Compressed-sensed-domain L_1-PCA Video Surveillance

Ying Liu and Dimitris A. Pados

Department of Electrical Engineering, State University of New York at Buffalo,
Buffalo, NY 14260

ABSTRACT

We consider the problem of foreground and background extraction from compressed-sensed (CS) surveillance video. We propose, for the first time in the literature, a principal component analysis (PCA) approach that computes the low-rank subspace of the background scene directly in the CS domain. Rather than computing the conventional L_2-norm-based principal components, which are simply the dominant left singular vectors of the CS measurement matrix, we compute the principal components under an L_1-norm maximization criterion. The background scene is then obtained by projecting the CS measurement vector onto the L_1 principal components followed by total-variation (TV) minimization image recovery. The proposed L_1-norm procedure directly carries out low-rank background representation without reconstructing the video sequence and, at the same time, exhibits significant robustness against outliers in CS measurements compared to L_2-norm PCA.

Keywords: Compressed sensing, convex optimization, feature extraction, L_1 principle component analysis, singular value decomposition, surveillance video, total-variation minimization.

1. INTRODUCTION

In video surveillance, video signals are captured by cameras and transmitted to a processing center where video streams are monitored and analyzed. With the advent of wireless multimedia sensor networks (WMSNs) in the last decade,[1] video surveillance through large-scale WMSNs has become a primary objective in research communities and industry. As a complement to existing wired video surveillance systems, wireless video surveillance enjoys advantages such as scalability and easy deployment. Wireless cameras in use can be miniature low-cost devices that capture, compress, and transmit video signals with low power consumption.

The recently introduced paradigm of compressive sensing (CS) acts as an enabler to WMSN technology. CS theory is an emerging bulk of work that deals with sparse signals of interest[2]-.[4] Rather than collecting an entire Nyquist ensemble of signal samples, CS performs signal acquisition by a small number of (random[4] or deterministic[5]) linear measurements. Successful signal reconstruction relies on effective sparse signal representations and appropriate recovery algorithms such as convex optimization,[6] linear regression,[7,8] or greedy procedures.[9] Wireless video surveillance via compressed sensing can capture and compress video signals simultaneously through simple linear measurements highly reducing, therefore, data acquisition time and power consumption.

In the reconstruction of CS surveillance video, of particular interest is the ability to detect anomalies or moving objects that stand out from the background. Since the video sequence can be considered as the sum of a low-rank component corresponding to the stationary background scene and a sparse component that represents moving objects in the foreground, the problem of foreground and background separation from CS measurements can be solved by minimizing the nuclear norm of the low-rank component and the L_1 norm of the sparse component.[10] Although this method utilizes the low-rank property to enhance signal sparsity, it offers good reconstruction quality only when a large number of frames is available introducing, therefore, large latency to the decoding monitoring system.

In this paper, we consider the background extraction problem in compressed-sensed surveillance video from a principal component analysis (PCA) viewpoint. PCA is arguably a most widely used statistical tool for data analysis and dimensionality reduction today. Given that the observed data matrix is the sum of a low-rank component and a perturbation component, extraction of the low-rank component can be solved by (conventional)

Y.L.: E-mail: yl72@buffalo.edu, Telephone: 1 716 645 1207

D.A.P: E-mail: pados@buffalo.edu, Telephone: 1 716 645 1150

L_2-norm based principal component analysis (L_2-PCA). Under the statistical assumption that perturbation data are independent, identically distributed (i.i.d.) Gaussian, the optimal L_2-PCA solution, known simply as the dominant-singular-value left singular vectors of the observed data matrix, corresponds to maximum-likelihood estimation (MLE). On the negative side, from a practical data driven point of view, L_2-PCA becomes quite sensitive to potential outlier values contained in the perturbation component that are numerically distant from the nominal data.

In recent years, there has been a growing interest in robust PCA methods to deal with the problem of outliers in principal-component design[11]-.[18] In several studies[11]-,[14] subspace decomposition is performed under an L_1-error minimization criterion and the problem is solved via convex optimization. In addition, a low-rank and sparse decomposition problem is solved via minimizing the nuclear norm of the low-rank component and the L_1-norm of the sparse component.[15] The robustness of these methods comes at high computational cost due to linear or quadratic programming. Recently there has been a growing documented effort to calculate robust subspace components by L_1 projection maximization[16]-.[18] The resulting principal components are the so-called L_1 principal components. For instance, a suboptimal iterative algorithm was presented for the computation of one L_1 principal component[16] and an iterative algorithm was presented for the joint computation of $d \geq 1$ L_1 principal components.[17] In particular, algorithms for exact calculation of L_1 principal components are developed[18] for the first time in the literature.

