学校编码：10384　　分类号_____密级_____
学号：23020141154643　　UDC_____

厦门大学

硕 士 学 位 论 文

视频监控系统中多目标跟踪算法的研究

Study on the Algorithms for Tracking of Multiple Objects in Video Surveillance System

KALISA Wilson

指导教师姓名： 雷 蕴 奇 教授
专 业 名 称： 计算机科学与技术
论文提交日期： 2016 年 月
论文答辩时间： 2016 年 月
学位授予日期： 2016 年 月

答辩委员会主席：_____
评 阅 人：_____

2016 年 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（计算机视觉）课题（组）的研究成果，获得（计算机视觉）课题（组）经费或实验室的资助，在（　雷蕴奇　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　）1.经厦门大学保密委员会审查核定的保密学位论文，于　　年　　月　　日解密，解密后适用上述授权。

（ √ ）2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

DEDICATION

This work is dedicated to my parents and my dear Brothers and Sisters. May God bless you for your unconditional love!

摘 要

在视频监控系统中，多种图像处理技术已经被用来检测和跟踪对象。本文主要研究的是基于 RGB 颜色空间的多目标跟踪算法。实时跟踪算法现已经应用到导航，确定目标的位置及其运动状态。RGB 颜色空间是用来计算每种颜色的阈值，阈值由被跟踪的对象设置。线性模型用于颜色识别。为了避开框架中不需要的对象，对二值图像进行形态学处理和块分析，有效促进了目标的检测。本文提供的算法能够在 50fps 的 640×360 视频图像中完成对目标的跟踪。

本文引用了一种对象目标密切跟踪的架构，并对其中使用的两种技术来进行改进。在目标跟踪中，以 VIBE 算法作为目标分割的压缩感知跟踪算法是跟踪的核心。压缩感知的执行速度更加适用于实时跟踪，但同时也会跟踪到目标的阴影。

在复杂的情况下，本文尝试加强并提高多目标跟踪算法的鲁棒性，尤其在目标色调变化明显的情形下。为此通过在视频中不断更新目标在运动状态中的颜色变化来完成。基于目标颜色概率分布，颜色直方图和反射投影并利用 VIBE 分割算法来获取目标区域。目标颜色采用基于颜色概率分布、颜色直方图和背投影的 VIBE 算法来进行获取。通过我们的实验可以得出，本文提出的算法比原来的方法更加有效。

关键字：阈值二值化，RGB 颜色空间，形态学操作，块分析，VIBE 背景减除算法，压缩感知算法

Abstract

Diverse algorithms based on image processing techniques have been used to detect and track the objects in video surveillance systems. In this work, RGB color based tracking for multiple objects is presented. For this, real-time tracking algorithm is proposed for the objects to navigate, and identify the current positions and their movements. RGB color space is used to compute the threshold values of each color. Threshold values can be set depending on the object to be tracked. Linear model is applied for color discrimination. To remove unwanted objects from the frames, morphological processing is applied on binary image and then Blob analysis technique is used to facilitate the object detection. Our proposed algorithm is capable to track objects at full frame size (640×480 pixels) at frame rate (50fps).

This thesis also proposes the hypothetical structure of techniques for object tracking closely analyzed. It is in this context that we use two of its techniques as the premise for our change. Compressive Tracking is one of the most important methods based segmentation in the object tracking. Its speed and performance is more suitable for real-time tracking. It is also suitable for tracking objects unique color.

We try to enhance the robustness of this strategy for multi-object tracking in an intricate scene, particularly the checking of objects of various hues. This, by concentrating on the clew provides the color of the object in motion in the video. The area of the object is done by selecting each object by using the ViBe segmentation based on the color probability distribution, the color histogram, and the back projection. It appears in our experiments that the proposed strategy is more effective than original's.

Keywords: Thresholding, RGB color space, Morphological operations, Blob analysis, ViBe Background Subtraction, Compressive tracking.

Table of contents

List of Figures

List of Tables

List of Abbreviation

HIS: Hue, Saturation, Intensity

HLS: Hue, Lightness and Saturation

HSD: Hue Saturation and Darkness

HSV: Hue, Saturation and Value

NTSC: Sequential Color with Memory

OpenCV: Open Source for Computer Vision

PAL: Phase Alternation Line

RGB: Red, Green and Blue

RIP: Property isometric restricted

ROI: Region of interested

SURF: Speeded up Robust Features

VSE: ViBe square error

ViBe: Visual Background Extractor

CT: Compressive Tracking

RIV-LBP: Rotation Invariant Volume-LBP

LBP-TOP: LBP - Three Orthogonal Planes

CS-LBP: Center Symmetries – LBP

HOG: Gradient Oriented Histograms

Chapter 1 Introduction

1.1 Background

Prioritizing and Classify the different methods of motion detection found in the literature is not an easy task. Indeed, the authors often use different terminology to describe the methods or algorithms that have the same goal. On the other hand, the same term in a particular section may designate a small well-defined algorithm module, while in another; It refers to a computer system which this module as well as many others. In these circumstances, it is necessary to define precisely the terms found to avoid confusion.

1.2 Terms definition

Modeling background, combining all motion detection methods, is to create a model of background of the scene filmed (without any moving object). This model can be created from a pixel in the image observed at different moments of the film, as in [1], a arithmetical model of writing, the supply function of gray scale or hue (color) at any time (this is by far the most common approach since [2]), or a characteristic base of images is a subspace in which we consider that the pixels that represents the background will find [3].

