

DESIGN AND IMPLEMENTATION OF VIDEO TRACKING SYSTEM BASED ON CAMERA FIELD OF VIEW

*Hassan E. El-Taher, **Kasim M. Al-hity, and Mubarak M.Ahmed***

*ZEC Company Cooperation with ***Sudan University of Science and Technology
**Omdurman Islamic University-Faculty of Engineering Science.

ABSTRACT

The basic idea of this paper is to design and implement of video tracking system based on Camera Field of View (CFOV), Otsu's method was used to detect targets such as vehicles and people. Whereas most algorithms were spent a lot of time to execute the process, an algorithm was developed to achieve it in a little time. The histogram projection was used in both directional to detect target from search region, which is robust to various light conditions in Charge Couple Device (CCD) camera images and saves computation time.

Our algorithm based on background subtraction, and normalize cross correlation operation from a series of sequential sub images can estimate the motion vector. Camera field of view (CFOV) was determined and calibrated to find the relation between real distance and image distance. The system was tested by measuring the real position of object in the laboratory and compares it with the result of computed one. So these results are promising to develop the system in future.

KEYWORDS

Camera field of view, Image segmentation, Normalize Cross Correlation, Otsu's method.

1. INTRODUCTION

Compared to radar and laser tracking systems, the applications of image processing are modern tracking technology. The image processing technology was used on the static images in the field of the meteorology application [1,2]. In the last decade, dynamic image processing was extensively used in different fields, especially, pattern recognition is very important for automation in industrial applications. The processing of static images and dynamic images is different. The processing of a series of images is the base of video image processing. So far, many methods for dynamic image processing such as cross-correlation matching, template matching, frame difference and motion vector estimation have been done. Visual servo can be considered the fusion of computer vision, robotics and control. There are two basic approaches to visual servo control: image-based visual servo and position-based visual servo [3]. Each approach brings the obvious vision problem: Feddema and Mitchell used the image Jacobian method based on the motion of points in the image to relate image-plane velocity of a point to the relative

velocity of the point with respect to the camera. With the progress of research in vision servo, each has its own advantages in specific application areas [4]. In this paper, a video tracking system based on Normalize cross correlation and background subtraction was described, which executes a series of operations from obtaining sub images and processing them to correspond with the target position.

The tracking algorithms used in the system is based on rejecting the background from a certain region of interest to highlight the target. This is achieved by manually placing a track-window around the target to specify the region of interest. The boundaries of the tracking-window are processed and the corresponding intensity values are marked as the background region. A histogram of the intensities for the region inside the tracking-window is made and everything inside it that has an intensity level different from the range of background intensities is nominated as the target [5].

Video tracking system is usually composed of image acquisition, image processing using MatLab, PC. The image acquisition includes MINI CCD camera and interface card (Frame Grabber). The cross-correlation algorithm was developed to get the motion vectors. System was used a MINI CCD camera which is restricted by video rates, e.g. 30 Hz (NTSC) or 25 Hz (PAL). In this paper, Camera field of view (CFOV) was determined and calibrated to find the relation between real distance and image distance.

2. IMAGE SEGMENTATION ALGORITHMS

Segmentation involves separating an image into regions (or their contours) corresponding to objects. Background can be identified by identifying differences between regions. A natural way to segment such regions is through thresholding to separate light and dark regions. Thresholding creates binary images from gray level ones by turning all pixels below some threshold to zero and all pixels above that threshold to one. If $g(x, y)$ is a thresholding binary image of $f(x, y)$ at some global threshold T [6].

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq Th \\ 0 & \text{other wise} \end{cases} \quad (1)$$

Extraneous pixels that aren't part of the desired region can be included, and isolated pixels within the region can be lost (especially near the boundaries of the region). These effects get worse as the noise gets worse, simply because it's more likely that pixels intensity doesn't represent the normal intensity in the region. Sometimes some information can be lost and sometimes getting too many extraneous background pixels. Another problem with global thresholding is that changes in illumination across the scene may cause some parts to be brighter and some parts darker. It can deal, at least in part, with such uneven illumination by determining thresholds locally. That is, instead of having a single global threshold, the threshold was allowing itself to smoothly vary across the image.

