ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 6, Issue 8, August 2017

A ROBUST STATISTICAL APPROACH FOR EXTRACTION AND TRACKING OF HUMANBODY FROM VIDEO

Malik KT, Mohammed Anvar P K

***Abstract*— — In video processing, tracking objects that are in motion has attracted lot of interest of the researchers all over the world. In numerous computer vision applications, such as monitoring traffic, remote video surveillance automation, human tracking, etc, moving object detection in video sequence is the major step of knowledge extraction .In this paper, we present a comparison of human tracking algorithms that have been applied to the tracking of people from video. first part,a human body detection algorithm based on the combination of moving information with shape information is proposed in the paper. Firstly, Eigen-object computed from three frames in the initial video sequences is used to detect the moving object. Secondly, the shape information of human body is used to classify human and other object. Furthermore, the occlusion between two objects during a short time is processed by using continues multiple frames. The advantages of the algorithm are accurately moving object detection, and the detection result doesn't effect by body pose. Moreover, as the shadow of moving object has been eliminated. second part, an effective human tracking system based on Daubechies Complex Wavelet Transform (DaubCxWT) combined with histogram of template is introduced. This transform is suitable to track a person in video sequences because of its approximate shift-invariance nature. Initially, DaubCxWT co-efficients associated to the person are computed. Then, in Daubechies complex wavelet domain, the energy of these co-efficients is compared to the neighbouring object, to carry out tracking in the consecutive frames. Histogram of template feature is used to extract the texture and information for the human detected. Daubechies Complex Wavelet co-efficients and histogram of template features are combined to form feature vector. In order to build feature vector for every pixel in that area, the calculated coefficients are utilised. Further, by making use of the generated feature vectors inside an adaptive search window, optimal search for the best match is performed. Search window adaption is employed to estimate the speed and direction of the person, in motion. DauCxWT method has shown appreciable results.**

***Index Terms*— Moving object detection, Background subtraction, Histogram of template, daubechies complex wavelet transfer.**

I. Introduction

There are three types of methods mainly used in moving object detection. These methods are the frame subtraction method, the background subtraction method and the optical flow method. In the Frame subtraction method the difference between two consecutive images is taken to determine the presence of moving objects. The calculation in this method is very simple and easy to develop. But in this method it is difficult to obtain a complete outline of moving object;
therefore the detection of moving object is not accurate. In the Optical flow method, calculation of the image optical flow field is done. The clustering processing is done according to the optical flow distribution characteristics of image. From this, the complete movement information of moving body is found and it detects the moving object from the quantity of calculation, poor antinoise performance makes it unsuitable for real-time applications. The background subtraction method is the method in which the difference between the current image and background image is taken for the detection
moving objects by using simple algorithm. But it is very sensitive to the changes which occur in the external environment and it also has poor anti interference ability. One advantage of this method is, it can provide the most complete object information in the case of the background is known. In the background subtraction method, in a single static camera
condition, the dynamic background modeling is combined with dynamic threshold selection method which depends on the background subtraction. The background is updated on the basis of accurate detection of object

Object tracking in video sequence is an important research topic in video analysis applications such as visual navigation, video surveillance, to analyze shopping behavior of customers in retail shop and military surveillance. Object tracking may be described as the process to initially detect the object of interest and then to continuously estimate the position and various significant information of the object in images against dynamic scenes. Moving object trajectory in video sequence can be obtained over time by locating its position in every image of the video, using object tracking process. The tracking algorithms that are reliable possess some issues such as, image noise which causes brisk changes in the appearance, variations in illumination, shape and size variations. Performance of any tracking system depends on the observation models and target representation. For many applications behaviour of persons are most interest such as for traffic surveillance, sports analysis or military surveillance. Thus people detection from visual observations is a very active research area. Most of the existing methods for multi human tracking are still restricted to specific application contexts.

In this paper a system which successfully detects and tracks people in video, is presented. Frames are generated from input video and pre-processing is performed for each frame. Preprocessing stage includes RGB to gray conversion, resizing and normalization of frames. Few of these frames are considered for generating background. Then the foreground object is extracted by subtracting the generated background from each frame. This foreground is sent to Adaboost to detect and validate human blobs. Each blob location is computed to crop the area of interest from the gray image. Daubechies Complex Wavelet Transform (DaubCxWT) is applied to this area to get the coefficients associated to human. Histogram of Template (HOT) features is also obtained. Daubechies Complex Wavelet coefficients and HOT features are combined to form feature vector. Combining features that best discriminate between object and background will improve the tracking performance . Further, humans are tracked by defining a search window. The computed feature vectors are used to update the search window.

The rest of the paper is organized as follows. Chapter 2 proposes the methodology of the proposed system. In chapter 3 experimental results are presented. Finally, conclusions and future work are discussed in chapter 4

II. LITERATURE SURVEY

Hong Han [1] formulated a new integration framework of texture and color information for background modeling, in which the foreground decision equation includes three parts. This framework is able to combine the advantages of texture and colour features. Here a block based method is used for background modelling. The limitation are some background scenes with same colour as humans can be misclassified and Computation time is high .

M. Archana and M. Kalaisevi Geetha [2] proposed Object detection is performed using background subtraction, optical flow and spatio-temporal filtering techniques .Due to frequent camera motion in broadcast tennis video detection and tracking of players and ball is a challenging task. Another problem is human body along with the tennis racket is not Detected accurately. In order to track the ball, logical AND operation is applied between the created background and image difference is performed, from that the ball candidates are detected by applying threshold values and dilated. Finally the ball is tracked. Player detection is performed from AND results by finding the biggest blob and filling the whole detected object by removing the small one and the players are tracked based on the contour

Zhe Lin and Larry S. Davis [3] coined Human detection technique based on shape is performed. A hierarchical part-template matching approach is employed to match human shapes with images to detect and segment humans simultaneously. The limitation of this shape based algorithm is it does not work well in dynamic situations and is unable to determine internal movements well.

