Robust regression for face recognition

Imran Naseem a,*,1, Roberto Togneri b, Mohammed Bennamoun c

a College of Engineering, Karachi Institute of Economics and Technology (KIET), Karachi 75190, Pakistan
b School of EEEC Engineering, University of Western Australia, WA 6009, Australia
c School of CSS Engineering, University of Western Australia, WA 6009, Australia

A R T I C L E   I N F O

Article history:
Received 13 October 2010
Received in revised form 18 April 2011
Accepted 5 July 2011
Available online 14 July 2011

Keywords:
Robust face recognition
Random pixel contamination
Robust estimation

A B S T R A C T

In this paper we address the problem of robust face recognition by formulating the pattern recognition task as a problem of robust estimation. Using a fundamental concept that in general, patterns from a single object class lie on a linear subspace (Barsi and Jacobs, 2003 [1]), we develop a linear model representing a probe image as a linear combination of class specific galleries. In the presence of noise, the well-conditioned inverse problem is solved using the robust Huber estimation and the decision is ruled in favor of the class with the minimum reconstruction error. The proposed Robust Linear Regression Classification (RLRC) algorithm is extensively evaluated for two important cases of robustness i.e. illumination variations and random pixel corruption. Illumination invariant face recognition is demonstrated on three standard databases under exemplary evaluation protocols reported in the literature. Comprehensive comparative analysis with the state-of-art illumination tolerant approaches indicates a comparable performance index for the proposed RLRC algorithm. The efficiency of the proposed approach in the presence of severe random noise is validated under several exemplary noise models such as dead-pixel problem, salt and pepper noise, speckle noise and Additive White Gaussian Noise (AWGN). The RLRC algorithm is found to be favorable compared with the benchmark generative approaches.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Increasing security threats have highlighted the need of efficient and foolproof authentication systems for sensitive facilities. In this regard, biometrics have demonstrated good performance. Among the other available biometrics, such as speech, iris, fingerprints, hand geometry and gait, face seems to be the most natural choice [2]. It is nonintrusive, requires a minimum of user cooperation and is cheap to implement. The importance of face recognition is highlighted for widely used video surveillance systems where we typically have facial images of suspects [3]. With the recent development in multimedia signal transmission and processing, we witness emerging applications in the paradigm of face recognition. Face recognition through Internet, for instance, being one of the latest implementations. However, these relatively new applications tend to equally signify the problem of robustness. Although training is conducted offline in a controlled laboratory environment, probe images are always vulnerable to distortion due to ambient luminance, sensor malfunctioning, channel noise, compression noise over the Internet medium, etc. [4,5].

In general, face recognition systems critically depend on manifold learning methods. A gray-scale face image of order \(a \times b\) can be represented as an \(ab\) dimensional vector in the original image space. Typically, in pattern recognition problems, it is believed that high-dimensional data vectors are redundant measurements of an underlying source. The objective of manifold learning is therefore to uncover this “underlying source” by a suitable transformation of high-dimensional measurements to low-dimensional data vectors. Therefore, at the feature extraction stage, images are transformed to low-dimensional vectors in a face space. The main objective is to find a basis function for this transformation, which could distinguishably represent faces in the face space. In the presence of noise, however, it is supposed to be an extremely challenging task [3,6]. It follows from coding theory that iterative measurements are more likely to safely recover information in the presence of noise [7], therefore working in a low-dimensional feature space maintaining the aspect of robustness is in fact an ardent problem in object recognition. A number of approaches have been reported in the literature for dimensionality reduction. In the context of robustness, these approaches have been broadly classified in two categories namely generative/reconstructive and discriminative methods [8]. Reconstructive approaches (such as PCA [9], ICA [10] and NMF [11,12])
are reported to be robust for the problem related to missing and contaminated pixels, these methods essentially exploit the redundancy in the visual data to produce representations with sufficient reconstruction property. Formally, given an input \( x \) and label \( y \), the generative classifiers learn a model of the joint probability \( p(x|y) \) and classify using \( p(y|x) \), which is determined using the Bayes’ rule. The discriminative approaches (such as LDA [13]), on the other hand, are known to yield better results in “clean” conditions [14] owing to the flexible decision boundaries. The optimal decision boundaries are determined using the posterior \( p(y|x) \) directly from the data [8] and are consequently more sensitive to outliers. Apart from these traditional approaches, it has been shown recently that unorthodox features such as downsampled images and random projections can serve equally well. In fact the choice of the feature space may no longer be so critical [15–17]. What really matters is the dimensionality of the feature space and the design of the classifier.

In the paradigm of face recognition, illumination variation is supposed to be a major robustness issue [3]. Several approaches have been proposed in the literature to tackle the problem. A sophisticated approach in [18] models images of a subject with a fixed pose but different illumination conditions as a convex cone in the space of images. Consequently a small number of training images of each face taken under various lighting conditions are used for the reconstruction of shape and albedo of the face. Although the approach has demonstrated some good results, in practice the computation of an exact illumination cone for a given subject is quite expensive and tedious due to a large number of extreme rays. Studies have shown that the facial images under varying luminance conditions can be modeled as low-dimensional linear subspaces [19]. Basis images for this images under varying luminance conditions can be modeled as low-dimensional linear subspaces [19]. Basis images for this case are used as basis vectors for low-dimensional linear space. Another line of action is to normalize/compensate the illumination effect by some kind of preprocessing such as histogram equalization, gamma correction and logarithm transform. However, these elementary global processing techniques are not of much help in the presence of nonuniform illumination variations [20]. Moreover some latest approaches such as the Line Edge Map (LEM) [21] and Face-ARG matching [22] have shown good tolerance under adverse illumination alterations. The use of geometrical/structural information of the face region justifies the implicit robustness of these approaches.

