Research Article

Optimization of the Regularization in Background and Foreground Modeling

Si-Qi Wang and Xiang-Chu Feng

School of Mathematics and Statistics, Xidian University, Xi'an 710126, China

Correspondence should be addressed to Si-Qi Wang; pearlwangxd@163.com

Background and foreground modeling is a typical method in the application of computer vision. The current general "low-rank + sparse" model decomposes the frames from the video sequences into low-rank background and sparse foreground. But the sparse assumption in such a model may not conform with the reality, and the model cannot directly reflect the correlation between the background and foreground either. Thus, we present a novel model to solve this problem by decomposing the arranged data matrix D into low-rank background L and moving foreground M. Here, we only need to give the priori assumption of the background to be low-rank and let the foreground be separated from the background as much as possible. Based on this division, we use a pair of dual norms, nuclear norm and spectral norm, to regularize the foreground and background, respectively. Furthermore, we use a reweighted function instead of the normal norm so as to get a better and faster approximation model. Detailed explanation based on linear algebra about our two models will be presented in this paper. By the observation of the experimental results, we can see that our model can get better background modeling, and even simplified versions of our algorithms perform better than mainstream techniques IALM and GoDec.

1. Introduction

Over the past several decades, the applications of monitoring and control system, including video surveillance, traffic monitoring and analysis, human detection and tracking, and gesture recognition in human-machine interface, are increasingly popular and widely used in our daily life. Moeslund et al. [1] and many scholars were committed to the related researches by using various methods.

Background subtraction is an algorithm proposed since 1999. Due to the underlying ease of implementation and effectiveness, foreground object detection by using background subtraction has been widely used in the video surveillance applications. This kind of approach can be traced back to Toyama et al. [2], the algorithm which predicts pixel intensity by using the Wiener filter to dig out the foreground region instead of maintaining a specific image. In the meanwhile, Stauffer and Grimson [3, 4] proposed a method based on the Gaussian mixture model (GMM) to track the dynamic

objects and give the trajectory of them. The GMM has gained great popularity among the computer vision community, and other scholars continue to revisit the method and propose enhanced algorithms [5, 6]. But modeling the high frequency variations in the background with 3–5 of Gaussian distribution is not accurate and even fails to achieve the sensitive detection. So, in the following studies, Elgammal et al. [7, 8] proposed a corresponding new background subtraction, using more short-term distributions to obtain better detection sensitivity. They put forward the statistical scene background by using general nonparametric kernel density estimation to subtract the background while retaining the foreground image. All the methods above are real-time algorithm; the estimated background at each pixel location is based on the pixels recent history and no spatial correlation is used between different or neighbouring pixel locations. So, Oliver built an eigenspace to model the background and used principal component analysis (PCA) to reduce the dimensionality of the space [9].

Over the last five years, there has been growing interest within the usage of PCA to model the background. Chandrasekaran put forward the "low-rank + sparse" model to achieve the purpose of background modeling and foreground detection in the video processing field [10]. In 2009, Candes et al. [11] proposed the robust principal component analysis (RPCA) method to decompose the large data matrix of video sequence into low-rank and sparse components. RPCA being used firstly in the field of computer vision was put forward by de la Torre and Black [12]. Almost at the same time as Candes, Wright et al. [13] also chose RPCA to model the background as a low-rank part of the frames in the video sequence. After that, many scholars threw themselves into the work to solve this model better. An alternating direction method (ADM) for sparsity and low-rank decomposition was studied by Yuan and Yang [14]. Lin et al. solved this optimization problem via the method of augmented Lagrange multipliers (ALM) in [15]. Zhou and Tao [16] introduced a special fast low-rank approximation based on bilateral random projections (BRP), which was the essence of their algorithm GoDec. Peng et al. [17] referred to RPCA problem as robust alignment by sparse and low-rank (RASL) decomposition and gave the outer and inner loop of RASL to solve the problem. All these improved PCA methods exploit the correlations between different pixels very well and no longer being local method like the previous background subtraction algorithms.

The general "low-rank + sparse" model in the articles available is as follows:

$$\min_{L,S} \frac{1}{2} \|D - L - S\|_F^2 + \lambda \|L\|_* + \mu \|S\|_1, \qquad (1)$$

where D is a data matrix whose columns are the pixels from each frame in the video sequence; L denotes a low-rank matrix which naturally corresponds to the stationary background; and S is a sparse matrix which contains the moving objects in the foreground. The Frobenius norm $\|\cdot\|_F^2$ is defined as the quadratic sum of the matrix elements; the nuclear norm $\|\cdot\|_*$ is defined as the sum of its singular values; the l_1-norm $\|\cdot\|_1$ is seen as the sum of the absolute values of the matrix elements [18, 19].

To address this convex optimization problem, numerous approaches have been explored and proposed in the academic papers over several decades. However, in my opinion, the "sparse" assumption in this model may not always conform to the reality. The foreground moving target may be a very small part of the full panoramic view, and it may also occupy a large proportion in the captured picture. And, under normal circumstances, the scholars usually add a related item onto (1) to constraint the correlation between foreground and background matrices. Therefore, we are trying to improve the PCA method in another way to find a new model to separate the background and foreground and reflect the correlation between them as well. In this paper, we proposed a "low-rank + dual" model to solve this foreground and background modeling problem. And we further use the reweighted dual function norm instead of the normal norms so as to get a better and faster model, as shown in Algorithm 2. Although our two models are all based on offline method, they provide the probability distribution function of the background within the consideration of the spatial correlation between neighboring pixels. Experiments on several scenes show that our methods have better performance and lower consumption compared with the "low-rank + sparse" methods, IALM and GoDec.

