

## Human Posture Recognition: Background Subtraction and Skeletonization

Rujuta Panvalkar<sup>1</sup>, Ami Vashi<sup>2</sup>, Ruhina Karani<sup>3</sup>

<sup>1</sup>(Computer Department, Dwarkadas J. Sanghvi college of Engineering/ Mumbai University, India)

<sup>2</sup>(Computer Department, Dwarkadas J. Sanghvi college of Engineering/ Mumbai University, India)

<sup>3</sup>(Computer Department, Dwarkadas J. Sanghvi college of Engineering/ Mumbai University, India)

---

**Abstract:** With the advent of new and improved technology in the market, it has become extremely important to gather information and create user friendly applications. Be it for security purposes or tracking purposes, monitoring human activities to create a human gesture-friendly interface is now a major branch of research. We hope to present one such method to observe and interpret human activities. This paper describes methods to obtain human activity records (in the form of images or videos) and perform suitable actions on these records to extract useful information out of them. Background subtraction is the initial one of the most important procedures performed on the image or video. Next comes preprocessing the record to improve quality. Then finally comes Skeletonisation. We focus on Background Subtraction method using the image pixels and Skeletonisation in this paper and provide the basic steps and strategies regarding it.

**Keywords:** Background Subtraction, Human activity recognition, Image Pixels, Pre-processing, Skeletonization

---

### I. Introduction

Human activity recognition has gained a lot of importance in the last few decades. Issues related to medical care, home security and video surveillance can be addressed using human activity recognition [1],[2]. For this, posture recognition and comprehension is an important phenomenon. Humans perform an array of diverse actions that command multiple postures. Activities like sitting, walking, running, sleeping undertaken by all humans in a similar fashion unless an impediment hinders them from following the natural course. For instance, while sleeping, a body will lie close to the ground and in a still mode. On the other hand, while running, a body will be up on the ground and will move at a constant speed. These distinctions aid in classifying the diverse structures that humans exhibit. The aim is to take footage of a human in a controlled environment, either through video surveillance or by a series of continuous photographs. This footage will be examined frame by frame and a series of progressions will be applied to each frame [3].

In this paper, we suggest methods to carry out the fundamental steps in posture recognition. Every movement the observed human makes, is captured in the video. There are a number of ways to separate and comprehend these movements. For every image that gets processed, it is necessary to initially perform the process of background subtraction. In this, the image is distinguished into two constituents- foreground and background. The foreground contains the human figure that is to be comprehended. The background, being the source of noise and the trifling part, is removed. Many Background Subtraction Algorithms have been designed up till now [4]. The process of Background Subtraction can be carried out by a simple Matlab code, but at the cost of blurring the foreground. In this paper, we have suggested a novel approach for doing so using the pixels distinguishing technique. In the next stage, pre-processing is suggested. It is necessary to remove the low frequency noise, to normalize the intensity of the image, to remove reflections. It can be used to convert an image to a grey scale image. Pre-processing an image makes it ideal for the further computational manoeuvres. After these initial steps of background subtraction and pre-processing the image is now ready for skeletonization. In skeletonization, the human image is reduced to an image border thus giving it the shape of a skeleton, hence the name skeletonization. The skeleton of the image is retained and the rest of the image is discarded. It is this skeleton that is used further for the classification of the various human activities described above. Many skeletonization algorithms have been researched till now [5], [6]. In this paper we mainly discuss the star skeletonization algorithm in detail. This algorithm considers the body extremities such as hands, legs, head and forms a star like pattern when joined with the centre of the body. Our paper focuses on this process as it is an interesting and integral part of image recognition. Later, Hidden Markov model can be used to classify all these postures [7].

The paper is divided as follows- Section [2] gives an overview of the Background Subtraction algorithm. Section [3] talks in detail about the process of star skeletonization. Section [4] concludes the paper with a summary and course of future works.

## II. Background Subtraction

We propose to perform Background subtraction using a three step method .

