Computer Science and Information Systems 18(4):1411–1426 https://doi.org/10.2298/CSIS200930029W

Background Modeling from Video Sequences via Online Motion-Aware RPCA

Xu Weiyao[1], Xia Ting[1], and Jing Changqiang[2⋆]

[1] Zaozhuang University
Zaozhuang 277160, Shandong Province, China
{xuweiyao_2008,xiayuxue121}@126.com
[2] Linyi University
Linyi 276000, Shandong Province, China
jingchangqiang@lyu.edu.cn

Abstract. Background modeling of video frame sequences is a prerequisite for computer vision applications. Robust principal component analysis(RPCA), which aims to recover low rank matrix in applications of data mining and machine learning, has shown improved background modeling performance. Unfortunately, The traditional RPCA method considers the batch recovery of low rank matrix of all samples, which leads to higher storage cost. This paper proposes a novel online motion-aware RPCA algorithm, named OM-RPCAT, which adopt truncated nuclear norm regularization as an approximation method for of low rank constraint. And then, Two methods are employed to obtain the motion estimation matrix, the optical flow and the frame selection, which are merged into the data items to separate the foreground and background. Finally, an efficient alternating optimization algorithm is designed in an online manner. Experimental evaluations of challenging sequences demonstrate promising results over state-of-the-art methods in online application.

Keywords: Computer vision, Background modeling, Online RPCA, Truncated nuclear norm.

1. Introduction

Background modeling aims to extract foreground objects from videos, which has been widely used in many fields, such as object detection [46][29][18], object localization [33], and image alignment [28]. It aims to initialize efficient and accurate background modeling from a series of video frames.

Many background modeling methods have been proposed in the past few years. Recently, RPCA [2][17][4]has attracted wide attention in the fields of video surveillance and computer vision, which are based on decomposition of the matrix into sparse and low rank components, and has shown improved performance in background modeling[43]. RPCA decomposes the observed video frame matrix into background and foreground[7][16].

Let $Z \in R^{m\times n}$ be the observational data, which can be represented in matrix form. RPCA attempts to decompose Z as the sum of a sparse matrix F and a low rank matrix B.

$$\min\ \mathrm{rank(B)}+\lambda||\mathrm{F}||_0 \quad \mathrm{s.t.} \quad \mathrm{Z}=\mathrm{B}+\mathrm{F} \tag{1}$$

⋆ Corresponding author

where $|| * ||_0$ denotes l_0 norm, i.e., the number of nonzero elements, which is the regularization term for promoting sparsity. λ is a regularization parameter. Unfortunately, due to the discontinuity of the rank function and the nonconvexity of l_0 norm , the optimization problem (1) is ordinarily NP-hard. Many researchers are seeking suitable alternatives to rank functions[8][6]. In addition, l_0 norm can convert problem (1) into a convex optimization problem. Wright et al. [2] proved that while the sparse matrix F is sufficiently sparse, the low-rank matrix Z can be recovered by solving the following problem:

$$\min \ ||B||_* + \lambda||F||_1 \quad \text{s.t.} \quad \mathrm{Z} = \mathrm{B} + \mathrm{F} \tag{2}$$

where $||B||_*$ is the nuclear norm of B.

The problem (2) is a convex optimization problem, and many algorithms have been proposed for this problem. Unfortunately, most of the algorithms solve this problem in a batch manner. Because all samples loaded in memory during the optimization procedure may have high storage costs, it is especially unacceptable for large-scale sample sets. In addition, while all the samples are collected by the streamlined way, these algorithms cannot update the low dimensional subspace efficiently while a new sample is adding. Each iteration needs to optimize every frame, which may seriously limit the scalability of streaming video.

To solve these problems, the online method of RPCA has recently been proposed. Many online mode algorithms for background modeling have been pursued. Shen et al. [30] adopted max norm as a substitute of the rank function in problem (2) to solve the RPCA problem in an online method.

However, the background in dynamic sequences may include multiple motions, which makes the accurate modeling more challenging. In most previous methods, objects are not moving. Smearing artifacts would be introduced while dealing with slow motion and stationary foregrounds. To be aware of motion, when the background scenes change gradually, many RPCA methods exhibit degraded performance, such as under changing lighting conditions. Moreover, existing methods such as Zhou et al. [46] produced an overwhelming outlier in the low-rank component, when the background was heavily occluded by foreground objects. Javed et al. [15] created a background model by using a modified version of RPCA to generate a low-rank matrix from a set of matrices.

Recently, X. Ye [42] proposed a motion-assisted matrix restoration (RMAMR) model for the separation of background objects from the foreground objects, in which the dense motion fields were incorporated into the framework of the RPCA. J. Yang [40] proposed an online motion-assisted RPCA model for back ground recovery from video sequences. This method is more efficient for memory and is scalable for the long video sequences, which are weighted by motion information.

In Hu [11], a novel norm called the truncated nuclear norm was proposed. The new nuclear norm is subtracted from the sum of several maximal singular values, achieving better rank approximation. Based on this method, F. Cao [3] proposed a new algorithm which was called low-rank and sparse decomposition based on the truncated nuclear norm(LRSD-TNN). B. Hong [10] proposed a novel and online robust principal analysis algorithm via truncated nuclear norm regularization and designed an online optimization scheme in which the matrices were updated alternately.

