International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Volume 4, Issue 2 (March-April, 2016), PP. 274-278

GESTURE RECOGNITION BY TREND BASED ANALYSIS

Abhishek Pratap Singh, Garima Chaudhary, Kumkum Yadav
Galgotia College of Engineering and Technology, Knowledge Park -2 Greater Noida Uttar Pradesh 201306
abhishek@cyroware.com, garimachaudhary125@gmail.com, kumkumyadav@gmail.com

***Abstract* — We present a method to recognise the motion path by using the standard sensors. The Idea is to recognise gestures based on motion. We propose to interpret gestures by trend based analysis. We directly use the angular momentum from the gyroscope and acceleration from the accelerometer to measure the motion path. A trend is a general direction in which something is developing or changing. A trend in a gesture can be explained as the motion trajectory observed from the values of the standard sensors.**

***Keywords* — Gesture, Handwritten words, Recognition, Accelerometer, Gyroscope, Interpretation, Retractable, Expandable.**

I. INTRODUCTION

We propose a way to recognise gestures observed from the sensors by identifying trend of the gesture. The use of gesture as a natural interface serves as a motivating force for research in modelling, analysing and recognition of gestures. A gesture can be consisting of one or more trends. A gesture can be broken down into combination of trends. We present a way to interpret gesture by using a trend based analysis. The method can be used to interpret different kinds of gestures that can be interpreted by the use of gyroscope and accelerometer. We will use raw values of acceleration and momentum to interpret the gesture trajectory.

The method we present can be used to interpret gestures and perform designated tasks. The gestures can be interpreted from any surface. When the gestures are performed with the sensors the raw value is received from the sensors the interpreter where by change we can predict whether a trend has changed or not by this we can predict the motion's path and perform adequate operations. The system can be used to interpret different patterns.
This methods has two levels which are characterised as follows:

- Hardware Interface
- Gesture Interpreter

These levels shows the hierarchy of the proposed method that we will be following but can be changed depending on the need of the system requirement. The data consisting of raw sensor values flows from the hardware interface to the gesture interpreter, where that data is divided into relevant datasets and irrelevant data sets. The relevant datasets are used to predict the current trend for the gesture and irrelevant gestures are not considered for interpretation and are just ignored. we will discuss about this in section (3). We will discuss about the hardware interface in the next section.

II. RELATED WORK

The devices that have been made use for gesture recognition use a mapping based technique to match the gesture that were preset for a certain task. by doing this the user can use the gesture on a particular surface to perform a task that needs to be mapped to the respective gesture for the system to work.

The Gesture recognition techniques used a static based techniques which are based on geometrical methods.a static geometry model is generated, from one or more images of a physical environment captured using a camera, using one or more static objects to model corresponding one or more objects in the physical environment. Interaction of a dynamic object with at least one of the static objects is identified by analysing at least one image and a gesture is recognised from the identified interaction of the dynamic object with the at least one of the static objects to initiate an operation of the computing device.

II. SENSORS

In this section we will discuss about the functions of the sensors and why we are using a particular sensor for our proposed method. The sensors are an integral part for this method to work and it is important to be keep in mind about the noise compensation from the sensors for an adequate output for the gestures interpreted.

The accelerometer measures the acceleration of the phone in three dimensions (or, equivalently, the forces acting on the phone). One main component of the measured acceleration is caused by gravitational forces. Additionally, the accelerometer measures any force applied to the phone that results in acceleration.

The gyroscope measures rotation about its own three axes. Most gyroscopes built into current smart phones use an oscillator to measure the Coriolis effect. The gyroscope measures angular velocity in rad per second.

III. HARDWARE

In this section we will discuss about the hardware structure for out approach of gesture recognition. Hardware consists of accelerometer and gyroscope, hardware provides the interpreter with values of acceleration along three axis and gyroscope provides angular momentum along three axis. For this proposal we suggest the use of raw values from the sensors.

