International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249 – 8958, Volume-4 Issue-6, August 2015

Lifting Based DWT for Object Tracking Using Variance Method

Resmi H. B, Deepambika V. A, M. Abdul Rahman

***Abstract* — Fast and accurate object tracking is very important for real time applications like video surveillance, traffic monitoring etc. In most of the conventional object tracking methods environmental changes, memory requirement and computation speed are the major constraints. This paper proposes an efficient object tracking method to compensate for all these challenges. Here a Lifting based Discrete Wavelet Transform (LDWT) has been used in order to compensate for fake motions and low memory requirement. Lifting based 9/7 Discrete Wavelet Transform is proposed to reduce the computational cost and preserve fine object boundaries. For fast object tracking variance method is adopted where maximum nonzero pixel value is considered. The experimental results show that the proposed method yields better result on the basis of computational time, memory requirement, speed of operation and environmental changes than the conventional DWT based approach.**

***Index Terms*— Object detection, Object Tracking, DWT, LDWT, Frame Differencing, Variance Method**

I. INTRODUCTION

The intelligent video surveillance system has become an active area of research topic in computer vision that tries to detect, recognize and track object from input video sequences. Successful motion detection in real time applications is a difficult task because of so many challenges which include fake motion, Gaussian noise in the background etc. that may lead to incorrect motion detection. The conventional approaches for motion detection include background substraction[1], temporal differencing[2] and optical flow method[3].In background subtraction approach the moving regions are detected by subtracting the current frame from background frame. It gives satisfactory motion mask data, but the problem with this approach is the dynamic scene changes due to lightening and other radiometric variations; therefore we need to frequently update the background frame. In temporal differencing approach the moving regions are extracted from consecutive frames of the video sequence. Even though it is suitable for dynamic environment, the problem with this approach is the incomplete motion extraction. In optical flow method, flow vector characteristics of the moving object over a time is used to detect the moving region. This method also gives satisfactory result but the computational complexity is very high. The most common problems with all of the above mentioned approaches are the sensitivity to noises, fake motion such as motion of tree leaves and memory requirement for the storage purpose and the speed.

Revised Version Manuscript Received on July 31, 2015.

Resmi H. B, Department of Electronics and Communication, LBS Institute of Science and Technology for Women, Trivandrum, India.

Deepambika V. A, Department of Electronics and Communication, LBS Institute of Science and Technology for Women, Trivandrum, India.

Dr. M. Abdul Rahman, Professor, Vice Chancellor of Kerala Technical University, Trivandrum, India.

DWT and Low resolution techniques like average filtering are used to reduce fake motions and computing cost. But the conventional DWT scheme has the disadvantages of complicated calculations. But using discrete wavelet transform reduces the size of the image and also removes the noise parts the input image. So DWT adopted as the preprocessing step of the object tracking.2D-DWT decomposes the image into four sub-band images.LL2 band is adopted for the further processing of tracking because fake motions are suppressed in the 2nd level low pass band. 2 –D Haar based DWT to leads to blur the object shape. To overcome this problem lifting scheme based discrete wavelet transform is used[12]. Low resolution images are suitable for object detection and tracking. It preserves the quality of low resolution images better than that of the other low resolution methods[4].Reducing the computation time is a challenging task in real time applications. The variance method is used for fast object detection and tracking. The rest of the paper is arranged as follows: the convolution based DWT approach is described in section 2. The proposed algorithm is described in section 3.In section 4 experimental result of the proposed algorithm is compared with conventional DWT based variance method. Section 5 provides the conclusion.

II. CONVOLUTION BASED DWT

Discrete wavelet transform is used in various fields such as signal processing, image processing, computer vision, video compression, biochemistry medicine etc. It gives a flexible multi-resolution image and decomposes the image into various sub-band images. During decomposition process, low pass filer L is represented as scaling functions and high pass filter H represented as wavelet filter. For example if the filter length is four the corresponding transfer functions is denoted as

$$H(z) = h_0+h_1z^{-1}+h_2z^{-2}+h_3z^{-3} \quad (1)$$

$$G(z) = g_0+g_1z^{-1}+g_2z^{-2}+g_3z^{-3} \quad (2)$$

The result is followed by down sampling operation. Here the input video sequence is decomposed into four components. Each sub-band image has its own feature. First input image is decomposed into high pass and low pass frequency images; again the low pass band image is decomposed into two sub-bands such as LL, LH, HL, and HH as shown in the Fig 1.The noise is suppressed in high frequency components. So the further processing is done on low frequency (LL2) band. But the multilevel decomposition images by using Haar-DWT the LL2 band become more blur. For detecting object more accurately lifting scheme based DWT is used in pre-processing steps.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