Badri Narayan Subudhi et.al [4] designed algorithm includes two schemes one for spatiotemporal spatial segmentation and the other for temporal segmentation. A combination of these schemes is used to identify moving objects and to track them.The limitation of this change information based algorithm is When the camera is dynamic(moving),object detection fails

Horesh Ben Shitrit et.al [5] launched global optimization framework for multi-people tracking that takes image-appearance cues into account, even if they are only available at distance time intervals.the limitations are It does better at preserving identity over very long sequences and Computation time is more.

Koen E. A. van de Sande and Theo Gevers and Cees G. M. Snoek [6] formulated Object detection based on colour descriptors is employed here .The limitations are When the intensity of light changes dynamically, the detection fails. This paper studies the invariance properties and the distinctiveness of colour descriptors in a structured way.

Rodrigo F. S. C. Oliveira and Carmelo J. A. Bastos-Filho [7] introduced Shape based algorithm to detect human beings in videos here aggregated channel features detector is used along with motion feature extractors(MBH,IMHCD).The limitations are Computation time required for detection is high Detection fails if full body part is not visible(suppose a person is sitting in a car and only top portion is visible).

Min-Hsiang Yang et.al [8] Here a coarse-to-fine detection theory algorithm to extract foreground objects based on nonparametric background and foreground models represented by binary descriptors. The limitations are Binary descriptors can only detect,tracking is not possible Cant used in event analysis such as fall detection,military surveilance tracking etc.

Risha K P and Chempak Kumar AI [9] advocated a novel moving objects detection in a video by using a method of optic flow with morphological operation. The limitations are Large quantity of calculations are required to obtain optical flow information. If moving camera is used optical flow method fails Can detect only single moving object.

Chunming Li et.al [10] launched an algorithm based on human shape information to detect and track the moving human from moving cameras, Firstly, human detector based on the human shape information is used to detect the standing and moving human in videos with moving cameras and backgrounds. Then, the moving person is tracked based on a small area tracking method. Finally, they test thier approach on real video experiments. From the experiment results we can see that the system provides a quick and robust tracking.In this work, a robust system for moving human detection and tracking from moving camera is proposed. Human body is detected by the human descriptor in the first; Then the head

detection is performed to find an appropriate tracking area; Finally, a small area based on head and centroid mean shift tracking method is used to track the moving person which makes the tracking result robust even with large changes of appearance and pose of the moving object.Our human detector and tracking method is a global method which makes it not very flexible to part appearance detection. In real-world scenes, part human appearance such as only head and shoulder, or a sitting person is also an important case needed to be considered. The limitation is to detect and track multiple moving person from the very beginning of the person move into the scene, without limitation of full or partial body visible.

B.S.M.Madhavi And M.V.Ganeswara Rao [11] proposed a new method to detect moving object based on background subtraction. First of all, we establish a reliable background updating model based on statistical and use a dynamic optimization threshold method to obtain a more complete moving object. And then,morphological filtering is introduced to eliminate the noise and solve the background disturbance problem. At last, contour projection analysis is combined with the shape analysis to remove the effect of shadow, the moving human body are accurately and reliably detected. The experiment results show that the method runs quickly, accurately and fits for the real-time detection. The occlusion is one of the most common events in object tracking and object centroid of each object is used for detecting the occlusion and identifying each object separately. Video sequences have been captured in the laboratory and tested with this algorithm. The algorithm works efficiently in the event of occlusion in the video sequences.

Ayushi Gahlot et.al [12] formulated an action recognition using Kinect technology by human skeletal tracking. Microsoft Kinect is one of the latest advancements in Computer Vision based HCI (Human Computer Interaction). Thework is focused on how the Kinect sensor captures the 3D information of a scene and recognizes the action being performed by the human body by retrieving the depth image information and real-time skeletal tracking. The Kinect technology has revolutionized the way humans interact with the machines. It has a wide range of applications areas. This work also covers one of the proposed approach to skeletal based action recognition using Kinect.

Jiude Li [13] proposed a tracking approach based on improved camshaft algorithm. Firstly, they introduce some common image noise reduction algorithms. By combination the frame difference and background subtraction methods, an improved moving target detection algorithm is proposed, by which the whole region of target can be detected. Then, with the analysis of particle filtering and traditional Cam-shift algorithm, we introduce a new human body tracking method that is able to choose the target automatically due to the detection result. On the basis of the detection and tracking results, the algorithm of motion parameter estimation is analyzed.

Cewu Lu et.al [14] propose an efficient sparse combination learning framework. It achieves decent performance in the detection phase without compromising result quality. The short running time is guaranteed because the new method effectively turns the original complicated problem to one in which only a few costless small-scale least square optimization steps are involved. This method reaches high detection rates on benchmark datasets at a speed of 140∼150 frames per second on average when computing on an ordinary desktop PC using MATLAB.

Lingyan Jianget.al [15] propounded systematic approach to detect feature points of human body automatically from its front and side images. Firstly, an efficient approach for silhouette and contour detection is used to represent the contour curves of a human body shape with Freeman‟s 8-connected chain codes. The contour curves are considered as a number of segments con-nected together. Then, a series of feature points on human body are extracted based on the specified rules by measuring the differences between the directions of the segments. In total, 101 feature points with clearly geometric properties (that rather accurately reflect the bump or turning of the contours) are extracted automatically, including 27 points cor-responding to the definitions of the landmarks about garment measurements. Finally, this approach was tested on ten human subjects and the entire 101 feature points with specific geography geometrical characteristics were cor-rectly extracted, indicating an effective and robust performance.

