https://dx.doi.org/10.22161/ijaers.4.2.14 *ISSN: 2349-6495(P) | 2456-1908(O)*

Dynamic and Static Gesture Recognition System Using Moments

Rajvardhan Thakare, Parvez Khan Pathan, Meghana Lokhande, Neha Waje

Department of Computer Engineering, Pimpri Chinchwad College of Engineering, Nigdi, Pune, India

Abstract—*Gesture recognition is the novel idea to enhance the signs recognition of those who have speech and hearing disability .Our project discusses an improved method for gesture recognition .The algorithm extracts the gestures from the video given to it and it detects the hand using HSV skin color segmentation in the intent to eliminate the other parts of the body and detect only hands .It distinguishes between static and dynamic gestures and extracts the appropriate feature vector .We used SPHINX parser to form word from set of letters. We strive to enhance the reliability and efficiency by using faster static gesture recognition algorithm.*

Keywords—SVM, ASL, OpenCv, Zernike moments.

I. INTRODUCTION

Gestures are an important aspect of humen interaction, everyone usually use the movements of hands and face while interating with others. A Gesture is a form of non-Verbal communication in which visible bodily actions communicate particular messages. Gesture recognition is one of the most famous techniques used in Video Games and in Sign Language Recognition. There are two type of gesture recognition system static and dynamic. Static gestures are that which are captured when actor is steady and dynamics are those which consist of set of actions like asking someone for water is the best example of dynamic gesture and saying someone thanks is the example of static gesture. It is then important to detect only hands out of humans whole body . HSV is simplified model for human skin detection which is used to identify the parts of the body, further more algorithms like Viola Jones can be used to detect and eliminate the face. OpenCV is the important technology and their packages can be easily imported to the frameworks like JAVA and Visual Studio, it can be easily used to detect and work on geometric means. Efficient detection and removal algorithms make the system faster and reliable. In order to reduce the cost of the system, we avoid using the gloves for the detection, no any sensor are required to detect and the gestures.

The input can be a static or dynamic gesture, so to make it as general as possible a video recording of two seconds is passed as the input at the rate of 6fps. The gesture is extracted and depending on its type (static/dynamic) certain features are extracted. These are then classified using pre-trained SVM classifiers.

The system is able to perform the classification of gestures as background cluttered HSV model is used, i.e. focus on the hands and face. Since the letters a-z does not involve the facial expressions and it is not needed, we eliminate the face using Viola-Jones face detection followed by subtraction of the detected region. We classify the gesture as static or dynamic by measuring the distance moved by the hand in subsequent frames. For static gestures, we use Zernike moments, a well-known shape descriptor in image processing. For dynamic gestures we extract a curve feature vector which shows high accuracy in uniquely identifying paths. These feature vectors are then classified using pre-trained SVM classifiers. [1]As discussed in [1] there are only few cues which requires lip moments; so the focus is not on the lip and shoulder movements.

II. LITERATURE SURVEY

Sign Language Recognition: This paper discusses the method to input gestures and the static and dynamic gesture recognition system using moments theory and support vector machine; the SPHINX parser is discussed to form the words from letters. In this paper they also presents a moments theory. And the system that they proposed is without using the sensors.

Static and dynamic gesture recognition in depth data using dynamic time wrapping: This paper discusses the recognition of the gesture using sensors. The dynamic gesture recognition system discussed in this paper can be applied to the gesture recognition system without sensors. They used the sensors; Microsoft Kinect. They also proposed a algorithm caleed K-Curvature algorithm. And the system they proposed is to improve the scanning time in order to identify the first pixel.

Pattern recognition (Journal) : It discusses dynamic gesture recognition using CoG and other methods to recognize the gestures with other methods

A study of hand gesture recognition using moments :It discusses the theory of the moments and ststic getsure recognition from it , it discusses the concepts like Zernike moments krawtchouk moments and geometric and orthogonal moments.

https://dx.doi.org/10.22161/ijaers.4.2.14 *ISSN: 2349-6495(P) | 2456-1908(O)*

III. RELATED WORK

The input can be a static or dynamic gesture, so to make it as general as possible a video recording of two seconds is passed as the input at the rate of 6fps. The gesture is extracted and depending on its type (static/dynamic) certain features are extracted. These are then classified using pre-trained SVM classifiers. The details of image extraction and classification is explained in the below section and not as an independent document. Please do not revise any of the current designations .There are two categories for vision-based hand gesture recognition. The 3-D hand model based method and appearance based method. The 3-D hand model based method works by comparing the input frames and the 2-D appearance projected by the 3-D hand model. However a huge database is required to deal with all the possible set of gestures in 3-D based model[1].Elastic graph matching was proposed to detect the gestures from the complex background but this approach was computationally complex and strict with its boundary conditions. The Zernike like moments can be easily used for gesture recognition to get the sample point from the required set of the gesture [2].

