On Video Background Subtraction Using Robust PCA Methods

Richard Al-Bayaty
School of Electrical and Computer Engineering
University of Florida
Gainesville, FL 32611
Email: ralbayaty@ufl.edu

***Abstract*—There are many application areas for background subtraction techniques, some of which being object tracking, image segmentation, human pose estimation, and change detection. Of the many ways by which background subtraction is performed, a robustification of the well known Principal Component Analysis (PCA) has been frequently adopted. In background detection, it is assumed that the dataset is composed of a low rank matrix and a sparse outlier matrix. This robust PCA (RPCA) method allows for the decomposition of a dataset into both the low rank and the sparse outlier portions, which models the background and the foreground respectively.**

***Index Terms*—Background subtraction, foreground detection, robust PCA, video processing.**

I. Introduction

In video analysis, it is often desired to parse video streams into both a foreground and background. The background of a video stream contains objects and scenery which do not necessarily change swiftly, while the foreground is taken to be objects which move more independently and regularly than the background. An example of this would be a surveillance camera pointed towards a parking lot. The parking lot, street lights, any grass or shrubbery as well as the painted lines should be considered as the background, while the foreground would be the cars, people, animals, and other objects as they come in and out of the lot. One of the methods used for performing foreground detection (or background detection) is to model the video as a dataset which can be decomposed as $X = L + S$, where X is the vectorized frames from the video scene into a data matrix, L is a low rank matrix which represents the aforementioned background and S is the sparse outlier matrix representing the foreground. Other methods incorporate an additional matrix, E, which aims to model additional noise from either the process itself or measurements. The optimization problem for the robust version of PCA (Principal Component Pursuit) is formulated by Candes [1] as:

$$\begin{aligned} \arg\min_{L,S} \quad & ||L||_* + \lambda||S||_1 \\ \text{s.t.} \quad X \quad &= \quad L + S \end{aligned} \tag{1}$$

The PCP framework allows for a convenient model towards background subtraction in that the total scene can be regarded as being composed of the L and S matrices. This work aims to highlight some of the more recent background subtraction algorithms developed using RPCA.

The rest of this paper is organized as follows: Section II covers the work related to this project and gives an overview of some of the currently used RPCA algorithms for background subtraction, Section III covers the qualitative and quantitative analysis, Section IV discusses the experiments performed including visualizations of experiment results, and Section V concludes the study with a summary and recommendations for future work.

II. Related Work

This work is comprised mainly of a brief introduction to the topic area as well as a superficial explanation of some of the algorithms used to accomplish the background subtraction task. More thorough investigations of methodologies and a rigourous attempt at utlizing quantitative measures for algorithm comparisons are performed by both Bouwmans [2] and Sobral [3]. The following sub sections elaborate on some non-RPCA methods as well as introduce the RPCA methods that are evaluated in Section IV.

A. Non-RPCA methods

There are other methods besides those modeled with RPCA that have been used successfully in previous years. Some of the techniques include Static Frame Difference, where a frame from the video sequence without foreground noise is chosen as being the background and subsequence frames are differenced from this image to produce the foreground. This technique is sensitive to both illumination and dynamic scenes, so an advancement to this framework was to use the previous frame instead of a static image for differencing. This technique was coined further Frame Difference.

Another methodology for this problem is to use statistical models such as Gaussian Mixture Models (GMM) to represent each pixel and calculate the likelihood of a particular pixel as being background/foreground for a given pixel value. This model can be updated in an online fashion using an expectation maximization algorithm.

Neural and neuro-fuzzy network models have also been used with great success. Each pixel is deemed as belonging to background/foreground by a Bayesian classifier while the background model is learned through the neural network.

B. Fast Principal Component Pursuit (FPCP) [4]

In fast principal component pursuit, a minor variation to the original PCP optimization problem is introduced and an alternating minimization technique is proposed. This variation is shown in equation 2.

$$\begin{array}{rl} \arg\min\limits_{L,S} & \frac{1}{2}||L+S-D||_F + \lambda||S||_1 \\ \text{s.t.} & rank(L) \;=\; t \end{array} \tag{2}$$

The alternating minimizations shown in equations 3 and 4 allow for an expedient solution recovery.

$$L_{k+1} = \arg\min_L ||L+S_k-D||_F \quad \text{s.t. } rank(L)=t \tag{3}$$

$$S_{k+1} = \arg\min_S ||L_{k+1}+S-D||_F + \lambda||S||_1 \tag{4}$$

The solution to equation 3 comes from computing a t-component Singular Value Decomposition (SVD) of $D - S_k$. Solving equation 4 is performed with an element-wise shrinkage, shrink($D - L_{k+1}, \lambda$).

