-max normalization. The implementation of the proposed scheme was accomplished using MATLAB tool, and the performance was tested using FERET b-series database. The evaluation metric, namely recognition rate was used to evaluate the performance of the proposed framework in comparison with the recent approaches. Experimental results have shown the excellent performance of the proposed method under a wide range of pose variations. From the simulation results, it is evident that the proposed technique achieved 100% and 96% and 94.5% recognition rates for moderate and wide pose variations, respectively. Although the proposed feature extraction method demonstrated excellent performance results in pose variations with −45° and ±60° yaw, it still cannot reach 100% accuracy. This is the limitation of the currently proposed method, which leads to further investigations. In the future, the possibility to improve the performance of the proposed method under large pose variations in semi-profile and profile face images will also be investigated.

Conflicting interest

Conflict of interest The authors declare that there is no conflict of interest.

References

1. Tan X, Chen S, Zhou Z-H, Zhang F (2006) Face recognition from a single image per person: a survey. Pattern Recognit 39(9):1725–1745
2. Ding C, Tao D (2016) A comprehensive survey on pose-invariant face recognition. ACM Transactions on intelligent systems and technology (TIST) 7(3):37
3. Zhang X, Gao Y (2009) Face recognition across pose: a review. Pattern Recognit 42(11):2876–2896
4. Wang H, Jiani H, Deng W (2017) Face feature extraction: a complete review. IEEE Access 6:6001–6039
5. Zhou L-F, Yue-Wei D, Li W-S, Mi J-X, Luan X (2018) Pose-robust face recognition with Huffman-LBP enhanced by divide-and-rule strategy. Pattern Recognit 78:43–55

6. Huang K-K, Dai D-Q, Ren C-X, Yu-Feng Y, Lai Z-R (2017a) Fusing landmark-based features at kernel level for face recognition. Pattern Recognit 63:406–415
7. Gao Y, Lee HJ (2017) Pose-invariant features and personalized correspondence learning for face recognition. Neural Comput Appl 31:607–616
8. Duan X, Tan Z-H (2018) A spatial self-similarity based feature learning method for face recognition under varying poses. Pattern Recognit Lett 111:109–116
9. Shao M, Zhang Y, Yun F (2018) Collaborative random faces-guided encoders for pose-invariant face representation learning. IEEE Trans Neural Netw Learn Syst 29(4):1019–1032
10. Ding C, Tao D (2015) Robust face recognition via multimodal deep face representation. IEEE Trans Multimed 17(11):2049–2058
11. Guo Y, Ding X, Xue J-H (2015) Milda: a graph embedding approach to multi-view face recognition. Neurocomputing 151:1255–1261
12. Cai X, Wang C, Xiao B, Chen X, Zhou J (2013) Regularized latent least square regression for cross pose face recognition. In: Twenty-third international joint conference on artificial intelligence
13. Wang W, Cui Z, Chang H, Shan S, Chen X (2014) Deeply coupled auto-encoder networks for cross-view classification. arXiv preprint arXiv:1402.2031
14. Sagonas C, Panagakis Y, Zafeiriou S, Pantic M (2017) Robust statistical frontalization of human and animal faces. Int J Comput Vis 122(2):270–291
15. Haghighat M, Abdel-Mottaleb M, Alhalabi W (2016) Fully automatic face normalization and single sample face recognition in unconstrained environments. Expert Syst Appl 47:23–34
16. Gao H, Ekenel HK, Stiefelhagen R (2015) Combining view-based pose normalization and feature transform for cross-pose face recognition. In: 2015 international conference on biometrics (ICB). IEEE, pp 487–492
17. Ding C, Tao D (2017) Pose-invariant face recognition with homography-based normalization. Pattern Recognit 66:144–152
18. Deng W, Jiani H, Zhongjun W, Guo J (2017) Lighting-aware face frontalization for unconstrained face recognition. Pattern Recognit 68:260–271
19. Zhang Y, Qian J, Yang J (2016) Robust face frontalization in unconstrained images. In: Chinese conference on pattern recognition. Springer, pp 225–233
20. Huang R, Zhang S, Li T, He R (2017) Beyond face rotation: global and local perception GAN for photorealistic and identity preserving frontal view synthesis. In: IEEE international conference on computer vision (ICCV). IEEE, pp 2458–2467
21. Kan M, Shan S, Chang H, Chen X (2014) Stacked progressive auto-encoders (SPAE) for face recognition across poses. In: Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, pp 1883–1890
22. Tran L, Yin X, Liu X (2018) Representation learning by rotating your faces. IEEE Trans Pattern Anal Mach Intell 41:3007–3021
23. Peng X, Yu X, Sohn K, Metaxas DN, Chandraker M (2017) Reconstruction-based disentanglement for pose-invariant face recognition. In: Proceedings of the IEEE international conference on computer vision, pp 1623–1632
24. Ding C, Chang X, Tao D (2015) Multi-task pose-invariant face recognition. IEEE Trans Image Process 24(3):980–993
25. Baltrusaitis T, Robinson P, Morency LP (2013) Constrained local neural fields for robust facial landmark detection in the wild. In: Proceedings of the IEEE international conference on computer vision workshops, pp 354–361
26. Štruc V, Pavešić N (2010) The complete gabor-fisher classifier for robust face recognition. EURASIP J Adv Signal Process 2010(1):847680
27. Liu C, Wechsler H (2002) Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition. IEEE Trans Image process 11(4):467–476
28. Navneet Dalal, Bill Triggs (2005) Histograms of oriented gradients for human detection. In: IEEE computer society conference on computer vision and pattern recognition, 2005. CVPR 2005, volume 1. IEEE, pp 886–893
29. Zhang Y, Li H-G, Wang Q, Peng C (2019) A filter-based bare-bone particle swarm optimization algorithm for unsupervised feature selection. Appl Intell 49(8):2889–2898
30. Zhang Y, Gong D, Gao X, Tian T, Sun X (2020) Binary differential evolution with self-learning for multi-objective feature selection. Inf Sci 507:67–85
31. Zhang Y, Song X, Gong D (2017) A return-cost-based binary firefly algorithm for feature selection. Inf Sci 418:561–574
32. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3(1):71–86
33. Sun Q-S, Zeng S-G, Liu Y, Heng P-A, Xia D-S (2005) A new method of feature fusion and its application in image recognition. Pattern Recognit 38(12):2437–2448
34. Phillips PJ, Moon H, Rizvi SA, Rauss PJ (2000) The FERET evaluation methodology for face-recognition algorithms. IEEE Trans Pattern Anal Mach Intell 22(10):1090–1104