Apart from the illumination problem, it has also been shown in the literature that traditional face recognition approaches do not cope well in the presence of severe random noise [17,23–26]. Most of the approaches in the literature, robust to random pixel noise, are variants of neural network classification. An important work is presented in [24] incorporating a robust kernel approach in the presence of severe noise. Encouraging results have been shown for two important problems of the additive noise (salt and pepper) and the multiplicative noise (speckle) compared to the traditional SVM approaches. Similarly in [26] it has been shown that neural network classifier outperforms traditional PCA [9], 2DPCA [27], LDA [13] and Laplacianfaces [28] approaches for the case of severe additive Gaussian noise. Apart from these neural network approaches, recently sparse representation classification (SRC) has been presented [17,25]. In the presence of noise (modeled as uniform random variable) the proposed approach has shown to outperform the traditional approaches of PCA, ICA I, LNMF and \( L^2 + NS \), however, other important noise models such as speckle and salt and pepper noise are not addressed.

In this research we propose a robust classification algorithm for the problem of face recognition in the presence of random pixel distortion. Samples from a specific object class are known to lie on a linear subspace [1,13]. In our previous work [15,16] we proposed to develop class specific models of the registered users thereby defining the task of face recognition as a problem of linear regression. In the work presented here, we extend our investigations to the problem of noise contaminated probes, where the inverse problem is solved using a novel application of the robust linear Huber estimation [29,30] and the class label is decided based on the subspace with the most precise estimation. The proposed approach, although being simple in architecture, has shown demonstrating results for two critical robustness issues of severe illumination variations and random pixel noise.

Although, in principle, compensation for contaminated/missing pixels can be addressed at learning or/and classification stages, robust regression for classification has rarely been addressed in the literature [31]. Efforts in this direction mainly focus on development of robust learning methods which aim to detect outliers in the learning phase and work on potentially uncorrupted learning data thereby resulting in a high breakdown point. The majority of these robust learning approaches are based on adapting a discriminative model by replacing the classical location and scatter matrix estimators by their robust counterparts such as MVE estimators [32], MCD estimators [33–35] and the projection pursuit approach [36], etc. Some robust variants of generative methods have been proposed which make use of robust metrics rather than a standard least-square computation [31,37]. An important implementation to the problem of robust object recognition is presented in [38] which improves the learning process of the traditional eigenspace method by introducing a hypothesize-and-test paradigm. Fundamentally a subsampling approach is introduced which works on subsets of image points, generating each hypothesis by the robust solution of a set of linear equations. Based on the Minimum Description Length (MDL) principle, competing hypotheses are further selected for the determination of the eigenspace coefficients. Recently a more sophisticated approach has been proposed in [31] which essentially combines the discriminative and reconstructive models to construct such a subspace which characterizes the discrimination power and reconstruction property of the two methodologies simultaneously. An important relevant work is presented in [39], where generative probabilistic image formation models are developed for the in-lier and out-lier processes. The task of face recognition is thereby formulated as a Maximum-a-Posteriori (MAP) estimation problem, the approach is applied to the problem of contiguous occlusion. To the best of our knowledge, however, it is for the first time that the problem of randomly missing/corrupted pixels has simply been formulated as a linear robust regression task.

The rest of the paper is organized as follows: The fundamental problem of robust estimation is discussed in Section 2 followed by the face recognition problem formulation in Section 3. Section 4 demonstrates the efficacy of the proposed approach for the problem of severely varying illumination followed by the experiments for random pixel corruption in Section 5. The paper finally concludes in Section 6.

2. The problem of robust estimation

Consider a linear model

\[
y = X\beta + e
\]

(1)

where the dependent or response variable \( y \in \mathbb{R}^{\times 1} \), the regressor or predictor variable \( X \in \mathbb{R}^{n \times p} \), the vector of parameters \( \beta \in \mathbb{R}^{p \times 1} \) and error term \( e \in \mathbb{R}^{\times 1} \). The problem of robust estimation is to
estimate the vector of parameters $\hat{\beta}$ so as to minimize the residual
\[ \mathbf{r} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{X}\hat{\beta} \]
$\hat{\mathbf{y}}$ being the predicted response variable. In classical statistics the
error term $\mathbf{e}$ is conventionally taken as a zero-mean Gaussian noise [40]. A traditional method to optimize the regression is to
minimize the least squares (LS) problem
\[ \arg \min_{\hat{\beta}} \sum_{j=1}^{q} r_j^2(\hat{\beta}) \]
where $r_j(\hat{\beta})$ is the $j$th component of the residual vector $\mathbf{r}$. However, in the presence of outliers, least squares estimation is inefficient and can be biased. Although it has been claimed that
classical statistical methods are robust, they are only robust in the
sense of type I error. Type I error corresponds to the rejection of
null hypothesis when it is in fact true. It is straightforward to note that
Type I error rate for classical approaches in the presence of outliers tend to be lower than the nominal value. This is often referred to as conservatism of classical statistics. However, due to
contaminated data, type II error increases drastically. Type II error is the error when the null hypothesis is not rejected when it is in
fact false. This drawback is often referred to as inadmissibility of
the classical approaches. Additionally, classical statistical meth-
ods are known to perform well with the homoskedastic data
model. In many real scenarios, however, this assumption is not
true and heteroskedasticity is indispensable, thereby emphasizing
the need of robust estimation.