C. Go Decomposition (GoDec) [5]

The GoDec algorithm expands upon the traditional PCP framework by presuming that the L and S are not exactly recoverable due to Gaussian noise corruption of the form $X = L + S + G$, and thus incorporates an L2 minimization as the optimal estimator for a model with additive Gaussian noise constraining the rank and the cardinality of L and S as shown in equation 5. To further approximate and boost the speed of this algorithm, the authors proposed replacing the SVD solving method in lieu of a bilateral random projection based low-rank approximation to the Naive GoDec algorithm.

$$\begin{array}{rlcl} \min\limits_{L,S} & ||X-L-S||_F^2 & & \\ \text{s.t.} & rank(L) & \leq & r \\ & card(C) & \leq & k \end{array} \tag{5}$$

D. Semi-Soft GoDec (SSGoDec) [5]

While GoDec applies a hard threshold to both the singular values of the low rank matrix L and the sparse entries of the matrix S, the Semi-Soft variant allows for a soft thresholding to the values in S. This adaptation provides a strength in that the parameter k to control the cardinality of S is now automatically driven by the soft thresholding parameter τ. In GoDec, if one chose the value of k to be too large then S could absorb some of the noise introduced by the additive noise matrix G. This soft-thresholding also allows for a boost to the computation speed of the algorithm.

E. Reinforced Robust Principal Component Pursuit (R^2-PCP)

To handle the need to model outliers in both the principal component space as well as the orthogonal component space of the transformed subspace, R^2-PCP offers a data model as follows in equation 6 :

$$\begin{array}{r} X = B + O + (1\mu^T + S)V_\perp^T + E \\ \text{with } rank(B) \leq r \end{array} \tag{6}$$

B is low rank, O is the sparse principal component outlier matrix, E is an additive noise matrix, and the remaining term, $(1\mu^T + S)V_\perp^T$, is the orthogonal space outlier matrix.This problem can be formulated as equation 7, where $P_O(O;\lambda_O)$ and $P_S(S;\lambda_S)$ are general sparsity penalties with λ_O and λ_S as regularization parameters.

$$\begin{array}{rl} \min\limits_{V_\perp,\mu,O,S} & \frac{1}{2}||(X-O)V_\perp - 1\mu^T - S||_F^2 + P_O(O;\lambda_O) + P_S(S;\lambda_S) \\ \text{s.t.} & V_\perp^T V_\perp \;=\; I \end{array} \tag{7}$$

Given μ, O, S, equation 7 is optimized for $V_\perp$ as an unconstrained problem on the Stiefel manifold.

III. Description

A vast amount of background subtraction datasets are available to test different algorithm performance. The dataset used for comparing these algorithms comes from the ChangeDetection.NET (CDNET) 2014 dataset [6]. A subset of the video sequences CDNET2014 were used, namely the baseline highway and the baseline pedestrians. This dataset contains many video sequences to test the robustness of algorithms in many ways, such as with night vision videos, panning/tilting/zooming, dynamic backgrounds, and bad weather. Since the function of this work is aimed at introducing the topic and algorithms, using the baseline videos were sufficient.

Each video sequence, highway and pedestrian, are split into frames, converted to grayscale to remove two-thirds the values, and vectorized to construct an n by p matrix, X, where n is the number of frames and p is $m * n$ (frame height times frame width). This matrix is then fed to the selected RPCA algorithms to obtain outputs consisting of a background (L), a foreground (S), and a hard-thresholded version of the foreground (O) to highlight the nature of the foreground outliers. To compare results quantitatively, the peak signal-to-noise ratio (PSNR) was used between the provided ground truth after being hard-thresholded (to remove intermediary gray levels) and the outputs of the algorithms. The algorithm with the highest PSNR is deemed to have reproduced the best foreground for the given video sequence.

IV. Evaluation

The code used to run FPCP, GoDec, and SSGoDec was obtained from "Low-Rank and Sparse Tools for Background Modeling and Subtraction in Videos" library created by Andrews Sobral [7]. The code for R^2-PCP was made available through the Multimedia, Communications, and Networking Lab from the dissertation of Dr. Shijie Li.

Tables I and II show the run times of the algorithms and the PSNR as compared to the ground truth, respectively. The GoDec algorithm stood out as being the fastest for both datasets, while the performance of the SSGoDec was superior.

