Proceedings of SPIE, Commercial + Scientific Sensing and Imaging Symp., Compressive Sensing Conf., Orlando, FL, April 2018, vol. 10211

Conformity Evaluation of Data Samples by L_1-Norm Principal-Component Analysis

Ying Liu[a] and Dimitris A. Pados[b]

[a]Department of Electrical Engineering, The State University of New York at Buffalo, Buffalo, NY 14260

[b]Department of Computer and Electrical Engineering and Computer Science & I-SENSE, Florida Atlantic University, Boca Raton, FL 33431

ABSTRACT

We describe an iterative procedure for soft characterization of outlier data in any given data set. In each iteration, data compliance to nominal data behavior is measured according to current L_1-norm principal-component subspace representations of the data set. Successively refined L_1-norm subspace data set representations lead to successively refined outlier data characterization. The effectiveness of the proposed theoretical scheme is experimentally studied and the results show significantly improved performance compared to L_2-PCA schemes, standard L_1-PCA, and state-of-the-art robust PCA methods.

Keywords: Conformity evaluation, dimensionality reduction, eigenvector decomposition, feature extraction, L_1-norm, outliers, principal-component analysis.

1. INTRODUCTION

In many real-world applications such as video surveillance, processing of remotely sensed images, and network traffic analysis, the dimensionality of data are usually very high, and direct handling of the high-dimensional data is computationally expensive. Fortunately, many high-dimensional data have intrinsic low-rank structures, that can be revealed through dimensionality reduction schemes. Over several decades, principal-component analysis (PCA) has been one of the most widely applied dimensionality reduction methods due to its simplicity and effectiveness. Given a dataset, PCA finds a set of projection vectors (the so-called "principal components") to maximize the variance of the projected data points, and the structure of the original data could be effectively preserved under the projection. The nominal compliance of each data sample can then be inferred by leveraging the principal components, and be used to perform tasks such as moving object extraction and outlier detection.

Nevertheless, the conventional PCA based on the L_2-norm (L_2-PCA) is prone to the presence of outliers because the effect of the outliers with a large norm is exaggerated by the use of the L_2-norm. To alleviate this problem, there has been a growing interest in robust PCA methods, such as L_1-error minimization[1]-,[4] non-negative matrix factorization via Manhattan distance minimization (MahNMF),[5] and the low-rank and sparse decomposition approach[6]-.[8] Another line of research has focused on robust subspaces calculation by explicit L_1 projection maximization[9]-.[13] The resulting principal components are called L_1 principal components. In particular, the exact calculation of L_1 principal components was developed[11] for the first time in the literature. Later, suboptimal algorithms were developed[12,13] for fast computation of the L_1 principal components. The L_1-PCA method has been successfully applied to a wide range of research fields such as direction of arrival (DoA) estimation[14] and robust face recognition[15]-.[17] Most recently, compressed-sensed-domain L_1-PCA methods were developed for low-rank background scene and sparse foreground moving objects extraction from compressed-sensed surveillance video sequences[18]-.[20]

In spite of their robustness, for a given data set with potential outliers, the existing L_1-PCA methods[9]-[20] merely calculate a "one-shot" L_1 subspace, which can still be away from the true nominal signal subspace of interest when the data sets are severely contaminated. In this paper, we propose an iterative approach that

Ying Liu: E-mail: yingliuub@gmail.com
Dimitris A. Pados: E-mail: dpados@fau.edu

iteratively generates a sequence of improved L_1 subspaces. In each iteration, nominal compliance of each sample is inferred by its position relative to the L_1 subspace calculated in the previous iteration and translated to a "weight". Samples with higher weights tend to be nominal samples and samples with lower weights are more likely to be the outliers. Weighted L_1-PCA calculation is then carried out in which the contribution of outlying samples in the data set is suppressed resulting in an improved L_1-subspace. The sample weights converge as the iteration number increases and the iterative algorithm terminates when the weights in the current and previous iterations are deemed close enough.

The remainder of this paper is organized as follows. In Section 2, we develop the iterative L_1-PCA algorithm, provide convergence analysis, and propose a stopping criterion for practical implementation of the algorithm. In Section 3, the effectiveness of the proposed algorithm is demonstrated through two experiments: (i) moving objects extraction in video surveillance, and (ii) unsupervised outlier detection of the UCI machine learning data sets.[21] Finally, a few conclusions are drawn in Section 4.

