International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 10, October 2016)

Multiple Moving Objects Identification for Smart Video Surveillance Systems

M. Vamshi Krishna[1], R. Raja Kishore[2], Dr. M. Narsing Yadav [3]

[1]*PG Scholar,* [2]*M.Tech, Asst. Prof.,* [3]*Ph.D., Prof. HOD, CERD, Department of Electronics and Communication Engineering, Mallareddy Institute of Engineering and Technology, Secunderabad, India,*

Abstract**—Autonomous video surveillance and monitoring has a rich history. A new method for detecting multiple moving objects has been proposed in this paper. In this we developed a new algorithm based on filtering the frames and removing the background. In detecting multiple objects there are many problems like speed of objects and size of objects. In this we developed a robust routine for detecting multiple moving objects. This proposed method is proved to be robust in various environments and different types of background scenes. The experimental results prove the feasibility of proposed method. Experiments on real scenes show that the algorithm is effective for object detection and tracking.**

Keywords**—tracking; moving objects; detection**

I. Introduction

Visual surveillance in dynamic scenes, especially for humans and vehicles, is currently one of the most active research topics in computer vision. It has a wide spectrum of promising applications, including access control in special areas, human identification at a distance, crowd flux statistics and congestion analysis, detection of anomalous behaviours, and interactive surveillance using multiple cameras, etc. In general, the processing framework of visual surveillance in dynamic scenes includes the following stages: modelling of environments, detection of motion, and classification of moving objects, tracking, understanding and description of behaviours, human identification, and fusion of data from multiple cameras. We review recent developments and general strategies of all these stages. Finally, we analyse possible research directions, e.g., occlusion handling, a combination of two and three-dimensional tracking, a combination of motion analysis and biometrics, anomaly detection and behaviour prediction, content-based retrieval of surveillance videos, behaviour understanding and natural language description, fusion of information from multiple sensors, and remote surveillance.

II. Paper Organization

This paper is organized as follows. Chapter III discusses about the Previous Works.

Features Extraction is discussed in Chapter IV. Experimental results are discussed in Chapter Conclusions are driven in Chapter VI.

III. Previous Works

Motion detection, tracking, behaviour understanding, and personal identification at a distance can be realized by single camera-based visual surveillance systems. Multiple camera-based visual surveillance systems can be extremely helpful because the surveillance area is expanded and multiple view information can overcome occlusion. Tracking with a single camera easily generates ambiguity due to occlusion or depth. This ambiguity may be eliminated from another view. However, visual surveillance using multiple cameras also brings problems such as camera installation (how to cover the entire scene with the minimum number of cameras), camera calibration, object matching, automated camera switching, and data fusion.

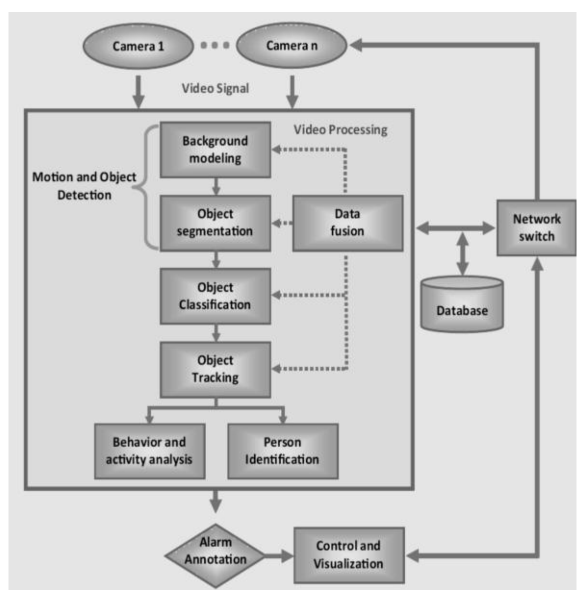

Fig 1 General frame work of an automated visual surveillance system

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 10, October 2016)

A. Wavelet Based Detection

Video tracking systems have to deal with variously shaped and sized input objects which often result in a massive computing cost of the input images. In this method a 2d Discrete Wavelet Transform can be used to decompose the image into four sub band images namely LL, LH, HL & HH. The DLLBS scheme processes only the part of LL band image due to the low computation cost and noise reduction issues.

