B Savinya* et al.
(IJITR) INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND RESEARCH
Volume No.7, Issue No.6, October-November 2019, 9367-9375.

High Throughput Robust Face Recognition using SVD

B.SAVINYA
M.Tech Scholar, Department of Electronics and Communications Engineering, Chaitanya Institute of Science & Technology, India.

S.SRIVIDYA
Assistant Professor, Electronics & Communications Engineering, Chaitanya Institute of Science & Technology, India.

Abstract: **In practice, there is no guarantee that the collected data would cover all different occlusions for all identities of interest. Here proposed an iterative method to address the face identification problem with block occlusions of two characteristics in order to model contiguous errors (e.g., block occlusion) effectively. The first describes a tailored loss function. The second describes the error image as having a specific low-rank image comparison structure. In this paper shown that joint characterization is effective for describing errors with spatial continuity. Our approach is computationally efficient due to the utilization of the alternating direction method of multipliers. Using of the fast iterative algorithm leads to the robust representation method, which is normally used to handle non-contiguous errors. Extensive results on representative face databases document the effectiveness of our method over existing robust representation methods with respect to both identification rates and computational time.**

Key Terms— **Face Identification; Robust Representation; Low-Rank Image; Iterative Reweighted Coding;**

I. Introduction

A facial recognition system is a technology capable of identifying or verifying a person from a digital image or a video frame from a video source. There are multiple methods in which facial recognition systems work, but in general, they work by comparing selected facial features from given image with faces within a database. It is also described as a Biometric Artificial Intelligence based application that can uniquely identify a person by analyzing patterns based on the person's facial textures and shape. This project was labeled man-machine because the human extracted the coordinates of a set of features from the photographs, which were then used by the computer for recognition. Using a graphics tablet, the operator would extract the coordinates of features such as the center of pupils, the inside corner of eyes, the outside corner of eyes, point of windows peak, and so on. From these coordinates, a list of 20 distances, such as the width of mouth and width of eyes, pupil to pupil, were computed. These operators could process about 40 pictures an hour. When building the database, the name of the person in the photograph was associated with the list of computed distances and stored in the computer.

In the recognition phase, the set of distances was compared with the corresponding distance for each photograph, yielding a distance between the photograph and the database record. The closest records are returned.Face Identification (FI) focuses on deducing a subject's identity through a provided test image and is one of the most popular problems in computer vision.

Typically, test images exhibit large variations, such as occlusions. Ideally, if the training set contains the same type of occlusion as the test image then identification becomes a rather straight forward task. Early works on face identification attempted to deal with illumination variations. The concept of 41-graph, which is robust to data noises and naturally sparse, was introduced in to encode the overall behavior of the training set in sparse representations. To handle more complex variations such as face disguise and expressions, sparse representation based classification models were proposed. The main idea in these approaches is that a subject's face sample can be represented as a linear combination of available images of the same subject. Then, the face class that yields the minimum reconstruction error is selected in order to classify the subject. One recent extension of the sparse representation based classification model is the class-wise sparse representation. In this method, the number of training classes is minimized to alleviate the problem of representing the query by samples from many different subjects. Another extension is the patch based classification approach. The patch based approach employs the sparse representation-based classifier to each patch of the face separately and the final decision is made by fusing the patch classification results.

Low-rank estimation has been considered in where a discriminative low-rank metric learning method was proposed that jointly learns a low-rank linear transformation matrix and a low-rank representation. A graph construction model, with robust similarity metric (low-rank representation, which is robust to noisy data) and balanced property for the application of semi-supervised learning.

In this work, we propose an iterative method to solve the FI problem with occlusions. We consider the same

scenario as in according to which we are given "clean" frontal aligned views with a block occlusion which appear in any position on the test image but is "unseen" to the training data. When corrupted training data are provided that are not frontally aligned (e.g., scenarios in an unconstrained environment) a tool such as RPCA and SLR is employed to separate outlier pixels and corruptions from the training samples as a pre-processing step. Then, the "clean" frontal aligned faces are used for training data to perform face identification with occluded test images.