Qing Ye et.al [16] proposed a human detection method based on motion object extraction and head–shoulder feature to complete human detection and statistics in video image sequences. Firstly, background subtraction based on adaptive threshold was used to extract foreground moving object information, then image erosion and image dilation were used to bypass the object shade and remove false object in order to optimize the results of motion object extraction. And finally, for realizing human moving object detection, we proposed the

object discrimination algorithm based on human head–shoulder feature to complete human detection and statistics. Experimental results show that the method can successfully realize human detection and statistics. The method is highly accurate and has good real-time and extensive applications. The identification rate is 86% through human video sequences to test. This method can detect human automatically and provide the theoretical and technological base for object detection in the intelligent surveillance system.

Jian Yao and Jean-Marc Odobez [17] introduced a fast method to detect humans from videos captured in surveillance applications. It is based on a cascade of LogitBoost classifiers

relying on features mapped from the Riemanian manifold of region covariance matrices computed from input image features. The method was extended in several ways. First, as the mapping process is slow for high dimensional feature space, they propose to select weak classifiers based on subsets of the complete image feature space. In addition, they propose to combine these sub-matrix covariance features with the means of the image features computed within the same subwindow, which are readily available from the covariance extraction process. Finally, in the context of video acquired with stationary cameras, they propose to fuse image features from the spatial and temporal domains in order to jointly learn the correlation between appearance and foreground information based on background subtraction.

Sapana K. Mishra and Kanchan .S Bhagat [21] proposed a human body detection algorithm based on the combination of moving information with shape information is proposed in the

paper. Firstly, Eigen-object computed from three frames in the initial video sequences is used to detect the moving object. Secondly, the shape information of human body is used to classify human and other object. Furthermore, the occlusion between two objects during a short time is processed by using continues multiple frames. The

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 6, Issue 8, August 2017

advantages of the algorithm are accurately moving object detection, and the detection result doesn't effect by body pose. Moreover, as the shadow of moving object has been eliminated. A shape-based approach for classification of objects is used following background subtraction based on frame differencing. The goal is to detect the humans for threat

assessment. The target intruder is classified as human or animal or vehicle based on the height to width ratio (H/W) of the moving object detected during background subtraction.
Then the general block diagram of shape based human detection is given by

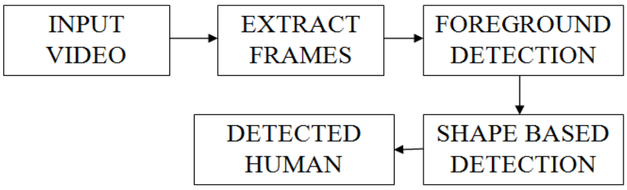

Fig.1. Block Diagram Of Shape Based Human Detection

i. Input video

The input video is an mp4 format video. Fig.3.1 shows the block diagram of shape based human detection. Two videos are taken in which first video contains a walking human and a

moving tyre and the second video contains multiple humans walking in different directions with varying speed. Video 1 is a real time video whereas video 2 is from dataset.

ii. Extraction Of Frames

Motion of pictures constitute a video. So for processing of video, first we have to convert the video in to frames .Frames are nothing but still images captured during a particular interval of time. Here the first video consisting of a walking human and a moving tyre consist of 200 frames and the second video consisting of multiple human beings walking in different directions and with varying speeds consist of 400 frames.

Fig.2 Extracted Frames Of Video 1

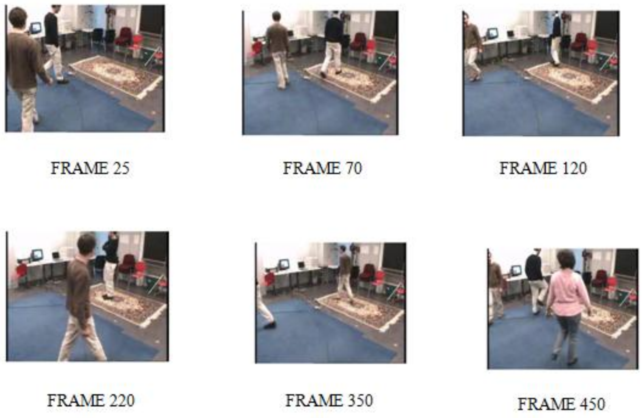

Fig.3 Extracted Frames Of Video 2

iii. Foreground Detection

Background subtraction, also known as Foreground Detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further processing .Generally an image's regions of interest are objects (humans, cars, text etc.) in its foreground. After the stage of image preprocessing which may include image denoising, post processing like morphology. Object localisation is required which may make use of this technique. Background subtraction is a widely used approach for detecting moving objects in videos from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called "background image", or

"background model". Background subtraction is mostly done if the image in question is a part of a video stream. Background subtraction provides important cues for numerous applications in computer vision, for example surveillance tracking or human poses estimation. However, background subtraction is generally based on a static background hypothesis which is often not applicable in real environments. With indoor scenes, reflections or animated images on screens lead to background changes. In a same way, due to wind, rain or illumination changes brought by weather, staticbackgroundmethods have difficulties with outdoor scenes.

