International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869, Volume-2, Issue-5, May 2014

Self Organizing Approach for Moving Object Detection and Tracking for Visual Surveillance

Ms. Neha Dhavale , Dr. Seema Biday, Mr. Saurabh Athalye

Abstract— **Visual surveillance is a very active research area in computer vision and moving object detection and tracking is often the first step in applications such as video surveillance. Here, we propose a methodology to detect and track human image based on self organization through artificial neural networks. We choose the HSV color space, relying on the hue, saturation and value properties of each color to represent each weight vector. A neural network mapping method is proposed to use a whole trajectory incrementally in time fed as an input to the network. The adopted artificial neural network is organized as a 2-D flat grid of neurons (or nodes). Each node computes a function of the weighted linear combination of incoming inputs, where weights resemble the neural network learning. Each node could be represented by a weight vector obtained, collecting the weights related to incoming links. An incoming pattern is mapped to the node whose model is "most similar" (according to a predefined metric) to the pattern, and weight vectors in a neighborhood of such node are updated. We would focus on combining contour projection analysis with shape analysis to remove the shadow effect.**

Index Terms— **HSV, SOBS, moving objects.**

I. INTRODUCTION

Visual surveillance systems include object detection, object classification, tracking, activity understanding, and semantic description. Keeping human watch 24x7 is not possible as we all know that humans can easily be distracted and a small distraction in very sensitive and highly secure area can lead to big loses. To overcome this human flaw in the area of monitoring, the concept of making monitoring automatic came into existence. Since, video surveillance has came in the market, researches have been taking place in order to make to more easy, accurate, fast and intelligent. The scientific challenge is to devise and implement automatic systems able to detect and track moving objects, and interpret their activities and behaviors. The detection of moving objects in video streams is the first relevant step of information extraction in many computer vision applications. Object tracking, in general, is a challenging problem. Difficulties in tracking objects can arise due to abrupt object motion, changing appearance patterns of the object and the scene, non rigid object structures, object-to-object and object-to-scene occlusions, and camera motion. Tracking is usually performed in the context of higher-level applications that require the location and/or shape of the object in every frame. Object tracking is an important task within the field of computer vision. There are three key steps in video analysis: detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior. We studied a real-time visual tracking system on a controlled pan-tilt camera. The input/output HMM (Hidden Markov Model) is employed to model the overall visual tracking system in the spherical camera platform coordinate. In order to fast detect and track targets on a moving camera at the same time, the optical flow is adopted to observe the different displacement in the image sequence. A new efficient moving target detection method is proposed which is a improved background subtraction to detect moving objects. Two significant advantages were the improved the background subtraction and increased algorithm's running efficiency and offset sensitive deficiency of the light changes. Background subtraction is a widely used approach for detecting foreground objects in videos from a static camera. Indoor surveillance applications such as home-care and health-care monitoring, a motionless person should not be a part of the background. A reference background image without moving objects is, therefore, required for such applications. In this paper, an ICA (Independent Component Analysis)-based background subtraction scheme for foreground segmentation is presented. The ICA model is based on the direct measurement of statistical independency that minimizes the difference between the joint PDF and the product of marginal PDFs, in which the probabilities are simply estimated from the relative frequency distributions.. Convergence of SEOS (Simultaneous Estimation of Optical flow and State dynamics) was evaluated for both the Gauss-Seidel and Jacobi iterative techniques. The SEOS converges for both Gauss-Seidel and Jacobi iterative schemes for any initial approximation. Background subtraction is an active researching field, because it can be used in many applications. An efficient background subtraction approach is the base which determines performance of the whole system.

Using frame differencing on frame-by-frame basis a moving object, if any, is detected with high accuracy and efficiency. Once the object has been detected it is tracked by employing an efficient Template Matching algorithm. The templates used for the matching purposes are generated dynamically. This ensures that any change in the pose of the object does not hinder the tracking procedure. To automate the tracking process the camera is mounted on a pan-tilt arrangement, which is synchronized with a tracking algorithm. As and when the object being tracked moves out

Manuscript received May 02, 2014.

Ms. Neha Dhavale, Department of Computer Engineering, Terna Engineering College, Nerul.

Dr. Seema Biday, Department of Computer Engineering, Terna Engineering College, Nerul.

Mr. Saurabh Athalye, Department of Electronics & Telecommunication Engineering, Rajendra Mane College of Engineering & Technology, Ambav.

.

of the viewing range of the camera, the pan-tilt setup is automatically adjusted to move the camera so as to keep the object in view.

Evaluation based on GT (Ground Truth) offers a framework for objective comparison of performance of alternate surveillance algorithms. Such evaluation techniques compare the output of the algorithm with the GT obtained manually by drawing bounding boxes around objects, or marking-up the pixel boundary of objects, or labeling objects of interest in the original video sequence. Manual generation of GT is an extraordinarily time-consuming and tedious task and, thus, inevitably error prone even for motivated researchers. Interpretation of evaluation results is based on the type of GT used for comparison.

