

# Semi-Supervised Human Action Recognition Using Modified Sift Algorithm And Decision Based Median Filter

Vamsi Krishna <sup>1</sup>, Smreeti Singh <sup>2</sup>, Mayank Raj Gupta <sup>3</sup>

Associate Professor, Department of Information Science and Engineering, Sri Krishna institute of technology,

Bangalore, Karnataka India<sup>1</sup>

U.G. Student, Department of Computer Engineering, Sri Krishna institute of technology, Bangalore, Karnataka, India<sup>2</sup>

U.G. Student, Department of Computer Engineering, Sri Krishna institute of technology, Bangalore, Karnataka, India<sup>3</sup>

**ABSTRACT:** Human activity identification is one of the most explored areas in the field of computer vision. The action judgment which ought to be learned from the input videos and the images which are still. The focus of the activity recognition is an automatic analysis of ongoing events and their context from video data. Application of human activity identification includes surveillance system, patient system and other systems which include interaction between human and the computer. Many program needs applications of high level activities, consisting of many simple action of human. In traditional method this is done using SIFT algorithm which is having low time complexity. To overcome the problem of Motion tracking in sophisticated space we adopted modified scaled invariant feature transform (MODIFIED-SIFT) as well applied modified decision based median filter (DBMF) that takes into consideration all those pixel consisting of noise and hence replaces the value of the pixel with that of the median value in the processing window this method works extremely fine with high noise density

**KEYWORD:** Modified-sift, modified decision based filter, noise density, and time complexity

## I. INTRODUCTION

Image processing is a way of performing basic operations on an picture, for an enhanced picture and to extract important features from the picture. Image processing is a part of signal processing in which image is loaded and produces either image or characteristics which are related to that particular picture. Here are two sorts of strategies utilized for picture preparing in particular, simple and advanced picture handling. Simple picture preparing can be utilized for the printed copies like printouts and photos. Picture examiners utilize different essentials of understanding while at the same time utilizing these visual procedures. Advanced picture preparing systems help in control of the computerized pictures by utilizing PCs. The three general stages that a wide range of information need to experience while utilizing computerized procedure are pre-handling, improvement, and show, data extraction. Human activity examination assumes a basic part in human focused registering, as in numerous applications we require to break down and comprehend human activities in the enormous video information, for example, video observation, social insurance and human- PC collaboration. Picture preparing essentially incorporates the accompanying three stages: Bringing in the picture by means of picture procurement instruments, Dissecting and controlling the picture and Yield in which result can be modified picture or report that depends on picture examination.

## II. RELATED WORK

### I. MODIFIED SIFT ALGORITHM

Cong Yang, Liu Waiyu and Zhaz Yauli's paper research of video tracking based on improved SIFT algorithm deals with a better algorithm which helps in feature extraction of details from individual image. The scale invariant feature

transform commonly known as SIFT was first devised by David lowi in 1999. There are 3 parts in which this algorithm works.

**IA. EXTREME POINTS DETECTION IN SCALE SPACE** In this 2d Gaussian kernel formula is taken from  $G(x,y,\sigma)$  where  $\sigma$  is the standard deviation. This important part is Gaussian blur. In real world Scale space is replicated as Gaussian pyramid. To get the extreme point in scale space we have to get difference of Gaussian [DOG] function which helps in finding extreme point.

**IB. MAIN DIRECTION DETERMINATION OF THE FEATURE POINT** For producing the descriptor rotation invariance, key point will be given using feature of an image. the magnitude and direction is given by DOG represented as  $n(x,y)$  and  $\Theta(x,y)$ . After this step, key point descriptor is produced with the help of magnitude and direction. for better matching result only 80% peak is saved as auxiliary direction of feature point.

**IC. FEATURE POINT DESCRIPTION** By performing the above two steps we will have scale, position, direction for every single keypoint. the descriptor vector for feature point will have 128 dimension. The descriptor contains feature point and neighbouring pixels too. It should have great uniqueness for matching accuracy to be very high.

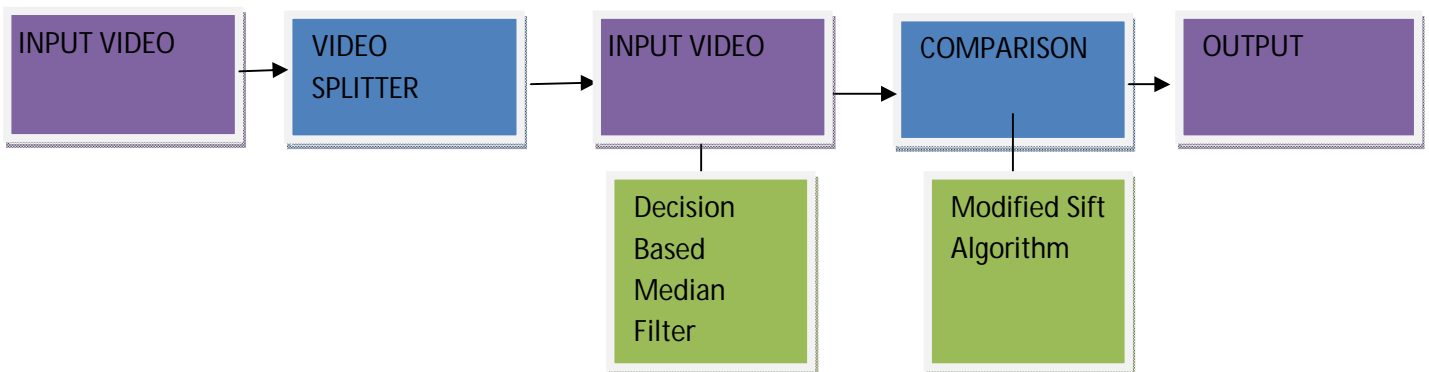
What's better about the M-SIFT algorithm is that the video which will be given for tracking can have both object and background moving. This shows how vigorous the algorithm.