Several approaches to robust estimation have been proposed such as $R$-estimators and $L$-estimators. However, $M$-estimators
have shown superiority due to their generality and high break-
down point [29,40]. Primarily $M$-estimators are based on mini-
mizing a function of residuals
\[ \hat{\beta} = \arg \min_{\hat{\beta} \in \mathbb{R}^p} \left\{ F(\hat{\beta}) \equiv \sum_{j=1}^{q} \rho(r_j(\hat{\beta})) \right\} \]
where $\rho(r)$ is a symmetric function with a unique minimum at
zero [29,30]
\[ \rho(r) = \begin{cases} \frac{1}{2}r^2 & \text{for } |r| \leq \gamma \\ \frac{1}{2}r^2 - \gamma^2 & \text{for } |r| > \gamma \end{cases} \]
$\gamma$ being a tuning constant called the Huber threshold. Many algorithms have been developed for calculating the Huber
$M$-estimate in Eq. (4), some of the most efficient are based on
Newton’s method [41]. $M$-estimators have been found to be robust and statistically efficient compared to classical methods
[42–44]. Although robust methods, in general, are superior to their classical counterparts, they have rarely been addressed in
applied fields [31,40]. Several reasons have been discussed in [40]
for this paradox, computational expense related to the robust
methods has been a major hindrance [42]. However, with recent
developments in computational power, this reason has become
insignificant. The reluctance in the use of robust regression
methods may also be credited to the belief of many statisticians
that classical methods are robust.

3. Robust Linear Regression Classification (RLRC) for robust
face recognition
Consider $N$ number of distinguished classes with $p_i$ number of
training images from the $i$th class such that $i = 1, 2, \ldots, N$. Each
grayscale training image is of an order $a \times b$ and is represented
as $\mathbf{u}_i \in \mathbb{R}^{a \times b}, i = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, p_i$. Each gallery image is
downsampled to an order $c \times d$ and transformed to a vector
through column concatenation such that $\mathbf{u}_i \in \mathbb{R}^{a \times b} \rightarrow \mathbf{w}_i \in \mathbb{R}^{c \times d}$, where $q = cd, \ ax \ ab$. Each image vector is normalized so that
the maximum pixel value is 1. Using the concept that patterns from
the same class lie on a linear subspace [1], we develop a class
specific model $\mathbf{X}_i$ by stacking the $q$-dimensional image vectors,
\[ \mathbf{X}_i = [\mathbf{w}_{i(1)}, \mathbf{w}_{i(2)}, \ldots, \mathbf{w}_{i(p_i)}] \in \mathbb{R}^{c \times p_i}, \quad i = 1, 2, \ldots, N \]
Each vector $\mathbf{w}_i$, $m = 1, 2, \ldots, p_i$, spans a subspace of $\mathbb{R}^q$ also
called the column space of $\mathbf{X}_i$. Therefore at the training level each
class $i$ is represented by a vector subspace, $\mathbf{X}_i$, which is also called
the regressor or predictor for class $i$. Let $z$ be an unlabeled test
image and our problem is to classify $z$ as one of the classes
$i = 1, 2, \ldots, N$. We transform and normalize the grayscale image
$z$ to an image vector $\mathbf{y} \in \mathbb{R}^{c \times d}$ as discussed for the gallery. If $\mathbf{y}$
belongs to the $i$th class it should be represented as a linear
combination of the training images from the same class (lying in
the same subspace) i.e.
\[ \mathbf{y} = \mathbf{X}_i \hat{\beta}_i + \mathbf{e}, \quad i = 1, 2, \ldots, N \]
where $\hat{\beta}_i \in \mathbb{R}^{q \times 1}$. From the perspective of face recognition
the training of the system corresponds to the development of the
explanatory variable ($\mathbf{X}_i$) which is normally done in a controlled
environment, therefore the explanatory variable can safely be
regarded as noise free. The issue of robustness comes into play
when a given test pattern is contaminated with noise which may
arise due to lumiance, malfunctioning of the sensor, channel
noise, etc. Given that $q \geq p_i$, the system of equations in Eq. (7) is
well-conditioned and $\hat{\beta}_i$ is estimated using robust Huber estimation
as discussed in Section 2 [30]
\[ \hat{\beta}_i = \arg \min_{\hat{\beta}_i \in \mathbb{R}^q} \left\{ F(\hat{\beta}_i) \equiv \sum_{j=1}^{q} \rho(r_j(\hat{\beta}_i)) \right\}, \quad i = 1, 2, \ldots, N \]
where $r_j(\hat{\beta}_i)$ is the $j$th component of the residual
\[ \mathbf{r}(\hat{\beta}_i) = \mathbf{y} - \mathbf{X}_i \hat{\beta}_i, \quad i = 1, 2, \ldots, N \]
The estimated vector of parameters, $\hat{\beta}_i$, along with the pre-
dictors $\mathbf{X}_i$ is used to predict the response vector for each class $i$
\[ \hat{\mathbf{y}}_i = \mathbf{X}_i \hat{\beta}_i, \quad i = 1, 2, \ldots, N \]
We now calculate the distance measure between the predicted
response vector $\hat{\mathbf{y}}_i$, $i = 1, 2, \ldots, N$ and the original response
target vector $\mathbf{y}$
\[ d_i(\mathbf{y}) = \| \mathbf{y} - \hat{\mathbf{y}}_i \|_2, \quad i = 1, 2, \ldots, N \]
and rule in favor of the class with minimum distance i.e.
\[ \min_{i} d_i(\mathbf{y}) \]

4. Case study: face recognition in presence of severe
illumination variations
The proposed RLRC algorithm is extensively evaluated on
various databases incorporating several modes of lumiance
variations. In particular we address three standard databases
namely Yale Face Database B [18], CMU-PIE database [45] and
AR database [46]. For all experiments images are histogram
equalized and transformed to logarithm domain.