This paper is organized as follows. Section 2 briefly details our motivation of proposing the rudimentary dual model, and the algorithm to this dual model will also be given; a further improved reweighted form model and the corresponding algorithm will be given in Section 3; and, in Section 4, the experiments of our methods will be presented and the results will be compared with IALM and GoDec; and a final summary will be given in Section 5.

2. Dual Norm Model

2.1. Functional with the Dual Norm Constraint. Inspired by Meyer for the usage of the dual norms to regular the cartoon and texture in the image decomposition problem [20], we try to use the advantages of such dual regularizations. So, we put forward our rudimentary dual norm model, "low-rank + dual" model, which has been improved from the general "low-rank + sparse" model, as shown in Algorithm 1. Here, we keep the low-rank hypothesis of the model for the background, which is similar to Oliver's concept of eigenspace, and use the nuclear norm as one regularization to realize this assumption. But, for the foreground part, we use the "dual" norm regularization instead of the "sparse" constraint to obtain the most uncorrelated foreground from the background. To illustrate this point, we need to recall the definition of the dual norm [21].

Definition 1. For any given norm $\|\cdot\|$ in an inner product space, there exists a dual norm $\|\cdot\|_d$ defined as

$$\|X\|_d := \sup \left\{ |\langle X, Y \rangle| : Y \in \mathbb{R}^{m \times n}, \|Y\| \le 1 \right\}$$
$$= \sup \left\{ \frac{|\langle X, Y \rangle|}{\|Y\|} : Y \in \mathbb{R}^{m \times n}, Y \ne 0 \right\}. \qquad (2)$$

Furthermore, the dual norm of $\|\cdot\|_d$ is again the original norm $\|\cdot\|$.

From the definition, we can obtain the following upper bound of the inner product:

$$|\langle X, Y \rangle| \le \|X\|_d \|Y\|. \qquad (3)$$

Then, due to the fact that $\forall a, b > 0$ always has $2\sqrt{ab} \le a + b$, we can further convert (3) into the following equivalent result:

$$2\sqrt{|\langle X, Y \rangle|} \le \|X\|_d + \|Y\|. \qquad (4)$$

Thus, minimizing the sum of the norms $\|X\|_d$ and $\|Y\|$ can get a smaller value of the inner product $|\langle X, Y \rangle|$, which leads to the most irrelevant X and Y. Back to the foreground and background modeling processing in the video sequence problems, we use the nuclear norm as a prior assumption of the background. Then, if we use the dual norm of the nuclear norm to the foreground and minimize the sum of the norms

(1) **Input**: D,
(2) **initialize**: $L^0 = \overline{D}, k = 0, \lambda > 0, \mu > 0$.
(3) **repeat**
(4) $k = k + 1$;
(5) $M^k = T_\tau(D - L^{k-1})$; [calculate τ with (17)]
(6) $L^k = S_\lambda(D - M^k)$.
(7) **until** a stopping criterion is satisfied.
(8) **output**: L^*, M^*.

ALGORITHM 1: Solving dual norm model NSMP with AIM.

(1) **Input**: D,
(2) **initialize**: $M^0 = 0, k = 0, \lambda > 0, \mu > 0, c > 0, \delta > 0$.
(3) **repeat**
(4) $k = k + 1$;
(5) $w_i^k \in c\sqrt{n}/(\Sigma_i(D - M^{k-1}) + \delta)$;
(6) $L^k = T_{\lambda w}(D - M^{k-1})$;
(7) $M^k = T_{\tau/w}(D - L^k)$. [calculate τ with (17)]
(8) **until** a stopping criterion is satisfied.
(9) **output**: L^*, M^*.

ALGORITHM 2: Solving WNSMP with reweighted dual function.

in the minimization problem, then we will be able to get the most unrelated foreground and background from the known data matrix.

Theorem 2 (Proposition 2.1 in [22]). *The dual norm of the nuclear norm $\| \cdot \|_*$ is the operator norm $\| \cdot \|$ (it is the spectral norm $\| \cdot \|_2$) in $\mathbb{R}^{m \times n}$.*

So, the minimization problem (1) can be improved as the following nuclear norm and spectral norm minimization problem (NSMP):

$$\text{NSMP}: \min_{L,M} \frac{1}{2}\|D - L - M\|_F^2 + \lambda\|L\|_* + \mu\|M\|_2, \quad (5)$$

where $\| \cdot \|_2$, the spectral norm, is the largest singular value of the matrix.

This is our rudimentary model, with the fidelity term, which makes the result being faithful and loyal to the original video data; with the regularization terms, the nuclear norm ensures the low-rank of the background, and the spectral norm ensures the extraction of the foreground.

2.2. Algorithm to the Dual Norm Model. Although (5) is a nonconvex problem, while its two subproblems, nuclear norm minimization (NNM) problem and spectral norm minimization (SNM) problem, are all convex. Thus, we can solve the subproblems alternatively instead of solving (5) directly.

2.2.1. Nuclear Norm Minimization Problem. For a fixed matrix M, the NNM problem can be expressed as the following subproblem:

$$\min_L J_1 = \arg\min_L \frac{1}{2}\|(D - M) - L\|_F^2 + \lambda\|L\|_*. \quad (6)$$

The solution to this NNM problem has already been given by Cai et al. in Theorem 2.1 of his work [23]. Suppose $D - M = U\Sigma V^T$ is the singular value decomposition (SVD) of the matrix $D - M$; then, the solution to (6) is

$$S_\lambda(D - M) := US_\lambda(\Sigma)V^T, \quad (7)$$

where the soft-threshold operator S_λ is defined as $S_\lambda := \max(|\cdot| - \lambda, 0)$ [24].