2. PROPOSED ITERATIVE L_1 PRINCIPAL-COMPONENT ANALYSIS

2.1 Background of L_1-PCA

We consider a $D \times N$ $(N < D)$ data matrix that has N data points $\mathbf{x}_i \in \mathbb{R}^D$, $i = 1, 2, \cdots, N$, that is

$$\mathbf{X} = [\mathbf{x}_1 \ \ \mathbf{x}_2 \ \ \cdots \mathbf{x}_N]. \tag{1}$$

The conventional L_2-PCA seeks r orthonormal projection vectors to describe the rank-r subspace of the data matrix $\mathbf{X}$ by

$$\mathcal{P}^{L_2} : \mathbf{P}_{L_2} = \arg \max_{\substack{\mathbf{P} \in \mathbb{R}^{D \times r} \\ \mathbf{P}^T \mathbf{P} = \mathbf{I}_r}} \|\mathbf{X}^{\mathrm{T}} \mathbf{P}\|_2, \tag{2}$$

and the solution is given by the r dominant-singular-value left singular vectors of the original data matrix $\mathbf{X}$.

Nevertheless, the L_2-norm metric in Problem $\mathcal{P}^{L_2}$ is sensitive to outlying samples that are numerically distant from the nominal data, leading to inaccurate low-rank subspace calculation and data conformity evaluation. Motivated by this observed drawback of L_2 subspace signal processing, subspace-decomposition approaches that are based on the L_1 norm were proposed for robust low-rank subspace computation. Replacing the L_2-norm in Problem $\mathcal{P}^{L_2}$ by L_1-norm, the so-called L_1-PCA calculates principal components in the form of

$$\mathcal{P}^{L_1} : \mathbf{P}_{L_1} = \arg \max_{\substack{\mathbf{P} \in \mathbb{R}^{D \times r} \\ \mathbf{P}^T \mathbf{P} = \mathbf{I}_r}} \|\mathbf{X}^{\mathrm{T}} \mathbf{P}\|_1. \tag{3}$$

Since projecting all the data to a subspace deviated by the outliers is less likely to generate a larger projection L_1-norm than projecting the data to the correct low-rank subspace, $\mathbf{P}_{L_1}$ in (3) is likely to be closer to the true nominal rank-r subspace than L_2-PCA. The r columns of $\mathbf{P}_{L_1}$ in (3) are the so-called r L_1 principal components that describe the rank-r subspace in which $\mathbf{X}$ lies. It was shown that the exact calculation of the L_1 principal components in Problem $\mathcal{P}^{L_1}$ can be recast as a combinatorial problem[11] and solved in exponential time. Besides, a polynomial-time algorithm is developed[11] for any fixed data dimension D, and bit-flipping based sub-optimal algorithms were also proposed for the fast calculation of the L_1 principal components,[12].[13]

The regular L_1-PCA problem in (3) seeks a rank-r subspace from the data matrix $\mathbf{X} \in \mathbb{R}^{D \times N}$ by one-shot calculation. Although the adopted L_1-norm maximization is less affected by outliers compared to L_2-norm maximization in L_2-PCA in (2), the produced L_1 subspace $\mathbf{P}_{L_1}$ can still be away from the true nominal signal low-rank subspace. In the following subsection, we propose an iterative method that generates a sequence of improved L_1 subspaces for the same data matrix $\mathbf{X}$.

2.2 Iterative Data Conformity Evaluation

We consider the calculation of r principal components $\mathbf{P}_{L_1} \in \mathbb{R}^{D\times r}$, $D > r > 1$. Initially, the direct L_1 subspace is computed via (3) and denoted by $\mathbf{P}_{L_1}^{(0)}$. Next, the distance of each sample $\mathbf{x}_n$ from subspace $\mathbf{P}_{L_1}^{(0)}$ is defined as the L_2 error between $\mathbf{x}_n$ and its rank-r surrogate

$$d_n^{(1)} = \|\mathbf{x}_n - \mathbf{P}_{L_1}^{(0)} \mathbf{P}_{L_1}^{(0)^{\mathrm{T}}} \mathbf{x}_n\|_2, \quad n = 1, ..., N. \tag{4}$$

We expect large $d_n^{(1)}$ if $\mathbf{x}_n$ is an "outlier" and small $d_n^{(1)}$ if $\mathbf{x}_n$ is a nominal sample. Therefore, the conformity of each sample can be measured as the reciprocal of its L_2 distance from the subspace, i.e.,

$$w_n^{(1)} = (d_n^{(1)})^{-1}, \quad n = 1, ..., N, \tag{5}$$

followed by normalization

$$\widetilde{w}_n^{(1)} = \frac{w_n^{(1)}}{\sum\limits_{n=1}^{N} w_n^{(1)}}, \quad n = 1, ..., N. \tag{6}$$

