A Robust Sparse Representation based Face Recognition System for Smartphones

Mahdi Abavisani[†], Mohsen Joneidi[♯], Shideh Rezaeifar[§], Shahriar Baradaran Shokouhi[◇]
[†]Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.
[♯]Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran.
[§]Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.
[◇] Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.

***Abstract*—Many research works have been done in face recognition during the last years that indicates the importance of face recognition systems in many applications including identity authentication. In this paper we propose an approach for face recognition which is suitable for unconstrained image acquisition and has a low computational cost. Since in practical applications such as in smartphones, imaging conditions are not limited to existing images in the database, robustness of the recognition algorithm is very important. Here a sparse representation framework is proposed which achieves some degree of robustness. Using double sparse representation the high computational cost of sparsity-based classifications is resolved. The experimental result indicates that the proposed method outperforms other well-known algorithms in terms of robustness and it is still fast enough for real-time applications in smartphones.**

***Index Terms*—Discriminative Dictionary Learning; Face Recognition; Fast Sparse Representation; Smartphones; Sparse Representation Based Classification.**

I. Introduction

Face recognition is beneficial in many applications including security applications such as identifying white and black lists and access controls, or other applications like tagging image albums in social networks. However, face recognition systems usually encounter different limitation and challenges to the extent that the database is captured under controlled conditions. Whereas in real situations, their performance has been shown to be brittle because of difference among training samples and what in fact the system faces in practice. This difference may occur due to irregular illumination, improper pose or perhaps occurrence of occlusions. These problems are usually caused by lack of user cooperation, uncontrolled illumination and target movements.

Existing approaches to deal with aforementioned problems in face recognition systems are mainly categorized into online and offline techniques. Methods in the first category, such as feature extraction from image derivatives, illumination normalization and 3D face modeling, confronts the face recognition challenges by estimating condition of the the image illumination. The major advantage of these methods is that there is no need for prediction and training of different possible conditions, hence, face recognition in new situations is feasible. On the other hand, these methods have a high level of time and memory complexity for the recognition process. Since the computational power is usually limited in smartphones, using online methods is troublesome and hardly feasible. Offline methods, instead, extend the range of possible image conditions for each person by utilizing a group of different image conditions for training system. Methods that are based on sparse representation, Eigenface methods [1], and Fisherface methods [2] fall into this category. Offline methods often provide a high level of robustness and accuracy. They transfer the computational complexity to the training stage and before practical usage of the system. Thus, due to the fact that if the training step is offline, it can be done on a stronger remote server, and a simpler process is required for the recognition step, using offline methods seems to be a better choice for smartphones. However, there exist another issue; training systems with sample data from available databases is straightforward and simple, but taking a group of different pictures with all possible conditions for making a comprehensive training set is in fact, hardly feasible. Therefore, a practical face recognition system with offline methods should be designed in a way that registering people becomes an easier procedure. This paper, proposes a novel illumination robust face recognition system that assign a low computational and memory complexity to the system for the recognition process while it uses only one image for each person to register.

The present paper proposes to build various possible image conditions artificially using 3D face modeling in [3] and then uses these images for training the system. Using sparse representation of data for classification, the proposed method is an efficient solution for classification of high dimensional data as in face recognition systems. It is shown that sparse representation based methods not only outperform conventional methods such as Nearest Neighborhood, Nearest Subspace and Support Vector Machine (SVM) under illumination variation but also is able to confront occlusion challenges in images due to accessories like scarves, sunglasses or beards [4] . Moreover, the proposed method uses the double sparse representation technique that makes the computational complexity reduced. This choice makes usage of the sparse representation based face recognition possible for smartphones.

The rest of this paper is organized as follows. Section II briefly investigates sparse representation and sparse based classification. In section III, existing challenges and proposed solutions are introduced. Section IV presents experimental results in the PIE database and under variate illuminations and poses. Finally, Section V concludes the paper with a brief summary and discussion.

