



## Proposed Multi-object Tracking Algorithm Using Sobel Edge Detection operator

A. M. Sallam<sup>1</sup>, M. Sakr<sup>2</sup>, Mohamed Abdallah<sup>3</sup>  
Egyptian Armed Forces Egypt

Received 20Dec, 2016; Accepted 17 Jan, 2017 © The author 2017. Published with open access at [www.questjournals.org](http://www.questjournals.org)

**ABSTRACT:** Tracking of moving objects that is called video tracking is used for measuring motion parameters and obtaining a visual record of the moving objects, it is an important area of application in image processing. In general there are two different approaches to obtain object tracking: the first is Recognition-based Tracking, and the second is the Motion-based Tracking. Video tracking system raises a wide possibility in today's society. This system is used in various applications such as military, security, monitoring, robotic, and nowadays in day-to-day applications. However the video tracking systems still have many open problems and various research activities in a video tracking system are explores. This paper presents an algorithm for video tracking of any moving targets with the uses of contour based detection technique that depends on the sobel operator. The proposed system is suitable for indoor and outdoor applications. Our approach has the advantage of extending the applicability of tracking system and also, as presented here improves the performance of the tracker making feasible high frame rate video tracking. The goal of the tracking system is to analyze the video frames and estimate the position of a part of the input video frame (usually a moving object), our approach can detect, tracked any object more than one object and calculate the position of the moving objects. Therefore, the aim of this paper is to construct a motion tracking system for moving objects. Where, at the end of this paper, the detail outcome and result are discussed using experiments results of the proposed technique.

**Keywords:** Image processing, Tracking system, image tracking, Contour based tracking, Video Tracking, Edge detection, Sobel operator.

### I. INTRODUCTION

The problem of object tracking can be considered an interesting branch in the scientific community and it is still an open and active field of research [1], [2]. This is a very useful skill that can be used in many fields including visual serving, surveillance, gesture based human machine interfaces, video editing, compression, augmented reality, visual effects, motion capture, medical and meteorological imaging, etc... [3], [4]. In the most approaches, an initial representation of the tope-tracked object or its background is given to the tracker that can measure and predict the motion of the moving object representation overtime. The most of the existing algorithms depends upon the thresholding technique or feature that extracted from the object to be tracked or combined it with the thresholding to try to separate the object from the background [5], [6], [7]. In this paper our proposed algorithm try to solve the tracking problem based on edge (contour) of the object to track it (i.e. we extracting the contour of the target and detect it among the whole sequence of frames using a mean of edge detection technique.

Object tracking is a very specific field of study within the general scope of image processing and analysis. Human can recognize and track any object perfectly, instantaneously, and effortlessly even the presence of high clutter, occlusion, and non-linear variations in background, target shape, orientation and size. However, it can be an overwhelming task for a machine! There are partial solutions, but the work is still progressing toward a complete solution for this complex problem [8].

The detection and classification of moving objects is an important area of research in computer vision. This problem assumes importance because of the fact that our visual world is dynamic and we constantly come across video scene that contains a finite large number of moving objects. To segment, detect, and track these objects from a video sequence of images is possibly the most important challenge that the vision experts con front today. These systems find applications in human surveillance, security systems, traffic monitoring, industrial vision, defense surveillance, home landing security, etc.

The remains of this paper, we will explain our literature review in Section 2. Then, in Section 3, we describe the Desirable system features and algorithms necessary for successful system. In section 4 we describe the system architecture (implementation environment of the system), and the proposed algorithm that will be used in our method. In Section 5, the experimental results of the proposed Multi-object based on contour tracking proposed. Finally, in Section 6 we will discuss and Analysis of the obtained results from section 5.

## II. TRACKING SYSTEM A LITERATURE REVIEW

In the recent times the vast number of algorithms has been proposed in the field of object tracking. An even greater number of solutions have been constructed from these algorithms, many solving parts of the puzzle that makes computer vision so complex. One technique proposed to use the small chromatic-space of human skin along with facial such as eyes, mouth and shape to locate faces in complex color images. Yang and Ahuja [10] investigated such an object localization techniques, where experimental results conclude

“Human faces in color image can be detected regardless of size, orientation or viewpoint.” In the above paper it was illustrated that the major difference in skin color across different appearances was due to intensity rather than color itself. McKinnon [11] also used in similar skin filtration based theory to implement a multiple object tracking system. McKinnon stated that his solution was often limited by the quality of the skin sample supplied initially. Further to this, in real-time environment the lack of or excessive level of light could cause the performance to suffer. The drawback of skin color systems is that they can only track objects containing areas of skin-color-like areas in the background may be confused with real regions of interest. As such they are not suitable for use in all applications and hence are often limited in their use [12]. The most two popular methods for image segmentation used in the object tracking field are temporal segmentation and background subtraction. Vass, Palaniappan and Zhuang presented a paper [13] that outlined a method of image segmentation based on a combination of temporal and spatial segmentation. By using interframe differences [14] a binary image was obtained showing pixels that had undergone change between frames. Temporal segmentation on its own fails for moving homogeneous regions, as such spatial segmentation was incorporated. Using a split and merge techniques an image are split into homogenous regions. Finally, by merging spatial and temporal information, segmentation of motion areas was achieved at a rate of approximately five frames per second; however a small amount of background was evident in the resulting segmented regions. Andrews [15] utilized background subtraction to create a system based on distance measures between object shapes for real-time object tracking. By acquiring an initial image of the operational environment free of moving objects, he was able to cleanly segment areas of change in future (object filled) frames. From this segmentation a model was created based on edge sets. One of the most drawbacks of image difference technique in the detection of moving objects is that it can only capture moving regions with large image frame difference. However, a region can have a small ImDiff even if it is the projection of a moving object due to the aperture problem [16].

$$\text{ImDiff} = \text{ImN} - \text{ImN-1}, (1)$$

**Where:** ImN is the current frame, ImN-1 is the previous frame, and ImDiff is the difference frame between the current frame and the previous Frame.

K. Chang and S. Lai [17] proposed an object contour tracking algorithm based on particle filter framework. It is only need an initial contour at the first frame and then the object models and the prediction matrix are constructed online from the previous contour tracking results automatically. This proposed algorithm builds two online models for the tracked object, the first gets the shape model and the other gets the grayscale histogram model. The grayscale histogram simply records the grayscale information inside the object contour region. Each of these two models is represented by a mean vector and several principle components, which are adaptively computed with the incremental singular value decomposition technique. E. Trucco, K. Plakas [18] introduce a concise introduction to video tracking in computer vision, including design requirements and a review of techniques from simple window tracking to tracking complex, deformable object by learning models of shape and dynamics. Sallam et al [9] proposed feature extraction based video object tracking depend on computing the features (mean, variance, length ...) of the object in 8 directions and compare it within a window around the object, but this system has a little drawback that the measured position has an error between  $\pm 12$  pixels from the exact trajectory of the object.

The systems that use the Motion-based tracking approach are the systems that track the fast and specific targets (e.g., a military aircraft, guided missile, etc.). This systems is developed based on the Motion-based predictive techniques such as (Kalman filter, Extended Kalman filter, Particle filter, etc.). In automatic image processing based object tracking systems, the target object entering the field of view (FOV) are acquired automatically without human intervention. In Recognition-based tracking, the object pattern is recognized in successive image frames and tracking is carried-out using its positional information, this approach uses in recorded video like (e.g., Traffic calculations, terrorism searching in videos, etc.) [19].

The detection and classification of moving objects is an important area of research in computer vision. This problem assumes importance because of the fact that our visual world is dynamic and we constantly come across video scene that contains a finite large number of moving objects. To segment, detect, and track these objects from a video sequence of images is possibly the most important challenge that the vision experts confront today.

In this paper we try to solve some of these problems by presenting it and the possible solution of its sub task in dynamic image analysis. These systems find applications in human surveillance, security systems, traffic monitoring, industrial vision, defense surveillance, home landing security, etc.

### III. EDGE

In the early stages of vision processing it is usual to identify features in images that are relevant to estimating the structure and properties of objects in a scene. Edges are one of such feature. Edges are characterized by significant local changes in the image and are important features for analyzing image content. Edges occur on the boundary between two different regions in an image, and edge detection is frequently the first step in recovering information from images. Due to its fundamental importance, the task of edge detection continues to be a very active research area [20].

#### A. How Can We Characterize the Edge?

A problem of fundamental importance in image analysis is edge detection. Edges characterize object boundaries and are therefore useful for segmentation, and identification of objects in scenes. An edge in an image characterized as a significant local change in the image intensity or the first derivative of the image intensity. Discontinuities in the image intensity can be either [20]:

- [1] Step discontinuities, where image intensity abruptly changes from one value on one side of the discontinuity to a different value on the opposite side,
- [2] Or line discontinuities, where the image intensity abruptly changes but then returns to the starting value within some short distance. However, step and line edges are rare in real images.

Because of low-frequency components or the smoothing introduced by most sensing devices, sharp discontinuities rarely exist in real images. Step edges become ramp edges and line edges become roof edges, where intensity changes are not instantaneous but occur over a finite distance. Illustrations of these edge profiles are shown in Figure 1.

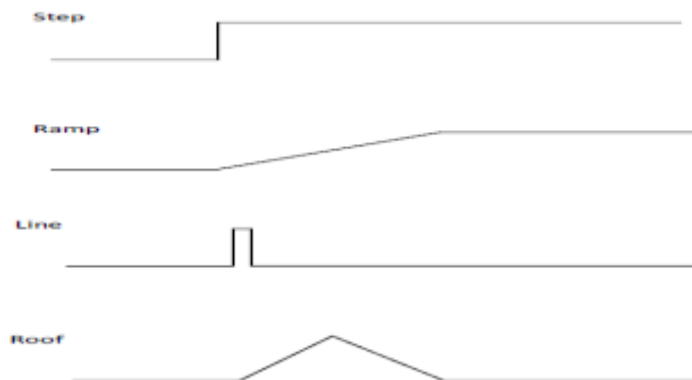


Figure 1. One- dimensional edge profiles

#### B. Edge detection general steps

Algorithms for edge detection generally contain three steps [21]:

- [1] **Filtering:** Since gradient computations based on intensity values of only two points are susceptible to noise and other vagaries in discrete computations, filtering is commonly used to improve the performance of an edge detector with respect to noise. However, there is a trade-off between edge strength and noise reduction. More filtering to reduce noise results in a loss of edge strength.
- [2] **Enhancement:** In order to facilitate the detection of edges, it is essential to determine changes in intensity in the neighborhood of a point. Enhancement emphasizes pixels where there is a significant change in local intensity values and is usually performed by computing the gradient magnitude.
- [3] **Detection:** We only want points with strong edge content. However, many points in an image have a nonzero value for the gradient, and not all of these points are edges for a particular application. Therefore, some method should be used to determine which points are edge points. Frequently, thresholding provides the criterion used for detection.

Many edge detectors have been developed in the last two decades. The most frequently used edge detection methods are used for image segmentation are Sobel, Prewitt, Robert, and Canny Edge detection operator [22, 23, 24, 25].

#### **IV. DESIRABLE SYSTEM FEATURES AND ALGORITHMS NECESSARY FOR SUCCESSFUL SYSTEM**

##### **A. Desirable System Features**

**The system should be designed with the following general performance measures in minds [20]:**

- 1- Ability for operate with complex scenes.
- 2- Ability to track more than one target within the same FOV and at the same time.
- 3- Adaptability to time-varying target and (slowly varying) background parameters.
- 4- Minimum probability of false alarm (false alarm means that the system detect a target (targets) from any noise in the scene and tracks it this phenomena appears in the tracking system that uses the automatic sensing of target and it can't be used in military, security, monitoring, etc. systems).
- 5- Minimum probability of loss of target (LOT), according to criterion:

$$\min\{ E[(B - b)^2] \} \quad (2)$$

**Where:** B is the actual target location,  
b is the estimated target location get from the tracking system.

##### **B. Algorithms Necessary for Successful System**

**The minimum Algorithms necessary for a successful system may be Sub-divided into four parts [26]:**

- 1- A target/background (T/B) separation or segmentation algorithm, which can segments the frame by classifying pixels (or groups of pixels) as members of either the target or background sets.
- 2- A tracking filter, to minimize the effects of noisy data which produce an inexact T/B separation that will effect on the estimated target location.
- 3- The used algorithm, which processes information from the just-segmented frame as well as memory information to generate raw estimates of the target centroid (target center).
- 4- An overall system control algorithm, to make the major system automatic decisions, and supervise algorithm interaction.

#### **V. TESTING VIDEO DESCRIPTION AND THE PROPOSED ALGORITHM**

##### **A. Testing Video Description**

Here we use real video sequences for testing the proposed video tracking algorithm for more than one target based on contour edge detection technique.

The testing video sequences used are two recorded video sequence that used to test the proposed tracking algorithm for multi-target (more than one target) to illustrate the ability of the algorithm to run with more than one target.

##### **B. The Proposed Video Tracking Algorithm**

Edge detection is one of the most commonly used operations in image analysis. The reason for this is that edges form the outline of an object. Objects are subjects of interest in image analysis and vision systems. An edge is the boundary between an object and the background. This means that if the edges in an image can be identified accurately, any object can be located. Since computer vision involves the identification and classification of objects in an image, edge detection is an essential tool [19].

The proposed Video Tracking Algorithm that we applied depends on extracting the contour of the target. The algorithm description can subdivided into the following steps:

- 1- 4- If the algorithm doesn't sense any target (targets), the algorithm goes into that loop until sensing any moving targets.
- 5- After sensing target, the second step the algorithm starts to separate the target from the background and tracking it by the following steps
- 6- For the tracked target we compute the center and the vertices that contains the target in between, we create a search window that contains the target and bigger than the target with twenty pixels in each of the four directions (top, bottom, right, and left).
- 7- We compute the total gradient of the current frame by the "Sobel" operator for each target within the search window of each target, applying the thresholding and the average filter within the search window of each

target only (to reduce the computation time and the complexity of the process to make the algorithm fast as possible).

- 8- After computing the ‘‘Sobel edge search window’’ for the target, a search module used to search in that window to get the current position of the target and compute the current vertices of the target that containing it and compute the center of it to get the whole trajectory of the target in the whole sequence.
- 9- The algorithm getting the target data, if the target never lost, the algorithm still getting the data of that target, but if the target lost during the tracking module more than 5 frames, the algorithm return to the search mode module again.
- 10- If the number of frames that lost the target exceeds more than 5 frames the algorithm use a predictor to try to predict the location of the target and return to the algorithm again as in figure 2.

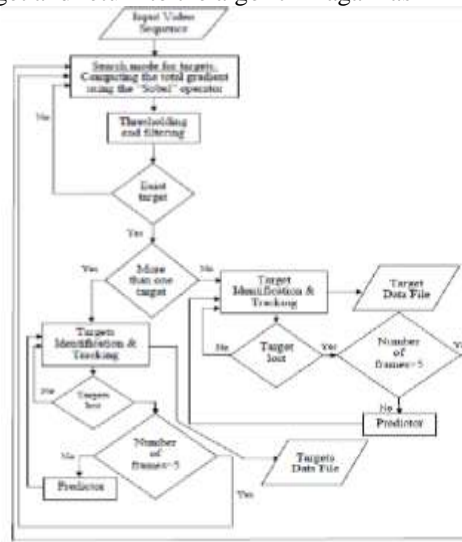


Figure 2. Proposed Contour based Target Tracking Algorithm

## VI. EXPERIMENTAL RESULTS AND COMPARISON FOR THE PROPOSED ALGORITHM, THE FEATURE EXTRACTION ALGORITHM AND THE TEMPORAL FILTRATION ALGORITHM

We used many real video sequences for testing the proposed video tracking algorithm we discuss two video sequences. According to Sallam [26] we use a recorded video sequences to compare the measured target (targets) position produced by the proposed algorithm with an exact target (targets) position to plot an error curves and compute the MSE (Mean Square error) for the two video sequences.

We compute the error in the x-position & y-position:

$$\xi_x (\text{Error in x-position}) = \text{Desired Target Position in (x)} - \text{Measured Target Position in (x)}. \quad (3)$$

$$\xi_y (\text{Error in y-position}) = \text{Desired Target Position in (y)} - \text{Measured Target Position in (y)}. \quad (4)$$

We can compute the Average Mean Square Error for the whole sequence by equation 5.

$$AMSE = \frac{MSE}{N} \quad (5)$$

**Where:** AMSE is the Average Mean Square Error,  
MSE is the Mean Square Error,  
N is the number of the frames in the sequence.

$$MSE = \sum_{n=1}^N \sqrt{(D_n(x_c, y_c) - M_n(x_c, y_c))^2} \quad (6)$$

**Where:**  $D_n(x_c, y_c)$  the desired trajectory at the center of the target for the frame n,



$M_n(x_c, y_c)$  the measured trajectory at the center of the target for the frame  $n$ .  
 $N$  the number of the frames at the whole video sequence.

The performance of different tracking algorithms can be evaluated by defining a normalized quantitative index  $q$ . The larger the value of  $q$  is, the better the performance the algorithm has. Thus  $q$  is defined as in equation (7), the bigger the value of  $q$  the better of the performance of the tracking algorithm. Not that AMSE is inversely proportional to  $q$ .

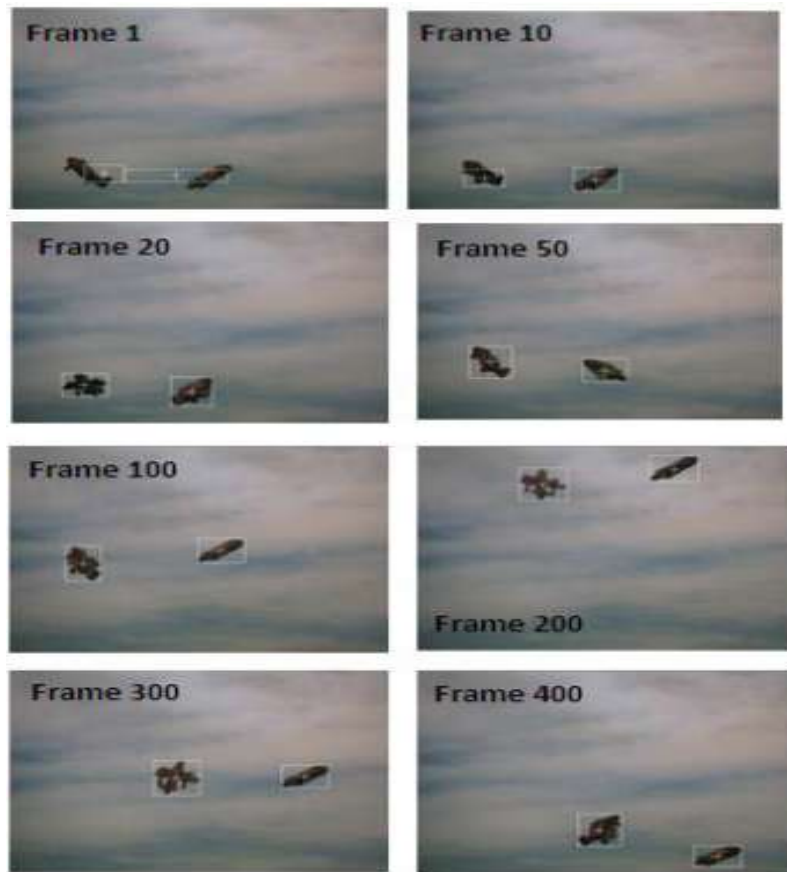
$$q = \frac{n_{success}}{N_f} \quad (7)$$

**Where:**  $n_{success}$  is the number of successfully tracked frames.

Figure 3 illustrates a sample of the detection results from the “2\_airplane1” video sequence using the proposed algorithm and the number of the frames of that sequence is 425 frames using the proposed algorithm. the proposed algorithm in the x-position, y-position, and the whole trajectory for the first target in the “2\_airplane1” video sequence.

Figure 5 illustrates the desired and measured trajectories by the proposed algorithm in the x-position, y-position, and the whole trajectory for the second target of the same video sequence.

Figure 6 illustrates a sample of the detection results from the “2\_airplane2” video sequence with more objects and more complicated background using the proposed algorithm with number of frames equal 480 frames.



**Figure 3.** The detection results of the “2airplane1” video sequence

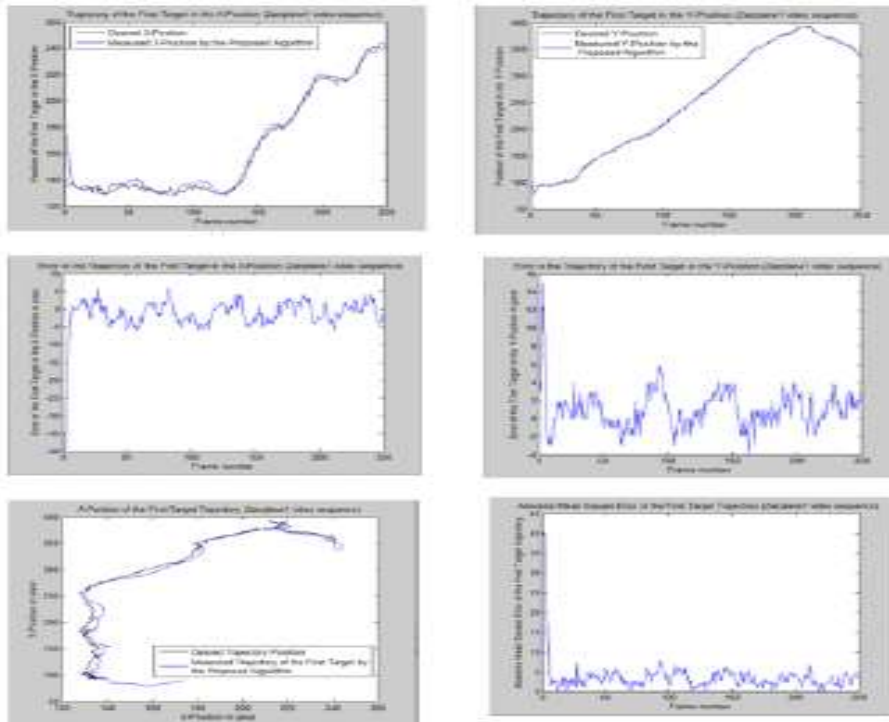


Figure 4. The error in the X&Y-Position Trajectories of the first target get

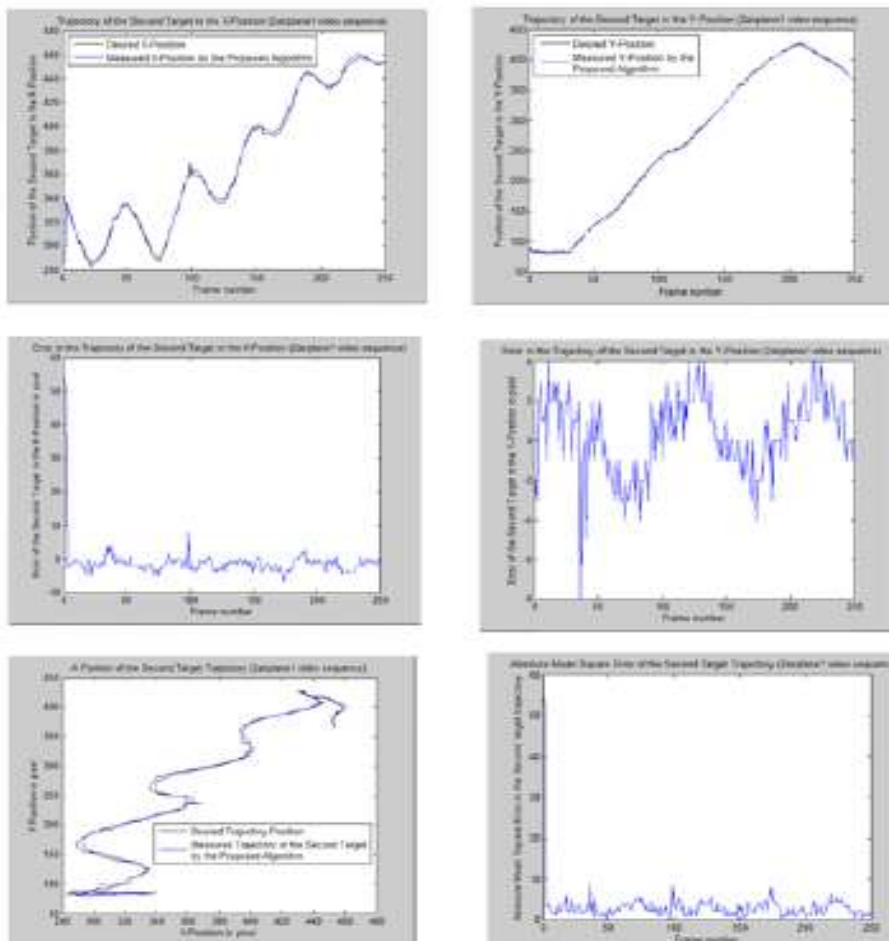
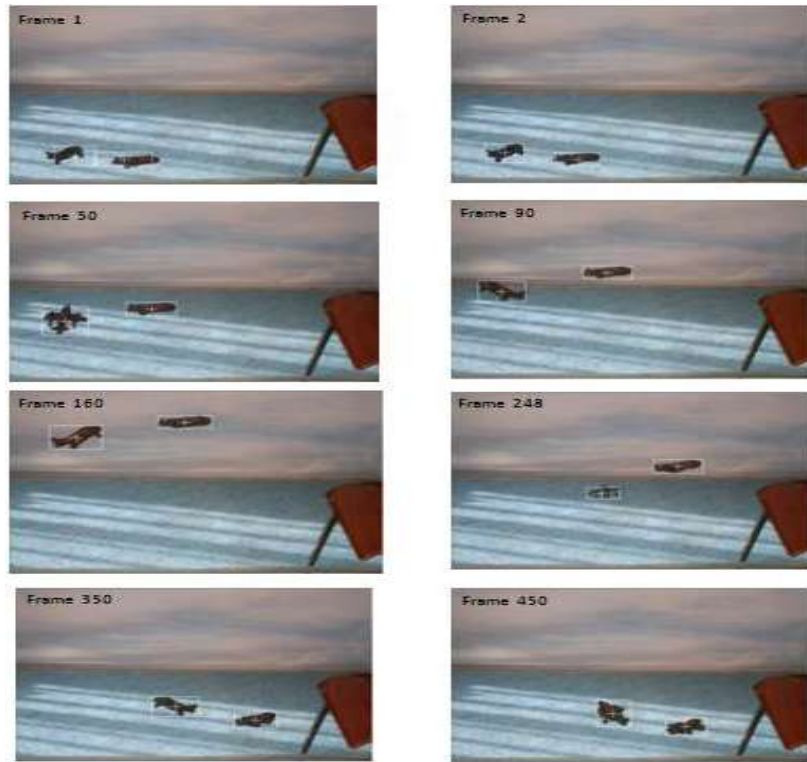


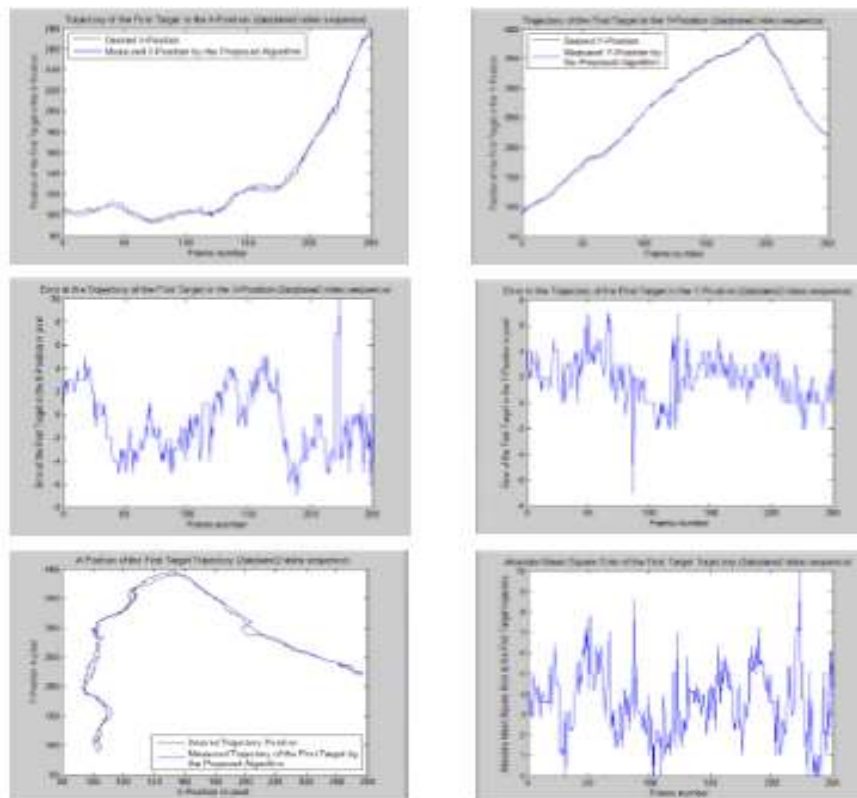
Figure 5. The error in the X&Y-Position Trajectories of the second target in the “2airplane1” video sequence



**Figure 6.** The detection results of the “2airplane2” video sequence

**Figure 7** illustrates the desired and measured trajectories by the proposed algorithm in the x-position, y-position, and the whole trajectory for the first target in the “2\_airplane2” video sequence.

**Figure 8** illustrates the desired and measured trajectories by the proposed algorithm in the x-position, y-position, and the whole trajectory for the second target of the same video sequence same .



**Figure 7.** The error in the X&Y-Position Trajectories of the first target in the “2airplane2” video sequence



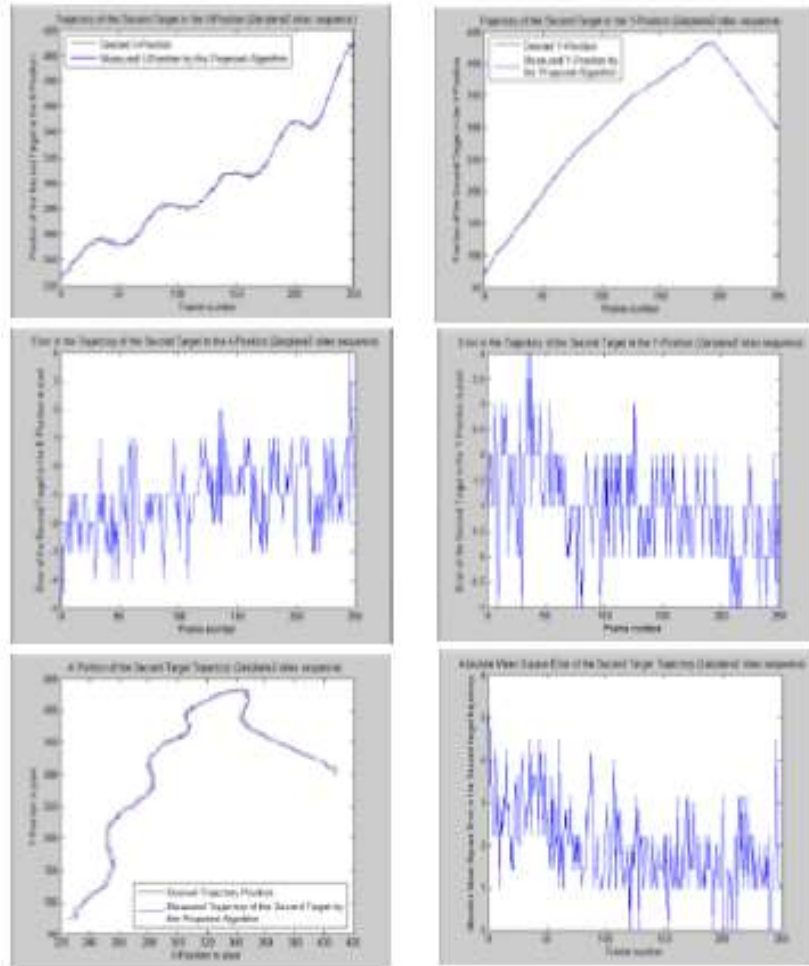


Figure 8. The error in the X&Y-Position Trajectories of the second target in the “2airplane2” video sequence

### VII. ANALYSIS OF THE OBTAINED RESULTS

From the experimental results and figure 3 through figure 8 and from table1 we found that:

- 1- Temporal filtration algorithm is difficult to handle shadow and occlusion.
- 2- The ASME and q that measured from the proposed contour-based algorithm for the two video sequence “2\_airplane1” and “2\_airplane2” are high for the 2 targets because the proposed contour based algorithm depends on the contour of the moving targets that produced by the sobel edge operator and it is more accurate in detection because it depends on the variation between the moving target and background.
- 3- The ASME and q for the “2\_airplane1” video sequence is better than the “2\_airplane2” video sequence because the first sequence has simple background than the second one and it means that the separation of the targets from the background is more accurate in the first sequence.
- 4- Contour-based video tracking algorithm can detect and track object within frame with better.
- 5- Contour-based video tracking algorithm can deal with the target at sudden changes in background or brightness.
- 6- Contour-based video tracking algorithm can deal with the target that moves in simple or difficult background because the algorithm depends on the edge detection technique.
- 7- The proposed algorithm for target tracking is able to track more than one target.
- 8- The obtained results in this paper showed the robustness of using Contour-based video tracking algorithm and showed that it produces a measured trajectory with high accuracy.
- 9- Table 1 showed that our proposed algorithm have minimum MSE that indicates high detection rate.

TABLE I. TEST RESULTS OF PROPOSED ALGORITHM FOR MORE THAN ONE TARGET BASED ON AMSE AND Q

Video Sequence		"?_airplane1"	"?_airplane2"
Target			
The First Target	AMSE*	0.1202	0.1203
	q*	0.967	0.96
The Second Target	AMSE	0.0881	0.11
	q	0.992	0.97

\*Low AMSE Lead to high detection, \*\* The larger the value of q is, the better the performance the algorithm has.