A. Accelerometer Based Approach

The Hardware Interface consists of a Gyroscope and Accelerometer sensor at the tip of the pen which allows it to read the rotational and accelerated motion and then transmitted to the interpreting device which interprets the handwritten gesture to the appropriate character. The device (For Testing Purposes) is connected to the Interpreter with a wired connection. The hardware attaches to the back of the tip of writing implement and measure the rotational and accelerated. The accelerometer measures the acceleration exerted on the axis (X,Y,Z) where a = (a_x, a_y, a_z). Once the gravity is estimated it can be rotated into vector a´= (a´$\boldsymbol{x}$, a´$\boldsymbol{y}$, a´$\boldsymbol{z}$). Where the -Z Cos($\boldsymbol{\theta}$) axis is opposite to the Gravity Which is not equal to the gravity applied on it so to nullify that the -Z Sin($\boldsymbol{\theta}$) when squared and added to the Z Cos($\boldsymbol{\theta}$) where $\boldsymbol{\theta}$ is the angle made my then device tip with the

writing material **θ** also helps in determining a threshold for the system by judging at what angle the writing implement will write on the material by monitoring that we can be sure when a user is actually writing.

The Z axis is taken in separately as the Z axis change in acceleration are used to find when the user is writing when the user has stopped writing.

The measurements of the accelerometer are given in $m/s^{2.}$

B. Gyroscope Based Approach

The gyroscope measures the rotation in its own three axis. The gyroscope measures the angular velocity in rad/s. The System uses the Gyroscope to measure the angular motion of a writing implement to measure the circular stokes and motion when the user uses it to write on any writable surface. The Gyroscope provides the interpreter a combination of acceleration and angular motion of that point which is sent through to the user via communication channel. The system uses the approach of vectors instead of static point values which describes the trend of that point.

$$(X_N, Y_N) = (x_0, y_0) + \sum_{n=1}^{N} (V_{nx}, V_{ny}) \qquad (1)$$

$$(V_{nx}, V_{ny}) = (0, 0) + \sum_{i=1}^{n} (a_i x, a_{iy}) \qquad (2)$$

Where n and i represent a particular point of the hand written gesture as represented in eq(1) and eq (2). While it's good in theory but double summations are prone to noise and errors. To avoid offsets and error this is done at the time of Calibration of the device where the all the offsets are nullified like the value of **θ** which decides when to take the gesture and when to avoid any false gestures it is crucial for the system to know when a user wants is writing on a paper.

The Hardware does not count the interpretation happening it only records the gestures and decides when it needs to record a gesture and when to avoid it. It is also responsible for the proper Calibration of the system with all the values to be decided for error correction and all the compensation for the offsets. If the communication channel is not established between the channel and the Interpreter the device will be unable to work at all once the combination is established the data is transferred in the form of packets of acceleration, angular motion through the communication channel with the interpreter.

C. Communication Channel For Hardware Interface

The Communication channel is responsible for establishing a reliable and consistent connection between the interpreter and hardware interface. All the data sent through the communication channel is done in the form of packet each packet consists of acceleration and angular motion of the device at that moment. The data packet is received by the interpreter through the communication channel which needs to established prior to the interpretation of the gesture. Once the communication is established the initial data packet is the most important in the sense that it initialise the starting point of the Interpretation. The proposed channels for communication are WIFI-Direct or Bluetooth Channel.

The communication channel works on continuous small packets transfer rather than the favourable bulk transfer as there is no possibility of storing the point on the Hardware Interface, As the hardware has to be less power consumable and efficient it is important to have packet transfer continuously for better interpretation.

D. *Filter for raw values*

The filter is applied to the sensors to convert the values of acceleration into angle for each axis so as to minimise the noise that can be perceived by the sensors as accelerometers are very noisy and can cause unexpected values and errors to occur to minimise this we will use the angle value on the accelerometer which represents the acc due to gravity on each axis thus providing values in terms of G.

IV.. GESTURE INTERPRETER

In this section we will discuss about how the system will interpret gestures performed by the user. The Gesture Interpreter gets the values from the hardware i.e angular momentum, acceleration which are belonging to each character in the system as it's unique properties that were assigned at the time of Calibration. These properties consists of:

- Angular velocity with respect to time
- Change in acceleration

These values help the interpreter to interpret the trends and perform actions according to the input gesture which helps the system to interpret the gesture better. The method doesn't require specific components to work. The interpreter consists of :

- Communication Channel
- Interpreter
- User Profile Manager
- Error Correction
- Graphical User Interface

A. Communication Channel For Gesture Interpreter

In this section we will discuss about how the system will Communicate with the hardware. The Communication Channel establishment between the Hardware Interface and Gesture Interpreter allows the Interpreter to interpret gesture using the Hardware Interface. Without the established and reliable communication channel existence between the Interface and Interpreter the device won't work.

