DEVELOPMENT OF MOTION BASED MULTIPLE HUMAN DETECTION AND TRACKING USING BACKGROUND SUBTRACTION ALGORITHM

S. Govardhan Naidu[1], S. Srilakshmi[2]

[1]*Department of ECE, MRITS, Hyderabad, India*

[2]*Assistant Professor, Department of ECE, MRITS, Hyderabad, India*

ABSTRACT

Recent research in computer vision has increasingly focused on building systems for observing humans and understanding their look, activities, and behavior providing advanced interfaces for interacting with humans, and creating sensible models of humans for various purposes. In order for any of these systems to function, they require methods for detecting people from a given input image or a video. Visual analysis of human motion is currently one of the most active research topics in computer vision. In which the moving human body detection is the most important part of the human body motion analysis, the purpose is to detect the moving human body from the background image in video sequences, and for the follow-up treatment such as the target classification, the human body tracking and behavior understanding, its effective detection plays a very important role. Human motion analysis concerns the detection, tracking and recognition of people behaviors, from image sequences involving humans. According to the result of moving detection research on video sequences. This paper presents a new algorithm for detecting moving objects from a static background scene to detect moving object based on background subtraction. We set up a reliable background updating model based on statistical. After that, morphological filtering is initiated to remove the noise and solve the background Interruption difficulty. At last, contour projection analysis is combined with shape analysis to remove the effect of shadow; the moving human bodies are accurately and reliably detected. At last the motion of each track is estimated by a Kalman filter. The filter is used to predict the track's location in each frame, and determine the likelihood of each detection being assigned to each track. Kalman Filter operates by Predicting the new state and its uncertainty and Correcting with the new measurement. The experiment results show that the proposed method runs rapidly, exactly and fits for the concurrent detection.

Keywords: *Background subtraction, Background updating, static background, moving object detection, Kalman filter.*

I. INTRODUCTION

An important stream of research within computer vision which has gained a lot of importance in the last few years is the understanding of human activity from a video. The growing interest in human motion analysis is strongly motivated by recent improvements in computer vision, the availability of low-cost hardware such as video cameras and a variety of new promising applications such as personal identification and visual surveillance. It aims to automatically guess the motion of a person or a body part from monocular or multi-view video images. Human

body motion analysis has been an interesting research for its various applications, such as physical performance, evaluation, medical diagnostics, virtual reality, and human–machine interface. In general, three aspects of research directions are considered in the analysis of human body motion: tracking and estimating motion parameters, analyzing of the human body structure, and recognizing of motion activities. At present methods used in moving object detection are mainly the frame subtraction method, the background subtraction method and the optical flow method. The presence of moving objects determined by calculating the difference between two consecutive images, in the frame subtraction method. Its calculation is simple and easy to implement. For a variety of dynamic environments, it has a strong adaptability, but it is generally difficult to obtain complete outline of moving object, responsible to appear the empty phenomenon, as a result the detection of moving object is not accurate. Optical flow method is to calculate the image optical flow field, and do clustering processing according to the optical flow distribution characteristics of image. This method can get the complete movement information and detect the moving object from the background better, however, a large quantity of calculation, sensitivity to noise, poor anti-noise performance, make it not suitable for real-time demanding occasions. The background subtraction method is to use the difference method of the current image and background image to detect moving objects, with simple algorithm, but very sensitive to the changes in the external environment and has poor anti- interference ability. However, it can provide the most complete object information in the case of the background is known. In this paper, in a single static camera condition, we combine dynamic background modeling with threshold selection method based on the background subtraction, and update background on the basis of exact detection of object, this method is effective to improve the effect of moving object detection. Any motion detection system based on background subtraction needs to handle a number of critical situations such as:

1. Noise image, due to a poor quality image source;
2. Gradual variations of the lighting conditions in the scene;
3. Small movements of non-static objects such as tree branches and bushes blowing in the wind;
4. Undeviating variations of the objects in the scene, such as cars that park (or depart after a long period);
5. Sudden changes in the light conditions, (e.g. sudden raining), or the presence of a light switch (the change from daylight to non-natural lights in the evening);
6. Movements of objects in the background that leave parts of it different from the background model;
7. Shadow regions that are projected by foreground objects and are detected as moving objects.
8. Multiple objects moving in the scene both for long and short periods;

The main objective of this paper is to develop an algorithm that can detect human motion at certain distance for object tracking applications. We carry out various tasks such as motion detection, background modeling and subtraction, shadow detection and removal.

