Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 3 Issue 9, September - 2016

Multiple Object Recognition

Sabri Mohammed Alamin
College of Computer Science & Information Technology
Sudan University of Science and technology
Khartoum, Sudan
sabri_alamin@hotmail.com

Othman O Khalifa
Electrical and Computer Engineering
International Islamic University Malaysia
Kuala Lumpur, Malaysia
Khalifa@iium.edu.my

***Abstract*—with the increased processing power and emerging an advanced technologies, industries challenging for more demands for human life. Many algorithms have been proposed to solve the recent unsolved problems of object recognition, however, it still lack of tracking multiple objects in real time. Object detection and recognition in noisy and crowded area is still a challenging difficulty in the area of computer vision. The objective of this system is to identify the different objects using some techniques. Mainly there are three basic steps in video analysis: Detection of objects of interest from moving objects, Tracking of that interested objects in consecutive frames, and Analysis of object tracks to understand their behavior. Simple object detection compares a static background frame at the pixel level with the current frame of video. The existing method in this domain first tries to detect the interest object in video frames. One of the main difficulties in object tracking among many others is to choose suitable features and models for recognizing and tracking the interested object from a video.**

Keywords—Object detection, Frame difference, Background subtraction, Object Tacking

I. INTRODUCTION

In today's extremely automated and advanced industries, extremely efficient methods are used for various inspection and production processes. Object detection and tracking is an important challenging task within the area in Computer Vision that try to detect, recognize and track objects over a sequence of images called video. It helps to understand, describe objects behavior instead of monitoring computer by human operators. It aims to locating moving objects in a video file or surveillance camera. Object tracking is the process of locating an object or multiple objects using a single camera, multiple cameras or given video file. Invention of high quality of the imaging sensor, quality of the image and resolution of the image are improved, and the exponential increment in computation power is required to be created of new good algorithm and its application using object tracking. In Object Detection and Tracking we have to detect the target object and track that object in consecutive frames of a video file.

A generic video processing framework for smart algorithm [1] is shown in Figure 1. This framework provides a good structure for Robust Vision-based Moving Target Detection and Tracking System because it can be consider as a surveillance applications. Each application needs different requirements to use smart video processing by efficient way.

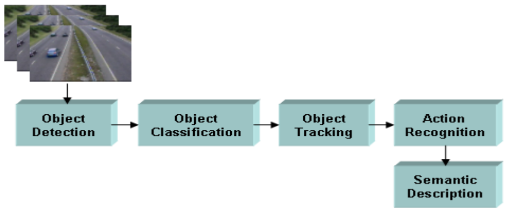

Fig. 1. *A generic framework for smart video processing algorithms.*

II. RELATED WORK

Tracking is the process to locating the interested object within a sequence of frames, from its first appearance to its last. The type of object and its description within the system depends on the application. During the time that it is present in the scene it may be occluded by other objects of interest or fixed obstacles within the scene. A tracking system should be able to predict the position of any occluded objects. In [2], the author suggests an algorithm to isolate the moving objects in video sequences and then presented a rule-based tracking algorithm. The preliminary experimental results demonstrate the effectiveness of the algorithm even in some complicated situations, such as new track, ceased track, track collision, etc. A tracking method without background extraction is discussed in [3]. Because while extracting background from video frame if there are small moving things in that frame they form a blob in thresholding which create confusion in case of tracking that blob as they aren't of any use that can be reduced here. The author introduces a video tracking in computer vision, including design requirements and a review of techniques from simple window tracking to tracking complex, deformable objects by learning models of shape and dynamics.

The [5] was present algorithm for real-time detection and tracking of moving targets in terrestrial scenes using a mobile camera. The algorithm consists of two modes: detection and tracking. In the detection mode, background motion is estimated and compensated using an affine transformation. The resultant motion rectified image is used for detection of the target location using split and merge algorithm. However this algorithm used gray images to recognize the object because the color-based recognition is complicated; hence, the system can be affected by the state of the surrounding environment.

The [6] presented a new method for vehicle detection, tracking and classification based colour. The detection stage has used the background subtraction technique, where the background image is subtracted from the current frame and the Pixel values greater than the set threshold are considered as part of the foreground. In the tracking stage Kalma filter was employed to track the detected vehicle's position from frame to frame. However this method can detect and track only one object (car) that near from the camera and ignore the rest objects.

Another detection technique applied in video surveillance and monitoring systems is an improved motion detection algorithm [7] based on an integrated algorithm consisting of the temporal frame differencing, optical flow, double background filtering, and morphological processing methods. Despite the good detection results that this algorithm gets, it is described as a very complicated and high time processing algorithm.