When computing the L_1 subspace, data samples with larger conformity should contribute more and samples with smaller conformity should be suppressed such that the resulting calculated L_1 subspace is more accurate. In this direction, we propose that each data sample $\mathbf{x}_n$ is scaled by its conformity $\widetilde{w}_n^{(1)}$. We then form a weight matrix with the conformity values,

$$\widetilde{\mathbf{W}}^{(1)} = \begin{bmatrix} \widetilde{w}_1^{(1)} & 0 & 0 & \cdots \\ 0 & \widetilde{w}_2^{(1)} & 0 & \cdots \\ & \vdots & & \\ 0 & 0 & \cdots & \widetilde{w}_N^{(1)} \end{bmatrix}, \tag{7}$$

and update the L_1 subspace as

$$\mathbf{P}_{L_1}^{(1)} = \arg \max_{\substack{P \in \mathbb{R}^{D\times r} \\ \mathbf{P}^T\mathbf{P} = \mathbf{I}_r}} \|(\mathbf{X}\widetilde{\mathbf{W}}^{(1)})^{\mathrm{T}} \mathbf{P}\|_1. \tag{8}$$

The above weighted L_1-subspace calculation can be performed iteratively. In the kth iteration, the weights are computed using the L_1-subspace $\mathbf{P}_{L_1}^{(k-1)}$ computed at the $(k-1)$th iteration, i.e.

$$d_n^{(k)} = \|\mathbf{x}_n - \mathbf{P}_{L_1}^{(k-1)} \mathbf{P}_{L_1}^{(k-1)^{\mathrm{T}}} \mathbf{x}_n\|_2, 1 \le n \le N, \tag{9}$$

$$w_n^{(k)} = (d_n^{(k)})^{-1}, \tag{10}$$

$$\widetilde{w}_n^{(k)} = \frac{w_n^{(k)}}{\sum\limits_{n=1}^{N} w_n^{(k)}}, \tag{11}$$

$$\widetilde{\mathbf{W}}^{(k)} = \begin{bmatrix} \widetilde{w}_1^{(k)} & 0 & 0 & \cdots \\ 0 & \widetilde{w}_2^{(k)} & 0 & \cdots \\ & \vdots & & \\ 0 & 0 & \cdots & \widetilde{w}_N^{(k)} \end{bmatrix}. \tag{12}$$

Subsequently, the L_1-subspace at the kth iteration is updated to

$$\mathbf{P}_{L_1}^{(k)} = \arg \max_{\substack{P \in \mathbb{R}^{D\times r} \\ \mathbf{P}^T\mathbf{P} = \mathbf{I}_r}} \|(\mathbf{X}\widetilde{\mathbf{W}}^{(k)})^{\mathrm{T}} \mathbf{P}\|_1. \tag{13}$$

Algorithm 1 Iterative Re-weighted L_1-norm Principal-Components Calculation (rank-r)

Input: $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_N] \in \mathbb{R}^{D\times N}$, r, β, ϵ.
Initialization: Find $\mathbf{P}_{L_1}^{(0)} \in \mathbb{R}^{D\times r}$ by (3).
Iterative L_1 Subspace Calculation:
for $k = 1, 2, \cdots$, **do**

1. Compute the L_2 error between $\mathbf{x}_n$ and its rank-r surrogate $d_n^{(k)} = \|\mathbf{x}_n - \mathbf{P}_{L_1}^{(k-1)}\mathbf{P}_{L_1}^{(k-1)^T}\mathbf{x}_n\|_2$.
2. Define $u_n^{(k)} = (d_n^{(k)})^{-1}$.
3. If $k = 1$, $w_n^{(k)} \leftarrow u_n^{(k)}$; if $k > 1$, update $w_n^{(k)}$ by (15).
4. Check stopping criterion: if $\|\mathbf{w}^{(k)} - \mathbf{w}^{(k-1)}\|_2 < \epsilon$, exit.
5. Compute $\widetilde{w}_n^{(k)}$ and $\widetilde{\mathbf{W}}^{(k)}$ by (11) and (12).
6. Update the L_1 subspace $\mathbf{P}_{L_1}^{(k)}$ by (13).

end for
Output: Rank-r L_1 subspace sequence $\mathbf{P}_{L_1}^{(k)} \in \mathbb{R}^{D\times r}$, $k = 1, 2, \cdots$.

2.3 A Convergent Weight Sequence

Inspired by,[22] we modify the weight update formula as follows to guarantee a convergent weight sequence for practical algorithmic implementation. In the kth iteration, we first compute the ℓ_2 error (distance) for each sample as in (9). Then we define

$$u_n^{(k)} = (d_n^{(k)})^{-1} \tag{14}$$

and update the weight $w_n^{(k)}$ based on $u_n^{(k)}$. If $k = 1$, let $w_n^{(k)} = u_n^{(k)}$. For $k > 1$, $w_n^{(k)}$ is updated by