2.2.2. Spectral Norm Minimization Problem. Similarly, for a fixed matrix L, the other subproblem can be expressed as the following form:

$$\min_M J_2 = \arg\min_M \frac{1}{2}\|(D - L) - M\|_F^2 + \mu\|M\|_2. \quad (8)$$

We will provide the singular value threshold method for the spectral norm minimization problem based on the linear algebra in this section. For convenience, we can reexpress the problem as follows:

$$\min_X J = \arg\min_X \frac{1}{2}\|Y - X\|_F^2 + \mu\|X\|_2. \quad (9)$$

Suppose we have the SVD of the matrix $Y = U_Y\Sigma_Y V_Y^T$; then, we put forward the following theorem.

Theorem 3. *The solution to the minimization problem* (9) *can be obtained by* $X = U_Y \Gamma^* V_Y^T$, *where* Γ^* *obeys*

$$\min_{\Gamma^*} J = \min_{\Gamma_\wedge} \frac{1}{2} \|\text{vec}\left(\text{diag}\left(\Sigma_Y - \Gamma_\wedge\right)\right)\|_2^2 \tag{10}$$
$$+ \mu \|\text{vec}\left(\text{diag}\left(\Gamma_\wedge\right)\right)\|_\infty.$$

vec(diag($\cdot$)) *denotes the vector obtained by the rearrangement of the diagonal elements of the matrix.*

Here, we assume $X = U_Y \Gamma V_Y^T$ directly and set Γ as any matrix of the same size with Σ_Y and let $\Gamma_\wedge$ denote the diagonal of the matrix Γ. For the reason that the Frobenius norm and the spectral norm are invariant under orthogonal transformation, (9) can be firstly converted into

$$\min_\Gamma J = \min_\Gamma \|\Sigma_Y - \Gamma\|_F^2 + \mu \|\Gamma\|_2. \tag{11}$$

On account of the fact that $\|\Gamma\|_2 = \|\Gamma_\wedge\|_2$ and $\|\Sigma_Y - \Gamma\|_F^2 \geq \|\Sigma_Y - \Gamma_\wedge\|_F^2$, (11) can be further converted into

$$\min_\Gamma J = \min_{\Gamma_\wedge} \|\Sigma_Y - \Gamma_\wedge\|_F^2 + \mu \|\Gamma_\wedge\|_2. \tag{12}$$

Finally, rewriting the above function into the vector form can get (10), which means the minimum of the solution reaches on the diagonal of the variable matrix.

The problem (10) has already been solved by Fadili and Peyre in Proposition 2 of his article [25]. To solve the problem, we introduce a sequence,

$$d[1] \leq d[2] \leq \cdots \leq d[N], \tag{13}$$

where $\{d[t]\}_{t=1}^N = \{\sigma_Y[i]\}_{i=1}^N$, $N = \min(m, n)$, means that the array $\{d[t]\}$ is the rearrangement of the singular values of the matrix Y. And sequence (13) can be extended into

$$d[1] \leq d[2] \leq \cdots \leq d[N]$$
$$= D[N] \leq D[N-1] \leq D[N-2] \leq \cdots \leq D[2] \leq D[1], \tag{14}$$

where $D[s] = \sum_{t=s}^N d[t]$ is the cumulated ordered values.

For $\mu \geq 0$ and $Y \in \mathbb{R}^{m \times n}$, the optimal solution to the problem (9) obeys

$$T_\tau(Y) = \begin{cases} 0, & \text{if } \|\sigma_Y\|_1 \leq \mu, \\ U_Y \cdot \text{diag}\left(\sigma_Y - S_\tau(\sigma_Y)\right) \cdot V_Y^T, & \text{otherwise,} \end{cases} \tag{15}$$

where

$$S_\tau(\sigma_Y)[i] = \max\left(1 - \frac{\tau}{|\sigma_Y[i]|}, 0\right) \cdot \sigma_Y[i], \tag{16}$$

and $\tau > 0$ is given by

$$\tau = d[t] + (d[t+1] - d[t]) \frac{\mu - D[t+1]}{D[t] - D[t+1]}, \tag{17}$$

where t satisfies $D[t+1] \leq \mu < D[t]$.

Such a threshold could keep most of the small singular value while changing the larger singular values as a constant, and the constant changes with the matrix Y. So, the solution to the problem (8) can be obtained by $T_\tau(D - L)$.

2.2.3. Alternating Iterative Algorithm. Based on aforementioned analysis, we present the alternating iterative method (AIM) to the minimization problem (5).

(i) *Stopping Criterion.* We suggest using the following value to judge whether to end iteration:

$$\min\left(\frac{\|L^k - L^{k-1}\|_*}{\|L^k\|_*}, \frac{\|M^k - M^{k-1}\|_2}{\|M^k\|_2}, \frac{\|D - L^k - M^k\|_F}{\|D\|_F}\right) \leq \varepsilon, \tag{18}$$

where $\varepsilon = 1e - 5$. And if the iterations exceed the maximum set $K = 30$, the method will also stop.

(ii) *Parameter λ.* We set the value of λ according to different video sequence as $\lambda = \delta_1 \sqrt{mn}$, where m and n are the sizes of the data matrix D and $\delta_1 \in (1, 10)$ is a small adjustable constant that ensures the good performance.

(iii) *Parameter μ.* We set μ according to $\mu = \delta_2 n \sqrt{mn}$, where $\delta_2 \in (0, 1)$ is also a small adjustable constant. We can see that the selection of parameters is basically multiplied by n, which is used to ensure the two regular items calculate in the same order of magnitude.

The choice of parameter values for the algorithm is very important, and there has two parameters, λ, and μ, which can be adjusted in our algorithm. Due to the slightly different parameter settings for different data, here we can only give out the setting range of these two parameters. The parameters will influence each other, thereby affecting the running of the algorithm. The selection of λ and μ cannot be too small; otherwise, the processing speed will slow down and the image will have great losses. The more complex the foreground of the frames, the greater λ and μ needed. A larger λ brings a cleaner background, and a larger μ causes a more complete extract of portrait. However, the value of the parameters can not just blindly large, too large λ will leads to the failure to extract the background, and too large μ will extract background into the foreground. The exact theoretical basis of the parameters selection will conduct a detailed study in the follow-up work.