II. ARCHITECTURE OF HUMAN TRACKING SYSTEM

This section describes the development of a visual surveillance system which has the ability to detect and track human in the camera field of view. The image capture from the camera needs to preprocess to reduce the noise. Background subtraction technique is used to subtract out the background and leave the moving object for

further analysis. Color histogram is calculates from the moving object at feature extraction stage. Last but not least, feature matching approach is used to establishing consistent labels across single camera FOV. Figure1 shows the architecture of the human tracking system

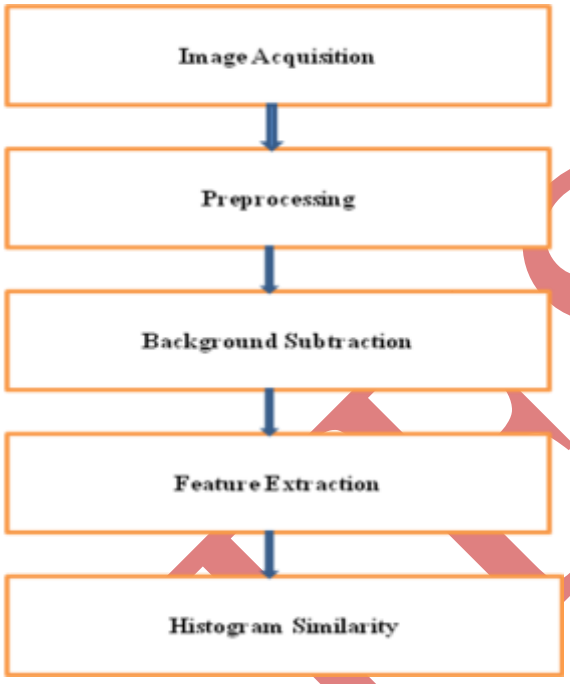

Figure 1.Architecture of Human Tracking System

A. Preprocessing

The image acquisition module captured the image from the camera. Gaussian noise also called normal noise usually arise due to factors such as poor illumination or high temperature. The acquired images are smooth with a Gaussian filter

B. Background Subtraction

Detection of moving human in videos from static camera is widely performed by background subtraction method. The origin in the approach is that of detecting the moving objects from the difference between the existing frame and a reference frame, frequently called the "background copy", or "background replica"[1]. **As** a baric, the background image must be a representation of the scene with no moving objects and must be kept regularly updated so as to adapt to the varying luminance conditions and geometry settings. More difficult models have extended the concept of "background subtraction" beyond its literal meaning. The background subtraction method is the common method of motion detection. It is a technology that uses the difference of the current image and the background image to detect the motion region, and it is generally able to provide data included object information. The key of this method lies in the initialization and update of the background image [2]. The effectiveness of both will affect the accuracy of test results. Therefore, this paper uses an effective method to initialize the background, and update the background in real time.

The process algorithm is described as follow:

1. Sequences of Video Frames
2. Frame Separation
3. Image Sequence
4. Separation of Image Sequence in Current Frame Image and Background Frame Image
5. Perform Background subtraction
6. Detection of Moving Object
7. Perform Background updating
8. Noise Removal
9. Shape Analysis

