Dynamic Hand Gesture Pattern Recognition Using Probabilistic Neural Network

Debasish Bal
Dept. of Electrical and Electronic Engineering
International Islamic University Chittagong
Chittagong-4318, Bangladesh
debasishopu@gmail.com

Asif Mohammed Arfi
Dept. of Computer Science and Engineering
University of Yamanashi
Takeda 4-4-37, Kofu, Yamanashi, Japan
arfieee@hotmail.com

Sujoy Dey
Dept. of Computer Science and Engineering
University of Science and Technology Chittagong
Chittagong-4202, Bangladesh
sujoy.ustc2011@gmail.com

2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) | 978-1-6654-4067-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/IEMTRONICS52119.2021.9422496

***Abstract*—Flex Sensor and Gyroscope based hand gloves are being used widely to detect gesture-based sign language patterns to provide aid for speech-impaired people. However, detecting dynamic gestures is not an easy task due to the variability and redundant implementation precautions. This article approaches a dynamic hand gesture recognition system by using Probabilistic Neural Network (PNN) deliberately focused on practical usage of dynamic gestures on day-to-day life. In this experiment, 10 gesture patterns have generated by using a data glove. For each gesture, 360 training input vectors are generated to train the PNN model, and the output has provided via a speaker.**

Keywords—flex sensor, gyroscope, dynamic gesture, PNN

I. Introduction

There are approximately 360 million people around the world who are deaf and mute [1]. Due to speech and hearing impairment, these people have to communicate by using sign language. By using a convenient Man-Machine Interface (MMI), it is possible to generate the record and validate such sign language gestures to mediate a communication system between mute and unmute people. Kinect sensor-based works are proposed in previous years to design gesture recognition prototype [2]–[6]. Kinect is a depth sensor that has designed by Microsoft. However, these systems are not portable and cheap. User needs stationary Kinect sensor on board to work with it. Authors have used Convolutional Neural Network (CNN) to detect static gestures from the user on [7], [8]. Static sensor-based work also has done using Time-of-flight (ToF) camera [9] and Dynamic Time Wrapping algorithm [10]. As these works are based on the camera sensor, they are highly dependent on the image quality, lighting condition, camera angle and many more factors. There is much approach to detecting various versions of sign language as well. Authors approached to detect Vietnamese, Indian and American Sign Language [11], [12]. Most of these approaches are not capable of detecting dynamic gestures.

This proposed system will approach a dynamic gesture recognition system to eliminate these problems at a much affordable solution as well as portability. This system also consists of a speaker module to generate Bangla speech as output for each corresponding gesture.

II. Methodology

A. Hand Gloves Design

The hand glove has designed using flex sensors attached with the fingers.

- Three flex sensors have used for gesture recognition.
- These three sensors have mounted in index, middle and ring finger.
- An MPU-6050 which is a three-axis and accelerometer and three-axis gyroscope has used. This device uses I2C protocol for communication and has 16-bit built-in ADC channel for high accuracy.
- A Raspberry Pi 3B has used for interfacing the flex sensors and the MPU-6050.
- An MCP-3008 ADC has used to interface the flex sensors with Raspberry Pi as Pi does not have inbuilt analog pins.

B. Mathematical Model

The mathematical models are used to process raw data from sensors and using them for gesture recognition purpose. The first 20 reading of sensors at the time *t* for generating a training set, can be expressed as:

$$\{x^n_{g,k,1}(t),\ x^n_{g,k,2}(t), x^n_{g,k,3}(t), \dots x^n_{g,k,20}(t)\} \tag{1}$$

Where x is the signal value, n is the user index, k is the replicate index, and t is the time.

A replicate of a gesture by a person can be expressed as:

$$\Pi_{n,g,k} = \{x^n_{g,k,1}(i.dt), x^n_{g,k,2}(i.dt), x^n_{g,k,3}(i.dt) \dots x^n_{g,k,20}(i.dt) : i \in \{0, 1, \dots D_{n,g,k}\}\} \tag{2}$$

Where dt is the interval between sampling of sensors data and $D_{n,g,k}$ is the last reading from sensor by n^{th} user for g^{th} gesture.

C. Overall Framework

The proposed gesture recognition system consists of three stages of processing. The framework consists of three main procedures where the value of gyroscope is preprocessed first. In the next stage, the raw data from flex is retrieved. After gathering these data, the sensors data have trained and tested by PNN further to analyze the recorded gesture. The PNN has specifically chosen to solve the classification problem for each dynamic gestures. The overall system has demonstrated in Fig 1.

D. Reading MPU-6050

The MPU-6050 is consists of a three-axis gyroscope and accelerometer integrated on a single chip. This sensor is also called 6 Degree of Freedom (DoF) device because of its six outputs. It measures the rotational velocity of the rate of change of the angular position over time in X, Y, and Z-axis.

Micro-Electro-Mechanical Systems (MEMS) technology has used inside this sensor.

