Recursive Sample Scaling Low-Rank Representation

Wenyun Gao, Xiaoyun Li, Sheng Dai, Xinghui Yin, and Stanley Ebhohimhen Abhadiomhen

1Nanjing LES Information Technology Co., LTD, Nanjing, China
2College of Computer and Information, Hohai University, Nanjing 211100, China
3Department of Computer Science, University of Nigeria, Nsukka, Nigeria

Correspondence should be addressed to Stanley Ebhohimhen Abhadiomhen; stanley.abhadiomhen@unn.edu.ng

Received 3 October 2021; Revised 1 December 2021; Accepted 11 December 2021; Published 27 December 2021

Copyright © 2021 Wenyun Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The low-rank representation (LRR) method has recently gained enormous popularity due to its robust approach in solving the subspace segmentation problem, particularly those concerning corrupted data. In this paper, the recursive sample scaling low-rank representation (RSS-LRR) method is proposed. The advantage of RSS-LRR over traditional LRR is that a cosine scaling factor is further introduced, which imposes a penalty on each sample to minimize noise and outlier influence better. Specifically, the cosine scaling factor is a similarity measure learned to extract each sample’s relationship with the low-rank representation’s principal components in the feature space. In order words, the smaller the angle between an individual data sample and the low-rank representation’s principal components, the more likely it is that the data sample is clean. Thus, the proposed method can then effectively obtain a good low-rank representation influenced mainly by clean data. Several experiments are performed with varying levels of corruption on ORL, CMU PIE, COIL20, COIL100, and LFW in order to evaluate RSS-LRR’s effectiveness over state-of-the-art low-rank methods. The experimental results show that RSS-LRR consistently performs better than the compared methods in image clustering and classification tasks.

1. Introduction

The limitations of classical feature learning techniques such as PCA [1] easily made the robust principal component analysis (RPCA) method an efficient choice for dealing with noise and outliers. Specifically, RPCA is focused on learning a low-rank subspace directly from the original high-dimensional data to preserve its geometric structure in a low-dimensional subspace. And this strategy has shown tremendous improvements in several applications [2–5]. However, as RPCA only seeks a single low-rank subspace, it may still be limited with noise damage since high-dimensional data are known to reside in multiple low-dimensional subspaces [6]. Thus, extending RPCA’s idea, Liu et al. [4, 7] proposed a method named “low-rank representation” (LRR). LRR’s main advantage over RPCA lies in its aim to learn data’s multiple low-dimensional subspaces and their membership. This approach makes LRR very robust to the negative effect of noise and outliers [8].

Therefore, considering LRR’s robustness mentioned above, several attempts were made in the literature such as references [9–13] to improve its performance. For example, to deal with the data from nonlinear subspaces, Tang et al. [13] proposed robust kernel LRR (RKLRR). Liu et al. [11] then adopted a fixed-rank strategy to accelerate LRR’s computation process. However, their performances could be reduced with insufficient or heavily corrupted samples. That is why Xiao et al. [14] had previously proposed the latent LRR (LatLRR) for joint subspace segmentation and feature selection. The idea behind LatLRR is to include hidden data in constructing the dictionary to improve robustness further. On the contrary, how to handle gross data damage remains unsolved. It is observed that the more the data corrupted by noise, the larger the degradation in the classification and clustering performance.

To address this issue, a recursive sample scaling low-rank representation (RSS-LRR) method is proposed in this paper. Since some data samples will be more damaged in gross data...
corruption, we estimate each data sample’s importance using a cosine scaling factor. This scaling factor measures the angle connecting each data sample and the low-rank representation’s principal components in feature space. In this way, we then iteratively subdue noisy data samples to overcome their effect. Thus, the proposed RSS-LRR can effectively obtain a good low-rank representation than existing methods. Our main contributions are summarized as follows:

1. We propose a novel method named “RSS-LRR,” which measures each data sample’s importance using a cosine scaling factor. This scaling factor is used in our model to extract each data sample’s relationship with the low-rank matrix’s principal components.