W. Hu [12] analyzed the problem of mocap data completion based on the truncated nuclear norm. In order to reduce the redundant frames, a simple joint motion detection

and frame selection operation was adopted[31]. Unfortunately, these methods deal with batch data only. In order to meet the needs of dynamic subspace, Hu et al.[5] Proposed an online optimization method to 60 deal with the static camera background scene in the video sequence of each sampling point.

Pan et al.[27] proposed a motion-assisted RPCA model based on matrix factorization, and designed an effective linear alternating direction multiplier method and matrix factorization algorithm to solve the proposed FM-RPCA 65 model. Hu et al.[13] Proposed a non-convex rank approximation RPCA model based on segmentation constraints. Firstly, the original video sequence is divided into three parts by low rank matrix decomposition. Then, a new non-convex function is proposed to constrain the low rank feature[9].

In this paper, we propose an online motion-aware RPCA with a truncated nuclear norm regularization (OM-RPCAT) framework for background modeling. The key idea is to extend the motion-aware low rank matrix approximation methods into an online model. In addition, to get better effect, we added a joint method of frame selection. Motion information from the video sequence is estimated by two methods in this paper. This method replaces the objective of rank minimization by minimizing the truncated kernel norm, which can be represented in a matrix factorization form. Then, we designed an efficient iterative optimization method for implementation [39][14].

The rest of this paper is organized as follows: In Section II, we present the OM-RPCAT scheme and two methods for obtaining the motion estimation matrix. In Section III, we design an efficient optimization algorithm to solve the optimization function. In Section IV, our algorithm is evaluated by experiments. Conclusions are made in Section V.

2. Background Modeling via OM-RPCAT

In this section, the OM-RPCAT model is proposed for background recovery. The main idea of our method is to make a more rigorous approximation to the rank operator and to exploit an online method for solving the optimization problem. Moreover, the background of recovery always suffers from smearing artifacts in areas covered by slow-moving objects. In order to overcome this defect, we combine the motion information and frame selection into the framework to separate the background from the moving objects[19][22][21].

2.1. The Proposed OM-RPCAT Model

In particular, let $Z \in R^{\mathrm{mxn}}$ be the observational data, with $Z = (\mathrm{z}_1,...,\mathrm{z}_n)$, and each z_i expresses a sample. Our goal is to decompose the matrix Z into the low-rank matrix and the sparse matrix. The traditional methods of recovering the two components B and F is solved by solving the following equation:

$$\min \frac{1}{2}||Z - B - F||_F^2 + \lambda_1||B||_* + \lambda_2||F||_1 \tag{3}$$

The truncated nuclear norm minimization is adopted as a more rigorous low-rank constraint on B. Therefore, the objective function for this method becomes:

$$\min \frac{1}{2}||Z - B - F||_F^2 + \lambda_1||B||_T + \lambda_2||F||_1 \tag{4}$$

where $||*||_F$ is the Frobenius norm, $||*||_T$ is the truncated nuclear norm, $||*||_1$ is the l_1 norm, λ_1 and λ_2 are the regularization parameters.

To solve the problem in (4), we usually use iterative optimization methods, such as the augmented Lagrangian multiplier [20] or the accelerated proximal gradient. Unfortunately, these methods are implemented in batch processing. Hence, huge data storage costs are incurred when large data are solved. To overcome this problem, we factorize B as $B=\mathrm{LR}^T$.

Given a matrix B, the relationship [11] between $||B||_T$ and $||B||_*$ is:

$$\begin{aligned} ||B||_T = ||B||_* - \max Tr(UBV^T) \\ ,UU^T = I, VV^T = I \end{aligned} \tag{5}$$

where $Tr(*)$ denotes the trace of the matrix and I stands for the identical matrix.

Then, the nuclear norm can be factorized as follows [30] :

$$||\mathrm{B}||_* = \min_{B=LR^T} \tfrac{1}{2}(||L||_F^2 + ||R||_F^2) \tag{6}$$

where $L \in R^{\mathrm{mxd}}$, $R \in R^{\mathrm{nxd}}$.

From this paper B. Hong [10] , we can obtain the following relationship:

$$\begin{aligned} ||B||_T &= ||B||_* - \sum_{i=1}^{T} \sigma_i(B) \\ &= \frac{1}{2}||L||_F^2 + \frac{1}{2}||R||_F^2 - Tr(UBV^T) \end{aligned} \tag{7}$$

Thus, the problem(4) can be transformed into the following constrained problem:

$$\begin{aligned} &\min \tfrac{1}{2}||Z - LR^T - F||_F^2 + \\ &\quad \lambda_1(\frac{1}{2}||L||_F^2 + \frac{1}{2}||R||_F^2 - Tr(ULR^TV^T)) + \lambda_2||F||_1 \\ &\quad \mathrm{s.t.} UU^T = I, VV^T = I \end{aligned} \tag{8}$$