2-D DWT is used to detect and track moving objects. The 2-D DWT can be used to decompose an image into four sub band images LL, LH, HL & HH. But it only processes the part of LL band image due to the consideration of low computing cost and noise reduction issues.

B. Contrast Change (CC) Model

In CC model objects are detected based on the Contrast Change feature in CC model. Contrast Change is not a definite feature such as absolute luminance but a relative salient feature. It means the effectiveness of CC model is decided not by its absolute value but by its relative salience in different environments.

IV. Objects Detection & Feature Extraction

In order to detect moving objects in a video we need to divide the video into frames. So, a video of multiple moving objects is selected. Then the video is divided into frames.

A. Noise Removal

Here we used a Weiner filter in order to remove the noise from the images. Wiener filter is a filter used to produce an estimate of a desired or target random process by linear time-invariant (LTI) filtering of an observed noisy process, assuming known stationary signal and noise spectra, and additive noise. The Wiener filter minimizes the mean square error between the estimated random process and the desired process. The goal of the Wiener filter is to compute a statistical estimate of an unknown signal using a related signal as an input and filtering that known signal to produce the estimate as an output.

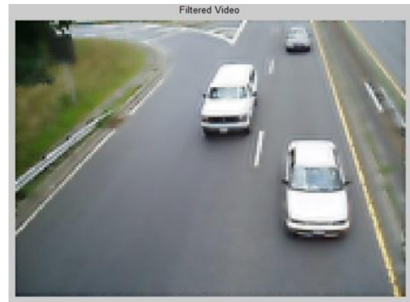

Fig 2. Filtered Image

B. Background Subtraction

A motion detection algorithm begins with the segmentation part where foreground or moving objects are segmented from the background. The simplest way to implement this is to take an image as background and take the frames obtained at the time t, denoted by I (t) to compare with the background image denoted by B. Here using simple arithmetic calculations, we can segment out the objects simply by using image subtraction technique of computer vision meaning for each pixels in I(t), take the pixel value denoted by P[I(t)] and subtract it with the corresponding pixels at the same position on the background image denoted as P[B].

$$P[F(t)] = P[I(t)] - P[B]$$

The background is assumed to be the frame at time t. This difference image would only show some intensity for the pixel locations which have changed in the two frames. Though we have seemingly removed the background, this approach will only work for cases where all foreground pixels are moving and all background pixels are static. A threshold "Threshold" is put on this difference image to improve the subtraction.

$$|P[F(t)] - P[F(t+1)]| > \text{Threshold}$$

Fig 3. Input frame

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 10, October 2016)

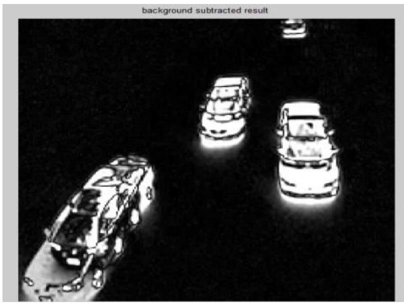

Fig 4. Background subtracted Frame

Fig 5. Input Image

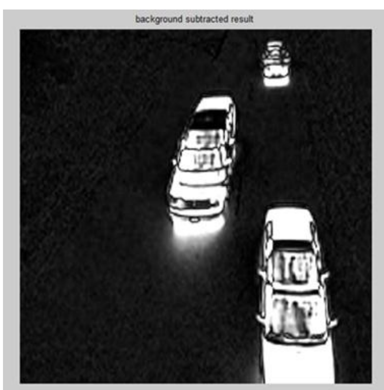

Fig 6. Background subtracted image

C. Object Labeling

First an image is taken and is converted to binary image. When converted into a binary image the image has some unwanted part in it. So we fill the holes of the binary image. Then it tends to look as binary image without noise. We calculate the background image by averaging the first 5 images. So in order to label the objects we first detect the centroid of the objects, and then calculate the number of elements in an array or subscripted array expression. We even measure the properties of image regions.