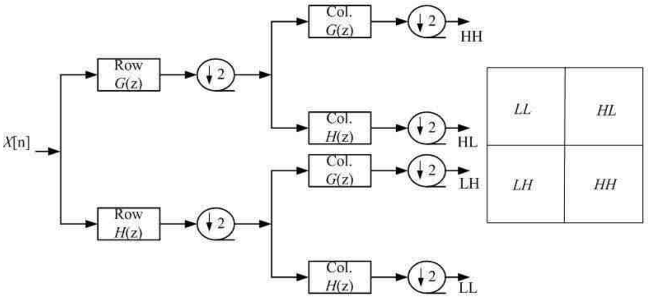

Fig. 1. Dimensional-DWT Decomposition

III. PROPOSED METHOD

For video compression and edge detection lifting scheme based discrete wavelet transform is used. For extracting moving regions between the consecutive frames, frame differencing method is adopted. Finally for fast object detection and tracking variance method is used. The detailed flow diagram of proposed method is shown in Fig.4

A. *9/7 Lifting –based discrete wavelet transform*

The lifting scheme based DWT[12] is used for image compression because lifting has many advantages over the traditional convolution based approach. As faster implementation and as the number of computation is half of the convolution based discrete wavelet transform, this property very useful for real time applications. As the lifting scheme based DWT (LDWT) does in place calculation, no auxiliary memory is needed for the operation. It also avoids the finite precision or rounding problem. The Euclidean algorithm is adopted to factorize the poly-phase matrix of a discrete wavelet transform filter, into a sequence of alternating upper and lower triangular matrices, and diagonal matrix. In equation 3, h (z) and g (z) denote the low pass and high pass filters respectively, which can be further decomposed into even and odd components to poly phase matrix P (z) as expressed by the following equations:

$$P(Z) = \begin{bmatrix} h_e(z) & g_e(z) \\ h_o(z) & g_o(z) \end{bmatrix} \tag{3}$$

The greatest common divisors of the even and odd parts of the original filters, can be obtained by recurrent operation of the Euclidean algorithm. As *h*(z) and *g*(z) form a complementary filter pair p(z) can be factorized into Equation 4.

$$\mathrm{P(z)}= \prod_{i=1}^{m} \begin{pmatrix} 1 & s_i(z) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_i(z) & 1 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & 1/k \end{pmatrix} \tag{4}$$

Where i=1 to m, m is the even degree of the polynomial $s_i(z)$ and $t_i(z)$ are Laurent polynomials, They correspond to the prediction and update steps respectively, and *k* is a nonzero constant. As shown in Fig. 3, a lifting-based scheme is divided into four stages: Split phase: Decomposition of original image into two disjoint components, the variable $X_{e(n)}$ and $X_{o(n)}$ which denote the set of even samples and the set of odd samples respectively. As the split phase does not decorrelate the data but only subsamples the signal into even and odd samples, it is also referred to as lazy wavelet transform.

$$X_{o(n)} = X_{(2n+1)} \tag{5}$$

$$X_{e(n)} = X_{(2n)} \tag{6}$$

Predict phase: To obtain the wavelet coefficients the predicting operator P is applied to the subset $X_{o(n)}$.This is shown in Equation 7.

$$d[n]= X_{o(n)}\, \mathrm{P}\times (X_{e(n)}) \tag{7}$$

Update phase: The combination of $X_{e(n)}$ and *d*[n] yields the scaling coefficients *s*[n] after an update operator. This is shown in Equation 8.

$$s[n]= X_{e(n)} + \mathrm{U} \times (d[n]). \tag{8}$$

Scaling: In the final step, the normalization factor is applied on *s*[n] and *d*[n] to obtain the wavelet coefficients. Equations 8 and 9 show high pass and low pass coefficients respectively.

$$H_n = {}^{1}/_{K} \times \mathrm{d[n]} \tag{9}$$

$$L_n = {}^{1}/_{K} \times \mathrm{s[n]} \tag{10}$$

By vertical and horizontal decomposition of 1-D coefficients, 2-D LDWT coefficients are obtained.

