Computer Vision and Image Understanding 146 (2016) 27–39

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

COROLA: A sequential solution to moving object detection using low-rank approximation☆

Moein Shakeri*, Hong Zhang

Department of Computing Science, University of Alberta, Edmonton, Alberta T6G2E8, Canada

ARTICLE INFO

Article history:
Received 10 July 2015
Accepted 13 February 2016
Available online 4 March 2016

Keywords:
Moving object detection
Online low rank approximation
Markov Random Fields
Online background modeling

ABSTRACT

Extracting moving objects from a video sequence and estimating the background of each individual image are fundamental issues in many practical applications such as visual surveillance, intelligent vehicle navigation, and traffic monitoring. Recently, some methods have been proposed to detect moving objects in a video via low-rank approximation and sparse outliers where the background is modeled with the computed low-rank component of the video and the foreground objects are detected as the sparse outliers in the low-rank approximation. Many of these existing methods work in a batch manner, preventing them from being applied in real time and long duration tasks. To address this issue, some online methods have been proposed; however, existing online methods fail to provide satisfactory results under challenging conditions such as dynamic background scene and noisy environments. In this paper, we present an online sequential framework, namely contiguous outliers representation via online low-rank approximation (COROLA), to detect moving objects and learn the background model at the same time. We also show that our model can detect moving objects with a moving camera. Our experimental evaluation uses simulated data and real public datasets to demonstrate the superior performance of COROLA to the existing batch and online methods in terms of both accuracy and efficiency.

1. Introduction

Moving object detection and background estimation are fundamental in various applications of computer vision and robotics such as visual surveillance [1], traffic monitoring [2], vehicle tracking and navigation [3], and avian protection [4]. Many methods have been proposed to extract objects from a sequence of images with a stationary camera [5,6] or with a moving camera [7–9]. These methods can be grouped into several categories. Motion-based methods [10,11] use motion information of the image pixels to separate the foreground from the background. These methods work based on the assumption that foreground objects move differently from the background. Therefore it is possible for these methods to classify pixels according to their movement characteristics even in the case of significant camera motion. However, these methods require point tracking to identify the foreground, which can be difficult especially with large camera motion [12]. In addition, they are limited in terms of dealing with dynamic background or noisy data [13].

Another popular category for moving object detection methods is background subtraction [14], which compares the pixels of an image with a background model and considers those that differ from the background model as moving objects. Thus, building a background model plays a critical role in background subtraction methods. Conventional algorithms for background modelling include single Gaussian distribution [15], Gaussian mixture model (GMM) [5], and kernel density estimation [16]. These methods model the background for each pixel independently and so they are not robust against global variations such as illumination changes.

Recently a new approach to background modelling, namely low-rank matrix approximation, has been developed [17,18]. Methods in this approach follow the basic idea from [19]. Oliver et al. [19] proposed Eigenbackground subtraction using PCA [20] (principal component analysis) to model the background and detect moving objects. It is based on the observation that the underlying background images should be unchanged and the composed matrix of vectorized background images can be naturally modeled as a low-rank matrix. Extending this idea, current methods exploit the fact that the background model in an image sequence can be defined by those pixels that are temporally linearly correlated [21]. By capturing the correlation between images one can naturally handle global variations. Algebraically speaking, if an image is vectorized

☆ This paper has been recommended for acceptance by Alper Yilmaz.

* Corresponding author.
E-mail addresses: shakeri@ualberta.ca (M. Shakeri), hzhang@ualberta.ca (H. Zhang).

http://dx.doi.org/10.1016/j.cviu.2016.02.009

in a column and all images are concatenated into a 2D matrix, then the columns are dependent and its low-rank approximation matrix represents the background model of the images. As a result, the background modeling problem is converted to the low-rank approximation problem. In general, by decomposing an input matrix of vectorized images into a low-rank matrix and a sparse matrix, the low-rank and sparse matrices correspond to the background model and the foreground objects in the image sequence respectively. Our COROLA algorithm described in this paper adopts the low-rank approximation approach. We will detail representative algorithms in this approach in Section 2.

Most of the existing background subtraction algorithms based on low-rank approximation operate in a batch manner; i.e., all images whose background model is too constructed are first collected and then used to build a data matrix whose low-rank approximation is computed. This unfortunately limits the application of the low-rank approximation approach in terms of its efficiency and accuracy. Although existing online methods via low-rank approximation have addressed the efficiency issue to some extent, they are not robust against dynamic and noisy background. In this paper, we offer an algorithm, COROLA, that performs low-rank approximation in a sequential manner so that its computational complexity does not grow with the number of images in the sequence. In addition, through image registration, our algorithm is able to handle the case of a moving camera due to the adaptive nature of the background model that is being learned. The main contributions of this paper are as follows.

