Single Sample Face Identification Utilizing Sparse Discriminative Multi Manifold Embedding

Fatemeh Shahali
School of Electrical and
Computer Engineering
Shiraz University
Shiraz, Iran
Email:
f.shalali@cse.shirazu.ac.ir

Azadeh Nazemi
School of Electrical and
Computer Engineering
Shiraz University
Shiraz, Iran
Email:
a.nazemi@cse.shirazu.ac.ir

Zohreh Azimifar
School of Electrical and
Computer Engineering
Shiraz University
Shiraz, Iran
Email:
azimifar@cse.shirazu.ac.ir

Abstract—This paper describes three methods to improve single sample dataset face identification. The recent approaches to address this issue use intensity and do not guarantee for the high accuracy under uncontrolled conditions. This research presents an approach based on Sparse Discriminative Multi Manifold Embedding (SDMME), which uses feature extraction rather than intensity and normalization for pre–processing to reduce the effects of uncontrolled condition such as illumination. In average this study improves identification accuracy about 17% compare to current methods.

Index Terms—**Face Identification, Sparse Discriminative Multi Manifold Embedding (SDMME), Single Sample dataset, Feature extraction, Self Quotient Image (SQI)**

I. INTRODUCTION

IN recent decade, face identification is the base of several researches in pattern recognition, image processing and computer vision area. Face identification is used in many security, recognition and monitoring applications. One of the most problematic challenges in this area is lack of sufficient training samples for each class. In the best case scenario the dataset for face identification contains normal, posed, illuminated, expressed and occluded images of each class. Since in real world providing such a dataset for different applications is not feasible, thus this research makes effort to use small sample size dataset for face identification. Using small sample size dataset in identification purposes leads to reduce computational costs, storage and processing time. Sparse Discriminative Multi Manifold Embedding (SDMME) is a recent method applied on extended Yale B as a small sample dataset [1]. Since SDMME method is based on dividing each image to patches, then it models the internal connections between patches. In the next step SDMME generates discriminate features for each class. Finally, it minimizes manifold variances. During the matching process global manifold distance is considered as a similarity regulator. Considering SDMME is a high accuracy method in small sample size dataset area, this research aims to modify it by using Local Binary Pattern (LBP) fracture extraction rather than patches intensity.

The rest of this paper was organized as the following:
Related works are described in Section II, Proposed method is available in Section III, Section IV represents Evaluations and experimental results and finally, Section V contains conclusion and further works.

II. RELATED WORKS

Generally, face identification systems contain various steps: As it is shown in Figure 1, the first step is pre-processing; the next step depends on used method extracts features from pre-processed facial images. All extracted features are sent to training step which leads to classify facial images. Using an appropriate method for feature extraction and selecting suitable classifier provide opportunity to achieve accurate face identification [2].

Based on related literature, face identification methods are categorized into Holistic and local approaches. Holistic methods learn a linear or nonlinear sub-space to extract low dimension features and model the relation between features such as Principal Component Analysis (PCA) method. They need a large training dataset consequently, cannot support face identification application with single sample due to face on overfit issues and lose generalization ability. Therefore, local methods are considered as alternatives[3].
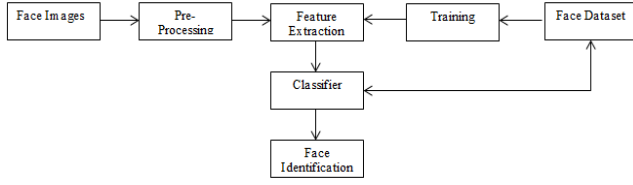
Figure. 1: Block diagram for face identification overview.

Local methods extract face components features such as eyes, nose and mouth using edge detection then classify them utilizing local statistics. The local method advantages include considerable reduction in feature dimensions and discriminative representation of an image with these features. Moreover, local methods are robust against illumination and corruption [4]. However, face identification based on only local feature causes to destroy holistic information. Therefore, it is essential to combine both local and holistic to get a better result.

An example for category of local methods is Discriminative Multi Manifold Analysis (DMMA) which learns discriminative features from image patches. It divides an image to local patches as data points. The target is to learn N various feature spaces from N classes [5].