II. Related Work

Rasied et al. displayed a created Face Recognition framework. The technique utilized particular esteemed disintegration as pictures highlight extractor and back spread neural systems as its classifier. The back spread preparing parameters are shifted so as to locate the best parameter with the best precision. The outcomes from tests have indicated that blends of the two strategies give great acknowledgment rate and along these lines considered as a compelling Face Recognition framework. Varieties in posture, demeanor and enlightenment circumstances make Face Recognition a significantly all the more testing and troublesome errand. Wang et al. introduced a Face Recognition approach by utilizing picture improvement and Gabor wavelets change. Logarithm change and standardization are performed in face pictures caught under different lighting conditions for Face Recognition. This incorporates convolving a face picture with a progression of Gabor wavelets at different areas, scales and areas and separating highlights from Gabor sifted pictures. Essential improvements are likewise seen when the preprocessing and Gabor separated pictures are utilized for include extraction rather than the first pictures. The methodology accomplishes 94.4% acknowledgment exactness utilizing just 160 highlights of a face picture. Results show that this technique propels Face Recognition execution utilizing this plan when preparing and testing on pictures caught under factor enlightenment and appearance.

. The freedom property of these Gabor highlights encourages the use of the PRM technique for highlights arrangement. Gabor changed face pictures display solid attributes of direction selectivity and spatial region. These pictures can, along these lines, produce notable neighborhood includes that are generally reasonable for Face Recognition. While, ICA further lessens excess and speaks to free highlights expressly. These free highlights are gotten valuable for other consequent example separation and affiliated review. Tests done on Face Recognition utilizing the FERET for example Face Recognition Technology and the ORL database sets, where the pictures differ in demeanor, scale, enlightenment, and posture, show the possibility of the IGF strategy. Specifically, the IGF technique increase 98.5% right Face Recognition exactness when utilizing 180 highlights for the FERET dataset, and 100% precision for the ORL dataset utilizing 88 highlights. Rida et al. exhibited a paper shows how a Face Recognition framework can be structured with counterfeit neural system. Note that the preparation procedure didn't comprise of a solitary call to a preparation work. Rather, the system was prepared a few times on different info perfect and uproarious pictures, the pictures that substance face. For this situation preparing a system on various arrangements of loud pictures constrained the system to figure out how to manage clamor, a typical issue in reality [2].

II. Recognition Metrics

Metrics driven in two characteristics and uses a tailored loss function based on M-estimators. The methodhandlescontiguouserrorsthatareconsideredlow-rank in comparison to the size of the image and is efficient in terms of computational cost. A special case of our method is also utilized to solve efficiently the robust representation problem for non-contiguouserrors.

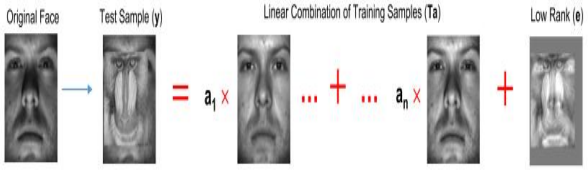


Fig. 2 Degradation Model

Thetestsamplewithocclusioncanberepresentedasthelinear combinationoftrainingsampleswithsomeintra-classvariations(e.g., lighting) plus the error **e**. The error **e** has two characteristics; it is considered low-rank in comparison of image size and follows a distribution described by a tailored potential lossfunction.

Let**y**∈R^d denotethefacetestsampleinacolumn-wise vectorized form where $j \times k = d$ is the size of theimage.

Let T=**[**T_i, … , T_c *]* € $R^{d \times n}$denote the matrix (dictionary) with the set of samples of c subjects stacked in columns.

T_i € $R^{d \times n_i}$ denotes the $n_i$set of samples of the ith subject, such that, $\sum_i ni$=n.

In Figure 2 we can represent the test sample with

occlusion as the superposition of training samples and a representation error **e**, thus, the degradation model is,

$$\mathbf{y} = \mathbf{Ta} + \mathbf{e}, \qquad (1)$$

Optimization

Let us now describe the iterative algorithm tofind efficiently a^{t+1} in problem (11b).In order to solve the proposed problem, first we let y − Ta = e and since we are interested in estimatingnonnegative coefficients for the representation vector we also introduce an additional variable **z** suchthat**a=z**. Then, the coding step (11b) is reformulatedas,

$$\underset{\mathbf{a},\mathbf{z},\mathbf{e}}{\text{minimize}} \ \|\sqrt{\mathbf{W}^{t+1}}\mathbf{e}\|_2^2 + \lambda_*\|T_M(\mathbf{e})\|_* + \vartheta(\mathbf{z})$$
$$\text{subject to } \mathbf{y} - \mathbf{Ta} = \mathbf{e}, \mathbf{a} = \mathbf{z}.$$

Problem (1) is solved efficiently with ADMM which is known for fast convergence to an approximate solution [2]. As in the method of multipliers, the problem takes the form of the augmentedLagrangian.