$$w_n^{(k)} = \begin{cases} w_n^{(k-1)}(1-\beta^{k-1}), \text{ if } u_n^{(k)} < w_n^{(k-1)}(1-\beta^{k-1}), \\ u_n^{(k)}, \text{ if } w_n^{(k-1)}(1-\beta^{k-1}) \leq u_n^{(k)} \leq w_n^{(k-1)}(1+\beta^{k-1}), \\ w_n^{(k-1)}(1+\beta^{k-1}), \text{ if } u_n^{(k)} > w_n^{(k-1)}(1+\beta^{k-1}) \end{cases} \tag{15}$$

where $0 < \beta < 1$ is a pre-defined parameter. Intuitively, we avoid updating the weights too aggressively by restricting the new weight $w_n^{(k)}$ to be within a small neighborhood of the weight in the previous iteration $w_n^{(k-1)}$. The size of the neighborhood depends on β. Subsequently, $w_n^{(k)}$ is normalized as in (11), followed by weight matrix construction in (12). The convergence of the weight sequence can be verified by

$$\lim_{k\to\infty} \beta^{k-1} = 0, \tag{16}$$

$$\lim_{k\to\infty} (w_n^{(k)} - w_n^{(k-1)}) = 0, \tag{17}$$

$$\lim_{k\to\infty} (\widetilde{w}_n^{(k)} - \widetilde{w}_n^{(k-1)}) = 0. \tag{18}$$

2.4 Stopping Criterion

In implementing the proposed iterative algorithm, we exit the algorithm when the difference between the weight vectors at the kth and $(k-1)$th iteration is smaller than a predefined threshold ϵ, that is,

$$\|\mathbf{w}^{(k)} - \mathbf{w}^{(k-1)}\|_2 < \epsilon, \tag{19}$$

where $\mathbf{w}^{(k)} = [w_1^{(k)}, w_2^{(k)}, \cdots, w_N^{(k)}]^T$ and $\mathbf{w}^{(k-1)} = [w_1^{(k-1)}, w_2^{(k-1)}, \cdots, w_N^{(k-1)}]^T$.

The complete algorithm is summarized in Algorithm 1.

3. EXPERIMENTAL STUDIES

In this section, we assess the effectiveness of the proposed iterative re-weighted L_1-PCA (IRW L_1-PCA) algorithm through two experimental studies: (i) moving object extraction in video surveillance, and (ii) unsupervised outlier detection of the UCI[21] data sets.

3.1 Moving Object Extraction

Consider a sequence of surveillance video frames $\mathbf{X}_t \in \mathbb{R}^{m\times n}$ with frame resolution of $m \times n$ pixels and time index $t = 1, ..., N$. A typical surveillance video sequence is consisted of a static background scene that can be modeled as a low-rank component, and sparse foreground moving objects superimposed on the static background scene that are regarded as the outliers. For security monitoring, the objective is to extract the moving objects. In our experiment, we perform block-by-block IRW L_1-PCA for low-rank background modeling and foreground extraction. We divide each frame $\mathbf{X}_t$ into J blocks $\mathbf{X}_t^j \in \mathbb{R}^{m_b\times n_b}$, $j = 1, ..., J$. We let $\mathbf{x}_t^j \in \mathbb{R}^D$, $D = m_b n_b$, represent vectorization of $\mathbf{X}_t^j$ via column concatenation. For each sequence of co-located blocks, $\mathbf{x}_t^j$, $t = 1, ..., N$, we model the static background scene as a low-rank component $\mathbf{z}_t^j$ and the foreground moving objects as an outlying component $\mathbf{s}_t^j$. That is,

$$\mathbf{x}_t^j = \mathbf{z}_t^j + \mathbf{s}_t^j, \quad t = 1, ..., N, \tag{20}$$

and $\mathbf{s}_t^j$ appears only a few times in the block sequence. In matrix-form representation of the jth block across N frames, $\mathbf{X}^j \triangleq [\mathbf{x}_1^j, ..., \mathbf{x}_N^j] \in \mathbb{R}^{D\times N}$ and

$$\mathbf{X}^j = \mathbf{Z}^j + \mathbf{S}^j. \tag{21}$$

To extract the low-rank background information, we carry out IRW L_1-PCA on $\mathbf{X}^j$ and obtain the rank-2 L_1 subspace $\mathbf{P}_{L_1}^j$ at convergence. Afterwards, the background blocks can be approximated by $\widehat{\mathbf{Z}}^j = \mathbf{P}_{L_1}^j \mathbf{P}_{L_1}^{j\,T} \mathbf{X}^j$ and the foreground blocks can be extracted as $\widehat{\mathbf{S}}^j = \mathbf{X}^j - \widehat{\mathbf{Z}}^j$, $j = 1, ..., J$.