The Hardware Interface uses the Communication Channel and transmits a packet consisting of the vector representation of the acceleration and angular velocity at each different point of motion and is received by the Interpreter.

In any communication channel lose of data and corruption of data can occur, to deal with such "rogue" data packets the Interpreter in the case of damage packets is said to irrelevant when the change between the previous point and present point is marked as irrelevant and moves on. The channel can be turned on by the Hardware Interface which on the side for the Gesture Interpreter when turned on will listen to request of Hardware Interpreter to connect once the connection is established then the interpretation can be performed.

The communication channel can wary from a wired connection to wireless connection channel that help the hardware to interact with the Gesture Interpreter to communicate. It is important for this method to work to have a reliable communication between the hardware and the gesture Interpreter so that the device can get the continuos packet.

B. Interpreter

In this section we will discuss about how the system will Interpret the gestures when connected to the Hardware

through communication channel. The Gesture interpreter receives vectors that represent value of acceleration along the axis (X, Y, Z) and Angular velocity along the axis (X, Y, Z). The value of the acceleration and angular velocity is used of any two point are used to find the change in the values between the two point if the change in value does not passes the threshold value then the next point is not considered relevant and is moved on to find another relevant point.

To interpret the gesture the system uses a "Trend based analysis" which is based on the changes in a specific trend. When a change in a specific trend is noticed then the trend is considered and different combinations of trends provide a complex gesture interpretation. The Trend based analysis allows the users to predict by the trajectory and motion of the device. When the user will use the device for the first time it is important for the user to use the device and perform gestures to provide the system with the type of style user might use for the future use. Once the device is initialised the system creates a digital imprint which creates a database of all the user gesture and value written at the time of initialisation which are stored in a vector form. when the user is using the device, the device listens and receives a lot of packets consisting of information about the points which are then filtered and relevant points are taken. These points are then used for matching and significantly reduce the list of the possible results stored at the time of initialisation these checks are made for only the relevant values and irrelevant values are ignored and not considered this allows fewer matching operations and faster prediction for the gesture recognition. The system creates a list of the possible result as the user starts using the device. As more and more relevant values are found the list is sorted for possible results. The vectors values represent a trend and change in trend when taken in combination. For the case of misinterpretation or a gesture not recognised the Graphical User Interface allows the user to fix misinterpretations by the device.

C. User Profile Manager

The purpose of User Profile Manager is to create the user imprint of the style of the user about how a user uses a device that can consists of angle, momentum, speed. This makes the profile of user to be unique and easily adapt to the user. This makes for better and dynamic interpretation of gestures.

The profiles of users are stored in the application and not on the cloud so that all the users data is stored on the device and is always available at any time.

D. Error Correction

The purpose of Error Correction is to compensate for any deviation that can occur by the user at the time of writing. It's impossible to maintain all the different styles a user ay write a character for that we propose the use of vectors representing checkpoints for any gesture the vector form is a relational representation of the trend observed by the system. As the user writes using the devices some offsets can be observed but the trend remains somewhat similar to the initial gesture. The offset can be observed in the actual character but it can be divided into the checkpoints that represent the trend of the character written which can be used to interpret the handwritten character. The system creates a list of all the vectors representing the trend of a character written at the time of initialisation. The system compares the trend at the time of writing the character which is done only at the relevant points knows as check point.

The Graphical User Interface provides an interface for the user to interact with the interpretation of the system for error in interpretation, The Graphical User Interface allows the user to interact with the system and inform system about the any mistakes happened in the interpretation of any gesture (which we will discuss in the next section). When the system is not able to find a proper match with the system it stores the trends in a temporary state if the user changes that state or system automatically correct.

E. Graphical User Interface

The GUI (Graphical User Interface) provides the user with an interactive interface to the user that provides the output as well as an interface to the user to edit the interpreted data. The GUI also saves the document on the basis of the Heading which gets followed by the initialiser symbol " # ". The system uses the heading for the management of document. It creates a folder and saves the document on the basis of Time-Stamp. This keeps a record of all the documents which can be used as a management tool as well as provides a tool to edit any interpreted document.

