Recognition of Human Activity using Incremental SVM

Mr. Rupesh V. Jagtap¹ & Mrs. M. A. Potey²
¹Student Department of Computer Engineering DYPCOE, Akurdi, Pune, India.
²Faculty Department of Computer Engineering DYPCOE, Akurdi, Pune, India.

Abstract: Human activity detection is recognize the human body activity from different variation in human body size, and different activity scope. This area of research is become more interesting because of various human body movement style. The main goal of human activity detection is to detect normal and abnormal action which helps to make human life more comfortable. The multi-class support vector machine is used in first step for initial classification and then each decision accept from new frame is used to update the classifier. In the field of human activity detection the detection of abnormal activity is still a research problem, with respect to large human movement. As in hospital monitoring patient and give special manpower to each patient is difficult and more costly hence the propose system helps to monitor abnormal patient activity and provide instant help to the patient. As there is lack of patient dataset we prepare a dataset and experiment the propose system.

Keywords- Human action; Incremental support vector machine; Multi-class SVM; Human features; bounding box.

1. Introduction

Detection and monitoring of the human actions is upcoming research area and become ambitious due to the strength of support and useful in many application such as patient monitoring at home i.e. smart nursing and it also relate with the many different field of study such as medicine, human computer interface or sociology. Many human action detection problem is still unsolved and if solve it gives rise to many application that make human life more comfortable. Action detection in video sequence is learn machine to recognize human body and different actions from extracting information from the video. The information such as motion information and different features present in the videos. The features like height and width of bounding box and instantaneous speed of the moving object is important for the detection of human and activity. For classification of different actions perform by the single person or different person performing single action different soft computing techniques are used for effective classification. Detection of human activity robustly emerging as an investigation tool push ahead by the growing need of the people. Action recognition is a method of detecting action that happen in video series given as input to system or any application [4]. There are many complications face by the researcher while experimenting the videos for the human action detection as background of the video is moving or static, if the video contain human as moving object then human is detected but system which detects the moving object also detect the moving leaf of tree and any moving object. So for this problem in the propose system concept of bounding box is used for detecting the human body. It draws the box around the human body. Bounding box specifies the human body very effectively and form the bounding box the height and width are calculated which is used as training feature for classification of the human activity. For the classification of actions the support vector classifier is used which gives the better result with respect to other classifier as it minimize the upper bound risk and effective kernel function are available and gives better accuracy in different classification problem. Human detection in real world finds many application in various domain such as intelligent video surveillance, customer attributes, shopping behavior analysis. Still the human action recognition is very difficult task due to changing background, situations and viewpoint variations [3].

Also human motion analysis is interesting research topic in computer vision. Mainly the task of assigning the different class labels to training data sequence encouraged by human activity has attracted the researchers. The focus of this research is real time monitoring of elderly peoples health and recognizing any abnormal activity for e.g. fall recognition. To take advantage of recent advances in information technologies and reduce the burden on hospitals and economy. Also, it is proved that people suffering from long term diseases feel much better in their home environment than hospitals. The human activity and posture recognition have been extensively studied during the past few years. A detailed survey of video based motion and activity recognition systems is discussed in this paper. A smart home system for elderly
peoples health care will be implement by using video monitoring for detect and monitor elderly people activities and generate an alert in emergency situation and provide instant help will save the many patient from getting the extra injury as they may get into problem[4].

2. Related Work

In [7] authors used method depend on detecting the concern point using SIFT (scale invariant feature transform) from each frame of the video A acceptable modification step is used to limit the number of concern points giving to the amount of details. Then the prevalent approach bag of video words is applied with a new standardization technique. This standardization technique unusually improves the results. Finally a multi class linear support vector machine (SVM) is applied for classification. Experiments were piloted on the KTH and Weizmann datasets. The results demonstrate that given approach beats most existing methods, succeeding accuracy of 97.89% for KTH and 96.66% for Weizmann.

In [8] used a new approach for human activity recognition in a video sequence by manipulating the key poses of the human shapes, and building a new classification model. The spatiotemporal shape dissimilarities of the human shapes are signified by separating the key poses of the silhouettes into a fixed number of lattices and cells, which leads to a noise free representation. The efficiency of this approach of activity representation and classification model is verified over three public data sets i.e. Weizmann, KTH, and ballet movement. The relative analysis shows that this method is greater in terms of recognition accuracy to comparable state-of-the-art methods.