Fig.4 Foreground Detected Frames Of Video 1 In Shape Based Detection

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 6, Issue 8, August 2017

Fig.5 Foreground Detected Frames Of Video 2 In Shape Based Detection

The foreground detected video are shown in Fig 3.4 and Fig 3.5. Foreground detection is one of the major tasks in the field of Computer Vision whose aim is to detect changes in image sequences. Many applications do not need to know everything about the evolution of movement in a video sequence, but only require the information of changes in the scene. Detecting foreground to separate these changes taking place in the foreground of the background. It is a set of techniques that typically analyze the video sequences in real time and are recorded with a stationary camera. Here we use GMM model (MinHsiang Yang2014) for foreground detection In statistics, a mixture model is a probabilistic model for representing the presence of subpopulations within an overall population, without requiring that an observed data set should identify the sub-population to which an individual observation belongs. Formally a mixture model corresponds to the mixture distribution that represents the probability distribution of observations in the overall population. However, while problems associated with "mixture distributions" relate to deriving the properties of the overall population from those of the sub-populations, mixture models are used to make statistical inferences about the properties of the sub-populations given only observations on the pooled population, without sub-population identity information. Some ways of implementing mixture models involve steps that attribute postulated subpopulation-identities to individual observations (or weights towards such sub-populations), in which case these can be regarded as types of unsupervised learning or clustering procedures. However, not all inference procedures involve such steps. Mixture models should not be confused with models for compositional data, i.e., data whose components are constrained to sum to a constant value (1, 100%, etc.). However, compositional models can be thought of as mixture models, where members of the population are sampled at random. Conversely, mixture models can be thought of as compositional models, where the total size of the population has been normalized to 1.

iv. Shape Based Algorithm

Cascade Object Detection algorithm is used in this detection.The cascade object detector uses the Viola-Jones algorithm to detect people's faces, noses, eyes, mouth, or upper body. The Viola–Jones object detection framework is the first object detection framework to provide competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones. Although it can be trained to detect a variety of object classes, it was motivated primarily by the problem of face detection. This algorithm is implemented in OpenCV. The problem to be solved is detection of faces in an image. A human can do this easily, but a computer needs precise instructions and constraints. Aggregated Channel Feature (Rodrigo F2011) based extractors are shape based algorithm which uses motion features to detect humans with less false positives. To make the task more manageable, Viola–Jones requires full view frontal upright faces. Thus in order to be detected, the entire face must point towards the camera and should not be tilted to either side. While it seems these constraints could diminish the algorithm's utility somewhat, because the detection step is most often followed by a recognition step, in practice these limits on pose are quite acceptable. Fig 6 and Fig 7 shows detected humans based on cascade object detection which is nothing but shape based detection.

- ☐ Robust – very high detection rate (true-positive rate) & very low false-positive rate always.
- ☐ Real time – For practical applications at least 2 frames per second must be processed.
- ☐ Face detection only (not recognition) - The goal is to distinguish faces from non-faces. (detection is the first step in the recognition process).

v. Blob Analysis

In computer vision, blob detection methods are aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to surrounding regions.

Informally, a blob is a region of an image in which some properties are constant or approximately constant; all the points in a blob can be considered in some sense to be similar to each other. Given some property of interest expressed as a function of position on the image, there are two main classes of blob detectors, one is differential methods which are based on derivatives of the function with respect to position and next is methods based on local extremal which are based on finding the local maxima and minima of the function. With the more recent terminology used in the field, these detectors can also be referred to as interest point operators, or alternatively interest region operators.There are several motivations for studying and developing blob detectors. One main reason is to provide complementary information about regions, which is not obtained from edge detectors or corner detectors. In early work in the area, blob detection was used to obtain regions of interest for further processing. These regions could signal the presence of objects or parts of objects in the image domain with application to object recognition and/or object tracking. In other domains, such as histogram analysis, blob descriptors can also be used for peak detection with application to segmentation. Another common use of blob descriptors is as main primitives for texture analysis and texture recognition. In more recent work, blob descriptors have found increasingly popular use as interest points for wide baseline stereo matching and to signal the presence of informative image features for appearance-based object recognition based on local image statistics. There is also the related notion of ridge detection to signal the presence of elongated object.

Fig.6 Shape Based Human Detection Of Video 1

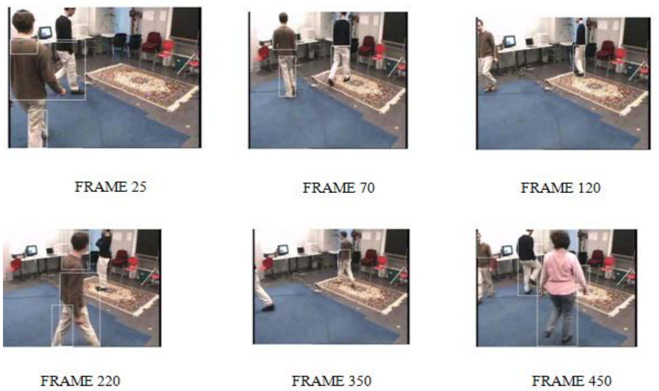

Fig.7 Shape Based Human Detection Of Video 2

III. PROPOSED SYSTEM

A. *Human Detection Based On Daubcxwt*

The block diagram of proposed system is illustrated in Figure.

Frames are generated from input video and pre-processing is performed for each frame. Fig.8 shows the block diagram of Daubechies based human detection algorithm. Preprocessing stage includes RGB to gray conversion, resizing and normalization of frames. Few of these frames are considered for generating background. Then the foreground object is extracted by subtracting the generated background from each frame. This foreground is sent to Adaboost to detect and validate human blobs. Each blob location is computed to crop the area of interest from the gray image. Daubechies Complex Wavelet Transform (DaubCxWT) is applied to this area to get the coefficients associated to human. Histogram of Template

(HOT) features is also obtained. Daubechies Complex Wavelet coefficients and HOT features are combined to form feature vector. Combining features that best discriminate between object and background will improve the tracking performance . Further, humans are tracked by defining a search window. The computed feature vectors are used to update the search window.