2. The proposed RSS-LRR method can effectively handle noisy data samples by iteratively restricting noisy data samples using the sample scaling factor to suppress noise so as to obtain a good low-rank representation.

3. Several experiments are performed with varying levels of corruption to evaluate RSS-LRR’s effectiveness. The experimental results show that RSS-LRR consistently outperforms other state-of-the-art methods in image clustering and classification tasks.

2. Related Work

This section presents a brief review of the baseline methods RPCA and LRR. First, in this paper, matrices are written in uppercase, e.g., \(X\). Thus, \(\|X\|_F\) and \(\|X\|_*\) denote the Frobenius norm and nuclear norm, respectively. \(\|X\|_2\) and \(\|X\|_{2,1}\) denote the vector norm and \(L_{2,1}\)-norm, which are defined as \(\|X\|_2 = \sum_{ij} |x_{ij}|\) and \(\|X\|_{2,1} = \sum_i \|x_i\|^2_{2}\).

2.1. Robust Principal Component Analysis. To recover a subspace structure from corrupted data, RPCA was proposed in [2]. Its strategy is to decompose a given data matrix into two components matrices by solving the following optimization problem:

\[
\min_{D,E} \|D\|_* + \lambda \|E\|_1 \quad \text{s.t.} \quad X = D + E,
\]

where the data matrix \(X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^{m \times n}\) is a matrix of \(n\) samples in the \(m\)-dimensional space, \(D\) is the low-rank matrix, \(E\) is a sparse error matrix, and \(\lambda\) is the regularized parameter to balance the effects of the two terms. Thus, RPCA’s main objective from the above formula is to obtain low-rank and sparse elements by combining the nuclear norm and \(L_1\)-norm. This approach is proven to be possible under some assumptions [16]. However, as RPCA assumes a single low-rank subspace, its performance can degrade easily.

2.2. Low-Rank Representation. Liu et al. [5] proposed LRR to tackle RPCA’s limitations. Specifically, LRR is focused on pursuing a data representation matrix with the lowest rank. It can achieve that by using data’s self-expressiveness property such that the given data itself are utilized as a self-dictionary. This way, each data sample is then represented as a linear combination of similar samples belonging to the same class. The optimal low-rank matrix obtained by LRR is defined as follows:

\[
\min_{Z,E} \|Z\|_* + \lambda \|E\|_p \quad \text{s.t.} \quad X = XZ + E,
\]

where the data matrix \(X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^{m \times n}\) denotes the self-dictionary \(E\) is used to capture the error components where \(\|\cdot\|_p\) denotes a certain norm, which can be determined based on the type of noise corruption. For example, while \(\|\cdot\|_F\) is a suitable candidate for data damaged by Gaussian noise, \(\|\cdot\|_1\) is good for random noise. Besides, \(\|\cdot\|_1\) is an efficient choice when only a part of data are contaminated.

Although LRR’s approach is shown to be very effective, particularly in noisy settings, its performance may degrade with insufficient samples. For this reason, Liu et al. [15] also proposed latent LRR that exploits both the observed and the hidden data to construct the self-dictionary. This strategy is most useful for image restoration [17]. Consequently, the work of [18] proposed a method for line pattern noise removal to address contaminated instances. It is realized in a transform domain by using a line pattern’s directional property. Besides, other efforts were made in references [11, 12, 19–31] to improve LRR’s discriminative capability.

