

# A Study on Computer Vision Systems for Real-Time Object Detection and Tracking

Daniel Mohammed

Jiangu University of Science and Technology, China  
School of Electronics and Engineering

Francis A. Amavi

Data Link Institute, Ghana  
School of Computer Sciences

## ABSTRACT

Computer Vision (CV) concentrates on the automatic extraction, examination and comprehension of valuable data from a solitary image or a group of images. Object tracking, one of the key areas in CV has received a lot of attention in recent times. Tracking objects is a systematic process of monitoring the movement of a target object from its initial state to the  $n^{\text{th}}$  state over a period of time using a camera. This technique is usually employed as a security feature in both military and civilian systems. However, prior studies has shown that tracking objects in motion is a very difficult task and is a hot research hotspot in the field of computer vision and machine learning. In this review paper we discuss various techniques in detection, tracking and some other related works of moving objects in video streams.

## Keywords

Automatic Extraction, Object Detection, Object Tracking, Computer Vision and Machine Learning.

## 1. INTRODUCTION

Object tracking a surrogate of computer vision which attempts to identify recognition and track series of images. Moreso, it makes an effort to understand and describe the parameters of objects to substitute the usual ways of aging of surveillance cameras by identified users [1]. Living and inanimate movable objects such as humans, animals, swaying tree, floating clouds and man-made movable objects such as vehicles, robots, traffic control and any other moving objects could be tracked [2]. As a rule, detection and tracking of moving objects in video stream is one of the continuous reviews in computer vision and numerous algorithms have been recommended. Vision is a complex process and it requires systematic steps for analysis as humans have a limited vision especially for fast moving objects.

Detection and recognition is very important in intelligent systems [3]. Generally moving object(s) are detected in a video stream as a first step, classified before tracking. Another challenging factor in tracking moving object in video stream is when the orientation of the object in question changes over time. Lighting conditions, shadow and occlusion are some of the challenges faced when tracking. In this paper we discussed the general algorithms deployed for detection and tracking of objects. Tracking systems are in a high demand at both private and public domain for example; residential areas, airports, train stations, schools, offices, subways and mass events places.

Detection is the process of identifying the presence of object concealed. As much as numerous algorithms have been suggested, detection and tracking problems such as occlusion and illumination have not been completely solved [4]. The general question is why do we track? also, is tracking

essential by any stretch of the imagination? The appropriate response is truly straightforward, for wellbeing. As wellbeing is the essential key that is considered as society anticipate forestalling catastrophe and wrongdoings in securing people [5]. Methodologies use in tracking needs intelligence to identify moving object automatically in a frame when the target object show up for the first time [1]. Computer systems for instance use hardware and software for tracking which are generally expensive, likewise installation and setup procedures are also complex.

Detection and tracking of fast moving objects are complex and time-consuming, as humans have constrained vision for example; (i) limited memory that is sometimes it's difficult for humans to remember a quickly flash images (ii) limited to visible spectrum and lastly (iii) illusion. In a video stream, the rate at which images are displayed is very fast at a minimum scanning frequency of 50 Hz which is very deceptive to the human eyes as seeing a steady moving image; anything lower and the eyes would see a flicker. Object detection in a video sequence can be grouped into sets of pixels such as foreground and background pixels [2]. Foreground objects in tracking can be described as dynamic objects meaning they change in direction or angle and moves whiles background objects such as building, road, wall, mountain and many others are static. Normally, it is difficult to select foreground from a background pixels of the same color or texture. Many methods have been suggested for detection of moving objects. For instance frame differencing, point detectors, segmentation, optical flow, temporal differencing and background subtraction [6, 7], [2, 8] likewise various approaches have been suggested for tracking objects such as point tracking, kernel tracking and silhouette.