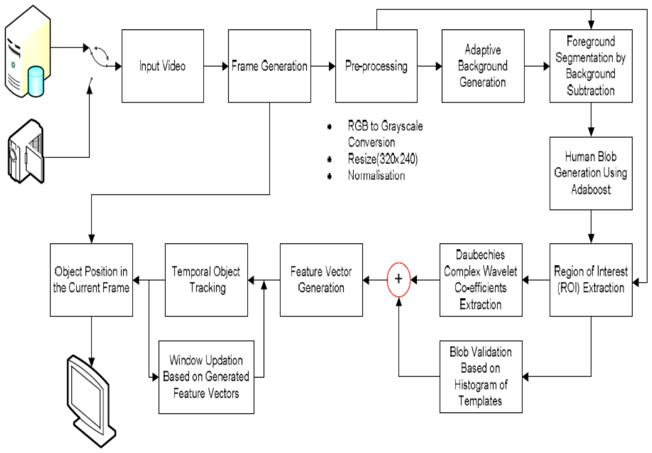

Fig.8 Block Diagram Of Daubechies Based Human Detection

i. *Forground Extraction*

In the area of video processing and computer vision,background subtraction is a real time technique used to extract foreground objects in an image or video stream. This technique assists in distinguishing moving people in videos with static cameras. The distinguishing is done by taking the absolute difference between the reference model and the current frame.Fig.9 and Fig.10 shows foreground detected frames of video1 and video2 respectively.

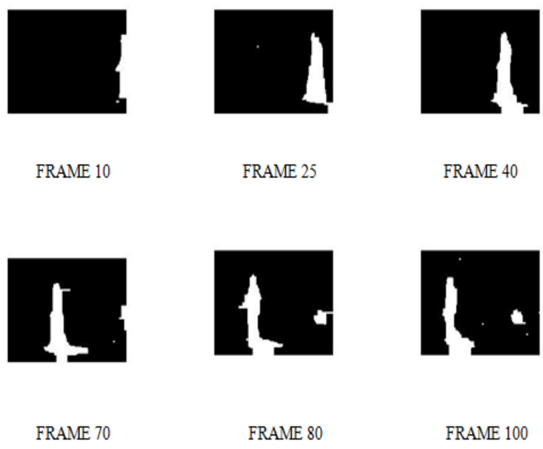

Fig.9 Foreground Detected Frames Of Video 1 In Daubcxwt Detection

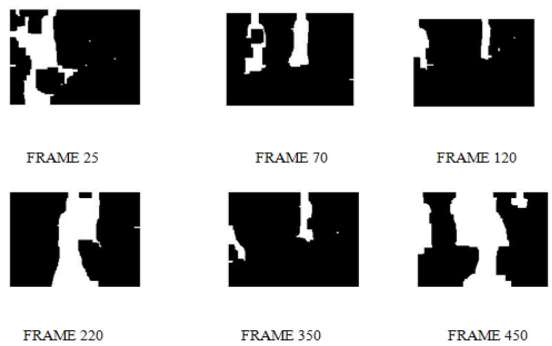

Fig.10 Foreground Detected Frames Of Video 2 In Daubcxwt Detection

ii. Human Detection Using Adaboost

After the separation of moving human from the background scene, human blobs are generated and validated, using AdaBoost algorithm. It is quite complicated to detect human in motion as human bodies have variety of poses and external appearance. Hence, to get a good result out of it AdaBoost algorithm is used. This algorithm is one which can detect people in motion, with varying orientation and sizes, under a challenging background.

iii. Daubechies Complex Wavelet Transform

DaubCxWT is a shift-invariant method. This nature makes this transform suitable to track objects in video stream. Hence it has been adapted in our system to track multiple humans efficiently. The scaling equation of multi-resolution theory is given by equation 1.

$$\emptyset(x) = 2\sum\_n \quad ak\ \emptyset(2x - k) \qquad (1)$$

This ak can be complex or real valued. Daubechies considered ak to be real valued only in order to provide general solution. DaubCxWT is got by considering complex values of both ak and phi Multiresolution analysis of L2(R) and the scaling function are used to define Daubechies's wavelet bases {phi, k(t)} in single dimension. . The wavelet ψ(t) is as given in equation 2.

$$\varphi(t) = 2\ \sum_n -1^n\ a_{1-n}\ \emptyset(2t - n) \qquad (2)$$

Following are the significant benefits of DaubCxWT noted based on the properties of DaubCxWT, for people tracking in video stream:

- The shape of the signal which is quite essential in object tracking applications is by the linear phase property of DaubCxWT. Thus it lessens the incorrect tracks of people..
- The real components of the DaubCxWT represent only few of the stronger edges.

whereas imaginary part represents every strong edge. This helps in retaining the edges. Hence DaubCxWT is rightly called as local edge detector. If an input signal shift causes an unpredictable change in the transform coefficients, then that transform is said to be shift-sensitive. Shift-sensitivity is reduced in DaubCxWT. As the tracker navigates through the video frames, the reconstruction by making use of real valued Discrete Wavelet transform (DWT) coefficients is altered remarkably. In the case of complex wavelet transform, all local shifts and orientation are reconstructed in the similar fashion. Therefore, boundaries of object in next frames can be found quickly and accurately by employing DaubCxWT.

iv. Histogram Of Template

The texture information is not taken into consideration for detection by gradient based feature extraction procedures. For complex scenes containing dynamic illumination variations (Kalpana-2015),Local Binary Pattern texture based object detection is used. The texture information are more accurate than shape based detection. These methods consider only gradient information. When both, texture and gradient information are considered, more precise detection result can be acquired. The method extracts features like gradient and intensity information, and both information are made homologous. Also, these features will encode the relationship of three pixels in one template.. Fig.11 shows the templates which describes spatial relationship of three pixels . It is expected that these templates can be used to reconstruct the human body. Hence, all the information of human body can be represented by these templates. In our work, for calculating features the templates (1) to (8) are employed. Some formulas to express human body shape make use of these templates. These formulas have reasonable computation complexity.