$$\begin{aligned} &L(\mathbf{e},\mathbf{a},\mathbf{z},\mathbf{u}_1,\mathbf{u}_2,\mathbf{w}^{t+1}) \\ &\quad = \|\sqrt{\mathbf{W}^{t+1}}\mathbf{e}\|_2^2 + \lambda_*\|T_M(\mathbf{e})\|_* \\ &\qquad + \vartheta(\mathbf{z}) + \mathbf{u}_1^T(\mathbf{y} - \mathbf{Ta} - \mathbf{e}) + \frac{\rho_1}{2}\|\mathbf{y} - \mathbf{Ta} - \mathbf{e}\|_2^2 \\ &\qquad + \mathbf{u}_2^T(\mathbf{a} - \mathbf{z}) + \frac{\rho_2}{2}\|\mathbf{a} - \mathbf{z}\|_2^2, \end{aligned}$$

whereρ_1>0 and ρ_2>0 are the penalty parameters, and
$\mathbf{u}_1$and$\mathbf{u}_2$arethedualvariables.TheADMMupdatesca nbe expressedas,

$$\begin{aligned} \mathbf{e}_{s+1} &= \underset{\mathbf{e}}{\text{argmin}}\, L(\mathbf{e},\mathbf{a}_s,\mathbf{u}_{1,s},\mathbf{w}^{t+1}), && (15a)\\ \mathbf{z}_{s+1} &= \underset{\mathbf{z}}{\text{argmin}}\, L(\mathbf{z},\mathbf{a}_s,\mathbf{u}_{2,s}), && (15b)\\ \mathbf{a}_{s+1} &= \underset{\mathbf{a}}{\text{argmin}}\, L(\mathbf{e}_{s+1},\mathbf{a},\mathbf{u}_{1,s},\mathbf{z}_{s+1},\mathbf{u}_{2,s}), && (15c)\\ \mathbf{u}_{1,s+1} &= \mathbf{u}_{1,s} + \rho_1(\mathbf{y} - \mathbf{Ta}_{s+1} - \mathbf{e}_{s+1}), && (15d)\\ \mathbf{u}_{2,s+1} &= \mathbf{u}_{2,s} + \rho_2(\mathbf{a}_{s+1} - \mathbf{z}_{s+1}), && (15e) \end{aligned}$$

where sdenotes the ADMM iteration and finally weset

$$\mathbf{a}^{t+1} = \lim_{s\to+\infty}\mathbf{a}_{s+1}$$

Algorithm 1 Fast & Low-Rank IRNNLS Algorithm

Inputs: $\mathbf{y}$, $\mathbf{T}$, λ_*, ρ_1, ρ_2, ϵ_1, ϵ_2 and ϵ_3.
Initialize $\mathbf{a}^1 = 1/n$, $\mathbf{u}_{1,1} = \mathbf{0}$, $\mathbf{u}_{2,1} = \mathbf{0}$, $t = 0$ and $s = 0$
Repeat
1) $t = t + 1$
2) Estimate the weights,

$$w_i^t = \phi'((\mathbf{y} - \mathbf{Ta}^t)_i)/(\mathbf{y} - \mathbf{Ta}^t)_i, \quad i = 1,\ldots,d$$

Repeat
3) $s = s + 1$
4) Find $\mathbf{e}_s$ using (22) (contiguous errors) or (32) (non-contiguous errors)
5) Find $\mathbf{z}_s$ using (24)
6) Update $\mathbf{a}_s$ using (26)
7) Update $\mathbf{u}_{1,s}$ and $\mathbf{u}_{2,s}$ using (15d), (15e)
Until *converge*
8) Set $\mathbf{a}^t = \mathbf{a}_s$, $\mathbf{u}_{1,1} = \mathbf{0}$, $\mathbf{u}_{2,1} = \mathbf{0}$, $s = 0$
Until *converge*
Output: The final estimates of $\mathbf{a}$ and $\mathbf{w}$.

III. Implementation

In this section we present experiments on four publicly available databases, AR, Extended Yale B, Multi-PIE and Labeled Facesin the Wild (LFW)

(a) (b) (c)

Fig. 1. The three artificial images used for the block occlusion experiments.

(a) Baboon.(b)Non-squareimage.(c)Dog.

to show the efficacy of the proposed method. We demonstrate identificationandreconstructionresultsundervariousartificial and real-world variations. We compare our framework with ten other FI algorithms, SRC [6], CR-RLS [9], LR^3 [2], L12,[6]and the robust algorithms SDR-SLR, HQ (additive form) [5], CESR [6], RRC_L1 [4], RRC_L2[4],SSEC [9].

We consider the following five FI cases:

1) Cases with contiguous variations such as random block occlusion with different sizes and objects, face disguise and mixture noise which is a combination of block occlusion and pixelcorruption,

2) Cases with non-contiguous variations such as illumination variations, pixel corruption, faceexpressions.