C. Feature extraction

A color histogram is a popular form of representation of the object appearance one of the human feature in object tracking [5]. It is also independent from scaling and rotation of the object and robust to the partial occlusion. This method makes a model of color histogram according to the appearance of the human shape extracted from the image. The color histogram can be calculated in a grey level space, in RGB or any other color space. In this paper, we proposed using Hue channel of HSV color space to build the color histogram model. The HSV color space corresponds closely with the way humans describe and interpret color. Hue (H) channel represents the "color". Saturation (S) channel is the "amount" of color and Value (V)is the brightness of the color. HSV model has their information stored in hue, saturation and value respectively. Hence, by using Hue channel (color information) is sufficient for the human tracking. The moving object is extracted out by using the binary image obtained from the background subtraction stage. The binary image will act like a mask to filter out non interested region and the histogram model will only build at the moving object region. If the pixel in the mask image is not zero, the pixel from the current frame will copy to a new image as the foreground. Expression is as follow:

$$I_{fore}(x,y) = I_{current}(x,y) \quad if\ I_{mask}(x,y) \neq 0 \tag{1}$$

Where *Ifore* is the foreground color image, *Icurrent* is the current color image and *Imask* is the binary image. The foreground image *Ifore* is converted to HSV color space.1D hue channel histogram h(rk*)* is calculated as representation space for the appearance of the object. The histogram hue Levels range from 0 until 255. Expression is as follow:

$$h(Rm) = Nm = 0,1,2,...255 \tag{2}$$

Where Rm is the mth hue level and *Nm* is the number of pixels in the hue image.

D. Histogram Similarity

Histogram similarity is the method of measuring the similarity between the histogram of a reference histogram and the histogram of the tracked image. There are few histogram comparison methods such as Chi-square distance, histogram intersection, Bhattacharyya distance and etc. Intersection works well for quick matching and

Chi-square and Bhattacharyya work best for slower but more accurate matches. In this test, Bhattacharyya distance is used to compare the histogram of the whole person region in the consecutive frame. After the matching, the similarity of moving object is determine by a threshold value *Ts*. Experiments with indoor and outdoor environment found out that the threshold value *Ts* of 0.4 is acceptable. A score that lower than the threshold value indicate a good match, a unique color boundary box will be constantly labeled at the moving object. Otherwise, the moving object will Label with a white boundary box. Expression of Bhattacharyya histogram comparison is as follows:

$$d_{Bhattacharyya}(H_1,H_2)=\sqrt{1-\sum_i \frac{\sqrt{H_1(i)\cdot H_2(i)}}{\sqrt{\sum_i H_1(i)\cdot \sum_i H_2(i)}}} \quad (3)$$

Where H,H2 are previous and current histograms respectively.

III. BACKGROUND SUBTRACTION ALGORITHM

The flowchart for background subtraction algorithm is given below.

3.1 Flow chart for Background Subtraction Algorithm

Background subtraction algorithm consists of several stages. Those are listed below.

1. Background Image Initialization
2. Moving Object Mining
3. Noise Removal
4. Extraction of Moving Human Body
5. Behaviour Understanding

The above stages are briefly explained as follows.

A. Background Image Initialization

There are many civilizations to obtain the initial background image. For instance, with the first frame as the background directly or, the average pixel brightness of the first few frames as the background or using a background image sequences without the hope of moving objects to approximate the background model parameters and so on. From these methods average method is commonly used for background Image initialization [3], but there are many shadow problems will occur which can be removed by median method, So the median method is selected in this paper to initialize the background.

Expression is as follows:

$$Binit(L, M) = \text{median } Fk(l, m)\ k = 1,2 \ldots n \quad (4)$$

Where Binit is the initial background, *n* is the total number of frames selected.

B. Moving Object Mining

Background subtraction is a popular technique to fragment out the interested objects in a frame. This technique involves subtracting an image that contains the object, with the previous background image that has no foreground objects of interest. The area of the image plane where there is a significant difference within these images indicates the pixel location of the moving objects [5].These objects, which are represented by groups of

pixel, are then separated from the background image by using threshold technique. After the background image B(l,m) is obtained, subtract the background image B(l,m) from the current frame Fk(l,m).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.