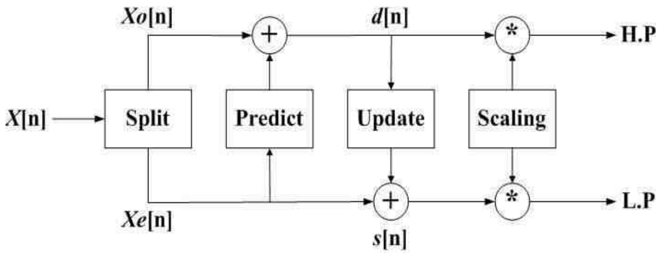

Fig. 2. Block Diagram of LDWT

Equations 11, 12 gives the 9/7 Daubechies decomposition filter coefficients high pass h(n) and low pass g(n) respectively.

h(n) = {0.02675, -0.01686,-0.07822, 0.26686, 0.60295, 0.26686,-0.07822,-0.01686,- 0.02675 } (11)

g(n)={0.09127, -0.05754, -0.59127, 1.11509, -0.59127, -0.05754, 0.09127} (12)

Lifting decomposition result is follows in equation.13

$$\mathrm{P(z)}= \begin{pmatrix} 1 & \propto(1+z-1) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \beta(1+z) & 1 \end{pmatrix} \begin{pmatrix} 1 & \gamma(1+z-1) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta(1+z) & 1 \end{pmatrix} \begin{pmatrix} \zeta & 0 \\ 0 & 1/\zeta \end{pmatrix} \tag{13}$$

where $\propto$=-1.586134342, β=-0.0529118, γ=0.882911075, δ=0.443506852, 1/ ζ=1.230174105.

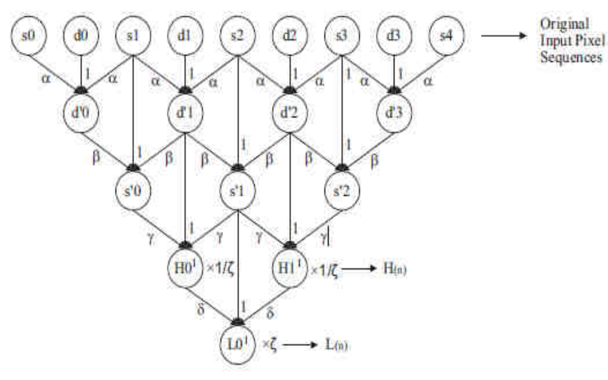

Fig. 3. 9/7 LDWT Diagram

Figure 3 shows the 9/7 lifting scheme diagram. Here first input sequence is divided into odd and even subset which is known as splitting phase. Then the output follows the predict, update and scaling phases.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

B. Background subtraction and frame differencing.

For steady state condition, background subtraction method is adopted. Here moving regions in the frames are not considered for reference or background frame[5] .The reference frame is subtracted from the current frame to get the targeting object in the frame. In frame differencing method, the moving regions are extracted from successive frames of the video sequence. Moving regions are detected from this method. Background subtraction[1] and frame differencing[2] approaches can be represented as in equation 3 and 4

$$\{BS(i,j)\}_k=\{F(i,j)\}_k - \{F(i,j)\}_{k=0} \tag{3}$$

$$\{FD(i,j)\}_k=\{F(i,j)\}_k - \{F(i,j)\}_{k-1} \tag{4}$$

where i =1,…….,m and j=1,……,n k =1,2,3………,k; kis the number of frames in the video. BS is the background frame and FD is the frame difference.

C. Binarization and OR operation

The process of conversion of gray image into black and white image is called binarization. The gray scale image consists of 256 intensity levels which are converted into two equal levels of either 0's(dark) or 1's (white).This can be done by taking the threshold value. For binarization, selection of threshold value is very important. From frame to frame the threshold value may change due to changes in the average frame intensity from one frame to another. Some parameters like mean U and standard deviation[1] σ is used to calculate the global threshold value[6] for a particular frame.

$$U= \frac{1}{r*c}\sum_{i=1}^{r}\sum_{j=1}^{c}FD(i,j) \tag{5}$$

$$\sigma=\sqrt{\frac{1}{r*c}\sum_{i=1}^{r}\sum_{j=1}^{c}(FD(i,j)-u)^2} \tag{6}$$

$$T=0.06*\sigma \tag{7}$$

Here FD (i,j) is the frame difference of successive frames, and [r, c] is the size of the input image. The threshold value is then obtained as shown in equation (7) and the constant value 0.06 is obtained through trial and error method. After binarization.[7] OR operation is carried out in order to detect the object while moving, and when object stops suddenly. OR operation can be given by:

$$\{RV(i,j)\}_k=\{BS(i,j)\}_k\mathrm{OR}\{FD(i,j)\}_k \tag{8}$$

RV is the resultant frame of OR operation. Since exact thresholding is not possible due to the presence of noise, some morphological operations[7] are also done on the binarized frame. To remove the noise, pixel opening operation is done first, since object pixels also get eliminated during this process. Closing operation is done after this to retrieve the lost pixel.

