Ain Shams University
Faculty of Computer and Information Sciences
Scientific Computing Department

Developing a Real-Time Object Tracking System

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Scientific Computing

By

Maha Mohamed El-Sayed Azab

B.Sc. in Computer and Information Sciences,
Department of Scientific Computing,
Faculty of Computer and Information Sciences, Ain Shams University.

Under Supervision of

Ashraf Saad Hussein

Professor,
Department of Scientific Computing,
Faculty of Computer and Information Sciences, Ain Shams University.

Howida Abdel-Fattah Shedeed

Associate Professor,
Department of Scientific Computing,
Faculty of Computer and Information Sciences, Ain Shams University.

Cairo 2014

Contents

List of Tables

List of Figures

List of abbreviations

BeSemiT: Beyond Semi-Supervised Tracker

BoostT: Online Boosting Tracker

EM: Expectation–Maximization

FN: False Negatives

FP: False Positives

FPR: False Positives Rate

FragT: Fragments-Based Tracker

HOG: Histogram of Oriented Gradients

KDE: Kernel Density Estimation

LBP: Local Binary Pattern

MoG: Mixture of Gaussians

PF: Particle Filter

SemiT: Semi-Supervised Tracker

SGM: Single Gaussian Model

TN: True Negatives

TP: True Positives

TPR: True Positives Rate

List of publication

Conference Publication:

Maha M. Azab, Howida A. Shedeed and Ashraf S. Hussein: A New Technique for Background Modeling and Subtraction for Motion Detection in Real-time Videos. IEEE 17th International Conference on Image Processing (ICIP), Hong Kong, pp. 3453–3456, 2010. No. of citations (9)

Journal Publication:

Maha M. Azab, Howida A. Shedeed and Ashraf S. Hussein: New Technique for Online Object Tracking-by-Detection in Video. IET Image Processing Journal, 2014. doi: 10.1049/iet-ipr.2014.0238 Impact Factor (0.9)

Acknowledgements

I would like to thank God for giving me the strength to finish this work. Without the help and inspiration of God, I would not be able to accomplish this work.

I would like to thank Prof. Ashraf Saad Hussein for supervising this research work. I would like to express my sincere gratitude to Dr. Howida Abdel-Fattah Shedeed for her continuous support and patience all the time. Without her, I would give up this work two years ago.

I would like to thank my colleagues in work, especially Emad Wahib, Walaa Husny and Gihan Haroun and my friends, especially Maiada Mohsen, Mona El Sayed and Mona Mahran, for their willing to help. I'm grateful for their continuous encouragement to afford working on this research, beside the hard workload.

Most importantly, nothing would have been possible without the patience and the love of my parents and my sister Soha. They are the reason for all what I have reached; Finally, I'd like to thank my fiancé Osama Abo Zied for his continuous love, support and encouragement to finish this work. His practical help was very invaluable especially, while executing the experiments and searching for specific publications, which I could not find.

Abstract

Object detection and tracking is an important task within the field of computer vision, due to its promising applications in many areas, such as video surveillance, human-computer interaction, intelligent transportation, etc. The availability of high quality and inexpensive video cameras and the increasing need for automated video analysis has generated a great deal of interest in the areas of motion tracking.

The commonly used technique to extract the foreground objects is through subtracting the background. Therefore, the basic operation needed is the separation of the moving objects from the background static information. The main goal of the object tracking is to estimate the locations and the motion parameters of the target, in an image sequence, given the initial position in the first frame.

Most of the current techniques modeled complex frameworks, with more accurate results, which are not suitable for real-time applications. On the other hand, simple detection and tracking techniques were proposed with low accurate results. There are many sources of uncertainty for the object locations, such as illumination variation, measurement noise, moving object's shadow, addition or removal of stationary objects and scene motion. There exists no single detection or tracking technique that can be successfully applied to all tasks and situations.

This research work addressed the problem of modeling a simple and robust technique for object motion detection and tracking in videos, capable of achieving high detection and tracking rates that can be successfully applied to most of the tasks and situations.