Another approach of local category is Sparse Neighborhood Preserving Embedding (SNPE). To determine the eternal connection between patches in an image, this method considers a parameter such as 'k' which denotes the number of patches in neighborhood. It aims to find optical embedding reconstructing sparse graph. Which leads to get similarity between neighborhoods [6]. As stated before, this research uses SDMME as the base method and claims to decrease recognition error rate. The SDMME implementation includes two major modules.

1) Making sparse graphs 2) Making SDMME

The first module consists of patching, normalization, initialization, optimization and construction. The second module is made of optimal embedding and extraction. The steps of each module are briefly described here:
 Initially, each training images is converted to M patches without any overlap then all patches are represented as vectors. In order, to obtain a unit norm all vectors are normalized. Then both inter and intra manifold dictionaries are initialized. Eventually, inter and intra manifold graph are created. In second module projection matrix is obtained by calculation both inter and intra manifold then all eigenvectors corresponding to non-zero eigenvalues are found and extracted. The next section describes methodology of this research based on SDMME in details.

III.   PROPOSED METHOD

This proposed method is divided into two main steps, which are explained here:

1) Representation Module:
Since the images under study in this research are illuminated under various light on conditions of different sides, then Self Quotient Image (SQI) is used to decrease the light's effect on images [7]. Initially face boundary is detected in cropped using Dlib library [8] and is aligned based on eyes and bottom lip. Then, aligned images are divided into M patches without any overlap. In the next step the Local Binary Pattern (LBP) features are extracted from each patch. These features determine some information regarding textures are extracted from each patch. In small size dataset each class has only one single sample, thus by dividing images into M patches, each class contains M patches. Considering the length of LBP features (1×256), each patch is converted into 1×256 features vector. In the statement $X_i = \{x_{i,j} \mid j = 1, \dots, M \}$, while $x_{i,j}$ is a single feature, if i, j respectively indicate the classes label and patches label.

A.1 Dictionary Construction:
Following feature extraction, two dictionaries are defined namely inter-class and intra-class. Assuming X denotes training dataset includes N classes with single sample and each class is represented by M patches. ($X = \{X_1, \dots, X_N\}$)

- Inter-class dictionary:
$$A_i^{inter} = X/X_i = [x_{1,1}, \dots, x_{i-1,M}, x_{i+1,1}, \dots, x_{N,M}]$$
$$\in \mathcal{R}^{d \times (N-1)M} \tag{1}$$
and
- Intra-class dictionary:

$$A_i^{intra} = X_i/x_{i,j} = [x_{i,1}, \dots, x_{i,j-1}, x_{i,j+1}, \dots, x_{i,M}]$$
$$\in \mathcal{R}^{d \times (M-1)} \tag{2}$$

A.2 Making Graph:
Dictionary definition supports algorithm to make two graphs which are namely inter-class and intra-class.
Intra-class graph expresses within-class compactness and inter-class graph expresses inter-class separability.
In order to initialize the graphs, the following optimization formula must be solved:

- *Inter-class graph optimization:*
$$\alpha_{i,j}^{inter} = \arg\ \min_i \parallel \alpha \parallel_1$$
$$subject\ to\ \parallel x_{i,j} - B_i^{inter}\beta \parallel_2 < \lambda \tag{3}$$

where, $B_i^{inter} = [A_i^{inter}, I] \in \mathcal{R}^{d \times [(N-1)M+d]}$ and $\beta = [\alpha; \alpha_e]$.

Therefore, j^{th} columnes of ith class of inter-class graph is:

$$W_{i,j,k}^{inter} = \begin{cases} |\alpha_{i,j,k}^{inter}| & 0 < k \leq (i-1)\,M \\ 1 & (i-1)M < k \leq iM \\ |\alpha_{i,j,k-M}^{inter}| & iM < k \leq NM \end{cases} \tag{4}$$

where, $W_{i,j,k}^{inter}$ and $\alpha_{i,j,k}^{inter}$ are, respectively, k^{th} element of $W_{i,j}^{inter}$ and $\alpha_{i,j}^{inter}$. The inter-class graph of the ith class is:

$$W_i^{inter} = \left[w_{i,1}^{inter}, \ldots, w_{i,M}^{inter}\right] \in \mathcal{R}^{NM \times M}$$