4.1. Yale Face Database B
Yale Face Database B consists of 10 individuals with 9
poses incorporating 64 different illumination alterations for each
pose [18]. The database has been used by researchers as a test-
bed for the evaluation of robust face recognition algorithms. Since
we are concerned with the illumination tolerant face recognition
problem, only the frontal images of the subjects are considered.
The images are divided into five subsets with respect to the angle
between the light source direction and the camera axis (see
Fig. 1), refer to Table 1. Interested readers may also refer to [18]
for further details of the database, all images are downsampled to
an order of 50/C2 50.

We follow the evaluation protocol as reported in [18,20,47–51].
Training is conducted using subset 1 and the system is validated on
the remaining subsets. A detail comparison of the results with some
latest approaches is shown in Table 2, all results are as reported in
[20]. Note that the error rates have been converted to the recogni-
tion success rates. Since subset 3 incorporates moderate luminance
variations, most of the state-of-art algorithms report error-free
recognition as shown in Table 2. For subset 4 with more adverse
illumination variations, the proposed algorithm achieves 100%
recognition which is either better than or comparable to all
the results reported in the literature. In particular the proposed
approach outperforms the Cones-attached, Illumination Ratio
Images and Quotient Illumination Relighting methods by 8.60%,
18.60% and 9.40%, respectively. It is also found to be fairly compar-
able to the latest Cone-cast and Gradient Angle approaches. Subset 5
represents the worst case scenario with angle between the light
source direction and camera axis being greater than 77°. The
proposed RLRC algorithm consistently achieves 100% recognition
for the severe alterations comparing favorably with all the reported
results in the literature beating the Quotient Illumination Relighting
method by more than 17%. Noteworthy is the fact that results
for this subset are not available in the literature for most of the
contemporary approaches.

4.2. CMU-PIE face database

4.2.1. Evaluation Protocol 1 (EP 1)

Extensive experiments were conducted on CMU-PIE database
[45]. We follow the evaluation protocol as proposed in [52]
randomly selecting a subset of database consisting of 65 subjects
with 21 illumination variations per subject, all images are resized
to an order 50 × 50. Fig. 2 represents 21 different alterations for a
typical subject each image being accordingly labeled. We follow
two experimental setups as proposed in [52], in the first set of
experiments the system is trained using images with near frontal
lighting and validation is conducted across the whole database.
A detailed comparison of the performance with the state-of-art
approaches is depicted in Table 3. The proposed RLRC algorithm is
found to be pretty much comparable with the latest approaches of
MACE filters and Corefaces, it also comprehensively outperforms
the IPCA, 3D linear subspace and Fisherfaces methods for various
case-studies of training sessions. For instance with train-
ing images labeled 7, 10 and 19, the proposed RLRC algorithm
achieves 99.93% recognition which is 63.83%, 49.03% and 26.63%
better than IPCA, 3D linear subspace and Fisherfaces methods,
respectively.

For the second set of experiments training is conducted on
images captured under extreme lighting conditions, the system is
again validated across the whole database. The proposed RLRC
algorithm is found to be comparable with the latest approaches as

**Table 1**

Details of the subsets for Yale Face Database B with respect to light source
directions.

<table>
<thead>
<tr>
<th>Subsets</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lighting angle (degrees)</td>
<td>0–12</td>
<td>13–25</td>
<td>26–50</td>
<td>51–77</td>
<td>&gt; 77</td>
</tr>
<tr>
<td>Number of images</td>
<td>70</td>
<td>120</td>
<td>120</td>
<td>140</td>
<td>190</td>
</tr>
</tbody>
</table>

**Table 2**

Recognition results for Yale Face Database B.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Subset 3 (%)</th>
<th>Subset 4 (%)</th>
<th>Subset 5 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No normalization [20]</td>
<td>89.20</td>
<td>48.60</td>
<td>22.60%</td>
</tr>
<tr>
<td>Histogram equalization [20]</td>
<td>90.80</td>
<td>45.80</td>
<td>58.90%</td>
</tr>
<tr>
<td>Linear subspace [18]</td>
<td>100.00</td>
<td>85.00</td>
<td>N/A</td>
</tr>
<tr>
<td>Cones-attached [18]</td>
<td>100.00</td>
<td>91.40</td>
<td>N/A</td>
</tr>
<tr>
<td>Cones-cast [18]</td>
<td>100.00</td>
<td>100.00</td>
<td>N/A</td>
</tr>
<tr>
<td>Gradient Angle [47]</td>
<td>100.00</td>
<td>98.60</td>
<td>N/A</td>
</tr>
<tr>
<td>Harmonic images [48]</td>
<td>99.70</td>
<td>96.90</td>
<td>N/A</td>
</tr>
<tr>
<td>Illumination Ratio Images [49]</td>
<td>96.70</td>
<td>81.40</td>
<td>N/A</td>
</tr>
<tr>
<td>Quotient Illumination Relighting [50]</td>
<td>100.00</td>
<td>90.60</td>
<td>82.50%</td>
</tr>
<tr>
<td>9PL [51]</td>
<td>100.00</td>
<td>97.20</td>
<td>N/A</td>
</tr>
<tr>
<td>Method in [20]</td>
<td>100.00</td>
<td>99.82</td>
<td>98.29%</td>
</tr>
<tr>
<td>RLRC</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Fig. 1. Yale Face Database B: starting from top, each row represents typical images from subsets 3, 4 and 5, respectively. Note that subset 5 (third row) characterizes the
worst illumination variations.

Fig. 2. CMU-PIE database: starting from top, each row represents typical images from subsets 3, 4 and 5, respectively. Note that subset 5 (third row) characterizes the
worst illumination variations.
shown in Table 4. The only erroneous recognition trial was for the case with the training images labeled 3 and 16. The error may be attributed to the fact that the system was trained using only two images which accounts for only a couple of regressor or predictor observations for each class in the context of the RLRC algorithm. Apart from insufficient information, it has to be noted that images 3 and 16 (Fig. 2) have adverse luminance conditions.