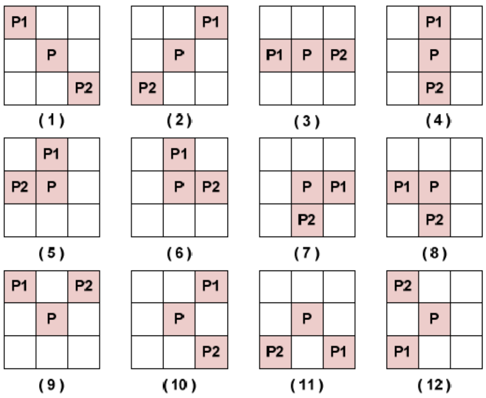

Fig.11 Spatial Relationship Of Three Pixel

The formulae are based on both texture and gradient information, which gives concrete definition of these feature. Using the below given two function, texture information can be obtained.

$$I(P) > I(P1)\ \&\&\ I(P) > I(P2) \qquad (3)$$

According to this function, for every template, a pixel P is said to have met the template, if the intensity value of P is more than the two neighbouring pixel P1 and P2. It is able to capture the pixels with larger value in one template, and the properties of local part of human body are well reflected by the histogram of pixels that satisfy every template in a sub window.

This histogram of pixels meeting different templates is calculated as feature. Histogram of pixels with 8 bins . Each template is represented as a bin in the histogram. The amount of pixels which meet corresponding template in the given region decides the value of each bin. The second function is:

$$k = \arg\max\ \{I(Pi) + I(P1i) + I(P2i)\} \qquad (4)$$

Function (6) means that for each template, if the gradient magnitude value of P is more than other two, we say P meets this template. For function (7), if P meets template, the sum of

gradient magnitude of three pixels in template k must be the biggest one in all templates. For each function, an m-dimensional vector can be extracted from a given region. m denotes number of templates used in feature extraction.

Here, an 8 dimensional vector is obtained for each function, since we are using 8 templates. The final feature vector is generated by integrating all these vectors together. For function (5) templates, the histogram contains 32bins. Compared with methods that use features like gradient information, such as HOG, the proposed feature shows more discriminative ability.

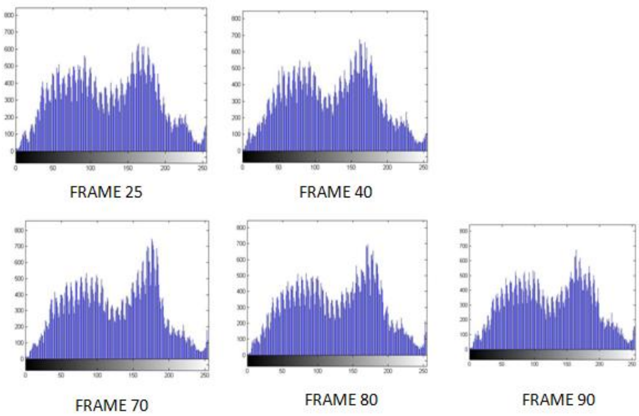

Fig.12 Histogram Of Template Of Video 1

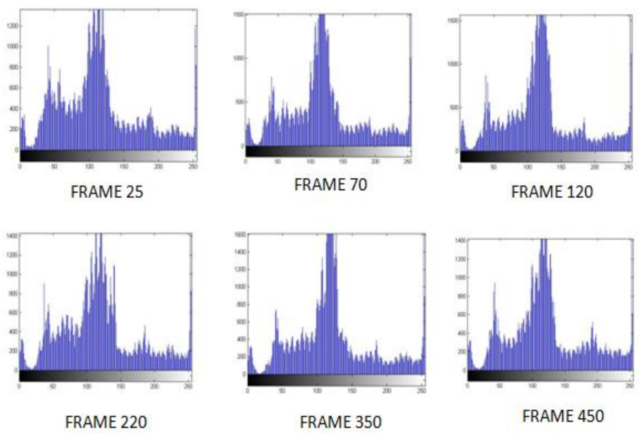

Fig.13 Histogram Of Template Of Video 2

v. Temporal Object Tracking

The generated DaubCxWT co-efficients and HOT features are combined to form feature vector (FV). The next stage is to track detected people in the successive frames. In our system, temporal tracking is used for object tracking, in which, at a reference frame, rectangle is drawn around the detected object. At the reference frame, the pixels in the rectangle must be tracked temporally. This is the final step to be done before locating the object in the next frame. The FV are used by the temporal tracking to the find the new pixel locations in the adaptive search window.Fig.14 and Fig.15 shows the tracking of the predicted frames.

Fig.14 Daubechies Based Human Detection Of Video 1

Fig.15 Daubechies Based Human Detection Of Video 2

IV. EXPERIMENTAL RESULTS

A. Inferences Of Human Detection Based On Shapeand tracked human of video1 and video2.

Fig.16 Shape Based Human Detection And Tracking Of Video 1

In this work, a human body detection algorithm based on the combination of temporal information and shape information is designed. Firstly, moving objects are detected using the proposed background elimination technique. Secondly, shape information is used to distinguish human body and other moving object and the outside rectangle of moving object is computed using the max width and height value of the moving regions.Furthermore, occlusion during a short time is handled by detecting the shape of moving object in continues frames. The evaluation cases shows the accurate detection of moving object and the detection result do not effect by the body pose. Also, the shadow of moving object has been eliminated in detection step, thus human can be detected.

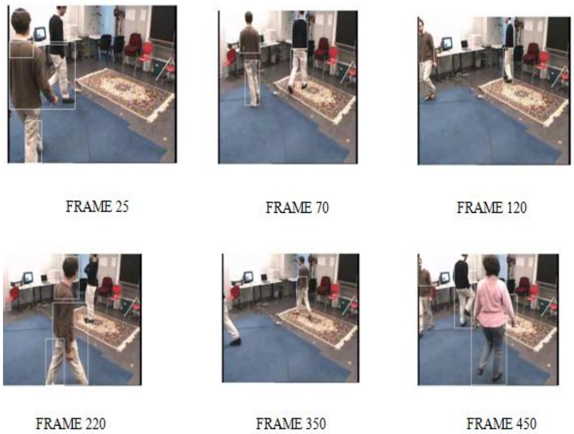

Fig.17 Shape Based Human Detection And Tracking Of Video 2

Overall Efficiency of the project is about 77.78%, i.e., most of cases gave the successful result by the first detection algorithm. The algorithm is robust to noise and can detect the human bodies under complex circumstance.

B. *Inferences On Human Detection Based On Daubechies*

Fig.18 Daubechies Based Human Detection And Tracking Of Video 1

The results of above work is presented in this section. The video clips considered here are of frame size 320 x240 with static background . The algorithms are processed on gray scale of every frame. Foreground is generated by taking absolute difference between generated background model and the gray frame. Adaboost is applied to detect humans in the frame bylocating 15 cells at around the human silhouette. Here multiple people are being tracked based on feature updation in DaubCxWT combined with HOT features are used to track the people inside a search window of size 90x90.Each feature is computed on a patch with random size and position from within the bounding box of detection. The HOT features are represented as histogram containing 8 bins and every bin corresponds to one template. Out of 12 diff templates are being used. The average accuracy attained is 90%.

Fig.19 Daubechies Based Human Detection And Tracking Of Video 2

C. *Performance Comparisons*

The main objective of my work is to compare the performance of two human detection algorithms in terms of detection accuracy, precision and recall. Fig.21 shows the idea behind precision and recall.

i. *Detection Accuracy*

Detection accuracy = Number of human detected frames / number of frames containing human

Table. 1 Detection Accuracy Of Both Human Detection Methods

Video	Shape Based Detection(%)	Human Detection Based OnDaubechie(%
VIDEO 1	59.37	97.35
VIDEO 2	27.96	89.83

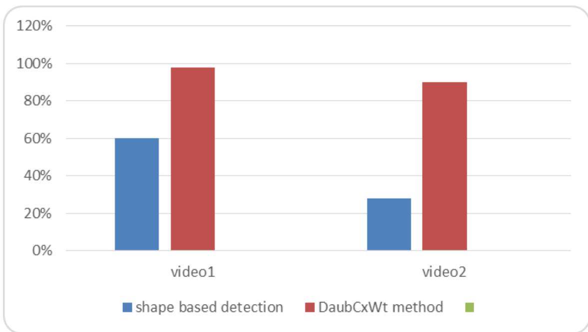

Fig.20 Detection Accuracy Comparison Of Both Algorithm

ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 6, Issue 8, August 2017

ii. Precision

Precision = Number of human detected / Total number of detections

= True positive / true positive + False positive

Where True positive = It is the quantity which is defined as the number of moving humans detected.

iii. Recall

Recall = Number of human detected / Total number of humans

=True positive / True positive + False negative

Table.2 Precission And Recall Of Multiple Human Detected Video

Performance Metrices	Shape Based Human Detection(%)	Daubechies Based Detection(%)
Precision	42.5	91.66
Recall	44.4	86

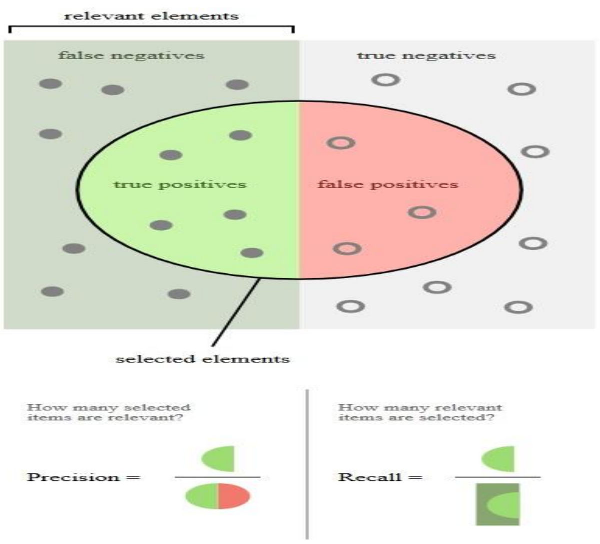

Fig.21 Precission And Recall

V. CONCLUSION

Human detection using shape based features and human detection using daubechies complex wavelet transform is performed in this work. In shape based human detection the shadows of human are also detected and also detection work on high clarity videos. Some of the other disadvantages of this shape based algorithm are detection works only when the camera is stable and also the edges of humans are not detected accurately. Shape based detection is very much dependent on intensity variations and detection result doesn't affect by body pose. In this daubechies based human detection algorithm human blobs are generated using adaboost machine learning algorithm. Histogram of template feature is used to extract both texture and gradient information. After performing the daubechies complex wavelet based human detection,the detection accuracy, precision and recall is more compared to shape based algorithm. The false positives are also less in the proposed human detection algorithm.

REFERENCES

[1] Hong Han, Jianfei Zhu, Shengcai Liao, Zhen Le and Stan Li), "Moving Object Detection Revisited Speed and Robustness", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 4, 2014.

[2] M.Archana, Kalaisevi geetha, "Object Detection and Tracking based on Trajectory in Broadcast Tennis Video", Procedia Computer Science,vol.58,p.p.225-232, 2015. C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, "Rotation, scale, and translation resilient public watermarking for images," *IEEE Trans. Image Process.*, vol. 10, no. 5, pp. 767-782, May 2001.

[3] Zhe Lin and Larry S. Davis, "Shape-Based Human Detection and Segmentation via Hierarchical Part-Template Matching", IEEE transactions on pattern analysis and machine intelligence, vol. 32, no. 4, april 2010.

[4] Badri Narayan Subudhi, Pradipta Kumar Nanda and Ashish Ghosh,, "A Change Information Based Fast Algorithm for Video Object Detection and Tracking" , IEEE transactions on circuits and systems for video technology, vol. 21, no. 7, july 2011.

[5] Horesh Ben Shitrit, Jerome Berclaz, Francois Fleuret, Pascal Fua, "TrackingMultiplePeopleunderGlobalAppearanceConstraints",Elsevier, vol.12,no.42, 2015.

[6] Koen E. A. van de Sande ,Theo Gevers ,Cees G. M. Snoek, "Evaluation of Colour Descriptors for Object and Scene Recognition", IEEE International conference on video and image processing, 2008 .

[7] Rodrigo F. S. C. Oliveira, Carmelo J. A. Bastos-Filho, "Human Detection in Digital Videos Using Motion Features Extractors", IEEE International conference on video processing, 2011.

[8] Min-Hsiang Yang, Chun-Rong Huang , "Binary Descriptor based Nonparametric Background Modeling for Foreground Extraction Using Detection Theory", IEEE transactions on circuits and systems for video technology, 2014.

[9] Risha K P, Chempak Kumar A , "Novel method of detecting moving object in video", International Conference on Emerging Trends in Engineering, Science and Technology,vol.24,p.p.1055-1060, 2016.

[10] Rucha Damle, Aditya Gurjar, Anuja Joshi, Kartik Nagre, "Human Body Skeleton Detection And Tracking", International Journal of Technical Research and Applications e-ISSN: 2320-8163, Volume 3, Issue 6, PP. 222-225 NovemberDecember, 2015.

[11] Chunming Li, Yunfeng Xu, Xinna Shi, Shaoguang Wu "Human Body Detection and Tracking From Moving Cameras", 5th International Conference on BioMedical Engineering and Informatics BMEI 2012, 2012.

[12] B.S.M.Madhavi, M.V.Ganeswara Rao, "A Fast And Reliable Motion Human Detection And Tracking Based On Background Subtraction", IOSR Journal of Electronics and Communication Engineering (IOSRJECE) ISSN : 2278-2834 Volume 1, Issue 1 (May-June 2012), PP 29-35,2012.

[13] Jiude Li, "Research On Camera-Based Human Body Tracking Using Improved CamShift Algorithm", International Journal On Smart Sensing And Intelligent Systems Vol. 8, No. 2, June 2015.

[14] CewuLu,JianpingShi,JiayaJia,"AbnormalEventDetectionat150FPSin MAT- LAB" ,IEEE International Conference on Computer Vision, 2013.

[15] Lingyan Jiang, Jian Yao, Baopu Li, Fei Fang, Qi Zhang, Max Q.H. Meng, "Automatic Body Feature Extraction from Front and Side Images", A Journal of Software Engineering and Applications, , 5, 94-100, 2012.

[16] Qing Ye,Rentaogu,Yuefeng Ji, "Human detection based on motion object extraction and head-shoulder feature", Optik124(2013)3880-3885,2013.

[17] Jian Yao Jean-Marc Odobez, "Fast Human Detection from Videos Using Covariance Features", This Paper Appeared In The Eccv Visual Surveillance Workshop (Eccv-Vs), Marseille, October 2008.

[18] Kishor B. Bhangale, Prof. R. U. Shekokar, "Human Body Detection using Histogram of Oriented Gradients and SVM", International Journal of Electronics Communication and Computer Engineering Volume5, Issue (4)July,Technovision-2014, ISSN2249–071X, 2014.

[19] Gundegoni Rajeshwari, Vara Prasad Rao, "Extraction Of Standing Human Bodies In Single Images Based On Multi Level Image Segmentation ", International Journal of Electronics Communication and

Computer Engineering Volume5, Issue (4)July,Technovision-2014, ISSN2249–071X, 2014.

Mr. Malik KT , completed B.Tech in Electronics And Communication Engineering from Calicut university in 2015. Currently pursuing M.Tech in Electronics and Communication Engineering from APJ Abdul Kalam Technological University.his research interests include Image Processing and video processing.

Prof. Mohammed Anvar PK, Completed B.Tech in Electronics And Communication Engineering from Calicut university .Completed M.tech in VLSI Design from Anna University Coimbatore.Currently working as an Assistant Professor in Electronics and Communication Engg.dept. of Al Ameen Engineering College, Palakkad, Kerala State, India..