3) Cases with random block occlusion and few training samples.

4) Cases in unconstrained environment and with corrupted testing and trainingsamples.

For all methods, we used the solvers provided by the authors of the corresponding papers. We chose to solve the ℓ_1 minimization problem in SRC and RRC_L1 with the Homotopy algorithm since it resulted in the highest accuracy in the performance comparison in with reasonable time execution. In our algorithms, we set $\rho_1 = 1$ and $\rho_2 = 0.1$. The convergence parameters were set equal to $\epsilon_1 = 10^{-2}$, $\epsilon_2 = 10^{-1}$, $\epsilon_3 = 10^{-2}$. For fair comparisons with respect to execution time and identification rates we set the same ϵ_3 for the RRC algorithm and the same maximum number of iterations (t=100). All face images were normalized to have unit ℓ_2-norm and all variables initialized to zero except for $\mathbf{a}^1 = 1/n$ as in[14].

A. Identification Under Block Occlusions

Experiments with occluded images were conducted on three datasets: Extended Yale B, AR and Multi-PIE.

B. Identification under Expressions & Face Disguise

In this experiment we tested our algorithms with face expressions and occlusion with real-world objects in three different scenarios: 1) faces with expressions 2) faces with sunglasses and 3) faces with scarves. The training set consists of the two neutral images (one from each session) from the AR dataset per subject. For the first scenario (face expressions) the 6 images per subject from sessions 1 and 2 with face expressions (smile, anger and scream) were selected for the testing set.

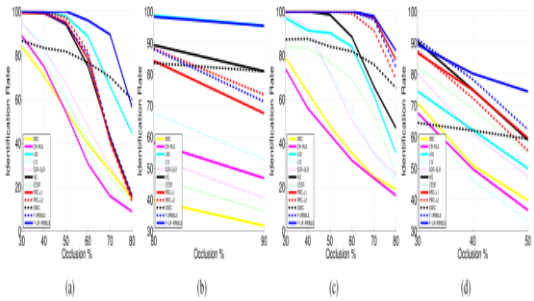
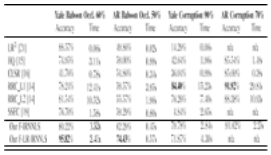


Fig. 2. Identification Rates on the Extended Yale B and AR under Block Occlusion with Baboon, Vase and Dog. (a) Yale B dataset with baboon image. Yale B dataset with vase image. (c) Yale B dataset with dog image. (d) AR dataset with baboon image

TABLE I

Average Run Time per Test Sample on the Extended Yale B and Ar Datasets under Different Variations

For the second scenario (faces with sunglasses) the



testing set consisted of the 6 images per subject with sunglasses

From sessions 1 and 2. In the third scenario (faces with scarves) the 6 images per subject with scarves from sessions 1 and 2 were chosen for the testing set. The images were resized to 60 x 43 pixels.

TABLE II

Identification Rates (%) Under Face Disguise On theAr Database

Case	Expressions	Sunglasses	Scarves
SRC [6]	82.33%	37.17%	35.17%
CR-RLS [9]	81.33%	33.83%	39.33%
LR^3 [21]	79.50%	80.50%	77.00%
L12	85.17%	40.33%	43.00%
SDR-SLR [27]	92.87%	36.67%	32.83%
HQ [15]	94.33%	66.67%	45.67%
CESR [16]	93.50%	66.50%	17.17%
RRC_L1 [14]	94.33%	83.00%	58.50%
RRC_L2 [14]	93.67%	84.17%	72.50%
SSEC [19]	56.17%	70.67%	75.00%
Our F-IRNNLS	**94.83%**	81.50%	53.67%
Our F-LR-IRNNLS	93.00%	**89.83%**	**78.83%**

Identification rates for the three scenarios are shown in Table II for the various methods.

For the face expressions experiment all robust algorithms achieved high performance. A key observation for this experiment is that modeling the error as low-rank does not improve the results since face expression errors do not in general form a contiguous area.

For the sunglasses experiment we observed that our F-LR-IRNNLS algorithm outperformed previous methods and was able to detect the outliers effectively. In this case modeling the error to be low-rank was adequate. This is due to the fact that the residual image consisted mainly of the sunglasses that made a contiguous error. SSEC performed poorly, perhaps, because the method does not capture well contiguous areas that are not square. A similar conclusion was drawn above with the random block occlusion experiment with a non-square image.

