ADVANTAGES OF ROBUST PRINCIPAL COMPONENT ANALYSIS OVER GAUSSIAN MIXTURE MODELS FOR BACKGROUND SUBTRACTION OF NOISY VIDEOS

Pooya Khorrami, Xingqian Xu, Mert Dikmen, Thomas S. Huang

University of Illinois - Urbana Champaign
Department of Electrical and Computer Engineering
Beckman Institute, 405 N. Matthews Avenue, Urbana IL, 61801

ABSTRACT

The abstract should appear at the top of the left-hand column of text, about 0.5 inch (12 mm) below the title area and no more than 3.125 inches (80 mm) in length. Leave a 0.5 inch (12 mm) space between the end of the abstract and the beginning of the main text. The abstract should contain about 100 to 150 words, and should be identical to the abstract text submitted electronically along with the paper cover sheet. All manuscripts must be in English, printed in black ink.

Index Terms— One, two, three, four, five

1. INTRODUCTION

In the field of video surveillance, the user typically wishes to extract meaningful and salient information from a video sequence in a completely automatic fashion. In several instances, the video sequences are captured using stationary cameras leading to a relatively static scene layout. The absence of camera motion implies that the background of the video sequence exhibits very little variation while dynamic changes in the scene represent the objects of interest. In such a case, the most common approach is to perform background subtraction to separate said dynamic regions (i.e. foreground) from the background of each frame.

When performing background subtraction, some naïve methods include frame differencing and approximate median [1]. While these algorithms are simple to use and quite efficient, the most popular technique, by far, is adaptive Gaussian Mixture Models (GMMs) [2]. Friedman et al. posit that each pixel in the background image can be represented by a probability distribution formed by a mixture of Gaussians. If a pixel greatly deviates from its corresponding model, then the pixel is labeled foreground. While the number of Gaussians used at each pixel is usually fixed, there has been some work [3] that adaptively selects the number of mixture components.

Although background subtraction via Gaussian Mixture Models enjoys widespread use in the computer vision community, it is not without drawbacks. As opposed to the frame differencing and approximate median techniques, GMMs possess several parameters that must be individually tuned. This implies that the algorithm is innately sensitive to different scene configurations. Therefore it should be no surprise that GMMs tend to perform rather poorly on noisy videos where the foreground objects are not immediately distinguishable.

When dealing with noisy video sequences, we advocate the use of Robust Principal Component Analysis (RPCA or Robust PCA) for background subtraction. Robust PCA [4] refers to how any matrix M can be represented as the sum of a low-rank matrix L and sparse matrix S by solving a convex optimization problem. If we form the matrix M by stacking each frame of the video sequence as a column, then we will see that the columns are highly correlated. This is expected considering that the video was obtained using a stationary camera and implies that the background of a video scene lies on a low-dimensional subspace. As a result, when Robust PCA is performed on the matrix M, the columns of the low-rank matrix L will correspond to the background of the frame and the columns of the matrix S will contain sparse deviations from the low-rank subspace.

While the aforementioned description of Robust PCA operates on a video sequence in batch, there is also the newly proposed Grassmannian Robust Adaptive Subspace Tracking Algorithm (GRASTA) by He et al. [5] that learns the low-rank subspace by subsampling the video frames and proceeds in an online fashion. Although it is not Robust PCA in its purest sense, GRASTA still assumes that each frame can be represented as the sum of data generated from a low-dimensional subspace and a sparse error term.

In this paper, we will show how Robust PCA achieves superior performance to GMMs when applied to noisy videos. We will also describe how performing Robust PCA on each of the color channels of a video sequence and subsequent thresholding of the wavelet coefficients provide additional improvements. Should the user require a real-time alternative, we will also show how GRASTA also outperforms GMMs on the same noisy data. The remainder of this paper is organized as follows. Section 2 will describe the improvements made to

Thanks to XYZ agency for funding.

the batch Robust PCA algorithm. Section 3 will present our experimental setup and findings. Section 4 will describe our conclusions and directions for future work.

2. METHOD DESCRIPTION

3. EXPERIMENTAL RESULTS

4. CONCLUSIONS

5. REFERENCES

[1] N. J. B. McFarlane and C. P. Schofield, "Segmentation and tracking of piglets in images," *Machine Vision and Applications*, vol. 8, no. 3, pp. 187–193, May 1995.

[2] Nir Friedman and Stuart Russell, "Image segmentation in video sequences: A probabilistic approach," 1997, pp. 175–181.

[3] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition*, aug. 2004, vol. 2, pp. 28 – 31 Vol.2.

[4] Emmanuel J. Candès, Xiaodong Li, Yi Ma, and John Wright, "Robust principal component analysis?," *CoRR*, vol. abs/0912.3599, 2009.

[5] Jun He, Laura Balzano, and Arthur Szlam, "Incremental gradient on the grassmannian for online foreground and background separation in subsampled video," in *CVPR*, 2012, pp. 1568–1575.