II. Sparse representation classification

Sparse representations have recently gained much attention in many image processing applications including image restoration, image de-noising , object classification, image segmentation and target tracking [5]–[11]. This is mainly due to the fact that many signals including natural images are sparse within a specific domain. Among all these application, sparse representation based classification has been widely used due to the discriminative nature of sparse representation. Suppose that we have,

$$\mathbf{x} = \sum_{i=1}^{m} \gamma_i \boldsymbol{\psi}_i = \boldsymbol{\Psi\gamma} \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the observation and we aim to factorize it to $\boldsymbol{\gamma}$ and $\boldsymbol{\Psi}$ where $\gamma \in \mathbb{R}^m, \boldsymbol{\Psi} \in \mathbb{R}^{n\times m}$.

$\boldsymbol{\gamma}$ is a representation of $\mathbf{x}$ in the domain which $\boldsymbol{\Psi}$ specifies its bases. We call $\boldsymbol{\gamma}$ as the sparse representation of $\mathbf{x}$ if the matrix $\boldsymbol{\Psi}$ be chosen in a way that $\boldsymbol{\gamma}$ has as few as possible non-zero coefficients.

For solving equation (1), if $m > n$ and rank of $\boldsymbol{\Psi}$ is equal to n. There is an infinite number of possible answers, but the most sparse one is desired. This additional constraint results a unique solution for an underdetermined system of linear equations under a set of certain conditions [12]. The problem of solving the sparsest answer of equation (1) can be written as:

$$\min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_0 \quad \text{s.t.} \quad \mathbf{x} = \boldsymbol{\Psi\gamma} \tag{2}$$

where $\|.\|_0$ is $\ell_0 - norm$ which is a function that indicates the number of non-zero entries. Function $\|.\|_0$ is non-convex and hence the $\ell_0 - norm$ minimization problem is known to be NP-hard. Greedy algorithms such as matching pursuit (MP), and orthogonal matching pursuit (OMP) have been proposed to solve the $\ell_0 - norm$ minimization problem. To obtain a robust decomposition and sparse representation, a bound on noise is considered, the equality constraint in (2) can be changed to $\|\mathbf{x} - \boldsymbol{\Psi\gamma}\|_2 < \epsilon$. Thus, the sparsest solution for a proper value of parameter λ can be recovered by,

$$\min_{\boldsymbol{\gamma}} \|\boldsymbol{\gamma}\|_0 + \lambda \|\mathbf{x} - \boldsymbol{\Psi\gamma}\|_2 \tag{3}$$

By proper selection of the regularization parameter λ, noise of recovery can be reduced. In addition, obtaining a well-represented sparse representation has a straightforward connection with fitting the bases to the data; hence, one method of choosing $\boldsymbol{\Psi}$ is using prototypes of signal $\mathbf{x}$. Assume that there are K classes and m_k training samples for the k^{th} class, i.e., $k = 1, \ldots, K$. Thus, by assembling m_k columns of samples for the k^{th} class the sub-dictionary $\boldsymbol{\Psi}_k \in \mathbb{R}^{m_k \times n}$ results. Finally, the dictionary $\boldsymbol{\Psi} \in \mathbb{R}^{m\times n}$ is formed by concatenating sub-dictionaries together as,

$$\boldsymbol{\Psi} = [\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, ..., \boldsymbol{\Psi}_K] \tag{4}$$

Therefore, if an observation $\mathbf{x}$ belongs to the k^{th} class, the model says that it can be well represented by linear combination of k^{th} class data samples (i.e. by using just the $\boldsymbol{\Psi}_k$). When the class of $\mathbf{x}$ is unknown, index of a $\boldsymbol{\Psi}_k$ that contains the largest group of non-zero coefficients determines class of the observation $\mathbf{x}$. In other words, the observation can be represented as follows,

$$\mathbf{x} = \boldsymbol{\Psi}_1\boldsymbol{\gamma}_1 + \boldsymbol{\Psi}_1\boldsymbol{\gamma}_2 + ... + \boldsymbol{\Psi}_K\boldsymbol{\gamma}_K, \tag{5}$$