While the algorithms for optimal L_1-PCA[18] find successful applications in the original signal space, in this work we propose for the first time a direct CS-measurement-domain L_1-PCA algorithm and apply the procedure to compressed-sensed surveillance video processing. For a surveillance video sequence, the low-rank property is preserved in the CS domain if each frame of the video is captured by the same compressed sensing matrix. Hence, L_1-PCA can be performed directly on the collected CS measurement vectors. Since the CS measurement vectors lie in a reduced dimensional space compared to the original pixel-domain data, the computational complexity for CS-L_1-PCA is dramatically lower. In the experimental studies, we not only demonstrate that the CS-L_1-PCA followed by regular CS image recovery can successfully extract the background scene for a surveillance video, but also illustrate the advantages of CS-L_1-PCA over CS-L_2-PCA when CS measurements are corrupted by outliers/faulty data.

The remainder of this paper is organized as follows. In Section 2, we present our motivation and establish the building blocks for our proposed procedure, that is exact computation of L_1-PCs and compressed-sensed image recovery based on total-variation (TV) minimization. In Section 3, the proposed CS L_1-PCA algorithm is developed and the overall foreground and background separation scheme is described in detail. Experimental results and performance analysis are presented in Section 4. Finally, a few conclusions are drawn in Section 5.

2. BUILDING BLOCKS OF THE PROPOSED ALGORITHM

2.1 Exact Computation of the L_1 Principal Components

Consider an observation data matrix $\mathbf{X} \in \mathbb{R}^{D \times N}$ that consists of a low-rank component $\mathbf{L} \in \mathbb{R}^{D \times N}$ and a perturbation matrix $\mathbf{E} \in \mathbb{R}^{D \times N}$, i.e.

$$\mathbf{X} = \mathbf{L} + \mathbf{E} \tag{1}$$

L_2-PCA refers to the problem of seeking the best low-rank-d ($d \leq \min\{D, N\}$) representation of $\mathbf{L}$ by solving

$$\mathcal{P}_1^{L_2} : (\mathbf{R}_{L_2}, \mathbf{S}_{L_2}) = \arg \min_{\substack{\mathbf{R} \in \mathbb{R}^{D \times d}, \mathbf{R}^T\mathbf{R} = \mathbf{I}_d \\ \mathbf{S} \in \mathbb{R}^{N \times d}}} \|\mathbf{X} - \mathbf{R}\mathbf{S}^T\|_2, \tag{2}$$

which is equivalent to the L_2 projection (energy) maximization problem

$$\mathcal{P}_2^{L_2} : \mathbf{R}_{L_2} = \arg \max_{\substack{\mathbf{R} \in \mathbb{R}^{D \times d} \\ \mathbf{R}^T\mathbf{R} = \mathbf{I}_d}} \|\mathbf{X}^T\mathbf{R}\|_2. \tag{3}$$

The optimal $\mathbf{R}_{L_2}$ solution (same in both $\mathcal{P}_1^{L_2}$ and $\mathcal{P}_2^{L_2}$) is known simply as the d dominant-singular-value left singular vectors of the data matrix $\mathbf{X}$.

When the perturbation matrix $\mathbf{E}$ may contain extreme outlier values (faulty measurements), L_1-PCA in the following form pursues a more accurate/robust subspace representation for $\mathbf{L}$ than L_2-PCA,

$$\mathcal{P}^{L_1} : \mathbf{R}_{L_1} = \arg \max_{\substack{\mathbf{R} \in \mathbb{R}^{D \times d} \\ \mathbf{R}^T \mathbf{R} = I_d}} \|\mathbf{X}^T \mathbf{R}\|_1. \tag{4}$$

The d columns of $\mathbf{R}_{L_1}$ in (4) are the so-called d L_1 principal components that describe the rank-d subspace in which $\mathbf{L}$ lies. Analysis shows that exact calculation of the L_1 principal components in Problem $\mathcal{P}^{L_1}$ can be recast as a combinatorial problem.[18] In short, when the rank of the nominal signal is $d = 1$, Problem $\mathcal{P}^{L_1}$ can be reduced to

$$\mathcal{P}^{L_1} : \mathbf{r}_{L_1} = \arg \max_{\substack{\mathbf{r} \in \mathbb{R}^D \\ \|\mathbf{r}\|_2 = 1}} \|\mathbf{X}^T \mathbf{r}\|_1, \tag{5}$$

and we can rewrite $\mathcal{P}^{L_1}$ as

$$\max_{\substack{\mathbf{r} \in \mathbb{R}^D \\ \|\mathbf{r}\|_2 = 1}} \|\mathbf{X}^T \mathbf{r}\|_1 = \max_{\substack{\mathbf{r} \in \mathbb{R}^D \\ \|\mathbf{r}\|_2 = 1}} \max_{\mathbf{b} \in \{\pm 1\}^N} \mathbf{b}^T \mathbf{X}^T \mathbf{r} \tag{6}$$

$$= \max_{\mathbf{b} \in \{\pm 1\}^N} \max_{\substack{\mathbf{r} \in \mathbb{R}^D \\ \|\mathbf{r}\|_2 = 1}} \mathbf{r}^T \mathbf{X} \mathbf{b} \tag{7}$$