Notably, Bing-Kun Bao et al. [23] used a fixed-rank approach so as to reduce LRR’s singular value decomposition cost. Zhang et al. [24] proposed two instantaneous methods: the first can reasonably handle noise interference by decomposing given data into two parts, namely, the low-rank sparse principal feature part and a noise-fitting error part. In [25], Tang et al. introduced a diversity regularization and a rank constraint to suppress the redundancy in different data views. In [26], Zhang et al. presented a method that adaptively preserves local information of salient features, thus guaranteeing a block-diagonal coefficient structure. Meanwhile, a compressive robust subspace clustering method was proposed in [27] for dimensionality reduction. However, because the subspace techniques, including LRR, do not provide linear dimensionality reduction (LDR) functionality, the feature selective projection (FSP) [28] was proposed. FSP combines feature extraction, feature selection, and LRR into a unified model to promote robust LDR. Likewise, a method was introduced in [29], which exploits a robust dictionary learning strategy to discover hybrid salient low-rank and sparse representations in a factorized compressed space. Furthermore, in an attempt to keep both similarity and local structures, the hierarchical weighted low-rank representation (H威尔E) [30] was proposed. Similarly, the more recent study [31] was focused on capturing cross-view information through an approach that preserves both diversity and consensus information of each data view.
3. The Proposed Method

3.1. RSS-LRR. Since real-world data are not always perfect and inevitably corrupted by noise in practice, most existing low-rank methods cannot guarantee robust performance. Therefore, noise interference must be carefully handled to resolve the present drawback. As a result, a rational solution is pursued in this paper, which uses a cosine scaling factor to estimate the importance of each data sample. Essentially, we suppose that the clean data samples will give high significant values, while the noisy ones would differ from the principal component of the data. Thus, the cosine scaling factor $D$ is introduced into the LRR formulation in equation (2) using the constraint $X D = X D Z + E$. The motivation behind our approach is straightforward: according to reference [32], $Z$ can be decomposed as $U \Sigma V^T$, where $U$ and $V$ are left and right low-rank singular vectors. In other words, $U$ or $V$ becomes the pursued projection vector such that $X U e R^m$ is the data projection in feature space. Therefore, $X U_1$ is chosen as our maximum projection direction where $U_1$, the maximum eigenvalue of $\Sigma$, is $U$’s column vectors. So, for an outlier data sample $x_j$, the angle between it and the principal component vector $X U_1$ would differ more than that of clean data $x_j$, as described in Figure 1(a). Hence, $d_i$ expressed in the following is used to estimate the importance of each data sample:

$$d_i = |\cos \theta_i| + \epsilon = \frac{x^T_i (X U_1)}{\|x_i\| \cdot (X U_1)} + \epsilon,$$

where $\epsilon$ is a constant that stops $d_i$ from 0. Thus, using the significant factor $d$, a given data matrix can be scaled to minimize the effect of noisy data samples, allowing the low-rank structure to be realized with clean data as shown in Figure 1(b). As such, the proposed sample scaling low-rank model is obtained as follows:

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1}$$

s.t. $X D = X D Z + E,$

where $Z$ denotes the low-rank matrix and $E$ is used to capture the noise elements, similar to that in equation (2).

Then, $D = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$ denotes the scaling factor of $n$ samples.

From the above formulation, one can easily check that the scaling factor $\pi/2$ suits our goal, as we can then detect the noisy points from closer angles to $\pi/2$, with lower values being assigned to such points. Illustratively, let us assume $\theta_i = \pi/2$, and $d_i$ is almost 0, so $x_i$ is subdued with $\tilde{x}_i = x_i d_i$. Thus, $\tilde{X} = X D$ is used to obtain new training data. Besides, using SVD, $\tilde{X} = X D = U \Sigma V^T$, the singular vectors $\tilde{U}$ and new projection vectors of sample space are obtained by suppressing the noisy data samples. As a result, our proposed method can then learn an optimal low-rank structure using new data $\tilde{X}$ where the points closer to the principal component vector are enhanced.

