Imperial Journal of Interdisciplinary Research (IJIR)
Vol-2, Issue-5, 2016
ISSN: 2454-1362, http://www.onlinejournal.in

Human Activity Recognition by using Bayesian Classifier on SURF Features

Sonia Sehgal[1] & Sonit Singh[2]
[1]Research Graduate, Department of Electronics and Communication Engineering
, [2]Assistant Professor ,Lovely Professional University, Punjab, India

Abstract: *In modern era, automation is done in every field to reduce the human efforts and manpower. So is the need of the systems able to process videos to retrieve necessary information out of it. Human activity recognition has many applications in many fields such as health care system, security systems etc. Human activity recognition is performed by human detection out of a video, feature extraction and activity classification. This paper presents a novel approach of activity recognition by background subtraction, feature extraction using SURF features and classification using Bayesian classifier.*

Keyword: *Background subtraction, Speeded up robust features, Bayesian classifier.*

1. Introduction

With increasing technology, no. of cameras installed everywhere, it is need of the hour that humans should be detected automatically in the video sequence and so is the recognition of human activities calling for human attention only when necessary. Here comes the video processing which is handling of image sequences out of video using mathematical operations. Human activity recognition contributes a lot in applications such as daily life activity monitoring, surveillance environments, human–machine interaction. Human activity recognition is done by three steps: human detection, feature extraction and activity classification.

In detection, human is detected out of the video and segmentation is achieved depending on whether camera is mobile or static. Nowadays methods used for segmentation are background subtraction[2], Gaussian mixture model and optical flow. If the background of a video is known, the entire object information could be known by just performing subtraction of current frame from the background frame. Optical flow method[1] is performed by doing clustering processing after computing optical flow field of subsequent frames.

Second, the human object is best represented by a set of features such as color, pose, silhouette etc. The features of detected human could be found by local descriptor algorithms such as HOG, SIFT, SURF and many more depending on space time volume, frequency and body modeling. Scale Invariant feature transform being invariant to scale and rotation represents the features by keypoint localization. Histogram of Oriented gradients is performed by taking gradient orientations and associated histogram bins. Speeded up robust features(SURF) provides the blob features associated with gray scale input image.

Third step is the activity classification which could be performed by different classifiers such as SVM, Random forest, Back Propagation Neural Network (BPNN) and Bayesian classifier. Support Vector Machine classifies the activities by obtaining the hyper-plane and support vectors. It works on spatio-temporal features. BPNN works by propagation in three layers: input layer, output layer and hidden layer. By using back-propagation it is able to form a feature space by generating complex decision boundaries. Random forest works by creating decision trees with randomly selected features out of feature space.

Figure 1.1: Block Diagram of Human Activity Recognition

The main focus of the paper is implementation of activity recognition by background subtraction, SURF features and Bayesian classifier.

2. Literature Review

The importance and emerging technologies in human activity analysis, representation and analysis has led to various researches which are described below:

Manoranjan Paul, Shah M E Haque and Subrata Chakrabort [1] proposed several methods for human detection in videos and its applications. They classified the detection process in two steps:object detection and object classification. Methods discussed for object detection are Background subtraction, optical flow and spatio-temporal filter. Object classification is done by Shape based method, Motion based method and Texture based method. The most accurate method for object detection is spatio-temporal filter and for object classification is Texture based methods. Many applications of human

activity recognition and various datasets of the same are discussed

Rupali S.Rakibe and Bharati D. Patil[2] presented human motion detection based on background subtraction algorithm. It is performed by taking a sequence of video frames, obtaining the background image and finding the difference between current frame and background. Depending on various frames, Background is updated. Shape analysis is performed for removing shadow effects. Hence, it works well for static background.

Pedro Canotilho Ribeiro and Jos´e Santos-Victor[3] proposed novel approach for human activity recognition. The activities of major concern are active, inactive, running, walking and fighting. Different sets of features are taken with Bayesian classifier. Features such as position, velocity and speed are grouped. Likelihood functions of training set are modeled as Gaussian mixtures.

Tao Zhao, Member, IEEE, and Ram Nevatia, Fellow, IEEE[4] presented an approach for tracking multiple humans in complex situations. The algorithm is developed by decomposing human motion into global motion and limb motion. For segmentation and tracking of multiple humans, 3D shape model is used. Persistent Occlusion could be tackled easily in shape based models. Tracking is performed in cases of shadow casting and reflection.

Javier Andreu and Plamen Angelov[5] discussed the need and applications of human activity recognition systems. Through different methods adopted by different researchers, efficiency of human activity recognition is discussed.

3. Dataset used

Out of all the available datasets of KTH dataset, Weizmann dataset, PETS dataset and INRIA XMAS dataset, Weizmann dataset[1] is used.

A total of nine activities are performed which are bend, run, jump, single hand wave, double hand wave, skip, side walk, jack and walk. There are total of 360 frames in the multiple activity video in which various nine activities are performed.40 frames of each activity are there.