4.2.2. Evaluation Protocol 2 (EP 2)

Under Evaluation Protocol 2 (EP 2) we follow the leave-one-out strategy on the 68 subjects of the CMU-PIE database as proposed in a recent work of generalized quotient image\[53\]. A detailed comparison with the best results in\[53\] is shown in Fig. 3. The proposed RLRC approach consistently attained high recognition accuracy for all leave-one-out experiments. In particular, apart from one recognition trail we attained an error-free performance index with 100% recognition accuracy. Only one error was reported for the seventh leave-one-out experiment where we achieved a recognition rate of 98.53%. It is appropriate to point out that performance curve for S-QI method in Fig. 3 is an approximation to the curve shown in [53].
4.3. AR database

The AR face database contains over 4000 color images taken in two sessions separated by two weeks [46]. The database characterizes various deviations from the ideal conditions including facial expressions, luminance conditions and occlusion modes. In particular, there are three lighting modes with left light on, right light on and both lights on. Fig. 4 represents these variations for the two sessions.

4.3.1. Evaluation Protocol 1 (EP 1)

We follow the evaluation protocol as proposed in [54], a subset of the database consisting of 118 randomly selected individuals is selected. Training is performed on images with nominal lighting conditions (Fig. 4(a) and (e)) while validation is conducted on images with adverse ambient lighting (Fig. 4(b), (c), (d), (f), (g) and (h)). Therefore altogether we have 236 (118 × 2) gallery images and 708 (118 × 6) probes.

All images are downsampled to an order of 180 × 180. The results are dilated in Table 5, the proposed RLRC algorithm outperforms the Locality Preserving Projections (LPP) method by a margin of 30.51% and is quite comparable to the Discriminant Locality Preserving Projections (DLPP) method. All results are as reported in [54].
4.3.2. Evaluation Protocol 2 (EP 2)

Under EP 2 we follow the experimental setup as proposed in [55]. We now have a subset of 121 subjects, training is done on Fig. 4(a) and the system is validated for adverse luminance variations of the same session i.e Fig. 4(b), (c) and (d). Therefore we have 121 gallery images and 363 \( \binom{121}{3} \) probes. The results are tabulated in Table 6, all the results are as reported in [55].

The proposed RLRC algorithm achieves a high recognition accuracy of 94.49% outperforming the latest 2D face model approach by a margin of 12.69%.

4.3.3. Evaluation Protocol 3 (EP 3)

Under Evaluation Protocol 3 (EP 3) we follow the experimental setup as proposed in recent works of Face-ARG matching [22] and Line Edge Map (LEM) [21]. These recent approaches use the geometric quantities and structural information of a human face and have therefore shown to be robust to severe illumination variations. We select a subset of AR database consisting

![Fig. 6. First row illustrates some gallery images from subsets 1 and 2 while second row shows some probes from subset 3.](image)

![Fig. 7. Probe images corrupted with (a) 20%, (b) 40%, (c) 60% and (d) 80% dead pixels.](image)

![Fig. 8. Recognition accuracy of various approaches for a range of dead pixel noise density.](image)

![Fig. 9. Probes with (a) 20%, (b) 40%, (c) 70% and (d) 90% salt and pepper noise density.](image)

![Fig. 10. Recognition accuracy curves in the presence of varying density of salt and pepper noise.](image)

![Fig. 11. Probe images corrupted with (a) 4, (b) 6, (c) 8 and (d) 10 variance speckle noise.](image)

| Table 8 |

| Verification results for dead-pixel noise. |
|---|---|---|---|---|
| 20% noise density | 40% noise density | 60% noise density | 80% noise density |
| Approach | EER (%) | Verif. (%) | Approach | EER (%) | Verif. (%) | Approach | EER (%) | Verif. (%) | Approach | EER (%) | Verif. (%) |
| PCA | 0.70 | 99.78 | PCA | 1.00 | 97.81 | PCA | 1.80 | 94.30 | PCA | 5.80 | 75.44 |
| ICA I | 0.60 | 100 | ICA I | 1.50 | 96.93 | ICA I | 3.70 | 92.11 | ICA I | 11.02 | 51.54 |
| RLRC | 0.20 | 100 | RLRC | 0.20 | 100 | RLRC | 0.20 | 100 | RLRC | 0.40 | 99.78 |
of 135 subjects. The system is trained using Fig. 4(a) while Fig. 4(b)–(d) serve as probes, altogether we have 135 gallery images and 405 (135 × 3) probes. The results are tabulated in Table 7, noteworthy is the fact that results in [21] are shown for 112 subjects.

The proposed RLRC approach shows a consistent performance across all illumination modes of the AR database. For the cases of “left light on” and “right light on”, recognition accuracies of 96.30% and 94.07% are achieved which are fairly comparable to the latest LEM and Face-ARG approaches as shown in Table 7. For

Fig. 12. Dead-pixel noise: (a)–(d) elaborate rank-recognition profiles while (e)–(h) show the Receiver Operating Characteristics (ROC) for 20%, 40%, 60% and 80% noise densities, respectively.
the most challenging problem of illumination with “both lights on”
the proposed RLRC approach attains 94.07% recognition which is
favorably comparable with the Face-ARG approach and outperforms
the LEM approach by a margin of approximately 20%. The conven-
tional methods of PCA and 1-NN reported in [22] are out of
discussion as they lag far behind these latest approaches.