For intra class graph:

$$\alpha_{i,j}^{intra} = \text{arg min}_i \parallel \alpha \parallel_1$$
$$subject\ to\ \parallel x_{i,j} - B_i^{intra}\beta \parallel_2 < \lambda \qquad (5)$$

where, $B_i^{intra} = \left[A_i^{intra}, I\right] \in \mathcal{R}^{d \times [(M-1)+d]}$ and $\beta = [\alpha; \alpha_e]$.
Therefore, j^{th} columnes of ith class of intra-class graph is :

$$W_{i,j,k}^{intra} = \begin{cases} |\alpha_{i,j,k}^{intra}| & 0 < k < j \\ 1 & k = j \\ |\alpha_{i,j,k-1}^{inter}| & j < k < M \end{cases} \qquad (6)$$

where, $W_{i,j,k}^{intra}$ and $\alpha_{i,j,k}^{intra}$ are respectively k^{th} element of $W_{i,j}^{intra}$ and $\alpha_{i,j}^{intra}$. The intra-class graph of the ith class is:

$$W_i^{intra} = \left[w_{i,1}^{intra}, \ldots, w_{i,M}^{intra}\right] \in \mathcal{R}^{M \times M}.$$
All equations from (1) to (6) were presented by Zhang et al[1].

A.3 The Graph Specification:

Graphs construction constraints three properties:

1) All graph points are connected and W values obtained by (4) and (6), indicate the weight of graph's edges.
2) All samples of classes are sparsely represented. That approves similarity between sample in the same class is more than others.
3) All patch features during graph construction are unique, to avoid unbalancing dictionary words.

Now, the graph structure obtained from previous section appears in manifold space without any modification. In this stage for each training sample data the optimum projection matrix must be found $p_i \in \mathcal{R}^{M \times D_i}$, then N projection matrices are represented $P = [p_1, \ldots, p_N]$. This projection matrix is obtained by solving the generalized eigenvalue:

$$\left(M_i^{intra} - M_i^{inter}\right)p_k = \lambda_k p_k \qquad (7)$$

where,

$$M_i^{intra} = \sum_{j=1}^{M}\left(x_{i,j} - X_i W_{i,j}^{intra}\right)\left(x_{i,j} - X_i W_{i,j}^{intra}\right)^T$$

and

$$M_i^{inter} = \sum_{j=1}^{M}\left(x_{i,j} - X W_{i,j}^{inter}\right)\left(x_{i,j} - X W_{i,j}^{inter}\right)^T .$$

Here, we practically assume that:
$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0 > \lambda_{d+1} \geq \cdots \geq \lambda_D$. It only needs to extract eigenvectors corresponding to positive eigenvalues as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$, then consider them as the projection matrix. Then all data pointes are projected to manifold space without any overlaps [1].

2) Classification Module:

Considering u_i is a test sample to the system to be matched with a trained class. Initially, it is converted to t patches as it stated before for training stage. Then, the sum of the distances of patches to $manifold_i$, $i = 1, \ldots, N$ are calculated. The shortest distance is related to the class closest to the test class. The label of this class is obtained by solving this optimization:

$$d(u, M_i) = \sum_{j=1}^{t} d\left(u_j, M_i\right)$$

where

$$d\left(u_j, M_i\right) = \text{min} \parallel p_i^T u_j - \sum_{l=1}^{M} \pi_l p_i^T m_{il} \parallel^2$$
$$subject\ to\ \sum_{l} \pi_l = 1$$

where $m_{i,l}$ is l^{th} vector of M_i and π_l is a weighted factor corresponding to $m_{i,l}$.

L is considered as the label of test sample if L is satisfied in:
$$L = \text{arg min}_i\ d(u, M_i)\ , i = 1, 2, \ldots, N \qquad (8)$$
Therefore L indicates the class of test sample in training dataset with single sample classes.

IV. EVALUATION AND EXPERIMENTAL RESULTS

This section compares six current methods such as PCA and $(PC)^2A$ [9], Locality preserving projection (LPP) [10], SNPE [4], DMMA [3] and SDMME+GD [1] with three methods proposed in this paper. All results are reported by applying different approaches on the same dataset. Extended Yale B dataset includes 16408 frontal face images of 28 classes with different illumination. This dataset was categorized into two main subsets: train and test. Subset A represents training dataset includes one normal frontal face of each class without any illumination as it is shown in Figure 2.