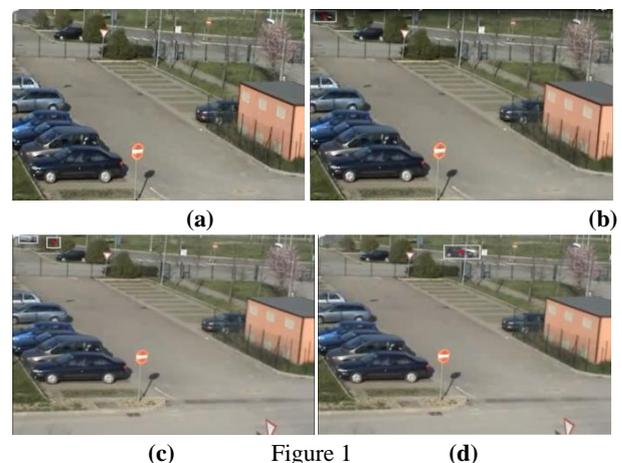


Fig. 1 shows overview of detecting and tracking of moving vehicle using a single static camera at a car park within a time period of 1 minute 52 seconds. As time  $t = 00:02s$  (a)

moving object was detected. A moving vehicle was detected (b) at the far left corner at the time  $t=00:03s$ . At the time  $t=00:04s$  two vehicles (c) were detected over a distance but the targets are very small. The rest of this paper is organized as follows: in section 2 we discussed the previous related works. Section 3 presents some of the approaches proposed for object detection, object tracking, benefits of tracking and some challenges in tracking. Finally, conclusion and the future works of this paper is discussed in Section 4.

## 2. RELATED WORKS

This section discusses some related works done on object detection and tracking in the arena of computer vision. Many research works have been done in object detection and tracking such as; in [22], Visual Object Tracking using Adaptive Correlation Filters, to track complex objects through rotations, occlusions and other distractions at over 20 times the rate of the current state-of-the-art techniques. Yilmaz, et al. [23] proposed a target tracking in FLIR imagery using mean-shift and global motion compensation. The approach uses the target intensity and the Gabor response distributions and computes a likelihood measure between the candidate and the model distributions by evaluating the Mean Shift Vector (MSV). When the MSV based tracker fails to locate the target due to large ego-motion, the system compensates the ego-motion using a multi-resolution scheme, which employs the Gabor responses of two consecutive frames, and assumes a pseudo-perspective motion model. Arbeláez, et al. [9]

Paper [24] proposed a novel design for region-based object detectors that integrates efficiently top-down information from scanning-windows part models and global appearance cues. Divya M [25] developed an algorithm for detecting, tracking of moving objects from the source by using LABVIEW vision module. An algorithm for object detection and tracking, where the current location is searched based on the histogram of the object in the previous image frame and result of mean shift is used to find the peak of probability density function near the object old position. The algorithm works well so long as the initial state of the target object is determined or located.

Dehghan and Shah [26] also proposed Binary Quadratic Programming for tracking of hundreds of people in extremely crowded scenes. The aim is to accomplish the computational complexity in the tracking of multiple objects in a high-density crowded scenes. Gyaourova, et al. [1] proposed a Block Matching Algorithm for object tracking. The goal of the proposed algorithm was in two folds; (i) to explore the abilities of the block matching algorithm on low resolution and low frame rate video and (ii) to improve the motion detection performance by the use of different search techniques during the process of block matching. The algorithm worked well on low resolution and low frame rate video data. Its aims at detecting the motion between two images in a block-wise sense.

Other algorithms such Kalman filter, Adaptive Kalman filter, Particle filter and others were proposed but few of them aid for the purpose of predicting the initial searching point and improving tracking performance for the next frame [27]. An experimental result in [27] shows that, adapting Kalman filter algorithm can achieve an extremely high accuracy of predicting parameters and hence a significant decrease in the risk of being confused by background interferences. As discussed in [28], in the field of Computer Vision the Forward Module use to develop in physics (radiometry, optics and sensor design)

## 3. OBJECT DETECTION

Detecting objects turn out to be difficult when we have to deal with extensive features, occlusion and illumination [10]. Primarily, object detection suitable for tracking depends on the framework or location in which the tracking is achieved. For an accurate tracking to be achieved, it depends on detection. If the object in question is not detected correctly this may affect tracking of the said object. In computer vision and pattern recognition analysis, the key challenge is object detection [11]. The following are some of the approaches to object detection;