3. Improved Dual Function Model

3.1. Model Creation. In the previous section, we put forward the preliminary dual model NSMP. And next, we will introduce an improved dual model, weighted function nuclear norm and spectral norm minimization model. This "low-rank + dual" model applies the weighted function nuclear norm to the low-rank background instead of the normal nuclear norm then find the dual representation to this weighted function nuclear norm to regular the foreground. The reason we put forward the improvement is that the well-known nuclear norm regularized problem is the convex relaxation of the rank minimization problem, but it is not a perfect approximation of the rank function. Although the minimization problem with the weighted function nuclear norm is nonconvex, fortunately it has a closed form solution

due to the special choice of the value of weights, and it is also a better approximation to the rank function. Thus, what we are going to do next is introduce the weighted function and optimize our model further. By this way, we can improve (5) into the following weighted function nuclear norm and spectral norm minimization problem (WNSMP):

$$\text{WNSMP} : \min_{L,M} \frac{1}{2}\|D - L - M\|_F^2 + \lambda\|w(L)\|_* + \mu\|w^{-1}(M)\|_2,$$
$$(19)$$

where $w(\cdot)$ denotes the weighted function which directly adds the weights onto the singular values of the matrix, and, for any matrix $X \in \mathbb{R}^{m \times n}$, weighted function norm is defined as

$$\|w(X)\|_* := \sum_{i=1}^{\min(m,n)} w_i \sigma_i(X), \qquad (20)$$

$$\|w^{-1}(X)\|_2 := \max_i \left(\frac{1}{w_i}\sigma_i(X) \right). \qquad (21)$$

If we still rearrange the singular values σ_i ($i = 1, \ldots, \min\{m, n\}$) of the matrix X into a vector σ, then we will have that $\|w(X)\|_*$ is actually the weighted l_1-norm of the vector σ. For any vector v, there exists

$$|\langle \sigma, v \rangle| = \left| \sum_{i=1}^{\min(m,n)} \frac{1}{w_i}\sigma_i \cdot w_i v_i \right| \le \sum_{i=1}^{\min(m,n)} \left| \frac{1}{w_i}\sigma_i \cdot w_i v_i \right|$$

$$\le \max_i \left| \frac{1}{w_i}\sigma_i \right| \cdot \sum_{i=1}^{\min(m,n)} |w_i v_i| = \max_i \left| \frac{1}{w_i}\sigma_i \right| \cdot \|v\|_{w,1}.$$
$$(22)$$

So, we can further get

$$\|\sigma\|_{w,1}^{\text{dual}} = \sup_{\|v\|_{w,1} \le 1} |\langle \sigma, v \rangle| \le \sup_{\|v\|_{w,1} \le 1} \max_i \left| \frac{1}{w_i}\sigma_i \right| \cdot \|v\|_{w,1}$$

$$\le \max_i \left| \frac{1}{w_i}\sigma_i \right| = \|\sigma\|_{w^{-1},\infty}, \qquad (23)$$

where $\|\sigma\|_{w^{-1},\infty}$ is actually the weighted function spectral norm $\|w^{-1}(X)\|_2$. Thus, we finally get that $\|w(X)\|_*$ and $\|w^{-1}(X)\|_2$ are dual form to each other under the same weight.

3.2. Algorithm to the Dual Function Model

3.2.1. Weighted Function Nuclear Norm Minimization Problem.
Here, we still use the alternating iterative algorithm to solve the problem (19). Fix the matrix M to get the first subproblem as

$$\min_L J_1 = \arg\min_L \frac{1}{2}\|(D - M) - L\|_F^2 + \lambda\|w(L)\|_*. \qquad (24)$$

For the convenience of discussion, we set $Z = D - M$ and substitute (20) into the functional; then, we can rewrite (24) as follows:

$$\min_L J_1 = \arg\min_L \frac{1}{2}\|Z - L\|_F^2 + \lambda \sum_{i=1}^{\min(m,n)} w_i \sigma_i(L). \qquad (25)$$

Here, we would impose increasing weight to ensure the global solution to (25), although $\sum_{i=1}^{\min(m,n)} w_i \sigma_i(L)$ is no longer a matrix norm. In our paper, we determine the adaptive weights by

$$w_i \in \frac{\delta\sqrt{n}}{(\sigma_i(L) + c)}, \qquad (26)$$

where $\delta = 1e + 2$ is an adjustable constant to ensure the value range of the weights including 1, n is the number of the frames included in the sample image sequence, $\sigma_i(L)$ denotes the singular values of the low-rank matrix L, and $c = 1e - 16$ is to avoid the initial value of zero.

Theorem 4 (Theorem 2.3 in [26]). *For any $\lambda \ge 0$, $Z \in \mathbb{R}^{m \times n}$, and $0 \le w_1 \le w_2 \le \cdots \le w_{\min(m,n)}$, a global optimal solution to the optimization problem (25) is given by the adaptive SVD soft-thresholding (ASVT) operator $L^* := T_{\lambda w}(Z)$, where $Z = U_Z \Sigma_Z V_Z^T$, and*

$$T_{\lambda w}(Z) := U_Z S_{\lambda w}(\Sigma_Z) V_Z^T,$$

$$S_{\lambda w}(\Sigma_Z) := \text{diag}\{(\Sigma_Z[i] - \lambda w_i)_+, i = 1, \ldots, \min(m, n)\}.$$
$$(27)$$

Thus, the solution to the problem (24) is $T_{\lambda w}(D - M)$.