in which, $\boldsymbol{\gamma}_k \in \mathbb{R}^{m_k}$. The observed signal then can be classified as follows,

$$\text{class}(\mathbf{x}) := Arg \max_k \|\boldsymbol{\gamma}_k\|_0 \tag{6}$$

This method is the ordinary approach in sparse based classification systems. As the the signal prototypes make the basis matrix $\boldsymbol{\Psi}$, when there are a huge number of prototypes the basis matrix can become too large; while, due to possibility of existence of high correlated columns, many of its columns might be useless for representing $\mathbf{x}$ signals. Thus, it is also possible to use an optimum learned dictionary instead of the basis dictionary for representing sparse codes [13], [14]. Furthermore, to adapt the approach for smartphone applications, further improvement is proposed in the next section.

III. Designing the proposed system

Since this system is proposed to be employed on smartphones, additional considerations are taken into account. First, due to the fact that image acquisition conditions are not limited in mobile devices, the system should provide high level of robustness. Second, because of limits in computing power of mobile devices, computational complexity should be as low as possible.

Robustness of the system is further classified into two categories of robustness to processing challenges, e.g., low quality images, and robustness to non-ideal conditions of image acquisition such as pose and illumination variation.

In this section each, at first, challeng and its proposed solution is mentioned. Then, the proposed face recognition system is explained.

A. Robustness

1) Non-ideal image acquisition conditions challenges: In practical situations, the image acquisition conditions are not fully under control, i.e. the direction of the face to the camera is not always strictly frontal, or also we might have lightening variation in images. As mentioned before, the approaches for dealing with these challenges are classified into online and offline methods. Offline methods are preferable and more suitable as the computation power of mobile devices is limited. Therefore, images that are taken in different acquisition conditions should be used for the training stage of the system. It is clear that taking various images from different angles and with

Fig. 1. Some examples of various illumination conditions that are artificially rendered for the the person with tag number of 4006 from PIE face database.

different illuminations for registering a person in the system is far away from practical implementation of a face recognition system. To overcome this challenge, a 3D face modeling is exploited for registration of the people in the system. At the first, a 3D face model is fitted to a single face picture of the person. Then, the 3D facial model is applied to make artificial images from different possible conditions for training set.

Many algorithms are proposed and investigated for 3D face modeling such as shape from shading (SFS) [15], and Morphable model [16] based algorithms, or Basel University's Face Model [3], which is to the best of authors knowledge, one of the state of the arts facial reconstruction methods. Figure 1 shows some samples of artificial lighting on a 3D face model that is fitted to a face image from PIE face database by using the Basel University's face model 3D facial reconstruction method.

2) Robust Classification: Since in practical applications such as in smartphones, images are usually not taken in studios and are not under controlled conditions, low quality of cameras, target movement or occlusion may cause problems in images. These challenges can be resolved using a robust classification method [4]. Recent studies have shown that sparse representation based classifications are of the most robust techniques [4], [17], [18]. Hence, this technique is employed in the classification stage of the proposed system.

B. Limited time and space power

Despite rapid growth of the mobile devices, their computational power is still not sufficient for implementation of highly-complex systems. In addition, the increase of hardware complexity will require spending more cost; hence, in practical applications, designing simpler systems is desirable. On the other hand, the necessity of having a robust face recognition system for implementation in smartphones, requires high level of computation. Although computation in sparse domain due to fewer numbers of non-zero coefficients is much less than computation in space domain, finding the sparse representation of data itself, needs an operation with high computational complexity. To reduce computation cost and memory requirement, a double sparse representation technique is proposed. This method is motivated by, and is combination of sparse dictionary learning, presented in [19], and the sparse representation technique presented in [20].