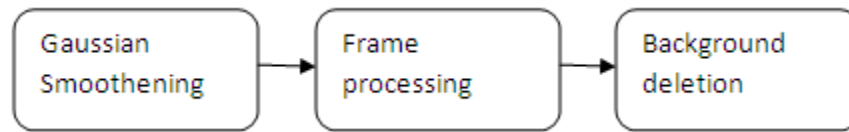


Fig. 1 Three step method for Background subtraction

### 2.1 Gaussian smoothening

Now a days most of the images are a myriad of gaussians the best way to normalize the image is to apply Gaussian smoothening. This aids in lessening the noise and limits the sharp vicissitudes in pixel density. Due to gaussian smoothening, fewer edges are detected. Additionally, even though the image is slightly blurred, the pixels with sharp intensity difference (human figure and the background) are substantially differentiated.

### 2.2 Frame Processing

After smoothening, the process of frame separation is applied to the footage. Multiple frames are computed and pre-processing is done on each individual frame. As the camera is a static that is, not moving, the background is assumed to be majorly static with the exception of a few dynamic apparatuses like fans, windows, etc. The major hurdles in the relatively easy to scrutinize static environment are the luminous changes i.e the brightness variations. To cope with these abnormalities, the arithmetic mean of the pixel brightness over a predetermined period of time is taken and is considered as standard for that particular experiment. Now, for the fairly static background, it will remain almost the same over the period of time.

$$P_B = P_{Br} + P_{Bb} + P_{Bg} \quad (1)$$

Where  $P_B$  = Pixel brightness of a particular pixel in a single frame.  $P_{Br}$ ,  $P_{Bb}$ ,  $P_{Bg}$  are the Pixel brightness of RGB vectors of the pixels.

### 2.3 Background Deletion

After adjusting the pixel brightness, the foreground image is now ready to be extracted. In other words the background can now be eliminated allowing us to obtain just the image of the human that we wish to observe[8].

Background deletion is a two-step process.

#### 1. Separation of background and foreground

Each pixel is given its semi-permanent value as follows

$$P_{Bf} = \mu P_{Br} + \mu P_{Bb} + \mu P_{Bg} \quad (2)$$

Where  $P_{Bf}$  = final Pixel brightness of the pixel and  $P_{Br}, P_{Bb}, P_{Bg}$  = mean RGB pixel brightness of the pixel calculated over a range of N frames.

A threshold value is set to distinguish background pixels from the foreground pixels. If the change in pixel colour is more than the set threshold, then the pixel is said to be in the foreground; else it lies in the background. Now we can easily distinguish between background and foreground.

#### 2. Deletion of background

For deletion of the background, each consecutive frame is compared with the threshold value and the pixels which lie below the determined threshold are deleted from the frame. In this process, the background which is relatively stable throughout is discarded, leaving behind the dynamic human figure i.e the foreground. Now we finally have the foreground image ready for skeletonization.

All these above processes are applied iteratively over a series of frames to get an array of processed frames from a running video footage.

