iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5106

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Human Activity Recognition using OpenCv & Python

Mr. Siddharth Sharma[1], Mr. Abhay Kumar[2]
[1, 2]*Galgotias University, Gautam Buddh Nagar, Uttar Pradesh*

Abstract: Human activities recognition has become a groundwork area of great interest because it has many significant and futuristic applications; including automated surveillance, Automated Vehicles, language interpretation and human computer interfaces (HCI). In recent time an exhaustive and in depth research has been done and progress has been made in this area. The idea of the proposed system is a system which can be used for surveillance and monitoring applications. This paper presents a part of newer Human activity/interaction recognition onto human skeletal poses for video surveillance using one stationary camera for the recorded video data set.

The traditional surveillance cameras system requires humans to monitor the surveillance cameras for 24*7 which is oddly inefficient and expensive. Therefore, this research paper will provide the mandatory motivation for recognizing human action effectively in real-time (future work). This paper focuses on recognition of simple activity like walk, run, sit, stand by using image processing techniques.

Keywords: Human Activity Recognition; Human computer Interface; Surveillance and Monitoring.

I. INTRODUCTION

Recognition of the actions involving a person from a video is the objective of action recognition. The primary objective of our underlying methodology is to improve the accuracy and meanwhile add some prediction models of tour prediction.

A. There Are Four Levels Of Understanding In Human Activity Recognition (HAR) Specifically

1) *Object-level:* Object-level recognizes the location and clout of the object.
2) *Tracking-level:* Tracking level recognizes the path and the avenues of the under observation object.
3) *Pose-level:* Pose level recognizes the pose of an Actor.(The newer models do not consider the change of the Actor as an issue using RGB-D models.)
4) *Activity-level:* Activity level recognizes the activity and the interaction of the Actor/Model.

B. The Current Problems Faced In Human Action Recognition (HAR) Are

1) *Camera View-point Anomaly:* movement of camera, Blurred Focus, Incomprehensible numerical data.
2) *Human Anomaly:* frequent changes in Actors due to difference in the shape and sizes,
3) *Spatial/action Anomaly:* Difference in actions performed by various actors in a state space,
4) *Temporal Anomaly:* unconformity in duration and shift, siege/blockage of main actor, Incomprehensible action due to non-visibility,
5) *Background:* shifty and irrelevant objects in background In our paper we present a technique which removes the above mentioned problems by using the concept of optical flow and Hidden Markov Model(HMM) instead of the Gaussian Markov Model.

II. LITERATURE REVIEW

Ren and Xu (2002) [1] presented a new system for Natural & complex Human action recognition in the smart classroom environment in place so as to realize incomprehensible cameraman and virtual mouse which will make it convenient to gather visual data . First, the system projects a Dynamic human model and makes use of a second order B-spline function to detect the body parts such as the two shoulder joints in the silhouette/Profile/delimited image to extract the basic motion norms/features including the motion of hands, parameters of legs, Facial action. Then, a primitive-based coupled hidden Markov model is projected for natural action recognition which is dependent on some context or rather context-driven. Last, some comparison experiments show that PCHMM is better than the Hidden Markov Model that we are using here.

The PCHMM is also better than the coupled Hidden Markov Model. Akilandasowmya and P.Sathiya [2] explain that HAR is a vital research area tilting to the visual field of computers.

There is a technique to discern and extract the definite events in Videos. Low level video contents(frames) are turned/translated to high-level video(frames) to sequence content is an interesting & growing research topic in recent years. Its application includes automated video surveillance schemes, intensive care units in hospitals(ICU), airports security checks and scans, analysis of physicals of people in high security zones, they even include H-CI or human computer interfaces.