$$= \max_{\mathbf{b} \in \{\pm 1\}^N} \|\mathbf{X} \mathbf{b}\|_2. \tag{8}$$

The optimal solution for (8) can be obtained by exhaustive search in the space of the binary antipodal vector $\mathbf{b}$. Since $\|\mathbf{X}\mathbf{b}\|_2 = (\mathbf{b}^T \mathbf{X}^T \mathbf{X} \mathbf{b})^{1/2} = (-\mathbf{b}^T \mathbf{X}^T \mathbf{X}(-\mathbf{b}))^{1/2}$, if $\mathbf{b}^*$ is an optimal solution, then $-\mathbf{b}^*$ is also an optimal solution. Hence, we can always set $\mathbf{b}(1) = 1$ and the complexity for exhaustive search on the other $N - 1$ coordinates of $\mathbf{b}$ is 2^{N-1}. Computation of $\|\mathbf{X}\mathbf{b}\|_2$ needs $D \times N$ multiplications, therefore the overall complexity for solving $\mathcal{P}^{L_1}$ is $2^{N-1} \times D \times N$.

When the rank of the nominal signal is $d > 1$, the problem $\mathcal{P}^{L_1}$ can be solved by

$$\max_{\substack{\mathbf{R} \in \mathbb{R}^{D \times d} \\ \mathbf{R}^T \mathbf{R} = I_d}} \|\mathbf{X}^T \mathbf{R}\|_1 \tag{9}$$

$$= \max_{\substack{\mathbf{R} \in \mathbb{R}^{D \times d} \\ \mathbf{R}^T \mathbf{R} = I_d}} \max_{\mathbf{B} \in \{\pm 1\}^{N \times d}} \mathrm{tr}(\mathbf{R}^T \mathbf{X} \mathbf{B}) \tag{10}$$

$$= \max_{\mathbf{B} \in \{\pm 1\}^{N \times d}} \|\mathbf{X} \mathbf{B}\|_* \tag{11}$$

where $\| \cdot \|_*$ stands for the nuclear norm. By Proposition 4,[18] to find exactly the optimal L_1-norm projection operator $\mathbf{R}_{L_1}$ in (4) we can perform the following steps:
1) Solve (11) to obtain $\mathbf{B}_{\mathrm{opt}}$.
2) Perform singular value decomposition (SVD) on $\mathbf{X}\mathbf{B}_{\mathrm{opt}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$.
3) Return $\mathbf{R}_{L_1} = \mathbf{U}_{:,1:d}\mathbf{V}^T$.
The complexity of the above algorithm is dominated by Step 1, which includes the exhaustive search on the binary matrix $\mathbf{B}^{N \times d}$ with complexity $\mathcal{O}(2^{N \times d})$, and the SVD per iteration of complexity $\mathcal{O}(\min\{D^2 d, D d^2\})$. Therefore, the overall complexity for finding d L_1 principal components via exhaustive search is $\mathcal{O}(2^{N \times d} \min\{D^2 d, D d^2\})$. For fixed data dimension D, a polynomial-time algorithm is developed[18] to solve optimally (11) with complexity $\mathcal{O}(N^{\mathrm{rank}(\mathbf{X})d-d+1})$, $\mathrm{rank}(\mathbf{X}) \leq D$. Moreover, a fast greedy approximation algorithm was proposed to solve (11) with complexity $\mathcal{O}(N^3)$.[19]

2.2 Compressed-sensed Video Recovery via Total-Variation (TV) Minimization

In this section we briefly review video frame acquisition by compressive sampling and recovery using sparse gradient constraints (TV minimization). If we consider the tth frame $\mathbf{X}_t \in \mathbb{R}^{m \times n}$ of the video sequence and

$\mathbf{x}_t \in \mathbb{R}^D$, $D = mn$, represents the vectorization of $\mathbf{X}_t$ via column concatenation, then CS measurements of $\mathbf{X}_t$ are collected in the form of

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{x}_t \tag{12}$$

with a linear measurement matrix $\mathbf{\Phi}_{P \times D}, P \ll D$. Under the assumption that images are mostly piece-wise smooth, it is natural to consider utilizing the sparsity of the spatial gradient of $\mathbf{X}_t$ for CS frame reconstruction[20]-[26]. If $x_{i,j}$ denotes the pixel in the ith row and jth column of $\mathbf{X}_t$, the horizontal and vertical gradients at $x_{i,j}$ are defined, respectively, as

$$D_{h;ij}[\mathbf{X}_t] = \begin{cases} x_{i,j+1} - x_{i,j}, & j < n, \\ 0, & j = n, \end{cases}$$

and

$$D_{v;ij}[\mathbf{X}_t] = \begin{cases} x_{i+1,j} - x_{i,j}, & i < m, \\ 0, & i = m. \end{cases}$$

The discrete spatial gradient of $\mathbf{X}_t$ at pixel $x_{i,j}$ can be interpreted as the 2D vector

$$D_{ij}[\mathbf{X}_t] = \begin{pmatrix} D_{h;ij}[\mathbf{X}_t] \\ D_{v;ij}[\mathbf{X}_t] \end{pmatrix} \tag{13}$$

and the anisotropic 2D-TV of $\mathbf{X}_t$ is simply the sum of the magnitudes of this discrete gradient at every pixel,

$$\begin{align} \mathrm{TV_{2D}}(\mathbf{x}_t) &\triangleq \sum_{ij} \left(|D_{h;ij}[\mathbf{X}_t]| + |D_{v;ij}[\mathbf{X}_t]| \right) \tag{14} \\ &= \sum_{ij} ||D_{ij}[\mathbf{X}_t]||_{\ell_1}. \end{align}$$

