Tracking of Objects in Video by Using Median Filters, Segmentation & Discrete Wavelet Transform

G. Naga Vasantha[1] K. Prathyusha[2]

1. PG Student, Dept of ECE, ALIET, Vijayawada, vasu4nag29@gmail.com
2. PG Student, Dept of ECE, ALIET, Vijayawada, katakam.prathyusha190@gmail.com

Abstract: This paper deals with the tracking of objects in video by using segmentation and median filter based method. Intelligent visual surveillance system can be used many different methods for detection of moving targets. Advantages and disadvantages of two common algorithms frequently used in the moving target detection: background subtraction method and frame detection method are analyzed and compared in this paper. Then based on the background subtraction method, a moving target detection algorithm is proposed. The background image used to process the next frame image is generated through superposition of the current frame image and the current background image with a certain probability. This algorithm makes the objects which stay long time to be a part of the background after a certain period of time, but not be detected as a part of foreground. The experimental results show that this algorithm can detect moving targets more efficiently and precisely.

Keywords: Median Filters, Segmentation, Visual Surveillance, Background Subtraction and Frame Detection, Daubechies.

1. INTRODUCTION

Background subtraction is the most common method for identifying the moving objects in a video sequence. It is the first step in many vision applications like video surveillance [1], traffic monitoring. The main objective of segmentation technique is to decompose a video into background and moving foreground objects. Several problems may occur while segmenting the video sequences because of changing background, varying lighting operations, automatic operations etc.

Segmentation technique is divided into six groups. 1. Threshold based techniques. 2. Pixel classification based techniques. 3. Range image segmentation. 4. Color image segmentation. 5. Edge detection based segmentation. 6. Techniques based on fuzzy set theory. A background adaptation technique is categorized as 1.Non-recursive 2. Recursive. A non-recursive technique estimates the background based on a sliding-window approach. There are various video object segmentation techniques can be found in literature such as Running Gaussian average [13], Temporal Median Filter [10], and Mixture of Gaussians [8]. These methods are time consuming and space consuming. Running Gaussian Average is a simple method to describe a background, and it can get a real-time performing. But the shortcoming is it is not so accurate one. Many methods have been introduced for improving this method. Sumer et al. in [15] combined this method with edge information to improve the quality and reliability of the results. The methods discussed above for the segmentation of moving objects suffer from the problem of either slow speed or inaccurate segmentation of moving object due to non-removal of noise in consecutive frames. The other limitations include detection of only moving object and the presence of ghosts like appearances in segmented object. Cheng et al. [16] proposed a discrete wavelet transform (DWT) based method for tracking and identification of different objects. In this technique, inter-frame differencing method is used for segmentation of moving object in DWT domain. Approximate Median Filter method in Daubechies complex wavelet domain is proposed in this paper. The Daubechies complex wavelet transform have advantages of shift invariance and better directional selectivity as compared to DWT.

The performance of the proposed method is evaluated and compared with other standard spatial domain methods. The various performance measures used for comparison include RFAM (relative foreground area measure), MP (misclassification penalty), RPM (relative position based measure), NCC (normalized cross correlation) and the various methods are tested on standard Pets dataset. Finally, based on performance analysis it is observed that the proposed method in complex wavelet domain is performing better in comparison to other methods. Performance of the proposed method is found better in terms of visual performance and a number of quantitative measures such as RFAM (Relative foreground area measure), MP (Misclassification Penalty), RPM (Relative position based measure), NCC (Normalized cross correlation).

2. METHODOLOGY

Here in this paper the proposed method is an median filter based method in Daubechies complex wavelet domain. It differentiate the frames of obtained video which gives the changed pixel value from consecutive frames. First decompose two consecutive frames (I_{n-1}

and I_n) using complex wavelet domain and then apply approximate median filter based method to detect frame difference.