The frames preprocessing include skin color segmentation and Zernike like moments to detect the shapes of the gesture.

3.1. Image Capturing:

The task of this phase is to acquire an image, or a sequence of images (video), which is then processed in the next phases. The capturing is mostly done using a single camera with a frontal view of the persons hand, which performs the gestures.

3.2. Preprocessing:

Skin color segmentation is used to detect the body parts as we want to detect only the hand gesture we detect and eliminate using voila johns algorithm.

3.3. Feature Extraction

Zernike moments (ZM) are in general used to describe shapes. Zernike moments is used to identify the orientation of the hand.

3.4. Classification:

3.4.1 k-Nearest Neighbors.

This classification method uses the feature-vectors gathered in the training to find the k nearest neighbors in a n-dimensional space.[6] The training mainly consists of the extraction of (possible good discriminable) features from training images, which are then stored for later classification. Due to the use of distance measuring such as the euclidian or Manhattan distance, the algorithm performs relatively slowly in higher dimensional spaces or if there are many reference features, an approximate nearest neighbors classification was proposed, which provides a better performance. [3]

3.4.2 Hidden Markov Models. The Hidden Markov Model (HMM) classifiers belong to the class of trainable classifiers. It represents a statistical model, in which the most probable matching gesture-class is determined for a given feature vector, based on the training data. In [6], HMMs were successfully based to distinguish up to 40 different hand gestures with an accuracy of up to 91.9%. In order to train the HMM, a Baum-Welch re-estimation algorithm, which adapts the internal states of the HMM according to some feedback concerning the accuracy, was used. Sphinx is tool which is used to make the words from the letters, this tool can be easily refered as the framework to .NET , JAVA and other technologies[2]

IV. SYSTEM OVERVIEW

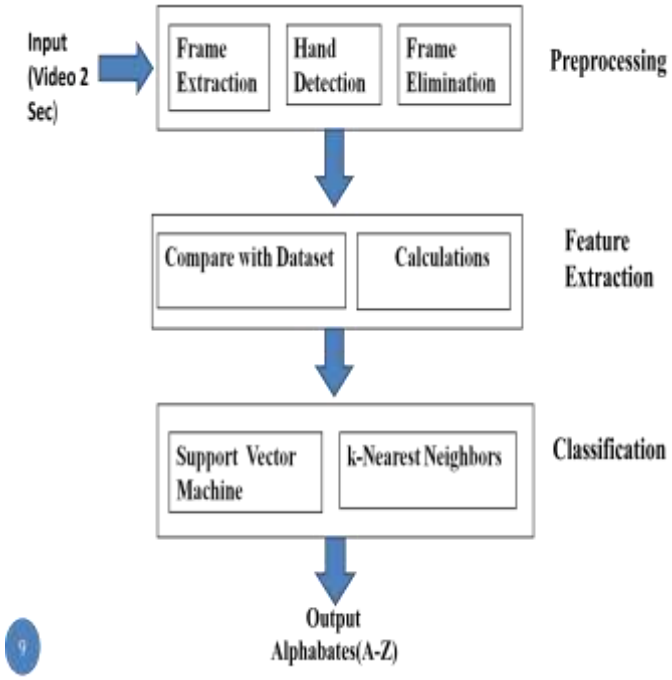

Fig.1: System overview

As discussed in previous section, the steps are pictured above. Now, the processing now involves the detection of the length and width of the palm . The length and width is found to compute the value of the total pixels from the frames and compare it with database .The second approach of processing is to detect the center of the palm as discussed [3] to make the angles using the samples points.

V. ALGORITHM AND PROPOSED SYSTEM

As discussed in [2] , the Zernike moments are used to find the shape and the moments of the gesture . By considering the different moments, we are going to apply our algorithm on it . The calculations will be formed out of the set of the gestures.

As discussed in [3] , after obtaining CoG of the gesture , we can easily find the angles from CoG to the moments . Our algorithm detects the height and width of gesture by finding

https://dx.doi.org/10.22161/ijaers.4.2.14 *ISSN: 2349-6495(P) | 2456-1908(O)*

the first and last most pixel of X co-ordinate (to find the width) , upper and lower cost pixel of gesture (to find the height). And accordingly the threshold will be calculated and compared with the dataset.