Arie et al. (2002) [3], projected a new method for Viewpoint -based HAR using videos. By keeping a small number of frames under observation, the identification of the valid activity could be extracted. The Motivation of multidimensional indexing technique helps in certainly identifying the action/interaction of a human from a sporadically sampled sequence of skeleton structure poses of humans acquired from videos. Davide Anguita et al.[4] presented an Activity-Based Computations aimed to seige the state of the user and its surrounding domain by making use of heterogeneous sensors so as to provide modification to external computational resources. When these sensors are tended to the Actors skeletal structure, they allow for perpetual monitoring of Actions/signals physical in nature surmounting to a logic behind the actions.

III. METHODOLOGY

Human activity Recognition can be done using one of the 2 techniques.

1) Template Matching Technique: The template matching technique convert an picture(image) sequence into a static shape pattern here instead of using GMM we will use HMM(Hidden Markov Model and optical flow For defining the sequence of the data in the separated frames.) and then compare the value of the static picture with that of the values previously stored in the trained data-set, when the value of the data set matches the value of the data the blobs displays the derived result. The advantage of using the template matching procedure is that it takes less computational power of the system but it is still reactive to the temporal anomaly discussed above.

2) State-Space Model defines each Stationary static pose as a single state. This stationary pose is relevant to each frame formed by HMM These states are connected by certain Possibilities such as the activities will all have a predefined number and other activities surrounding that number will form a chain of events likely to happen and hence increasing the probability of recognition and also making prediction a reality. Any motion sequence taken into account as a tour going through these states. Joint expectation is to be calculated through all these tours and the value cost maximum and closest to the values in the data-set is chosen as the criteria for classifying activities. In such a scenario, temporal anomaly of motion does not raise any issue because each state on loop visits itself in repetition. Hence this method of state-space The model is reliable against temporary anomalies. below are the broad steps of the projected technique :

a) Pre processing
b) Feature Extraction
c) Human activity Recognition

Figure 3.1 Broad Overview of the Technique

IV. IMPLEMENTATION

This whole model is based on python openCv2 (CvHMM version) which makes use of the Hidden Markov Model. The pretrained data set is taken into consideration from Microsoft Kinects which in broad sense just involves basic movements. Also the KTH data set seemed useful so we made an adjustment to even use it. In the Specificity of the order of the technique we have around 5 steps which are mentioned below.

1) Grabbing Video
2) Preprocessing
3) Construction of frames
a) Testing frame (in accordance to data set)
4) Feature Extraction
5) Human Activity Recognition
a) Classification of human activity(in acc. To data-set)

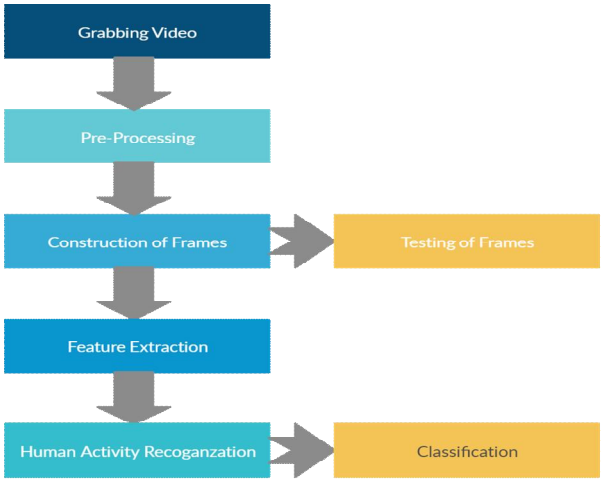


Figure 4.1 In-depth Overview Of the Proposed Technique.

A. Video Grabbing

The video data from the dataset or recorded surveillance vedios is taken into consideration. It is a finding that if the data is supervised the results will be better than that of the unsupervised data(video).

B. Pre-Processing

The process leads on with the first step of importing necessary packages of numpy, argparse, imutils, sys, opencv2, after which the construction of the argument parser to parse the arguments takes place, using cvHMM version will eventually provide us with the preconfigured code settings for the dataset.