III. PROPOSED DESIGNED AND METHODOLOGY

Detection of object and tracking of that object is an important task in the area of computer vision application. In object detection we locate or detect interested object in consecutive frames of a video file. Tracking is a process to locate moving interested object or multiple objects in a video file or camera with respect to time. Technically, in object tracking we estimate or define the trajectory or path of an interested object in the frame plane as it moving around the image plane. Because of technology increasing in computational power, availability of good quality and low cost video camera and the need of automated video system people are sowing the more interest in object tracking algorithm. In a video analysis, there are three basic or main steps are there: Detection of interested object from moving objects, Tracking of that interested objects in consecutive frames, and Analysis of trajectory of object to understand the behavior of interested object.

Generally, there are five phases involved in this framework: Phase one: Object segmentation, Phase two: Object recognition, Phase three: Object representation, Phase four: Object tracking and Phase 5: Object Visualization. These phases are described in the Figure bellow

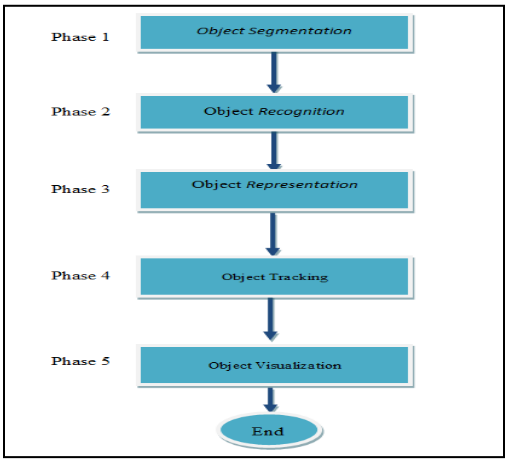

Fig. 2. *shown the phases of the framework*

IV. FRAMEWROK DESIGN

As we have seen above framework that contents 5 main phases. The following subsections will describe each phase briefly.

A. *Phase 1: Object Segmentation*

Segmentation is the process of separating the foreground objects from the background of the video sequence by using simple background subtraction based on Gaussian Mixture Models [1].

Background subtraction is a common technique used for motion segmentation in static scenes. To detect moving regions this technology subtracts the current image pixel-by-pixel from a reference background image that is created by averaging images over time in an initialization period [4]. The pixels where the difference is above a threshold are classified as foreground. After creating a foreground pixel map, some processing operations are performed to reduce the effects of noise and enhance the detected regions. With new images over time the reference background is updated to adapt to dynamic scene changes [4].

The simple version of this scheme where a pixel at location (x, y) in the current image It is marked as foreground if

$$|I_t(x,y) - B_t(x,y)| > T \quad (1)$$

Is satisfied where T is a predefined threshold. The background image Bt is updated by the use of an Infinite Impulse Response (IIR) filter as follows [1]:

$$B_{t+1} = \alpha I_t + (1 - \alpha) B_t \quad (2)$$

The foreground pixel map creation is followed by morphological closing and the elimination of small-sized regions.

In Gaussian Mixture Model, every pixel in a frame is modeled into Gaussian distribution. First, every pixel is divided by its intensity in RGB color space. Every pixel is computed for its probability whether it is included in the FG or BG with:

$$P\,(X_t) = \sum_{i=1}^{k} \omega_{i,t} . \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (3)$$

- X_t : current pixel in frame t
- K: the number of distributions in the mixture
- $\omega_{i,t}$: the weight of the kth distribution in frame t
- $\mu_{i,t}$: the mean of the kth distribution in frame t
- $\Sigma_{i,t}$: the standard deviation of the kth distribution in frame t
- Where $\eta(X_t, \mu_{i,t}, \Sigma_{i,t})$ is probability density function (pdf):

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \, exp^{-\frac{1}{2}(X_t - \mu)\Sigma^{-1}(X_t - \mu)} \quad (4)$$

B. Phase 2: Object Recognition Or Detection

Recognition or Detection is the process of identifying which of the segmented foreground objects are of interest to the tracking based on simple blob detection [5]. Blob detection refers to mathematical methods that are aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to areas surrounding those regions [6][8]. Blob detection was used to obtain regions of interest as the "recognized" object to be tracked.

C. Phase 3: Object Representation

This phase takes the results of phase 2 object recognition; after that computes a representation for each recognized object to be tracked. Object representation extracts the Bounding Box and size of each detected object and store it in list (array) where each element of array is data structure used to store properties of objects that are tracked in a video.

This phase is also responsible for maintaining the list of objects being tracked and the establishment of the correspondence between the recognized objects, and those that currently tracked. If the list size is greater than zero so the recording is start and will be continue until list size is equals to zero.

D. Phase 4: Object Tracking

This phase used to trace detected object by using Kalman filter method depend on the phase 3 results to track the position and extent of the object being tracked [5]. The Kalman filter works by estimating an unobservable state which is updated in time with a linear state update and additive Gaussian noise. A measurement is provided at each time step, which is assumed to be a linear function of the state plus Gaussian noise.

Kalman filter is a Point Tracking method used to detected a single object in a sequence of frames based on this rule: if f and h are linear functions and the initial state X and noise have a Gaussian distribution then the optimal state estimate is given by the Kalman Filter: prediction and correction

Prediction: this step uses the state model to predict the new state of the variables:

$$\bar{X}^t = DX^{t-1} + W$$

$$\bar{\Sigma}^t = DX^{t-1}D^T + Q^t \quad (5)$$

Where D is the state transition matrix which defines the relation between the state variables at time t and t-1, And Q is the covariance of the noise W.

Correction: this step used the current observations Zt to update the object's state:

$$K^t = \bar{\Sigma}^t M^T [\, M\bar{\Sigma}^t M^T + R^t \,]^{-1} \quad (6)$$

$$X^t = \bar{X}^t + K^t \underbrace{[\, Z^t - M\bar{X}^t \,]}_{v}$$

$$\Sigma^t = \bar{\Sigma}^t - K^t M \bar{\Sigma}^t \quad (7)$$

Where V is called the innovation, M is the measurement matrix, K is the Kalman gain.

E. Phase 5: Object Visualization

The final phase in the processing of the framework is the visualization. This phase useful to provide a minimal visualization to see and follow the performance of tracking, This phase was implemented by simply displays the current frame indicating the current detection of tracked objects by using rectangle with yellow color. In addition this phase also display label describe the number of objects that are currently being tracked.

V. Implementation And Results

Depending on the framework we used Matlab 2013a with three videos recorded from surveillance cameras; the yellow rectangle shows the position of tracked objects in the video and there also a counter at the top left of the video shown the number of tracked objects in current frame.

The First video is recorded under laboratory conditions. The video contains a single moving target: a red car. The videos are taken from a number of viewpoints and are specifically designed so that background subtraction can work well on them. The results in Figure 3 Shows the results movement of the car can be detected perfectly.

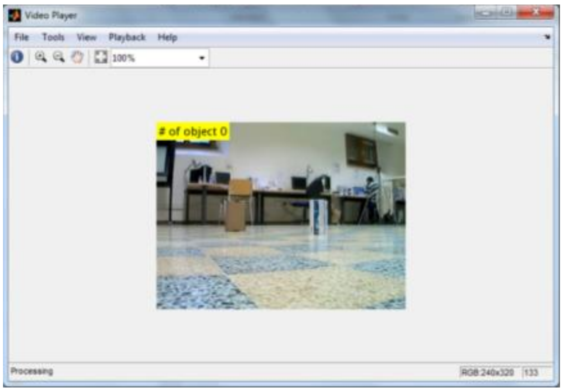

Fig. 3. a. shows the background image

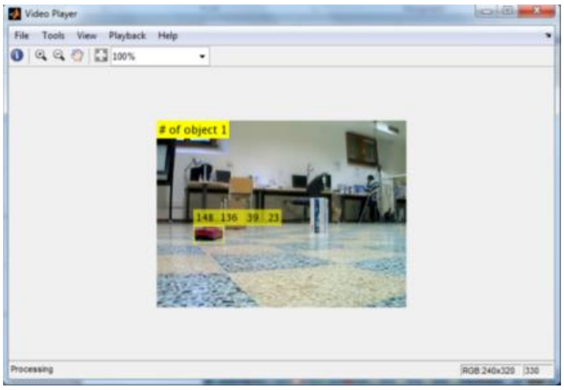

Fig. 3. b. shows the detect object

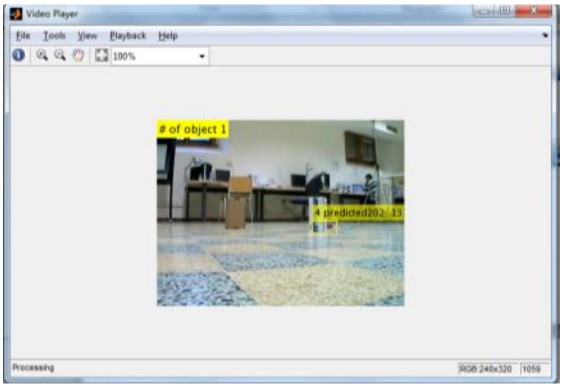

Fig. 3. c. shows the move object can be detected even a part of objects hidden behind another object

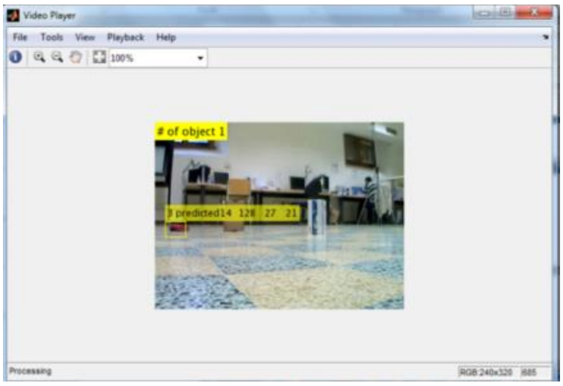

Fig. 3. d. shows the detect object

Figure 3 Shows the movement of car can be detect perfectly. Figure 3.a shows the background image and Figure 3 b,c,d shows the detect object by using yellow rectangle and in top left corner number of object (1) because this video contain only one object.

The second video is recorded from a traffic surveillance camera in the public street. The video contains a multiple target movement (cars). The videos are taken from a street without specifically designed for the environment. The results in Figure 4 Shows the framework can be used detected single or multiple objects in same times.

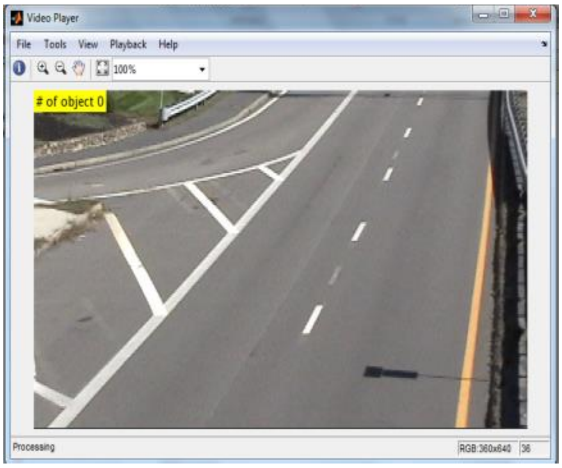

Fig. 4. a. shows the background image

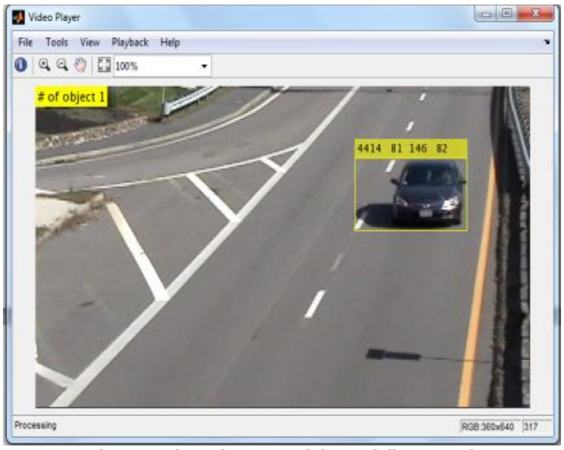

Fig. 4. b. shows the detect object (first car)

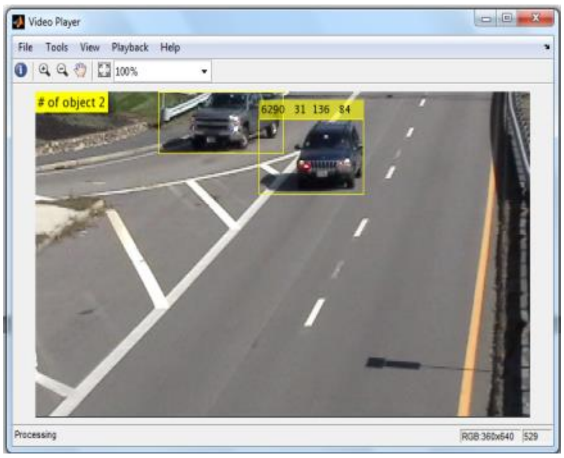

Fig. 4. c. shows the two detected objects (two cars)

