Robust Background Subtraction via Online Robust PCA using Image Decomposition

Sajid Javed∗
School of Computer Science and Engineering
Kyungpook National University
80 Daehak-ro, Buk-gu, Daegu, 702-701, Republic of Korea
sajid@vr.knu.ac.kr

Seon Ho Oh
School of Computer Science and Engineering
Kyungpook National University
80 Daehak-ro, Buk-gu, Daegu, 702-701, Republic of Korea
shoh@vr.knu.ac.kr

JunHyeok Heo
IDIS Co., Ltd.
IDIS Tower 1F, 344, Pangyo-ro, Bungdang-gu
Seongnam-si, Gyeonggi-do, 463-400, Republic of Korea
jhheo@idis.co.kr

Soon Ki Jung†
School of Computer Science and Engineering
Kyungpook National University
80 Daehak-ro, Buk-gu, Daegu, 702-701, Republic of Korea
skjung@knu.ac.kr

ABSTRACT
Accurate and efficient background subtraction is an important task in video surveillance system. The task becomes more critical when the background scene shows more variations, such as water surface, waving trees and lighting conditions, etc. Recently, Robust Principal Components Analysis (RPCA) shows a nice framework for moving object detection. The background sequence is modeled by a low-dimensional subspace called low-rank matrix and sparse error constitutes the foreground objects. But RPCA presents the limitations of computational complexity and memory storage due to batch optimization methods, as a result it is hard to apply for real-time system. To handle these challenges, this paper presents a robust background subtraction algorithm via Online Robust PCA (OR-PCA) using image decomposition. OR-PCA with image decomposition approach improves the accuracy of foreground detection and the computation time as well. Comprehensive simulations on challenging datasets such as Wallflower, I2R and Change Detection 2014 demonstrate that our proposed scheme significantly outperforms the state-of-the-art approaches and works effectively on a wide range of complex background scenes.

∗Mr. Javed is a first author.
†Prof. Jung is a corresponding author.

Categories and Subject Descriptors
I.4.9 [Image Processing and Computer Vision]: Applications.

General Terms
System, Algorithm

Keywords
Online Robust PCA, Image Decomposition, Low-rank matrix, Foreground detection.

1. INTRODUCTION
Background subtraction (also known as foreground detection) is one of the most important preprocessing step in many computer vision applications. Typically, the background subtraction process forms the first stage in automated visual surveillance systems, as well as other applications such as motion capture, object tracking and augmented reality.

Many algorithms have been developed to handle the problem of background subtraction in videos [1]. In recent years, Robust Principal Component Analysis (RPCA) based low-rank matrix decomposition algorithms have been used for background subtraction [2]. RPCA decomposes the original data matrix A as a sum of low-dimensional subspace called low-rank matrix L and correlated sparse outliers S. Fig. 1 shows an example of background subtraction using RPCA.

As RPCA based approaches provide a nice framework for background modeling, but it currently faces two major difficulties. Traditional RPCA based approaches use batch optimization, e.g. in order to decompose low-rank and sparse components, a number of samples are required to store. Therefore, it suffers from high memory cost and computational complexity.

In order to tackle these issues, this paper presents a robust background subtraction algorithm via Online Robust
sparse matrix. Semi-Soft GoDec [14] method is an extension to separate the dom projections Zhou et al literature to accelerate the PCP algorithms. For example, real-time system. Pursuit (PCP) via batch optimization is not acceptable for complexity problems occur. Therefore Principal Component optimization manners, as a result memory storage and time sparse error and Zhou et al [15] of GoDec which is four times faster than GoDec. In [15], accelerates RPCA algorithm via PCP using bilateral ran-

...low-rank representation (DECOLOR) method, which accel-

2. RELATED WORK

Over the past few years, excellent methods have been pro-

posed for foreground object detection and tracking. Among them, RPCA [5] shows promising results for background modeling. Excellent surveys on background modeling using RPCA can be found in [2, 7]. All these RPCA approaches, such as Augmented Lagrangian Multiplier (ALM) and Singular Value Thresholding (SVT) etc. discussed in [2], solve the sub-optimization problem in each iteration under defined convergence criteria, to separate the low-rank matrix and sparse error. All these methods work in a batch optimization manners, as a result memory storage and time complexity problems occur. Therefore Principal Component Pursuit (PCP) via batch optimization is not acceptable for real-time system.

