Volume 5, Issue 5, May 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Analysis of Effective Moving Object Detection Using Background Subtraction in Different Types of Wavelet

Pritee Gupta[1], Dr. Yashpal Singh[2], Dr. Manoj Kumar Gupta[3]
[1]Ph.D Student, Bhagwant University, Ajmer & Assistant Professor, CSE Deptt,,
I.T.S Engg. College, Gr. Noida (U.P), India,
[2]Associate Professor & Head, CSE Deptt , Bundelkhand Institute of Engg. &Technology, Jhansi, India,
[3]Associate Professor, CSE Deptt , School of Engineering and Technology, Sharda University , Gr. Noida (U.P), India

Abstract— Visual surveillance has been a very active research topic in the last few years due to its growing importance in security, law enforcement, and military applications. This work presents moving object detection based on background subtraction under different types of wavelet transform domain for video surveillance system. The object detection in frequency domain will be approached to segment objects from foreground with absence of background noise. Initially it starts with background initialization by choosing start frame or taking initial few frames with approximate median method. Then, wavelet transform is applied to both current and initialized background frame generates subbands of low and high frequencies. Frame differencing will be done in this subbands followed by edge map creation and image reconstruction. In order to remove some unwanted pixels, morphological erosion and dilation operation is performed for object edge smoothness. The proposed approach has some advantages of background noise insensitiveness and invariant to varying illumination or lighting conditions. After the object detection, performance of method will be measured (between frame ground truth and obtained result) through metrics such as sensitivity, accuracy, correlation and peak signal to noise ratio. This object detection also helps to track detected object using connected component analysis. The simulated result shows that used methodologies for effective object detection has better accuracy and with less processing time consumption rather than existing methods. The experimental results show the most suitable method that runs quickly, accurately and fits for the real-time detection.

Keywords Input video, Frame separation, wavelet transform, Subband differencing, Morphological Filtering, Motion detection, Performance measurement

I. INTRODUCTION

Surveillance is the monitoring of behavior. Systems surveillance is the process of monitoring the behavior of people, objects or processes within systems for conformity to expected or desired norms in trusted systems for security or social control. The word surveillance is commonly used to describe observation from a distance by means of electronic equipment or other technological means.

Fig. 1 Example of CCTV camera

At a basic level, computers are a surveillance target because large amounts of personal information are stored on them. Anyone who can access or remove a computer can retrieve information. If someone is able to install software on a computer system, they can turn the computer into a surveillance device. CCTV is a collection of video cameras used for video surveillance. CCTV is generally used in areas where there is an increased need for security, such as banks, airports and town centers. A basic CCTV system comprises of the Camera, lens and power supply. Recording device, VCR or a digital video recorder and monitor. Closed-circuit television (CCTV) is the use of video cameras to transmit a signal to a specific place, on a limited set of monitors. The main tasks in visual surveillance systems include motion detection, object classification, tracking. Our focus here is on the detection phase of a general visual surveillance system using static cameras. The usual approach for 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 background image is not fixed but must adapt to: Illumination changes, sudden (such as clouds), Motion changes, camera oscillations, high-frequencies background objects (such as tree branches, sea waves, and similar) Changes in the background geometry.

II. L LITERATURE SURVEY

Several problems arise while segmenting the video sequences because of changing background, clutter, occlusion, varying lighting conditions, automatic operation, adverse weather conditions such as fog, rain, snow, camera angle, and real time processing requirements etc. [1-7]. Zhang [4] divided the segmentation techniques into six groups: - Threshold based techniques, Pixel classification based techniques, Range image segmentation, Colour image segmentation, Edge detection based segmentation and techniques based on fuzzy set theory. According to Cheung and Kamath [6], background adaptation techniques could also be categorized as: 1) no recursive and 2) recursive. A non-recursive technique estimates the background based on a sliding-window approach. Various interesting video object segmentation techniques can be found in literature [1-5] such as Running Gaussian Average , Temporal Median Filter , and Mixture of Gaussians. These methods are either too time consuming (like GMM with online EM algorithm) or too space consuming (like Temporal Median Filter proposed in .All the methods discussed as 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.* [11] proposed a discrete wavelet transform (DWT) based method for approach, inter-frame differencing method is used for segmentation of moving object in DWT domain. DWT based methods are shift-sensitive. Any shift sensitive method will not give good results for video applications because in video applications, objects are present in shifted form. Motivated by these facts, a new method for video segmentation using Discrete wavelet domain is proposed in this paper. The Discrete wavelet transform have advantages of shift invariance and better directional selectivity as compared to DWT. The performance of the proposed method is compared with other standard methods available in literature such as Frame Difference, Background subtraction, SOBS [12-15].