The GUI provides the user to export the document into a text editable file and provides a management tool as well to keep a record of all the documents interpreted. These provide an easy management of records on the basis of Time Stamp and Heading. Heading becomes a bundle which holds the different documents on the basis of Time-Stamp records. The GUI also consists of auto-correction tool which serves a purpose of managing spelling mistakes and rectifying mis-interpreted values. The Interpreter feeds the interpreted words to the GUI as an output similar to any input device like Keyboard which make it easier for the system to easily correct any mistakes.

The GUI works as a separate module that all other module. It can be used as a text editor and interpretation. All the modules work parallel to other modules to provide better performance and faster interpretation.

The GUI can be present either on a users device or can be used to control other devices. The GUI can be used to create more gestures or check the results of gestures.

V. IMPLEMENTATION

For this method we implemented this method in a Pen Gesture Recognition which implement this method to predict user written character on any surface. The Characters can be written on any surface and the trend of that character gets identified for that character and the resultant character is printed on the screen.

The accelerometer is used to predict the values of gravity in degrees in terms of G. The values change as the trend changes. A filter is applied to convert raw values into angles along X, Y, Z axis. The sensors are used to predict the nature of trends and different combinations of trends form a character. The device gets accustomed to one user and ones set has to reseted to be calibrated according to user again. This functionality has not been implement but is a future scope. The device is adjusted to a particular user style and as the user uses the device to write anything on an writing implement it gets interpreted to a digital document which can be saved as a file. The device consists of two sensors exiting in one board i.e.

- Accelerometer
- Gyroscope

The sensor board used is BerryIMU which is a inertial Measuring Unit which measures measures and reports on velocity, orientation and gravitational forces, using a combination of an accelerometer, gyroscope and a magnetometer. BerryIMU is also fitted with a barometric sensor (BMP180) which can be used to calculate altitude. A temperature sensor is also included but we wont be using it in this project.

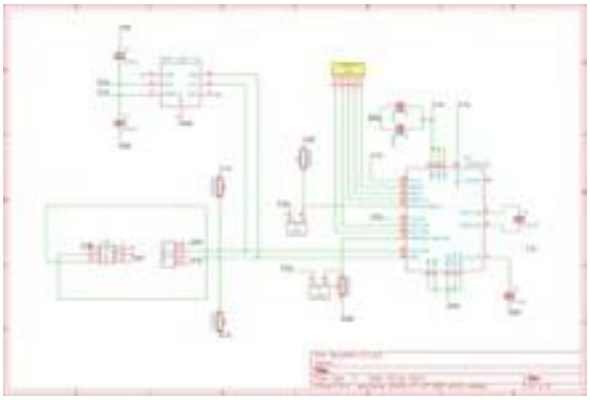

Fig. 4.1 BerryIMU Structure(This Fig shows the structure of the IMU)

For the interpreter we are using a raspberry pi B as it provides an environment to directly communicate with the sensor and work in the Python Development Environment to implement intelligence for the interpreter. We have implement Kalman Filter for the sensor value filtering.

The Kalman filter, also known as linear quadratic estimation (LQE), is an algorithm that uses a series of measurements observed over time, containing noise (random variations) and other inaccuracies, and produces estimates of unknown variables that tend to be more precise than those based on a single measurement alone. More formally, the Kalman filter operates recursively on streams of noisy input data to produce a statistically optimal estimate of the underlying system state. The filter is named for Rudolf (Rudy) E. Kálmán, one of the primary developers of its theory. Kalman filter is: "It's a method of predicting the future state of a system based on the previous ones." and was used to assist with navigation in the Apollo 11 moon landing. The Raspberry Pi communicates with the hardware with the use of i2C smbus which communicates with the system bus present in the raspberry pi.

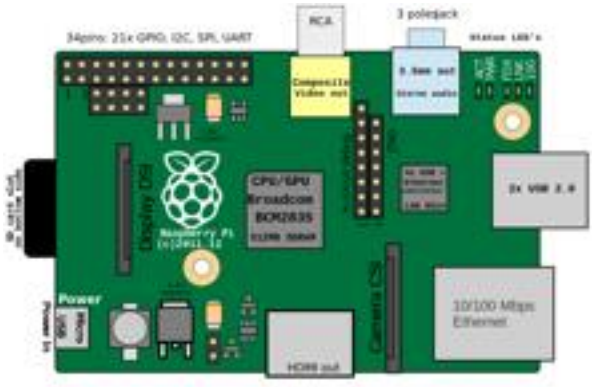

Fig. 4.2 Raspberry Pi

After the implementation it was observed that the error varying from 0% to 20% which is a time dependent when used for prolonged used the error increases to 20%. So we can conclude that 80% accuracy is observed.