TABLE I
ALGORITHM RUN TIMES (SECONDS)

	FPCP	GoDec	SSGoDec	R^2-PCP
Highway	249.7573	81.3508	865.2944	Unk
Pedestrian	89.2926	45.1151	608.7915	Unk

TABLE II
ALGORITHM PSNR (DECIBELS)

	FPCP	GoDec	SSGoDec	R^2-PCP
Highway	10.9708	7.4582	11.6149	Unk
Pedestrian	15.7586	15.7529	15.9369	Unk

In Figures 1 and 2, the ground truth for a single frame of the two video sequences are shown on the left-most column, while the remaining columns show the outputs for that particular frame in the three algorithms that provided outputs. It is difficult to make a qualitative comparison in these results since they are seemingly very similar. A comparison between the input video sequence frame and the resultant background and foreground from the algorithms is shown in Figures 3 and 4.

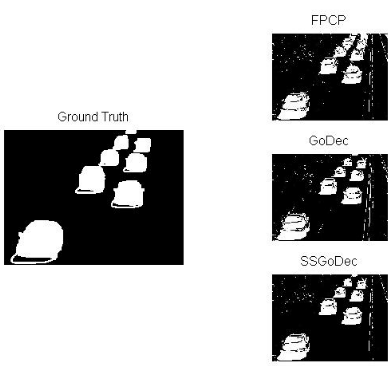

Fig. 1. Comparison to ground truth for highway sequence

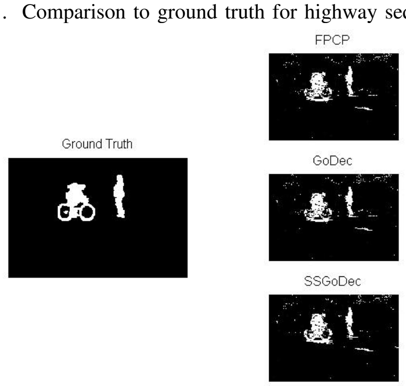

Fig. 2. Comparison to ground truth for pedestrian sequence

Fig. 3. The input compared to the L and S outputs for highway

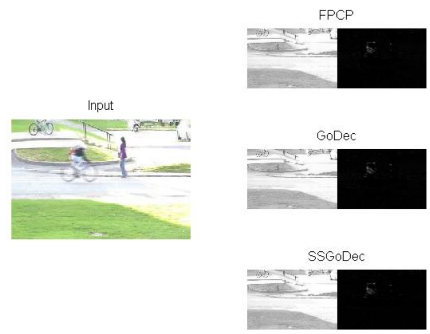

Fig. 4. The input compared to the L and S outputs for pedestrians

V. CONCLUSION

The task of splitting a video sequence into two separate sequences, a background and foreground, using methodologies from PCP was introduced. A handful of currently used algorithms were explained and run on a reputable dataset in order to make comparisons between the algorithms' run times and their output quality. It was shown for one run of the algorithms that GoDec succeeded in the fastest run time while SSGoDec achieved best PSNR for foreground detection, given the ground truth of the foreground provided in the CDNET 2014 dataset. Unfortunately, the novel and new algorithm R^2-PCP was not capable of processing the datasets in any acceptable amount of time due to the nature of the optimization problem and the computing limitations for this experimentation. In order to test the performance of this algorithm, a more compact and fast solver must be developed or found to be implemented for use with large datasets such as the ones used in this experimentation.

References

[1] Cands, Emmanuel J., Xiaodong Li, Yi Ma, and John Wright. “Robust principal component analysis?.” Journal of the ACM (JACM) 58, no. 3 (2011): 11.

[2] Thierry Bouwmans, El Hadi Zahzah, “Robust PCA via Principal Component Pursuit: A review for a comparative evaluation in video surveillance”, Computer Vision and Image Understanding, Volume 122, May 2014, Pages 22-34, ISSN 1077-3142, http://dx.doi.org/10.1016/j.cviu.2013.11.009.

[3] Andrews Sobral, Antoine Vacavant, “A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos”, Computer Vision and Image Understanding, Volume 122, May 2014, Pages 4-21, ISSN 1077-3142, http://dx.doi.org/10.1016/j.cviu.2013.12.005.

[4] Rodriguez, P.; Wohlberg, B., “Fast principal component pursuit via alternating minimization,” Image Processing (ICIP), 2013 20th IEEE International Conference on , vol., no., pp.69,73, 15-18 Sept. 2013 doi: 10.1109/ICIP.2013.6738015

[5] Zhou, Tianyi, and Dacheng 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. 2011.

[6] Y. Wang, P.-M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, “CDnet 2014: An Expanded Change Detection Benchmark Dataset”, in Proc. IEEE Workshop on Change Detection (CDW-2014) at CVPR-2014, pp. 387-394. 2014 (http://www.changedetection.net)

[7] Sobral, A. and Baker, C. G. and Bouwmans, T. and Zahzah, E, “Incremental and Multi-feature Tensor Subspace Learning applied for Background Modeling and Subtraction”, International Conference on Image Analysis and Recognition (ICIAR’14), Oct. 2014