The sparse K-SVD dictionary learning algorithm trains a dictionary, namely a dictionary tantamount to $\mathbf{\Psi}$, in a way that the dictionary is product of a basic dictionary and a sparse dictionary. Its optimization problem can be written as:

$$\min_{\boldsymbol{A},\boldsymbol{\Gamma}} \|\boldsymbol{X} - \boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{\Gamma}\|_F^2 \; s.t. \begin{cases} \forall i \; \|\gamma_i\|_0 \leq t \\ \forall j \; \|\boldsymbol{a}_j\|_0 \leq p, \; \|\boldsymbol{\Phi}\boldsymbol{a_j}\|_2 = 1 \end{cases} \tag{7}$$

where $\mathbf{\Phi}$ and $\boldsymbol{A}$ are the basic and the sparse dictionaries, and $\mathbf{\Gamma}$ is approximation of sparse codes of signals over $\mathbf{\Psi} = \mathbf{\Phi}\boldsymbol{A}$. γ_i and $\boldsymbol{a}_j$ denote columns of $\mathbf{\Gamma}$ and $\boldsymbol{A}$ respectively. The basic dictionary is a fixed and dense implicit dictionary like DCT or Wavelets and the sparse dictionary is an explicit sparse dictionary learned by training data. Further details about the algorithm is available in [19]. Suppose that the basic dictionary $\mathbf{\Phi}$ and sparse dictionary $\mathbf{A}$ are given. Thus, if we define a dictionary $\mathbf{\Psi}$ as the product of these two dictionaries, representing a signal over $\mathbf{\Psi}$ is tantamount to representing it over the basic dictionary, $\mathbf{\Phi}$, at first point, and then representing the resulting sparse codes over the second dictionary, the sparse dictionary $\mathbf{A}$.

Now, assume that the sparse dictionary is trained using training samples, and we aim to classify matrix $\mathbf{X}$. Each column of this matrix represents an observation. We can obtain a sparse matrix $\tilde{\mathbf{X}}$ using sparse representation of $\mathbf{X}$ over the basic dictionary $\mathbf{\Phi}$. As sparse matrix $\tilde{\mathbf{X}}$ can be further represented sparser over dictionary $\boldsymbol{A}$, we will have:

$$\mathbf{X} \approx \mathbf{\Phi}\tilde{\boldsymbol{X}} \approx \mathbf{\Phi}\boldsymbol{A}\mathbf{\Gamma} \tag{8}$$

or for a test sample:

$$\mathbf{x} \approx \mathbf{\Phi}\tilde{\mathbf{x}} \approx \mathbf{\Phi}\boldsymbol{A}\boldsymbol{\gamma} \tag{9}$$

Therefore, the problem in (3), as illustrated in Figure 2 (b), will be changed to finding the sparsest solution of a set of sparse signals over $\mathbf{\Phi}$, over the sparse dictionary $\boldsymbol{A}$:

$$\min_{\gamma} \|\mathbf{x} - \mathbf{\Phi}\boldsymbol{A}\boldsymbol{\gamma}\|_2 \quad s.t. \quad \|\boldsymbol{\gamma}\|_1 \leq \epsilon \tag{10}$$

As an observation sample can itself be represented over $\mathbf{\Phi}$, the following problem has approximately an equal solution to the solution of (10) [22]:

$$\min_{\boldsymbol{\gamma}} \|\tilde{\mathbf{x}} - \boldsymbol{A}\boldsymbol{\gamma}\|_2 \quad s.t. \quad \|\boldsymbol{\gamma}\|_1 \leq \epsilon' \tag{11}$$

It has been shown that the resulting optimization problem can be reduced to a computationally cheaper problem as many of multiplications can be disregarded due to the sparsity of $\mathbf{A}$