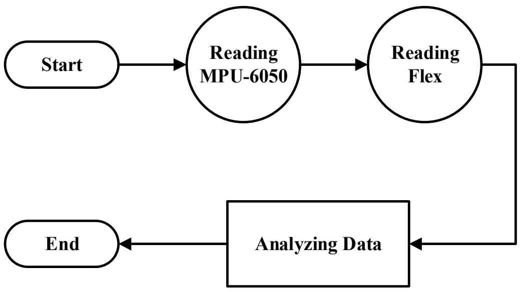


Fig. 1. Overall demonstration of the system

The outputs of the gyroscope are measured in degrees per second. The unit for measuring acceleration is meter per second squared. To read the sensor, we have used the wire library after that the by resetting the power management register; the sensor was reset. The register address is as follows:

TABLE I. MPU-6050 REGISTER ADDRESS

Register	B7	B6	B5	B4	B3	B2	B1	B0
0x6B	DEVICE_RESET	SLEEP	CYCLE		TEMP_DIS	CLKSEL[2:0]		

As the datasheet says, the data of each axis is stockpiled in two registers. The addresses of these registers are shown in Table 2.

TABLE II. MPU-6050 ACCELEROMETER DATA REGISTERS

Register	B7	B6	B5	B4	B3	B2	B1	B0
0x3B	ACCEL_XOUT[15:8]							
0x3C	ACCEL_XOUT[7:0]							
0x3D	ACCEL_YOUT[15:8]							
0x3E	ACCEL_YOUT[7:0]							
0x3F	ACCEL_ZOUT[15:8]							
0x40	ACCEL_ZOUT[7:0]							

To read all the registers value correctly, at first the value of X, Y and Z-axis for first six register are requested. After that, we read the second register, perform 2's complement and combine them appropriately to generate the actual values. Here, the roll φ and pitch θ angles are calculated for normalized accelerometer reading G_p.

$$\frac{G_p}{\left\|G_p\right\|} = \begin{pmatrix} -sin\theta \\ cos\theta sin\varphi \\ cos\theta cos\varphi \end{pmatrix} \Rightarrow \frac{1}{\sqrt{G_{px}^2 + G_{py}^2 + G_{pz}^2}} \begin{pmatrix} G_{px} \\ G_{py} \\ G_{pz} \end{pmatrix} = \begin{pmatrix} -sin\theta \\ cos\theta sin\varphi \\ cos\theta cos\varphi \end{pmatrix} \tag{3}$$

$$tan\varphi_{xyz} = \left(\frac{G_{py}}{G_{pz}}\right) \tag{4}$$

$$tan\theta_{xyz} = \frac{-G_{px}}{\sqrt{G_{px}^2 + G_{pz}^2}} \tag{5}$$

Similarly, the register values of Gyroscope are:

TABLE III. MPU-6050 GYROSCOPE DATA REGISTERS

Register	B7	B6	B5	B4	B3	B2	B1	B0
0x43	GYRO_XOUT[15:8]							
0x44	GYRO _XOUT[7:0]							
0x45	GYRO _YOUT[15:8]							
0x46	GYRO _YOUT[7:0]							
0x47	GYRO _ZOUT[15:8]							
0x48	GYRO _ZOUT[7:0]							

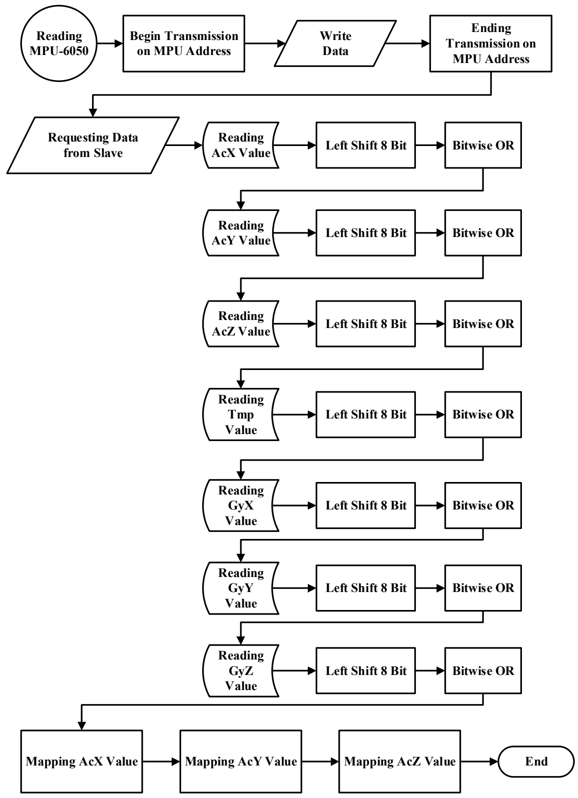


Fig. 2. MPU-6050 Data Reading Procedural

E. *Reading Flex Sensors*

Flex sensors are ideal for measuring the amount of deflection. Generally, these sensors can be capacitive, fiber optic or conductive ink-based.