Many noticeable improvements have been found in the literature to accelerate the PCP algorithms. For example, Zhou et al. [14] proposed Go Decomposition (GoDec) which accelerates RPCA algorithm via PCP using bilateral random projections (BRP) scheme to separate the low-rank and sparse matrix. Semi-Soft GoDec [14] method is an extension of GoDec which is four times faster than GoDec. In [15], Zhou et al. proposed Detecting Contiguous Outliers in the low-rank representation (DECOLOR) method, which accelerates PCP algorithm by integrating the object detection and background learning into a single process of optimization. In [11], a fast PCP algorithm is proposed, which reduces the SVD computational time in inexact ALM (IALM) by adding some constants in the minimization equation of low-rank and sparse, but it is based on PCP.

Incremental and online robust PCA methods are also developed for PCP algorithms. Fore example, in [8], He et al. proposed Grassmanian Robust Adaptive Subspace Tracking Algorithm (GRASTA), which is an incremental gradient descent method on Grassmannian manifold for solving the RPCA problem in online manner. Results are encouraging for background modeling, but no theoretic guarantee of the algorithm convergence for GRASTA is provided. Therefore, in [10], an online learning method for sparse coding and dictionary learning is proposed which efficiently solves the smooth nonconvex objective function over a convex set. A real-time processing is achieved, but it does not require learning rate tuning like regular stochastic gradient descents . In [6], Guan et al. proposed non-negative matrix factorization (NMF) method which receives one chunk of samples per step. But, using a buffering strategy both time complexity and space remains the issue in this approach.

Therefore, Feng and Xu [4] recently proposed Online Robust PCA (OR-PCA) algorithm which processes one sample per time instance using stochastic approximations. A nuclear norm objective function is reformulated in this approach and therefore, all the samples are decoupled in optimization process for sparse error separation. In this paper, OR-PCA is modified to be adapted for robust background subtraction.

3. METHODOLOGY

In this section, we discuss our scheme for background modeling in detail. Our methodology consists of three stages: decomposition, background modeling and integration (see Fig. 2). Initially, the input video frames are decomposed into Gaussian and Laplacian images using a set of two Gaussian kernels. Then, OR-PCA is applied for each of Gaussian and Laplacian image with different parameters to model the background, separately. In the background modeling stage, we have proposed an alternative initialization scheme to speed up the stochastic optimization process. Finally, the integration stage, which combines low-rank and sparse components obtained via OR-PCA to recover the final background and foreground image, is performed. The reconstructed sparse matrix is thresholded to get the binary foreground mask. In the following sections, we will describe each module in detail.

3.1 Decomposition

In the first stage, two separate spatial Gaussian kernels are designed to decompose input image into Gaussian and Laplacian images. First, Gaussian kernels are applied on input image to get the Gaussian images. In the first case, we choose the standard deviation $\sigma$ on the Gaussian kernel as 2 with a filter size of $5 \times 5$ to get the first Gaussian image. In the second case, we apply Gaussian kernel with a same $\sigma$ value on the first blurred image due to its smoothing properties for reducing false alarms. Since the difference of Gaussians is approximately same as Laplacian of Gaussian, Laplacian image can be obtained by the difference of two Gaussian images.
Every input video frame is decomposed into Gaussian and Laplacian images using the method discussed above. As Gaussian image is robust against background dynamics and Laplacian image provides enough edge features for small pixels variations. Therefore, the false alarms are reduced from foreground region as a result, our methodology provides accurate foreground detection.

3.2 Background Modeling

Online Robust PCA \[4\] is used to model the background from Gaussian and Laplacian images. OR-PCA decomposes the nuclear norm of the objective function of the traditional PCP algorithms into an explicit product of two low-rank matrices, i.e., basis and coefficient. Thus, OR-PCA can be formulated as

\[
\begin{align*}
\min_{L \in \mathbb{R}^{p \times n}, R \in \mathbb{R}^{n \times r}, E} & \left\{ \frac{1}{2} \| Z - LR^T - E \|_F^2 \right. \\
+ & \left. \frac{\lambda_1}{2} (\| L \|_F^2 + \| R \|_F^2) + \lambda_2 \| E \|_1 \right\}.
\end{align*}
\]

where \( Z \) is an input data, \( L \) is a basis, \( R \) is a coefficient and \( E \) is a sparse error. \( \lambda_1 \) controls the basis and coefficients for low-rank matrix, whereas \( \lambda_2 \) controls the sparsity pattern, which can be tuned according to video analysis. In addition, basis and coefficient depend on the value of rank.

In particular, the OR-PCA optimization consists of two iterative updating components. Firstly, the input video frame is projected onto current initialized basis and we separate the sparse noise component, which includes the outliers contamination. Then, the basis is updated with a new input video frame. More details can be found in \[4\].

The background sequence for each image is then modeled by a multiple of basis \( L \) and its coefficient \( R \), whereas the sparse component \( E \) for each image constitutes the foreground objects.

