IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

Moving Object Detection, a Succinct Review

Ruchira Susar[1], Mahendra Dongare[2]

M.E. Student & Research Scholar, Department of Electronics, Amrutvahini College of Engineering, Sangamner, India [1]

Professor, Department of Electronics, Amrutvahini College of Engineering, Sangamner, India [2]

Abstract: For many computer vision applications such as surveillance, vehicle navigation, even for autonomous robot navigation moving object detection is very important and challenging task. Detecting the objects in the video and tracking their motion to identify their characteristics has been emerging as a demanding research area in the domain of Image Processing and Computer Vision. Video surveillance in a dynamic environment, especially for humans and vehicles, is one of the current challenging research topics in computer vision. It is a very significant technology to fight against terrorism, crime, public safety. The work involves designing of the efficient video surveillance system in complex environments. In video surveillance, detection of moving objects from a video is important for object detection, target tracking, and behaviour understanding. Detection of moving objects in video streams is the very first and basic step for further processing.

Keywords: Video surveillance, moving object detection, background subtraction, optical flow.

I. INTRODUCTION

In the last years, smart surveillance has been one of the most active research topics in computer vision because of the wide spectrum of promising applications. Its main point is about the use of automatic video analysis technologies for surveillance purposes.

In general, a processing framework for smart surveillance consists of a preliminary motion detection step in combination with high level reasoning that allows automatic understanding of evolutions of observed scenes. In general a visual surveillance system with real-time moving object detection, classification, tracking and activity analysis capabilities is desired.

There have been a number of surveys about object detection, classification, tracking and activity analysis in the literature. The survey present here covers only that work which are in the same context as our study. However, for comprehensive completeness, we have also given brief information on some techniques which are used for similar tasks that are not covered in our study.

A generic video processing framework for smart algorithms is shown in Fig 1.This framework provides a good structure for this survey.

Fig 1: A generic framework for smart video processing algorithms.

Each application that benefit from smart video processing has different needs, thus requires different treatment. However they have something common: moving objects. Thus, detecting regions that correspond to moving objects such as people and vehicles in video is the first basic step of almost every vision system. Since it provides a focus of attention and simplifies the processing on subsequent analysis steps. Due to dynamic changes in natural scenes such as sudden illumination and weather changes, repetitive motions that cause clutter (tree leaves moving in blowing wind), motion detection is a difficult problem to process reliably. Frequently used techniques for moving object detection are background subtraction statistical methods, temporal differencing and optical flow whose descriptions are given below.

In a video sequence analysis for moving objection detection various methods are depicted in the following fig. 2.

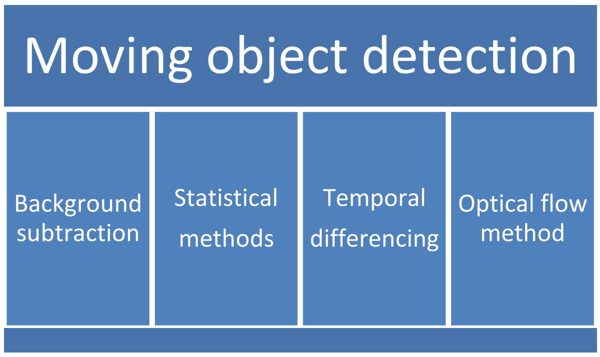

Fig 2: Taxonomy of moving object detection methods

A. Background Subtraction

Background subtraction is particularly a commonly used technique for motion segmentation in static scenes [4]. It attempts to detect moving regions by subtracting the current image pixel-by-pixel from a reference background image that is created by averaging images over time in an initialization period. The pixels where the difference is above a threshold are classified as foreground. After creating a foreground pixel map, some morphological post processing operations such as erosion, dilation and closing are performed to reduce the effects of noise and enhance

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

the detected regions. The reference background is updated with new images over time to adapt to dynamic scene changes. There are different approaches to this basic scheme of background subtraction in terms of foreground region detection, background maintenance and post processing.

In [7] Heikkila and Silven uses the simple version of this scheme where a pixel at location (x, y) in the current image It is marked as foreground if

$$|I_t(x, y) – B_t (x, y)| > Th$$

is satisfied where Th is a predefined threshold. The background image BT is updated by the use of an Infinite Impulse Response (IIR) filter as follows:

$$B_{t+1} = \alpha I_t + (1 - \alpha) B_t$$

The foreground pixel map creation is followed by morphological closing and the elimination of small-sized regions. Although background subtraction techniques perform well at extracting most of the relevant pixels of moving regions even they stop, they are usually sensitive to dynamic changes when, for instance, stationary objects uncover the background (e.g. a parked car moves out of the parking lot) or sudden illumination changes occur.

B. Statistical Methods

More advanced methods that make use of the statistical characteristics of individual pixels have been developed to overcome the shortcomings of basic background subtraction methods. These statistical methods are mainly inspired by the background subtraction methods in terms of keeping and dynamically updating statistics of the pixels that belong to the background image process. Foreground pixels are identified by comparing each pixel's statistics with that of the background model. This approach is becoming more popular due to its reliability in scenes that contain noise, illumination changes and shadow [9].