II. PROPOSED METHOD

Visual surveillance systems include object detection, object classification, tracking, activity understanding, and semantic description. The detection of moving objects in video streams is the first relevant step of information extraction in many computer vision applications. Object tracking, in general, is a challenging problem. The usual approach to moving object detection is through background subtraction that consists in maintaining an up-to-date model of the background and detecting moving objects as those that deviate from such a model. The main problem in moving object detection and tracking is its sensitivity to dynamic scene changes, and the consequent need for the background model adaptation via background maintenance.

The proposed method Self Organizing Background Subtraction (SOBS) is to adopt a biologically inspired problem-solving method based on visual attention mechanisms. Currently, methods used in moving object detection are mainly the frame subtraction method, the background subtraction method and the optical flow method. We focus on overcoming all the problems these methods such as light changes, moving background, cast shadows, bootstrapping, and camouflage. The objective is to detect the objects that keep the user attention in accordance with a set of predefined features, including gray level, motion and shape features. The approach defines a method for the generation of an active attention focus to monitor dynamic scenes for surveillance purposes. The idea is to build the background model by learning in a self-organizing manner for many background variations, i.e., background motion cycles, seen as trajectories of pixels in time. Based on the background model through a map of motion and stationary patterns, the algorithm can detect motion and selectively update the background model. A neural network based method is proposed to use a whole trajectory incrementally in time fed as an input to the network. This makes the network structure much simpler and the learning process much more efficient. The neural network is organized as a 2-D flat grid of neurons (or nodes) and, similarly to self-organizing maps (SOMs) or Kohonen networks, allows to produce representations of training samples with lower dimensionality, at the same time preserving topological neighborhood relations of the input patterns (nearby outputs correspond to nearby input patterns).

The algorithm can be explained as follows:-
Input: pixel value p_t in frame I_t ,t=0…..Last frame
Output: background/foreground binary mask value $B(p_t)$
1. Initialize model C for pixel p_0 and store it into A
2. for t=1, LastFrame
3. Find best match c_m in C to current sample p_t
4. if (c_m found) then
5. $B(p_t)=0$ //background
6. update A in the neighborhood of c_m
7. else if(p_t shadow) then
8. $B(p_t)=0$ //background
9. else
10. $B(p_t)=1$ //foreground

The proposed method can be explained diagrammatically as shown in figure 2.1

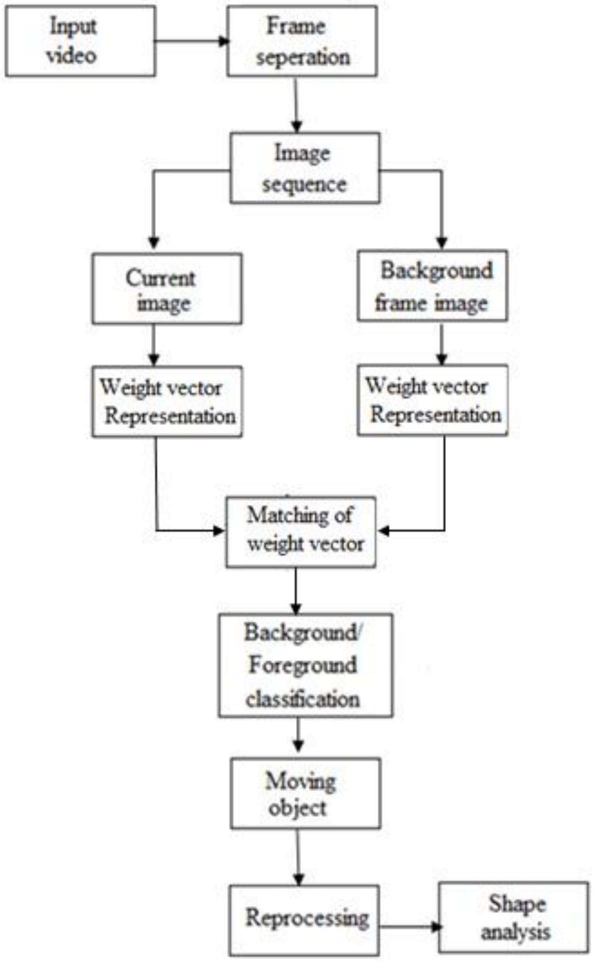

Figure 2.1: Flow chart for human motion detection & tracking.