We give the summary of our model’s main characteristics as follows:

3.2. Optimization. In this section, we propose an optimization algorithm to solve equation (4). First, following standard practice, we introduce a variable $J = Z$ to relax equation (4) further. Thus, equation (4) can be recast as

$$\argmin_{Z,E} ||J||_* + \lambda \|E\|_{2,1}$$

s.t. $X D = X D Z + E, Z = J.$

The augmented Lagrangian function equation (5) is given as

$$L(J,E,Z,\mu) = J_\star + \lambda E_{2,1} + \text{tr}(M_1^T (X D - X D Z - E))$$

$$+ \text{tr}(M_2^T (Z - J))$$

$$+ \frac{\mu}{2} (\|X D - X D Z - E\|^2_F + \|Z - J\|^2_F),$$

where $M_1$ and $M_2$ are Lagrange multipliers, $\mu > 0$ is a penalty parameter, and $\| \cdot \|_F$ denotes the Frobenius norm of a matrix. Many concepts for convex optimization have been developed [33–35], which rely on nuclear-norm regularization. And the optimization problem can be solved via the method in [36].
3.2.1. Computation of $J$. According to references [37, 38], nuclear-norm minimization methods have a stable performance. For computing $J$, we rewrite equation (6) as

$$J^* = \arg\min_J \|J\|_* + \frac{1}{2} \|\| - \left( Z + \frac{M_2}{\mu} \right)^2_F. \quad (7)$$

3.2.2. Computation of $Z$. By fixing $J$ and $E$ and substituting $\tilde{X} = XD, Z$ can be updated using the following formula:

$$Z^* = \left( I + \tilde{X}^T \tilde{X} \right)^{-1} \left( \tilde{X}^T \tilde{X} - \tilde{X}^T E + J + \frac{\tilde{X}^T M_1 - M_2}{\mu} \right). \quad (8)$$

3.2.3. Computation of $E$. With $\tilde{X}, Z,$ and $J$ fixed, $E$ can be solved as follows:

$$E^* = \arg\min_E \frac{\lambda}{\mu} \|E\|_{2,1} + \frac{1}{2} \|E - \left( \tilde{X} - \tilde{X} Z + \frac{M_1}{\mu} \right)\|_F^2. \quad (9)$$

Following reference [39], equation (9) can be solved by the following lemma.

**Lemma 1** (see [40]). Let $Q = [q_1, q_2, \ldots, q_n]$ be a given matrix. If $W^*$ is the optimal solution, then

$$W^* = \min_{u, \lambda} \|W\|_{2,1} + \frac{1}{2\|\|W - Q\|_F^2}. \quad (10)$$

Thus, the $i$th column of $W^*$ is

$$W^*(i,:)=\begin{cases} q_i - \frac{\lambda}{\|q_i\|}, & \text{if } \lambda < \|q_i\| \\ 0, & \text{otherwise}. \end{cases} \quad (11)$$

Based on Lemma 1, supposing we have a matrix $Q$, $W^*$ can be reached directly using the above formula, making the $E$ computation process very efficient.

The complete solution is given in Algorithm 1.

3.3. Complexity Analysis. This section gives an analysis of the computational cost of Algorithm 1. In Step 6 of Algorithm 1, the values of $J_k, Z_k, E_k$ are the same as $J_{k-1}, Z_{k-1}, E_{k-1}$ in the inner loop. Therefore, as the number of iterations $k$ increases, the time complexity in pursuing the low-rank variable $Z_k$ decreases faster. From the above discussion, the computing time in $t_k$ is far shorter than that in $t_1$. The cost of SVQ is $O(n^r)$, where $n$ denotes the number of data vectors. Besides, $Z$ costs $O(n^r m + n^r)$ for computation, where $m$ is the data dimension. $E$ in Step 10 then costs $O(dn)$. Furthermore, supposing that the $D$ subblock requires $k$ times until convergence, the $J, Z, E$ subproblems will then be calculated for $t_k$ iterations. Therefore, based on the number of iterations, the combined cost is $O(t(n^r + m^r + d))$, such that $t$ is a representation of the number of iterations. Thus, when $n \geq m$, the cost’s upper bound would be $O(t n^r)$. Accordingly, the overall computational cost of the proposed method is $O(ktn^r)$.