Expression is as follows:

$$Dk(l, m) = 1 \quad \text{if } Fk(l,m)-Bk(l,m) \geq T$$
$$0 \quad \text{Otherwise} \qquad (5)$$

The flow chart of moving human body mining is shown in (Fig.1)**:** Human body detection is to identify the corresponding part of human from the moving region. But the extracted moving region may correspond to different moving objects, such as vehicles and other such birds, floating clouds, swaying tree and other moving objects [8]. Hence we use the shape features of motion regions to further determine whether the moving object is a human being or not. Judging criteria are as follows:

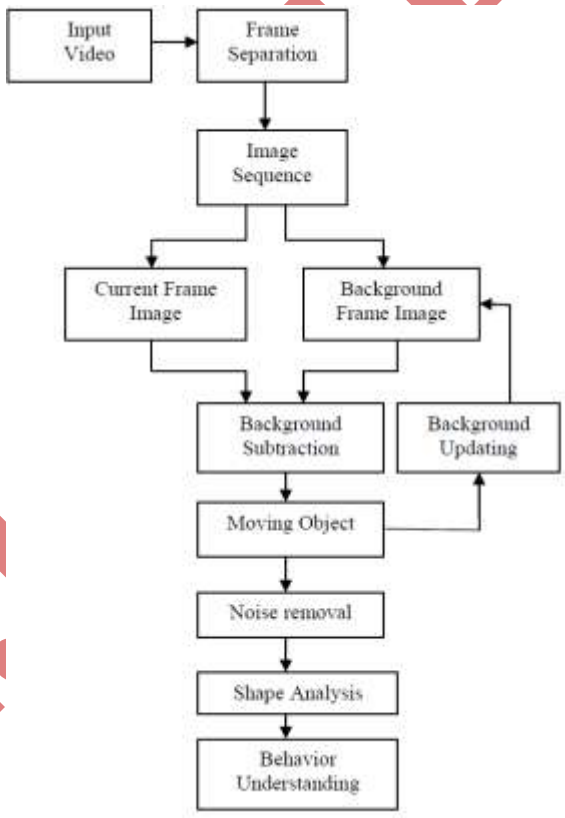

Figure 2: Flow chart for background subtraction algorithm

1. The object area is larger than the set threshold

2. The aspect ratio of the object region should conform to the set ratio.

If these two conditions are met, the moving object is the moving human body or not a human body is conform.

C. Noise Removal

Since the difficulty of the background, the discrepancy image obtained contains the motion region as well as large number of noise. These noises might be included in the image due to environmental factors, illumination changes, during transmission of video from the camera to the further processing. Therefore, noise needs to be removed. In this paper we adopts median filter with the 3 X 3 window for filtering noise.

As we know, motion region just not includes human being, but also it may include moving cars, flying birds, flowing clouds and swaying trees and other non body parts. Morphological methods are used for further processing. Corrosion operation is taken to effectively filter out non-human activity areas and by using the development operation they can filter out most of the non-body motion regions while preserving the shape of human motion without injury. After expansion and corrosion operations, some inaccessible spots of the image and some intrusion of small pieces are eliminated, and we get more accurate human motion region.

D. Extraction of Moving Human Body

Some accurate edge regions will be got after median filtering, corrosion and expansion operations, but the region belongs to the moving human body could not be determined. Through inspection, we can find out that when moving object appears, shadow will appear in some regions of the scène. Accurate mining of the moving object affected by the presence of shadow. By analyzing the characteristics of motion detection, we combine the projection operator with the previous methods.

Based on the results of the methods above, height of the motion region will get detected by adopting the method of combining horizontal with vertical projection. This can eliminate the impact of the shadow to a certain level. Then we analyze the vertical projection value and set the threshold value to remove the pseudo-local maximum value and the pseudo-local minimum value of the vertical projection to determine the number and width of the body in the motion region, we will get the moving human body with precise edge. This paper assumes that people in the scene are all in upright-walking state.

E. Behavior Understanding

After successfully detecting the moving humans from one frame to another in an image sequence, the problem of understanding human behaviors from image sequences follows naturally. Behavior understanding involves action identification and description. Human behavior understanding can guide the development of many human motion analysis systems. Behavior understanding is to analyze and recognize human motion patterns, and to produce high-level description of actions and interactions. The behavior understanding will be the most important area of future research in human motion analysis.