3.2.1 Initialization

The number of subspace basis is randomly determined using improper value of rank, and no initialization method is considered for OR-PCA in \[4\]. As a result, the algorithm converges slowly to the optimal solution and outliers appear in the low-rank matrix, which effects the sparse component as well as foreground mask for background modeling case, as shown in Fig. 3.

In order to meet the time complexity, the basis for low-dimensional subspace is initialized using first \( N \) video frames with a good selection of rank. Since we are applying OR-PCA on two images in our scheme, the basis for each image is initialized according to this scheme. In this case, the rank is a tunable parameter for each image, that will be discussed in the later section.

By this technique, the OR-PCA converges to the low-dimensional subspace faster than original one. The outliers are also reduced and good computational time is achieved without sacrificing the quality of foreground in surveillance case.
3.3 Integration

The low-rank and sparse components are obtained from each decomposed image after applying OR-PCA. Gaussian and Laplacian low-rank and sparse components are integrated in this step. We use different parameters setting for OR-PCA on each decomposed image. $\lambda_2$ for Laplacian and $\lambda_2'$ for Gaussian image are selected according to background scene, for obtaining enough sparsity pattern for each decomposed image.

Since Laplacian image provides enough edge features for small variations in background scene, therefore $\lambda_2$ value must be smaller than $\lambda_2'$ of Gaussian image. After integrating both components of each image, the binary foreground mask is then obtained by thresholding the integrated sparse component.

4. EXPERIMENTAL RESULTS

In this section, we present experimental results in detail. We have tested our algorithm on three challenging datasets namely, Change Detection 2014 [5], Wallflower [13] and I2R [9] dataset.

Due to the space limitations, we present some specific sequences from each dataset. Our algorithm is compared with three most widely used methods namely, Mixture of Gaussians (MOG) [12], Semi-Soft GoDec [14] and DECOLOR [15].

The algorithm is implemented in Matlab R2013a with 3.40 GHz Intel core i5 processor with 4 GB RAM. Additionally, 5 × 5 median filtering is applied as a post-processing step on binary mask.

We determined the regularization parameters experimentally in Eqn. [1]. Therefore, we use different parameter settings for two images as follows: in case of static backgrounds, we use rank as 1 and $\lambda_2 \in (0.01, 0.04)$ for both images. However in case of highly dynamic backgrounds, such as waving trees, water surface and fountain, the rank $r \in (1, 3)$ and $\lambda_2 \in (0.01, 0.03)$ for Laplacian image, whereas the rank $r' \in (1, 3)$ and $\lambda_2' \in (0.01, 0.04)$ for Gaussian image, respectively. Similarly, for Bootstrapping case, the rank $r \in (1, 5)$ and $\lambda_2 \in (0.01, 0.05)$ for Laplacian image and rank $r' \in (5, 10)$ and $\lambda_2' \in (0.03, 0.06)$ for Gaussian image.

From Change Detection 2014 dataset [5], five sequences namely, highway and office from category baseline, and canoe, fountain2 and overpass from category dynamic backgrounds are selected. The image size of each sequence is $320 \times 240$. Fig. 4 shows the visual results of change detection dataset. Waving trees, camouflage and light switch sequences are taken from Wallflower dataset [13]. The frame size of each sequence is $160 \times 120$. Fig. 5 shows the qualitative results of wallflower dataset. Similarly, moving curtain, water surface, shopping mall sequences are tested from I2R dataset [9]. Each sequence contains a $160 \times 128$ of frame size. Fig. 6 shows results of I2R dataset.

We also evaluate our algorithm quantitatively with other methods. F-measure is computed in percentage for all sequences, by comparing our results (pixel-by-pixel) with their corresponding ground truths data.

Table 1-2-3 and 4 show the achieved performance on three datasets. In every case our algorithm outperforms with other state-of-art methods e.g on average F-measure of 87.79%, 87.34% and 86.76% in each dataset, respectively.

The computational time is also investigated during our experiments. The computational time is observed frame by frame in seconds. In our algorithm, time is proportional to the value of rank. Table 4 shows the comparison of computational time. These good experimental evaluations are the consequences of proposed initialization scheme in OR-PCA via image decompositions as well as parameters setting without loosing the quality of foreground. In addition, these valuable performance indicate that our proposed scheme makes a worth to be implemented in real time application.