Our OM-RPCAT model is proposed to separate the foreground and background by joining motion information in the video sequences. We introduce the matrix W to represent motion information, $W \in [0, 1]$. The elements in the matrix W indicate whether the pixels in Z belong to the background. Therefore, the final form of the problem is

$$\begin{aligned} &\min \tfrac{1}{2}||W \circ (Z - LR^T - F)||_F^2 + \\ &\quad \lambda_1(\frac{1}{2}||L||_F^2 + \frac{1}{2}||R||_F^2 - Tr(MR^T)) + \lambda_2||F||_1 \\ &\quad \mathrm{s.t.} UU^T = I, VV^T = I \end{aligned} \tag{9}$$

where $\circ$ denotes the Hadamard product, which is known as the element-wise product. Let $M=\mathrm{V}^TUL$, and we use the fact that Tr(XYZ)=Tr(ZXY):

$$Tr(ULR^TV^T) = Tr(V^TULR^T) \tag{10}$$

For each sample z_i, we can get the approximate value of each sample under dictionary L through Lr_i + f_i. The problem (9) can be decomposed into the sample form:

$$\begin{aligned}&\min\tfrac{1}{2}||w_i \circ (z_i - Lr_i - f_i)||_2^2 + \\ &\quad \lambda_1(\frac{1}{2}||L||_F^2 + \frac{1}{2}\sum_{i=1}^{n}||r_i||_2^2 - \sum_{i=1}^{n} m_i^T r_i) + \lambda_2 \sum_{i=1}^{n}||f_i||_1 \\ &\quad \text{s.t.} UU^T = I, VV^T = I\end{aligned} \tag{11}$$

To simplify this problem, we introduce the $l(z_i, L)$ function:

$$\begin{aligned}l(z_i, L) = \min &\frac{1}{2}||w_i \circ (z_i - Lr_i - f_i)||_2^2 + \\ &\frac{\lambda_1}{2}||r_i||_2^2 - \lambda_1 m_i^T r_i + \lambda_2||f_i||_1\end{aligned} \tag{12}$$

The problem (11) can be solved by minimizing the following loss function:

$$f_n(L) = \frac{1}{n}\sum_{k=1}^{n} l(z_i, L) + \frac{\lambda_1}{2n}||L||_F^2 \tag{13}$$

where z_i is the current frame, r_i represents the coefficient under the dictionary L, and w_i is the ith row of W. $l(z_i, L)$ is the loss function for each sample.

2.2. Motion Estimation

In this section, the motion matrix W is generated by computing the video sequence Z. We use two methods to construct the weighting matrix W. The first method is to adopt the optical flow [1], and the second method involves joining the motion detection and frame selection [31][34][35][36].

A. The first method involves using optical flow to obtain the motion estimation matrix.

Assume that z_i and z_{i+1} are the two consecutive frames of a sequence Z. Then, the horizontal V^x and vertical V^y motion vector components can be obtained by estimating the optical flow field. The motion map W is constructed as follows[44][45]:

$$w_{i,j} = \begin{cases} 0, if \sqrt{(v_{i,j}^x)^2 + (v_{i,j}^y)^2} \geq \tau \\ 1, otherwise \end{cases} \tag{14}$$

where τ is the threshold of motion magnitude according to the average intensity of the motion field determined experimentally. $v_{i,j}^y$ and $v_{i,j}^x$ are entries of V^y and V^x in the vertical motion fields and horizontal motion fields respectively.

B. The second method is to add the frame selection.

This scheme aims to reduce the number of redundant frames [31] . The index of the relevant frames is given as follows:

$$y_i = \begin{cases} 1, if |\hat{z} - \hat{\mu}| \geq \tau \\ 0, otherwise \end{cases} \tag{15}$$

where $\hat{\mu}$ denotes the mean value of the vector $\hat{z}$ which is gradiented and regularized for each frame [31]. τ controls the threshold. When the label y_i is 1, the corresponding frame

will be selected. Due to the removal of invalid data frames, this method can restore the background of the video more accurately[23][24][26].

While the frame selection process is complete, the motion estimation of the selected frames is determined by:

$$w_k(i,j) = \begin{cases} 0, if \frac{1}{2}(D_k(i,j))^2 \geq \beta \\ 1, \qquad otherwise \end{cases} \tag{16}$$

Differing from this method [31] , β is the thresholding parameter, D_k is the difference between the two consecutive frames, and z_t denotes the t-th frame:

$$D_t = \sqrt{(Z_t - Z_{t-1})^2} \tag{17}$$

3. Optimization Method.

We propose an algorithm to minimize the average cost, which could solve our OM-RPCAT model in detail. The coefficient r_t, sparse error f_t, motion estimation matrix w_t, and basis L are optimized using alternative methods. The optimization procedure is divided into two steps[37][38][41]:

In the first step, we optimize coefficient r_t, sparse error f_t and motion estimation matrix w_t. The w_t has been illustrated in Section 2.2.

$$\begin{aligned} \{r_t, f_t\} = \arg\min \tfrac{1}{2}||w_i \circ (z_t - L_{t-1}r - f)||_2^2 + \\ \lambda_1(\frac{1}{2}||r||_2^2 - m_t^T r) + \lambda_2||f||_1 \end{aligned} \tag{18}$$

where w_i denotes the motion map for the current frame z_t.