VI. CONCLUSION

The device allows the interpretation of gestures by using the trend based analysis. The analysis is based over motion and dividing the trends. On the basis of the Trends for the purpose of interpreting the trends allows the device to remove limitations. It allows the system to be used on any surface and to be used by anywhere. The vectors representing each gestures are derived according to user's style which are based and stored in a relational vector representation of a gesture. By dividing the input gesture into data sets allows the interpretation to be done by not mapping a gesture to a stored existing gesture limits the different styles a user may use the device of a user may vary which can result in inconsistencies. The proposed system removes these limitations and introduces the interpretation of gestures using motion and Trend-Based analysis allows the interpretation of pen gestures based on motion to be more dynamic and allows the user to use any gestures on any surface. The GUI allows the user to edit the interpreted gesture. The system serves as a tool for management and gesture interpretations.

VII. ACKNOWLEDGEMENT

I acknowledge with gratitude to assistant professor, Mr.Manish Kumar Sharma, my respective teacher, who has been sincere and helpful in making me understanding the different systems of research and problems in my proposal paper

I would like to thank Ms.Garima Chaudhary and Ms.Kumkum Yadav for providing the support and making efforts in making this paper possible.

VIII. FUTURE SCOPE

We will implement the method to use neural network based approach to increase accuracy and better evolution cycles to interpret the user handwriting, In this paper we have not covered the elements of GUI and how GUI will look as in now its a command line prompt system that present the user with the words interpreted. We are using a standard wired connection for the communication channel. The method can be used to control other devices as well by using a device that acts as a controller for multiple devices and performs actions based on different gestures.

REFERENCES

[1] Jeen-Shing Wang and Fang-Chen Chuang An Accelerometer-Based Digital Pen With a Trajectory Recognition Algorithm for Handwritten Digit and Gesture Recognition, *IEEE Sens. J., vol. 59, no. 07, pp. 1543–1551, July 2012.*

[2] Miss Leena mahajan ,and Prof G,.A.Kulkarni Digital Pen for Handwritten Digit and Gesture Recognition Using Trajectory Recognition Algorithm Based On Triaxial Accelerometer- A Review

[3] Mr. Kunal J. Patil1, Mr. A. H. Karode, and Mr. S. R. Suralkar (2014) Gesture Recognition of Handwritten Digit using Accelerometer Based Digital Pen, *Vol 3, Issue 4.*

[4] Ouais Alsharif, Tom Ouyang, Fracoise Beaufays, Shumin Zhai, Thomas Breuel and Johan Schalkwyk (2015) Long Short Term Memory Neural Network For Keyboard Gesture Decoding.

[5] Ronald J Williams and David Zipser, "A learning algorithm for continually running fully recurrent neural net- works," *Neural computation*, vol. 1, no. 2, pp. 270–280, 1989.

[6] Thomas Deselaers, Danial Keysers, Jan Hosing and Henry A. Rowley (2015) GyroPen: Gyroscopes For Pen-Input With Mobile Phones. *IEEE Transactions On Human-Machine Systems, Vol. 45, No. 2.*

[7] Vincent Vanhoucke, Matthieu Devin, and Georg Heigold, "Multiframe deep neural networks for acoustic modeling," in *Proceedings of ICASSP*, 2013.

[8] Yoshua Bengio, Patrice Simard, and Paolo Frasconi, "Learning long-term dependencies with gradient de- scent is difficult," *IEEE Transactions on Neural Net- works*, vol. 5, no. 2, pp. 157–166, 1994.

[9] Z. Dong, U. C. Wejinya, and W. J. Li, "An optical-tracking calibration method for MEMS-based digital writing instrument," *IEEE Sens. J., vol. 10, no. 10, pp. 1543–1551, Oct. 2010.*

[10] Zhenzhen Kou, Daisy Stanton, Funchun Peng, Fracoise Beaufays, and Trevor Strohman (2015) Fix It Where It Fails: Pronunciation Learning By Mining Error Corrections From Speech Logs.