The W4 [8] system uses a statistical background model where each pixel is represented with its minimum (M) and maximum (N) intensity values and maximum intensity difference (D) between any consecutive frames observed during initial training period where the scene contains no moving objects. A pixel in the current image It is classified as foreground if it satisfies:

$$|M(x, y) - I_t(x, y)| > D(x, y) \text{ or}$$
$$|N(x, y) - I_t(x, y)| > D(x, y)$$

After thresholding, a single iteration of morphological erosion is applied to the detected foreground pixels to remove one-pixel thick noise. In order to grow the eroded regions to their original sizes, a sequence of erosion and dilation is performed on the foreground pixel map.

Also, small-sized regions are eliminated after applying connected component labelling to find the regions. The statistics of the background pixels that belong to the non-moving regions of current image are updated with new image data.

As another example of statistical methods, Stauffer and Grimson [5] described an adaptive background mixture model for real-time tracking. In their work, every pixel is separately modelled by a mixture of Gaussians which are updated online by incoming image data. In order to detect whether a pixel belongs to a foreground or background process, the Gaussian distributions of the mixture model for that pixel are evaluated.

C. Temporal Differencing

Temporal differencing attempts to detect moving regions by making use of the pixel-by-pixel difference of consecutive frames (two or three) in a video sequence. This method is highly adaptive to dynamic scene changes; however, it generally fails in detecting whole relevant pixels of some types of moving objects. A sample object for inaccurate motion detection is shown in Fig 3.

The mono colored region of the human on the left hand side makes the temporal differencing algorithm to fail in extracting all pixels of the human's moving region. Also, this method fails to detect stopped objects in the scene. Additional methods need to be adopted in order to detect stopped objects for the success of higher level processing.

(a) (b)

Fig 3: Temporal differencing sample. (a) A sample scene with two moving objects. (b) Temporal differencing fails to detect all moving pixels of the object on the left hand side since it is uniform colored. The detected moving regions are marked with red pixels.

Lipton et al. presented a two-frame differencing scheme where the pixels that satisfy the following equation are marked as foreground [3].

$$|I_t(x, y) – I_{t-1}(x, y)| > Th$$

In order to overcome shortcomings of two frame differencing in some cases, three frame differencing can be used [6]. For instance, Collins et al. developed a hybrid method that combines three-frame differencing with an adaptive background subtraction model for their VSAM project [1]. The hybrid algorithm successfully segments moving regions in video without the defects of temporal differencing and background subtraction.

D. Optical Flow

Optical flow methods make use of the flow vectors of moving objects over time to detect moving regions in an image. They can detect motion in video sequences even from a moving camera, however, most of the optical flow methods are computationally complex and cannot be used real-time without specialized hardware [6].

II. CONCLUSION

This paper provides a short and snappy study on the proposed techniques which have used for moving object detection. Frequently used techniques for moving object detection such as background subtraction; statistical methods, temporal differencing and optical flow are explained here. However, no object detection algorithm is perfect, since it needs improvements in handling darker shadows, sudden illumination changes and object occlusions. Higher level semantic extraction steps would be used to support object detection step to enhance its results and eliminate inaccurate segmentation.

ACKNOWLEDGMENT

I would like to express my true sense and a sincerest gratitude to my guide **Prof M.P.Dongare** for his dynamic and valuable guidance. I am grateful to him for constant encouragement in the fulfilment of this paper. This work is a result of combined efforts put in by my guide and me. I would also like to thank him for providing me with all necessary infrastructure and facilities to complete the paper.

REFERENCES

[1] R. T. Collins et al. A system for video surveillance and monitoring: VSAM final report. Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000.
[2] I. Haritaoglu, D. Harwood, and L.S. Davis. W4: A real time system for detecting and tracking people. In Computer Vision and Pattern Recognition, pages 962–967, 1998.
[3] A.J. Lipton, H. Fujiyoshi, and R.S. Patil. Moving target classification and tracking from real-time video. In Proc. of Workshop Applications of Computer Vision, pages 129–136, 1998.
[4] A.M. McIvor. Background subtraction techniques. In Proc. of Image and Vision Computing, Auckland, New Zealand, 2000.
[5] C. Stauffer and W. Grimson. Adaptive background mixture models for realtime tracking. In Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, page 246252, 1999.
[6] L. Wang, W. Hu, and T. Tan. Recent developments in human motion analysis. Pattern Recognition, 36(3):585–601, March 2003.
[7] J. Heikkila and O. Silven. A real-time system for monitoring of cyclists and pedestrians. In Proc. of Second IEEE Workshop on Visual Surveillance, pages 74–81, Fort Collins, Colorado, June 1999.
[8] A. Amer. Voting-based simultaneous tracking of multiple video objects. In Proc. SPIE Int. Symposium on Electronic Imaging, pages 500–511, Santa Clara, USA, January 2003.
[9] H.T. Chen, H.H. Lin, and T.L. Liu. Multi-object tracking using dynamical graph matching. In Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 210–217, 2001.
[10] R. T. Collins et al. A system for video surveillance and monitoring: VSAM final report. Technical report CMU-RI-TR-00-12, Robotics Institute, CarnegieMellon University, May 2000.