C. Construction of The Frames/blobs

2D blobs are the most commonly used feature (low level) for recognition of human activity, that is why we generally come across it as the first stage. The dilution in blob is for the enhancement of the frame, dilution can be done easily via masking or by applying a filter it is only after dilation that we obtain a 2D blob. Blob segments the frame (here we are taking one frame of the video set in consideration) into foreground and Background & the net median numerical video. Blobs are multidimensional arrays or data.

1) *Testing Of The Frames:* after loading the contents of the class label, it is advisable to define the sample duration that is defining the number if frames for classification and sample size just to save the computational costs. loading it into human activity recognition model in order to test the data, after this it would provide a better gui experience for the user as well.

D. Feature Extraction

After the classification of the segments in the blob the next stage in the sequel is of feature extraction, here the numerical median of the blob in motion is taken into consideration as the value for the recognition of activity is best described by the blob rather than the colour or the size of the actor. Here the feature of" Motion/Activity/Movement" of the actor in the blob is done.

Here as previously mentioned to go from one video frame to another we use optical flow which is nothing but the usage of the HMM in between of the frames, following are the popular methods for finding optical flow

1) Horn-Schunck Technique

2) Lucas-Kanade Technique

Horn Schunck technique is used for floating point input & Lucas-Kanade for otherwise (I.e for fixed point input.) Here in this paper we have made use of the Lucas Kanade method.

E. Action Recognition

This in Sequence is after the 'Feature Extraction" where the activity/Movement which was the median numerical number of the blob is extracted, here then by using the optical flow of the Lucas kanade Method & also for human activity recognition we use Hidden Markov Model.

1) *Action Classification:* The "Activity/Motion/Movement" is classified due to the median of the blob which is then compared to the already stored numerical values of the pre-trained data-set. each activity has a corresponding numerical value to it, which when matched with the value given by the blobs results in itself classifying the activity in observation.

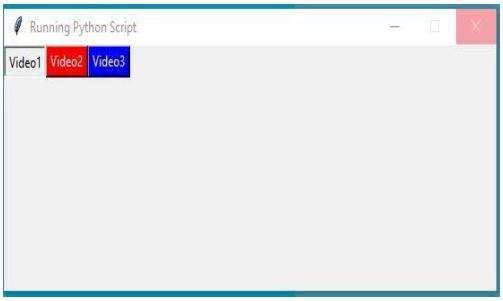


Figure 4.2 The GUI of the data (videos).

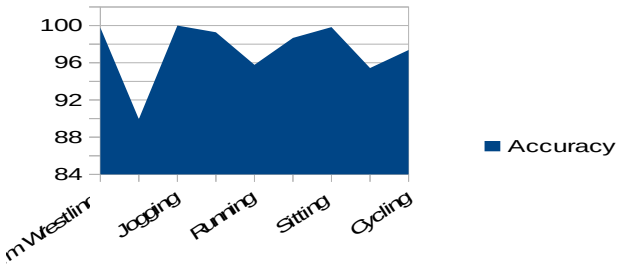
Figure 4.3 An example of how the Proposed model effectively recognizes the activity.

V. CONCLUSION

The approach provided by this paper is based for action recognition. It has a 2D blob using Lucas kanade method of optical flow . the motion parameters are transformed into symbol sequence using HMM