To reconstruct $\mathbf{X}_t$, we can solve the convex program

$$\widehat{\mathbf{x}}_t = \arg\min_{\mathbf{X_t}} \mathrm{TV_{2D}}(\mathbf{x}_t) \qquad \text{subject to} \quad \mathbf{y}_t = \mathbf{\Phi}\mathbf{x}_t. \tag{15}$$

However, in practical situations the measurement vector $\mathbf{y}_t$ may be corrupted by noise. Then, CS acquisition of $\mathbf{x}_t$ can be formulated as

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{x}_t + \mathbf{n}_t \tag{16}$$

where $\mathbf{n}_t$ is the noise vector. To recover $\mathbf{X}_t$, we can use 2D-TV minimization as in (15) and formulate the following unconstrained optimization problem

$$\widehat{\mathbf{x}}_t = \arg\min_{\mathbf{x_t}} \mu\mathrm{TV_{2D}}(\mathbf{x}_t) + \frac{1}{2}\|\mathbf{y}_t - \mathbf{\Phi}\mathbf{x}_t\|_{\ell_2}^2, \tag{17}$$

where μ is a non-negative weight controlling the sparsity level.

3. CS-DOMAIN L_1-PCA FOR COMPRESSED-SENSED SURVEILLANCE VIDEO

3.1 CS Surveillance Video Acquisition

In this work, we consider a practical CS surveillance video acquisition system that performs pure, direct compressed sensing on each video frame. In the simple compressive video encoding block diagram shown in Fig. 1, each frame $\mathbf{X}_t$ of size $m \times n$ is viewed as a vectorized column $\mathbf{x}_t \in \mathbb{R}^D$, $D = mn$, where t is the frame index. Compressive sampling is performed by projecting $\mathbf{x}_t$ onto a $P \times D$ ($P < D$) measurement matrix $\mathbf{\Phi}$,

$$\mathbf{y}_t = \mathbf{\Phi}\mathbf{x}_t, \tag{18}$$

where $\mathbf{\Phi}$ is generated by randomly permuting the columns of an order-k, $k \geq D$ and multiple-of-four, Walsh-Hadamard (WH) matrix followed by arbitrary selection of P rows from the k available WH rows (if $k > D$, only D arbitrary columns are utilized). This class of WH measurement matrices has the advantage of easy implementation (antipodal ± 1 entries), fast transformation, and satisfactory reconstruction performance as we will see later on. A reacher class of matrices can also be found,[27],[28] For practical implementation, $\mathbf{\Phi}$ is generated once and fixed for all frames in the video sequence. The resulting CS measurement vectors $\mathbf{y}_t$, $t = 1, 2, ...$, are then transmitted to the decoder.
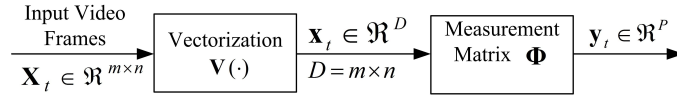
Figure 1. A compressed-sensing (CS) video encoder.

3.2 L_1-PCA for Background Extraction

We can collect a sequence of N CS measurement vectors $\mathbf{Y} \triangleq [\mathbf{y}_1 \ \mathbf{y}_2 \ \dots \mathbf{y}_N]$ of the form

$$\mathbf{Y} = \mathbf{\Phi X} + \mathbf{O} \qquad (19)$$
$$= \mathbf{\Phi}(\mathbf{L} + \mathbf{E}) + \mathbf{O} \qquad (20)$$
$$= \mathbf{Y_L} + \mathbf{\Phi E} + \mathbf{O} \qquad (21)$$

where $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \mathbf{x}_N]$ is a matrix consisting of the N corresponding video frames that can be viewed as a sum of the low-rank background $\mathbf{L}$ and the sparse moving objects in the foreground $\mathbf{E}$. Assuming $\mathbf{L}$ is a rank-d matrix, the CS-domain observed data matrix $\mathbf{Y_L} = \mathbf{\Phi L}$ is also of rank-d and represents the compressed-sensed background scene. To deal with the practical issue of possible faulty data, we assume that the observed CS measurements may be corrupted by outliers $\mathbf{O}$ due to acquisition failures. By applying L_1-PCA to $\mathbf{Y}$, we can extract the background scene information, i.e.

$$\mathbf{R}_{L_1} = \arg \max_{\substack{\mathbf{R} \in \mathbb{R}^{P \times d} \\ \mathbf{R}^T \mathbf{R} = \mathbf{I}_d}} \|\mathbf{Y}^T \mathbf{R}\|_1. \qquad (22)$$