We test the method on the *Daniel_light* video sequence with 80 frames, each of 120×160 pixels. We process $N = 8$ successive frames at a time. To mitigate the "blockiness" artifact, we divide each frame into $J = 1296$ overlapping blocks of size 15×20, and apply the proposed IRW L_1-PCA method independently to each group of co-located blocks across 8 frames. The final background and foreground scenes are obtained by averaging the extracted background pixels (as well as the foreground pixels) for which multiple results are available.

Fig. 1 displays the background and foreground extracted at multiple distinct time slots $t = 5, 15, 35, 47$ by the proposed IRW L_1-PCA, DECOLOR,[7] MahNMF,[5] and the regular L_1-PCA[11] methods. In addition to low-rank and sparse decomposition, the DECOLOR[7] approach also uses Markov random-field (MRF) modeling to improve the accuracy of detecting contiguous outliers. The results show that the foreground scenes extracted by DECOLOR are somewhat affected by the desks and computers that belong to the background scene, and the regular L_1-PCA extracted foreground scenes suffer from severe "ghosts" that are formed when an object initially in the background begins to move. Although MahNMF performs as well as the proposed algorithm for $t = 15$ and 35, it incorrectly recovers the background scene for $t = 5$ and 47 and results in problematic foreground extraction. In contrast, the proposed IRW L_1-PCA offers much clearer foreground scenes than the other three methods consistently across multiple frames.

3.2 Unsupervised Outlier Detection of UCI Data Sets

The problem of unsupervised outlier detection is to determine data samples that significantly deviate from the majority of data samples. It has been widely studied due to its numerous applications in intrusion detection, land resource exploration, criminological investigation, and medical diagnostics. Different outlier detection paradigms were proposed in the literature, such as density-based methods[23,24] , feature bagging,[25] and subspace methods[26,27] . Among these approaches, the classic local outlier factor (LOF) algorithm[23] compares the density of each data sample with that of its local neighbors, considering that an outlier has low density compared to its local neighborhood. The state-of-the-art LODES algorithm[27] combines spectral embedding with LOF. It embeds data samples in a low-dimensional subspace where similar samples are pulled close together. Subsequently, LOF algorithm is applied to the embedded data samples such that outliers can be more easily distinguished from regular objects.

In this section, we apply the proposed IRW L_1-PCA algorithm to detect the outliers in two real-world data sets taken from the UCI machine learning repository,[21] *Satimage* and *Glass*. The original Statlog (Landsat Satellite) data set is a multi-class classification data set generated from multi-spectral scanner image data. Each sample has $D = 36$ attributes and each class is one land type. In our study, class 2 (cotton crop) is down-sampled to 71 outliers, while all the other classes (red soil, grey soil, damp grey soil, soil with vegetation stubble, and