Results with the scarves experiment demonstrate that all methods robust to contiguous errors performed well, as expected (since the scarf occlusion is contiguous). Our F-LR-IRNNLS method achieved the best performance with 78.83% identification rate while our non-contiguous F-IRNNLS method achieved only 53.67%. This result emphasizes the fact that exploiting the spatial correlation in contigu- ous variations, such as scarves, is beneficial. Time performance is not reported here since all methods run very fast (less than a second) due to the fact that the training dictionary in this experiment was relatively small (200 training samples).

C. Identification Under Mixture Noise

In this experiment we evaluate the performance of our algorithm for the case of mixture noise. In this case, both pixel corruption and block occlusion degraded the testing images. An example image with this degradation is shown in Figure 12. This experiment was conducted with two datasets, Extended Yale B and AR. Similarly to the previous Extended Yale B settings, Subsets 1 and 2 of Extended Yale B were used for training and Subset 3 was used for testing. With the AR dataset we chose the 700 non-occluded AR images for training from session 1 and the 700 non-occluded images for testing from session 2. In both datasets, for each testing image a percentage of randomly chosen pixels was corrupted. Corruption was performed by replacing those pixel values with independent and identically distributed samples from a uniform distribution between [0, 255]. Then, we placed the baboon square image on each corrupted test image. In Yale B dataset we performed this experiment with 30% pixel corruption and 60% occlusion. With the AR dataset, experiments were conducted with 20% pixel corruption and 70% occlusion. Identification rates are shown in Table III for the various methods.

F-LR-IRNNLS outperformed all previous methods which indicates that in the mixture noise case, our two error constraints capture the error term effectively. SSEC performed poorly due to the presence of pixel corruption. RRC_L1, RRC_L2 and HQ were robust to pixel corruption, however, their performance remained low since they were not effective on describing the occlusion part. Our F-LR-IRNNLS had a good balance on detecting the corrupted pixels and capturing the occlusion part with the employment of the weighted and nuclear norms. However, although F-LR-IRNNLS achieved higher performance than the previous methods, the actual accuracy was relative low with 63.08% in YaleB and 57.29% in AR. The result may indicate that in mixture of noises further investigation about modeling the error is required.

TABLE III

Identification Rates (%) & Time Performance (S) Under Mixture Noise: Yale B 30% Corruption & 60% Occlusion, Ar 20% Corruption & 50% Occlusion

Dataset	Yale B		AR
	Accuracy	Time	Accuracy
SRC [6]	26.81%	1.04s	27.86%
CR-RLS [9]	14.73%	0.02s	28.29%
LR^3 [21]	44.62%	0.06s	35.43%
SDR-SLR [27]	32.75%	0.50s	40.00%
HQ [15]	42.20%	3.06s	42.00%
CESR [16]	23.96%	0.86s	40.86%
RRC_L1 [14]	43.08%	14.58s	47.00%
RRC_L2 [14]	41.54%	9.65s	38.14%
SSEC [19]	14.95%	1.72s	11.86%
Our F-IRNNLS	45.27%	1.45s	49.00%
Our F-LR-IRNNLS	**63.08%**	6.33s	**57.29%**

C. Identification UnderIllumination

Experiments with variations in illumination were conducted on the Multi-PIE dataset. As in the block occlusion experiments, we used all 249 subjects in Session 1.

TABLE IV

Identification Rates onthe Multi-Pie under Illumination

Sessions	Session 2 Accuracy	Session 3 Accuracy	Session 4 Accuracy	Session 4 Time
SRC [6]	95.48%	92.13%	95.71%	0.63s
CR-RLS [9]	95.30%	90.56%	94.46%	0.02s
LR^3 [21]	90.00%	83.44%	87.89%	0.04s
L12	93.92%	88.69%	92.91%	0.04s
HQ [15]	95.84%	94.63%	97.14%	2.21s
CESR [16]	94.64%	92.94%	96.06%	1.94s
RRC_L1 [14]	95.84%	95.13%	97.14%	6.88s
RRC_L2 [14]	97.11%	94.19%	97.37%	31.11s
SSEC [19]	86.75%	76.13%	83.09%	5.40s
Our F-IRNNLS	**97.29%**	**96.00%**	**98.00%**	1.05s
Our F-LR-IRNNLS	96.87%	94.13%	97.83%	1.83s

As in [9], we used 14 frontal images with 14 illuminations and neutral expression from Session 1 for training, and 10 frontal images from Session 4 for testing. Identification rates are shown in Table IV for the various methods.

Our first observation is that all methods achieved high identification rates. Simple SRC approaches performed well while robust methods only slightly improved the results. The reason with respect to our method is that for illumination variations modeling the error image as low-rank does not hold in this case. Similar observations can also be deduced from results of the LR^3 method in this case.