## III. Skeletonization

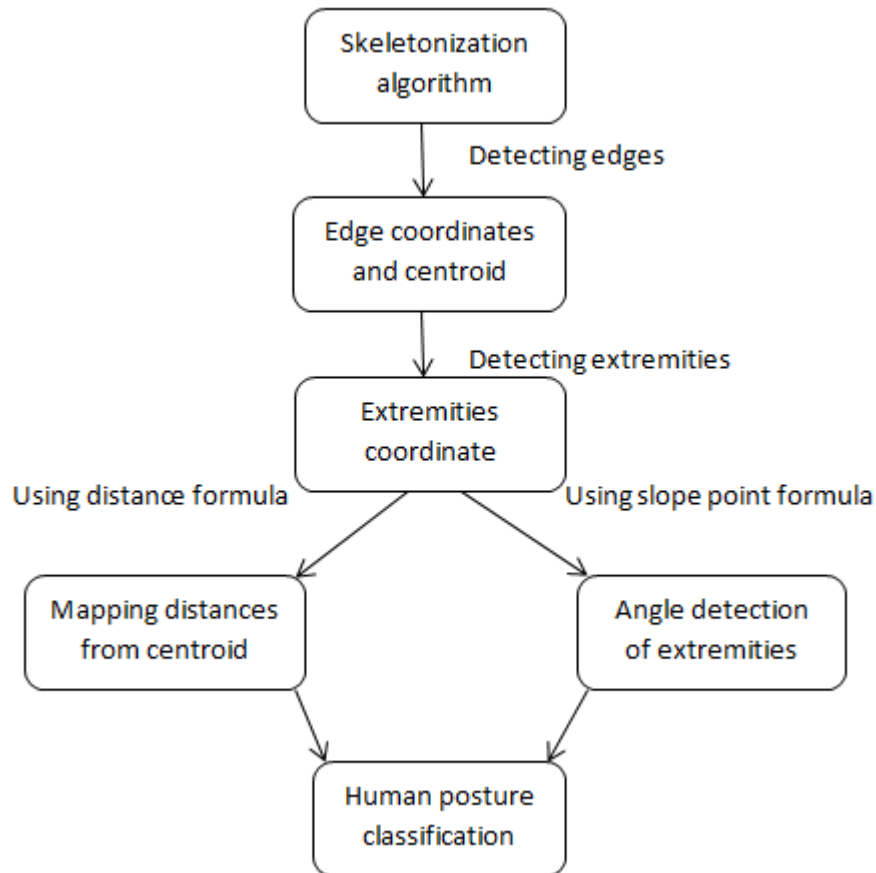
Skeletonization is a very important task when it comes to human activity recognition. Skeletonization is used to outline the dynamic human figure and to provide detailed insight in the mechanics and working of the figure. The method we suggest here is the Star Skeletonization.

Star skeletonization works by identifying the coordinates of the edges of the human skeleton and mapping them with respect to its centroid. The extremities (head, hands and feet) of the human form are then analysed and are joined with the centroid. This depiction forms a star shape. The distance of each extremity

from the centroid is calculated and is utilized for further processing. We do carry forward from the details explained in [9].

In addition to the above information, the following information is also tracked using skeletonization-

1. The height and width of the human figure. This is used to calculate the aspect ratio which is defined as the width divided by the height.
2. The angles between the extremities.
3. The vertical and the horizontal gradient.



**Fig. 2** shows the flowchart of the proposed Skeletonization process

The human figure detection can be done by classifying it into an array of positions [10]. The five main positions are explained as follows-

- a. Standing- The aspect ratio is less than 1. Both the vertical and horizontal gradient remain constant to a large extent.
- b. Sitting- The aspect ratio can range from 0.5 to 1. The horizontal gradient changes when the person shifts from a standing position to a sitting position.
- c. Standing up(from a sitting position)- The aspect ratio becomes less than 1. The horizontal gradient changes while the vertical gradient remains almost constant.
- d. Falling down- The aspect ratio increases and becomes more than 2. The vertical gradient Increases
- e. Sleeping- The aspect ratio is more than 2.

Skeletonization can be used for angle detection between the extremities.

- a. Running- The angle between the leg extremities will change fast and will always be less than 90 degrees about the centroid.
- b. Walking- The angle between the leg extremities will change at a slower pace and will always be less than 90 degrees about the centroid.
- c. Standing- There will be no change in angles of the leg and the angle measure will remain limited to less than approximately 30 degrees.

- d. Falling- The angle of head with respect to the body becomes less than 180 degrees. The vertical gradient of the body changes drastically over the gamut of consecutive frames. The aspect ratio of the body reaches between 0.5-1 (bending) and eventually becomes more than 2 (flat on ground). Over the next series of frames, all the components remain fairly constant.

### **3.1 Advantages of Star Skeletonization.**

1. It does not implement an iterative continuous computation. This makes this method computationally cheap. It accentuates the time factor and the hence has a decent time complexity. The only computation is detection of the boundary point and centroid coordinates and then employing the standard distance formula which reduces the computation complexity as well.
2. Star skeletonization just requires the boundary pixels and not the entire human body image. This greatly diminishes the number of pixels involved in the computation process. Thus saving us a lot of memory.
3. It gives information about the position of the extremities as well as the angles between them for distinguishing the activities like spinning, rotating, etc.

