HUMAN ACTION CLASSIFICATION USING OPTICAL FLOW

K.Sai Baba; K.Sai Nehatha; P.Hemadri Rao
Guide-N. Dhanalakshmi
Electronics and Communication Engineering
VNR Vignana Jyothi Institute of Engineering & Technology Affiliated to JNTU Hyderabad-500090,India Accredited by NBA and NAAC with A++ Grade

ksaibaba98@gmail.com sainehatha@gmail.com hemadri.pandrapragada.rao@gmail.com
dhanalakshmi_n@vnrvjiet.in

***Abstract*: Human action classification is a process of classifying the image sequences with labels . The human activity classification is a key enabler in the development of many applications, such as smart rooms, interactive virtual reality systems, people monitoring and environment modeling. Two essential categories of approaches of human action recognition can be distinguished: Wearable sensors_based and vision_based methods. This work is going to address the classification of human action based on vision. This vision based approach uses human body information which will be extracted from video sequences for classifying human actions. There are many methods to classify human action. Here, we are applying feature extraction train video ,extract interested human regions and then apply spatiotemporal features description on them. These histograms are used for training to classify action using support vector machine and Hidden Markov Model as classifiers. We provide comparison of the approaches based on the Support Vector Machine and Hidden Markov model.**

1. Introduction

Human Action Recognition is a developing field in computer vision and machine learning. Our aim is to perceive the game being played in the video. There are different procedures like extracting feature vector or key-points for Human Action Recognition. Optical flow can be calculated, which gives the estimation of movement, of a pixel in a sequence of frames. Human action recognition using the optical flow is one of the finest and best way to locate and track the human behaviour and find out there accuracy by using necessary classifiers by giving input action video .

Optical flow detects humans in the video and divide into frames and places the co-ordinate points in the given video. Here optical flow detects the motion of human or any movable objects in the video. The main function of optical flow is it would calculate the difference between the two images marking with the co-ordinates points. The general definition of an optical flow would calculate angle and magnitude of pixel movements between any two consecutive frames in the video. As the video is divided into frames and also calculating the distance of the change of pixel is been represented in the bin format which we can call as the histogram of optical flow(HOF) .

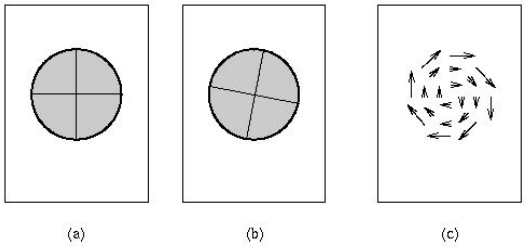

Fig.1.1 Optical Flow, a)time t_1 b)time t_2 c)Optical flow

Our main aim is to recognize action of humans in the given input video and to calculate the accuracy levels using different classifiers by training with the different action videos . The action video comprises of the jumping, walking, boxing etc . Here these action dataset can be downloaded through the internet . We have to train the classifiers by giving this action videos as an input . We are training two classifiers for the human action that is SVM and HMM. A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After giving an SVM model sets of labelled training data for each category, they're able to categorize new text.For HMM we are choosing the time sequence varying for the better results and better performance as the result is static. HMM makes the strong in this classification we can predict the future sequence . The present state has no impact on the future state also using hmm classifier. Here after train the input video the next stage is the prediction stage here where the new data and the train video data is been classified based on the search of the data the classifier gives us the accuracy level of the new data and the trained data.

2. Literature Survey

Understanding human action in videos has received significant research attention in the field of video analysis. Most applications are in summarization, video content retrieval, and human computer interfaces. For the testing of the video the video contains some negative and positive samples first step we have to train the video (i.e) to train the input video here the input video can be in the action format like running, boxing, jumping and many more any video which performs any action or change in the objects can be consider as the input video. The input video should be train having the combination of the both positive and negative video can be determine by the SVM classifier. Diffrent from the pervious approach it does not require human action detection and tracking as input there are some methods using detection methods to track the video contents . Based on the input video after the first step we complete after training the video the next step is the classifiers here they are different types of classifiers here the classifiers determine the efficiency of the input video. The main tools we are using to determine the optical flow is the python and the openCV. Python has the function of the in-built function of the classifier it can be any classifier that is SVM and HMM classifier . SVM which it determines the separation of the two data sets the two data sets can be may be both positive but different actions, can positive and negative , can be both negative but with different action .

Fig 1.2 Different Action Video

The first step we have to train the machine here we are giving the input action video to train we are training the action dataset to find out the accuracy . Recent years the best methods where which employ machine learning techniques such as SVMs and Ada Boost, this technique provides the incorporation information in a training sets. As the SVM that is support vector machine is quite simple method to train the data and find the accuracy level of the train data .