4. Experiments

In this section, our proposed RSS-LRR method’s effectiveness is evaluated by comparing it with similar methods such as LRR [5], latent LRR [15], LRRRC, GLRR [8], and GODEC+ [41] in background modeling from video [42] and image denoising [43].

4.1. Background Modeling from Surveillance Video

4.1.1. Experimental Settings. This experiment is performed using surveillance video with various illumination settings. It is composed of a chain of 200 grayscale frames of $32 \times 32$ dimensions. Thus, each algorithm’s effectiveness is evaluated using precision, recall, and F-score metrics, and their parameters are tuned according to the corresponding literature. Precisely, background modeling [44] is measured by manually quoting out the activities. In this experiment, 50% of frames are randomly selected as the training set, while the remaining are treated as the testing set.

4.1.2. Experimental Results. In Figure 2, we show each algorithm’s background recovery and activity segmentation performance. Additionally, it can be observed from Table 1 that RSS-LRR outperforms other methods in activity segmentation as it can better generalize to the testing frames. This result further substantiates our sample scaling factor approach’s effectiveness such that a more reliable low-rank object is obtained than that from the compared methods.

4.2. Image Clustering with Varying Levels of Noise

4.2.1. Experimental Settings. Here, several experiments are performed on three well-known image datasets, namely, ORL, CMU PIE, and COIL20, to evaluate the effectiveness of the proposed algorithm on image clustering. Each algorithm’s parameters are tuned according to the corresponding literature using the grid search strategy. We give a brief description of each one of these datasets as follows:

- **ORL**: it contains face images of ten individuals, with each of them contributing forty distinct images under various conditions such as facial details and different facial expressions.
- **CMU PIE**: it is a face image repository with images of sixty-eight individuals with different settings. It includes thirteen different poses, four different expressions, and forty-two different illuminations.
- **COIL20**: it is an object image dataset consisting of twenty separate objects. Each object contributes 72 grayscale images, amounting to a total of 1440 images.

For each dataset, the images are resized to $32 \times 32$ dimensions in our experiments. As illustrated in Figure 3,
these datasets are then corrupted with 5%, 10%, 15%, and 20% random pixel noise to demonstrate each algorithm’s robustness to noise. Thus, a spectral clustering algorithm is applied to the similarity matrix of each algorithm to obtain the clustering results with ten multiple tries to ensure fairness [45].

4.2.2. Experimental Results. Tables 2–4 display the clustering results of each algorithm concerning the accuracy evaluation metric on ORL, CMU PIE, and COIL20, respectively. Therefore, it is obvious to see that the accuracy of our proposed RSS-LRR method consistently beats those of the compared methods in all three datasets. For example, on the ORL dataset, the accuracy of our proposed method is about 1% higher than that of its closest competitor to clean data. Then, gradually increasing the noise level, one may notice that all the algorithms had reduced performance. However, our proposed method shows more robustness than the other methods, especially with 20% noise damage. Let us take, for instance, the clustering result of LRR, which is the best robustness of RSS-LRR on image recognition under different levels of contiguous occlusions in images [19], we randomly add 6 × 6 and 8 × 8 46 block occlusions to each dataset, as illustrated in Figure 5. Therefore, 50% of samples in each dataset are selected as the training set and the rest as the testing set. Besides, we compare our proposed method’s performance with that of similar ones: LRR, LRRLC, latent LRR (LLRR), GLRR, and GODEC+, by adopting relevant experimental settings in Section 4.2. Thus, each algorithm’s classification accuracy is evaluated using the K nearest neighbor (KNN) classifier.

4.3. Image Recognition with Contiguous Occlusion

4.3.1. Experimental Settings. In order to evaluate the robustness of RSS-LRR on image recognition under different levels of contiguous occlusions in images [19], we randomly add 6 × 6 and 8 × 8 46 block occlusions to each dataset, as illustrated in Figure 5. Therefore, 50% of samples in each dataset are selected as the training set and the rest as the testing set. Besides, we compare our proposed method’s performance with that of similar ones: LRR, LRRLC, latent LRR (LLRR), GLRR, and GODEC+, by adopting relevant experimental settings in Section 4.2. Thus, each algorithm’s classification accuracy is evaluated using the K nearest neighbor (KNN) classifier.