## **IV. Conclusion**

This paper describes the detailed procedure of human activity recognition. Where in the extraction of the foreground image is efficiently carried out using background subtraction. Furthermore, skeletonisation is performed to determine the human posture. Emphasis is placed on the method of star skeletonisation. Human Activity recognition is broadly used in areas where elderly people are needed to be monitored around the clock. This procedure provides a form of health care which was not available in earlier times. The existence of smart environments has made it possible to achieve this feat. Another use of posture recognition is in Human-Computer Interaction systems which is highly robust and has a very high level of recognition capacity.

The applications of this system are- 1. Assistance for elderly people living alone. 2. Can be used for scanning at security agencies 3. Human gesture recognition for blind people. 3. Can be used for detecting future human activity along with the present activity detection. The applications of this subject are wide and apart. Much research needs to be done. The main advantage of using the approach explained in the paper is that it is a very easy, computationally cheap and universal approach.

Further work needs to be done to create a system where real-time postures and gestures are recognised with minimum delay. Faster and better systems need to be developed which have faster computation and lower computational complexities. These systems need to show superior results when applied to several video samples. Work needs to be done to overcome problems which are caused due to occlusion of certain parts of the body. Partial occlusion or hiding of certain part of the body may lead to confusion regarding the posture of the body. This confusion needs to be eradicated.

## **Acknowledgements**

We thank our college, DJ Sanghvi College of Engineering and the department of computer engineering for their constant support in terms of the resources provided, and some valuable advise that helped us throughout us research.

## **References**

### **Journal Papers:**

- [1] V. Vishwakarma, C. Mandal and S. Sural , Automatic Detection of Human Fall in Video, Proceeding PRMI'07 Proceedings of the 2nd international conference on Pattern recognition and machine intelligence, pp. 616-623.
- [2] Wadley, V., Okonkwo, O., Crowe, M., Ross-Meadows, L. A. 2007, Mild Cognitive Impairment and everyday function: Evidence of reduced speed in performing instrumental activities of daily living. *American Journal of Geriatric Psychiatry*, 16(5): 416-424.
- [3] L. Chen, C. Nugent, J. Biswas, J. Hoey, Activity Recognition in Pervasive Intelligent Environments, Chapter 1 , Activity Recognition: Approaches, Practices and Trends, ATLANTIS PRESS, 2011, 1875-7669
- [4] M. Piccardi, Background subtraction techniques: a review, 2004 IEEE International Conference on Systems, Man and Cybernetics, vol.4, 2004, pp. 3099- 3104.
- [5] Sungkuk Chun, Kwangjin Hong, and Keechul Jung, 3D Star Skeleton for Fast Human Posture Representation,. *International Science Index, World Academy of Science, Engineering and Technology Vol:2 2008-08-27*
- [6] Deenbandhu Singh, Akhilesh Kumar Yadav, Vivek Kumar, Human Activity Tracking using Star Skeleton and Activity Recognition using HMM's and Neural Network, *International Journal of Scientific and Research Publications*, Volume 4, Issue 5, May 2014.
- [7] Rabiner, L. and Juang, B. H. 1986., An introduction to hidden Markov models, *IEEE Acoust. Speech Signal Proc.*, Mag. 3, 4-16
- [8] pande r.p.\*, mishra n.d., gulhane s. and joshi a., Detection of moving object with the help of motion detection alarm system in video surveillance, *Journal of Signal and Image Processing*
- [9] H. Fujjyoshi, A. Lipton, Real-time human motion analysis by image skeletonization, *Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision*, WACV '98, 1998, pp. 15-21.
- [10] Monica-Andreea Dragan, Irina Mocanu, Human Activity Recognition in Smart Environments, Published in *Control System and Computer Science(CSCS)*, 2013 19<sup>th</sup> International Conference, 2013