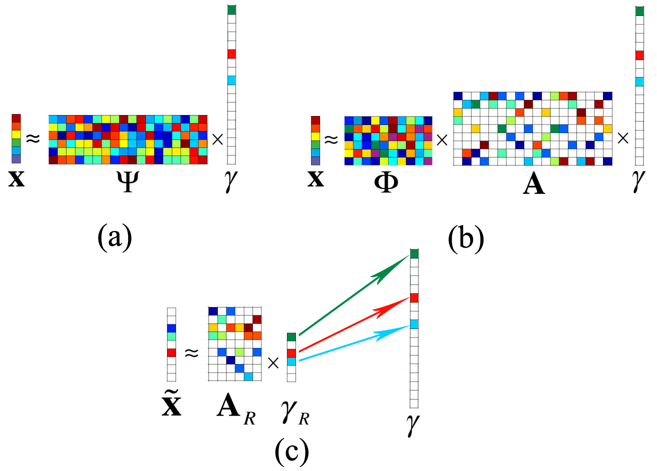

Fig. 2. Graphical illustration of sparse representation problems. (a) A regular sparse representation problem; $\mathbf{\Psi}$ can be a group of prototypes, or a learned dictionary. (b) A sparse representation using the double sparsity model for dictionaries; $\boldsymbol{A}$ is tantamount to a $\mathbf{\Psi}$ (c) Reduced sparse representation problem; $\boldsymbol{\gamma_R}$ can be easily mapped to $\boldsymbol{\gamma}$.

and $\tilde{\mathbf{X}}$ [19]. The approach is to locate non-zero entries in each column of $\tilde{\mathbf{X}}$. i.e. each sample data $\tilde{\mathbf{x}}$. Then, we should only look for those columns in $\mathbf{A}$ that has at least one non-zero element in the corresponding non-zero rows of vector $\tilde{\mathbf{x}}$. In other word, columns that have not the same supports as $\tilde{\mathbf{x}}$, have no role in representing $\tilde{\mathbf{x}}$, and therefore thwy should be simply ignored. Given relevant columns in $\mathbf{A}$, reduced matrices $\mathbf{A_R}$, $\tilde{\mathbf{x}}_R$ and $\gamma_\mathbf{R}$ are built correspondingly. As a result, optimization problem in (10) is reduced to a much lower dimensional problem as:

$$\min_{\gamma_R} \|\tilde{\mathbf{x}} - \boldsymbol{A}_R\gamma_R\|_2 \quad s.t. \quad \|\gamma_R\|_1 \leq \epsilon'' \qquad (12)$$

The current problem is graphically illustrated in Figure 2 (c). Solving the reduced optimization problem instead of (10) leads to significant decrease in computation and memory cost because its matrices are small and also sparse. As selected columns of $\boldsymbol{A}$ are known for each specific $\tilde{\mathbf{x}}$, mapping the solution of (10), $\gamma_\mathbf{R}$, to an equal γ is straightforward. After obtaining the sparse representation of an observed sample, which is in a strictly discriminative space, any classification method such as SVM can be used for classifying the sparse code.

C. The proposed scheme

According to the challenges and solutions that outloned, and described, a proper face recognition system should be able to generate a 3D face model given a 2D input image, and then use the generated 3D facial model to capture all possible conditions that in an observation can be occurred. Afterward, all the captured images are used to train a sparse dictionary using the double sparsity dictionary learning method [19] that is compatible to the next step, proposed double sparse representation technique. Assuming that a proper basic dictionary such as $\mathbf{\Phi}$ is given, the sparse code of the observed face image can be obtained by proposed double sparse representation technique. Having the sparse code, finding its class is strate forward using a classifier. Figure (3) illustrates whole process of the proposed face recognition system.