Fig. 13. Salt and pepper noise: (a)–(d) represent rank-recognition curves while (e)–(h) show Receiver Operating Characteristics (ROC) for 50%, 60%, 70% and 80% noise
densities, respectively.
4.4. FERET database

The FERET database is arguably one of the largest publicly available database with two versions [56], gray FERET database and color FERET database. The database addresses several challenging issues such as expression variations, pose alterations and aging factor, etc. For the case of varying illumination there is only one evaluation protocol recognized as “faco” within the framework of gray FERET database. The methodology utilizes only one gallery image for each of the 1196 subjects, the gallery size is therefore 1196. The probe set consists of 194 images, refer to [56] for further details on the FERET evaluation methodology. It is worthy to note that recognizing a person from a single gallery image is itself an independent ardent issue within the paradigm of face recognition [57] and as such not the focus of the presented research. However, to evaluate the efficacy of the proposed algorithm with a single gallery image per subject, we conducted extensive experiments as shown in Fig. 5. The proposed RLRC algorithm outperformed 13 of the reported 14 algorithms and lagged only one algorithm tagged as USC MAR 97 in [56]. Fig. 5 illustrates receiver operating characteristics for three best reported algorithms (in the sense of recognition accuracy). The proposed RLRC algorithm achieves a verification accuracy of 70.10% at 0.001 FAR which lags 9.28% as compared to the best result reported for USC MAR 97, the proposed RLRC algorithm, however, comprehensively outperforms the other 13 algorithms beating UMD MAR 97 and EF HIST DEV M12 by a margin of approximately 30% and 50% (at 0.001 FAR), respectively. It has to be noted that for higher values of FAR the proposed RLRC algorithm is better than the best reported method. In particular the RLRC algorithm achieves better performance from 0.017 FAR onwards with a good equal error rate of 0.03 (approximately) which is better than 0.05 as reported for USC MAR 97 method [56].

5. Case study: face recognition in presence of random pixel noise

Extensive experiments were carried out using the Extended Yale B database [18,19]. The database consists of 2414 frontal-face images of 38 subjects under various lighting conditions. Subsets 1 and 2 consisting of 719 images under normal-to-moderate lighting conditions were used as gallery. Subset 3 consisting of 456 images under severe luminance alterations was designated as probes. Sample gallery and probe images are shown in Fig. 6. The choice of training and testing images is specifically to isolate the effect of noise.

The proposed approach was validated for a range of exemplary noise models specific to image data. For all experiments the location of noisy pixels is unknown to the algorithm. Fig. 7 reflects the probe images corrupted with various degrees of dead-pixel noise.

The proposed Robust Linear Regression Classification approach comprehensively outperforms the benchmark reconstructive algorithms of PCA [9] and ICA I [10] showing a high breakdown point as shown in Fig. 8. Even for the worst case scenario of 90% corrupted pixels, the proposed approach achieves 94.52% recognition accuracy outperforming PCA and ICA I by 49.13% and 75.44%, respectively. In the presence of outliers the L1-norm computation is reported to be more efficient compared to the usual euclidean distance measure [42,43]. Also in the literature median filtering has shown to improve image understanding in the presence of noise [5]. Therefore for an appropriate indication of the performance index for the proposed RLRC algorithm we also conducted experiments using median filtering and L1-norm calculations for the PCA. Note that the median preprocessing is indicated by letter “M” before the corresponding approach. The performance curves are shown in Fig. 8, the proposed RLRC algorithm consistently outperforms these robust variants of the benchmark approaches.

In particular, for the case of 90% corruption we note a major performance difference with the RLRC algorithm outperforming the L1-norm and M-PCA methods by 89.04% and 91.67%, respectively. Noteworthy is the fact that standard preprocessing and robust calculations are of no use in such severe noise conditions.

For a comprehensive comparison with the benchmark generative approaches, extensive verification experiments were conducted, the similarity scores were normalized on the scale (0 1). Results are dilated in Fig. 12 and Table 8. Rank recognition profiles for various degree of dead-pixel contamination show an excellent performance index for the proposed RLRC approach. With the increasing noise intensity the proposed approach shows good tolerance as compared to PCA and ICA I.

Specifically with 80% noise density, the proposed RLRC approach achieves an excellent rank-1 recognition accuracy of 99.78% comprehensively beating PCA and ICA I rank-1 recognition results by 28.29% and 48.68%, respectively. The RLRC approach achieves an excellent equal error rate (EER) of 0.40% as compared to 5.80% and 11.02% for PCA and ICA I, respectively. Verification rate of 99.78% at a typical 0.01 FAR, as indicated in Table 8, also comprehensively outperforms the benchmark approaches.

![Fig. 14. Recognition accuracy of various approaches in the presence of speckle noise for different variances. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)

<table>
<thead>
<tr>
<th>Approach</th>
<th>50% noise density</th>
<th>60% noise density</th>
<th>70% noise density</th>
<th>80% noise density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EER (%)</td>
<td>Verif. (%)</td>
<td>EER (%)</td>
<td>Verif. (%)</td>
</tr>
<tr>
<td>PCA</td>
<td>2.12</td>
<td>95.61</td>
<td>4.10</td>
<td>87.28</td>
</tr>
<tr>
<td>ICA I</td>
<td>2.16</td>
<td>93.86</td>
<td>4.25</td>
<td>83.99</td>
</tr>
<tr>
<td>RLRC</td>
<td>0.21</td>
<td>99.78</td>
<td>2.18</td>
<td>97.37</td>
</tr>
</tbody>
</table>

Table 9
Verification results for salt and pepper noise.
In the next set of experiments we contaminate the probe images with *data drop-out* and *snow in the image* noise simultaneously, commonly referred to as salt and pepper noise [58]. This fat-tail distributed noise, also called impulsive noise or *spike noise* [58], can be caused by analog-to-digital converter errors and bit errors in transmission [59,60]. Fig. 9 reflects probes distorted with various

![Graphs showing rank recognition curves and receiver operating characteristics.](image)

*Fig. 15*. Speckle noise: (a)–(d) represent rank-recognition curves and (e)–(h) show receiver operating characteristics for noise densities with variances 2, 4, 6, and 8, respectively.
degrees of salt and pepper noise. In the overall sense, the proposed RLRC approach is favorably comparable with the benchmark reconstructive approaches as depicted in Fig. 10.