4.3.2. Experimental Results. In Table 5, the average classification accuracies are obtained on the ORL dataset in two levels of contiguous occlusions. From Table 5, we can see that the accuracy of RSS-LRR is only slightly higher than that
Figure 2: Results obtained by all algorithms on surveillance video.

Table 1: Activity results of all algorithms on surveillance video.

<table>
<thead>
<tr>
<th></th>
<th>LRR</th>
<th>GLRR</th>
<th>GODEC+</th>
<th>RSS-LRR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>Test</td>
<td>Train</td>
<td>Test</td>
</tr>
<tr>
<td>Precision</td>
<td>0.9153</td>
<td>0.4872</td>
<td>0.9204</td>
<td>0.8859</td>
</tr>
<tr>
<td>Recall</td>
<td>0.7897</td>
<td>0.7358</td>
<td>0.7937</td>
<td>0.7478</td>
</tr>
<tr>
<td>F-score</td>
<td>0.8478</td>
<td>0.5862</td>
<td>0.8523</td>
<td>0.8110</td>
</tr>
</tbody>
</table>
of GODEC+. On the contrary, RSS-LRR’s accuracy is significantly better than that of LRR, LLRR, and LRR by about 5%, 4%, and 12%, respectively, under $6 \times 6$ occlusion. For example, the classification accuracy of RSS-LRR is 0.6015 and 0.4891 in LRR. This indicates that our proposed RSS-LRR can effectively obtain a good low-rank representation than existing methods.

From Table 6, we can see that the accuracy of RSS-LRR is better than that of LRR by about 6%. Under $6 \times 6$ occlusion, the accuracy of RSS-LRR is better than that of GODEC+ and GLRR by about 2%. Thus, our proposed RSS-LRR demonstrates approximately outstanding effectiveness on classification accuracies among all algorithms. In Table 7, the accuracy of RSS-LRR is higher than that of LLRC and LRR by about 3% and 6% under $6 \times 6$ occlusion. For example, the classification accuracy of GLRR is 0.6985 and 0.7130 in RSS-LRR, while the result of LRR is 0.6522 under $6 \times 6$ occlusion.

Figures 6(a)–6(f) illustrate the variations of classification accuracies with increasing feature dimensionalities in different occlusions on ORL, CMU PIE, and COIL20 datasets. From Figures 6(a) and 6(b), we can see that RSS-LRR achieves the highest accuracies among all algorithms. In Figures 6(c) and 6(d), we can see that the performance of RSS-LRR becomes better than that of others on the CMU PIE dataset when the feature dimensionality is over 70. From Figures 6(e) and 6(f), on the COIL20 dataset, we can see that the accuracies of RSS-LRR gradually show its superiority when the feature dimensionalities are more than 70.

From the above discussion, it can be seen that our proposed RSS-LRR shows better performance in classification accuracy compared with the other algorithms.

### 4.4 Experiments on Large-Scale Dataset

#### 4.4.1. Experimental Settings

In this experiment, RSS-LRR effectiveness is further evaluated on two larger datasets, namely, COIL100 and LFW. Also, the relevant settings from previous sections are adopted to perform the large-scale experiment. We give a brief description of each dataset as follows:

- **ORL Dataset**: The ORL dataset contains 400 images of 40 subjects, with 10 images per subject. The images are of size 92x112 pixels. The dataset is divided into training and test sets.

- **CMU PIE Dataset**: The CMU PIE dataset contains 411 images of 68 subjects, with 6 images per subject. The images are of size 144x144 pixels. The dataset is divided into training and test sets.