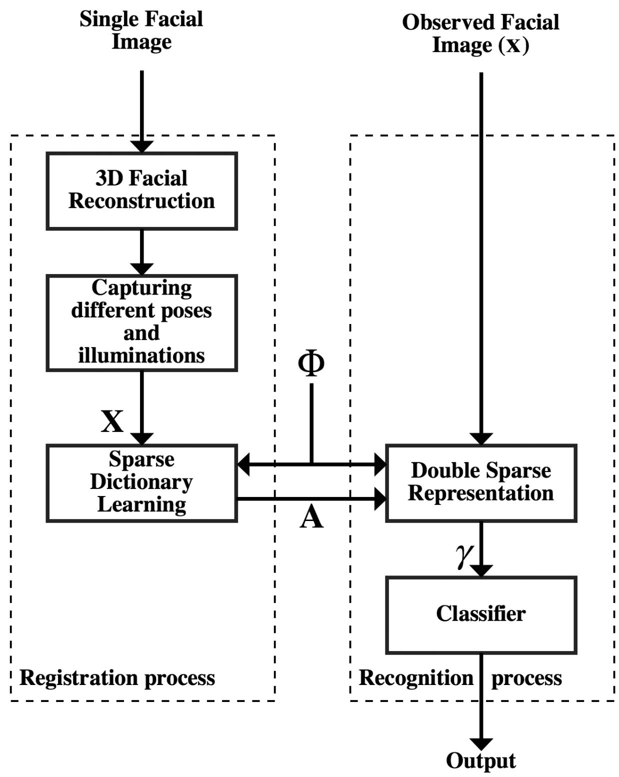

Fig. 3. The block diagram of the proposed face recognition system.

IV. Experiments

In this section, we evaluate the performance of the proposed system and compare it with seven well-known face recognition algorithms on the smartphones. The comparison is made in two categories for comparison of system robustness and running time. At the end, effect of using an artificial training set is also tested. All the experiments is carried out on a desktop computer with dual core $3GHz$ CPU and $2.0GB$ RAM.

A. System Robustness Comparison

In this section, we aim to evaluate the recognition accuracies with respect to illumination variations. Under the absence of face database captured from mobile devices, we make use of the Extended Yale B face database [21], which contains face images for 38 people captured under 64 different illumination conditions. Limited numbers of images taken under normal conditions are used for training the system and other images are used as test images.

We choose one of the images as the reference image and the other images will be divided into 5 subsets of images according to the light angle difference from the reference. The first subset covers the angular range from $0°$ to $5°$, the second one covers $10° - 20°$, the third subset covers $25° - 35°$, the fourth subset covers $50° - 70°$, and the remaining images are in the fifth subset. The first subset of images is selected

(a)

(b)

(c)

(d)

Fig. 4. Four subsets form YaleB face database. (a) Subset one that is used for training step of test. (b) Second subset that covers light angle range of $10° - 20°$. (c) Third subset of the database that covers light angle range of $25° - 35°$. (d) images of fourth subset that covers light angles from $50°$ to $70°$.

TABLE I
COMPARING ROBUSTNESS OF PROPOSED CLASSIFICATION METHOD, GENERAL SPARSE REPRESENTATION CLASSIFICATION METHOD, AND STATE OF THE ART FACE RECOGNITION METHODS FOR SMARTPHONES.

Database / Method	2^{nd}subset of YaleB database	3^{rd}subset of YaleB database	4^{th}subset of YaleB database
PCA-SVM	99.78%	74.47%	20.30%
LDA-SVM	94.3%	64.47%	15.79%
RP-NN	98.25%	52.37%	6.95%
ERP-NN	99.27%	54.47%	11.22%
ELM	13.6%	4.21%	3.76%
LRIC	100%	85.53%	36.84%
General SRC	100%	98.2%	65.3%
Proposed method	100%	93.5%	59.4%

for training and the fifth one are not used, whereas the other images are the test samples used for evaluation.

The result of comparison of our proposed method with those of conventional methods is presented in Table I. Table I exhibits that as long as the difference of light angle is small, all methods, excluding ELM, have almost high accuracies. As the light angle increases, performance of those conventional methods decline dramatically, whereas the proposed method is much less sensitive to illumination variation and maintain a good performance in fourth subset.

B. Running time comparison

Under computational limitation in mobile devices, the computational complexity of system should be as low as possible. In addition, the training time of method is highly related to computational complexity of system. Therefore, calculation of the running time for a test image is a good criterion for complexity evaluation. In this section, the running time of proposed method is compared with the conventional methods mentioned in [22]. Obviously, comparison should be arranged in the equal conditions with the same hardware and software. Similar to [22], comparison of running time for a test image is made using the CMU- PIE database [23], which contains a large number of samples. Large number of samples in PIE database can clearly challenge the complexity of system. Figure 5 exhibits some example of a person's face images from PIE face database.