The complexity of solving (22) is $O(2^{N-1} \times P \times N)$ by the exhaustive search algorithm and $\mathcal{O}(N^{P \times d - d + 1})$ by the polynomial-time algorithm.[18] Compared to pixel-domain L_1-PCA computation described in (9), the complexity is significantly reduced since the vector length is reduced from D to P due to compressed sensing. By projecting the observed CS measurements $\mathbf{Y}$ onto $\mathbf{R}_{L_1}$, we can obtain the compressed-sensed low-rank component

$$\widehat{\mathbf{Y}}_{\mathbf{L}}^{L_1} = \mathbf{R}_{L_1} \mathbf{R}_{L_1}^T \mathbf{Y}. \qquad (23)$$

Afterwards, the background scene can be reconstructed by performing CS recovery on the columns of $\widehat{\mathbf{Y}}_{\mathbf{L}}^{L_1}$, i.e. $\widehat{\mathbf{y}}_{\mathbf{L},t}^{L_1}$, $t = 1, 2, ..., N$. Here, we propose and use TV minimization introduced in Section 2 of the following form,

$$\widehat{\boldsymbol{\ell}}_t = \arg \min_{\boldsymbol{\ell}_t} \mu \text{TV}_{2\text{D}}(\boldsymbol{\ell}_t) + \frac{1}{2} \|\widehat{\mathbf{y}}_{\mathbf{L},t}^{L_1} - \mathbf{\Phi}\boldsymbol{\ell}_t\|_2^2. \qquad (24)$$

For comparison purposes, in parallel we introduce L_2-norm based CS-domain PCA calculation (SVD) by

$$\mathbf{R}_{L_2} = \arg \max_{\substack{\mathbf{R} \in \mathbb{R}^{P \times d} \\ \mathbf{R}^T \mathbf{R} = \mathbf{I}_d}} \|\mathbf{Y}^T \mathbf{R}\|_2. \qquad (25)$$

Similar to (23), the background scene can be obtained by projecting $\mathbf{Y}$ onto the L_2 principal components,

$$\widehat{\mathbf{Y}}_{\mathbf{L}}^{L_2} = \mathbf{R}_{L_2} \mathbf{R}_{L_2}^T \mathbf{Y}, \qquad (26)$$

followed by TV minimization as in (24). Since L_2-PCA is sensitive to outlier values, the performance of CS-L_2-PCA is anticipated to be inferior to CS-L_1-PCA in the presence of faulty/corrupted data.

3.3 Moving Objects Extraction

To extract the sparse moving objects in the foreground, we first perform frame-by-frame CS reconstruction in the form of

$$\widehat{\mathbf{x}}_t = \arg \min_{\mathbf{x}} \mu \text{TV}_{2\text{D}}(\mathbf{x}) + \frac{1}{2} \|\mathbf{y}_t - \mathbf{\Phi x}\|_2^2. \qquad (27)$$

With the recovered video frames $\widehat{\mathbf{X}} = [\widehat{\mathbf{x}}_1 \ \widehat{\mathbf{x}}_2 \ \ldots \widehat{\mathbf{x}}_N] \in \mathbb{R}^{D \times N}$ the sparse foreground can be recovered as

$$\widehat{\mathbf{e}}_t = \widehat{\mathbf{x}}_t - \widehat{\boldsymbol{\ell}}_t \tag{28}$$

for $t = 1, 2, ..., N$.

4. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

In this section, we demonstrate the proposed algorithmic developments on two test surveillance video sequences, *PETS2001* and *Airport*. For each video sequence, 10 frames are selected to form a video volume. Each frame is compressed-sensed independently using the same randomly permutated partial Walsh-Hadamard matrix. The number of CS measurements per frame is 37.5% of the total number of pixels in the video frame. Fig. 2 shows a representative frame from each video sequence. Fig. 3 displays background and foreground extraction

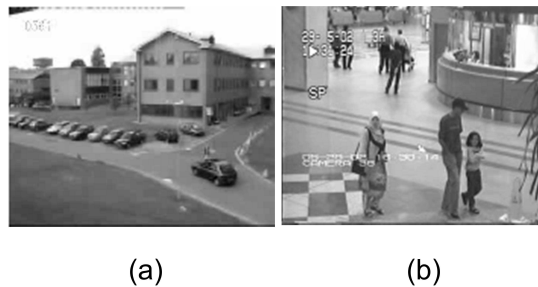(a)   (b)

Figure 2. Representative frames of (a) *PETS2001* and (b) *Airport* sequences.

results for the representative *PETS2001* frame shown in Fig. 2(a) when the CS measurements *do not contain outliers*. Rows (i) and (ii) are the results of CS-L_1-PCA* background and foreground, respectively, extraction with $d = 1, 2$, and 3 principal components (columns (a), (b), and (c), correspondingly). Rows (iii) and (iv) repeat the study for CS-L_2-PCA. It is observed that in the absence of outliers, both CS-L_1-PCA and CS-L_2-PCA correctly extract the background and the moving objects with one principal component ($d = 1$). When d increases (rank mismatch), CS-L_1-PCA maintains superior performance, while CS-L_2-PCA rapidly deteriorates.

