

A Survey on Motion Detection, Tracking and Classification for Automated Video Surveillance

Vidyashree H M¹, Dr.G Raghavendra Rao²

^{1,2}Department of ISE

^{1,2}National Institute of Engineering, Mysuru, India

Abstract- Moving object recognition and tracking motion has been the foundation source to dig out fundamental information about moving objects from sequences in uninterrupted image based surveillance systems. The assignment of target tracking is a key section of video surveillance and monitoring systems. Achieving a complete detection of an moving object which is robust against changes from the surrounding is presented here, methods used is pixel dependent and non parameterized based on first frames to construct a model. It provides input to high-level processing such as recognition, access control or re-identification or is used to initialize the analysis and classification of human activities. The proposed algorithm has been test implemented on several open source videos by imposing single set of variables to overcome shortcomings of relevant and recently developed techniques.

Keywords- Object Tracking; Motion Detection; Background Subtraction; Ghost Object; Video Surveillance; Background Modelling, Object Classification.

I. INTRODUCTION

Video surveillance has extensively be in use to observe security sensitive areas such as banks, department stores, highways, busy public places and borders. The progress in computing power, accessibility of huge-capacity storage space devices and high speed network infrastructure cemented the means for cheaper, multi sensor video surveillance systems. Conventionally, the video outcomes are done online by human operators and are generally saved to tapes for afterward use only subsequent to a forensic event. The enhance in the amount of cameras in regular surveillance systems filled to capacity both the human operators and the storage devices with high volumes of data and made it infeasible to ensure proper monitoring of sensitive areas. In order to sort out unnecessary information generated by an array of cameras, and enlarge the reaction time to forensic events, supporting the human operators with discovery of important events in video by the use of “smart” video surveillance systems has become a significant requirement. The construction of video surveillance systems “smart” must have fast, consistent and healthy algorithms for moving object detection, classification, tracking and activity analysis.

Identification and Tracking of object is an essential aspect in study of video in a surveillance system. It gives the withdrawal of the information from frames and video sequences which can be various processor vision applications for example, CCTV based surveillance, understanding an action in focal point, analysing flow of traffic, classifying and tracking an object. This exhibits that identifying and tracking an object is an essential field of research in computer vision and its applications in a variety of surveillance systems. CCTV based surveillance has turn out to be a challenging technology due to amplify in terrorist threats, increase in public\private safety concerns, increase in crime rate, efficient management of public properties and diverse modes of transportation.

Earlier period work in the field of 'Video Surveillance', have existing a variety of different methods every one having different pros and cons. In existent word scenario Alan generation in a tracked/traced event is dependent upon accuracy of these proposed models on a variety of researches. By far the most accurate methodology [1] developed has an correctness rate of more than fifty percent; this method does a two phased background detection using parametric method, for optimum results. Also it employs background elimination method so as to decrease processing load. The main restriction is the rectification of ghost background. This methodology can be taken forward with a mix of earlier discussed method to achieve a better accuracy rate and hence making it reliable system.

Recognition of object and its motion with slightest amount of processing is required, which is compromised in researches every time high accuracy is met. Whenever processing is compromised it says more number of CPU with high & speedy processing capacity, so as not to end up in a scenario where generating an alarm after a long .

Colour based isolation of background and object and both the objects becomes outdated in a scenario where occlusion or repetition occurs, as in case of tracking a white car the system may switch its focus to another white car overtaking the marked car.

In certain researches compression has been utilized for fewer processing, but compression causes integration of pixels of similar colours which makes detection and segregation of objects difficult. Most researches do not accomplishing accommodate for shadow casted and change in light over the background or object. Such changes may cause complete abstraction of object or appearance of new object.

II. RELATED WORK

A sequence of methods for detection of motion have been designed and developed in the past decade. One of the generally recent and significant high precision technique being temporal differencing which is well thought-out to be most advanced of all.

The temporal differencing technique employs pixel-by-pixel difference between consecutively incremental frames, from then on the threshold is determined based on averaging of differences to set up the foreground object. The mentioned technique has a shortcoming that if it is exposed to any illumination or any structural variations then it causes interminable ghosts when first frame displays any moving object.

There are other techniques for video analysis which are based on Gaussian Mixture Models which are used in complex scenes such as road traffic where each pixel is a result of multi-dimensional Gaussians mixtures. This Gaussian mixture model is resistant to the issue's caused due to modification in illumination or repeated motion due to close proximity of the moving object creating a cluttered object scenario and scene changes over extended duration.