A. Frame Difference

This method is through the difference between two consecutive images to determine the presence of moving objects.

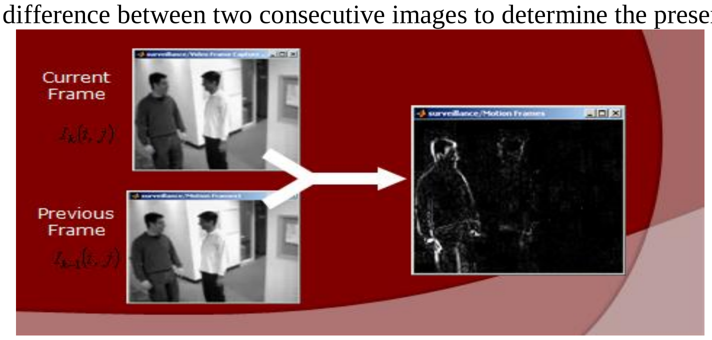


Fig. 2 Example of frame difference method

The Frame difference is arguably the simplest form of background subtraction. Frame differencing, also known as temporal difference, uses the video frame at time t-1 as the background model for the frame at time t. This technique is sensitive to noise and variations in illumination, and does not consider local consistency properties of the change mask. This method also fails to segment the non-background objects if they stop moving. Since it uses only a single previous frame, frame differencing may not be able to identify the interior pixels of a large, uniformly-colored moving object. This is commonly known as the aperture problem. It has strong adaptability, but it is generally difficult to obtain a complete outline of moving object, liable to appear the empty phenomenon, as a result the detection of moving object is not accurate.

B. Background subtraction

The basic scheme of background subtraction is to subtract the image from a reference image that models the background scene. Background modelling constructs a reference image representing the background. Threshold selection determines appropriate threshold values used in the subtraction operation to obtain a desired detection rate. Subtraction operation or pixel classification classifies the type of a given pixel, i.e., the pixel is the part of background (including ordinary background and shaded background), or it is a moving object.

After background image B(X, Y) is obtained, subtract the background image B(X, Y) from the current frame FK(X, Y). If the pixel difference is greater than the set threshold T, then determines that the pixels appear in the moving object, otherwise, as the background pixels . The moving object can be detected after threshold operation.

Its expression is as follows:

$$\mathrm{DK(X,Y)} = \begin{cases} \text{if } (|\mathrm{FK(X,Y)} - \mathrm{B(X,Y)}\,| > T) \\ \qquad\qquad 0 \text{ others} \end{cases} \qquad (1)$$

Background subtraction method is very sensitive to the changes in the external environment. The methods with a background model based on a single scalar value can guarantee adaptation to slow illumination changes, but cannot cope with multi-valued background distributions. As such, they will be prone to errors whenever those situations arise. Processing time required to detect the object using this technique is low but accuracy may not be good enough.