- i. **Frame Differencing:** it is also called Adjacent Frame Difference Method [12], is extensively use to capture images using static cameras [13]. The goal of the approach is to determine the difference in image from frame to frame of the moving object.
- ii. **Optical Flow:** One of the features selection for detection of object for tracking is optical flow. It defines the change in position for a feature displayed in two succeeding images [14]. This approached shows absolute movement of object compared to others with a high accuracy rate [11]. The drawback of this approach is that it exceptionally affects ability to noise and poor anti-noise execution [15], which make it less appropriate for real-time detection and tracking.
- iii. **Background Subtraction:** This is a well known approach for foreground segmentation in the field of computer vision. The problem of illumination variation, background clutter and shadows can be handled by background subtraction [1]. Tracking become more and more difficult when the target object and the background appears in the same color or texture. Any little change in the image position from the background shows the moving object [11]. Article [1, 6] discussed the various approaches for background subtraction as Spatial Information (SI), Hidden Markov Models (HMM) and Eigen Space Decomposition (ESD).
- iv. **Point Detector:** This is one of the detection method used in tracking when object first appears in a video. They provide interested points in images which are expressive texture in their respective localities . This method has been used long ago in solving tracking problems. In [6] Harris detector method was employed to analyze the first order image derivations,  $(I_x, I_y)$  where  $x$  and  $y$  guide to highlight the directional intensity variations, whiles the second moment matrix which encodes the variation is estimated for individual pixel in a small zone as; the algorithm is well explained in paper [6, 16].
- v. **Segmentation:** This is by separating image into its perceptual similar groups, for example intensity, color, textures, patterns etc. But in real time conditions issues for example, limited spatial resolution, noise, poor contrast features, overlapping intensities and many more creates difficulties in segmentation task [17]. To beat this, [18] suggested Silhouette-based posture analysis. Segmentation becomes more difficult especially in areas where foreground and background are identical in terms of color or texture.
- vi. **Temporal differencing:** This technique uses at least two sequential frames to remove moving locales. This strategy is powerless and inclined to false detection if

the fleeting changes zone unit created by noise or light change because of climate conditions.

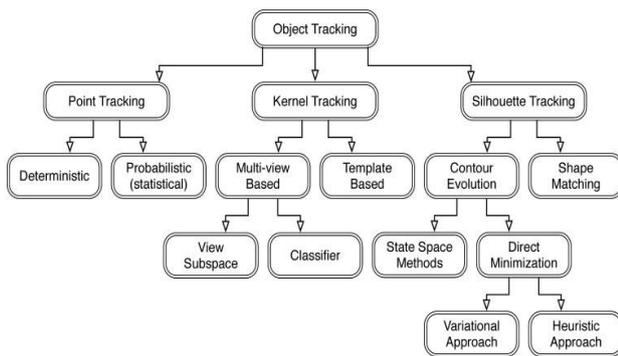


Figure 2: shows the overview of object tracking

### 3.1. Object Tracking

One of the on-going topics being discussed in the area of computer vision is object tracking. Tracking could be described as studying the path of the object, from its initial state as it moves around in different directions [19]. The purpose of tracking an object is to generate the route for an object over time by finding its position in every single frame of the video [11] [2].

Object detection and tracking are interrelated because tracking usually starts with detection while detecting an object repeatedly in a subsequent image sequence is often necessary to help and verify tracking. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video [2]. In general, the three basic tracking methods such as; point-based tracking, kernel-based tracking and silhouette-based tracking method and many more have been suggested for tracking moving object. Paper [10] also explained object tracking as while as background model which is employed in tracking for a successful segmentation of background and tracking of moving object in video stream.



(a) Figure 3 (b)

In figure 3 (a) and (b) shows moving cars that has been detected and tracked as they appears in a domain of tracking at a car pack. The reference video is taken by a camera when the system detect a moving object in a scene.