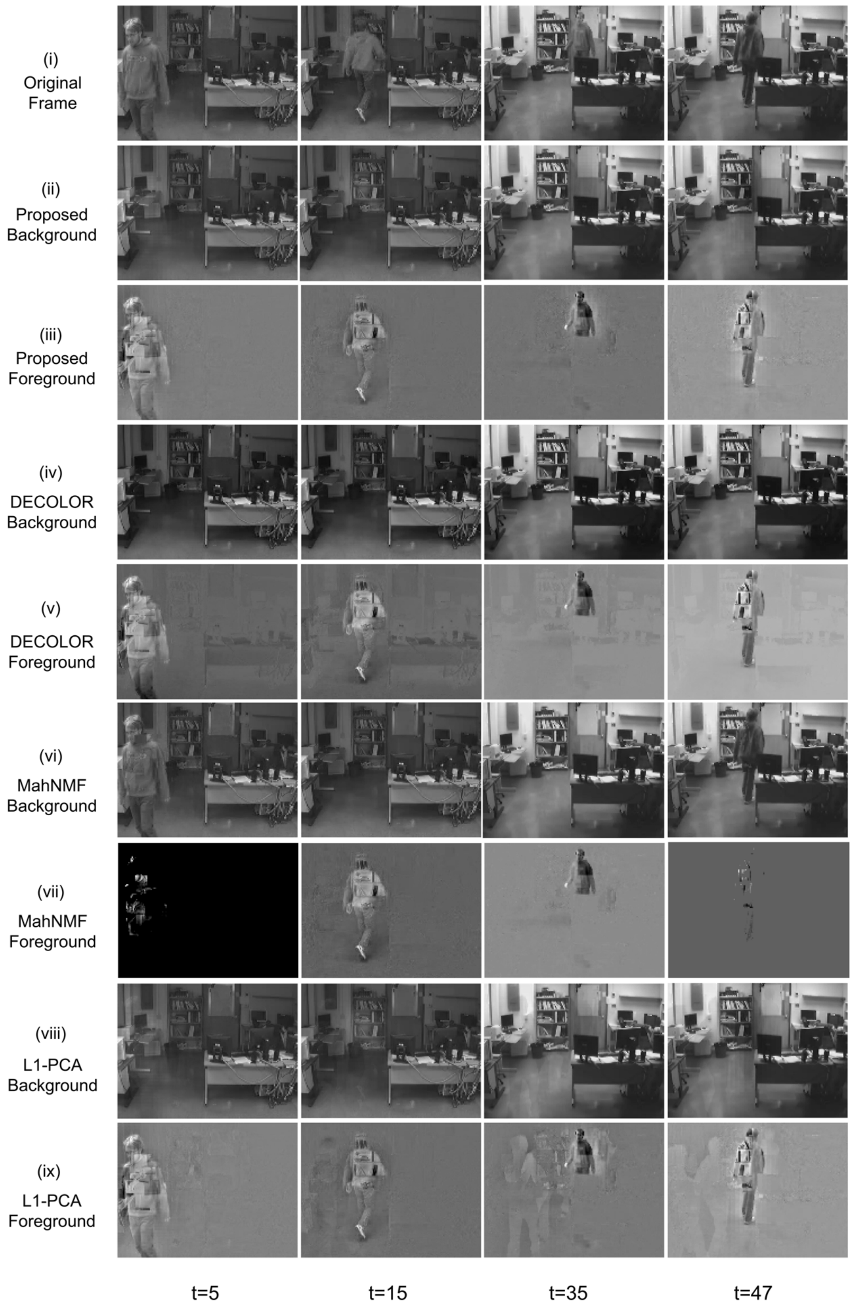
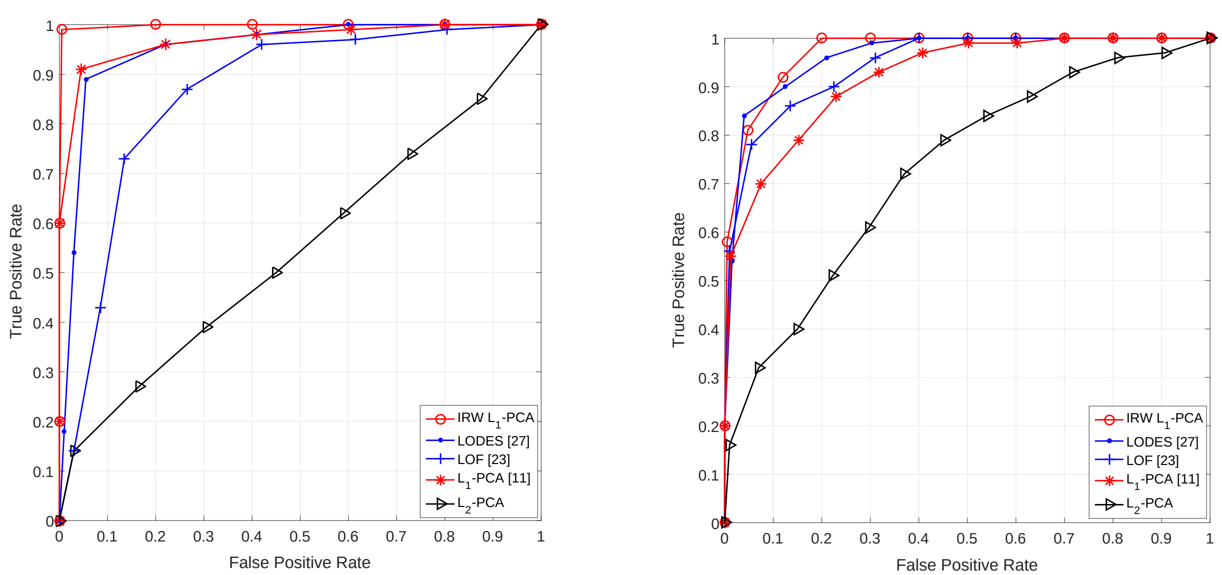


Figure 1: *Daniel_light* sequence: Original frame [row (i)] of time slot $t = 5, 15, 35$, and 47; proposed IRW L_1-PCA reconstructed background and moving objects [rows (ii) and (iii)]; DECOLOR[7] reconstructed background and moving objects [rows (iv) and (v)]; MahNMF[5] reconstructed background and moving objects [rows (vi) and (vii)]; regular L_1-PCA[11] reconstructed background and moving objects [rows (viii) and (ix)].