### III. DECISION BASED MEDIAN FILTER FOR NOISE REMOVAL

In Raghuram Kunsoth and Mantosh Biswas paper on Modified Decision Based Median Filter for Impulse Noise Removal explains about One of the most concerning issue in computerized picture preparing is disposing of the clamour from pictures. The pixels in the given picture can get debased while transmission in view of noise. This paper introduces an adjusted choice based middle channel for taking out commotion. There are different diverse channel for this, for example, Standard median filter (SMF), Adaptive median filter (AMF), weighted middle filter (WMF) and Decision based middle filter (DBMF).

In SMF, AMF and WMF the pixels gets exchanged by the median value in their surrounding area without caring about local feature of the image. In this proposed System the decision based median filter will only care about noisy pixel and replacing its value with median value present in the processing window. In this process window size is made variable and can be altered based on requirement. This method is much better than other median filter and works with higher accuracy when the intensity of noise is more. This modified algorithm solves two problems which are present in the present techniques. The first one is the creation of blurred image when the size of window is increased and other one is below average quality of the picture when the pixel are corrupted in a small Window. this method works with high density of noise and solves the problem of poor quality images where the pixels are corrupted by processing as much of the image details.

### IV. PROPOSED METHODOLOGY



The proposed system here consists of five different stages. The elements here can be explained as below

- A. Input video: A 10 sec video is fed as an input to the given system. The video which is feed may or may not contain any noise. The given video will be further used for feature extraction. The given video can be of any format as described.
- B. Video Splitter: a video splitter here is used in order to separate the given video into smaller frames and further into still images. Each individual image is called a frame which is measured as frames per second (FPS). There are various split which can be incorporated into the system for separating the frames accurately.
- C. Filter: This is one the most important part of the given system. Here a decision based Median Filter is Used so in order to remove even very high density of noise. The split image is passed through this given filter. The poor quality of the output image is resolved even when the given pixels are corrupted in a small window. This filter works efficiently with large noise density and preserves as much image details as possible.
- D. Comparison: This is the part in the system where the action recognition takes place. Once the given images are passed through the given filter, noise removed images are compared with each other. Activity recognition can be ordered as: stream based methodologies, spatio-fleeting shape layout based methodologies, following based methodologies and intrigue focuses based methodologies. In stream based methodologies optical stream calculation is utilized to portray movement, it is delicate to clamor and can't uncover the genuine movements. Spatio-transient shape layout based methodologies regard the activity acknowledgment issue as a 3D question acknowledgment issue and concentrates highlights from the 3D volume. The removed highlights are exceptionally tremendous so the computational cost is unsuitable for ongoing applications. Following based methodologies experience the ill effects of similar issues. Intrigue focuses based methodologies have the upside of short element vectors; thus low computational cost. The algorithm used hence reduces time complexity as well as compares the image faster than SIFT approach.
- E. Output: The given system after comparing produces the result about the recognized feature of the given human action. It explains about what the Human is doing in the video.

## V. CONCLUSION AND FUTURE WORK

This work exhibits a human activity acknowledgment framework that is quick and basic. The framework is made out of five phases: Taking the input, passing it through the video splitter, then the given frame through the filter, and through the comparison part and finally getting the output. The 3rd part works on decision based median filter and the 4th one is based upon Modified scale invariant transform The proposed standardization technique works faster than the given traditional approach. Future work incorporates applying the proposed framework on various complex datasets, for example, games and genuine activities ones. These datasets are more perplexing than the ones utilized here and the framework may require a few changes to accomplish satisfactory acknowledgment rate. Likewise the utilization of an arrangement of various activities to fragment it at that point perceive each activity is another purpose of research later on work

## REFERENCES

- [1] R. Cai, Y. Wu, M. Wang, and Q. Wu, "Survey of visual object tracking algorithms," Video Application & Project, vol.34, no.12, pp. 135-142, December 2010.
- [2] T. Cao, J. Liu, L. Zhang, W. Pu, P. Yu, Y. Liu, C. Zhao, "Implantation and application of radio transmitter for wild tracking Burmese Python," Natural Science Journal of Hainan University, vol. 32, no. 4, pp. 363-367, December 2014.
- [3] R. Cai, K. Lei, X. Zhang, Y. Wang, "Design of video tracking simulation platform based on .Net," Computer Simulation, vol. 24, no. 12, pp. 181-184, December 2007.
- [4] Y. Cao, "Research on application of face recognition technology in video surveillance system," Southwestern University of Finance and Economics, March 2012.
- [5] B. Yang, "Vision-based navigation research for aircraft approaching and landing," Shanghai Jiao Tong University, February 2013
- [6] Raghuram Kunsoth, Mantosh Biswas, "Modified Decision Based Median Filter for Impulse Noise Removal," Wireless Communications, Signal Processing and Networking (WiSPNET), International Conference on 2016