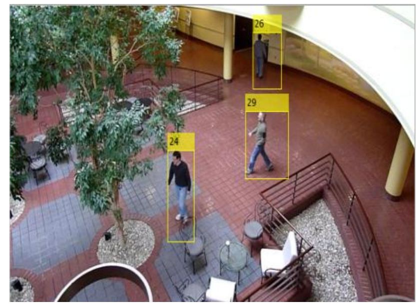

Fig 7. Moving Objects Labelled

D. Bounding Box

The labeled difference image is scanned pixel by pixel from right to left, left to right, top to bottom, and bottom to top to find the bounding box of moving objects. After the scanning process is completed, the top left coordinates, height, and width of each bounding box is calculated.

Object identification is the last stage. The RGB color is used as the color information of the moving objects. The first color information calculated is mean value of each color object. The mean value is calculated for each color component of RGB space.

We even applied the Kalman filter to get the predicted and actual centroid and velocity. We design the Kalman filter using

kalmanx = centroidx(i) - xcorner
kalmany = centroidy(i) – ycorner

Where i is the number of frames in the video. We use the former state to predict the new centroid.

Fig 8. Box around moving object

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 10, October 2016)

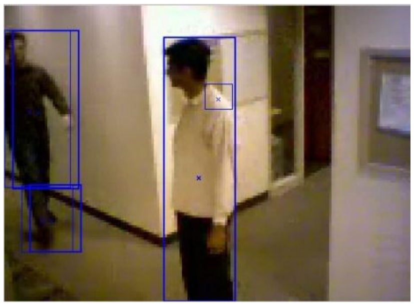

Fig 9. Bounding Box around the moving objects

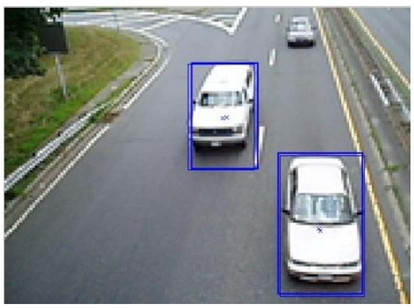

Fig 10. Detected objects

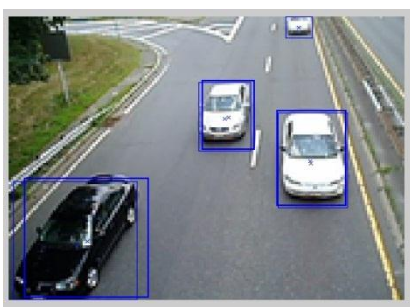

Fig 11. Tracking of moving objects

V. EXPERIMENTAL RESULTS

In this section, we did several experiments to prove the feasibility of the proposed tracking and identification method. We used an entry-level video camera and capture card to capture test sequences in different campus, and simulated several cases of condition for moving objects, such as single object indoors, single object in outdoor environment with static background, single object in outdoor environment with varied background, and multiple objects outdoors.

In order to test the proposed tracking and identification method for multiple objects, we took a video which has multiple moving objects. In each instance the no.of objects moving are being changed. Even other test video sequences are also used during experiments.

We choose to address the problem of tracking people. This is one of the most difficult problems in visual tracking because people can meet, form groups, cross-over etc. It is clear that simple connected component methods that extract several isolated regions and label each region as a person will not be adequate in such situations.

Although there are some pixels misclassified in the segmented results, the objects are correctly classified and are successfully tracked even when they are subject to occlusions. The proposed method successfully tracked both persons when they pass across and occluded each other.

Table 1
The results of objects identifying & tracking

S.No	No.of frames	Moving objects	Objects detected	Rate %
1	71	10	9	90.26
2	80	50	49	98.54
3	152	195	194	99.65
4	171	200	180	90.89
5	280	22	20	90.95
6	120	100	99	99.26

Although we could find some drawbacks in the proposed system, the objects are detected and tracked successfully. The proposed method successfully tracked the objects in the video. Even if the persons are close to each other it is able to detect the persons or objects successfully. The above table is used to evaluate the performance of the proposed model, where each sequence presents a different type of difficulty that a practical task may meet.