With respect to time performance, our algorithm output-forms the previous robust methods. In particular the execution time in our approaches is around 1 second per test image while for RRC_L2 is around 30 seconds. Notice that although this was an experiment with a large training dictionary, our method retains very low running time.

E. Identification Under Pixel Corruptions

Experiments under pixel corruption were conducted on two datasets: Extended Yale B and AR.

As in [6], [4] we used the non-occluded faces of Subsets 1 and 2 of the Extended Yale B for training and Subset 3 for testing. Images were resized to 96 x 84 pixels. In the AR dataset, to make the experiment more challenging

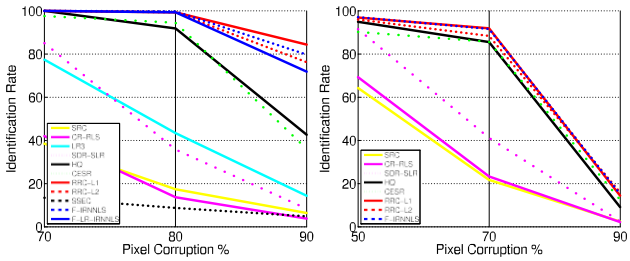


Fig. 3. Identification Rates on the Extended Yale B and AR datasets under Pixel Corruptions. (a) Yale B dataset. (b) AR dataset.

F. Identification Under Unconstrained Environment and Occluded Testing andTraining Samples

1) Identification Under Unconstrained Environment: Thus far, we have assumed that training samples are "clean" frontal aligned views and without large variations of the same identity. In an unconstrained environment this assumption does not hold and often face images of the same identity exhibit large variations in pose, illumination, expression and occlusion.

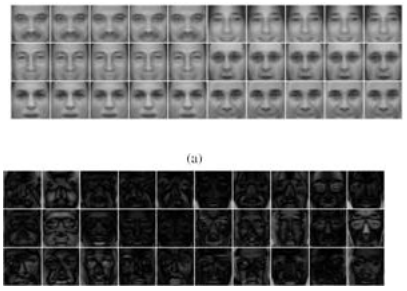


Fig. 4. Sample images from the LFW-a dataset and the SDR-SLR decomposition applied to the dataset. (a) Training sample images from the class-specific dictionary **A**. (b) Training sample images from the variation dictionary **B**.

Furthermore, testing images may not contain the same variations and occlusions as the training images.

To examine the robustness of our method on an unconstrained environment we evaluate its performance in the LFW database. The dataset contains images of 5,749 different subjects and in this work we used the LFW-a [47], which is an aligned version of LFW based on commercial face alignment software. We used the subjects that include no less than ten samples and we constructed a dataset with 158 subjects from LFW-a. For each subject, we randomly chose 5 samples for training resulting in a dictionary of 790 faces) and 5 samples for testing. The images were resized to 90x 90.

To deal with such environment we utilize the SDR-SLR algorithm as a pre-processing step as explained in Section II-F. Sample images from dictionaries **A** and **B** estimated on LFW-a dataset are illustrated in Figure 8(a) and (b). More specifically, in Figure 8(b) images cover variations which are not class- specific and are used to represent complex variations of the query. These variations may be represented by the component **B** using images that do not belong to the same identity of the test image. Any remaining variations that cannon be described by **B** are captured by the term **e**.

2) Identification Under Occluded Testing and Training Samples: In order to further investigate the scenario where we are given corrupted testing and training data in a constrained environment this time,

TABLE V

Identification Rates (%) In The Unconstrained Dataset Lfw-A In The Left Column. In The Right Column We Demonstrate Identification Rates (%) In Multi-Pie Dataset Under Occluded Training Samples

	LFW-a	Multi-PIE
SRC [6]	68.35%	14.86%
CR-RLS [9]	64.68%	6.91%
LR^3 [21]	62.53%	6.00%
SDR-SLR [27]	76.20%	39.03%
HQ [15]	71.00%	17.83%
CESR [16]	59.87%	11.43%
RRC_L1 [14]	63.42%	23.20%
RRC_L2 [14]	72.53%	37.66%
SSEC [19]	32.41%	44.14%
Our F-IRNNLS	72.91%	23.89%
Our F-LR-IRNNLS	74.81%	44.06%
Our F-LR-IRNNLS (SLR)	**77.47%**	**47.71%**