Table 1: AUC values of the ROC curves

	Proposed	LODES[27]	LOF[23]	L_1-PCA[11]	L_2-PCA
Satimage	0.998	0.951	0.861	0.968	0.545
Glass	0.971	0.963	0.946	0.920	0.730

very damp grey soil) are combined to form an inlier class. The modified data set is referred to as *Satimage*. Each time we took $N_o = 5$ random samples from the outlier class to construct the outlier data set $\mathbf{X}_o \in \mathbb{R}^{D\times N_o}$, and $N_i = 10$ random samples from the inlier class to construct the inlier data set $\mathbf{X}_i \in \mathbb{R}^{D\times N_i}$. The overall data set is formed as $\mathbf{X} = [\mathbf{X}_i \quad \mathbf{X_o}] \in \mathbb{R}^{(D=36)\times(N=15)}$. The proposed IRW L_1-PCA algorithm with $r = 2$ principal components is applied to $\mathbf{X}$ to obtain the L_1 subspace $\mathbf{P}_{L_1} \in \mathbb{R}^{(D=36)\times(r=2)}$ at convergence. Subsequently, the ℓ_2-norm distance from data sample $\mathbf{x}_n$ to the proposed L_1 subspace $\mathbf{P}_{L_1}$, that is, $\|\mathbf{x}_n - \mathbf{P}_{L_1}\mathbf{P}_{L_1}^T\mathbf{x}_n\|_2$ is calculated for each data sample as the outlier score, and the samples with highest scores are detected as outliers, i.e. samples that belong to the cotton crop land type. We carried out 20 independent experiments, and plot the receiver operating characteristic curves (ROC) of the proposed algorithm and four existing algorithms, LODES,[27] LOF,[23] L_1-PCA[11] and L_2-PCA. The ROC plots the true positive rate versus the false positive rate, which is a common measure to evaluate the performance of outlier detection methods. As shown in Fig. 2a, the proposed IRW L_1-PCA method demonstrates the ability to capture the outliers significantly earlier than all other algorithms. A summary measure of the detection accuracy is the area under the ROC curve (AUC) shown in table 1. Again we observe that the proposed method achieves the highest AUC value for the *Satimage* data set among all algorithms in comparison.



(a) (b)

Figure 2: The receiver operating characteristics (ROC) curves of (a): the *Satimage* data set, and (b) the *Glass* data set.

We carried out a similar study for the *Glass* data set. For this data set, we are interested in identifying tableware glasses (outlier class) from building window glasses (inlier class). The study of glass-type identification was motivated by criminological investigation. At the scene of the crime, the glass left can be used as evidence, if correctly identified. In our study, each time we took $N_o = 5$ random samples from the outlier class and $N_i = 20$ random samples from the inlier class to form a data set of size $(D = 9) \times (N = 25)$. We applied the proposed IRW L_1-PCA method with $r = 2$ L_1 principal components to the data set, and calculate the outlier scores for outlier detection with the same method used for the *Satimage* data set. The ROC curves are shown in Fig. 2b for the algorithms in comparison, and the corresponding AUC values are shown in table 1. Again we observe that the proposed algorithm achieved the maximum true positive rate 1 at the lowest false positive rate 0.2, while the LODES and LOF schemes achieve the maximum true positive rate 1 at a higher false positive rate 0.4.

4. CONCLUSION

In the presented work, an iterative re-weighted L_1 principal-component analysis algorithm is developed to measure the compliance of data samples to the nominal low-rank structure. In each iteration, the L_1 subspace computed in the previous iteration is utilized to infer the conformity of each data sample, and the resultant conformity value is used to weigh the data sample, followed by L_1-subspace update. The procedure generates successively refined L_1-norm subspaces, which present significantly enhanced performance in tasks such as moving object extraction and outlier detection, compared to the regular L_1-PCA and state-of-the-art robust PCA methods.