D. Variance method

It is used to detect and track the moving region. In this step the variance values of the entire row and column[8] of the frame is calculated .It gives image intensity variations corresponding to each row and column. Variance is given[9] as:

$$\sigma^2=\frac{1}{n2}\sum_{x=1}^{m}RV^2\ (i,j) \tag{9}$$

Where, RV is the sub image and n is the number of pixels in the sub image. First, the intersection point of row and column where the value of variance is maximum is considered. Then a window is formed around it and the same process is carried out in the window resulting in a new intersection point. The window is shifted around the new intersection point until same point is obtained as intersection in successive iteration .Here the process converges to a point. Finally, for showing the location of the object in the input frame, a bounding box is placed around the object. This will help in tracking the object in the video.

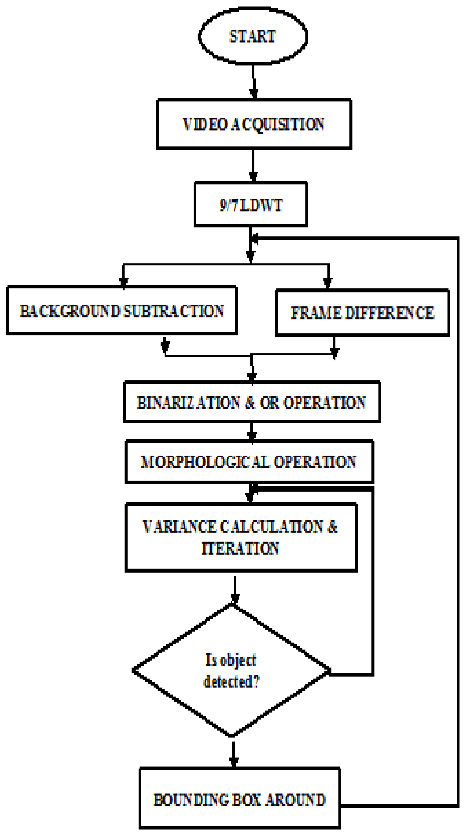

Fig. 4. Flow Diagram of Proposed Method.

IV. EXPERIMENT RESULT

This section gives the comparative result of the various methods done. The tests were carried out on an Intel(R) Core(TM) i3 processor with 2.10GHz clock frequency and 4GB of RAM. CAVIAR[11] project at INRIA labs database[12]are used. The details about the videos are given in Table1.Comparing the result of LDWT and DWT from the Fig (7) and Fig (8), we can see that the objects are accurately detected using LDWT method where three objects are detected from the input frame; but in the case of DWT method only partial results are obtained, leaving all other objects eliminated from the frame. So DWT may cause incomplete detection. For tracking, we calculate the variance of the image where the pixel with maximum nonzero value is considered. The computational time of LDWT is also lesser compared to DWT as shown in Table 1 .Fig 12 shows the graphical representation of comparison of computational time for object tracking.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

Table 1. Input Video Details

Video Names	Input Video Details		Computation Time for 50 frames (in seconds)	
	No. of Frames	Duration of Video (sec)	LDWT-Variance Method	DWT-Variance Method
Browse1.mpg	1040	41	75.56	86.34
Walk1.mpg	610	24	71.44	89
Rest_In Chair.mpg	1007	42	73.34	87.57
LeftBag.mpg	1439	56	71.98	89.78
Women.mp4	201	9	170	205

Fig. 5. Input Frame.

Fig. 6. Background Frame.