3.2.2. Weighted Function Spectral Norm Minimization Problem.
Fixing the matrix L, we can express WNSMP by the following subproblem:

$$\min_M J_2 = \arg\min_M \frac{1}{2}\|(D - L) - M\|_F^2 + \mu\|w^{-1}(M)\|_2. \qquad (28)$$

For any $\mu \ge 0$, the solution to the problem (28) is $T_{\tau/w}(D - L)$, and, for any matrix, $X = U_X \Sigma_X V_X^T$ has

$$T_{\tau/w}(X) = \begin{cases} 0, & \text{if } \|\sigma_X\|_1 \le \mu \\ U_X \cdot \text{diag}(\sigma_X - S_{\tau/w}(\sigma_X)) \cdot V_X^T, & \text{otherwise,} \end{cases}$$
$$(29)$$

where $S_{\tau/w}(\sigma_X)[i] = (\sigma_X[i] - (\tau/w_i))_+$.

3.2.3. Alternating Iterative Algorithm.
In order to enable the values of weights to be updated in each iteration process, we set w_i^k in the kth iteration associated with the low-rank matrix L^k. But the problem is that, in the kth iterative, the matrix L^k is not available. So, we need to use the matrixes Z^{k-1} to estimate the value of the matrix $\hat{L}^k$ as

$$\Sigma_i(\hat{L}^k) = \Sigma_i(Z^{k-1}) = \Sigma_i(D - M^{k-1}). \qquad (30)$$

We always rearrange the value included in the set $\{\delta\sqrt{n}/(\Sigma_i(D - M^{k-1}) + c), i = 1, \ldots, \min(m, n)\}$ to ensure the weights to be nondecreasing. Stopping criterion and parameter setting are the same as in NSMP.
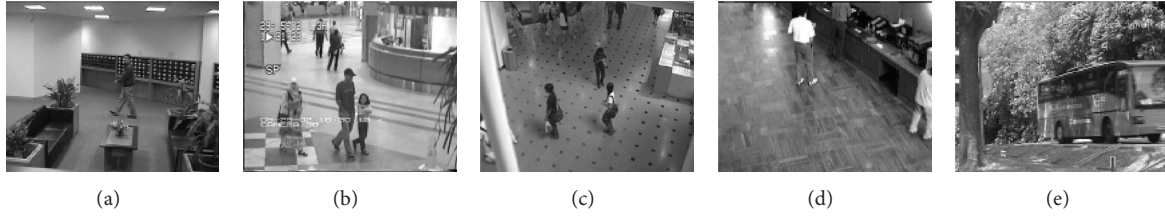
FIGURE 1: Test samples from video sequence: lobby (a), airport (b), shopping mall (c), restaurant (d), and campus (e).
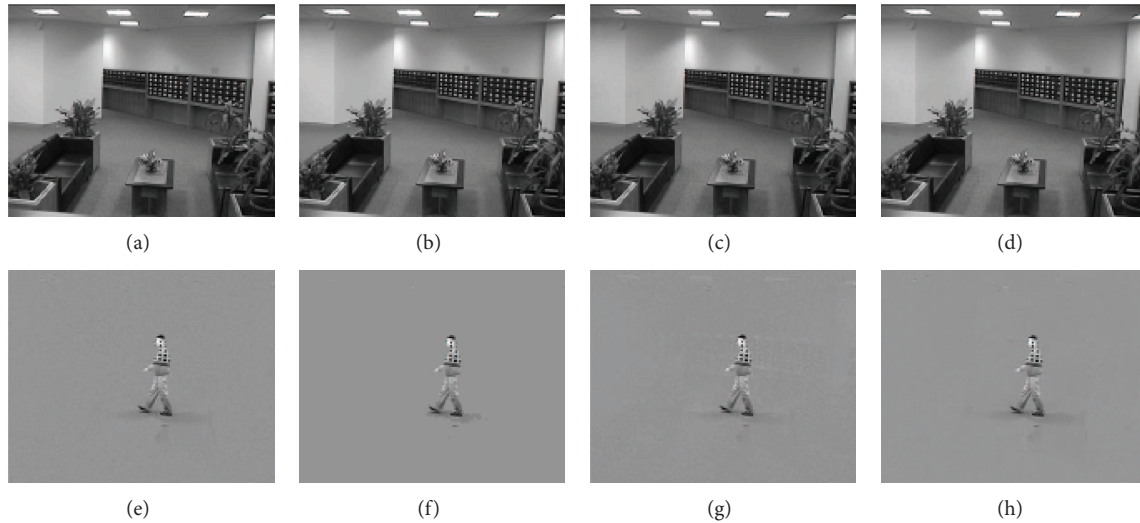

FIGURE 2: Lobby: (a)–(d) background and (e)–(h) foreground from IALM, GoDec, NSMP, and WNSMP.

4. Experimental Results

In this section, we present the numerical experiments corroborating our main results. We practice IALM, GoDec, NSMP, and WNSMP in a wide variety of surveillance video sequences, including five different scenarios: lobby, restaurant, airport, shopping mall, and campus, and all the datasets can be downloaded in http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html. We stack each 200 frames from one scenario to be rearranged as the columns of one data matrix and use these four algorithms to decompose the data matrix into different background and foreground. All the selected frames from the five video sequences are listed in Figure 1. We can see that the foreground part in Figure 1(a) is just a single person, while, in Figures 1(b), 1(c), and 1(d), there are a lot of people in motion; further to Figure 1(e), the moving foreground is expanded into objects.

Figure 2 shows us the segmentation results in the case that only one person is moving in the video, and we can see that the difference of the performance is inconspicuous, four kinds of methods all doing well.