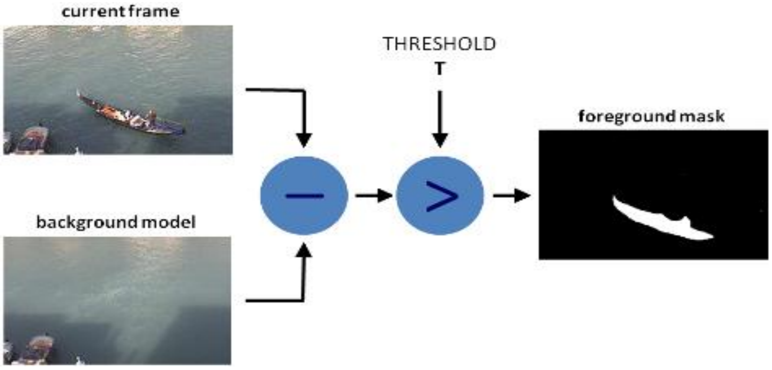


Fig. 3 Example of Background subtraction method

C. SOBS (Self organizing background subtraction method)

This is biologically inspired problem-solving method based on visual attention mechanisms. This 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 many background variations, i.e., background motion cycles, seen as trajectories of pixels in time. Based on the learnt background model through a map of motion and stationary patterns, this can detect motion and selectively update the background model. Each node computes a function of the weighted linear combination of incoming inputs, where weights resemble the neural network learning. Doing so, each node could be represented by a weight vector, obtained collecting the weights related to incoming links. In the following, the set of weight vectors will be called a *model*. An incoming pattern is mapped to the node whose model is "most similar" to the pattern, and weight vectors in a neighbourhood of such node are updated. Therefore, the network behaves as a competitive neural network that implements a winner take- all function with an associated mechanism, that modifies the local synaptic plasticity of the neurons, allowing learning to be restricted spatially to the local neighbourhood of the most active neurons. For each colour pixel, consider a neuronal map consisting of n*n weight vectors. Each incoming sample is mapped to the weight vector that is closest according to a suitable distance measure, and the weight vectors in its neighbourhood are updated. The whole set of weight vectors acts as a background model, that is used for background subtraction in order to identify moving pixels [1-10].

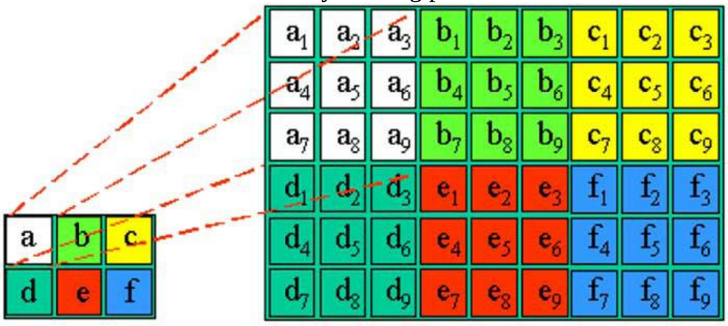


Fig. 4 (Left) Simple image and the (right) neuronal map structure

III. PROPOSED METHOD

The proposed method is Background subtraction based effective moving object detection using Wavelet transform.

The process algorithm is described as follow:

1. Input video
2. Frame Separation
3. Image Sequence
4. Separation of Image Sequence in Current Frame Image and Background Frame Image
5. Apply wavelet transform for both background and current image
6. Subband Differencing
7. Soft threshold
8. Inverse wavelet transform
9. Threshold for foreground detection
10. Noise removal
11. Morphological filtering
12. Detection of Moving Object
13. Performance measurement

(Similarity/MSE/PSNR/Correlation)

A. Proposed Method Block diagram

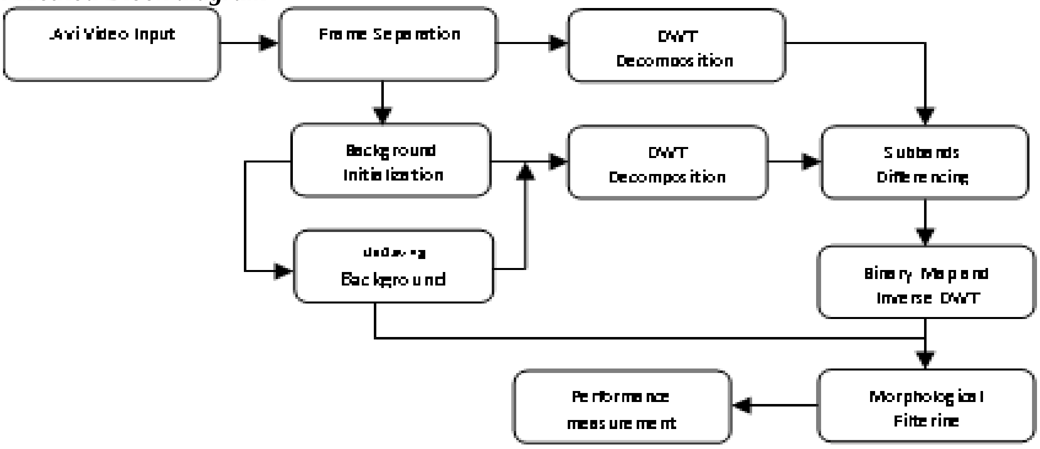


Fig.5 block diagram of proposed method

The proposed method is an approximate median filter based method in Discrete wavelet domain. It uses frame differencing for obtaining video object planes which gives the changed pixel value from consecutive frames. First, we decompose two consecutive frames (I_{n-1} and I_n) using Discrete wavelet transform and then apply approximate median filter based method to detect frame difference. For every pixel location (i, j) the co-ordinate of frame