Table II provides comprehensive comparison of recognition time, as a criterion for system complexity. This table shows that the regular sparse representation based method often provides high computational complexity in comparison to other methods. However, in the proposed method duo to using double sparse representation technique, the computational complexity is reduced in a way that is completely comparable to other existing face recognition algorithms for smartphones.

Fig. 5. Some face images of a person in PIE face database.

As Table II exhibits, the recognition time of the proposed system is almost independent of the size of training set. This is because in fact, we have assigned major of calculations to the training step, and increase in training set just affects training time, not the test time. As the sparse dictionary, $\boldsymbol{A}$, is already trained in the training step for the recognition step, we have to just calculate the sparse code of the observed image over the learned dictionary, $\boldsymbol{\Phi}$, and then, project it over the reduced sparse dictionary, $\boldsymbol{A}_R$. Therefore, owing to the assumption that the number of non-zero elements in each columns of $\boldsymbol{A}$, and also in the resulting sparse code, γ, is constant, increase in the number of training samples has no effect on the running time. Note that, though the number of training samples will affect the training time, but the training process can be done on a stronger remote server.

C. Registration

In order to examine effect of artificial training set on the result of face recognition system, following test is designed. A random face image from image collection of each person from CMU-PIE database is selected as an import image for registering person in the system. Afterward, on one hand, groups of artificial facial images are rendered and used for training step of the proposed approach. On the other hand, a basis matrix consist of selected images for registration is defined for using in a general sparse representation classification based face recognition system. Then, a groups of real facial images from CMU-PIE database is used for testing both face recognition systems. Table III shows how the artificial training set affects performance of face recognition system. As the test shows, performance of a face recognition system based on a single images themselves is completely poor, while in the case that a training set is rendered with their 3D reconstructed models, it can be used to gain a favorable system.

TABLE II

TEST TIME EVALUATION. ALL TIMES ARE IN SECONDS AND RELATED TO CLASSIFYING 100 TEST SAMPLES.

Methods Number of training samples	General SRC	Proposed method	LRIC	ELM	ILDA	CCI-PCA	LDA-SVM	PCA-SVM
680	0.84	2.12	0.023	0.009	0.23	0.45	0.035	0.062
1360	6.04	2.23	0.021	0.009	0.41	0.92	0.053	0.089
2040	19.01	2.41	0.019	0.009	0.55	1.49	0.071	0.109
2720	40.9	2.45	0.018	0.01	0.78	1.25	0.129	0.132
3400	71.49	1.96	0.017	0.009	0.93	1.96	0.088	0.212
4080	130.65	2.36	0.018	0.009	1.05	2.7	0.121	0.188
4760	248.17	2.63	0.017	0.009	1.14	2.44	0.288	0.229
5440	425.34	2.12	0.03	0.01	1.21	2.57	0.182	0.241
6120	663.2	2.28	0.025	0.01	1.54	3.19	0.152	0.257
6800	962	2.36	0.019	0.009	1.73	3.80	0.177	0.288
7480	1619	2.34	0.026	0.01	1.89	3.98	0.178	0.249
8160	3361	2.47	0.025	0.012	2.08	4.45	0.221	0.447

TABLE III

EFFECT OF ARTIFICIAL TRAINING SET ON ACCURACY RATE OF THE FACE RECOGNITION SYSTEM. RECONSTRUCTED TRAINING SET STRONGLY IMPROVES THE ACCURACY OF THE SYSTEM.

Method	Accuracy rate
General SRC(single training image based)	17.34%
Proposed method	92.14%

V. CONCLUSION

Due to the popularity of smartphones and mobile devices, implementing face recognition systems on them has gained much attention. The proposed system is designed to have good performance under practical conditions and existing limitations. In particular, practical conditions, robustness of system to processing challenges and computational complexity are considered in designing the proposed system. The experiments on proposed method confirm the remarkable performance of system in real situations. The first experiment investigates the accuracy of system in which its results validate high robustness of system. Moreover, we conduct experiment on the running time of proposed method, which indicates feasibility of the system for implementing on smartphones. Thus, the suggested method provides high accuracy and robustness, while its computational complexity maintains comparable to conventional methods.