III. PERFORMANCE EVALUATION

The proposed system can be experimented with different settings of adjustable parameters which can be used for performance evaluation.

a) Processing time

We calculate the elapsed time using tic (Timer on) and toc (Timer off).

tic and toc functions work together to measure elapsed time. We will evaluate the elapsed time for three methods viz. Frame subtraction, Background subtraction and the proposed method. Based on the processing time, we would determine which system is fastest.

b) Accuracy

After classification of pixels as background & foreground pixels, we will retrieve color image from it. In retrieved image, we will check accuracy of retrieved object which is being tracked and represent it in percentage by comparing with original image. We can compare this accuracy measure with output of other object tracking algorithms. For measuring accuracy there are different metrics viz. Recall, Precision, and Similarity.

1) Recall

Recall also known as detection rate, gives the percentage of detected true positives as compared to the total number of true positives in the ground truth.

$$\text{Recall} = \frac{tp}{tp+fn} \qquad (3.1)$$

Where,

t_p = total number of true positives

$(t_p + f_{n)}$= total number of false negatives, and indicates the total number of items present in the ground truth.

2) Precision

Precision, also known as positive prediction, gives the percentage of detected true positives as compared to the total number of items detected by the method.

$$\text{Precision} = \frac{tp}{tp+fp} \qquad (3.2)$$

Where,

f_p = total number of false positives

$(t_p + f_{p)}$= total number of detected items.

Using the above mentioned metrics, generally, a method is considered good if it reaches high Recall values, without sacrificing Precision.

3) Similarity

The pixel-based similarity measure is defined as-

$$\text{Similarity} = \frac{t_p}{t_p + f_n + f_p} \qquad (3.3)$$

IV. Experimental Results

Experimental results for moving object detection using the proposed approach have been produced for several image sequences. Here, we describe three different sequences, that represent typical situations critical for video surveillance systems, and present qualitative results obtained with the proposed method.

1) Sequence Walk1

Sequence Walk1 of the CAVIAR Project comprises 31 frames of 512 * 512 spatial resolutions, captured at a frequency of 25 fps. The scene consists in a laboratory where a man comes in, walks around, and leaves on the left side. This is an example of hard sequence.

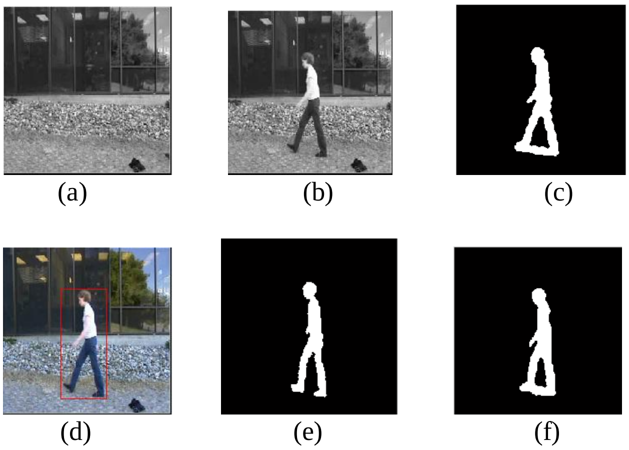

Figure 4.1: Results of SOBS algorithm on sequence Walk1: (a) Background image; (b) Current image; (c) SOBS result; (d) Tracking result; (e) Ground truth; (f) Actual output

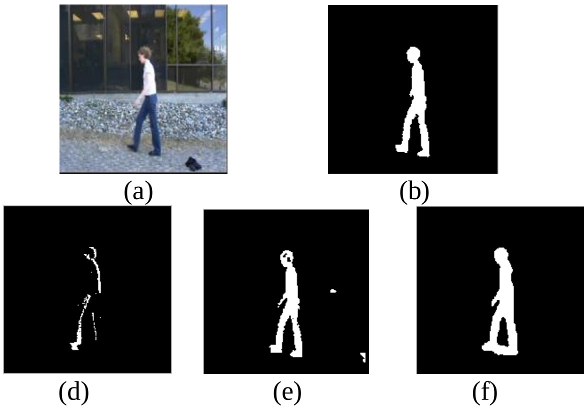

Figure 4.2: Segmentation of sequence Walk1: (a) Test image; (b) Ground truth; (c) Frame subtraction result; (d) Background subtraction result; (e)SOBS result

The accuracy values of Pixel based accuracy values for sequence Walk1 can be observed in Table I.

TABLE I

Pixel based accuracy values for sequence Walk1

Parameter	**Frame subtraction**	**Backgroun d subtraction**	**SOBS**
Recall	0.1400	0.9501	0.9845
Precision	0.6169	0.6933	0.7838
F_1 metric	0.2282	0.8017	0.8728
Similarity	0.1288	0.6690	0.7743

2) Sequence Hall monitor

Sequence Hall monitor is an indoor sequence consisting of 287 frames of 320 * 240 spatial resolution, acquired at a frequency of 30 fps (frames per second). The scene consists hall, where a man comes out, leaves a bag on the floor, and then goes in the room. While the first man passes, the another man comes into the hall and moves towards the room. It represents an example of easy sequence, in that lighting conditions are quite stable and moving objects are well contrasted with the background (there is no camouflage); however, strong shadows cast by moving objects can be observed in the entire sequence.