Update r_t: Given the current basis L, the coefficient r_t can be obtained by the following formula:

$$f(r) = \tfrac{1}{2}||w_i \circ (z_t - L_{t-1}r - f_t^k)||_2^2 + \lambda_1(\tfrac{1}{2}||r||_2^2 - m_t^T r) \tag{19}$$

Let $\partial f / \partial r = 0$, and we can obtain the following solution:

$$\begin{aligned} & r_t^{k+1} = \\ & (L_{t-1}^T \hat{W}_t^T \hat{W}_t L_{t-1} + \lambda_1 I)^{-1}(L_{t-1}^T \hat{W}_t^T \hat{W}_t(z_t - f_t^k) + \lambda_1 m_t) \end{aligned} \tag{20}$$

where $\hat{W}$ is a diagonal matrix formed by placing the elements on the diagonal.

Update f_t: The objective formula to optimize f_t induced from problem(18) is:

$$g(f) = \tfrac{1}{2}||w_i \circ (z_t - L_{t-1}r_t^{k+1} - f)||_2^2 + \lambda_2||f||_1 \tag{21}$$

We used the common approach presented in E. J. Candes [28] to solve this problem and obtained the following closed formula:

$$f_t^{k+1} = S_{\lambda_2}[w_i \circ (z_t - L_{t-1}r_t^{k+1})] \tag{22}$$

where $S_\tau[*]$ is a shrinkage operator, which is defined as

$$S_\tau[x] = sign(x)\max(|x| - \tau, 0) \tag{23}$$

In the second step, we optimize the basis matrices L_t, V_t, and U_t under the previously obtained r_k, f_k, and w_k. Based on problem(9), the objective function is derived as follows:

$$\begin{aligned}\{L_t, U_t, V_t\} = &\min \tfrac{1}{2}||W \circ (Z - LR^T - F)||_F^2 + \\ &\lambda_1(\frac{1}{2}||L||_F^2 - Tr(MR^T))\end{aligned} \tag{24}$$

Because of the Hadamard product, it is difficult to calculate the matrix multiplication for this formula. We introduce an additional variable Y:

$$Y_k = Z_k - L_k R_k^T - F_k \tag{25}$$

The above equation is converted to the following equation:

$$\begin{cases} Y_k^{l+1} = \arg\min\limits_{Y} \frac{1}{2}||W_k \circ Y_k^l||_F^2 + \lambda_1(\frac{1}{2}||L_k^l||_F^2 - \\ Tr(M_k R_k^T)) + \frac{\lambda_3}{2}||Y_k^l - Z_k + L_k^l R_k^T + F_k||_F^2 \\ L_k^{l+1} = \arg\min\limits_{L} \frac{1}{2}||W_k \circ Y_k^{l+1}||_F^2 + \lambda_1(\frac{1}{2}||L_k^l||_F^2 - \\ Tr(M_k R_k^T)) + \frac{\lambda_3}{2}||Y_k^{l+1} - Z_k + L_k^l R_k^T + F_k||_F^2 \end{cases} \tag{26}$$

where λ_3 is a constrained parameter. After removing the irrelevant items, we can get the following:

Update Y:

$$\begin{aligned}Y_k^{l+1} = &\arg\min_{Y} \tfrac{1}{2}||W_k \circ Y_k^l||_F^2 + \\ &\tfrac{\lambda_3}{2}||Y_k^l - Z_k + L_k^l R_k^T + F_k||_F^2\end{aligned} \tag{27}$$

We can compute Y in a pixel-wise manner:

$$y_{ik} = \frac{\lambda_3}{\lambda_3 + w_{ik}^2}(z_{ik} - f_{ik} - l_i r_k) \tag{28}$$

where l_i denotes the ith row of matrix L.

Update L:

$$\begin{aligned}L_k^{l+1} = &\arg\min_{L} \lambda_1(||\tfrac{1}{2}L_k^l||_F^2 - Tr(M_k R_k^T)) + \\ &\tfrac{\lambda_3}{2}||Y_k^{l+1} - Z_k + L_k^l R_k^T + F_k||_F^2\end{aligned} \tag{29}$$

where $M = V^T U L \in R^{n \times d}$.

The matrix L can be updated via the closed-form solution of the least square problem in Eq.(29):

$$L_k = (\lambda_3(Z - F - Y)R + U^T V R)(\lambda_1 I + \lambda_3 R^T R) \tag{30}$$

U_t is optimized by the following formula [10]:

$$U_t = \arg\ \max \mathrm{Tr}(\mathrm{U}\mathrm{L}_t\mathrm{R}_t^T\mathrm{V}_{t-1}^T) \text{ s.t. } \mathrm{U}\mathrm{U}^T = I \tag{31}$$

Similarly, V_t is optimized by the following formula [10]:

$$V_t = \arg\ \max \mathrm{Tr}(\mathrm{VR}_t\mathrm{L}_t^T\mathrm{U}_t^T) \text{ s.t. } \mathrm{VV}^T = I \tag{32}$$

The algorithm is shown as Algorithm 1.