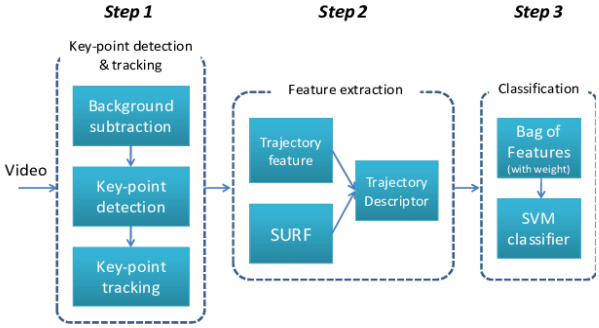

Fig 1.3 Block Diagram of SVM classifier

So next classifier that is HMM is the hidden markov method classifier. Hidden Markov Model (HMM) is a statistical Markov model in which the system being modeled to be assumed. HMM is a model that tells us something about the probabilities of sequences of random variables, states, each of which can take on values from some set. It can b.

HMM makes the strong assumptions that if we want to predict the future sequence, all it matters in the current state. The states before the present state haven't any impact on the longer term except via the present state. Our main task is for assuming the temperature. We'll simplify this weather task by assuming there are only two kinds of days: cold (C) and hot (H) . .The two hidden states (H and C) correspond to hot and cold weather, and the observations (drawn from the alphabet O = {1,2,3}) correspond to the number of ice creams eaten by person on a given day.

Markov Assumption: p(qi |q1...qi−1) = P(qi |qi−1)

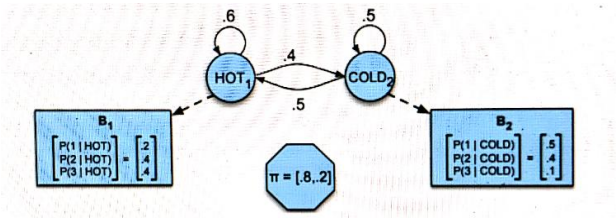

Fig 1.4 Example of the ice cream eaten by the person by relating to weather .

3. Methodology

HOFs for each of the videos in the dataset are obtained to train for action classification. The HOFs are obtained from optical flow by applying Fernaback Optical Flow algorithm. The HOFs, as training input, given to machine learning algorithms SVM(Support Vector Machine) and HMM(Hidden Markov Model) and compared for performance

A. Dataset

The video dataset has been taken from KTH Royal Institute of Technology database. The daraset contains videos in six categories namely walking, running,

jogging,boxing, handclapping and handwaving, each category containing 100 videos. These 100 videos inturn 25 cattegories, each category having 4 ssample videos of a person doing action in different orientations both in outdoor and indoor.

Fig. 3.1 Dataset Screenshots

B. Software and Hardware

The entire training is done on Python programming. OpenCV, an open source computer vision library, is main library along with numpy, matpltlib, sklearn, hmmlearn and smutils installed on a computer having minimum of 2,2 GHz clock, 4GB RAM and minimum graphical properties.

C. Calculating HOFs

Calculating Optical Flow in videos is heart of this project. Before passing the image for optical flow calculation, to reduce the load on CPU, the images are converted to grayscale. Optical flow is calculated between every two frames in video. It is done using Gunnar-Farneback algorithm. The resultant of this function is complex number which indicates both angle and magnitude of pixel change in video. From this histogram elements are filtered to get histogram matix. This process is repeated for every video in the dataset. The resultant of this whole process gives a huge amount of 2-D matrices which are used for training.

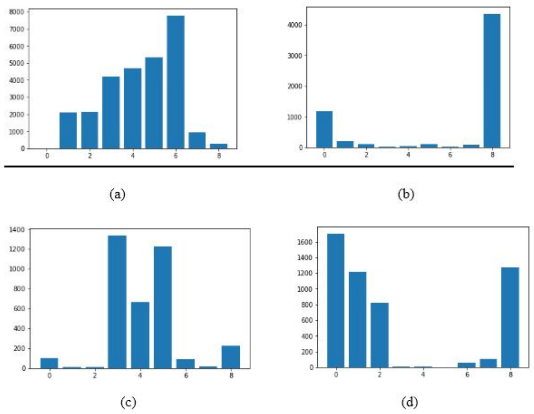

Fig.3.2 Histogram Plots

D. Training

To train the model we are interest in SVM and HMM machine learning algorithms. The reason to choose these algorithms is to compare the performance of these two completey diverse algorithms. SVM is a classification model which uses multi-class training datasets simultaneous. On the otherhand HMM is a sequential algorithm which predicts probability of given output by checking the individual stages in the sequence. In this work, for optimum result, we split the data into 85% of training data and 15% of testing data. And SVM parameters kernel, C and gamma are varied depending upon preformance. Similarly for HMM model with its training parameter algorithm as “forward_backward”.

