ISSN 2250-3153

A Survey on Object detection and Object tracking in Videos

S. Rosi*, W. Thamba Meshach, J.Surya Prakash****

*Computer Science and Engineering, Prathyusha Institute of Technology and Management.
** Computer Science and Engineering, Prathyusha Institute of Technology and Management

Abstract- This paper presents survey on moving object detection and tracking methods is presented by classifying them into different categories and identify new trends. This survey shows moving object detection and tracking using different and efficient methodologies. Object detection and object tracking is used to track the object type(such as human, vehicles) and detect the movement of the object(such as moving, standing).This survey shows various methodologies for object detection and tracking such as background subtraction, background modeling, intensity range based background subtraction. The simulated result shows that used methodologies for effective object detection has better accuracy and with less processing time consumption rather than existing methods.

Index Terms- Background modeling, Background subtraction, Object detection, Object tracking.

I. Introduction

Video surveillance is the monitoring behavior, activities (or) other changing information usually of the people for the purpose of influencing, managing, directing or protecting them. Video surveillance is useful to government and law enforcement to maintain social control, recognize and monitor threats and prevent/investigate criminal activity.

Object detection and tracking in video is a challenging problem. It has application in numerous fields such as video compression, video surveillance, video indexing and retrieval etc. Object detection and tracking are two closely related processes. Object detection can be performed through various approaches such as region based segmentation, background subtraction, temporal differencing etc. The goals of object tracking are segmenting a region of interest from a video Scene and keep tracking of its motion, positioning, and occlusions. The object detection and object classification are preceding steps for tracking object in sequence of images. Object detection is performed to check existence of objects in video and to precisely locate that object. The detected object can be classified in various categories such as humans, vehicles, birds, floating clouds, swaying tree, and other moving object. Object tracking is performed using monitoring objects spatial and temporal changes during a video sequence includes its presence, position, shape, size etc. The direct LL mask band scheme (DLLBS) is used for moving object detection and tracking using low resolution image [1].The DLLBS method can effectively reduce this noise, with low computing cost, in both indoor and outdoor environments. Together with DLLBS and CPR (characteristic point recognition), the problems associated with occlusions are alleviated. With a combination of DLLBS and CPR can accurately track various types of obstructed movements. This DLLBS is suitable for real time video surveillance system applications such as object classification and the descriptive behavior of objects.

II. MOVING OBJECT DETECTION AND TRACKING TECHNIQUES

Video surveillance system for the purpose of security has been developed. Many researches try to develop intelligent video surveillance systems [2]. . The intelligent video surveillance can detect moving object in its initial stage and subsequently process the function such as object classification, object tracking, and object behavior descriptions. The accurate location of the moving object does not only provide a focus of attention for post processing but can also reduce the redundant computation for the incorrect motion of the moving objects. In this paper on application involving the surveillance of people or vehicles, and include full range of surveillance methods. Surveillance applications involving people or vehicles include the following.

A. Access control in special areas

When somebody is about to enter in some sensitive locations such as military based and important government units, the system could automatically obtain the visitor’s features, Such as height, facial appearance and walking gait from images taken in real time, and then decide whether the visitor can be cleared for entry.

B. Person-specific identification in certain scenes

Personal identification at a distance by a smart surveillance system can help the police to catch suspects. The system automatically recognize and judge whether or not the people in view are suspects. If yes, alarms are given immediately, such systems with face recognition have already been used at public sites, but the reliability is too low for police requirements.

C.Anomaly detection and alarming:

To analyze the behaviors of people and vehicles and determine whether these behaviors of people are normal or abnormal. For example, visual surveillance system set in parking lots and supermarkets could analyze abnormal behaviors indicative or theft. Normally, there are two ways of giving an alarm. One way is to automatically make a recorded public announcement whenever any abnormal behavior is detected; the other is to contact the police automatically.

D.Interactive surveillance using multiple cameras:

For social security, cooperative surveillance using multiple cameras could be used to ensure the security of an entire community, for example by tracking suspects over a wide area by using the cooperation of multiple cameras. For traffic management, interactive surveillance using multiple cameras can help the traffic police discover, track and catch vehicles involved in traffic offences.

III. BACKGROUND SUBTRACTION AND SHADOW DETECTION

The background subtraction algorithm is used to detect foreground objects and apply shadow detection algorithm to remove shadows [3]. The background subtraction and shadow detection is used in both indoor and outdoor environments and does not require color cameras. Background detection techniques may use gray scale or color images, shadow detection methods make use of chromaticity information. The car tracking system of Koller et al. [4] used an adaptive background model based on monochromatic images filtered with Gaussian and Gaussian derivative (vertical and horizontal) kernels. McKenna et al. [5] proposed a background model that combines pixel RGB and chromaticity values with local image gradients. In their W4 system, Haritaoglu and collaborators [6] used grayscale images to build a background model, representing each pixel by three values; its minimum intensity value, its maximum intensity value and the maximum intensity difference between consecutive frames observed during the training period. Elgammalel[7] proposed a nonparametric background model based on kernel based estimators, that can be used to both color and grayscale images. KaewTrakulPong and Bowden [8] used color images for background representation. In their method, each pixel in the scene is modeled by a mixture of Gaussian distributions (and different Gaussians are assumed to represent different colors). Cucchiara's group [9] temporal median filter in the RGB color space.

Shadow detection algorithms have also been widely explored by several authors, mostly based on invariant color features that are not significatively affected by illumination conditions. McKenna et al. [5] used each channel pixel and edge information of the normalized RGB color space (or rgb) to detect shadowed pixels. Elgammal [7] also used the normalized rgb color space, including a lightness measure to detect cast shadows. Prati's and Cucchiara'sgroups [9, 10] used the HSV color space, classifying as shadows those pixels having the approximately the same hue and saturation values compared to the background, but lower luminosity. KaewTrakulPong and Bowden [8] used a chromatic distortion measure and a brightness threshold in the *RGB* space to determine foreground pixels affected by shadows. Salvador et al. [11] adopted the $c1c2c3$ photometric invariant color model, and explored geometric features of shadows. A few authors [12, 13, 14, 15] was showed shadow detection in monochromatic video sequences, used in mind applications such as indoor video surveillance and conferencing. Basically, they detect the penumbra of the shadow, assuming that edge intensity within the penumbra is much smaller that edge intensity of actual moving objects. Clearly, such hypothesis does not used for video sequences containing low-contrast foreground objects (especially in outdoors applications). More about background subtraction and shadow removal can be found in [16, 17]. Also, there are several shadow detection algorithms to remove undesired segmentation of cast shadows in video sequences. However, in accordance with other authors [9, 6], we chose to use a background model based on median filtering, because it is effective and requires less computational cost than the Gaussian or other complex statistics. More specifically, we improved the background model proposed in [6], and included a novel shadow detection algorithm that is effective for both indoor and outdoor applications.

IV. BACKGROUND SUBTRACTION USING INTENSITY RANGE

Wren proposed a method to model the background independently at each pixel location which is based on computation of Gaussian probability density function (pdf) on the previous pixel values [18]. Stauffer and Grimson developed a complex procedure to accommodate permanent changes in the background scene [19]. Here each pixel is modeled separately by a mixture of three to five Gaussians. TheW4model presented by Haritaoglu is a simple and effective method [6]. It uses three values to represent each pixel in the background image namely, the minimum intensity, the maximum intensity, and the maximum intensity difference between consecutive frames of the training sequence. Jacques brought a small improvement to the W4 model together with the incorporation of a technique for shadow detection and removal [3]. McHugh proposed an adaptive thresholding technique by means of two statistical models [20].One of them is nonparametric background model and the other one is foreground model based on spatial information.

In Vibe, each pixel in the background can take values from its preceding frames in same location or its neighbor [21]. Then it compares this set to the current pixel value in order to determine whether that pixel belongs to the background, and adapts the model by choosing randomly which value to substitute from the background model. Kim and Kim introduced a novel background subtraction algorithm for dynamic texture scenes [22]. This scheme adopts a clustering-based feature, called fuzzy color histogram (FCH), which has an ability of greatly attenuating color variations generated by background motions while highlighting moving objects. Instead of segmenting a frame pixel-by-pixel, Reddy used an overlapping block-by-block approach for detection of foreground objects [23]. The scheme passes the texture information of each block through three cascading classifiers to classify them as background or foreground. The results are then integrated with a probabilistic voting scheme at pixel level for the final segmentation.

Generally, shadow removal algorithms are employed after object detection. Salvador developed a three step hypothesis based procedure to segment the shadows [11]. It assumes that shadow reduces the intensities followed by a complex hypothesis using the geometrical properties of shadows. Finally it confirms the validity of the previous assumption. Choi in their work of [24] have distinguished shadows from moving objects by cascading three estimators, which use the properties of chromaticity, brightness, and local intensity ratio. Liu[12] proposed a novel method for shadow removal using Markov random fields (MRF),where shadow model is constructed in a hierarchical manner. At the pixel level, Gaussian mixture model (GMM) is used, whereas at the global level statistical features of the shadow are utilized.

V. CONCLUSION

In this paper, the detection of moving object and tracking of moving object was discussed. From this study, background subtraction using complex wavelet transform domain outperforms with all the above mentioned techniques based on moving object detection provides more accurate in segmentation under various illuminations, less time consuming process, flexibility in background updating model, less sensitive to background noise. In Background subtraction using intensity range compares the initial background object compares with all foreground objects and find the moving object in videos, these does not provide more accurate result of object. Likewise many techniques are used to detect the moving object but it is not efficient. In complex wavelet transform domain the foreground object is compared with the entire background object using background updating model. The result of this domain based object detection provides more accurate result of the object. Accuracy can be estimated based on sensitivity, root mean square error, peak signal to noise ratio, correlation coefficient. Based on this study complex wavelet transform domain is recommended for moving object detection.

REFERENCES

[1] Chih hsien hsia and Jen-shiun chiang "Real time moving object detection and detection with direct LL-mask band scheme, vol no:43, July 2011,

[2] W.-M. Hu.T.-N. Tan, L.Wang and S. Maybanks, A survey on visual surveillance of object motion and behaviors, IEEE trans. On systems, Man and Cybernetics-Part C: Applications and Reviews, vol.34, no.3, pp.334-352, 2004.

[3] J.Jacques, C. jung, and S.musse, "Background subtraction and shadow detection in gray scale video sequences" in eighteenth brazillian symp. Computer graphics and image processing. Oct.2005,pp. 189-196.

[4] D. Koller, J.Weber, and J. Malik. Robust multiple car tracking with occlusion reasoning. In European Conference on Computer Vision, pages A:189–196, 1994.

[5] S. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H.Wechsler. Tracking groups of people. Computer Vision and Image Understanding, 80(1):42–56, October 2000.

[6] I. Haritaoglu, D. Harwood, and L. Davis. W4: Real time surveillance of people and their activities. IEEE Transactions on Pattern Analysis and Machine Intelligence,22(8):809–830, August 2000.

[7] A. Elgammal, R. Duraiswami, D. Harwood, and L. Davis.Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proceedings

Of the IEEE, 90(7):1151–1163, 2002.

[8] P. KaewTrakulPong and R. Bowden. A real time adaptive visual surveillance system for tracking low-resolution color targets in dynamically changing scenes. Image and Vision

Computing, 21(9):913–929, September 2003.

[9] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting moving objects, ghosts, and shadows in video streams. IEEE Transactions on Pattern Analysis and Machine Intelligence,

25(10):1337–1342, October 2003.

[10] A. Prati, I. Mikic, C. Grana, and M. M. Trivedi. Shadow detection algorithms for traffic flow analysis: a comparative study. In IEEE International Conference on Intelligent

Transportation Systems, pages 340–345, 2001.

[11] E. Salvador, A. Cavallaro, and T. Ebrahimi. Cast shadow segmentation using invariant color features. Computer Vision and Image Understanding, 95(2):238–259, August

2004.

[12] S.-Y. Chien, S.-Y.Ma, and L.-G. Chen. Efficient moving object segmentation algorithm using background registration technique. IEEE Transactions on Circuits and Systems for

Video Technology, 12(7):577–586, 2002.

[13] P. L. Rosin and T. Ellis. Image difference threshold strategies and shadow detection. In 6th British Machine VisionConf, Birmingham, pages 347–356, 1995.

[14] J. Stauder, R. Mech, and J. Ostermann. Detection of moving cast shadows for object segmentation. IEEE Transactions on Multimedia, 1(1):65–76, 1999.

[15] D. Xu, X. Li, Z. Liu, and Y. Yuan. Cast shadow detection in video segmentation. Pattern Recognition Letters, 26(1):5–26, 2005.

[16] J. J.Wang and S. Singh. Video analysis of human dynamics: a survey. Real-time imaging, 9(5):321–346, 2003.

[17] L.Wang,W. Hu, and T. Tan. Recent developments in human motion analysis. Pattern Recognition, 36:585–601, 2003.

[18] C.Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real time tracking of the human body," IEEE Trans. Patt. Anal. Mach. Intell., vol. 19, no. 7, pp. 780–785, Jul. 1997.

[19] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," in IEEE Computer. Soc. Conf. CVPR, 1999, pp.246–252.

[20] J.McHugh, J. Konrad,V. Saligrama, and P. Jodoin, "Foreground-adaptive background subtraction," IEEE Signal Process.Letters, vol. 16, no.5, pp. 390–393, May 2009.

[21] O. Barnich and M. Van Droogenbroeck, "Vibe: A universal background subtraction algorithm for video sequences," IEEE Trans. Image Process., vol. 20, no. 6, pp. 1709–1724, Jun. 2011.

[22] W. Kim and C. Kim, "Background subtraction for dynamic texture scenes using fuzzy color histograms," IEEE Signal Process. Lett., vol.19, no. 3, pp. 127–130, Mar. 2012.

[23] V. Reddy, C. Sanderson, and B. Lovell, "Improved foreground detection via block-based classifier cascade with probabilistic decision integration," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 1, pp.83–93, Jan. 2013.

[24] J. Choi, Y. Yoo, and J. Choi, "Adaptive shadow estimator for removing shadow of moving object," Comput. Vis. Image Understand., vol. 114,no. 9, pp. 1017–1029, 2010.

[25] Z. Liu, K. Huang, and T. Tan, "Cast shadow removal in a hierarchical manner using MRF," IEEE Trans. Circuits Syst. Video Technol., vol.22, no. 1, pp. 56–66, Jan. 2012.

AUTHORS

First Author –Rosi.S, Pg scholar, Prathyusha Institute Of Technology and Management, rosi.june22@gmail.com.
Second Author –Thamba Meshach.W, Associate Professor, Prathyusha Institute Of Technology and Management, thambameshac.cse@prathyusha.edu.in.
Third Author – Surya Prakash.J, Assistant Professor, Prathyusha Institute Of Technology and Management,suriya.engineer@gmail.com.

Correspondence Author – Rosi.S, rosi.june22@gmail.com, 8344414411