- **COIL20 Dataset**: The COIL20 dataset contains 20 images of 20 objects, with 100 images per object. The images are of size 100x100 pixels. The dataset is divided into training and test sets.
Figure 4: Clustering variation of different algorithms with varying noise levels on (a) ORL, (b) PIE, and (c) COIL20 datasets.

Figure 5: Example of the original image and corrupted one with different occlusion sizes on the COIL20 dataset.

Table 5: Classification accuracies on ORL under $6 \times 6$ and $8 \times 8$ occlusions.

<table>
<thead>
<tr>
<th></th>
<th>LRR</th>
<th>LRRLC</th>
<th>LLRR</th>
<th>GLRR</th>
<th>GODEC+</th>
<th>RSS-LRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$6 \times 6$</td>
<td>0.4891</td>
<td>0.5488</td>
<td>0.5534</td>
<td>0.5721</td>
<td>0.5953</td>
<td>0.6015</td>
</tr>
<tr>
<td>$8 \times 8$</td>
<td>0.3899</td>
<td>0.4046</td>
<td>0.4337</td>
<td>0.4473</td>
<td>0.4620</td>
<td>0.4762</td>
</tr>
</tbody>
</table>

Table 6: Classification accuracies on CMU PIE under $6 \times 6$ and $8 \times 8$ occlusions.

<table>
<thead>
<tr>
<th></th>
<th>LRR</th>
<th>LRRLC</th>
<th>LLRR</th>
<th>GLRR</th>
<th>GODEC+</th>
<th>RSS-LRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$6 \times 6$</td>
<td>0.7835</td>
<td>0.7980</td>
<td>0.8054</td>
<td>0.8247</td>
<td>0.8249</td>
<td>0.8415</td>
</tr>
<tr>
<td>$8 \times 8$</td>
<td>0.7589</td>
<td>0.7773</td>
<td>0.7821</td>
<td>0.8081</td>
<td>0.7919</td>
<td>0.8112</td>
</tr>
</tbody>
</table>
Figure 6: Variations of classification accuracies with increasing feature dimensionalities of $6 \times 6$ and $8 \times 8$ occlusions on (a, b) ORL, (c, d) CMU PIE, and (e, f) COIL20 datasets.
COIL100: it has 7200 images of 100 objects, which amount to 72 images for each object with each image taken at pose intervals of 5 degrees.

LFW: the Labeled Faces in the Wild (LFW) dataset originally contains more than 13000 face images, mainly from Internet sources. However, 2484 face images were extracted from 38 classes in our experiments due to fewer samples in some categories. Each image was resized to 64×64 pixels, yielding 4096 features per image.

4.4.2. Experimental Results. From the classification results in Tables 8 and 9, it can be noticed that RSS-LRR performance is consistently better than that of the compared methods. For instance, while RSS-LRR’s performance of 0.6052 on the COIL100 dataset corrupted with 6×6 block occlusion (Table 8) is slightly better than that of the second-best GODEC+ by over 1%, it is far better by over 2% under 8×8 block occlusion. Similar results are obtained on the LFW dataset (Table 9), where the proposed method’s performance is also better than that of the GODEC+ method, which follows closely, except that more margin of over 4% is obtained under 6×6 block occlusion.

From the clustering results displayed in Tables 10 and 11, the following can be observed:

(i) Although all methods obtained comparative performances on both datasets, their accuracies degrade significantly by increasing noise. This, however, is not unexpected because more corruption levels would mean that more discriminative data features are destroyed, making it difficult to accurately group similar data samples in the same cluster.

(ii) Furthermore, while the relatively newer method GODEC+ shows more robustness to noise than the older methods, its clustering accuracy on clean data is slightly lower than that of the other methods on the LFW dataset, perhaps due to class imbalance in this dataset.