Background subtraction follows logically the background modeling for motion detection. If the pattern of the background is an image, a difference in absolute value between the model and the current image is made for motion detection. When there is a statistical model, the probability that each pixel is calculated, which belongs to the background by testing the value observed in the model; the importance of the observed movement varies in the opposite direction of the computed probability.

Motion estimation includes the previous by adding the constraint that the results must be quantitative. This term is mostly used by authors who are working on the optical flow estimation; in this case, the result of the estimate is a vector field representing the card in two dimensional projection of the three-dimensional image of the real movement taking place in the scene. For more details and a comparative study of the traditional methods of optical flow calculation, see [4].

Detection of movement consists to find what is in the picture, an object of movement. An algorithm with this lens generates a variable quantitative (pulse) or qualitative (Boolean) for any pixel in each image input. All the methods presented below fall into this category.

Motion [-based] segmentation goes beyond the detection of movement, because they are segmenting each image into regions with homogeneity of the apparent motion. This operation is usually carried out separately from the estimation of optical flow [5] or spatio-temporal derivatives of the luminous intensity [64].

1.2.1 Areas covered

Most of the work published so far on the motion detection, are strongly oriented towards a specific area of application, and then included in others that are broader, dealing with comprehensive monitoring of moving objects.

For example, reference [65] presents a categorization (classification) on the comprehensive interpretation of the movement of human beings. Motion detection is discussed as a technique that can be used to segment each image between people and the land, as well as segmentation based on motion or on a morphological or chromatic model person.

In 2003, reference [66] present in the same area, similar method with a significant portion devoted to motion detection and segmentation of moving regions.

Reference [67] presented a classification of different CCTV systems. Motion detection is presented as the first step in any CCTV system.

In 2005, an overview of methods for modeling and subtraction of the background applied to traffic scenes is proposed in reference [68].

Not limited to a particular application domain, Piccardi [69] also reviewed several methods of subtracting the background and compared in terms of speed, memory requirements and accuracy. In reference [70], the authors present their own video surveillance system comprising a module for motion detection, a module for extracting regions of interest and a tracking module , and compare their motion detection algorithm nine others from the literature , which provides a good overview of known methods before 1999 . Furthermore, the authors identify seven types of difficulties that undermine the robustness of the best known algorithms, and derive five general principles that,

according to them, any algorithm modeling the background should follow to manage these difficulties:

- ✓Differentiation semantic objects should not be managed by the motion detection module;
- ✓Segmentation of objects must be correct as they appear in the scene;
- ✓It is necessary to define invariants across the pixel to characterize the values belonging to the background;
- ✓Model of the background must adapt as well to sudden changes than to gradual changes in the appearance of the background;
- ✓Model of the background must take into account changes that may occur at different scales of observation (pixel, region, and image).

Reference [71] presented five methods for local modeling background in the context of surveillance of outdoor scenes where the background may make movements that are not to be detected. By - modeling locally, the authors agree that each pixel has its own background model.

Reference [72] is to date the most comprehensive categorization of the automatic object tracking. The authors propose a hierarchical convincing description of the task in which the subtraction of the background is presented as a technique for detection of objects, as well as the extraction of points of interest, learning the appearance of objects from multiple views, or the segmentation of static images. In addition, the distinction between the different methods of motion detection is based on the observation scale used to build models. They differ primarily local modeling methods, that is to say the methods that create a statistical model at any point in the image without the notion of spatial neighborhood intervene. This family includes methods based on the difference between consecutive images and those who build a Gaussian or multi- Gaussian model at any point in the image. The second type of methods could be called semi- local. In this case, the neighborhood of pixels is taken into account during the construction process of the local background model. This is particularly the case of methods using a texture model to characterize the image points [73]. To a lesser extent, the articles offer a post- processing step to regularize the result of detection by standardizing measurements obtained in the same region [74] or to eliminate false positives by analyzing the connected components obtained after thresholding [75] can be classified in this category. The third family is that of global methods modeling (called methods -

holistic - in [72]).

In reference [76], we are talking about - change - detection between two views of the same scene. This task includes motion detection, when the two views are taken at consecutive times, but can go beyond where this is not the case. In the latter situation, the techniques used are not necessarily suitable for processing video sequences due to real-time constraints imposed in this case. Methods for motion detection are all presented methods of statistical modeling background.

In summary, among the articles which propose a classification of methods for detecting movement, there is a great disparity due essentially to the main purpose of the article. When the goal is to conduct a comparative study on a set of assessment orders , the number of methods presented is necessarily limited and it is not possible to " establish a hierarchical classification.

In the articles that aim to document a specific area of application, such as sequence analysis of traffic or interpretation of human movements, the proposed categories are often linked to the results provided by different algorithms on sequences of interest.

1.3 Overview and Motivation

Tracking of objects is an important task in the field of computer vision. The proliferation of large powerful computers, the accessibility of high quality and great video cameras markets and the developing requirement for computerized video examination has created a considerable measure of enthusiasm for article tracking algorithms. Tracking of multiple items in video surveillance systems is a critical undertaking for many applications, where developing the algorithms for fast image processing is one of the key research topics. Different algorithms based on image processing techniques have been applied to detect multiple objects and those often used methods include mean shift, optical flow, Kalman filter, etc [6]. And their emerging applications whether domestically, commercially, or militarily are becoming more and more sophisticated. By using image processing and computer vision techniques, the advanced camera appears to have more adaptable capacities to track objects. In [7] tracking of the vehicles is presented using a low-angle camera which was placed high above the ground to minimize the effect of occlusion. Based on image processing, Morphological and Blob analysis techniques have been commonly used in monitoring applications, pedestrian and human detection and in traffic surveillance systems [8]. In [9] Gaussian