Algorithm 1: OM-RPCAT Algorithm

Input:
The observed data:$Z = [z_1, z_2, ..., z_n] \in R^{m\times n}$
matrix: $L_0 \in R^{m\times d}$, $U_0 \in R^{s\times m}$, $V_0 \in R^{s\times n}$
regularization parameters:$\lambda_1, \lambda_2, \lambda_3 \in \mathrm{R}$
motion matrix:$W \in R^{m\times n}$
number of frames: t
Output: L_n, R_n

```
for t = 1 to n do
    coefficient r_t = 0
    sparse error f_t = 0
    m_t is is the t-th row of V_{t-1}^T U_{t-1} L_{t-1}
    compute motion matrix w_t by Eq.(14) or Eq.(16)
    Stage 1: compute r_t and f_t
    while not converged do
        Update the coefficient r_t by Eq.(20).
        Update the sparse error f_t by Eq.(22).
    end
    Stage 2: compute L_t, U_t and V_t
    while not converged do
        Update the additional matrix Y_t by Eq.(28).
        Update the matrix L_t by Eq.(30).
        Update the U_t by Eq.(31).
        Update the V_t by Eq.(32).
    end
end
```

4. Experimental Results and Discussions

In this section, we investigate the performance of the proposed method. All the experiments were implemented on a PC with an Intel Core i5 CPU at 2.4 GHz and with 16 GB of memory. Simulations were performed using MATLAB 2014a.

Experiments are performed on scene background initialization (SBI) dataset [25][3]. The SBI dataset, consisting of fourteen image sequences with ground truth images and a set of commonly adopted metrics, with a wide range of complex backgrounds and different situations, including a variety of sequences, such as HallMonitor, Board, CAVIAR1, CaVignal and HumanBody2, and it has been adopted by many existing and new background initialization methods. For the convenience of comparison, the ground truth images of the background are also provided.

[3] http://sbmi2015.na.icar.cnr.it/

For detailed qualitative comparisons, the proposed method is evaluated by comparing it with the online RPCA method [32] , OMA-RPCA method [40] , and OTNNR method [10] . These methods are state-of-the-art techniques for online background modeling.

We implemented the following settings: $\lambda_1 = 0.1, \lambda_2 = 1, \lambda_3 = 0.1$, the convergence error is set to 10e-4, the rank of the low dimensional subspace is 5, the threshold τ is set by the average intensity of the optical flow. We implemented these four algorithms on the more challenging moving background video sequences, i.e., HallMonitor, Board, CAVIAR1, CaVignal and HumanBody2. Table 1 describes of the information of datasets.

Table 1. Brief introduction of the datasets

Dataset	Size	Number of frames
HallMonitor	352*240	299
Board	200*164	227
CAVIAR1	384*256	609
CaVignal	200*136	257
HumanBody2	320*240	740

4.1. Qualitative Results and Comparison

In this paper, we use two methods to get the motion matrix. The first method is to adopt the optical flow [1] , and the second method is to apply joint motion detection and frame selection [31]. For these two methods of obtaining the motion matrix, the results of our proposed method and the existing methods are shown in Fig. 1 and Fig. 2 respectively.

Due to the slow moving foreground takes up many areas of the frame, it used to be severely tailed by using the previous methods to recover the true background, such as online RPCA, OMA-RPCA and OTNNR. The slow moving foreground may be considered as part of the background. From the results as shown in Fig. 1 and Fig. 2, we can see that our method is the closest to the real background and is superior to other algorithms. Fig. 1 and Fig. 2 show the background modeling results of frame selection mode and optical flow mode respectively. Because of abandoning the redundant frames in the frame selection technology, the effect of background modeling is much better than that of optical flow methods. But for video object with long static foreground, the effect of the two methods are both not very good.

4.2. Quantitative Evaluations and Analysis

In order to quantitatively analyze the performance of the algorithm and verify the rationality of the algorithm, we measured these methods with the PSNR(Peak Signal-to-Noise Ratio) and RRE(Relative Reconstruction Error), which is defined as follows:

$$RRE = ||\hat{A} - A||_F / ||A||_F \tag{33}$$

Fig. 1. Background modeling results by frame selection method: (a) true backgrounds, (b) OM-RPCAT(Ours), (c) OMA-RPCA, (d) online RPCA, (e) OTNNR. From top to bottom, the recovered backgrounds for HallMonitor, Board, CAVIAR1, CaVignal and HumanBody2 are presented respectively.