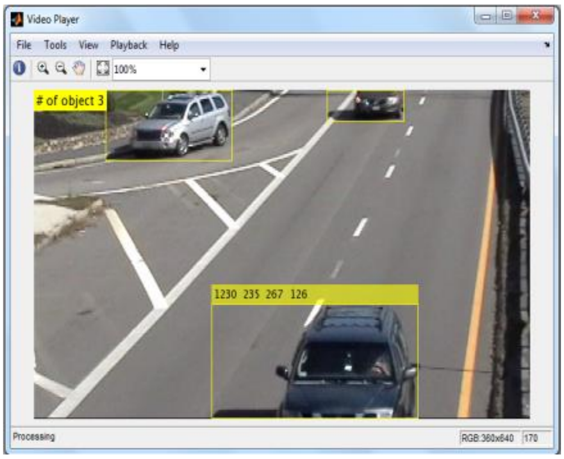

Fig. 4. d. shows the three detected cars

Figure 4 Shows the movement can be detect from surveillance camera in hotel, Figure 4.a shows the background image and Figure 4 b,c,d shows the detect objects by using yellow rectangle and in top left corner number of object depend on number of tracked objects in frame.

The third video is recorded from a surveillance camera in the street. They contain a multiple target movement two persons and two cars. The videos are taken from a street where two persons walking on the sidewalk; without specifically designed for the environment. The results in Figure 5 Shows the movement can be detected one person at the beginning and then detect the movement of a car.

Fig. 5. a. shows the background image

Fig. 5. b. shows the detect object (first person)

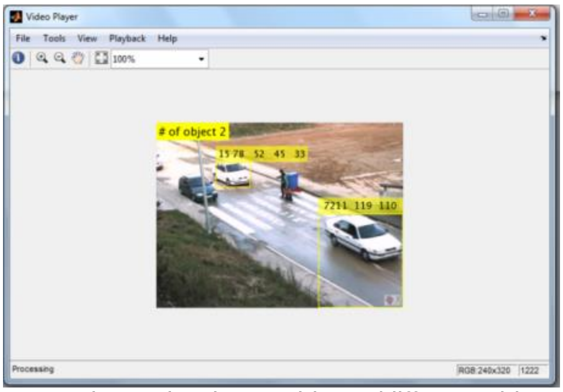

Fig. 5. c. shows the detect objects (different objects)

Fig. 5. d. shows the detect objects

Figure 5 Shows the movement can be detect from surveillance camera in street, Figure 5.a shows the background image and Figure 5 b,c,d shows the detect objects by using yellow rectangle and in top left corner number of object depend on number of tracked objects in frame.

VI. CONCLUSION

In this paper review of object detection and tracking is an important task in computer vision field. In object detection and tracking it consist of two major processes, object detection and object tracking based review analysis. Object detection in video image obtained from single camera with static background that means fixing camera is achieved by background subtraction approach. In this work we used proposed frame work to detect and recognize multiple objects in surveillance video. The results have shown the proposed framework can work well on single and multiple object recognition at videos of surveillance systems and give good results.

REFERENCES

[1] Y. Dedeoglu, “Moving object detection, tracking and classification for smart video surveillance,” Master's Thesis, Dept. of Computer Eng. Bilkent University, Ankara, 2004.

[2] J. Ba, V. Mnih, and K. Kavukcuoglu, “Multiple object recognition with visual attention,” arXiv:1412.7755, 2014.

[3] A. M. Cheriyadat, B. L. Bhaduri, R. J. Radke, “Detecting multiple moving objects in crowded environments with coherent motion regions,” in Computer Vision and Pattern Recognition Workshops 2008, vol. 0, IEEE Computer Society, 2008, pp. 1- 8.

[4] M. Piccardi, “Background subtraction techniques: a review,” in Proc. IEEE Int. Conf. Systems, Man, Cybernetics, 2004, pp. 3099–3104.

[5] A. Behrad, A. Shahrokni, S.A. Motamedi. A robust visionbased moving target detection and tracking system. In the Proceeding of Image and Vision Computing Conference, University of Otago, Dunedin, New Zealand, November 2001.

[6] Habibu Rabiu," Vehicle Detection Tracking and Colour-based classification in Video ", in IJAI, Volume 2, Issue1 , January 2013.

[7] Nan Lu, Jihong Wang, Q.H. Wu and Li Yang," An Improved Motion Detection Method for Real- Time Surveillance,” Proc.of International Journal of Computer Science, IAENG, vol. 35, January 2008, pp. 1-10.

[8] Y. C. Wang, C. C. Han, C. T. Hsieh, and K. C. Fan, "Vehicle type classification from surveillance videos on urban roads," in Ubi-Media Computing and Workshops (UMEDIA), 2014 7th International Conference on, 2014, pp. 266-270.