In the simple situations, the advantage of our algorithm is not very obvious, but, in a relatively complex case, the superiority will be immediately apparent. The backgrounds of IALM in Figures 3(a), 4(a), and 5(a) always exist incomplete processed moving shadows in the results. And the artifacts problem in both foreground and background of

GoDec gets worse. Characterized by three kinds of situations, backgrounds always have the unnatural artifacts shown in Figure 4(b); as shown in Figure 5(b), there exist incomplete processing moving shadows in the processed background just like IALM in Figure 5(a); Even more worse, due to the interference by the other frames, unexpected shadow appears in the results of the current frame, as shown in Figure 5(f). While, in these processes, our algorithms always have stable performance with clear background and complete foreground.

Figure 6 shows the far more complex situations. Although the results do not live up to the expectation of greater progress, our method still has a relatively good processing results. The backgrounds in our methods are always very clear, while GoDec has a little artifact, and IALM is even worse.

We can easily conclude from the observed image, Figures 2–6, that our model can get a better and clearer background with low-rank. Therefore, in order to show that our model is actually good enough to model the background, we display the precision-recall (PR) curves and the receiver operator characteristic (ROC) curves in Figure 7 to evaluate the binary decision problem about the decomposition of the foreground and background. Furthermore, we list the rank of the final modeling background of the four methods in Table 1.

Under the noiseless situation, "low-rank" is the unique characteristic of the background in the video sequence,
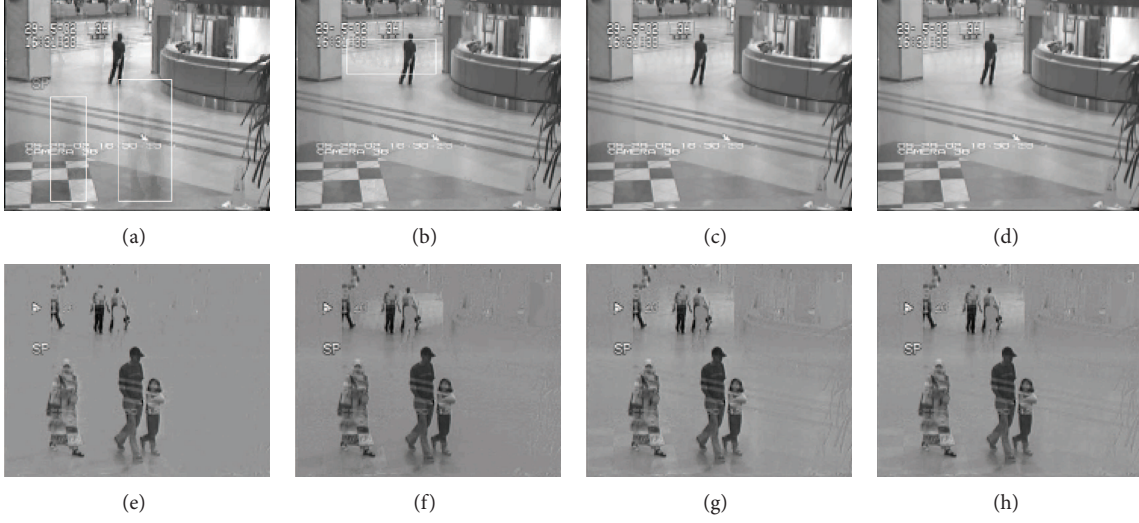
FIGURE 3: Airport: (a)–(d) background and (e)–(h) foreground from IALM, GoDec, NSMP, and WNSMP.
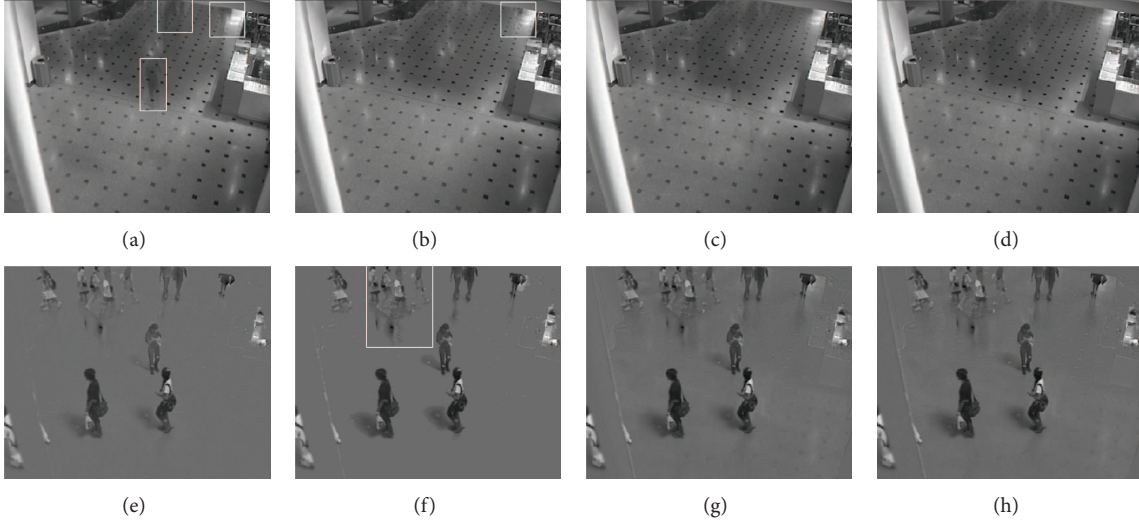

FIGURE 4: Shopping mall: (a)–(d) background and (e)–(h) foreground from IALM, GoDec, NSMP, and WNSMP.

and the small correlation can separate the nonlow-rank foreground from the background. Thus, our "low-rank + dual" model not only keeps the background but also extracts the foreground part of the video. Additionally, the correlation of the foreground and background images has been given in Table 1. Here, we adopt the following ways to measure the correlation of the foreground and background images:

$$\text{Cor}\,(L, M) = \max\left(\frac{\left|l_i^T m_i\right|}{\|l_i\|_2 \|m_i\|_2}\right), \tag{31}$$

where l_i and m_i denote each column in the matrixes L and M, which are the foreground and background images from the same frame to pull into columns.