Testing dataset was divided into five sub-datasets (B, C, D, E and F) and contains frontal face with various illumination angles, that illumination increasing respectively from B to F as illustrated in Figure 3. In addition, the top level of each sub-dataset in Figure 3 illustrates illuminated image prior to pre-processing and bottom level of each shows sub-dataset of testing samples pre-processed by SQI.

Table I indicates results of applying nine methods on five sub-datasets. Evidently, the accuracy is improved by this research approaches specifically on sub-dataset E and F. the accuracy in obtained by using following formula:

$$\text{Accuracy} = \left(\frac{\text{the number of correct recognition}}{\text{the total number of testing sample}}\right) * 100.$$

Experimental results confirm the following actions considerably improve accuracy and performance:
- Normalization to decrease illumination effect on frontal face images.
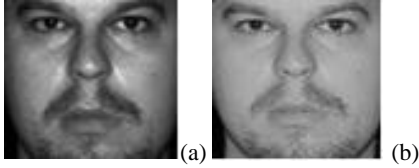- Creating dictionary using LBP descriptor rather than intensity.

Figure 2: Training image: (a), normal face image of prior to SQI normalization and (b), normal face following SQI normalization.

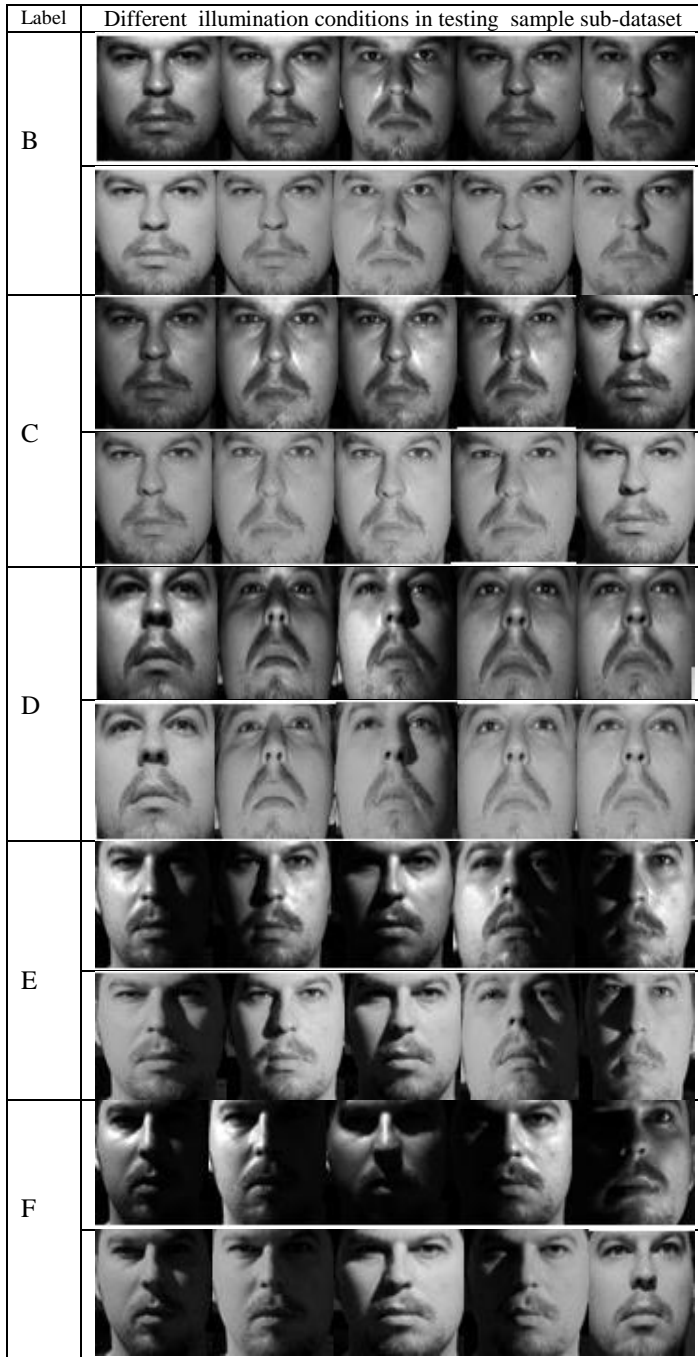| Label	Different illumination conditions in testing sample sub-dataset
B	
C	
D	
E	
F	

Figure 3: various testing sub-dataset. Top: illuminated images prior to pre-processing by SQI. Bottom: pre-processed images by SQI.