If two gestures found with same or somewhat same threshold value then angle will be found from CoG to different moments of the gesture. And accordingly the result will be displayed. We used SIFT for feature extraction.

The figure below shows the angle between CoG and the fingertips (obtained from Zernike moments).

Fig.2: Angle between CoG and fingertips

Fig.3: Zernike Moments

Dynamic gesture is discussed below. But Aditya ramamoorthy's "Dynamic gesture recogniton"[6] . static and dynamic gestures can be classified using task specific transitions . The time interval plays very important role in recognition if it is greater than specified time then the gesture is static else it is dynamic

5.1 Algorithm:

1. Start.
2. Input Video and Extract Frame.
3. Detects and eliminate face using Viola-Jones Algorithm.
4. Detect Hands using HSV color method.
5. Calculate Length and width of Gesture (Static).
6. Obtain CoG from the image.
7. Compare the distance covered by CoG between two set of frames if it is minimum then Gesture is static else it is dynamic

Gesture Recognition scheme can be broadly classified into two groups. In the first approach a gesture is modeled as a time sequence of states. One uses Hidden Markov Model, Discrete Finite State machine and variants of there of gesture recognition. In the second approach one uses dynamic time wrapping to compensate for the speed variations.

A dynamic hand gesture comprises a sequence of hand shapes with associated spatial transformation, parameters (such as translation, rotation, scaling/depth variations) that describes the hand trajectory.[4][5]

A task specific state transition machine is used to detect and differentiate between static and dynamic gestures. Dynamic Gestures are represented by combination of Cartesian space features and polar space features and recognized using an HMM based Framework.

5.2 Support Vector Machine and Classification

SVM is learning model of supervised learning that analyzes the data and also recognizes the patterns which are used for regression and classification , An SVM training algorithm also builds a model that particularly assigns the different examples into one category or any other and make is any other non probalistic binary linear classifier . This classifier gives the most probable points of the binary image .Different data points are there after the moments calculation and each point is belong two one of many classes then the main aim of the SVM is to decide which class a new data point will be in . Sample points are used to for the feature extraction further

5.3 Feature Extraction for Processing

The system can work maximum for 12 frames , so it is then important to find whether it is static or dynamic and the key frames for the comparison

Feature extraction involves following features:[1]

1. Trajectory length/location feature:

It is the distance between Cog and the sample points , it can be used to form the angles with different sample points and for comparison of it

2. Number of significant changes in hand orientation:

It helps to find whether the gesture is static or dynamic.

5.4 Gesture Recognition Challeneges:

1. **Latency**: Image Processng can be significantlyslow creating unacceptable latency for video games and other similar applications.
2. **Lack of Gesture Language:** Different users make gestures differently, causing difficulty in identifying motions.
3. **Performance:** Image processing involved in gesture recognition is quite resource intensive and the applications may found difficult to run on resource constraines device.

https://dx.doi.org/10.22161/ijaers.4.2.14 *ISSN: 2349-6495(P) | 2456-1908(O)*

VI. CONCLUSION

The system ease the processing time for recognizing static gestures using moments. There are further areas of improvement such as increasing system performance under robust and unfavorable environment. Dynamic gesture recognition can have alternate ways to improvements and it varies system to system. The system can achieve average performance of moderate time for one or two hand static and dynamic gesture, with which it is able to deal simultaneously.

REFERENCES

[1] Anup Kumar, Karun Thankachan and Mevin M. Dominic " Sign Language Recognition".
[2] Guillaume Plouffe , Ana-Maria Cretu , Pierre Payeur "Natural Human-Computer Interaction Using Static and Dynamic Hand Gestures".
[3] [3]Guillaume Plouffe and Ana-Maria Cretu, Member, IEEE "Static and Dynamic Hand Gesture Recognition in Depth Data Using Dynamic Time Warping".
[4] S. Padam Priyal and Prabin.K.Bora "A Study on Static Hand Gesture Recognition using Moments".
[5] Manisha M. Ingle Neelam K. Gilorkar. A review on feature extraction for indian and american sign language, 2014.
[6] Chih-Chung Chang and Chih-Jen Lin. A review on feature extraction for indian and american sign language, 2013.
[7] Aditya Ramamoorthya, Namrata Vaswania, Santanu Chaudhurya;∗, Subhashis Banerjeeb "Recognition of dynamic hand gestures" Pargomon journal 2016.
[8] Real-Time Palm Tracking and Hand Gesture Estimation Based on Fore-Arm Contour, 20 II .