Fig. 7. Result of Background Substraction using LDWT

Fig. 8. Result Background Substraction Using DWT

Fig. 9. Motion Tracking at 124th Frame using LDWT

Fig. 10. Motion Tracking at 124th Frame using DWT

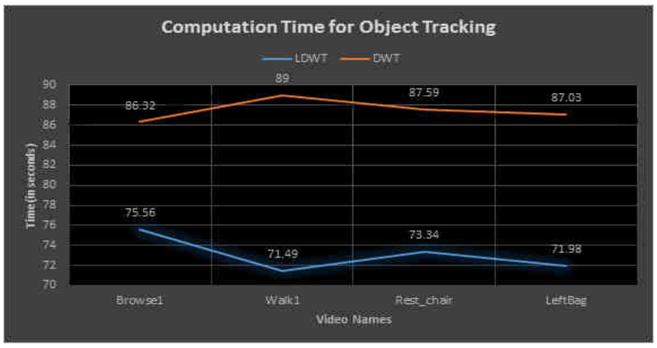

Fig. 11. Computational time for object tracking

V. CONCLUSION

A combined LDWT and variance method has been used for object tracking. Since LDWT preserves the fine shape of the object, an accurate detection can be achieved; here the memory requirement is also less. The use of variance method for tracking increases the computational speed to a great extent. The comparative results show that proposed method is much better than the conventional approach in detection and tracking time. On account of these advantages, the proposed method is very suitable for real time applications such as video surveillance, traffic monitoring system etc.

REFERENCES

[1] Ya Liu, Haizhou Ai, and Guang-you Xu. "Moving object detection and Tracking based on background substraction".In Multispectral Image Processing and pattern recongnition, pages 62-66.International Society for Optics and Photonics,2001

[2] Muyun Weng Guoce Huang,Xinyu Da, "A New Interframe Difference Algorithm for Moving Target Detection,"3rd International Congress on Image and Signal Processing(CSIP 2010),2010.

[3] Sepehr Aslani, Homayoun Mahdavi-Nasab World Academy of Science,Engineering and Technology International Journal of Electrical, Computer,Electronics and Communication Engineering Vol:7, No:9, 2013

[4] Sugandi, B.; Kim, H.; Tan, J. K. & Ishikawa, S. (2007). Tracking of moving objects by using a low resolution image, International Conference on Innovative Computing, Information and Control, (September 2007) pp. 408-408.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

[5] Cl´audio Rosito Jung. Efficient background subtraction and shadow removal for monochromatic video sequences.Multimedia,IEEE Transactions on, 11(3):571–577, 2009.
[6] Xiaoshi Zheng, Yanling Zhao, Na li, and Huimin Wu, "An Automatic Moving Object Detection Algorithm for Video Surveillance Applications,"International Conference on Embedded Software and Systems, 2009
[7] Arnab Roy, Sanket Shinde, and Kyoung-Don Kang, An Approach for Efficient Real Time Moving Object Detection
[8] Vishal R. Satpute, Kishor D. Kulat, and Avinash G. Keskar, "A novel approach based on variance for local feature analysis of Facial images," IEEE-RAICS International Conference on Recent Advances in Intelligent Computational Systems (RAICS), 2011, pp – 210-215
[9] Vishal R. Satpute, Kishor D. Kulat, and Avinash G. Keskar, "Variance Method for Finding Local Feature Points on Facial images,"International Conference on Signal, Image and Video Processing (ICSIVP*)* 2012, IITPatna, 2012, pp – 148 -153.
[10] For video database: http://groups.inf.ed.ac.uk/vision/ CAVIAR/ CAVIARDATA1/
[11] For video database: http://www.eecs.qmul.ac.uk/~avss2007_d.html
[12] Varshney, H.; Hasan, M. & Jain, S. (2007). Energy efficient novel architecture for the lifting based discrete wavelet transform, IET Image Process, vol. 1, no. 3, (September 2007),pp.305-310.

Resmi H. B, is currently pursuing M.Tech. Degree in Signal Processing with the Department of Electronics and Communication Engineering, LBS Institute of Technology for women, Poojappura, Trivandrum, Kerala. She received B. Tech degree from the University of Calicut,in 2012 in Electronics and Communication Engineering.

Deepambika V. A, currently working as Assistant Professor, in department of Electronics and Communication Engineering at LBS Institute of Technology for Women, Trivandrum, Kerala, India. Has obtained B.Tech and M.Tech degrees from University of Kerala and is currently pursuing research in the area of Computer Vision.

Prof. (Dr). M. Abdul Rahman is the Pro Vice Chancellor of Kerala Technological University. He was the Director of AICTE, Ministry of Human Resources Development, Govt of India. He holds PhD in Computer Science & Engineering and also possesses MBA and PG Diploma in Human Resource Management. He is a Professor in Computer Science & Engineering and has many publications in to his credit. He is author of many books in Technical Education. He had also associated with Prime Minister's Scholarship scheme for the State of Jammu & Kashmir with the aim to encourage the youth of J&K to pursue higher studies throughout the country and mingle different cultures and thus bringing them to the integral part of the Country.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.