The major contributions of this research work can be summarized as follows: (a) Design a robust background modeling and subtraction technique for object motion detection, in unconstrained environments. It is capable of processing a real-time video and achieving high detection rates, using a stationary camera. (b) Design a robust object motion tracking-by-detection technique in unconstrained environments. It is capable of processing a real-time video and achieving high tracking rates, using the proposed detection technique. Finally, high performance records are achieved in the proposed technique, compared to other state-of-the-art techniques that targeted the same problems, on the same test case(s).

With the all-pervasive application of the object motion detection and tracking techniques in computer vision, this work is expected to offer useful contribution to the research community.

Chapter One: Introduction

This thesis is concerned with the problem of modeling a robust technique for object motion detection and tracking in videos, capable of achieving high detection and tracking rates, using a stationary camera, in unconstrained environments.

Our investigations for the recent trends of the object detection and tracking techniques revealed many limitations. These limitations include modeling complex techniques, with more accurate results, which are not suitable for real-time applications. On the other hand, modeling simple detection and tracking techniques, with low accurate results. There are many sources of uncertainty for the object locations, such as illumination variation, measurement noise, moving object's shadow, addition or removal of stationary objects and scene motion.

Therefore, a new robust technique for object motion tracking-by-detection is proposed, using a simple model that makes use of the generative model and the discriminative boosting classifier, in a particle filtering framework. It solved the following problems, which exist in the current state-of-the-art techniques: avoided the limitations of using only one feature, separated the foreground target efficiently (classification uncertainty), improved the adaptability of the detection and the tracking phases, and decreased both the computational complexity and the computing time of the technique.

1.1 Object Motion Detection and Tracking

Video analysis and video surveillance are active areas of research in the field of computer vision. A video surveillance system can monitor either immediate unauthorized behavior or long term suspicious behavior. Hence, it alerts the human operator for deeper investigation of the event. The key areas of visual surveillance are video-based detection and tracking, video-based person identification, and large-scale surveillance systems [1]. Visual object detection and tracking has drawn a significant attention, due to its enormous worth in surveillance, beside other applications like human-computer interaction, intelligent transportation, medical image processing, etc. [2].

Foreground object detection and segmentation is the first stage in the visual surveillance system. Background subtraction, temporal differencing, and optical flow are the main approaches used to achieve the foreground object detection and segmentation stage. The commonly used technique to extract the foreground objects is through subtracting the background [3]. Therefore, the basic operation needed is the separation of the moving objects from the background static information.

The main goal of the object tracking is to estimate the locations and the motion parameters of the target in an image sequence given the initial position in the first frame [4]. Object tracking can be classified into three major categories: Point Tracking, Kernel Tracking and Silhouette Tracking [5]. Point trackers are suitable for tracking very small objects, which can be represented by a single point. Kernel and Silhouette trackers provide more accurate description for objects with complex shapes.

Although, it has been studied for several decades, automatically detecting and tracking a variable number of objects, in complex scenes, from a single camera, remains a challenging problem [3] [4] . Despite the large number of background subtraction and tracking techniques [6] [7] [8] that have been proposed, the task remains challenging as there are many sources of uncertainty for the object locations, such as illumination variation, measurement noise, moving object's shadow, addition or removal of stationary objects and scene motion.

In order to cope with those difficulties, tracking-by-detection techniques have become increasingly popular, due to the recent progress in object detection [9]. Such techniques involved the continuous application of a detection algorithm in individual frames and the association of the detections across the frames. In contrast to background modeling based trackers, they are generally robust to changing background and moving cameras. The main challenge when using an object detector for tracking is that the detector output is unreliable and sparse [9].

A wealth of researches has been conducted for following objects robustly. Among these algorithms, Particle Filter, known as sequential Monte Carlo technique, exhibits its outstanding capability in tackling non-linear and/or non-Gaussian reality of visual tracking [2] [10]. The basic idea of Particle Filter (PF) is that the posterior density is approximated by a set of discrete samples (called particles) with associated weights [2].

The combination of object detectors and particle filtering resulted in algorithms that are more suitable for time-critical online applications, than the data association based tracking techniques [11]. State of the art object detectors built up some form of confidence density as one stage of their pipeline. It could