Performances in trackers are evaluated by distinct measures; that is accuracy and robustness. Tracking systems must be design to be robust to produce accurate result.

i. **Accuracy:** Accuracy is critical in systems, for example, recognition and tracking of objects in the field of CV, and can be characterized as the measure of vulnerability in an estimation as for a flat out standard. As defined in [20], the accuracy of a tracking system can be named by the precision of the angle detection process nor by considering the result of the position approximation process. Some factors that need to be deliberated in

tracking accuracy are noise, illumination and distortion. Detection accuracies are increased when object classification techniques are well deployed [21].

ii. **Robustness:** Robustness can be described as the ability of a control system to handle errors throughout execution [22] and manage with invalid input. Object tracking systems are developed to be stout to handle errors and increase system performance likewise robustness can be increase in trackers. A robust system must remain functional irrespective of the existence of faults in its component subsystems or parts. When there is a failure in any part of the system, performance reduced until the faults are fixed. Robustness consists of reliability and availability; reliability measures the frequency of failures and availability replicates the fraction of the time when the system is debugged.

### 3.2. Benefits of Object Tracking

There are several benefits in object tracking. This have serve as a biggest security measure in many places for example; airport, train stations, schools, subways, shops, companies, traffic controls and many more.

- i. It is helpful to detect articulated objects
- ii. It provide surveillance, augmented reality, traffic control and medical imaging
- iii. It provide state-of-the-art security measures in both military and civilian systems
- iv. Manage and control movement.
- v. It enhances safety
- vi. Tracking help when detection fails
- vii. Tracking conserves identity

### 3.3. Challenges in Tracking

In early 1970's when computer vision began, it was seen as a visual perception component of an ambitious agenda to simulate human intelligence and endow robots with intelligent behavior [23]. Computer vision and its algorithms are trying to solve the inverse problem to recover some unknowns given insufficient information to fully specify a solution [23]. Although many studies have research bottomless in the field of computer vision and different methodologies proposed in detection and tracking of objects, there are many challenges face by trackers in tracking, and are not fully addressed by existing approaches. Some of these approaches introduced the use of multiple methodologies or techniques and there are combinations and intersections among different methodologies as most of the proposed algorithms has limitation. The goal of tracking is to identify all foreground regions as long as they remain visible in a frame [24]. Stated below are some of the challenges;

- i. Some of the target objects sizes often appear very small in the range of 20–30 pixels. A small number of pixels on a target object makes it significantly difficult to distinguish it from the background and noise. To solve this, an update mechanism needs to be integrated to handle the dynamic changes in appearance, size and shape models.
- ii. Another challenging problem in tracking is occlusion. Occlusions which is caused by the terrain features such as trees, buildings, bridges, etc. When a target object occluded and reappears again in a different scene, tracking becomes difficult. For example, tracking target object becomes more complex or tracking normally lost when tracking

object moves under trees, behind buildings or near other moving target object. This can be achieved through feature matching, when an object that reappears in the scene are treated as a new object where the spectral features of the new target are compared with those of previous targets.

- iii. Thirdly, other issues related to tracking such as, automatic initialization, automatic recovery after the loss of track, or tracking beyond known positions. Preferably tracking should begin immediately a target object appears in a scene. If tracking objects are not detected, systems used for tracking should have a method to automatically re-initialize the tracking process.

#### 4. CONCLUSION

Object tracking is a key task in computer vision which is an open research ongoing. Tracking is important in terms of video monitoring for both civilian and military systems. In this paper we discussed the numerous approaches for detection and tracking of moving object. There are various challenges and benefits in tracking. The future improvement of this paper will aid to track moving objects using Global Positioning System (GPS) and application will be improved to function efficiently and effectively.

#### 5. ACKNOWLEDGMENTS

First and foremost we offer our sincerest gratitude to the Almighty God for His protection and guidance and making it possible for us to live and study throughout this years. We would like to express our gratitude to our families and our employers.