Through the observation of the visual and data results, it is easy to draw the advantages and disadvantages of each method. IALM has a relatively fast speed, but it offers a blind separation of the low-rank and sparse data. Therefore, once the foreground becomes more complex, the detection of background will be obvious interference by the artifacts. More importantly, every time the result of the low-rank approximation matrix basically has reached the rank up to around 100, which in some sense no longer can be called a low-rank matrix. GoDec has two adjustable parameters, "rank" and "card," in the algorithm. Presetting "rank" can guarantee the approximation matrix to be necessarily low-rank, but, as the rank has been fixed already, it can not adaptively adjust according to the situation of scene. And the correct selection of "card" is also very important, which has a great influence on the results. The larger the value of "card" the greater time the algorithm runs, and the background would become messy; the smaller the value of "card" the lower time consumption, but the foreground appears lack of information. Both our methods can adaptively choose

FIGURE 5: Restaurant: (a)–(d) background and (e)–(h) foreground from IALM, GoDec, NSMP, and WNSMP.

TABLE 1: The data comparison of background modeling and foreground detection.

Method	Lobby 128 ∗ 160	Airport 144 ∗ 176	Shopping mall 256 ∗ 320	Restaurant 120 ∗ 160	Campus 128 ∗ 160
IALM					
Time	30.69	42.55	145.97	30.46	30.98
Iteration	33	35	34	34	34
Rank	111	103	91	103	101
Correlation	0.1822	0.3107	0.1900	0.2166	0.1567
GoDec					
Time	40.41	61.30	182.84	56.34	66.59
Iteration	max_iter	max_iter	max_iter	max_iter	max_iter
Rank	2	2	2	3	2
Correlation	0.2518	0.3867	0.2485	0.3135	0.2600
NSMP					
Time	10.87	18.66	60.21	11.21	12.90
Iteration	6	8	7	7	7
Rank	3	4	5	4	5
Correlation	0.1082	0.1059	0.0935	0.1743	0.0819
WNSMP					
Time	9.06	15.79	50.81	11.20	12.63
Iteration	5	7	6	7	7
Rank	3	3	4	4	4
Correlation	0.0895	0.0622	0.0517	0.0917	0.0639

the appropriate low-rank and ensure the most irrelevant foreground and background to be separated. The parameter setting of our methods has been already stated in Section 2, and the parameters of IALM and GoDec are set as in [15, 16].

5. Discussion and Future Work

This paper proposes a "low-rank + dual" model to decompose the static background and the moving foreground in each frame of the video surveillance sequence. Our model decomposes the stack data matrix D into two parts, L low-rank background and M moving foreground, and we upgrade this model further into a reweighted function norm form. Through the experiment, we can clearly observe that our two models can not only achieve clearer backgrounds but also have the lower consumption and correlation. Our "low-rank + dual" model is a better choice for the background modeling, which can always extract clear and complete background in any case. However, since the focus of this paper is to propose the new model and show the feasibility of this model,
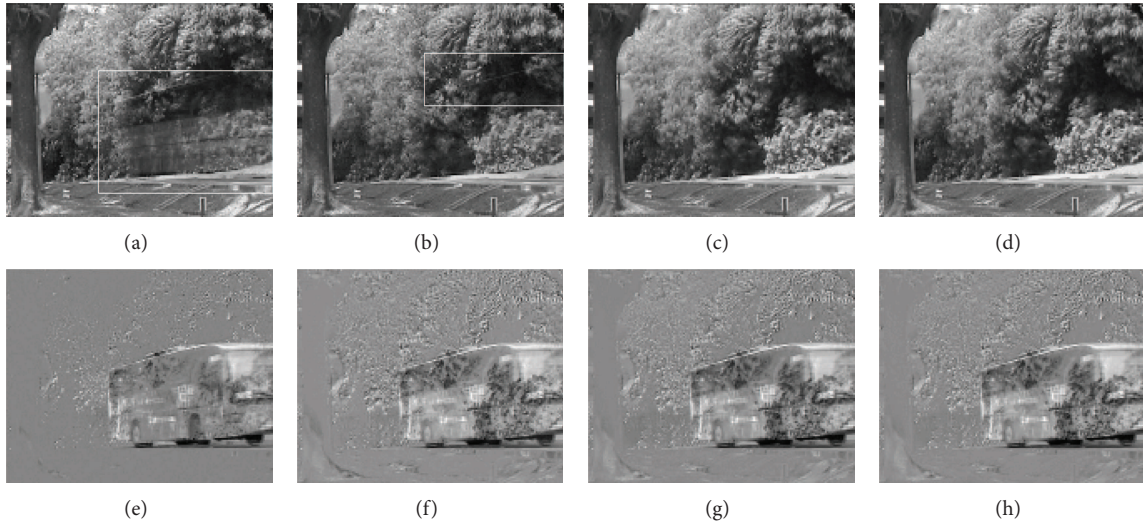
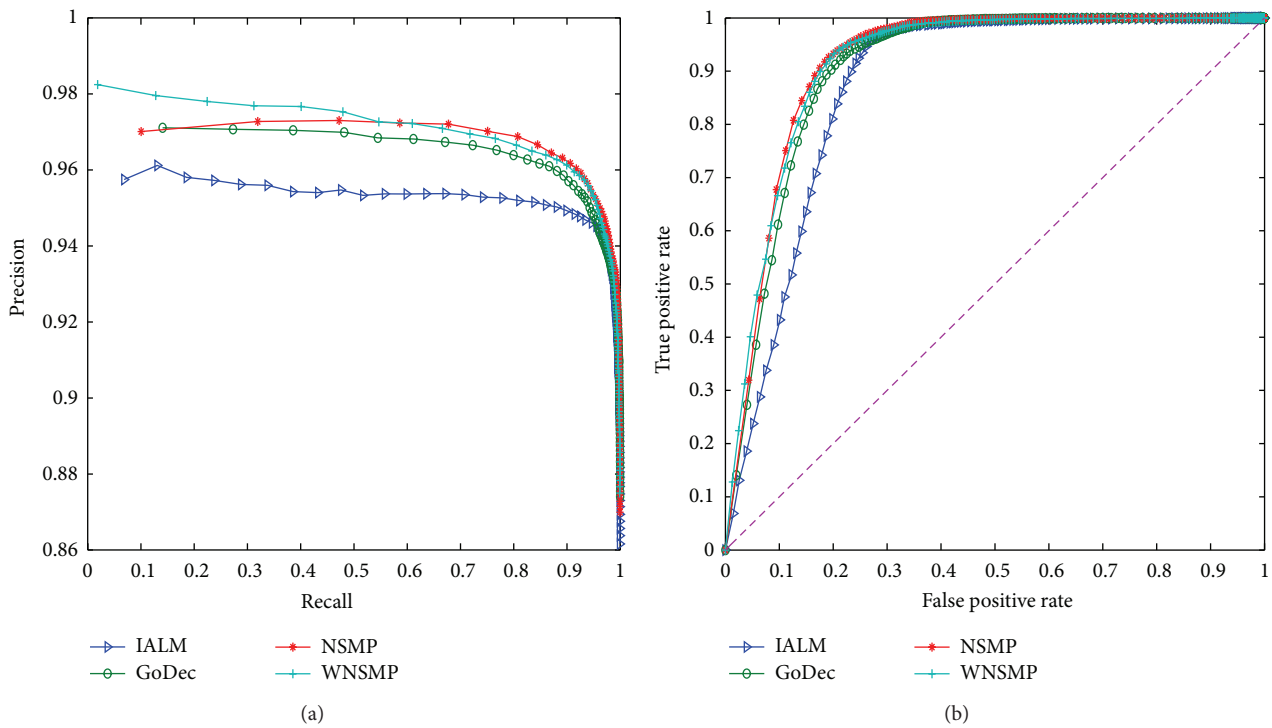FIGURE 6: Campus: (a)–(d) background and (e)–(h) foreground from IALM, GoDec, NSMP, and WNSMP.