F. Kalman Filter

The filter is used to predict the track's location in each frame, and determine the likelihood of each detection being assigned to each track. A Kalman filter estimates the state of a dynamic system recursively at each time in the linear minimum mean square error sense. The motion of each track is estimated by a Kalman filter. The filter is used to predict the track's location in each frame, and determine the likelihood of each detection being

assigned to each track. Kalman Filter operates by Predicting the new state and its uncertainty and Correcting with the new measurement.

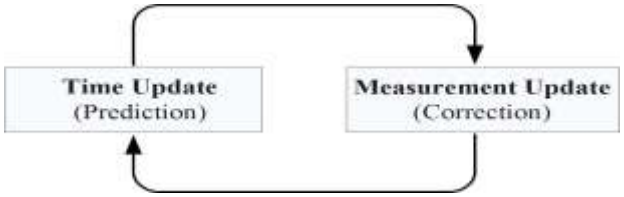

Figure3: Prediction and Correction updating

IV. RESULTS:

The below given results shows how to detect and track people in a video sequence with a stationary background using the following process:

1. Use the first few frames of the video to estimate the background image.
2. Separate the pixels that represent the people from the pixels that represent the background.
3. Group pixels that represent individual people together and calculate the appropriate bounding box for each person.
4. Match the people in the current frame with those in the previous frame by comparing the bounding boxes between frames

Figure 4 a) Original or input video (b) Background Image (c) Detected image (d) Tracked image

V. CONCLUSION

The system presented herein is successful as a robust detector of human motion. The proposed method has described a technique for tracking the movement of Humans in consecutive video frames. The proposed method of moving object detection will help to find the moving object perfectly in the approved manner. To minimize or avoid

the problems approaching in moving object detection, we use threshold method to detect moving object, background initialization and update the background in real time. At last, shadow effect removed by combining projection analysis with shape analysis. This method has also a very good effect on the elimination of noise and shadow, and be able to extract the complete and accurate picture of moving human body.

VI. FUTURE IMPROVEMENT

In this paper, we have considered static background; in future it can be improved for changing/non-static background.

REFERENCES

[1]Neeti A. Ogale, "A Survey of Techniques for Human Detection from Video".

[2]Prithviraj Banerjee Somnath Sengupta, "Human Motion Detection and Tracking for video Surveillance".

[3]Xiaofei Ji, Honghai Liu, "Advances in View-Invariant Human Motion Analysis: A Review," in IEEE Trans.on Systems, Man, Cybernetics, vol. 40, no.1, 2010.

[4]Murat Ekinci "Silhouette Based Human Motion Detection and Analysis For Real Time Automated Video Surveillance, "in Turk J Elec Engin, vol.13, no.2, 2005.

[5]HaziWang and David suter,"Background Subtraction Based on a Robust Consensus Method," Monash University, Clayton Vic. 3800, Australia.

[6]SumerJabir, Zoran Duric, Harry Wechsler, Azriel Rosenfeld "Detection and Location of People in Video Images Using Adaptive fusion of color and edge information."

[7]Jing Li, Zhaofa Zeng, Jiguang Sun and Fengshan Liu, "Through Wall Detection of Human Being's Movement by UWB Radar," in IEEE Geoscience and Remote Sensing Letters,Vol.9, no.6, Nov 2012.

[8]EyupGedikli,"Human Motion Detection, Tracking and Analysis for Automated Surveillance".

[9]Wang Weiqiang, Yang Jie and Gao Wen, "Modeling Background and Segmenting Moving Objects from Compressed Video," Circuits and Systems for Video Technology, IEEE Transactions on Vol.18, Issue 5, May 2008, pp.670 - 681.

[10]Qi Zang, Reinhard Klette "Object Classification and Tracking in Video Surveillance."

Biographical Notes

Mr. S. Govardhan Naidu is presently pursuing M.Tech. Final year in Electronics & Communication Engineering from Malla Reddy Institute of Technology and Science, Hyderabad, India.

Ms. S. Srilakshmi is working as a Assistant Professor in Electronics & Communication Engineering Department, from Malla Reddy Institute of Technology and Science, Hyderabad, India.