At a noise density of 70% for instance, the RLRC algorithm gives 9.21% and 18.64% better recognition accuracy compared to PCA and ICA I, respectively. However, under high noise densities of 80% and 90%, PCA seems to be better than either ICA I or RLRC. It should be noted that under such severe conditions of salt and pepper noise, although PCA gives a better comparative performance index, the recognition accuracy achieved (for e.g. 48.25% at 80% noise density) is itself far from satisfactory. For salt and pepper noise we note that the median preprocessing results in a significant improvement both for PCA and RLRC approaches. For instance at 70% noise density M-PCA achieves 93.86% recognition which is 22.15% better than simple PCA. Similarly the M-RLRC shows an improvement of almost 20% compared to simple RLRC achieving a maximum recognition of 100%. The L1-norm calculation is of little benefit as M-RLRC consistently performs better than all competing approaches.

Verification experiments were also conducted for salt and pepper noise, results are shown in Fig. 13 and Table 9. For noise density up to 70%, the proposed RLRC algorithm shows better performance index, both in recognition and verification, compared to PCA and ICA I. For instance at 60% contamination level RLRC achieves a high rank-1 recognition accuracy of 94.52%, outperforming PCA and ICA I by a difference of 12.50% and 16.45%, respectively (refer to Fig. 13(b)). A low EER of 2.18% for the proposed RLRC approach is also favorable compared to 4.10% and 4.25% of PCA and ICA I, respectively. Note the major performance difference of the receiver operating characteristics in Fig. 13(f). An excellent verification rate of 97.37% at standard 0.01 FAR outstandingly outperforms the benchmark approaches (refer to Table 9). However, for a noise density greater than 70% the verification results for PCA are better than either RLRC or ICA I approaches. For instance at 80% noise contamination, PCA achieves a verification rate of 53.07% at a typical 0.01 FAR which is better than ICA I and RLRC by 12.53% and 13.16%, respectively, also the EER performance for PCA is superior compared to both approaches as shown in Table 9. The superior performance of PCA for severe noise density is, however, undone by the fact that it is unable to reach satisfactory performance in the absolute sense as 53.07% success rate is not reliable by any standard. For low to moderate salt and pepper noise, the proposed RLRC remains the best choice.

Speckle noise is regarded as a major interference in digital imaging and therefore forms another important robustness issue. The proposed approach was extensively evaluated by adding varying multiplicative speckle noise to probes as shown in Fig. 11. Speckle noise is efficiently modeled as a zero-mean uniform random variable.

The proposed RLRC approach showed a good performance index as shown in Fig. 14, consistently achieving a high recognition accuracy for a wide range of error variance. The effect of speckle noise with a variance of 6 is shown in Fig. 11(b), the image is badly distorted and traditional reconstructive approaches fail to produce competitive results. The proposed RLRC approach attains a high recognition accuracy of 91.67% outperforming PCA and ICA I approaches by 16.67% and 23.69%, respectively. Noteworthy is the tolerance and consistency of the proposed approach for highly corrupted data. The best recognition results are reported for the RLRC approach with median filtering (M-RLRC indicated by red dashed line in Fig. 14). For instance for the worst case of speckle noise with variance 10 the M-RLRC approach achieves a high recognition success of 97.37% comprehensively outperforming all competing approaches, the best competitor being the RLRC without any preprocessing achieving 86.40% (solid blue line in Fig. 14). Median filtering variants again show significant improvement of approximately 12% compared to the unprocessed computations. Note that the L1-norm robust calculations are of no much help in such adverse conditions.

Results for verification experiments are shown in Fig. 15 and Table 10. In particular, in the presence of noise density with variance 8, the RLRC approach achieves a high rank-1 recognition accuracy of 89.04% outperforming PCA and ICA I by margins of 19.52% and 22.15%, respectively. At a typical 0.01 FAR the proposed RLRC reaches a verification rate of 94.74% with an EER of only 3.07% substantially outperforming both contesting approaches (refer to Fig. 15(d) and (h)).

Due to the fact that all imaging systems acquire images by counting photons, the detector noise (modeled as Additive White Gaussian Noise) is always an important case-study in the context of robustness [61]. The probes were distorted by adding zero-mean Gaussian noise with a wide range of error variance as shown in Fig. 18. Since classical statistical methods are known to be efficient in the presence of Gaussian noise, we also conducted experiments by solving Eq. (7) using the least squares (LS) approach. To harness redundant measurements in the presence of noise, all experiments for LS and RLRC were conducted in the original image space.

Results shown in Fig. 17 reflect the superiority of the proposed approach. The RLRC approach consistently outperformed all other approaches for a wide range of error variance. In particular, with an error variance of 0.8 the RLRC approach beats PCA, ICA I and LS methods by margins of 7.89%, 16.88% and 48.48%, respectively. Even with a severe additive noise of 0.9 variance a reasonable 93.64% recognition accuracy was achieved. The LS approach showed an interesting behavior for low variance noise the performance of LS is pretty much comparable to the RLRC approach. However, with low SNR the LS method substantially lags the Robust Linear Regression Classification. The best recognition results are obtained for the RLRC with median filtering (M-RLRC), for instance with 0.8 error variance a recognition accuracy of 99.78% is reported which is 3.73% better than plain RLRC approach. Median filtering also substantially improved the performance of PCA, however, the two top performance curves are obtained for the M-RLRC and RLRC methods.