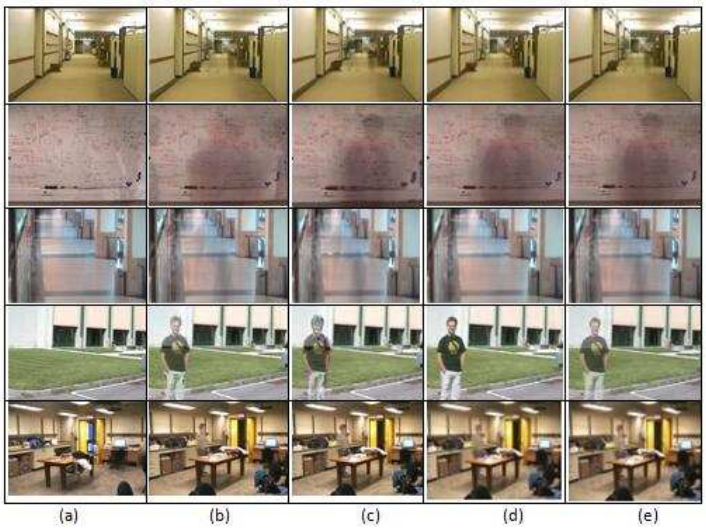

Fig. 2. Background modeling results by optical flow method:(a) true backgrounds, (b) OM-RPCAT(Ours), (c) OMA-RPCA, (d) online RPCA, (e) OTNNR. From top to bottom, the recovered backgrounds for HallMonitor, Board, CAVIAR1, CaVignal and HumanBody2 are presented, respectively.

$$PSNR = 10 * \log_{10}(255^2/MSE) \tag{34}$$

where the MSE is the mean square error of the real background and the reconstruction background.

Table 2. Quantitative background modeling results by frame selection

	PSNR	RRE
OM-RPCAT(OURS)	83.6	0.014
OMA-RPCA[40]	83.4	0.015
OTNNR[10]	75.8	0.021
Online RPCA[32]	76.1	0.019

Table 3. Quantitative background modeling results by optical flow

	PSNR	RRE
OM-RPCAT(OURS)	79.3	0.024
OMA-RPCA[40]	74.1	0.023
OTNNR[10]	75.7	0.027
Online RPCA[32]	76.3	0.020

In order to ensure the fairness of verification, we all use the HallMonitor dataset. Based on the results of Table 2 and Table 3, our method can get the best PSNR. In addition, the frame selection method obtain greater performance than optical flow.

4.3. Implementation Details and Computational Time

The solution of the proposed models require a set of parameters including λ_1, λ_2, λ_3, τ and d. The d represents the rank of the low dimensional subspace. In order to update the model quickly, in the previous experiment, we set the rank d to 5. Now we analyze the performance impact of rank d on background modeling, HallMonitor dataset was used in this experiment. The frame selection method is selected for the motion estimation.

It can be seen from Fig. 3 that with the change of rank d, the PSNR also changes. When the d is set to 5, the performance of background modeling is the best.

In the experiment, we also study the time complexity problem. To compare the overall computational time, we selected a short sequence named HallMonitor. For a fair comparison with the other methods, the time is recorded in seconds. The frame selection method is selected for the motion estimation. Fig. 4 presents the performance in terms of computational time. Compared with the previous algorithms, although our method does not achieve the most promising results, considering the final performance, our algorithm is better.

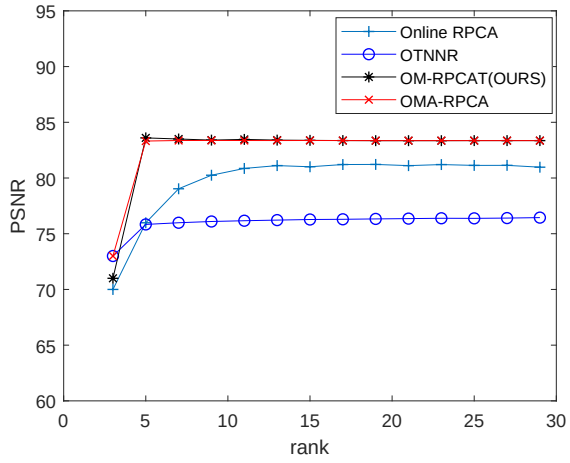

Fig. 3. PSNR analysis of the rank of the low dimensional subspace d

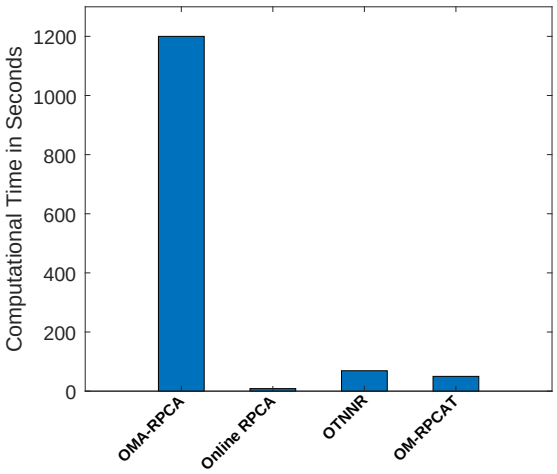

Fig. 4. Comparison of computational time in seconds

5. Conclusions

In this paper, we propose a novel model for background modeling from a given sequence of video frames and adopt the truncated nuclear norm as the convex optimization framework. In order to perceive motion information in video, we use motion information as the weighting matrix. In addition, the low-rank approximation is optimized to online mode, which is more efficient for the online video. Additionally, to achieve better effects and improve the processing efficiency, we introduced two methods to get the motion estimation matrix. The first method is to adopt the optical flow, and the second method involves joining the motion detection and frame selection which can use the frame selection technique to abandon the redundant frames. We further designed an online optimization scheme to solve the matrix decomposition problem with weighted matrix. Experimental results demonstrate that the proposed algorithm outperforms existing online RPCA algorithms significantly. Considering the limitations of RPCA, our future work will focus on designing more efficient background modeling algorithms.