REFERENCES

[1] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no.1, vol. 3, 1991.

[2] P. Belhumeur, J. Hespanda, and D. Kriegman, "Eigenfaces versus fisherfaces: Recognition using class specific linear projection," *IEEE Transaction on Pattern Analysis Machine Intelligence*, vol. 19, no. 17, pp. 711-720, 1997.

[3] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter, "A 3D face model for pose and illumination invariant face recognition," *Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 296-01, 2009.

[4] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 2009.

[5] W. Dong, L. Zhang; G. Shi; X. Li, "Nonlocally Centralized Sparse Representation for Image Restoration," *IEEE Transactions on Image Processing*, , vol.22, no.4, pp.1620-1630, 2013.

[6] M. Joneidi, M. Sadeghi, M. Sahraee-Ardakan, M. Babaie-Zadeh, C. Jutten, "A study on clustering-based image denoising: From global clustering to local grouping," *2014 Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)* , vol.1, no.5, pp.1657-1661, 2014

[7] S. Minaee, A. Abdolrashidi and Y. Wang, "Screen Content Image Segmentation Using Sparse-Smooth Decomposition", 49th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove CA, Nov 2015.

[8] S. Minaee and Y. Wang "Screen Content Image Segmentation Using Least Absolute Deviation Fitting" Internation Conference on Image Processing 2015.

[9] Q. Wang,F. Chen, W. Xu, M. Yang, "Object Tracking With Joint Optimization of Representation and Classification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.25, no.4, pp.638-650, 2015.

[10] S. Manaffam and A. Seyedi, "Synchronization probability in large complex networks," *IEEE Transactions on Circuits and Systems II*, vol. 60, no. 10, pp. 697–701, Oct. 2013.

[11] A. Zarezade, H. R. Rabiee, A. Soltani-Farani, A. Khajenezhad, "Patch-wise Joint Sparse Tracking With Occlusion Detection," *IEEE Transactions on Image Processing,* , vol.23, no.10, pp.4496-4510, 2014.

[12] Y. Shen, W. Hu, M. Yang, B. Wei, S. Lucey.; C. T. Chou, "Face recognition on smartphones via optimised Sparse Representation Classification," *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, IPSN-14* , vol.15, no. 17, pp. 237-248, 2014.

[13] J. Golmohammady, M. Joneidi, M. Sadeghi, M. Babaie-Zadeh, C. Jutten, "K-LDA: An algorithm for learning jointly overcomplete and discriminative dictionaries," *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)* , vol.1, no.5, pp.775-779, 2014.

[14] I. Ramirez, P. Sprechmann, G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* , vol.13, no.18, pp.3501-3508, 2010.

[15] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. "Shape-from-shading: a survey." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.21, no. 8 , pp. 690-706, 1999.

[16] I. Kemelmacher-Shlizerman, and R. Basri. "3d face reconstruction from a single image using a single reference face shape." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.33, no.2, pp. 394-405, 2011.

[17] Ptucha, R. and Savakis, A.E. "LGE-KSVD: Robust Sparse Representation Classification," *Image Processing, IEEE Transactions on*, vol.23, no.4, pp. 1737-1750, 2014.

[18] M. Rahmani and G. Atia, "Randomized Robust Subspace Recovery for Big Data," *IEEE International Workshop on Machine Learning for Signal Processing*, Boston MA, USA, Sept 2015.

[19] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double sparsity: Learning sparse dictionaries for sparse signal approximation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1553-1564, 2010.

[20] J. B. Huang and M.-H. Yang, "Fast sparse representation with prototypes," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3618-3625, 2010.

[21] A. Georghiades, D. Kriegman, and P. Belhumeur, "From few to many: generative models for recognition under variable pose and illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no.6, pp. 643-660, 2001.

[22] K. Choi, K. A. Toh, and H. Byun, "Realtime training on mobile devices for face recognition applications," *Pattern Recognition*, vol. 44, no. 2, pp. 386-400, 2011.

[23] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.25, no.12, pp. 1615-1618, 2003.