REFERENCES

1. Q. Ke and T. Kanade, "Robust subspace computation using L_1 norm," *Technical Report CMU-CS-03-172*, Carnegie Mellon Univ., http://ra.adm.cs.cmu.edu/anon/usr0/ftp/usr/anon/2003/CMU-CS-03-172.pdf, Aug. 2003.
2. Q. Ke and T. Kanade, "Robust norm factorization in the presence of outliers and missing data by alternative convex programming," in *Proc. IEEE Conf. Comput. Vision Pattern Recog. (CVPR)*, San Diego, CA, Jun. 2005, pp. 739-746.
3. A. Eriksson and A. v. d. Hengel, "Efficient computation of robust low-rank matrix approximations in the presence of missing data using the L_1 norm," in *Proc. IEEE Conf. Comput. Vision Pattern Recog. (CVPR)*, San Francisco, CA, Jun. 2010, pp. 771-778.
4. L. Yu, M. Zhang, and C. Ding, "An efficient algorithm for L_1-norm principal component analysis," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 1377-1380.
5. N. Guan, D. Tao, Z. Luo, and J. S. Taylor, "MahNMF: Manhattan non-negative matrix factorization," *arXiv:1207.3438*, July 2012.
6. E. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, pp. 1-37, May 2011.
7. X. Zhou, C. Yang, and W. Yu, "Moving object detection by detecting contiguous outliers in the low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 597-610, Mar. 2013.
8. T. Zhou and D. Tao, "GoDec: Randomized low-rank & sparse matrix decomposition in noisy case," in *Proc. Int. Conf. Machine Learning*, Bellevue, WA, 2011.
9. N. Kwak, "Principal component analysis based on L_1-norm maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, pp. 1672-1680, Sept. 2008.
10. F. Nie, H. Huang, C. Ding, D. Luo, and H. Wang, "Robust principal component analysis with non-greedy ℓ_1-norm maximization," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, Barcelona, Spain, July 2011, pp. 1433-1438.
11. P. P. Markopoulos, G. N. Karystinos, and D. A. Pados, "Optimal algorithms for L_1-subspace signal processing," *IEEE Trans. Signal Process.*, vol. 62, pp. 5046-5058, Oct. 2014.
12. S. Kundu, P. P. Markopoulos and D. A. Pados, "Fast computation of the L_1-principal component of real-valued data," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, Florence, Italy, May 2014, pp. 8028-8032.
13. P. P. Markopoulos, S. Kundu, S. Chamadia, and D. A. Pados, "Efficient L_1-norm principal-component analysis via bit flipping," *IEEE Trans. Signal Process.*, vol. 65, pp. 4252 - 4264, Aug. 2017.
14. P. P. Markopoulos, N. Tsagkarakis, D. A. Pados, and G. N. Karystinos, "Direction finding with L_1-norm subspaces," *SPIE Compressive Sensing Conference, Defense, Security, and Sensing (SPIE DSS 2014)*, Baltimore, MD, May 2014, vol. 9109, pp. 0J1-0J11.
15. M. Johnson and A. Savakis, "Fast L_1-eigenfaces for robust face recognition," in *Proc. IEEE Western NY Image and Signal Process. Workshop*, Rochester, NY, Nov. 2014, pp. 1-5.

16. F. Maritato, Y. Liu, D. A. Pados, and S. Colonnese, “Face recognition with L_1-norm subspaces,” in *Proc. SPIE Commercial + Scientific Sensing and Imaging*, Baltimore, MD, Apr. 2016.
17. F. Maritato, Y. Liu, S. Colonnese, and D. A. Pados, “Cloud-assisted individual L_1-PCA face recognition using wavelet-domain compressed images,” in *Proc. European Workshop on Visual Information Process. (EUVIP)*, Marseille, France, Oct. 2016.
18. Y. Liu and D. A. Pados, “Compressed-sensed-domain L_1-PCA video surveillance,” in *Proc. SPIE Defense, Security, and Sensing (DSS)*, Baltimore, MD, Apr. 2015.
19. Y. Liu and D. A. Pados, “Compressed-sensed-domain L_1-PCA video surveillance,” *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 351-363, Mar. 2016.
20. M. Pierantozzi, Y. Liu, D. A. Pados, and S. Colonnese, “Video background tracking and foreground extraction via L_1-subspace updates,” in *Proc. SPIE Commercial + Scientific Sensing and Imaging*, Baltimore, MD, Apr. 2016.
21. M. Lichman, (2013). UCI Machine Learning Repository, Irvine, CA: University of California, School of Information and Computer Science, [Online]. Available: http://archive.ics.uci.edu/ml.
22. Y. W. Park and D. Klabjan, “Iteratively reweighted least squares algorithms for L_1-Norm Principal Component Analysis,” in *Proc. Int. Conf. Data Mining (ICDM)*, Barcelona, Spain, Dec. 2016, pp. 430-438.
23. M. Breunig, H.-P. Kriegel, R. Ng, and J. Sander, “LOF: identifying density-based local outliers,” in *Proc. ACM SIGMOD Int. Conf. Management of Data*, Dallas, TX, May 2000, pp. 93-104.
24. H.-P. Kriegel, P. Kröger, E. Schubert, A. Zimek, “LoOP: Local outlier probabilities,” in *Proc. ACM Conf. Information and Knowledge Management (CIKM)*, Hong Kong, China, Nov. 2009, pp. 1649-1652.
25. A. Lazarevic and V. Kumar, “Feature bagging for outlier detection,” in *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery in Data Mining*, Chicago, IL, Aug. 2005, pp. 157-166.
26. F. Keller, E. Müller, and K. Böhm, “HiCS: High contrast subspaces for density-based outlier ranking,” in *Proc. IEEE Int. Conf. Data Engineering (ICDE)*, Washington D. C., Apr. 2012, pp. 1037-1048.
27. S. Sathe and C. Aggarwal, “LODES: Local density meets spectral outlier detection,” in *Proc. SIAM Int. Conf. Data Mining (ICDM)*, Barcelona, Spain, Dec. 2016, pp. 171-179.