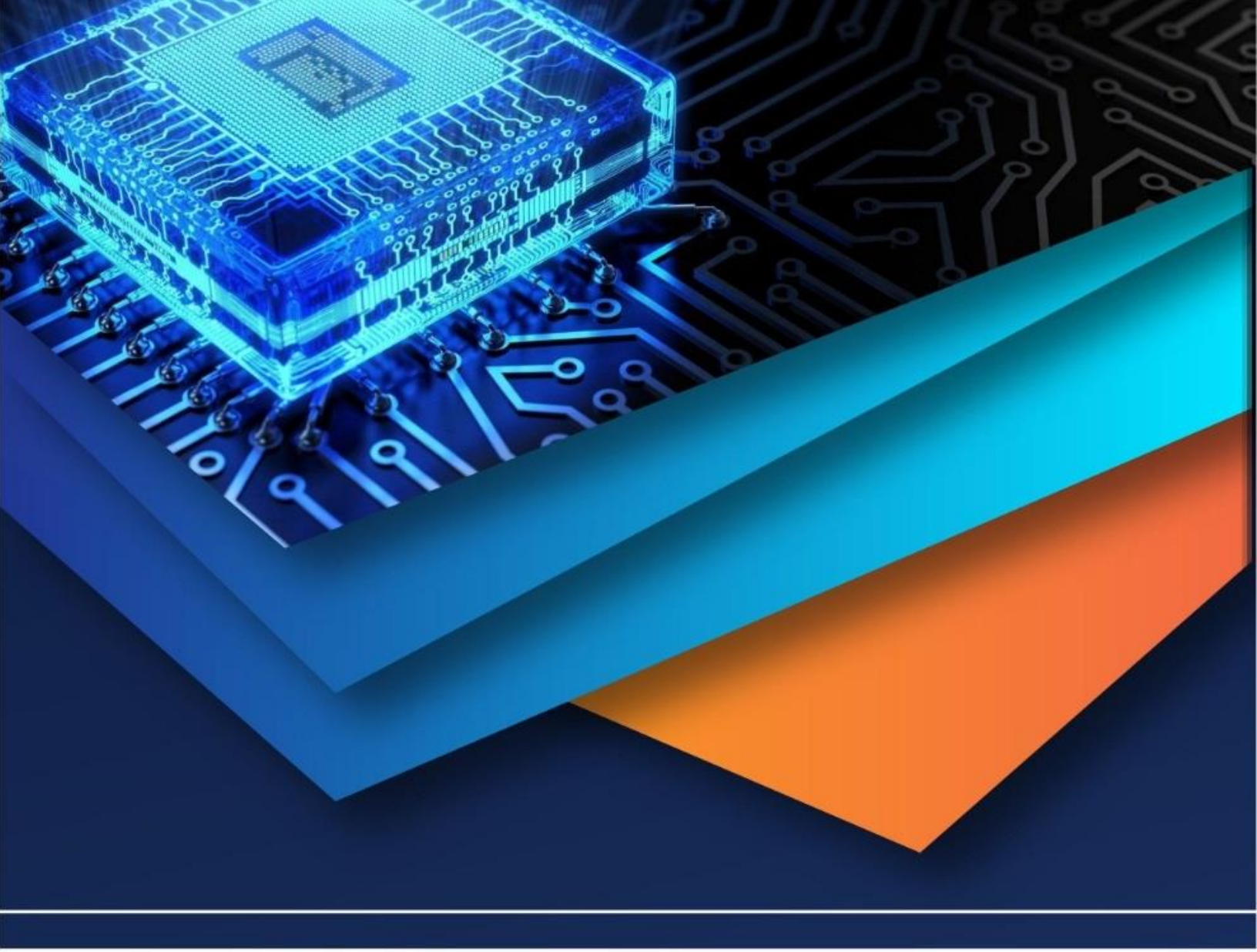
The Hmm is then trained to get the maximum likelihood of the model; this is selected as a recognition result. The average success using these data sets using this technique for various activities is given in the chart below.

Types of Sequence	Accuracy
Arm Wrestling	99.78
Boxing	89.94
Jogging	100
Cigarette	99.28
Running	95.78
Walking	98.67
Sitting	99.83
Drinking	95.43
Cycling	97.37



REFERENCES

[1] Guangyou Xu, Human action recognition in smart classroom , IEEE International conference
[2] G.Akilandasowmya, P.Sathiya, P.AnandhaKumar Human action recognition in research area of computer vision , IEEE International Conference on automatically detect and retrieve semantic events in video, 2015 Seventh International Conference on,15-17 Dec.2015
[3] Arie et al. Human activity recognition using multidimensional indexing IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 24,Issue: 8, Aug 2002.
[4] Davide Anguita et al. Human Activity Recognition on Smartphone's Using a Multiclass Hardware-Friendly Support Vector Machine volume 7657.nterfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY