International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 11 | Nov 2019 www.irjet.net p-ISSN: 2395-0072

Recognition of Theft by Gestures using Kinect Sensor in Machine Learning

Surbhi Kasat[1], Kanchan Tupe[2], Vaishnavi Kshatriya[3], Roshani More[4], Ms. Prachi S. Tambe[5]

[1,2,3,4]*Student, Computer Department, SVIT, Nashik, Maharashtra, India*
[5]*Assit. Prof. Computer Department, SVIT, Nashik, Maharashtra, India*

Abstract - Surveillance cameras have been popular as measures against thefts. They are useful for deterrence of crimes such as theft and murders, it is possible to detect crimes by installing artificial intelligence to surveillance cameras. However, a large amount of behavior data must be trained for anlyzing a gesture pattern of a person by image recognition.

In this worlds mostly crimes related to money .The security of bank, jwellery shops, houses and malls must be properly secured from theft peoples. Survilliance cameras or any camera like kinect sensor have been measurely used at many places. The accuracy of image recognition by artificial intelligence and machine learning has especially increased. That is possible to detect crimes by using kinectcamera. A large amount of gesture and posture data is used to analyze the behavior of person that means its a Theft or not by the image processing and the data size is larger as they are image data, For that Dynamic time wrapping is used to analyse the object and person. We propose an idea to judge a behavior gesture of a person using Kinect sensor in order to detect crimes at real-time.

***Key Words*: Abnormal behaviour, Kinect, Posture Recognition, Speech to Text, Alarm, SMS, Image Processing.**

1. INTRODUCTION

In recent years, cases such as theft and murder have been steadily occurring, and the number of occurrences is also increasing. The number of placed surveillance cameras has been increasing in order to deter crimes or identify suspects after the incident. In addition, due to the development of computer technology such as image processing and artificial intelligence, suspicious people can be detected. The camera will recognize the action being performed by the user, it gives skeleton of human body when user stand in front of kinectsensor. This actions are campared with action stored it dictionary. A dictionary is maintain with all gestures and related speech are stored. If match is found then the SMS and alarm is generated.

1.1 Literature Survey

Rajvi Nar, Alisha Singal, Praveen Kumar, ”Abnormal Activity Detection for BankATM Surveillance”, 2016 In this paper Posture recognition is one of the most interesting in computer vision because of its numerous applications in various elds. The problem we face with simple camera can be solved by the usage of a 3D camera. In this project, we explore a technique of using skeleton information provided by Kinect 3D camera for posture recognition for active real-time ATM intelligent monitoring. To achieve posture recognition we can use kinect to track bone joints and their positions. By analyzing the position information, the system detects abnormal behaviors.[1]

Min Yi.”Abnormal Event Detection Method for ATM Video and its Application”, 2011. In this paper the development of video surveillance system experiences three stages: analog video surveillance system (VCR), semi-digital video surveillance system (DVR/NVR) and digital video surveillance system[1-2]. With the development of network and multimedia technologies, intelligent video content analysis has become a key component in video surveillance system. Meanwhile, the video and image forensics has been become an indispensable function in security area. More and more digital video surveillance systems are used, more and more complex criminal scene investigation and evidence collection related to video analysis have become. [2]

2. METHODOLOGY

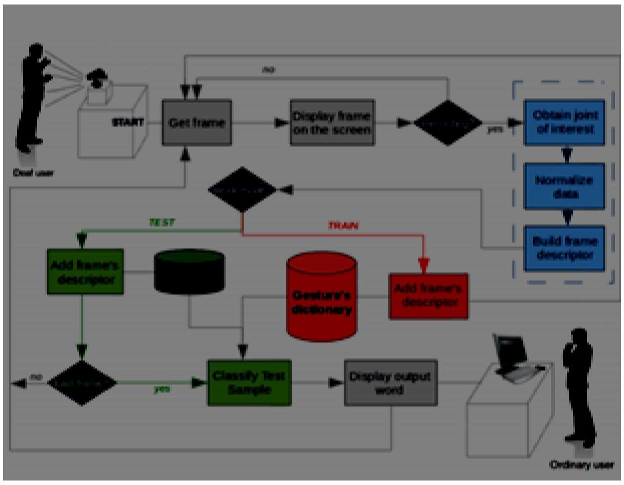

Fig -1:Methodology of System

In this system, the kinect Sensor is used to recognition of gestures by detecting the bone joints in the form of skeleton. In the dictionary of database some common gestures of thefts (Holding gun gesture)and ordinary people (Hands up gesture) are stored, If theft is come in front of kinect camera then the gestures will match found using dictionary gestures ,then the alarm or SMS can be generated and send to Police and neighbour or owner.

3. System Architecture

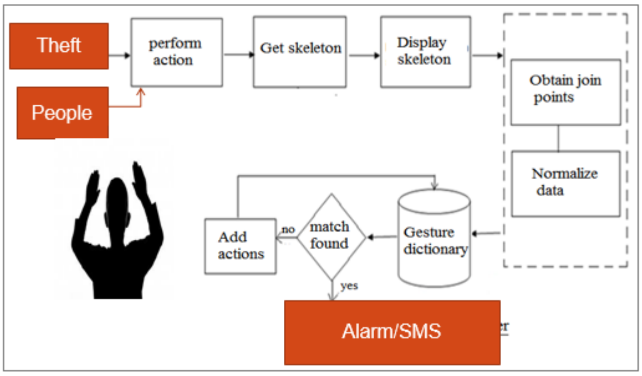

Fig -2:**System Architecture**

Consider the block diagram from Fig 2. The deaf user is in front of the camera doing a sign or getting ready to do so. With a frame rate of 20fps, a new frame is obtained and the video stream is updated with the skeleton of the user overlapped onto it. At that point, if the user wants to record a sequence (otherwise, the system asks the camera to get the next frame), three main blocks are executed: the first block consists of obtaining the data of the joints of interest (JoI) required for the frame descriptor, the second block consists of normalizing these data, and the third one consists of building the frame descriptor. Then, if the working mode is set to TRAINING (meaning that the user is adding a new sign to the training set), the frame descriptor is added to the correspondent file of the dictionary. Otherwise, if the mode is set to TESTING (meaning that the user wants to translate the sign that is been done), the frame descriptor is added to the current test sample. Then, the system checks if the current frame is the last frame of the sign. After a sign is finished and if the working mode is TESTING, the test sign is compared using a classifier with the signs from the dictionary and the corresponding output is displayed so that the ordinary user will know the corresponding word in the spoken language. After that, the system keeps going with the next frame and the flow of the block diagram is repeated again.

4. Implementations

Skeleton Tracking: Kinect used for skeleton tracking in this paper was the second generation Kinect v2. We acquired skeleton coordinates with Sensor, Source, Reader, Frame and Data in order [10]. First of all, Sensor defines the Sensor interface to handle Kinect, then Sensor is acquired and opened. Next, Source defines the Source interface for the Color frame, and then Source is acquired from the Sensor. Next, Reader defines the Reader interface for Color frame, and then Reader is opened from Source. Next, Frame and Data set the size of the Color image and the data, and then prepare the Open CV to handle the Color image.

When data starts to be acquired, the examinee turns in a direction of Kinect, and then the examinee moves the action to the behavior of suspicious behavior or normal behavior. For this reason, samples immediately after starting data acquisition are not used for evaluation, and samples after the beginning of the action are used for it. We decided the time 5 s since it is necessary for 5 s at the longest to change the motion to the action. Here, the acquisition interval of samples is 2 s, and the time required to acquire one sample is about 1 s. Based on that, in our experiments, we set the number of samples taken at once to 25, and we repeated it twice to acquire 50 samples of data. Therefore, the number of acquisitions of joint data by Kinect is assumed to be 27, and the first two samples are not used for the evaluation, and the 25 samples after that are used.

Shaping of the Joint Coordinate Data:

Specifically, joint coordinates that are in contact with the floor and joint coordinates that are in contact with the door knob are not changed during the shaping. That is, WRIST_RIGHT, HAND_RIGHT, WRIST_LEFT, HAND_LEFT, THUMB_RIGHT, HAND_TIP_RIGHT, THUMB_LEFT, HAND_TIP_LEFT, ANKLE_RIGHT, FOOT_RIGHT, ANKLE_LEFT and FOOT_LEFT are the same values as original values. We ignored the foot size since there was no correlation between height and foot size. When other joint coordinates are shaped, the magnification is obtained from HEAD, FOOT_RIGHT and FOOT_LEFT as mentioned above. In this shape, Y coordinate is changed since FOOT_RIGHT and FOOT_LEFT are in contact with the floor and hands are in contact with the door knob. The changed values of the coordinates are calculated with the magnification. Finally, they are obtained by adding the sum of the average of FOOT_RIGHT and FOOT_LEFT. ELROW_RIGHT and ELBOW_LEFT are calculated in the same way with HAND_RIGHT and HAND_LEFT.

Data processing and Feature extraction: With the skeleton tracked by Kinect, joint positions are obtained. Since, joint vector has 3 coordinates and a skeleton consist of 20 joints, the feature vector has 60 dimensions. Apart from the feature described above, another feature can be extracted by calculating the joint angles. While working with postures, we observed that ten joints, namely Torso, Neck, Head, Left shoulder, Left elbow, Left wrist, Right shoulder, Right elbow, Right wrist, Left hip and Right hip, are the most important joints for representing postures.

From these joints, we can calculate different sets of angles. Subsequently, we defined the following angle based features to recognize desired postures: angle between vector joining left hand to left elbow and vector joining left elbow to left shoulder, angle between vector joining left elbow to left shoulder and left shoulder to torso, angle between vector joining right hand to right elbow and right elbow to right shoulder, and angle between vector joining right elbow to right shoulder and right Shoulder to torso. The calculation of the subject's posture is based on the fundamental consideration that the orientation of the subject's torso is the most characteristic quantity of the subject during the execution of any action and for that reason it could be used as reference.

Normal and Abnormal Postures

For demonstration purposes, we have selected the following three postures to be considered abnormal, and have used the same as our training set data –

- Aggressive posture (fig. 3)
- Fiddling with the camera posture (fig. 4)
- Peeping posture (fig. 5)

An alarm is raised if the current posture of the person matches with any of the postures mentioned above. We have considered any posture that does not fit the abnormal posture category as a normal posture. No alarm is raised in this case.

Fig 3:-Aggressive Pose by People

Fig 4: Fiddling with the camera posture

Fig 5: peeping posture

5. Conclusions

In this project we showed that posture recognition can be used to detect abnormal behavior of a person in an ATM or Banks and Surveillance Areas. We achieved this by using skeleton data that can be extracted from the depth image provided by a 3D camera like Kinect.

We used Processing language (built on C#) to write our code. A Machine Learning algorithm, Logistic Regression was used to calculate the probability of the current pose of the person under surveillance being abnormal. This was achieved by calculating the weights from a training data set. We used angles between different bones as features which we processed using a program to calculate the weights for the algorithm, Image Extraction, Capture number of frames related to the movement of the hand, Extract Joint points and normalize points. For each sequence Si, in database we calculate Euclidian distance Si with captured sequence.

We used gradient descent method to calculate the optimum value of weights. And this way we were able to make the machine learn the abnormal behavior using algorithm and data. The Alarm or Notification message is been send by the system in the accordance if the condition occurs with buzzer and SMS to authorized Person.

REFERENCES

1) Rajvi Nar Alisha Singal Praveen Kumar," Abnormal Activity Detection for Bank ATM Surveillance", August 2016

2) Min Yi," Abnormal Event Detection Method for ATM Video and its Application. Communications in Computer and Information Science", 2011.

3) Antonio Chella, HarisDindo, IgnazioInfantino,"People Tracking and Posture Recognition for Human-Robot Interaction.Vision Based Human-Robot Interaction", March 2006.

4) Bernard Boulay," Human Posture Recognition for Behaviour, Universite Nice Sophia Antipolis", 2007. English.

5) Min Yi, Abnormal Event Detection Method for ATM Video and its Application. Communications in Computer and Information Science, Springer 2011.

6) Antonio Chella, Haris Dindo, Ignazio Infantino, People Tracking and Posture Recognition for Human-Robot Interaction. Vision Based Human-Robot Interaction, Palermo, Italy, March 2006.

7) Bernard Boulay, Human Posture Recognition for Behaviour. [cs.OH]. Universite Nice Sophia Antipolis, 2007. English.

8) Greg Borenstein, Making things see. Maker Media Inc., 2012.