#### 6. REFERENCES

- [1] Rout, R.K., *A survey on object detection and tracking algorithms*. 2013.
- [2] Parekh, H.S., D.G. Thakore, and U.K. Jaliya, *A survey on object detection and tracking methods*. International Journal of Innovative Research in Computer and Communication Engineering, 2014. **2**(2): p. 2970-2979.
- [3] Chen, Y.-L., et al., *A real-time vision system for nighttime vehicle detection and traffic surveillance*. IEEE Transactions on Industrial Electronics, 2011. **58**(5): p. 2030-2044.
- [4] P, P.R.K., *Detecting Abandon Objects Fastly Through Blob Analysis*. IJREAT International Journal of Research in Engineering & Advanced Technology, , 2014. **Volume 2**(Issue 2).
- [5] Gyaourova, A., C. Kamath, and S. Cheung, *Block matching for object tracking*. Lawrence livermore national laboratory, 2003.
- [6] Yilmaz, A., O. Javed, and M. Shah, *Object tracking: A survey*. Acm computing surveys (CSUR), 2006. **38**(4): p. 13.
- [7] Dedeoglu, Y., *Moving object detection, tracking and classification for smart video surveillance*. 2004, bilkent university.
- [8] Ranipa, K.R. and K. Bhatt, *Illumination Condition Effect on Object Tracking: A Review*. Global Journal of Computer Science and Technology, 2014. **14**(5-F): p. 9.
- [9] Arbeláez, P., et al. *Semantic segmentation using regions and parts*. in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. 2012. IEEE.
- [10] Chen, X., et al. *Detect what you can: Detecting and representing objects using holistic models and body parts*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [11] Sukanya, C., R. Gokul, and V. Paul, *A Survey on Object Recognition Methods*. International Journal of Science, Engineering and Computer Technology, 2016. **6**(1): p. 48.
- [12] Alex, D.S. and A. WAHI, *BSFD: Background Subtraction Frame Difference Algorithm For Moving Object Detection And Extraction*. Journal of Theoretical & Applied Information Technology, 2014. **60**(3).
- [13] Singla, N., *Motion detection based on frame difference method*. International Journal of Information & Computation Technology, 2014. **4**(15): p. 1559-1565.
- [14] Helgesen, H.H., *Object detection and tracking based on optical flow in unmanned aerial vehicles*. 2015, NTNU.
- [15] Joshi, U. and K. Patel, *Object tracking and classification under illumination variations*. 2016.
- [16] Roberts, W., et al. *Vehicle tracking for urban surveillance*. in *SPIE Defense and Security Symposium*. 2008. International Society for Optics and Photonics.
- [17] Gould, S., T. Gao, and D. Koller. *Region-based segmentation and object detection*. in *Advances in neural information processing systems*. 2009.
- [18] Haritaoglu, I., D. Harwood, and L.S. Davis, *W 4: Real-time surveillance of people and their activities*. IEEE Transactions on pattern analysis and machine intelligence, 2000. **22**(8): p. 809-830.
- [19] Zhang, R. and J. Ding, *Object tracking and detecting based on adaptive background subtraction*. Procedia Engineering, 2012. **29**: p. 1351-1355.
- [20] Pentenrieder, K., P. Meier, and G. Klinker. *Analysis of tracking accuracy for single-camera square-marker-based tracking*. in *Proc. Dritter Workshop Virtuelle und Erweiterte Realitt der GIFachgruppe VR/AR, Koblenz, Germany*. 2006.
- [21] Mishra, P.K. and G. Saroha, *A Study on Classification for Static and Moving Object in Video Surveillance System*. 2016.
- [22] Fernandez, J.-C., L. Mounier, and C. Pachon. *A model-based approach for robustness testing*. in *IFIP International Conference on Testing of Communicating Systems*. 2005. Springer.
- [23] Szeliski, R., *Computer vision: algorithms and applications*. 2010: Springer Science & Business Media.
- [24] Bhattacharya, S., et al., *Moving object detection and tracking in forward looking infra-red aerial imagery, in Machine Vision Beyond Visible Spectrum*. 2011, Springer. p. 221-252.