For every pixel location (i, j) the co-ordinate of frame.

$$(i, j)=WI_n (i, j)-WI_{n-1} (i, j) \qquad (1)$$

Where WI_n (i, j) & WI_{n-1} (i, j) are wavelet and coefficients of frame I_n (i, j) and I_{n-1} (i, j) respectively. The corresponding result may contain some noise by applying some thresholding the can be removed. In the presence of noise equation (1) can be expressed as

$$FD_n^{'} (i, j) = FD_N (i,j)- \lambda \qquad (2)$$

Where FD_n (i,j) is frame difference without noise. Where λ represent corresponding noise components. After noise removal, Sobel edge detection operation is applied on $FD_n^{'}$ (i, j) to detect the strong edges of significant difference pixels in all sub-bands as follows:

$$DE_n (i,j) = sobel (FD_n^{'} (i,j)) \qquad (3)$$

After finding edge map DE_n (i,j) in wavelet domain, inverse wavelet transform is applied to get moving object segmentation in spatial domain i.e. E_n. The obtained segmented object may include a number of disconnected edges due to non-ideal segmentation of moving object edges. Therefore, some morphological operation is needed for post-processing of object edge map to generate connected edges. Here a binary closing morphological operation described in [18] is used.

After applying the morphological operator M (E_n) is obtained which is the segmented moving object, and finally temporal updating of the background model is needed in order to adapt the changes in background and in lighting are:

$$\text{If } (I_n (i,j) > I_{n-1} (i,j)) \qquad (4)$$
$$I_{n-1} (i,j) = I_{n-1} (i,j) + 1 \qquad (5)$$
$$\text{Else } I_{n-1} (i,j) = I_{n-1} (i,j) – 1 \qquad (6)$$

Here, I_n (i,j) is the value of $(i,j)^{th}$ pixel of n^{th} frame and I_{n-1} (i,j) is the value of $(i,j)^{th}$ pixel of n^{th} frame.

Performance Evaluation: It can be observed from the results that none of the previous proposed segmentation algorithms give accurate segmentation result as compared to ground truth frames. In this paper, the performance of the proposed method have been compared quantitatively with other state-of-the-art methods viz. method used by McFarlane and Schofield[4],method used by Kim *et al.*[4], method used by Oliver et al.[6], method used by Liu *et al.* [7], method used by Stauffer and Grimson's [8], method used by Zivkovic [9], method used by Lo and Velastin [10],method used by Cucchiara *et al.*[11], method used by Bradski [12], and method used by Wren *et al.[*13]. For quantitative comparisons, four different performance metrics have been considered such as RFAM (Relative foreground area measure), MP (Misclassification Penalty), RPM (Relative position based measure), and NCC (Normalized cross correlation) defined as follows:

A. RELATIVE FOREGROUND AREA MEASURE (RFAM)

Good image segmentation should yield accurate measurements on the object properties such as area and shape [19]. Relative foreground area measure is calculated between ground-truth frame and segmented frame. The value of RFAM will be in the range [0, 1]. If the RFAM value is 1 then it shows perfect segmentation.

B. MISCLASSIFICATION PENALTY (MP)

The misclassified pixels estimated in the segmentation results which are farther from the actual object boundary (ground-truth image) are penalized more than the misclassified pixels which are closer to the actual object boundary. Misclassification penalty value lies in the range [0, 1]. Zero value of misclassification penalty means perfect segmentation and performance of segmentation methods degrades as the value of misclassification penalty becomes higher.

C. RELATIVE POSITION BASED MEASURE(RPM)

Relative position measure is defined as the centroid shift between ground-truth and segmented object mask [19]. It is normalized by parameter of the ground-truth object. Value of Relative Position based Measure (RPM) will be 1 for a perfect segmentation and decreases with the centroid shift increases.

D .NORMALIZED CROSS CORRELATION (NCC)

Normalized Cross Correlation can be used as a measure for calculating the degree of similarity between two images. NCC value lies in the range [0, 1].Higher value of NCC means good normalized cross correlation while lower value indicates poor normalized cross correlation.

Table1 show the values of Relative Foreground area measure (RFAM), Misclassification Penalty (MP), relative position measure, Normalized Cross Correlation for the proposed method.

The proposed method has higher values of relative foreground area, relative position measure, Normalized Cross Correlation, and lower value of misclassification penalty as compared to other methods [4-13] which are favorable to the proposed method. Since the value of one quantitative performance measure is not sufficient to take any conclusion. To arrive at any conclusion, it is necessary to see values of different quantitative measures simultaneously. Therefore, after observing values of four quantitative measures, it can be concluded that the proposed method using Daubechies complex wavelet transform give better results as compared to other methods in consideration.

4. RESULTS

The proposed method for segmentation of moving object as described in section 3 has been applied on people video sequences. People video sequences are performed in outdoor environment, in this video multiple objects are present (Human beings and cars) and partially and fully occlusions also occurs between human beings. For the segmentation of the video object by various methods, the numbers of frames taken in to consideration at a time include 125, 150, 175, 200, 225 and 250 frames. The obtained results have been compared with other standard methods as proposed in [4-13] in terms of various performance measures.

The proposed technique also handles the problem of ghosts and shadows unlike segmented frame obtained by other methods such as proposed in [4,8,9,10,11,13] which fail to handle these issues.

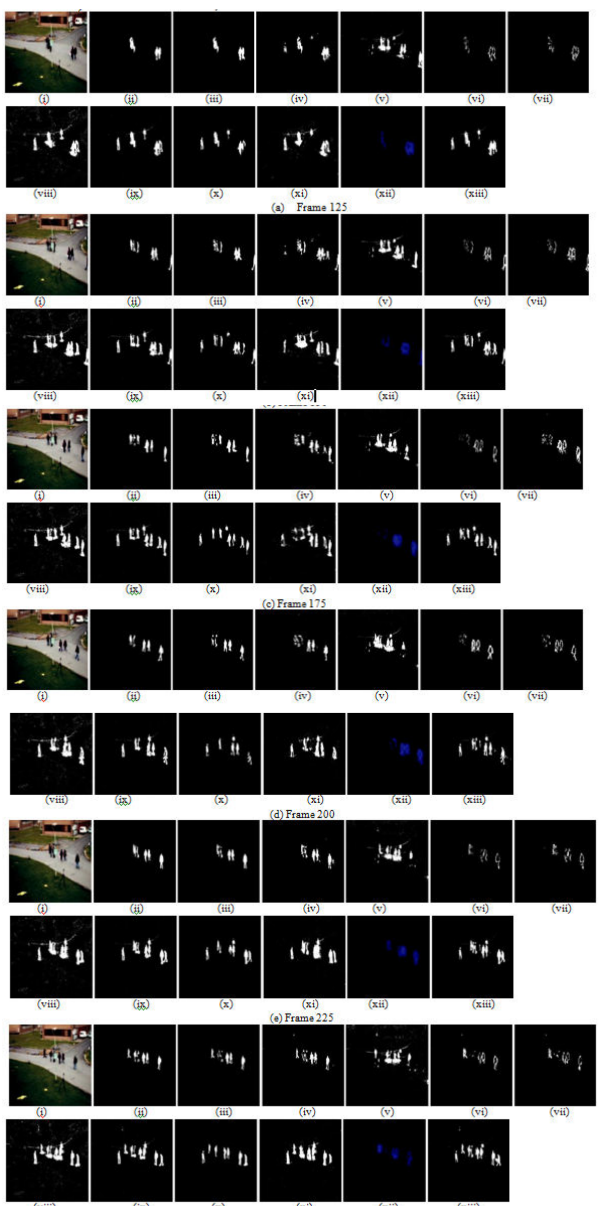

Fig. 1. Segmentation results for People video sequence corresponding to (a) Frame 125, (b) frame 150, (c) frame 175, (d) frame 200, (e) frame 225 and (f) frame 250 (i) original frame, (ii) ground truth segmented frame, and the segmented frame obtained by various methods such as: (iii) the proposed method, (iv)McFarlane and Schofield[4], (v) Kim *et al.*[5], (vi)Oliver *et al.*[6], (vii) Liu *et al.* [7], (viii) Stauffer and Grimson's[8], (ix)Zivkovic[9], (x)Lo and Velastin[10], (xi) Cucchiara*et al.*[11], (xii)Bradski[12], and (xiii)Wren *et al.*[13].

5. CONCLUSION

In the present work, a robust and efficiently computed method for segmentation of moving objects using approximate median filter based method in Daubechies complex wavelet transform domain have been proposed. The performance of the proposed method have been evaluated and compared with other standard methods in consideration [5-13] in terms of various performance metrics. From the obtained results and their qualitative and quantitative analysis, it can be concluded that the proposed method is performing better in comparison to other methods as well as it also capable of alleviating the problems associated with other spatial domain methods such as ghosts, clutters, noises etc.

REFERENCES:

[1] W. Hu, T. Tan, "A Survey on Visual Surveillance of Object Motion and Behaviors", IEEE Trans. Systems, Man, and Cybernetics, Vol. 34, No. 3, pp. 334-352,2006.

[2] Y. Zhang "An Overview of Image and Video Segmentation in the Last 40 Years," Book Vol. 1, pp. 1-16, 2006, IGI Global, USA.

[3] S.C. Cheung and C. Kamath, "Robust techniques for background subtraction in urban traffic video," Video Communications and Image Processing, SPIE Electronic Imaging, UCRL Conf. San Jose, vol.200706, Jan 2004.

[4] N. McFarlane, C. Schofield, "Segmentation and tracking of piglets in images", Machine Vision Application, Vol. 8, No. 3, pp. 187-193, 1995

[5] K. Kim, T. H. Chalidabhongse, D. Harwood, L. Davis, "Real Time Foreground Background Segmentation using Codebook Model", Real Time Imaging, Vol. 11, No. 3, June 2005, pp. 172-185.

[6]N. Oliver, B. Rosario, and A. Pentland, "A bayesian computer vision system for modeling human interactions", IEEE PAMI, Vol. 22, pp. 831–843, 2000

[7]M.-Y. Liu , Q.-H. Dai , X.-D. Liu, G.-H.Er, "Automatic extraction of initial moving object based on advanced feature and video analysis", proceeding of Visual Communications and Image Processing, Vol. 5960, pp. 160-168, 2005.

[8] C. Stauffer , W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," In Proc. Computer Vision and Pattern Recognition conference,pp. 246-252, 1999.

[9]Z. Zivkovic and F. van der Heijden," Efficient adaptive density estimation per image pixel for the task of background subtraction", Pattern Recognition Letters, Vol. 27 No.7, pp. 773–780, 2006.

[10] B.P.L. Lo and S.A. Velastin, "Automatic congestion detection system for underground platforms," Proc.ISIMP 2001, pp. 158-161, 2001.

[11]C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "P nder: real-time tracking of the human body", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19 No. 7, pp. 780–785, 1997.

[12]F-H. Cheng, Y-L. Chen, "Real time multiple

objects tracking and identification based on discrete wavelet transform", Elsevier journal of Pattern Recognition, Vol. 39, No. 6, pp. 1126–1139, 2006.
[13]. A. Khare, U. S. Tiwary, W. Pedrycz, M. Jeon, "Multilevel adaptive thresholding and shrinkage technique for denoising using Daubechies complex wavelet transform", The Imaging Science Journal, Vol. 58, No. 6, Dec. 2010, pp. 340-358.
[14] S. Sridhar, "Digital image processing', Oxford Publication, 3rd edition, 2008.
[15] Y. Gao-bo, Z. Zhao-yang, "Objective Performance Evaluation of video segmentation algorithms with Ground-Truth", Journal of Shanghai University (English Edition), Vol. 8, No. 1, pp. 70-74.