We also conducted an experiment on the Multi-PIE dataset in the scenario that corrupted testing and training data are provided in a constrained environment. To simulate the cor- rupted (occluded) training data we considered the same train- ing and test sets as in Multi-PIE block-occlusion experiment described above. In particular, we used the 6 frontal images with 6 illuminations and neutral expression from Session 1 for training. For each of the 249 subjects we randomly chose half of the training images to be occluded. In each image we replace a random block with the square baboon image with occlusion chosen randomly from 30% to 60%. We chose the 10 frontal images from Session 4 for testing. In each test image, we replace a random block with the square baboon image and the occlusion was randomly chosen from 30% to 60%. Identification rates are shown in Table V(right column). From the results we observe that as expected SDR-SLR perform way better than the SRC and CR-RLS. However, not all block-occlusions are sufficiently covered by **B**. This is due to the fact that occlusion appears in random places and sizes in the query as well as in training data. Therefore, it might be very unlikely that the occlusion on the query and training images will be of the same type. Thus, when SDR-SLR method is combined with the F-LR-IRNNLS algorithm the best accuracy is reported since the proposed modeling of the error term is robust to handle occlusions of the query. Finally, we observe that our F-LR-IRNNLS algorithm outperform the other approaches even when SDR-SLR is not utilized. The reason may be that our approach chooses the non-occluded training samples to represent the query since occlusion is effectively captured by the representation error image **e**.

3) Comparison With Inductive Methods: In this section we compare our method with two inductive methods, namely IRPCA [29] and IDNMD [28]. Both methods are able to handle new data meaning that given a new sample an under- lying learnt projection matrix **K** can be used to efficiently remove corruptions and occlusions. A test face is recovered from occlusions by computing **Ky**. Then, the "clean" test face **Ky** is provided as an input to a classifier to identify the subject. In this work we use the SRC classifier [6] for these methods in order to make direct comparisons with our method.[15] Also, in this case we do not perform any pre-processing step such as SDR-SLR to clean the training data for identification in our and other methods to make fair comparisons. There are two main differences between IRPCA, IDNMD and our method; i) Our method does not require the same class of occlusions to be present in the training and test data while inductive methods do to perform well. ii) Our method describes the error image by using two metrics, namely weighting and nuclear norms.

G. Identification Under Few Training Samples

To examine the robustness of our method under few training samples per subject we conducted experiments on the Multi- PIE database [4] with uncorrupted training data. As in the experiments under occlusion, we used 6 frontal images with 6 illuminations and neutral expression from Session 1 for training, and 10 frontal images[16] from Session 4 for testing. Then, we

randomly selected 2 or 4 samples per subject to perform experiments under fewer training examples. In each test image, we replace a random block with the square baboon image, where each block randomly covered between 30% and 60% of the image area.

H. Weight Map Estimations

Figure 5 shows the estimated weight maps between RRC_L1, RRC_L2 and our F-LR-IRNNLS in experiments with occlusions. Black values (close to zero) represent detected outliers by the various methods. We observe that

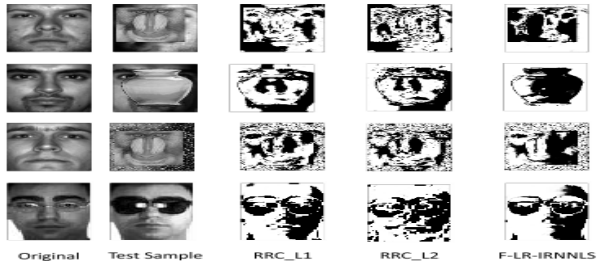


Fig. 5. Estimated weight maps for three iterative reweighted coding methods. The weight maps of our F-LR-IRNNLS method capture the outlier object of interest more accurately. With the other methods, a number of inlier pixels are detected as outliers.

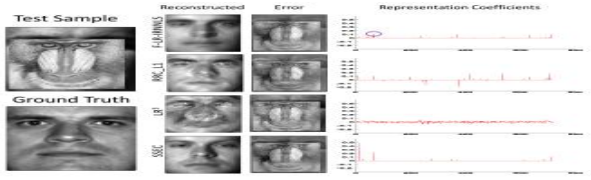


Fig. 6. Reconstruction results in 70% block occlusion.

F-LR-IRNNLS detected the outlier objects more effectively than the other methods. Most of the black regions in the weight maps are concentrated on the occluded area. In particular, we observe in Figure 10 (first row) that for the F-LR-IRNNLS method small weights are only assigned to the occluded (baboon) region as desired. On the other hand, the weight maps of RRC_L1 and RRC_L2 are not as accurate since outliers were detected in important pixels of the face. The reason is that with these methods there is no spatial correlation constraint between the weights. Similar conclusions can be drawn from all other examples in Figure 6.