In [9] authors used multi-category classifications of human actions are usually achieved by solving many one versus rest binary SVM classification tasks. Though, it hints to the class inequity problem. Furthermore, because of environmental problems and inherent noise of spatiotemporal features, videos of similar actions may hurt from huge intra class variations, author discourse these problems by presenting the energy based least square twin support vector. Scientist investigate the performance of this methods on Weizmann, KTH dataset. In [2] author used a system framework is presented to recognize different types of activities from videos by an SVM multi-class classifier with a binary tree architecture. The framework is collected of three functionally cascaded modules

1) Detecting and tracing people by background subtraction approach.
2) Extracting various features such as local features from the minimum bounding boxes of human body in each frames and a recently defined global one that is contour coding of the motion energy image (CCMEI).
3) Recognizing human activities by SVM multi-class classifier whose arrangement is determined by a clustering process.

In [1] authors used the incremental support vector algorithm which updates the new information contain in the every upcoming frame in the video. It solve the problem of adaptive ability of the human action and the training step depend only on few images, it uses the mathematical property of the support vector machine to update specific part. The initial step is the extraction of features from the video based on color and texture feature i.e. appearance of the person. The system is evaluated on the CASIA gait dataset which contain 20 people’s actions.

3. Human Activity Recognition

Various research papers are published in human activity detection but in this paper a novel idea related to abnormal activity of patient is discuss. Patient having mobility problem or patient suffering from disease like Parkinson’s are usually face the problem like sudden falling on the floor and no one is present for help at that time of instant, therefor they may get injured. Hence for avoid such type of situation the human activity detection system is proposed.

In the proposed system video surveillance camera provides the input to the system as video. Then the human is detected by the image differencing method. In the image differencing method frame wise subtraction is done also called as background subtraction. In this moving object is detected and then from the bounding box various feature are extracted from the video. The different information present in the video are extracted and according to that training is done. The support vector machine is used for the classification of different human activity. The SVM is train on the basis of different information in the video sequence such as height and width of the bounding box, instantaneous speed of the moving object, contour of the moving object i.e. human body. For the fast, mathematically stable and robust implementation the incremental support vector approach is come into picture. Incremental SVM uses the mathematical properties of multi-class SVM and update the system by new information coming from each frame.

4. Proposed System

The patient having long term disease have to spend more time in the hospital but patient feels more comfortable at home than the hospital hence if possible patient further treatment is given at home but no one is to monitor whole night and day for the
the patient. It is not possible but important to keep watch on the patient. Our system monitor the patient room and keep watch on the patient activity and if patient performs the abnormal activity such as falling down from bed or falling while taking water or in the bathroom system detects and immediately alarm the person who is nearby the patient and provide instant help to the patient which avoid the further damage to patient.

In system first the video sequence is given as a input to system. Next step is to subtraction of stable background in the video frame is perform. Then the different feature is extracted from the information in given video such as shape, motion etc. In the final step the classification of the different position carried out and action is detected. The preprocessing steps are discuss in the system. In the supervise learning the support vector machine is very efficient as it is discriminative classifier. Discriminative classifier gives better result than the generative classifier. The traditional classifier such as Naive Bayes, BP network and KNN may not give effective output when the limited sample is given for the training purpose, while SVM does. In statistical learning support vector machine provides the structural risk minimization gives better result on the testing dataset[6].

A) Foreground extraction:

The video sequence is given as input to the system this is the first step of the system. Background subtraction method detects human body. The fig 2 shows the human detected in the given video. The video consists of number of frames, background subtraction method separate moving object from the present in the frames. The some problem is occur while detecting the moving object in the video such as light condition, shadow, sudden change in the brightness etc[2].

B) Features extraction:

The detected human body is represented by the bounding box and the features are extracted from the bounding box as local feature and global feature for the whole activity sequence.

1) Local Features:

Local feature are extracted from the bounding box around the human body there are three concept that select the bounding box as feature foundation

1. It is easy to obtain the bounding box to the human body as compare to other type of computational model.
2. The information which is extracted from the bounding box is quite easy to process and make further computation.
3. The important fact is that the information of the dimension obtain from the minimum bounding box are more effective than the information obtain from those directly obtain from the detected binary images.