4. Results

For ease of making we trained models in SVM to classify only two classes each time. Running and walking classes give 90% accuracy, handclapping and handwaving classes gave 91% accuracy wheras boxing and jogging classes gave 100% accuracy. Coming to HMM, as HMM is generative model we train a separate model for each class of action. The resultant models gave 70% for running, 50% for walking, 80% jogging,, 80% for boxing, 85% for handclapping

and 100% for handwaving. For testing each video is in testing database is tested with the trained n models and using the likelihood approach we get the accuracy. The resultant for 3 classes (walking, jogging and ruuning) combined is 70% in HMM. Although HMM gives less accuracy in some cases, it takes less time than SVM. Keeping this in mind HMM is preffered, in some cases with time constraints, over SVM.

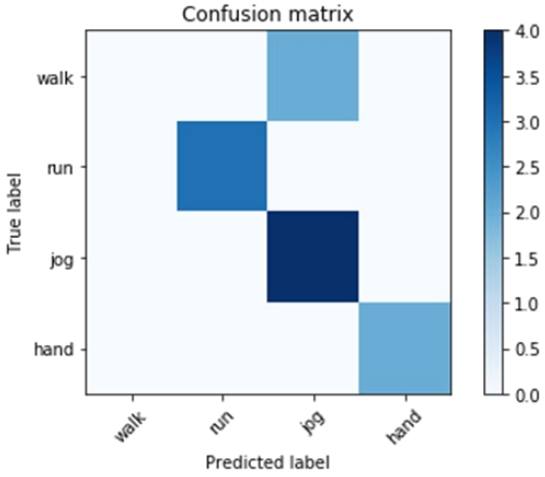

Fig 1 Viewing the confusion matrix for the training of 4 classes .

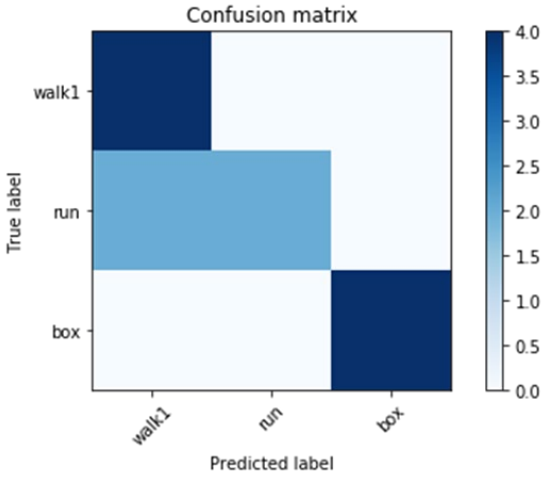

Fig 2 putting examples on view of the training of 3 classes

5. Conclusions and Future Scope

In recent years, video surveillance systems have become very popular due to heightened security concerns and low hardware costs. These types of systems are widely used in many applications such as nursing care institutions, law enforcement, building security, and traffic analysis. With the growth of technology and image processing cameras are widely used , the efficiency and accuracy of human operators have reached the limit. Therefore, there is a need for a real-time automated system that that detects the abnormal events or any abnormal moment actions . In this way, detecting unusual or suspicious,activities,uncommon behaviors, or irregular events in a scene is the primary objective of an automated video surveillance system. By introducing this classifiers we can resolve the issues.

`6. Bibliography

J. K. Aggarwal and Q. Cai. Human motion describe about a review on Computer Vision and Image Understanding.

S.Ali, A. Basharat, and M. Shah. Chaotic invariants for human action recognition.

A.Bissacco, A.Chiuso, Y. Ma, and S. Soatto. Recognition of human gaits .

Chiuso,and S.Soatto. Describes Classification of dynamic models.

A.Chan and N.Vasconcelos. Done research of with Probabilistic kernels for the classification of auto-regressive visual processes.

A. Chan and N. Vasconcelos. Classifying video with kernel dynamic textures. Describes about the pattern recognition

R. Chowdary and R. Vidal. Supplemental material for “Binet-Cauchy kernels on nonlinear dynamical systems for the recognition of human actions”.

ORIGINALITY REPORT

17%	12%	4%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Rensso Victor Hugo Mora Colque, Carlos Caetano, Matheus Toledo Lustosa de Andrade, William Robson Schwartz. "Histograms of Optical Flow Orientation and Magnitude and Entropy to Detect Anomalous Events in Videos", IEEE Transactions on Circuits and Systems for Video Technology, 2017 Publication	4
	www.ijert.org Internet Source	3%
	ijraset.com Internet Source	2%
	it.scribd.com Internet Source	2%
	monkeylearn.com Internet Source	2%
6	Submitted to Higher Education Commission Pakistan Student Paper	2

7 Submitted to IUBH - Internationale Hochschule Bad Honnef-Bonn 2

Student Paper

Exclude quotes	On	Exclude matches	< 30 words
Exclude bibliography	On		