I. Face Reconstruction Results

Figure 7 illustrates the face reconstruction results and the associated representation coefficients by the four methods. F-LR-IRNNLS and SSEC had the best reconstruction performance. The reconstructed face by LR^3 was poor mainly due to the choice of the regularizer for the representation coefficients (4_2 norm). Similar reconstruction performance for the LR^3 method was encountered in almost all of our conducted experiments.

More reconstruction results for various methods are presented in Figure 7. With mixture noise, our F-LR-IRNNLS achieved the best performance which demonstrates that our modeling was more effective in this case than the other methods.

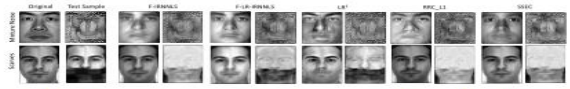
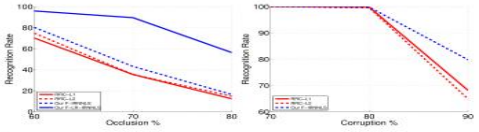


Fig. 12. Reconstruction results for various methods: The left image demonstrates the reconstructed face and the right image shows the estimated error for each of the methods tested.



Dataset	Yale (60% occl.) Accuracy	Time	Yale (90% corr.) Accuracy	Time
RRC_L1 [14]	70.11%	9.16	68.13%	8.75
RRC_L2 [14]	74.73%	3.69	64.84%	3.77
Our F-IRNNLS	80.22%	**1.52**	**79.78%**	**2.84**
Our F-LR-IRNNLS	**95.82%**	2.41	n/a	n/a

Fig. 13. Performance of RRC (t = 25 iterations) and our method (t = 100 iterations) under Block Occlusion and Pixel Corruption.

Face reconstruction was adequate for the case with scarves occlusion for all methods which validates the identification rates reported in Table II.

J. Time Performance Between Our Method and RRC

In this experiment we evaluate the identification rates in RRC [4] for the case where the maximum reweighted iterations t= 25. In other words, we want to investigate the performance degradation of RRC by keeping its execution time similar to our method. In our method we kept t =100. As shown in Figure 13, we observe that the computational time for RRC is now more competitive (although still higher than our method). However, the identification rates dropped significantly in both pixel corruption and block occlusion cases for RRC with t = 25.

K. Regularization of the Coefficients

In Table VII we report performance comparisons of our method with different regularizations of the representation coefficients. Our main take away from

the results is that sparsity is overall slightly better than the two other regularizes (non-negative and 4_2) in terms of identification rates. However, the non-negative regularizer provided a better balance between computational cost and identification rates.

Finally, there is significant difference in time performance between the RRC and our method regardless of the regularization of the coefficients. The efficiency of our method gives rise to robust face recognition systems for which computational time is a critical factor.

V. Conclusion

Finally, this project presents two characteristics. The first fits to the errors a distribution described by a tailored loss function. The second describes the error image as structural (low-rank). Our approach is computationally efficient due to the utilization of ADMM and PCA. The extensive experimental results support the claim that the proposed modeling of the error term can be beneficial and more robust than previous state-of-the-art methods to handle occlusions across a multitude of databases and in different scenarios.

References

[1] S. Albelwi and A. Mahmood. A deep architecture for face recognition based on multiple feature extraction techniques. In IEEE International Conference on Signal and Image Processing Applications, pages 390–395, Sept. 2017.

[2] A. Beck and M. Teboulle.A fast iterative shrink age thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences, 2(1):183–202, 2009. 4

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):711–720, Jul. 1997. 1

[4] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli. On the use of SIFT features for face authentication. In IEEE Conference on Computer Vision and Pattern Recognition Workshop, pages 35–35, Jun. 2006. 1

[5] S. P. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning, 3(1):1–122, 2011. 4

[6] P. Combettes and V. Wajs. Signal recovery by proximal forward-backward splitting. Multiscale Modeling & Simulation, 4(4):1168–1200, 2005. 4

[7] W. Deng, J. Hu, and J. Guo. Extended SRC: Undersampled face recognition via intraclass variant dictionary. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(9):1864–1870, Sept. 2012. 2

[8] W. Deng, J. Hu, and J. Guo. Face recognition via collaborative representation: Its discriminant nature and superposed representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(10):2513–2521, Oct. 2018. 2

[9] X. Ding, L.He, and L. Carin. Bayesian robust principal component analysis. IEEE Transactions on Image Processing, 20(12):3419–3430, Dec. 2011. 1

[10] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6):643– 660, Jun. 2001. 5