Acknowledgments. This work is supported by the Foundation of Zaozhuang University (No. 102061808 and No. 102061804), in part by Key Research and Development Project of Shandong Province (No. 2019JZZY010134) and National Natural Science Foundation of China (No. 61901206).

References

1. Brox, T., Malik, J., Fellow, IEEE: Large displacement optical flow: Descriptor matching in variational motion estimation. IEEE Transactions on Pattern Analysis & Machine Intelligence 33(3), 500–513 (2011)
2. Candes, E.J., Xiaodong, L.I., Yl, M.A., Wright, J.: Robust principal component analysis? Journal of the ACM (JACM) 58(3) (2011)
3. Cao, F., Chen, J., Ye, H., Zhao, J., Zhou, Z.: Recovering low-rank and sparse matrix based on the truncated nuclear norm. Neural Networks 85 (2017)
4. Fu, H., Gao, Z., Liu, H.Z.: Fast robust pca on background modeling. In: Chinese Intelligent Systems Conference (2018)
5. Fu, H., Wang, B., Liu, H.Z.: Online RPCA on Background Modeling: Volume II (2019)
6. Garg, K., Ramakrishnan, N., Prakash, A., Srikanthan, T.: Rapid and robust background modeling technique for low-cost road traffic surveillance systems. IEEE Transactions on Intelligent Transportation Systems PP(99), 1–12 (2019)
7. Guangming, S., Tao, H., Weisheng, D., Jinjian, W., Xuemei, X.: Robust foreground estimation via structured gaussian scale mixture modeling. IEEE Transactions on Image Processing pp. 1–1 (2018)
8. Guian, Zhang, Zhiyong, Yuan, Qianqian, Tong, Qiong, Wang: A novel and practical scheme for resolving the quality of samples in background modeling. Sensors (2019)
9. Guo, Wenzhong, Zhang, Qishan, Chen, Yuzhong, Kun, Qiu, Qirong: Community discovery by propagating local and global information based on the mapreduce model. Information Sciences: An International Journal (2015)
10. Hong, B., Wei, L., Hu, Y., Cai, D., He, X.: Online robust principal component analysis via truncated nuclear norm regularization. Neurocomputing 175(JAN.29PT.A), 216–222 (2016)
11. Hu, Yao, Zhang, Debing, Ye, Jieping, Li, Xuelong: Fast and accurate matrix completion via truncated nuclear norm regularization. IEEE Transactions on Pattern Analysis & Machine Intelligence 35(9), 2117–2130 (2013)

12. Hu, W., Wang, Z., Liu, S., Yang, X., Yu, G., Zhang, J.J.: Motion capture data completion via truncated nuclear norm regularization. IEEE Signal Processing Letters pp. 1–1 (2017)
13. Hu, Z., Wang, Y., Su, R., Bian, X., Wei, H., He, G.: Moving object detection based on non-convex rpca with segmentation constraint. IEEE Access 8, 41026–41036 (2020)
14. Huang, Z., Yu, Y., Gu, J., Liu, H.: An efficient method for traffic sign recognition based on extreme learning machine. IEEE Transactions on Cybernetics 47(4), 920–933 (2016)
15. Javed, S., Mahmood, A., Bouwmans, T., Jung, S.K.: Motion-aware graph regularized rpca for background modeling of complex scenes. In: 2016 23rd International Conference on Pattern Recognition (ICPR) (2017)
16. Le, H., Wei, W., Li, G., Mathematics, S.O.: Compressed video background/foreground recovery and separation based on ptv-tv tensor modeling. Journal of South China University of Technology(Natural Science Edition) (2019)
17. Li, H., Miao, Z., Li, Y., Wang, J., Zhang, Y.: Background subtraction via online box constrained rpca. In: 2018 International Conference (2018)
18. Li, Y., Liu, G., Liu, Q., Sun, Y., Chen, S.: Moving object detection via segmentation and saliency constrained rpca. Neurocomputing 323 (2018)
19. Lin, B., Guo, W., Lin, X.: Online optimization scheduling for scientific workflows with deadline constraint on hybrid clouds. Concurrency and Computation: Practice & Experience (2016)
20. Lin, Z., Chen, M., Ma, Y.: The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices (2010)
21. Liu, GG, Huang, Guo, WZ, Niu, YZ, Chen, GL: Multilayer obstacle-avoiding x-architecture steiner minimal tree construction based on particle swarm optimization. IEEE T CYBERNETICS 2015,45(5)(-), 989–1002 (2015)
22. Liu, G., Guo, W., Niu, Y., Chen, G., Huang, X.: A pso-based timing-driven octilinear steiner tree algorithm forvlsi routing considering bend reduction. Soft Computing 19(5), 1153–1169 (2015)
23. Luo, F., Guo, W., Yu, Y., Chen, G.: A multi-label classificationalgorithm based on kernel extreme learning machine. Neurocomputing 260(oct.18), 313–320 (2017)
24. Ma, T., Liu, Q., Cao, J., Tian, Y., Al-Rodhaan, M.: Lgiem: Global and local node influence based community detection. Future Generation Computer Systems 105 (2019)
25. Maddalena, L., Petrosino, A.: Towards benchmarking scene background initialization. In: International Conference on Image Analysis and Processing (2015)
26. Niu, Y., Chen, J., Guo, W.: Meta-metric for saliency detection evaluation metrics based on application preference. Multimedia Tools & Applications 77(20), 1–19 (2018)
27. Pan, Peng, Wang, Yongli, Zhou, Mingyuan, Sun, Zhipeng, Guoping: Background recovery via motion-based robust principal component analysis with matrix factorization. Journal of Electronic Imaging 27(2), 23034.1 (2018)
28. Peng, Y., Ganesh, A., Wright, J., Xu, W., Ma, Y.: Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images. IEEE Transactions on Pattern Analysis & Machine Intelligence 34(11), 2233–46 (2012)
29. Sajid, Javed, Arif, Mahmood, Somaya, Al-Maadeed, Thierry, Bouwmans, Ki, S., Jung: Moving object detection in complex scene using spatiotemporal structured-sparse rpca. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society (2018)
30. Shen, J., Xu, H., Li, P.: Online optimization for max-norm regularization. Machine Learning 2(3), 419–457 (2017)
31. Sobral, A., Bouwmans, T., Zahzah, E.H.: Comparison of matrix completion algorithms for background initialization in videos. In: Scene Background Modeling and Initialization (SBMI) Workshop in conjunction with ICIAP 2015 (2015)
32. Song, W., Zhu, J., Li, Y., Chen, C.: Image alignment by online robust pca via stochastic gradient descent. IEEE Transactions on Circuits & Systems for Video Technology 26(7), 1241–1250 (2016)