To set a global threshold or to adapt a local threshold to an area, usually one have to look at the histogram to see if one can find two or more distinct modes for the foreground and for the background. Converting a grayscale image to monochrome is a common image processing task. Otsu's method, named after its inventor Nobuyuki Otsu, is one of many binarization algorithms

[6]. This section describes how the algorithm works and provides a Matlab implementation, which can be easily ported to other languages [7,8,9,10].The result of this algorithm is shown in Figure (1).

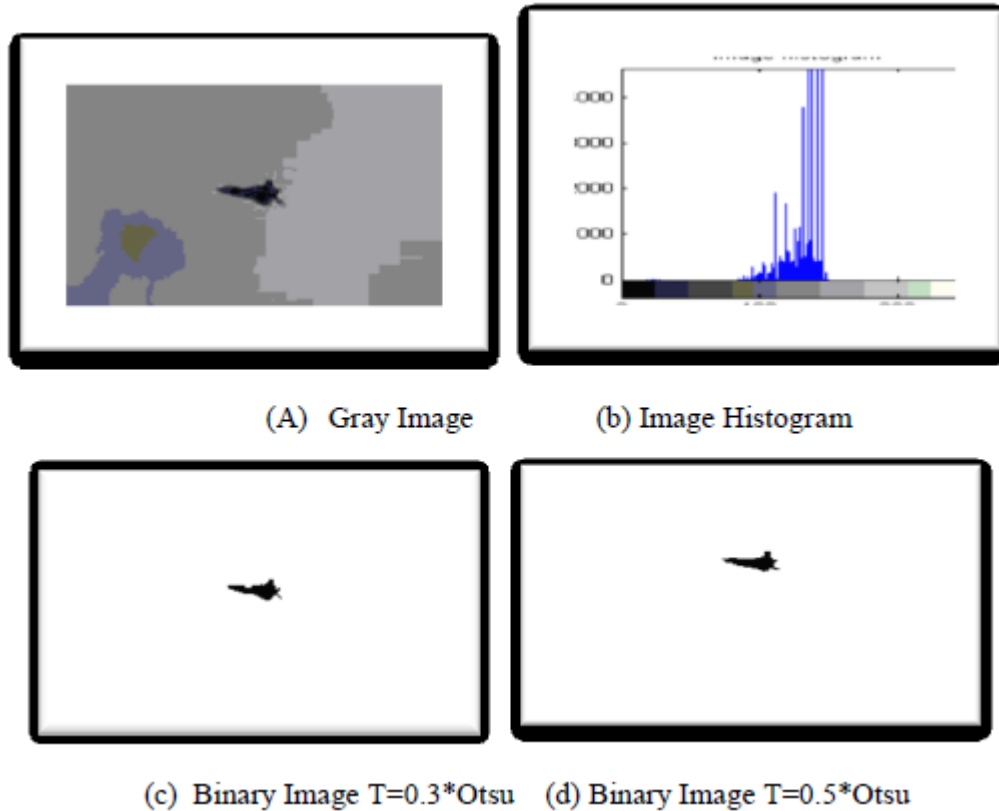


Figure (1) Difference values Otsu threshold and image histogram (Plane target)

3. TRACKING ALGORITHMS

3.1 Background Subtraction Algorithm

The basic idea underlying the background subtraction is image differentiation between the present frame image which contains the subject of interest and the reference images [11]. Thus, a set of background image frames will be functioned as a reference set of images. In this work, a better technique involves the background model as mentioned in [12] is applied. Next, each of the sequence frames from the input video that consist of object motion will be processed to extract the object boundaries. The previous background model is used to segment out the foreground pixels from the background scene. By taking the Otsu thresholding value, this was calculated in section (2). It is well known that thresholding is an important aspect in background subtraction procedure and it is quite crucial to determine the most suitable threshold value. Therefore, morphological processes are applied for a more perfect silhouette. Alternatively, morphology is a broad set of image processing operations that process images based on shapes [13]. In this work,

morphological operations namely opening and closing were applied. Both of these operations can be manipulated using single or combination of specific structuring elements.

The main purpose of this step is to filter the presence of noisy pixels in the foreground image as shown in Figure (2).