Fig. 4 repeats the study of Fig. 3 for the same data set with corrupted CS measurements. In the experiment, 75% of the CS measurements of three randomly selected frames are corrupted by outliers. The presence of outliers in three frames modifies/increases the effective SVD rank of the background from $d = 1$ to 4. Naturally, when we use $d = 1$, both L_1- and L_2-PCA cannot recover the background/foreground scenes. When $d \geq 2$, L_1-PCA shows remarkable resistance to outliers and recovers the background and the moving objects well with L_1 rank choice $d = 2$ (or above, $d = 3, 4, 5$). On the other hand, L_2-PCA needs specifically $d = 4$ principal components (number of corrupted frames plus one) to recover the low-rank outlier corrupted background scene and its performance decreases quickly when $d > 4$.

For increased credibility, a similar experiment is performed on the *Airport* sequence (Figs. 5 and 6). In Fig. 6, 50% CS measurements of three randomly selected frames are corrupted by outliers. Similar conclusion to the studies of Figs. 3 and 4 can be drawn: CS-L_1-PCA offers superior robustness in rank selection and background/moving objects extraction in clean and outlier corrupted video sequences.

Finally, to place our proposed CS-domain PCA work (L_2 or L_1) in a broader context, we contrast our findings against traditional convex optimization algorithm for compressed-sensed surveillance video process.[10] In Fig. 7, we perform compressed-sensed low-rank and sparse decomposition[10] on non-corrupted CS measurements of the same set of video frames studied in Fig. 5. The results show that the background cannot be well reconstructed even when all CS measurements are received correctly. The reason is that convex-optimization-based compressed-sensed signal reconstruction relies on the sparse representation, but using the nuclear norm as the regularization

---

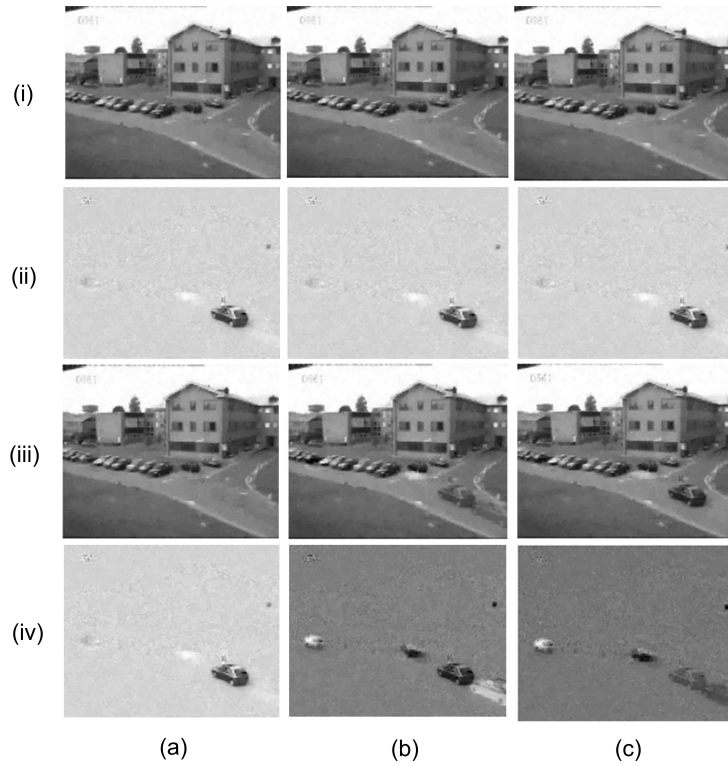*For L_1-PCA calculation, the fast algorithm[19] is adopted.

Figure 3. *PETS2001* sequence (clean CS measurements): CS-L_1-PCA reconstructed background and moving objects (row (i) and (ii), respectively) and CS-L_2-PCA reconstructed background and moving objects (row (iii) and (iv), respectively) with $d = 1$, $d = 2$, or $d = 3$ principal components (columns (a), (b), and (c), respectively).
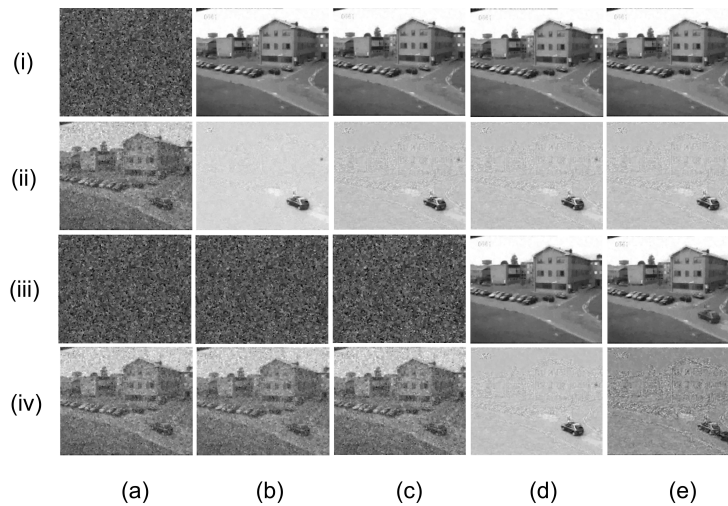