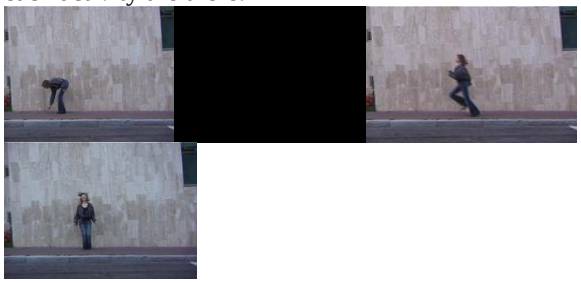

Figure 3.1: Weizmann dataset

4. Research Methodology

4.1. Background Subtraction

To recognize human activity recognition, human being is to be detected out of the video sequence. So background subtraction is performed which is based on following steps:

a. Video is read in to the MATLAB.
b. Background modeling is done by mean filter i.e. static background of the frames in video is calculated.
c. Difference of current frame from background image is taken by setting up a threshold.
d. Human is detected out of the video and is shown with the help of bounding box around detected human.

4.2. Speeded up robust features(SURF):

Once the human is detected out of the video, next comes the step to extract features of detected human by application of a specific algorithm.

Speeded up robust features(SURF) algorithm is applied for feature extraction. After performing background subtraction in previous steps, following steps are used for SURF feature extraction.

a. Call the function which finds SURF features, Since the algorithm extracts the features of grayscale input frames, so all frames of video are converted into grayscale.
b. Using the SURF algorithm on detected human, features obtained are Scale, sign Of Laplacian, Orientation, Metric and Location.
c. Strongest features are selected out of all the features.
d. All features are stored in a vector.

4.3. Bayesian Classifier:

Obtaining the features required, next comes the step for classification of the activity performed in video out of all the nine activities which are Bend, Run, Jump, Single hand wave, Double hand wave, Walk, Skip, Jack and Side-walk. Bayesian classifier is used

to classify the all the activities and recognizes the performed activity in video.

a. Divide the dataset in to 2 parts: Training dataset and testing dataset.
b. Use 70% of data for the training purpose and 30% of data for testing the classifier.
c. Store the names of all activities in a vector so that the classifier could sort out accordingly that which activity is being performed.
d. Call the function which trains the classifier with the features of the training data and activity vector.
e. Obtain the features of testing dataset by SURF algorithm.
f. After training the classifier, predict the activities of testing data.
g. Plot the confusion matrix.

5. Results

A Video consisting of 9 activities was loaded in MATLAB software. Total no. of frames are 360 and each activity had 40 frames. Size of every image sequence in video was 144 x180 x3.
After doing background modeling, background obtained is given below

Figure 5.1: Background

After getting background image, background subtraction is performed through which human is detected from a video with a bounding box around.

Figure 5.2: Detection of human in video

Feature extraction is performed using Surf features and features extracted for all the 9 activities are given below:

Table 5.1- SURF features for 9 activities

Bend	2.1	2.6	6.6	2.9	5.7
Run	2.2	2.8	4.8	4.4	3.8
Single wave	3.7	2.6	6.1	2.5	3.8
Double Wave	6.4	6.9	7.2	4.8	2.4
Walk	2.8	3	4.6	4.4	4.5
Jack	2.2	2.6	6.2	3.6	4.2
Side walk	1.7	2.2	3.2	2.8	3.8
Skip	1.8	2.9	3.4	4.1	4.6
Jump	6.1	2.5	2.9	3.8	5.8

Strongest features out of all the features of scale for all the activities are found.
Activity classification is performed using Bayesian classifier on all the 9 activities and most of the activities are classified accurately, as total frames in testing data were taken 108.12 frames of each activity are there and confusion matrix is found.

Table 5.2- Confusion matrix for 9 classes

12	0	0	0	0	0	0	0	0
0	11	1	0	0	0	0	0	0
0	1	10	1	0	0	0	0	0
0	0	1	11	0	0	0	0	0
0	0	0	0	11	0	0	1	0
0	0	0	0	0	12	0	0	0
0	0	0	0	0	0	12	0	0
0	0	0	0	1	0	0	10	1
0	0	0	0	0	0	0	1	11

6. Conclusion

Human activity is recognized by classification on the feature set of detected human in a video. 91.67% accuracy is achieved. Using SURF features with Bayesian classifier works well when the background in video is static. Algorithm worked well for the selected features.

7. Future work

In our work, we have chosen dataset with static background. In future, the work could be carried on by using dynamic background. Also, different feature extraction algorithms could be used with different classifiers to achieve high accuracy with minimum error possible

8. Acknowledgement

We express our sincere gratitude to all the staff of Lovely Professional University.

9. References

[1] Manoranjan Paul, Shah M E Haque and Subrata Chakraborty**,"**Human detection in surveillance videos and its applications ",Paul et al. EURASIP Journal on Advances in Signal Processing 2013, 2013:176.
[2**]** Rupali S.Rakibe, Bharati D.Patil,"Background subtraction algorithm based human motion detection", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013
[3] Pedro Canotilho Ribeiro & Jos´e Santos-Victor," Human Activity Recognition from Video: modeling, feature selection and classification architecture", International workshop on human activity recognition and modelling.
[4] Tao Zhao, Member, IEEE, and Ram Nevatia, Fellow, IEEE," Tracking multiple humans in complex situations", IEEE transactions on pattern analysis and machine intelligence, VOL. 26, NO. 9.
[5] Javier Andreu • Plamen Angelov," Towards generic human activity recognition for ubiquitous Applications", J Ambient Intell Human Comput (2013) 4:155–156.