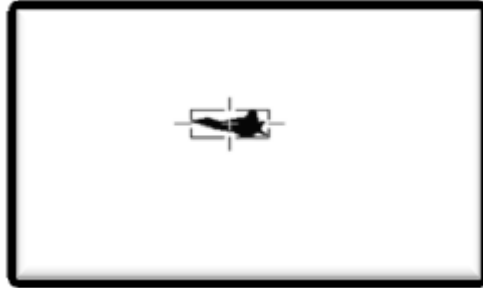


Figure (2) Background subtraction tracking algorithm

3.2 Normalize Cross Correlation (NCC)

The template matching method is simplest and widely used for detecting targets in image information, which is obtained from camera. The Template matching method is searching for the most similar frame target in the image with a template image obtained from prearranged information. It is important what similarity index is used for comparing the target with input frame [14]. The correlation process entails the correlation of frame and target. The Region of Interest (ROI) is a 101x101 pixel area from a given image segment and target is a 34x34 pixel area from an adjacent overlapping image segment.

When correlating the frame and target window using the correlation equation, a similarity in feature information will exist [15]. The correlation equation is as shown in equation (2).

$$NCC(T, f) = \frac{\sum_{x,y}(f(x,y) - \bar{f}_{u,v})(T(x-u,y-v) - \bar{T})}{\sqrt{\sum_{x,y}(f(x,y) - \bar{f}_{u,v})^2 \sum_{x,y} \sum_{x,y}(T(x-u,y-v) - \bar{T})^2}} \quad (2)$$

Where $f(x,y)$ denote the intensity value of the ROI of size at point (x,y) , $x \in \{0, \dots, -1\}$, $y \in \{0, \dots, \}$. The target is represented by a given template T of size .Common way to calculate the position $()$ of the target in the ROI is to evaluate the normalized cross correlation value NCC at each point (u,v) for f and the target T , which has been shifted by u steps in the x direction and by v steps in the y direction[16].

$\bar{f}_{u,v}$ denotes the mean value of $f(x,y)$ within the area of target T shifted to (u,v) . When feature information is identical within the frame and target, the correlation coefficient will be 1, which is denoting an exact matching. Figure (3) shows flow chart of algorithm.

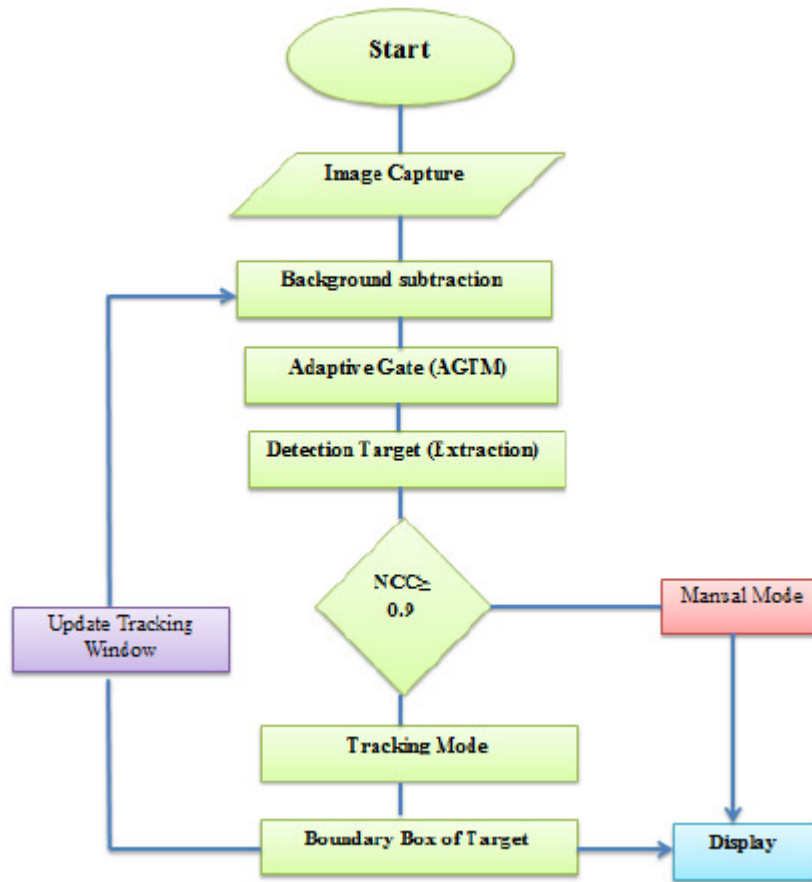


Figure (3) Flow chart of NCC algorithm

4. MOTION VECTOR

Because the position $[d_x(t_n + \Delta t), d_y(t_n + \Delta t)]$ is the horizontal component and the vertical component of the orthogonal coordinate, a transformation is used to transform the motion vector into the moving angles for motion control. The transformation in the pan direction is depicted in Figure (4) where the image plane is parallel to central plane of lens and the target moves in any direction. The locations of moving target are points A and B at time $t_n - \Delta t$ and t_n , respectively.

The target moving in real distances D_r so, to calculate the moving target in image using the formula:

$$d_y = D * \tan(\theta_{Tilt}) \quad (3)$$

$$d_x = D * \tan(\theta_{pan})$$

Angles θ_{Tilt} and θ_{pan} were calculated during the tracking mode equation (7) & (8) show that, the target was controlled at central point of image plane. In general case, the moving target is

faraway from the central plane of lens; therefore, the $\theta_{pan}(t_n + \Delta t)$ can be treated as θ . Fig.6 are the errors of tracking deviation between motion vectors and real moving distances in x and y directions, respectively. The significant meaning of these deviations is that the core of the target dose not matches the specific sampling points. Especially, the error deviations in the y direction are more than in the x direction. Exploring the reason, we find out that the illumination in x and y direction is different, which causes the diameter change of the target image in the different moving directions. The transformation in the tilt direction has the same fact as in pan direction.

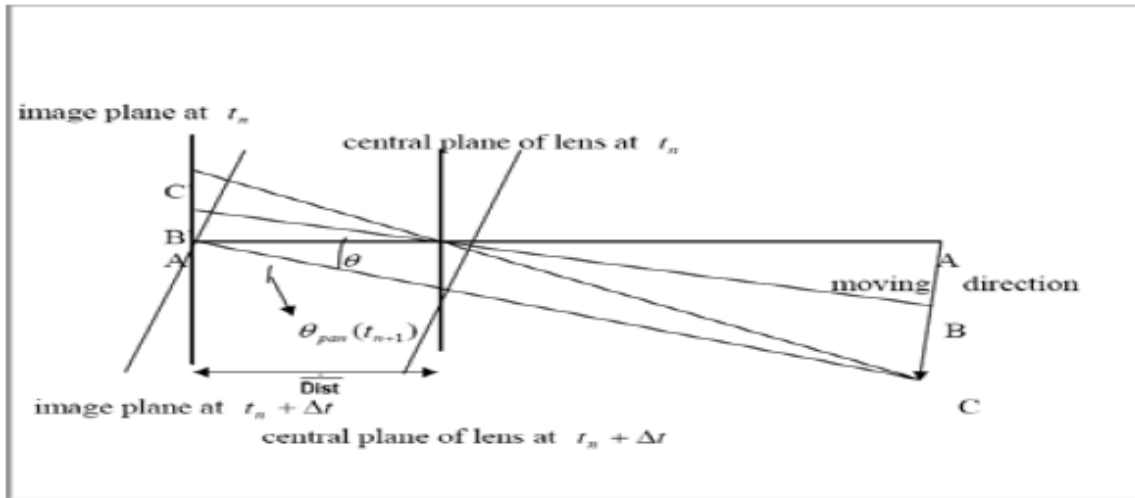


Figure (4) Image plane, central plane of lens, and target moving direction

4.1 Camera Field of View (CFOV)

Camera field of view is shown in Figure (5) the distance between camera and object D, the height and width of object in image plan is W_H and H_V , so the vertical field of view is given as:

$$2 \tan \theta_{VFOV} = \left(\frac{H_V}{2D} \right) \quad (4)$$

$$\theta_{VFOV} = 2 \tan^{-1} \left(\frac{30}{2 \times 45} \right) = 36.86 \text{ degree} \quad (5)$$

Where: θ_{VFOV} is angle vertical field of view.

The ratio of VFOV and image height as:

$$\sigma_{TR} = \frac{\text{imgheight}}{\theta_{VFOV}} \quad (6)$$

Same as in horizontal field of view.

The ratio of HFOV and image width as:

$$\sigma_{PR} = \frac{\text{imgWidth}}{\theta_{HFOV}} = \frac{640}{56.3} = 11.36 \text{ pixels/degree} \quad (7)$$

The ratio of VFOV and image width as:

$$\sigma_{TR} = \frac{\text{imgheight}}{\theta_{HFOV}} = \frac{480}{36.86} = 13.02 \text{ pixels/degree} \quad (8)$$

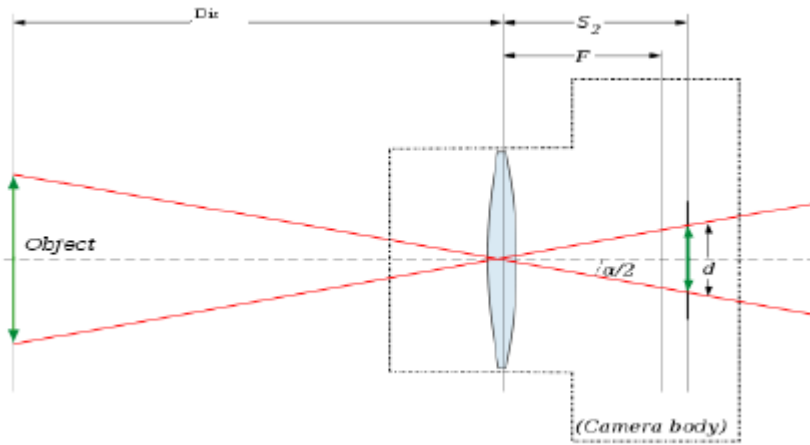


Figure (5) Shown camera field of view in vertical direction

Pan and tilt angles can be calculated as:

$$\theta_{Pan}((t_n + \Delta t)) = \left(\frac{d_x(t_n + \Delta t)}{\sigma_{PR}} \right) \quad (9)$$

$$\theta_{Tilt}((t_n + \Delta t)) = \left(\frac{d_y(t_n + \Delta t)}{\sigma_{TR}} \right) \quad (10)$$

4.2 Target Moving in Image Plane

Table (1) shows the transformation of target in image plane the tracking mode get the distance measurement in image plane compare it with the real distance in world space.

Table (1) the motion angle, distance and actual moving angle and distances of the tracking target

Sample points	Moving direction	Motion in X- direction			Motion in Y- direction			Real distance(cm)	
		Measurement			Measurement			Measurement	
		Pixels number	Distance cm	angle ^o	Pixels number	Distance cm	angle ^o	x-dist	y-dist
p_1	$p_1 \text{ top}_2$	20	10.75	1.53	42	25.95	3.69	10	25
p_2	$p_2 \text{ top}_3$	34	18.24	2.61	40	24.61	3.52	20	25
p_3	$p_3 \text{ top}_4$	25	-13.41	-1.92	-48	-29.55	-4.22	-15	30
p_4	$p_4 \text{ top}_3$	-17	-9.11	-1.30	-59	-36.35	-5.19	-10	35
p_5	$p_1 \text{ top}_3$	54	29.00	4.15	40	24.61	3.52	30	25
p_6	$p_2 \text{ top}_4$	-54	-29.00	-4.15	-40	-24.61	-3.52	-30	-25
p_7	$p_3 \text{ top}_1$	-20	-10.72	1.53	42	-14.75	-2.11	-10	-25

Note : (-) negative sign means the opposite direction

$\sigma_{\text{TiltRatio}} = 13.02$ pixels per degree, $\sigma_{\text{PanRatio}} = 11.36$ pixels per degree, σ_{TR} is a pixels Ratio in x direction and σ_{PR} is a pixels Ratio in y direction. Target distance from Camera = 400 mm.

5. EXPERIMENTAL RESULTS

To test the normalize cross correlation on the scene, different experiments were carried out to test the proposed algorithm to track object in the image frames to identify any position of target in the scene (static scene static camera(video file.avi)). Then some testes were carried out to track objects in real time (static camera moving object).

5. 1. Static Scene Static Camera

This test was used to detect object based on Otsu threshold, because the frame difference method cannot detect static objects. Estimated positions of the object (based on camera field of view) with respect to real target positions are summarized in Table.1. All values are measured with respect to center of camera view. Table.2 and Table.3 show that the measurement positions ($\theta_{\text{pan_measurement}}$, $\theta_{\text{Tilt_measurement}}$) are close to real positions ($\theta_{\text{pan_real}}$, $\theta_{\text{Tilt_real}}$). The algorithm can detect and segment the background object form foreground. The specified area from multiple moving objects are appears in the scene. The Otsu algorithm implemented to the segmented object at each frame. Estimated position (degree) in (x, y) directions used the ratio between number of pixels and camera field of view using equation (9) and equation (10). The real time object tracking has special performance of equipment, this system used low cost mini CCD camera it has wide field of view and short rang it can't detect and track target moving in long rang, it can detect target moving in short rang also this system has special performance it can work in and out door.

Although system has high accuracy to detect and track target in short rang, it was compared with the result based on the camera calibration. The results of this system better and fewer requirements than system used camera calibration.

To detect and track target it should make target moving in track-window in this time system received command form user to start tracking target, if target moving in high speed may be the algorithm failure to tracking the target because target moving out of track-window in this case it better to zoom the track-window until the target found, before that reduce the track-window again.

Table (2) Summary for estimated and measured Angles values (angles)

Angle in X direction			Angle in Y direction		
Measurement angle ^o	Estimate angle ^o	Error angle ^o	Measurement angle ^o	Estimate angle ^o	Error angle ^o
1.53	1.43	0.10	3.69	3.57	0.12
2.61	2.86	0.25	3.52	3.57	0.05
-1.92	-2.14	0.25	-4.22	-4.28	0.06
-1.30	-1.43	0.13	-5.19	-5.00	0.19
4.15	4.15	0.13	3.52	3.57	0.05
-4.15	-4.28	0.13	-3.52	-3.57	0.05

Table (3) Summary for estimated and measured position values (distance)

Position in X direction			Position in Y direction		
Measurement distance (cm)	Estimate distance (cm)	Error distance (cm)	Measurement distance (cm)	Estimate distance (cm)	Error distance (cm)
10.75	10	0.75	25.8	25	0.80
18.24	20	1.76	24.6	25	0.40
-13.41	-15	1.59	-29.55	-30	0.45
-9.11	-10	0.89	-36.35	-35	1.35
29	30	1.00	24.61	25	0.39
-29	-30	1.00	-24.61	-25	0.39