VI. CONCLUSION

A new method for detecting and tracking multiple moving objects is proposed. The moving objects are identified by their color and background. The experimental results prove the feasibility and usefulness of the proposed method. The correct identification rate highly depends on the correctness of the moving object detection and feature representation. The proposed method has been tested and validated by a significant number of experiments.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 10, October 2016)

The proposed model has proved to be robust in various environments (including indoor and outdoor scenes) and different types of background scenes. We showed that the proposed method can successfully segment and track people through occlusion, using both outdoors and indoors video sequences.

REFERENCES

[1] Detection and Tracking of Multiple Moving Objects with Occlusion in Smart Video Surveillance Systems in Intelligent Systems (IS), 2010 5th IEEE International Conference.

[2] S. J. McKenna et al., Tracking groups of people, Computer Vision Image Understanding, 2000, pp. 42–56.

[3] W. Hu et al., "A survey on visual surveillance of object motion and behaviors," IEEE Trans. Syst. Man Cybernetics–Part C, vol. 3, 2004, pp. 334–352.

[4] Fang-Hsuan Cheng and Yu-Liang Chen, "Real time multiple objects tracking and identification based on discrete wavelet transform," Pattern Recognition, 2006, pp. 1126 – 1139.

[5] Kaiqi Huanga, LiangshengWanga, Tieniu Tana and Steve Maybankb, "A real-time object detecting and tracking system for outdoor night surveillance," Pattern Recognition, 2008, pp. 432 – 444.

[6] Wing Wu and Ting Yu, "A field model for human detection and tracking," IEEE Transactions on pattern analysis and machine intelligence, Vol. 28, No.5, 2006.

[7] Ge, W.; Gao, L.-Q. & Sun, Q. (2007). A method of multi-scale edge detection based on lifting scheme and fusion rule, International Conference on Wavelet Analysis and Pattern Recognition, Vol. 2, (November 2007) pp. 952-955.

[8] Gonzalez, R. C. & Woods, R. E. (2001). Digital image processing, Addison-Wesley Longman Publish Co., Inc., Boston.

[9] Cheng, F.-H. & Chen, Y.-L. (2006). Real time multiple objects tracking and identification based on discrete wavelet transform, *Pattern Recognition*, Vol. 39, No. 3, (June 2006) pp. 1126-1139.

[10] Chih-Hsien Hsia and Jing-MingGuo., "Efficient modified directional lifting-based discrete wavelet transform for moving object detection", Signal Processing, article in press.

[11] Victoria Yanulevskaya, Jasper Uijlings, Jan-Mark Geusebroek, "Salient object detection: From pixels to segments", Image and Vision Computing Vol. 31, Pg. No. 31–42, 2013.

[12] P. S. Hiremath and Jagadeesh Pujari, "Content Based Image Retrieval using Color, Texture and Shape features,"Proceedings of the 15th International Conference on Advanced Computing and Communications, 2007 IEEE computer Society.

[13] Kart-Leong Lim, Hamed Kiani Galoogahi, "Shape Classification Using Local and Global Features" proceedings of Fourth Pacific-Rim Symposium on Image and Video Technology, 2010.

[14] A. Elgammal and L.S. Davis, "Probabilistic framework for segmenting people under occlusion," Proceedings of the IEEE Eighth International Conference on Computer Vision, 2001, pp. 145–152.

[15] A. Elgammal et al., "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," Proc.IEEE, vol. 7, 2002, pp. 1151–1163.

[16] Hanzi Wang and David Suter, "A consensus-based method for tracking," Pattern Recognition, vol. 40, 2007, pp. 1091-1105.

[17] I. Daubechies, Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math, 1988, pp. 909–996.

[18] M.T. Hanna and S.A. Mansoori, "The discrete time wavelet transform: its discrete time fourier transform and filter bank implementation," IEEE Trans. Circuits Systems II: Analog Digital Signal Process, vol. 2, 2001, pp. 180–183.

[19] A.S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transforms," IEEE Trans. Image Process, vol. 2, 1992, pp. 244–250.

[20] F. Marino, "Two fast architectures for the direct 2-D discrete wavelet transform," IEEE Trans. Signal Process, 2001, pp. 1248–1259.