(iii) Overall, RSS-LRR’s robustness to noise grows stronger with the increase in noise level. For example, its performance on COIL100 under 0% noise is merely 1% better than that of its closest competitor GODEC+, but it is over 2% better under 20% noise. The same can be said on the LFW dataset, where RSS-LRR’s clustering accuracy is only about 2% better than that of LRR on clean data, whereas it is more than 4% better than that of GODEC+, which is the closest result under the 20% noise level.

Additionally, in order to demonstrate more novelty, RSS-LRR’s clustering and classification performances on large-scale datasets are further compared with those of two more recent state-of-the-art (SOTA) methods, namely, nonnegative sparse discriminative low-rank representation (NSDLRR) [47] and low-rank and collaborative representations for hyperspectral anomaly detection (LRCRD) [48]. From the classification results displayed in Figure 7, it can be observed that all three methods obtain correlative results on both datasets. However, RSS-LRR shows more robustness, especially on the LFW dataset. Similarly, the clustering results are shown in Figure 8, and the results are also close with those of the proposed method displaying the best overall performance.
4.5. Convergence Study. Based on established standards of inexact augmented Lagrange multiplier (IALM) optimization strategy in several pieces of literature [5, 49–51], the objective function values of the $J$, $Z$, and $E$ subproblems were expected to decrease monotonically in each iteration until convergence. Besides, the $Z$ subproblem is known to have a closed-form solution [51] since it has a convex function. Also, the convergence of the $J$ and $E$ subproblems is confirmed by references [36, 52–54], respectively. Thus, a slow convergence is supposed when everything is put
together. Luckily, as shown in Figure 9, the proposed algorithm has a strong convergence property, as it converges within 150 iterations on COIL100 and ORL datasets.

5. Conclusion

In this paper, we propose a recursive sample scaling low-rank representation method named “RSS-LRR.” Different from the existing methods, each data sample’s importance is estimated by introducing a cosine scaling factor. This scaling factor is used to extract each sample’s relationship with the low-rank representation’s principal components in the feature space. Thus, our proposed model can effectively reduce the noise effect by iteratively reducing the importance of noisy samples in learning the robust low-rank matrix. Several experimental results on well-known benchmark datasets demonstrate that RSS-LRR performs better in clustering and classification tasks than the state-of-the-art methods. It includes various experiments conducted on gross corrupted data. Therefore, we will extend RSS-LRR’s idea to multiview data in future work.

Data Availability

The datasets used in this study are open benchmark datasets that are allowed for use in research. The following is a description and links to each one of them. Each can be accessed using the corresponding link. ORL contains face images of ten individuals, with each of them contributing forty distinct images under various conditions such as facial details and different facial expressions: http://cam-orl.co.uk/facedatabase.html. CMU PIE is a face image repository with images of sixty-eight individuals with different settings. It includes thirteen different poses, four different expressions, and forty-two different illuminations: https://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html. COIL20 is an object image dataset consisting of 20 separate objects. Each object contributes 72 grayscale images, amounting to a total of 1440 images: https://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php. COIL100 has 7200 images of 100 objects, which amount to 72 images for each object with each image taken at pose intervals of 5 degrees: https://www.kaggle.com/jessicali9530/coil100. The Labeled Faces in the Wild (LFW) dataset originally contains more than 13000 face images, mainly from Internet sources. However, 2484 face images were extracted from 38 classes in our experiments due to fewer samples in some categories. Each image was resized to 64 × 64 pixels, yielding 4096 features per image: http://vis-www.cs.umass.edu/lfw/.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Authors’ Contributions

Wenyun Gao and Stanley Abhadiomhen conceptualized the research idea and reviewed and edited the paper. Xiaoyun Li and Sheng Dai were involved in data investigation. Sheng Dai ran the software. Stanley Abhadiomhen wrote the original draft. Wenyun Gao, Xinghui Yin, and Stanley Abhadiomhen supervised the project. Wenyun Gao was involved in funding acquisition.

Acknowledgments

This research was funded in part by the National Key Research and Development Program of China, Grant no. 2020YFC1511800.

References