Figure 4. *PETS2001* sequence (75% outliers in CS measurements of 3 randomly selected frames): CS-L_1-PCA reconstructed background and moving objects (row (i) and (ii), respectively) and CS-L_2-PCA reconstructed background and moving objects (row (iii) and (iv), respectively) with $d = 1$, $d = 2$, $d = 3$, $d = 4$, or $d = 5$ principal components (columns (a), (b), (c), (d), and (e), respectively).
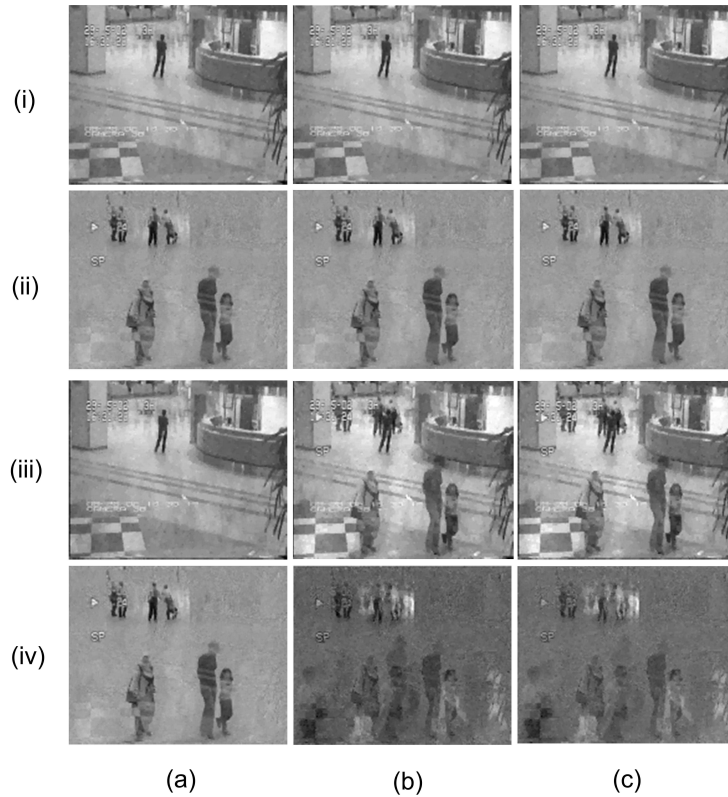
Figure 5. *Airport* sequence (clean CS measurements): CS-L_1-PCA reconstructed background and moving objects (row (i) and (ii), respectively) and CS-L_2-PCA reconstructed background and moving objects (row (iii) and (iv), respectively) with $d = 1$, $d = 2$, or $d = 3$ principal components (columns (a), (b), and (c), respectively).
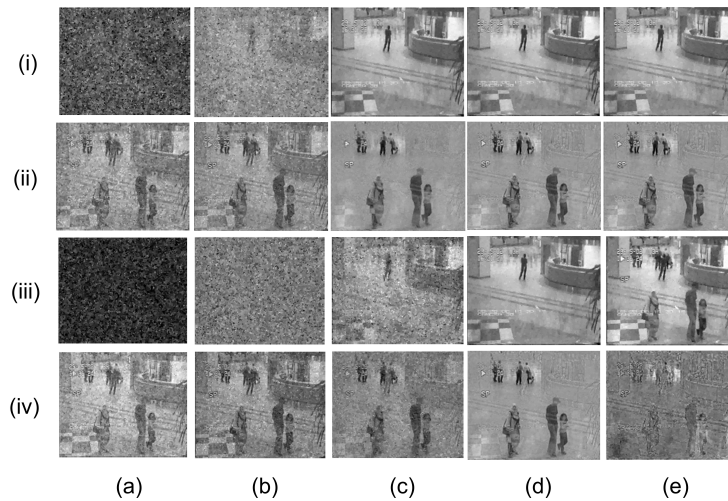

Figure 6. *Airport* sequence (50% outliers in CS measurements of 3 randomly selected frames): CS-L_1-PCA reconstructed background and moving objects (row (i) and (ii), respectively) and CS-L_2-PCA reconstructed background and moving objects (row (iii) and (iv), respectively) with $d = 1$, $d = 2$, $d = 3$, $d = 4$, or $d = 5$ principal components (columns (a), (b), (c), (d), and (e), respectively).

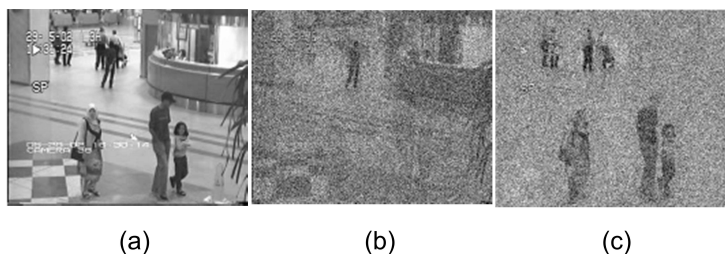(a)   (b)   (c)

Figure 7. *Airport* sequence frame reconstructed via CS low-rank and sparse decomposition[10] (no outliers in CS measurements): (a) Original frame; (b) reconstructed background; (c) reconstructed moving objects.

penalty for the low-rank component is not sparse enough when only a small number of frames are considered (10 frames in our experiment). In,[10] 100 frames are considered with a different compressed sensing matrix for each frame to guarantee the restricted isometry property (RIP)[2] required for successful sparse signal recovery from the CS measurements.