Verification results for various SNR case-studies of AWGN are shown in Fig. 16 and Table 11. In particular, for the worst case scenario of 0.9 variance Gaussian noise, the proposed RLRC approach achieves high verification rate of 98.68% at 0.01 FAR comprehensively outperforming PCA, ICA I and the LS methods by 12.50%, 12.28% and 51.75%, respectively (see Fig. 16(h)). The huge

<table>
<thead>
<tr>
<th>Noise variance</th>
<th>Approach</th>
<th>EER (%)</th>
<th>Verif. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>PCA</td>
<td>3.20</td>
<td>87.94</td>
</tr>
<tr>
<td></td>
<td>ICA I</td>
<td>4.60</td>
<td>86.84</td>
</tr>
<tr>
<td></td>
<td>RLRC</td>
<td>0.52</td>
<td>99.34</td>
</tr>
<tr>
<td>4</td>
<td>PCA</td>
<td>4.09</td>
<td>81.80</td>
</tr>
<tr>
<td></td>
<td>ICA I</td>
<td>5.57</td>
<td>77.63</td>
</tr>
<tr>
<td></td>
<td>RLRC</td>
<td>1.00</td>
<td>98.90</td>
</tr>
<tr>
<td>6</td>
<td>PCA</td>
<td>5.71</td>
<td>79.82</td>
</tr>
<tr>
<td></td>
<td>ICA I</td>
<td>6.36</td>
<td>70.61</td>
</tr>
<tr>
<td></td>
<td>RLRC</td>
<td>2.61</td>
<td>96.27</td>
</tr>
<tr>
<td>8</td>
<td>PCA</td>
<td>6.62</td>
<td>73.90</td>
</tr>
<tr>
<td></td>
<td>ICA I</td>
<td>8.27</td>
<td>67.76</td>
</tr>
<tr>
<td></td>
<td>RLRC</td>
<td>3.07</td>
<td>94.74</td>
</tr>
</tbody>
</table>
Fig. 16. Gaussian noise: (a)–(d) represent rank-recognition curves while (e)–(h) show Receiver Operating Characteristics (ROC) for noise densities with variances 0.5, 0.7, 0.8, and 0.9, respectively.

A careful analysis of the above reported results highlight an interesting behavior in presence of high intensity noise. For performance difference of more than 50% compared to the LS approach signifies the importance of robust regression for the particular case-study of face recognition. Also in terms of EER the proposed approach attains excellent performance index of 1.63% while other approaches substantially lag behind (refer to Table 11).
instance in Fig. 17 (Gaussian noise with 0.9 variance) the proposed RLRC approach gives almost equivalent performance compared to benchmark subspace techniques. This relates to the issue of low breakdown point of robust estimation in presence of severe noise. Classically speaking, for severely corrupted data adequate preprocessing increases the breakdown point [5]. Therefore the median processed RLRC (referred to as M-RLRC) shows excellent performance achieving approximately 100% recognition accuracy for even the worst case scenario of 0.9 variance in Fig. 17.

6. Conclusion

In this paper we present a novel robust face recognition algorithm based on the robust Huber estimation approach. It is for the first time that the problem of robust face recognition has been formulated as a robust Huber estimation task. The proposed Robust Linear Regression Classification (RLRC) algorithm has been evaluated for two case-studies i.e. severe illumination variation and random pixel corruption. For the case of illumination invariant face recognition, we have demonstrated results on three standard databases incorporating adverse luminance alterations. A comprehensive comparison with the state-of-art robust approaches indicates a comparable performance index for the proposed RLRC approach. We demonstrate, for the first time, an error-free recognition for the most challenging subset 5 of the Yale Face Database B. In addition we report a comparable evaluation for the RLRC algorithm on the CMU-PIE, AR and FERET databases under standard evaluation protocols as reported in the literature.

In addition, the problem of random pixel corruption is also addressed. The proposed RLRC approach has shown good results for various noise models comprehensively outperforming the reconstructive benchmark approaches. In particular the proposed approach attains a high verification rate of 99.78% at 0.01 FAR for the important case-study of probes contaminated with 80% dead-pixel noise. This performance is appreciable considering that the benchmark approaches are unable to provide satisfactory results under such severe noisy conditions. Similarly in the presence of severe AWGN the proposed RLRC approach beats the traditional generative approaches by a margin of an order of 12%. It has also been experimentally shown that the classical LS approach is extremely inefficient in the presence of severe AWGN and lags the proposed RLRC algorithm by more than 50%. For a fair comparison the robust variants of the base systems are also evaluated and a comprehensive comparison demonstrates the efficacy of the proposed approach.

It is adequate to point out that the proposed RLRC algorithm has shown demonstrating results for severe illumination and random pixel corruption. Accordingly it should be considered as an important module of a hierarchical, unconstrained face recognition system where other challenges such as severe pose variations, contiguous occlusion, etc. should be considered separately.

Apart from the good performance index of the proposed approach, there are several interesting outcomes of the presented research. In the paradigm of view-based face recognition, the choice of features for a given case-study has been a debatable topic. Recent research has, however, shown competency of the unorthodox features such as downsampled images and random projections, indicating a divergence from the conventional ideology [15–17]. The proposed RLRC approach in fact conforms to this emerging belief. Good results for randomly distributed noisy pixels are encouraging enough to extend the proposed algorithm for the problem of contiguous occlusion where contaminated pixels are known to have a connected neighborhood. Moreover due to good noise tolerance of reconstructive approaches, an important future direction is to explore efficient fusion of RLRC with PCA.

### References