FIGURE 7: PR curves and ROC curves for the background modeling of IALM, GoDec, NSMP, and WNSMP.

the algorithm itself has not been optimized which still uses singular value decomposition to get the solution. So, the computation complexity of our algorithm is still $O(n^3)$, and the space complexity is $O(m^2)$. Using the outstanding fast SVD method or other effective methods to speed up our algorithms will be our next work. Also, the further study in the choice of the optimal parameters and weights can be a follow-up process.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, no. 2-3, pp. 90–126, 2006.

[2] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: principles and practice of background maintenance," in *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV'99)*, vol. 1, pp. 255–261, September 1999.

[3] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99)*, vol. 2, pp. 246–252, Fort Collins, Colo, USA, June 1999.

[4] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.

[5] D.-S. Lee, "Effective Gaussian mixture learning for video background subtraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 827–832, 2005.

[6] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, pp. 28–31, Cambridge, UK, August 2004.

[7] A. Elgammal, D. Harwood, and L. S. Davis, "Non-parametric model for background subtraction," in *Computer Vision—ECCV 2000*, pp. 751–767, Springer, Berlin, Germany, 2000.

[8] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151–1162, 2002.

[9] N. M. Oliver, B. Rosario, and A. P. Pentland, "A bayesian computer vision system for modeling human interactions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 831–843, 2000.

[10] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, "Sparse and low-rank matrix decompositions," in *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton '09)*, pp. 962–967, October 2009.

[11] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, article 11, 2011.

[12] F. de La Torre and M. J. Black, "Robust principal component analysis for computer vision," in *Proceedings of the 8th International Conference on Computer Vision (ICCV '01)*, pp. 362–369, July 2001.

[13] J. Wright, A. Ganesh, and S. Rao, "Robust principal component analysis: exact recovery of corrupted low-rank matrices by convex optimization," in *Proceedings of the Neural Information Processing Systems*, vol. 3, 2009.

[14] X. Yuan and J. Yang, "Sparse and low-rank matrix decomposition via alternating direction methods," *Pacific Journal of Optimization*, vol. 9, no. 1, pp. 167–180, 2013.

[15] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," *NIPS*, vol. 2, article 6, 2011.

[16] T. Zhou and D. Tao, "GoDec: randomized low-rank & sparse matrix decomposition in noisy case," in *Proceedings of the 28th International Conference on Machine Learning (ICML '11)*, pp. 33–40, July 2011.

[17] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2233–2246, 2012.

[18] J. W. Demmel, *Applied Numerical linear Algebra*, SIAM, Philadelphia, Pa, USA, 1997.

[19] E. Süli and D. F. Mayers, *An Introduction to Numerical Analysis*, Cambridge University Press, 2003.

[20] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, American Mathematical Society, Boston, Mass, USA, 2001.

[21] A. N. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis*, Courier Dover, 1999.

[22] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.

[23] J. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[24] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[25] J. M. Fadili and G. Peyre, "Total variation projection with first order schemes," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 657–669, 2011.

[26] K. Chen, H. Dong, and K. Chan, "Reduced rank regression via adaptive nuclear norm penalization," *Biometrika*, vol. 100, no. 4, pp. 901–920, 2013.