### REFERENCES

- [1]. V. Manohar, P. Soundararajan, H. Raju, D. Goldgof, R. Kasturi, J. Garofolo, "Performance Evaluation of Object Detection and Tracking in Video," LNCS, vol. 3852, 2(2006), pp.151-161.
- [2]. Y. T. Hsiao, C. L. Chuang, Y. L. Lu, J. A. Jiang, "Robust multiple objects tracking using image segmentation and trajectory estimation scheme in video frames", Image and Vision Computing, vol. 24, 10(2006), pp. 1123-1136.
- [3]. T. Ellis, "Performance metrics and methods for tracking in surveillance", 3rd IEEE Workshop on PETS, Copenhagen, Denmark (2002), pp. 26-31.
- [4]. P. Pérez, C. Hue, J. Vermaak, M. Gangnet, "Color-based probabilistic tracking", Conference on Computer Vision, LNCS, vol. 2350 (2002), pp. 661-675.
- [5]. T. Schoenemann, D. Cremers, "Near Real-Time Motion Segmentation Using Graph Cuts", Springer, DAGM 2006, LNCS 4174, 2006, pp. 455-464.
- [6]. Hu, W., Tan T., Wang L., Maybank S., "A Survey on Visual Surveillance of Object Motion and Behaviours" IEEE Transactions on Systems, Man, and Cybernetics, Vol. 34, no. 3, August 2004. R. Polana and R. Nelson, "Low level recognition of human motion." Proceedings IEEE Workshop Motion of Non-Rigid and Articulated Objects, Austin, TX, 1994, pp. 77-82.
- [7]. N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," IEEE Trans. Pattern Anal. Machine Intell., vol. 22, pp. 266-280, Mar. 2000. Burky B. ÖRTEN, "Moving Object Identification and Event Recognition in Video Surveillance Systems", Master of Science, The Graduate School of Natural and Applied Sciences of Middle East Technical University, July 2005.
- [8]. Javed Ahmed, M. N. Jafri, J. Ahmed, M. I. Khan, "Design and Implementation of a Neural Network for Real-Time Object Tracking", World Academy of Science, Engineering and Technology, June 2005.
- [9]. Sallam et al. "Real Time Algorithm for Video Tracking", AL-Azhar Engineering Eleventh International Conference AEIC-2010 , Cairo, Egypt, December 21-23, pp. 228-235, 2010.
- [10]. M. Yang and N. Ahuja, "Detecting Human Faces in Color Images", Proceedings IEEE International Conference on Image Processing, IEEE Computer Soc. Press, Los Alamos, Calif., pp. 127-139, 1998.
- [11]. David N. McKinnon, "Multiple Object Tracking in Real-Time", Undergraduate Thesis, Univ. Queensland, St Lucia, Dept. Computer Science and Electrical Engineering, 1999.
- [12]. Daniel R. Corbett, "Multiple Object Tracking in Real-Time", Undergraduate Thesis, Univ. Queensland, St Lucia, Dept. Computer Science and Electrical Engineering, 2000.
- [13]. J. Vass, K. Palaniappan, X. Zhuang, "Automatic Spatio-Temporal Video Sequence Segmentation", Proc. IEEE International Conference on Image Processing V3, IEEE Computer Soc. Press, Los Alamos, Calif., pp.958-962, 1998.
- [14]. Sallam et al, "Object Based Video Coding Algorithm", Proceedings of the 7th International Conference on Electrical Engineering, ICEENG 2010, May 2010.
- [15]. Robert Andrews, "Multiple Object Tracking in Real-Time", Undergraduate Thesis, Univ. Queensland, St Lucia, Dept. Computer Science and Electrical Engineering, 1999.
- [16]. Berthold. K. P. Horn, "Robot Vision", Mc Graw-Hill Book Company, New York, 1986.
- [17]. K. Chang, S. Lai, "Adaptive Object Tracking with Online Statistical Model Update", Springer, ACCV 2006, LNCS 3852, 2006, pp. 363-372.
- [18]. E. Trucco, K. Plakas, "Video Tracking: A Concise Survey", IEEE Journal of Oceanic Engineering, Vol. 31, No. 2, April 2006.
- [19]. Tinku Acharya, Ajoy K. Ray, "Image Processing Principles and Applications", Wily-Interscience, 2005.
- [20]. A. M. Sallam, "Real Time Modern Target Tracking Techniques", Ph.D thesis, Military Technical College, Cairo, Egypt, 2012.
- [21]. N. Senthilkumaran, R. Rajesh, "Edge Detection Techniques for Image Segmentation – A Survey of Soft Computing Approaches", International Journal of Recent Trends in Engineering, pp. 250-254, Vol. 1, No. 2, May 2009.
- [22]. S. Al-amri, N. Kalyankar, Khamitkar S.D, "Image Segmentation by Using Edge Detection", International Journal on Computer Science and Engineering, pp. 804-807, Vol. 02, No. 03, 2010.
- [23]. E. Nadernejad, S. Sharifzadeh, "Edge Detection Techniques: Evaluations and Comparisons", Applied Mathematical Sciences, pp. 1507-1520, Vol. 2, no.31, 2008.
- [24]. R. Maini, H. Aggarwal, "Study and Comparison of Various Image Edge Detection Techniques", International Journal of Image Processing (IJIP), pp. 1-12, Volume (3) : Issue (1).
- [25]. Parker, J.R., "Algorithms for Image Processing and Computer Vision", Wiley Computer Publishing , 1997.
- [26]. Sallam et al, "Contour Based Algorithm for Object Tracking", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 9, No. 7, July 2011.