33. Stauffer, C.: Adaptive background mixture model for real-time tracking. Proc Cvpr 2, 2246 (1998)
34. Wang, J., Zhang, X.M., Lin, Y., Ge, X., Han, Q.L.: Event-triggered dissipative control for networked stochastic systems under non-uniform sampling. Information Sciences p. S0020025518301749 (2018)
35. Wang, S., Guo, W.: Robust co-clustering via dual local learning and high-order matrix factorization. Knowledge-Based Systems 138 (2017)
36. Wu, T.Y., Chen, C.M., Wang, K.H., Meng, C., Wang, E.K.: A provably secure certificateless public key encryption with keyword search. Journal of the Chinese Institute of Engineers pp. 1–9 (2019)
37. Xia, Y., Chen, T., Shan, J.: A novel iterative method for computing generalized inverse. Neural Computation 26(2), 449C465 (2014)
38. Xia, Y., Wang, J.: Low-dimensional recurrent neural network-based kalman filter for speech enhancement. Neural Networks 67, 131–139 (2015)
39. Xing, H., Guo, W., Liu, G., Chen, G.: Fh-oaos: A fast four-step heuristic for obstacle-avoiding octilinear steiner tree construction. Acm Transactions on Design Automation of Electronic Systems 21(3), 1–31 (2016)
40. Yang, J., Yang, J., Yang, X., Yue, H.: Background recovery from video sequences via online motion-assisted rpca. In: 2016 Visual Communications and Image Processing (VCIP) (2017)
41. Yang, L.H., Yang, L. H., K.S.: Multi-attribute search framework for optimizing extended belief rule-based systems. Information Sciences (2016)
42. Ye, X., Yang, J., Sun, X., Li, K., Hou, C., Wang, Y.: Foreground-background separation from video clips via motion-assisted matrix restoration. IEEE Transactions on Circuits & Systems for Video Technology 25(11), 1721–1734 (2015)
43. Yinxiao, Zhan, Ting, Liu: Weighted rpca based background subtraction for automatic berthing (2019)
44. Zhang, S., Xia, Y.: Two fast complex-valued algorithms for solving complex quadratic programming problems. IEEE Transactions on Cybernetics 46(12), 2837–2847 (2015)
45. Zhang, S., Xia, Y., Wang, J.: A complex-valued projection neural network for constrained optimization of real functions in complex variables. IEEE Transactions on Neural Networks & Learning Systems 26(12), 3227–3238 (2015)
46. Zhou, X., Yang, C., Yu, W.: Moving object detection by detecting contiguous outliers in the low-rank representation. IEEE Transactions on Pattern Analysis & Machine Intelligence 35(3), 597–610 (2013)

Xu Weiyao received the master's degree in communication and information systems from Nanjing University of Posts and Telecommunications. He is currently pursuing the Ph.D. degree with the Beijing University of Posts and Telecommunications. His current research interests include around machine learning and human action recognition.

Xia Ting received the master's degree in Communication and Information Systems from Hangzhou University of Electronic Science and Technology. She is currently a lecturer at Zaozhuang University. Her current research interests include machine learning and human action recognition.

Jing Changqiang received Ph.D. degrees from Kwangwoon University, Korea in 2015. His research interests include computer vision, and wireless sensor network.

Received: September 30, 2020; Accepted: April 18, 2021.