Background Subtraction

Theory and Practice

Synthesis Lectures on Computer Vision

Synthesis Lectures on Computer Vision is edited by Gérard Medioni of the University of Southern California and Sven Dickinson of the University of Toronto. The series publishes 50- to 150 page publications on topics pertaining to computer vision and pattern recognition.

Background Subtraction: Theory and Practice

Ahmed Elgammal

ISBN: 978-3-031-00685-2 paperback
ISBN: 978-3-031-01813-8 ebook

DOI 10.1007/978-3-031-01813-8

A Publication in the Springer series
SYNTHESIS LECTURES ON COMPUTER VISION

Lecture #6
Series Editors: Gérard Medioni, *University of Southern California*
Sven Dickinson, *University of Toronto*
Series ISSN
Print 2153-1056 Electronic 2153-1064

Background Subtraction

Theory and Practice

Ahmed Elgammal
Rutgers University

SYNTHESIS LECTURES ON COMPUTER VISION #6

ABSTRACT

Background subtraction is a widely used concept for detection of moving objects in videos. In the last two decades there has been a lot of development in designing algorithms for background subtraction, as well as wide use of these algorithms in various important applications, such as visual surveillance, sports video analysis, motion capture, etc. Various statistical approaches have been proposed to model scene backgrounds. The concept of background subtraction also has been extended to detect objects from videos captured from moving cameras. This book reviews the concept and practice of background subtraction. We discuss several traditional statistical background subtraction models, including the widely used parametric Gaussian mixture models and non-parametric models. We also discuss the issue of shadow suppression, which is essential for human motion analysis applications. This book discusses approaches and tradeoffs for background maintenance. This book also reviews many of the recent developments in background subtraction paradigm. Recent advances in developing algorithms for background subtraction from moving cameras are described, including motion-compensation-based approaches and motion-segmentation-based approaches.

For links to the videos to accompany this book, please see https://sites.google.com/a/morganclaypool.com/backgroundsubtraction/.

KEYWORDS

background subtraction, segmentation, visual surveillance, gaussian mixure model, kernel density estimation, moving object detection, shadow detection, figure-ground segmentation, motion segmentation, motion compensation, layered scene segmentation

To my wife for her emotional support,
to my parents for all what they have given me,
and to my kids for the joy they bring to my life.

– Ahmed

Contents

Preface

With cameras around us everywhere, a continuos stream of a vast amount of videos comes that need to be processed and analyzed. This explosion of videos drives the research and development in the computer vision community at large. The detection of foreground objects is the first step in several computer vision processing pipelines. The are several solution paradigms for detecting foreground objects in videos, with varying assumptions depending on the type of imaging sensor, the motion of the imaging platform, whether the video is available online or offline, and the characteristics of the applications. Among these solutions comes *Background Subtraction* as one of the widely used paradigms to detect foreground regions in videos. The success of Background Subtraction algorithms led to the development of the camera-based automated surveillance industry, as well several video analysis applications such as sports analysis, event detection, activity recognition, marker-less motion capture systems, and others.

The Background Subtraction paradigm has an over two-decades history in the computer vision field, with roots in image processing and even deeper roots extending to the earlier years of photography. With over 500 papers published in conferences, workshops, and journals related to background subtraction, it is very hard for a student to get hold of the different developments in this area, and put them in the right historical perspective. The goal of this book is to give the reader a birds-eye view of this topic, with enough details about the major approaches developed over the last two decades, without overwhelming the reader with all variants of these approaches. This book is not intended to be a comprehensive survey of every paper on the subject, but rather a gentle guide to exploring the topic.

This book is organized in three chapters. Chapter 1 introduces the problem of foreground detection in videos and explores solutions ranging from foreground detectors, motion segmentation, and video segmentation to background subtraction. The goal is to highlight the basic assumptions behind each of these paradigms and shed light on where background subtraction fits within the big picture. Chapter 2 reviews some of the traditional statistical background models, including parametric and non-parametric models, which are the bases for many algorithms for background subtraction. Chapter 2 also discusses issues related to detecting shadows and maintaining the background representation. Chapter 3 discusses the extensions of background subtraction to deal with videos captured from moving cameras. The chapter describes motion-compensation-based approaches, layered-motion segmentation, and motion-segmentation-based approaches for background subtraction from a moving platform.

Ahmed Elgammal
November 2014

Acknowledgments

I would like to thank Gérard Medioni and Sven Dickinson, the editors of this "Synthesis Lectures on Computer Vision" series for giving me this opportunity to serve the community by writing this book. I would like to thank Ali Elqursh for helping with some of the materials in Chapter 3. I greatly appreciate the help of the editor of the series, Diane Cerra. I would like to thank the researchers who granted me permission to use figures from their papers in this book, including: Edward H. Adelson, Thanarat H. Chalidabhongse, Bohyung Han, Ivan Huerta, Kyungnam Kim, Anurag Mittal, Yaser Sheikh, and Jürgen Stauder.

Ahmed Elgammal
November 2014

Figure Credits

Figure 2.1 From Guéziec, A. Tracking pitches for broadcast television. *Computer*, 35(3):38–43. Copyright © 2002, IEEE. DOI: 10.1109/2.989928. Used with permission.

Figures 2.2, 2.5, 2.9 From Elgammal, et al. Nonparametric background model for background subtraction. In *ECCV '00 Proceedings of the 6th European Conference on Computer Vision-Part II*: Pages 751–767. Copyright © 2000, Springer-Verlag London, UK. ISBN: 3-540-67686-4. Used with permission.

Figures 2.3, 2.14 Based on Elgammal, et al. Nonparametric background model for background subtraction. In *ECCV '00 Proceedings of the 6th European Conference on Computer Vision-Part II*: Pages 751–767. Copyright © 2000, Springer-Verlag London, UK. ISBN: 3-540-67686-4.

Figure 2.8 Based on Stauder, et al. Detection of moving cast shadows for object segmentation. *IEEE Transactions on Multimedia*, (Volume: 1, Issue: 1), page(s): 65–76. Copyright © 1999, IEEE. DOI: 10.1109/6046.748172. Courtesy of Jürgen Stauder.

Figure 2.10 Based on Horprasert, et al. (2000). A robust background subtraction and shadow detection. In *Proceedings of the 4th Asian Conference on Computer Vision*, pages 983–988. Courtesy of Thanarat Horprasert.

Figure 2.11 Based on Huerta, et al. Detection and removal of chromatic moving shadows in surveillance scenarios. *2009 IEEE 12th International Conference on Computer Vision*, pages 1499–1506. Copyright © 2009, IEEE. DOI: 10.1109/ICCV.2009.5459280. Courtesy of Ivan Huerta.

Figure 2.12 From Horprasert, et al. A robust background subtraction and shadow detection. In *Lecture Notes in Computer Science*: *Proceedings of the 4th Asian Conference on Computer Vision*, pages 983–988. Copyright © 2000, Springer Heidelberg Dordrecht London New York. Used with permission.

Figure 2.13 From Huerta, et al. Detection and removal of chromatic moving shadows in surveillance scenarios. *2009 IEEE 12th International Conference on Computer Vision*, pages 1499–1506. Copyright © 2009, IEEE. DOI: 10.1109/ICCV.2009.5459280. Used with permission.

Figure 2.15 Based on Kim, et al. Background updating for visual surveillance. *Advances in Visual Computing: First International Symposium, ISVC 2005*, Lake Tahoe, NV, USA, December 5-7, 2005. Copyright © 2005, Springer Berlin Heidelberg. DOI: 10.1007/11595755_41. Courtesy of Kyungnam Kim.

Figure 3.1 From Mittal, A. and Huttenlocher, D. Scene modeling for wide area surveillance and image synthesis. *IEEE Conference on Computer Vision and Pattern Recognition, 2000: Proceedings*. Copyright © 2000, IEEE. DOI: 10.1109/CVPR.2000.854767. Used with permission.

Figure 3.2 Based on Wang, J. and Adelson, E. Representing Moving Images with Layers. *IEEE Transactions on Image Processing* (Volume: 3, Issue: 5). Copyright © 1994, IEEE. DOI: 10.1109/83.334981. Courtesy of Edward H. Adelson.

Figure 3.3 Based on Sheikh, et al. Background Subtraction for Freely Moving Cameras. *2009 IEEE 12th International Conference on Computer Vision*. Copyright © 2009, IEEE. DOI: 10.1109/ICCV.2009.5459334. Courtesy of Yaser Sheikh.

Figure 3.4 Based on Kwak, et al. Generalized Background Subtraction Based on Hybrid Inference by Belief Propagation and Bayesian Filtering. *2011 IEEE International Conference on Computer Vision (ICCV)*. Copyright © 2011, IEEE. DOI: 10.1109/ICCV.2011.6126494. Courtesy of Bohyung Han.

Figures 3.5, 3.6, 3.7 From Elqursh, A. and Elgammal, A. Online moving camera background subtraction. *Computer Vision – ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI*. Copyright © 2012, Springer-Verlag Berlin Heidelberg. DOI: 10.1007/978-3-642-33783-3_17. Used with permission.