The accuracy values of Pixel based accuracy values for sequence Hall monitor can be observed in Table II.

TABLE II
Pixel based accuracy values for sequence Hall monitor

Parameter	Frame subtraction	Background subtraction	SOBS
Recall	0.0659	0.7908	0.9192
Precision	0.6244	0.4258	0.7758
F_1 metric	0.1192	0.5535	0.8415
Similarity	0.0633	0.3827	0.7264

3) *Sequence water surface*

Like the previous one, also sequence water surface consisting 60 frames of 160 * 120 spatial resolutions, captured at a frequency of 15 fps. Here, it has been chosen in order to test our method ability to cope with moving background. The outdoor scene includes (moving) waves of water in the background and, finally, a man passing in front of the camera; here we are not interested in the waving water, but only in extraneous moving objects (the man).

The accuracy values of Pixel based accuracy values for sequence Water surface can be observed in Table III.

TABLE III
Pixel based accuracy values for sequence Water surface

Parameter	Frame subtraction	Background subtraction	SOBS
Recall	0.1624	0.8308	0.6854
Precision	0.9845	0.4475	0.8954
F_1 metric	0.2789	0.5817	0.7764
Similarity	0.1620	0.4101	0.6346

V. CONCLUSION

We have implemented a new self-organizing method for modeling background by HSV model which allows foreground/background separation for scenes from stationary cameras, strongly required in video surveillance systems. Unlike existing methods viz.frame subtraction and background subtraction that use individual flow vectors as inputs, our method learns background motion trajectories in a self organizing manner; this makes the neural network structure much simpler. Experimental results, using different sets of data and comparing different methods, have demonstrated the effectiveness of the proposed method, which proves also robust to noise, moving backgrounds, gradual illumination changes, and cast shadows, and has no bootstrapping limitations.

REFERENCES

[1] J. M. Ferryman, Ed, *"Evaluation of Tracking and Surveillance"* in *Proc.* 9th IEEE Int. Workshop on Performance, 2006.
[2] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L.Wixson, "*A system for video surveillance and monitoring*," Tech.Rep. CMU-RI-TR-00-12, The Robotics Inst., Carnegie Mellon Univ., Pittsburgh, PA, 2000.
[3] Alper Yilmaz, Omar Javed, Mubarak Shah, "*Object Tracking: A Survey*", ACM Computing Surveys, Vol. 38, No. 4, Article 13, Publication date: December 2006.
[4] Cheng-Ming Huang, Yi-Ru Chen , Li-Chen Fu, *"Real-Time Object Detection and Tracking on a Moving Camera Platform"*, ICCAS-SICE, National Conference,2009.
[5] Lianqiang Niu, Nan Jiang ,"A Moving Objects Detection Algorithm on Imporved Background Subtraction ", Eighth International Confrence, Intelligent Systems Design and Applications,2008,ISDA '08
[6] Tsai DM, Lai SC, *"Independent Component Analysis-Based Background Subtraction for Indoor Surveillance"*, IEEE Trans Image Process, 2009 Jan; 18(1):158-67, doi: 10.1109/TIP.2008.2007558.
[7] Kinoshit, K.Enokida, M.Izumid, M.Murakami, *"Tracking of a Moving Object Using One-Dimensional Optical Flow with a Rotating Observer"*,9th International Conference ,Control, Automation, Robotics and Vision, 2006, ICARCV '06.
[8] Bauer, N.J, Pathirana, P.N, *"Object Focused Simultaneous Estimation of Optical Flow and State Dynamics"*, International Conference, Intelligent Sensors, Sensor Networks and Information Processing, 2008,ISSNIP 2008.
[9] Zhen Tang, Zhenjiang Miao , *"Fast Background Subtraction and Shadow Elimination Using Improved Gaussian Mixture Model"*, Haptic, Audio and Visual Environments and Games, 2007. HAVE 2007, IEEE International Workshop.
[10] Karan Gupta1, Anjali V. Kulkarni "*Implementation of an Automated Single Camera Object Tracking System Using Frame Differencing and Dynamic Template Matching*", Indian Institute of Technology, Kanpur, India.
[11] J. Black, T.J,Ellis, and P. Rosin, "A *NovelMethod for Video Tracking Performance valuation*",IEEE Workshop on Performance Analysis of Video Surveillance and Tracking (PETS'2003), pages 125–132, October 2003.
[12]http://perception.i2r.astar.edu.sg/bk_model/bk_index.html
[13] http://www.cvg.rdg.ac.uk/PETS2006/data.html