As it shown in table I, **S**QI + SDMME + GD(our method)

represent accuracy SQI normalization and intensity.

SDMME + GD + LBP(our method)   represent using LBP features without SQI normalization.

SQI + SDMME + GD + LBP(our method)   represent both SQI normalization and LBP features.

TABLE I
The accuracy comparison between nine different approaches for face identification on the extended Yale B

Method	Subset B%	Subset C%	Subset D%	Subset E%	Subset F%	Accuracy average %
PCA	99.34	95.39	76.3	32.24	19.08	64.47
$(PC)^2A$	96.05	85.53	83.5	36.18	40.79	68.41
LPP	99.34	98.68	82.2	35.53	33.55	69.86
SNPE	97.37	96.05	94.7	47.37	38.16	74.73
DMMA	96.71	97.37	98.0	40.79	34.21	73.42
SDMME + GD	99.34	98.68	98.0	58.55	42.76	79.46
SDMME + GD + LBP (**our method**)	**99.44**	**98.42**	**98.0**	**62.67**	**48.89**	**81.48**
SQI + SDMME + GD + LBP (**our method**)	**99.44**	**99.47**	**99.2**	**68.67**	**59.78**	**85.31**
SQI + SDMME + GD (**our method**)	**99.44**	**98.42**	**98.4**	**65.33**	**55.56**	**83.54**

V. CONCLUSION AND FUTURE WORKS

The research, studied three approaches to address the issue regarding face identification with single sample training dataset. The approaches are implemented using extended Yale B a dataset when applying SQI normalization in order to reduce illumination. In addition to constructing dictionaries, intensity points are replaced with LBP features extracted from image patches. Since this study is based on SDMME, all images are divided into certain number of patches without any overlap, then each patch is considered as a data point. To model these data points, each class is laid on an individual manifold. Besides, global manifold distance is utilized for classification purpose. Excremental result indicate the proposed method causes to improve accuracy by 85.31% for face identification under uncontrolled condition with single sample dataset. Further developments will be to extend the proposed methods and using other descriptors to challenge with occulted faces.

REFERENCES

[1]   P. Zhang, X. You, W. Ou, C. L. Philip Chen, and Y. M. Cheung, 'Sparse discriminative multi-manifold embedding for one-sample face identification,' Pattern Recognit., vol. 52, pp. 249–259, 2016.

[2]   V. P. Kshirsagar, M. R. Baviskar, and M. E. Gaikwad, "Face recognition using Eigenfaces," *2011 3rd Int. Conf. Comput. Res. Dev.*, vol. 2, pp. 302–306, 2011.

[3]   K. Choudhary and N. Goel, "A review on face recognition techniques," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 8760, no. 7, p. 87601E, 2013.

[4]   G. D. P. Carlose, H. Pedrini, and W. R. Schwartz, "Classification schemes based on Partial Least Squares for face identification ," 2015.

[5]   J. Lu, Y. P. Tan, and G. Wang, "Discriminative Multimanifold Analysis for Face Recognition from a Single Training Sample per Person," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 39–51, Jan. 2013.

[6]   B. Cheng *et al.*, "Learning With ` 1 -Graph for Image Analysis," vol. 19, no. 4, pp. 858–866, 2010.

[7]   H. Wang, S. Z. Li, Y. Wang, and J. Zhang, "Self quotient image for face recognition," in *Image Processing, 2004. ICIP '04. 2004 International Conference on*, 2004, vol. 2, p. 1397–1400 Vol.2.

[8]   "dlib." [Online]. Available: https://pypi.python.org/pypi/dlib. [Accessed: 01-Apr-2017].

[9]   J. Wu and Z. Zhou, "Face Recognition with One Training Image per Person," vol. 23, no. 14, pp. 1711–1719, 2002.

[10]   Y. Xu, A. Zhong, J. Yang, and D. Zhang, "LPP solution schemes for use with face recognition," *Pattern Recognit.*, vol. 43, no. 12, pp. 4165–4176, 2010.