A simple mathematical algorithm like Extended Kalman filter to distinguish specific dangerous events at a railway station. This intended methodology had a strike rate of 6%-22% for different sized object, hence failure to recognize object on the first hand will nullify the probability of event detection and might result into either false alarm or no alarm generation. An additional technique is one technological invent which has high accuracy of detection of objects which tracks human hands and head based on 2-D modelling and is subjected to very frequent limitation of not being capable to differentiate between backgrounds and object in case of near pixel match of colours, fails upon dynamic change in background and it makes assumptions to function in a domain specific environment.

Improvement was made over this in later research. Clarification came out as using 3D modelling over 2D modelling for topological layering. W4 gave a new approach

by not using colour cues but utilizing the grey scale monochromatic picture or feed from infrared camera to make 3D model of the subject, but this approach is way too slower as 3D modelling of an object or all subjects in view is too heavier load for CPU to process in real time practical scenario. A unique study has been prepared with referencet to human behaviour where escape actions of human indicates and confirms classification of event. This is also Bayesian based methodology. Presently this has restricted scope of function while this approach may lead way into expansion of solution to address other fields of surveillance and challenges. Another IEEE publication based on educating real time motion pattern of specific object, identified as a subject of study; this utilizes mix of Gaussians method for analysis. This method is susceptible to background randomness.

In 2010 Pheng Ann Heng(Sr. Mem.IEEE), Qian Chen and Group Track Two-Stage object, a technique based on Kernal and Active Contour. They can situate an object efficiently in composite situation with camera motion, limited occlusions, clutter etc the dispersion snake is used to change the object contour in classify to improve the tracking precision. In the first object localization stage, the primary target position is forecasted and estimated by the Kalman filter and the Bhattacharyya coefficient, respectively. In the contour evolution stage, the lively contour is evolved on the origin of an object feature image generated with the color information in the initial object region. In the process of the evolution, similarities of the objective region are compared to make sure that the object outline evolves in the right way. They having following disadvantages

1. This method is time consuming.
2. Method can't effectively track the object when the color feature of the object is very much related to that of the background.

In 2010 Michael, Fabian Bastian and Group of IEEE member Track and Detect "Online Multi-Person from a single Uncalibrated Camera using Partial Filter."The algorithm detects and tracks a huge number of dynamically moving persons in composite scenes with occlusion. Their algorithm uses the uninterrupted assurance of pedestrian detectors and online trained, instance-specific classifiers as a graded inspection model. Thus, common object category knowledge is balanced by instance-specific information. They analyze the influence of different algorithm mechanism on the robustness. It requires a more sophisticated structure than Partial Filtering. In 2011 Amir Salarpour, Arezoo Salarpour, Mahmoud and MirHossein Track "A Vehicle using Kalman Filter and Features." It identify all moving objects. The technique can differentiate and tracking all vehicles independently and work

in clutter scenes satisfactory. They sense all moving objects, and for tracking of vehicle they use the Kalman filter and color feature and distance of it from one frame to the next. So the technique can differentiate and tracking all vehicles individually. This method has tracking problem such as .Appearance

Object Detection

Robust and reliable detection and tracking has been focused a lot of consideration in recent years, motivated by applications such as pedestrian protection, vehicle platooning and autonomous driving. This is a complicated problem, which becomes even harder when the sensors (e.g. Optical sensors, radar, laser scanners) are build up on the vehicle rather than being fixed, such as in traffic monitoring systems. Effective detection and tracking wants accurate measurements of object situation and motion, even when the sensor itself is moving. Unfortunately, merely deducting out ego-motion does not remove all the possessions of motion because the apparent object's shape emerge to change as different aspects of the object come into view, and this change can easily be get the wrong idea about as motion. Plus, the supposed appearance of an object depends on its pose, and can also be affected by close by objects. Finally, difficult outdoor environments regularly engage cluttered backgrounds, irregular interaction between traffic participants, and are difficult to control. The subsequent figure illustrate that a person gets detected through a moving camera or it can be detected by using stationary camera.



Figure 1. Example of moving Object Detection from a moving Camera

Object tracking



Figure 2.