1. We propose an online formulation of the low-rank approximation algorithm for foreground object detection. The proposed formulation enables online application without requiring an entire image sequence, as in the batch formulation and is more robust than existing online methods for dynamic background scene or noisy environment.
2. COROLA uses a fixed window of images to perform low-rank approximation and so it is appropriate for continuous operation, which cannot be achieved by the batch formulation due to matrix decomposition and memory storage.
3. In the case of significant camera motion, a batch formulation has the limitation that the first and the last images of a sequence must be similar to find the low-rank matrix. However, in the case of a moving camera, there is in general no similarity between the first and the last images in a sequence. Our proposed COROLA algorithm does not require a stationary background.

The remainder of the paper is organized as follows. Related works on foreground detection via low-rank and sparse decomposition are summarized in Section 2. Section 3 explains the details of COROLA for foreground detection and background estimation, followed by the introduction of our online formulation via greedy bilateral sketch [22]. Experimental results and discussion are presented in Section 4, and concluding remarks in Section 5.

2. Foreground detection via low rank and sparse decomposition

In recent years, many algorithms have been developed for foreground detection based on low-rank matrix approximation with robust principal component analysis (RPCA) [18]. RPCA decomposes a given matrix D into low-rank matrix L and sparse matrix S called outliers. Different techniques exist for low-rank approximation including principal component pursuit (PCP) [21], augmented Lagrangian multiplier (ALM) [23], linearized alternating direction method with an adaptive penalty (LADMAP) [24], and singular value thresholding (SVT) [25]. All of these techniques need all the data in order to perform batch optimization that computes the low-rank matrix and the sparse outliers. Due to batch processing, the following two problems occur: *memory storage* and *time complexity*. In continuous monitoring tasks or video processing, if matrix D is built with a large number of images memory storage will be a problem [26]. In addition, by increasing the size of the input matrix D, time complexity for the matrix decomposition is also increasing.

To address the problem of time complexity, some efficient algorithms have been proposed [22,27,28]. Rodrigues and Wohlberg [28] proposed a fast PCP algorithm to reduce the computation time of SVD in inexact ALM (IALM). The “Go Decomposition” (GoDec) method, proposed by Zhou et al. computes RPCA using bilateral random projections (BRP) [28]. Semi-Soft GoDec (SSGoDec) and Greedy SSGoDec methods [22] are extensions of GoDec to speedup it. Although these algorithms reduce the computation time of low-rank approximation, they still are not satisfactory for applications such as visual surveillance and robot navigation due to their batch formulation. In many applications, online processing is critical and batch methods are infeasible. One of the best known batch processing algorithms is the “detecting contiguous outliers in the low-rank representation” (DECOLOR) method [29]. This method uses a priori knowledge of the foreground objects that they should be connected components of relatively small size. Using this constraint in the method, DECOLOR provides promising results; however, due to batch processing, it still suffers from memory storage and time complexity problems. Furthermore, in the case of a moving camera, the current image is no longer similar to the first images in matrix D, and therefore DECOLOR is not able to detect foreground appropriately. In general, batch processing methods cannot operate on a continuous basis and cannot deal with a moving camera. Although DECOLOR has introduced an implementation for moving camera, it only works for short video sequences with small camera motion.

To overcome the limitations of batch processing methods, incremental and online robust PCA methods have developed. He et al. [30] proposed Grassmannian robust adaptive subspace tracking algorithm (GRASTA),which is an incremental gradient descent algorithm on Grassmannian manifold for solving the robust PCA problem. This method incorporates the augmented Lagrangian of l_1-norm loss function into the Grassmannian optimization framework to alleviate the corruption by outliers in the subspace update at each gradient step. Following the idea of GRASTA, He et al. [31] proposed transformed GRASTA (t-GRASTA), which iteratively performs incremental gradient descent constrained to the Grassmann manifold in order to simultaneously decompose a sequence of images into three parts: a low-rank subspace, foreground objects, and a transformation such as rotation or translation of the image. This method can be regarded as an extension of GRASTA and RASL [32] (Robust Alignment by Sparse and Low-Rank decomposition) by computing the transformation and solving the decomposition with incremental gradient optimization framework. To improve the accuracy of online subspace updates especially for dynamic backgrounds, Xu et al. [33] developed an online Grassmannian subspace update algorithm with structured-sparsity (GOSUS) via an alternating direction method of multipliers (ADMM).

To deal with noisy conditions and dynamic background scene, Wang et al. [34] proposed a probabilistic approach to robust matrix factorization (PRMF) and its online extension for sequential data to obtain improved scalability. This model is based on the empirical Bayes approach and can estimate better background model than GRASTA. Recently, Feng et al. [35] proposed an online robust principal component analysis via stochastic optimization (OR-PCA). This method does not need to remember all the past samples and uses one sample at a time by a stochastic optimization. OR-PCA reformulates a nuclear norm objective function by decomposing to an explicit product of two low-rank matrices, which can