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

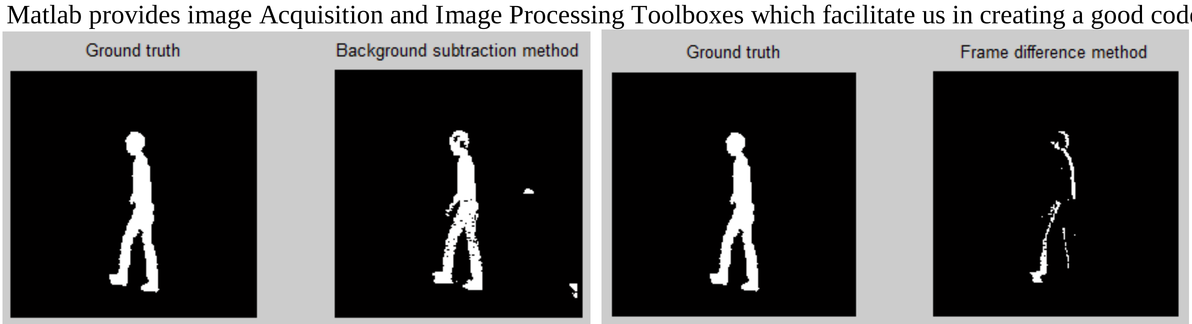
where $WI_n(i,j)$ and $WI_{n-1}(i,j)$ are wavelet coefficients of frame $I_n(i,j)$ and $I_{n-1}(i,j)$ respectively. Obtained result may have some noise. Applying soft Thresholding to remove noise. In presence of noise, equation is expressed as:

$$FD_n'(i, j) = FD_n(i, j) - \lambda \qquad (2)$$

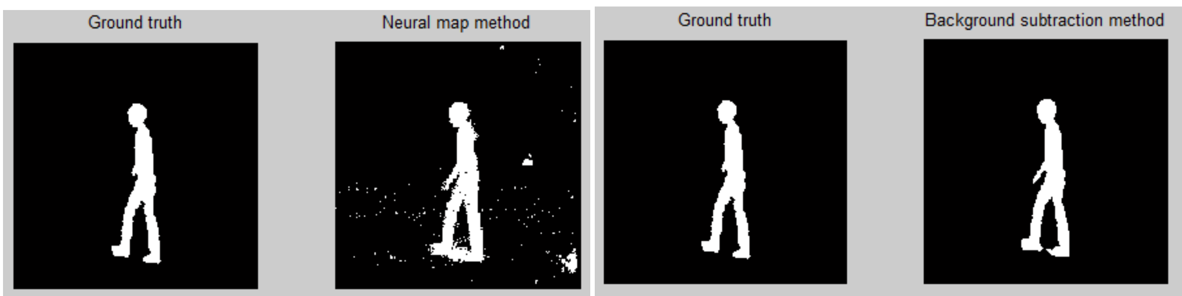
where $FD_n'(i,j)$ is frame difference without noise, λ represent corresponding noise components. For de-noising, soft thresholding technique in wavelet domain is used for estimation of frame difference $FD_n'(i,j)$**.** Inverse wavelet transform is applied to get moving object segmentation in spatial 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 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 conditions.

IV. EXPERIMENTS AND RESULTS

In this work the aim is to build such a surveillance system, which will detect motion even if the moving background, gradual illumination variations and camouflage and shadow into the background, thus achieves robust detection for different types of videos taken with stationary cameras. To fulfill this aim , a strong computing software called Matlab is used. Matlab provides image Acquisition and Image Processing Toolboxes which facilitate us in creating a good code.



(a) Result of Background subtraction method (b) Result of Frame difference method



(c) Result of Neural map method (d) Result of Proposed method

Fig.6 Results for Input video sequence corresponding to Frame no 29

Experimental results for moving object detection using the proposed methods have been produced for input video, that represent typical situations critical for video surveillance systems, and present qualitative results obtained with the proposed method and other three methods also. Efficiency in terms of accuracy is better than other methods.