Object tracking is an significant component of numerous computer vision systems. It is broadly used in video surveillance, robotics, 3D image reconstruction, medical imaging, and human computer interface .The expansion of high-performance computers, the convenience of high quality yet inexpensive video cameras, and the growing need for automated video analysis has caused a vast deal of interest in object tracking algorithms. There are three main steps in video analysis: detection of exciting moving objects, tracking of such objects from frame to frame, and investigation of tracks to recognize their behavior. The object tracking is pertinent in the tasks of:

- Motion-based recognition, that is, human identification based on gait, automatic object detection, etc.
- Automated surveillance that is, screening a scene to detect suspicious activities or unlikely events
- Video indexing, that is, automatic annotation and retrieval of the videos in multimedia databases
- Human-computer interaction, that is, gesture recognition, eye gaze tracking for data input to computers,
- Traffic monitoring, that is, real-time assembly of traffic statistics to direct traffic flow
- Vehicle navigation that is, video-based path setting up and obstacle avoidance capabilities.

Object classification

Object classification step categorizes detected objects into predefined classes such as human, vehicle, animal, clutter, etc. It is essential to make a distinction of objects from each other in order to track and examine their proceedings reliably. The classification algorithm put together use of the shape of the detected objects and tracking results to successfully categorize objects into pre-defined classes like human, human group and vehicle. Proposed method make use of the aspect ratio of the blob information to classify detected objects.

III. PROPOSED ALGORITHM

Most of the method initialize their model by make use of sequencing of frames such as described in [1], [2]. Sequencing of frames has verified to be very precise for statistically modelling the estimation of pixels in the background from the temporal spatial distribution perspective. The technique has verified to be resistant to the environment's noise this happens because the approach provides sufficiency of information with position to further sequential frames. Upon subjecting to illumination changes during the early sequence the turns out to be a failure. In such a case of alternative illumination the background model is discarded and it is necessary to construct another background model. Another inadequacy comes into picture if the measurement lengthwise of analysis video is shorter than the training video, this happens due to insufficient information.

Further work by authors of [3] rolls out to be a resolution to this problem. The solution is to make use of a single frame for initialization of background model. Each pixel in the spatial proximity is selected for initializing the model. This causes to share the information with adjacent pixel as is the case in temporal distribution as single frame is not able to retain temporal information. The collection of neighbouring pixels is based on the theory to be able to accumulate information of uniquepixels. This corrective action leads segmentation of foreground from the frame subsequent to initial or first frame.

It is apparent that one frame cannot hold entire and precise details color scheme and pattern of brightness in the succeeding frames, which results into a inadequacy of this method to be non-resistant to noise from the environment. There are a number of other techniques proposed where ghost is inserted into the first frame while object is in motion, as in [4].

The anticipate or proposed algorithm integrates Background subtraction with normalized graph cut segmentation which is robust against any changes in the illumination of the frame or the "ghosts" which are left by removed or extracted objects. Further, the algorithm performs a step by step tracking and classification of the detected object with more accuracy and within minimum processing time. Fig. 1 describes the methodology with functional blocks of proposed algorithm in the form of a flowchart, as shown below:

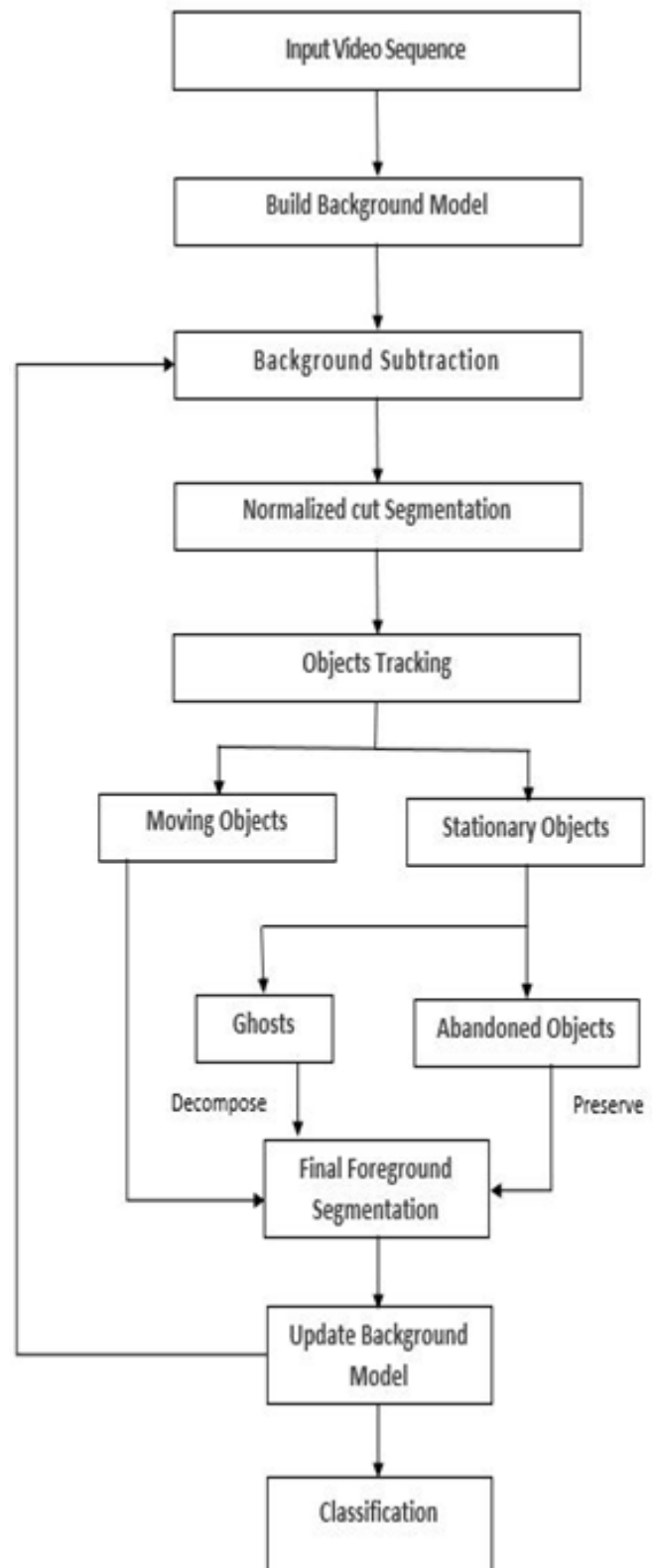


Figure 3. Flow chart of proposed algorithm

IV. CONCLUSION

The proposed method presents real-time firmness of video flow with high reliability to detect and track the object in the frame sequence. It describes an advanced algorithm

based framework which is capable of producing background with almost null or “NO” noise pixels. The proposed method overcomes the trails of artificial ghosts. This method has high efficiency in terms of accuracy and low or reduced processing requirements

[11] Christopher Wren, Ali Azarbayejani, et.al, "Ptinder: Real-Time Tracking of the Human Body", Proceedings of the 2nd international IEEE conference on Automatic Face and Gesture Recognition, pp. 51-56, 14-I60ct. 1996.

REFERENCES

- [1] Shih-Chia Huang, "An advanced motion detection algorithm with video quality analysis for video surveillance systems", *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 21, no. I, pp. 1-14, Jan. 2011.
- [2] T. Darrell, G. G. Gordon, M. Harville, and J. Woodfill, "Integrated person tracking using stereo, color and pattern detection," *Int. J. Comput. Vis.*, vol. 37, pp. 175-185, Jun. 2000.
- [3] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 22, no. 8, pp. 809-830, Aug. 2000.
- [4] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Trans. Syst.*,
- [5] Man, *Cybern. C, Appl. Rev.*, vol. 34, no. 3, pp. 334-352, Aug. 2004.
- [6] Lucia Maddalena and Alfredo Petrosino, "A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications", *IEEE Trans. on Image Processing*, vol. 17, no.7, pp. 1168-1177, July 2008.
- [7] Gonzalo Pajares, "A Hoptfield Neural Network for Image Change Detection", *IEEE Trans. on Neural Networks*, vol. 17, no. 5, pp. 1250-1264, Sept. 2006.
- [8] Ismail Haritaoglu, David Harwood and Larry S. Davis, ": Real- Time Surveillance of People and Their Activities", *IEEE Trans. on Pattern Analysis And Machine Intelligence*, vol. 22, no. 8, pp.809-830, Aug. 2000.
- [9] S. Dockstader and M. Tekalp, "Multiple camera tracking of interacting and occluded human motion," *Proc. IEEE*, vol. 89, no. 10, pp. 1441- 1455, Oct. 2001.
- [10] S. Park and J. Aggarwal, "A hierarchical bayesian network for event recognition of human actions and interactions," *Multimedia Syst.*, vol. 10, no. 2, pp. 164-179, Aug. 2004.