Table 1: Change Detection Dataset: Comparison of F-measure in percentage (direct one-to-one correspondence with Fig. 4).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>MOG</th>
<th>SSGoDec</th>
<th>DECOLOR</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highway</td>
<td>88.05</td>
<td>1.94</td>
<td>94.47</td>
<td>92.63</td>
</tr>
<tr>
<td>Office</td>
<td>60.48</td>
<td>56.02</td>
<td>57.30</td>
<td>88.50</td>
</tr>
<tr>
<td>Canoe</td>
<td>51.14</td>
<td>30.91</td>
<td>16.03</td>
<td>85.05</td>
</tr>
<tr>
<td>Fountain2</td>
<td>79.68</td>
<td>31.38</td>
<td>82.41</td>
<td>88.67</td>
</tr>
<tr>
<td>Overpass</td>
<td>50.95</td>
<td>55.17</td>
<td>35.73</td>
<td>83.84</td>
</tr>
<tr>
<td>Avg.</td>
<td>66.06</td>
<td>35.08</td>
<td>57.18</td>
<td>87.79</td>
</tr>
</tbody>
</table>

Table 2: Wallflower Dataset: Comparison of F-measure in percentage (direct one-to-one correspondence with Fig. 5).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>MOG</th>
<th>SSGoDec</th>
<th>DECOLOR</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waving Trees</td>
<td>66.39</td>
<td>18.29</td>
<td>88.45</td>
<td>87.20</td>
</tr>
<tr>
<td>Light Switch</td>
<td>16.86</td>
<td>26.71</td>
<td>24.20</td>
<td>84.18</td>
</tr>
<tr>
<td>Camouflage</td>
<td>74.21</td>
<td>66.31</td>
<td>38.56</td>
<td>90.66</td>
</tr>
<tr>
<td>Avg.</td>
<td>52.48</td>
<td>37.10</td>
<td>50.40</td>
<td>87.34</td>
</tr>
</tbody>
</table>

Table 3: I2R Dataset: Comparison of F-measure in percentage (direct one-to-one correspondence with Fig. 6).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>MOG</th>
<th>SSGoDec</th>
<th>DECOLOR</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curtain</td>
<td>77.09</td>
<td>49.44</td>
<td>87</td>
<td>88.13</td>
</tr>
<tr>
<td>Water Surface</td>
<td>86.23</td>
<td>44.73</td>
<td>90.22</td>
<td>91.75</td>
</tr>
<tr>
<td>Shopping Mall</td>
<td>67.78</td>
<td>65.54</td>
<td>68.22</td>
<td>80.51</td>
</tr>
<tr>
<td>Avg.</td>
<td>77.03</td>
<td>51.23</td>
<td>81.81</td>
<td>86.79</td>
</tr>
</tbody>
</table>

Table 4: Comparison of computational time in seconds.

<table>
<thead>
<tr>
<th>Cases</th>
<th>Frame Size</th>
<th>OR-PCA</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static</td>
<td>576 × 720</td>
<td>0.39</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>288 × 360</td>
<td>0.09</td>
<td>0.01</td>
</tr>
<tr>
<td>Dynamic</td>
<td>240 × 360</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>128 × 160</td>
<td>0.01</td>
<td>0.004</td>
</tr>
<tr>
<td>Bootstrap</td>
<td>256 × 320</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>128 × 160</td>
<td>0.01</td>
<td>0.005</td>
</tr>
</tbody>
</table>

5. CONCLUSION

In this paper, a robust background modeling method against challenging background scenes is presented using OR-PCA via Gaussian and Laplacian image decomposition. OR-PCA via image decomposition shows a nice potential for background subtraction. We first decompose the input image then OR-PCA with initialization scheme is applied on
Figure 4: Qualitative comparisons of Change Detection Dataset. From left to right: (a) input, (b) ground truth, (c) MOG, (d) Semi Soft GoDec, (e) DECOLOR and (f) our algorithm.

Figure 5: Qualitative comparisons of Wallflower Dataset. From left to right: (a) input, (b) ground truth, (c) MOG, (d) Semi Soft GoDec, (e) DECOLOR and (f) our algorithm.

each image with parameters tuning. As a result, computational complexity is reduced and foreground quality is improved due to spatial Gaussian kernels. Experiment evaluations and comparisons with other state-of-the-art methods show the effectiveness and robustness of our proposed scheme.

However, we just applied OR-PCA on two decomposed images and parameters are tuned manually. Therefore, our future work will focus more on brief analysis of each layer of hierarchical image decomposition with dynamical parame-
ters setting, which adapts changes according to background dynamics.

6. ACKNOWLEDGMENTS

This work is supported by the World Class 300 project, Development of HD video/network-based video surveillance system(10040370), funded by the Ministry of Trade, Industry, and Energy (MOTIE, Korea).

7. REFERENCES