The local feature consist of two type of feature one is shape feature and second is motion feature. The shape feature gives rise to another two type feature one is unitary height and another is ratio of height against width [2]. The height of the bounding box is represented as

\[ h(t) = \frac{H(t)}{H_{\text{max}}} \]  

(1)

Where \( H(t) \) gives the value of height of bounding box at time \( t \), and \( H_{\text{max}} \) is the maximum value of height of bounded box of whole action sequence. The another feature is the height width ratio is given as

Patient monitoring system helps to solve many problem of patient and help patient to improve their daily life. In the given system first step is to perform background subtraction by image differencing technique. In second step different feature is extracted from the video frame such as local feature global feature etc. Then the classification is done with the help of support vector machine. Operator need not watch monitor at whole time system awake an alarm for the adverse situation.
Imperial Journal of Interdisciplinary Research (IJIR)  
ISSN: 2454-1362, http://www.onlinejournal.in

\[ R(t) = \frac{H(t)}{W(t)} \]  
(2)

Where \( W(t) \) represent the width of the bounding box at time \( t \).

II) Motion feature:
The area of the human body gives the information about the motion feature. Let us consider the \( f(x,y) \) is an input image and the geometric moment of \( p + q \) rank is given as \[2].

\[ m_{pq} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} x^p y^q f(x,y) \]  
(3)

Where \( M \) and \( N \) are the height and width of the input image.

The coordinates that represent the centroid of the human body is given as,

\[ (x_0,y_0) = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \]  
(4)

Coordinates \( x_0(t) \) and \( y_0(t) \) is calculated from the minimum bounding box at time \( t \). The motion feature consists of instantaneous speed \( v(t) \), the speed with respect to \( x \) axis is denoted as \( v_x(t) \) and instantaneous speed with respect to \( y \) axis is denoted by \( v_y(t) \) of the centroid and defined as,

\[ v_x(t) = \frac{x_0(t) - x_0(t-1)}{1} \]
\[ v_y(t) = \frac{y_0(t) - y_0(t-1)}{1} \]
\[ v(t) = (v_x(t), v_y(t)) \]  
(5)

III) Global Feature:
The global feature contains the contour coding of the motion energy image and the binary version of the motion history image is the motion energy image. The contour consist of the hollow portion of the human body. The motion energy image is represented the motion energy image. Temporal extent of the motion energy is calculated which means the starting and end point of the specific activity sequence. The contour of the motion energy image is shown in fig 3.

IV) Classification of the human activity:
Classification of the human activity is done on the basis of the extracted feature. The feature like local, motion, global etc. are used to train the specific activity sequence of human. Height and width ratio is used to detect action.

4.1. Results

<table>
<thead>
<tr>
<th>Video dataset (action)</th>
<th>Walk</th>
<th>Fall</th>
<th>Jogging</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk</td>
<td>96%</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fall</td>
<td>0</td>
<td>95%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jogging</td>
<td>0</td>
<td>0</td>
<td>90%</td>
<td>0</td>
</tr>
<tr>
<td>Run</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>91%</td>
</tr>
</tbody>
</table>

Our dataset is compare with the CAVIER dataset which contains the large number of activity sequence. Also we compare our system with the KTH dataset for the walking activity for specifying normal and abnormal activity. The KTH dataset consist of 2391 activity sequence perform by 25 peoples. The video sequence are down sample to 160×120 pixel and length of videos are maximum 41 second. All the video are having static background. In our dataset the 20 different activity sequence perform by the 4 people.

5. Conclusion and Future Scope

Human activity recognition is important field in Human Computer Interface (HCI), artificial intelligence. It is used in the real time hence become more interesting as it is can used in daily life situation. Human for the effective patient monitoring is an important field as patient in critical situation needs to observe his every action. It became challenging because of recognition of human action from different body proportions and dissimilar feature of human. Human action detection seized the attention of numerous computer science communities because it provide personalized support for number of application and connection to many field such as medical, human computer interface. In future this
system used in hospital to detect the abnormal activity of patient and provide help to patient and also enhance the system by evaluating on the different types of dataset.

1. References