V. PERFORMANCE EVALUATION

It can be observed from the results that none of the previously proposed segmentation algorithms give accurate segmentation result as compared to ground truth frames. In this paper, the performance of the proposed method has been compared quantitatively with other state-of-the-art methods. The proposed system is experimented with different settings of adjustable parameters which can be used for performance evaluation. Performance of the proposed method is found better in terms of visual performance and a number of quantitative measures viz. MSE, PSNR, correlation coefficient and similarity This section outlines the set of performance evaluation metrics that have implemented in order to quantitatively analyze the performance of object detection methods. For measuring accuracy we adopted different metrics,

A. Mean Square Error (MSE) and Peak Signal to Noise ratio(PSNR)

The Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR) are the two error metrics used to compare image compression quality. The MSE represents the cumulative squared error between the compressed and the original image, whereas PSNR represents a measure of the peak error. The lower the value of MSE, the lower the error. To compute the PSNR, the block first calculates the mean-squared error using the following equation:

$$MSE = \frac{\sum_{M,N} [I_1(m,n) - I_2(m,n)]^2}{M * N}$$

In this equation, M and N are the number of rows and columns in the input images, respectively. Now the block computes the PSNR using the following equation:

$$PSNR = 10\log_{10}\left(\frac{R^2}{MSE}\right)$$

The PSNR between two images having 8 bits per pixel or sample in terms of decibels (dBs) is Generally when PSNR is 40 dB or greater, then the original and the reconstructed images are virtually indistinguishable by human observers . In this equation, *R* is the maximum fluctuation in the input image data type. For example, if the input image has a double-precision floating-point data type, then *R* is 1. If it has an 8-bit unsigned integer data type, *R* is 255, etc.

B. Correlation coefficient:

This gives statistical relationship between two or more random variable or observed data values. This computes the correlation coefficient between A and B, where A and B are matrices or vectors of the same size.

C. Similarity:

We considered the pixel-based *Similarity* measure as **Similarity** = $\frac{tp}{tp+fn+fp}$
Greatest value of similarity shows accurate detection of moving object.

Table 1. Values of MSE, PSNR, correlation coefficient and Similarity for frame no 29

Input	inp1.avi
GT :	29.bmp

S.No	Parameters	Methods				
		Background subtraction	Frame difference	neural map	DWT(db5) +BS	DWT(haar) +BS
1	MSE	0.0129	0.0487	0.0245	0.0098	**0.0092**
2	PSNR	67.0219	61.2563	64.2436	68.2	**68.4709**
3	correlation coefficient	0.8676	0.2783	0.8041	0.9032	**0.9079**
4	Similarity	0.9871	0.9513	0.9755	0.9902	**0.9908**

VI. CONCLUSION

In video surveillance, there are many interference factors such as target changes, complex scenes, and target deformation in the moving object tracking. In this paper moving objects detection using discrete wavelet transform domain have been proposed. The performance of the proposed method have been evaluated and compared with other standard methods in consideration 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. used . Future work will address on techniques to get better results to improve the human detection methods and occlusion handling in surveillance applications.

REFERENCES

[1] 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.

[2] L 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

[3] Tao Jianguo, Yu Changhong, " Real-Time Detection and Tracking of Moving Object,"Intelligent InformationTechnology Application, 2008. UTA '08. Second International Symposium on Volume 2, 20-22 Dec. 2008 Page(s):860 – 863

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

[5] M. Piccardi, "Background subtraction techniques: a review," in *Proc. IEEE Int. Conf. Systems, Man, Cybernetics*, 2004, pp. 3099–3104.

[6] 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.

[7] L. Maddalena and A. Petrosino, "A self-organizing approach to detection of moving patterns for real-time applications," in Proc. 2nd Int. Symp. Brain, Vision, and Artificial Intelligence, 2007, pp. 181–190, Lecture Notes Comput. Sci. 4729

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

[9] 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.

[10] 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,

[11] 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.

[12] A. Khare, U. S. Tiwary, W. Pedrycz, M. Jeon, "Multilevel adaptive thresholding and shrinkage technique for denoising using Daubechies Discrete wavelettransform", The Imaging Science Journal, Vol. 58, No. 6, Dec. 2010, pp. 340-358.

[13] Mahesh C. Pawaskar1, N. S.Narkhede2,Saurabh S. Athalye1,"Detection Of Moving Object Based On Background Subtraction," International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 3, May-June 2014

[14] Alok Singh Kushwaha, Rajeev Srivastava,"Complex wavelet based Moving Object segmentation using approximate median filter based method for video surveillance,"2014 IEEE International advance computing conference pp. 973-978, 2014

[15] Pritee Gupta, Dr. Yashpal Singh, "Implementation of Dynamic Threshold Method for Human Motion Detection in Video surveillance application," International Journal of Computers and Technology,2014 Vol. 13, No. 8