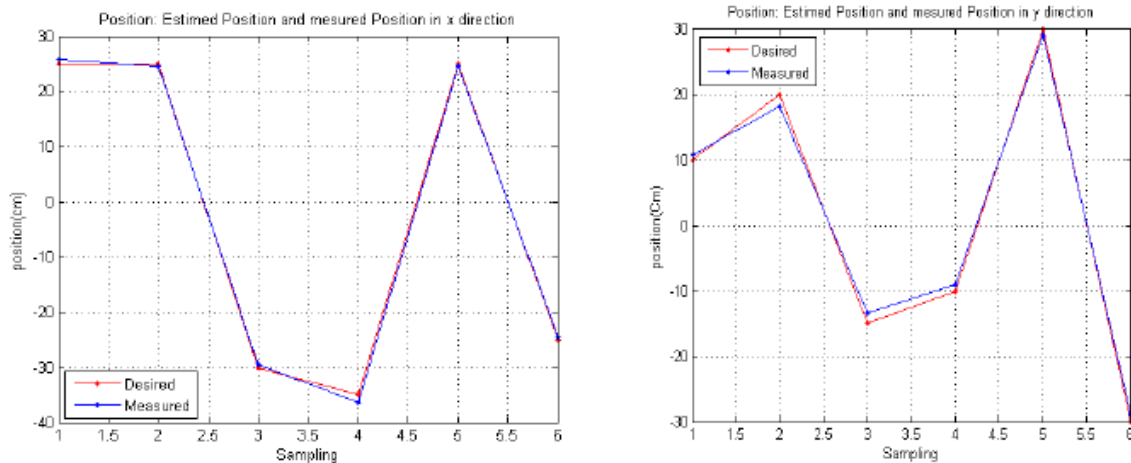


Figure (6) Desired position and measured position in X , Y direction

6. CONCLUSIONS

The real time object tracking based on camera field of view was achieved. The algorithms merge the histogram projection and normalize cross correlation technique to detect, segment and track static and moving object in real time. The ratio $\sigma_{TiltRatio}$ and $\sigma_{PanRatio}$ was calculated and implemented to determine the target distance in world coordinates. Different tests are carried to verify the validity of the proposed algorithm such as, static scene static camera and moving object static camera. The results show that, the static and moving object can be detected and segmented without noise and false detections in the scene. The average processing time was observed when we used track-window is better than process full frame. It is depending on the position of the object with respect to the camera view. The key factors to limit the application of this system in specific areas are the FOV limit of CCD camera and one target detected by the cross-correlation method at one time. Because of illumination effects on the moving object, it is not easy to get more precise tracking.

7. FUTURE WORK

These results are promising to develop the system in future it needs to include software and hardware to get high performance of system.

It is better to use open or close loop control with shaft encoder to know the position of motor when system connected to other system. Embedded system has high response, speed of processing and less weight.

The GPS data can be used in one time by interring manually, GPS and digital compass to know the directional and position.

The System can be integrated with Laser Rang Finder (LRF) to get target distance at each time. Matlab software was used to implemented algorithms, it has good features to test all algorithms, but in real time application need to other platform software.

REFERENCES

- [1] Bluestien H.B, Prediction of Satellite Cloud Patterns Using Spatial Fourier Transforms, S.B, Massachusetts Institute of Technology [1971].
- [2] Parcak H. Sarah, Satellite Remote Sensing for Archaeology, London & Newyork [2009].
- [3] Shiao S.Ying, Design and Implementation of Real-Time Tracking System Based on Vision Servo Control, Tamkang Journal of Science and Engineering, Vol. 4, No. 1, pp. 45-58 [2001].
- [4] Gonzalez C. Rafael & Woods R. E, Digital Image Processing. Addison-Wesley Publishing Company[1993].
- [5] Dr. Greensted A, Digital Image Processing (Java Software). Wikipedia, july [2010].
- [6] Ahmed N & Roa K.R, Orthogonal Transforms for Digital Signal Processing, New York Spring [1975].
- [7] Gonzales C. Rafael, Digital Image Processing Using Matlab, Math Work [2004].
- [8] Liang Xu, Threshold Selection Method from Gray-Level Histograms, Electro-Technical Laboratory, Tokyo University(2007), Tokyo, Japan, IEEE Transactions on System, Man, and Cybernetics. [1979].
- [9] Sezgin M. & Sankur B, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation", Journal of Electronic Imaging [2003].
- [10] Baldini G, Campadelli P, Cozzi D & Lanzarotti R, A Simple and Robust Method for Moving Target Tracking, roceedings of the International Conference Signal Processing, Pattern Recognition and Applications (SPPRA2002), Crete, Greece June [2002].
- [11] John C.Russ, The Image Processing Hand Book 3rd Edition, ACRC Published in cooperation with IEEE Press [1998].
- [12] Padmini J, Abhishek S & Suman K Mitra, Background Subtraction in Videos Using Bayesian Learning with Motion Information, Dhirubhai Ambani Institute of Information and Communication Technology, India [2007].
- [13] Jong S. Bae & Taek L.Song, Image Tracking Algorithm Using Template Matching and PSNF-m, International Journal of Control, Automation, and Systems [June 2008].
- [14] JaneenM.Carter, Correlation Evaluation of Intensity and Color Band Images, Morgan State University [Jan 23 1995].
- [15] Kai B & Hanebeck U, Template Matching Using Fast Normalize Cross Correlation, institute of Automatic control Engineering, technishche university Munchen , Germany [2001].

AUTHORS

HASSAN ELPASERI ELAMIN ELTAHER was born in July 5, 1982, Khartoum, Sudan. He received his M.Sc. in image processing, College of Science 2012 Sudan University of Science & technology, Sudan, high diploma in general physic 2007, Sudan University of Science & technology, B.Sc. in physic Science 2005 Sudan University of Science & technology, Sudan. Currently he is DIP & Thermal Imaging System Researcher in institute of Laser, Sudan University.