5. CONCLUSIONS

We proposed a compressed-sensing-domain L_1-norm maximization principal component analysis scheme for compressed-sensed surveillance video processing. The algorithm computes a low-rank subspace via L_1-PCA for the background scene directly in the CS domain and enjoys significantly lower computational complexity than pixel-domain L_1-PCA. Background reconstruction is then performed by projecting the CS measurement vectors onto the L_1 principal components followed by regular CS image recovery. For foreground extraction, each obtained background image is subtracted from its corresponding CS recovered frame. Experiments demonstrate that the CS-L_1-PCA algorithm performs better than L_2-norm based CS domain PCA when the CS measurements are corrupted by outliers. Even in clean CS data operation, CS-L_1-PCA offers exceptional performance and robustness in background subspace-rank selection compared to CS-L_2-PCA.

REFERENCES

1. I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Comput. Netw.*, vol. 51, no. 4, pp. 921-960, Mar. 2007.
2. E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.
3. D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.
4. E. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21-30, Mar. 2008.
5. K. Gao, S. N. Batalama, D. A. Pados, and B. W. Suter, "Compressive sampling with generalized polygons," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4759-4766, Oct. 2011.
6. E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure and Appl. Math.*, vol. 59, no. 8, pp. 1207-1223, Aug. 2006.
7. R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267-288, 1996.
8. B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, pp. 407-451, Apr. 2004.
9. J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.
10. H. Jiang, W. Deng, and Z. Shen, "Surveillance video processing using compressive sensing," *Inverse Problems and Imaging*, vol. 6, no. 2, pp. 201-214, 2012.
11. Q. Ke and T. Kanade, "Robust subspace computation using L_1 norm," *Internal Technical Report Computer Science Dept.*, Carnegie Mellon Univ., Pittsburgh, PA, USA, CMU-CS-03172, 2003.

12. 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, USA, Jun. 2005, pp. 739-746.

13. 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, USA, Jun. 2010, pp. 771-778.

14. 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.

15. E. Candès, X. Li, Y. Ma, and J. Wright, "Robust Principal Component Analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, pp. 1-37, May 2011.

16. N. Kwak, "Principal component analysis based on L_1-norm maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, pp. 1672-1680, Sep. 2008.

17. 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.

18. P. P. Markopoulos, G. N. Karystinos, and D. A. Pados, "Optimal algorithms for L_1-subspace signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5046-5058, Oct. 2014.

19. S. Kundu, P. P. Markopoulos and D. A. Pados, "Fast computation of the L_1-principal component of real-valued data," in *Proc. IEEE Intern. Conf. on Acoustics Speech and Signal Process. (ICASSP)*, Florence, Italy, May. 2014, pp. 8028-8032.

20. E. Candès and J. Romberg, "ℓ1-magic: Recovery of sparse signals via convex programming," URL: www.acm.caltech.edu/l1magic/downloads/l1magic.pdf.

21. M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 6, pp. 1182-1195, Dec. 2007.

22. S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *Proc. IEEE Conf. Comp. Vis. and Pattern Recogn.*, Anchorage, Alaska, Jun. 2008, pp. 1-8.

23. C. Li, "An efficient algorithm for total variation regularization with applications to the sigle pixel camera and compressive sensing," *Master's Thesis*, Dept. of Comp. and Applied Math., Rice University, Houston, TX, 2009.

24. M. R. Dadkhah, S. Shirani, and M. J. Deen, "Compressive sensing with modified total variation minimization algorithm," in *Proc. IEEE Intern. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, Dallas, TX, Mar. 2010, pp. 1310-1313.

25. Y. Liu and D. A. Pados, "Decoding of framewise compressed-sensed video via interframe total variation minimization," *SPIE J. Electron. Imaging, Special Issue on Compressive Sensing for Imaging*, vol. 22, no. 2, pp. 1-8, Apr.-Jun. 2013.

26. Y. Liu and D. A. Pados, "Rate-adaptive compressive video acquisition with sliding-window total-variation-minimization reconstruction," in *Proc. SPIE, Compressive Sensing Conf., SPIE Defense, Security, and Sensing*, Baltimore, MD, vol. 8717, May, 2013, pp. 1-13.

27. H. Ganapathy, D. A. Pados, and G. N. Karystinos, "New bounds and optimal binary signature sets - Part I: Periodic total squared correlation," *IEEE Trans. Comm.*, vol. 59, pp. 1123-1132, Apr. 2011.

28. H. Ganapathy, D. A. Pados, and G. N. Karystinos, "New bounds and optimal binary signature sets - Part II: Aperiodic total squared correlation," *IEEE Trans. Comm.*, vol. 59, pp. 1411-1420, May 2011.