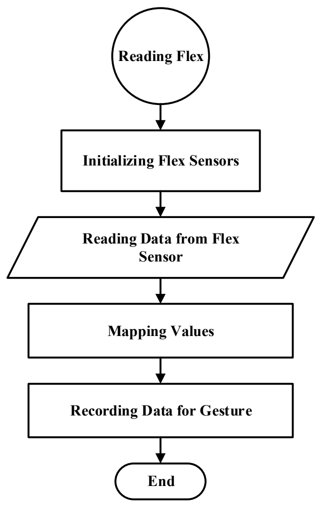


Fig. 3. Flex Sensor Data Reading Procedural

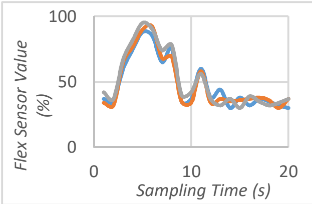
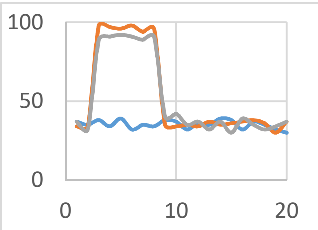
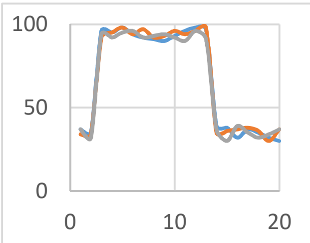
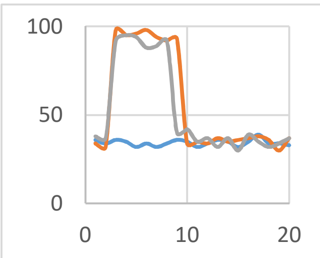
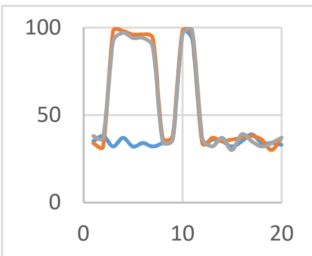
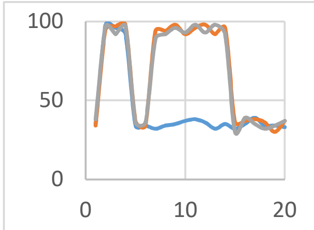
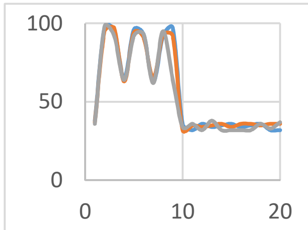
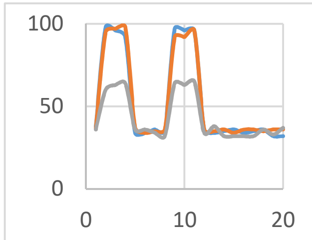
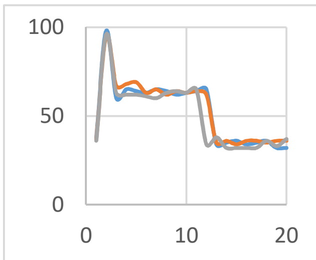
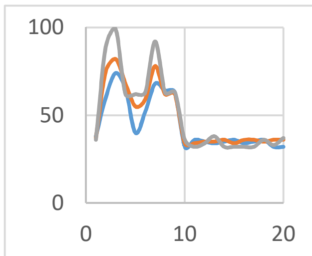
We have used conductive ink-based flex sensors here. Flex sensor works as a variable pin resistor. The output from a flex sensor can be determined from the following equation.

$$V_{out} = V_{in}\left(\frac{R_1}{R_1 + R_2}\right) \quad (6)$$

F. Building Gestures

Gestures have a changeable property, which changes dynamically every time we execute a particular gesture. Therefore, it is necessary to record a gesture for more than once to identify it correctly. The gesture properties are plotted below.

We noted the parameters by recording the values for flex sensors and gyroscope from 88 volunteers for 20 times. We have trained each gesture for 1760 times using a Probabilistic Neural Network algorithm. The algorithm implemented in TensorFlow 1.15 with Python 3.5 GPU support on Nvidia 720m card. The learning rate was 0.1 at 180 epochs.

No	Bangla Speech	Flex Sensor Data
1	আসসালামুয়ালাইকুম (Greetings!)	
2	আমার নাম অপু (My name is Opu.)	
3	আমি চট্টগ্রাম থা□ । (I live in Chittagong.)	
4	আপনার নাম □ ? (What's your name?)	
5	আমি ভালো আছি। (I'm fine.)	
6	আপনি □ মন আছেন? (How are you?)	
7	এই দি□ আসুন। (Please come here.)	
8	এখানে বসুন। (Please Sit down.)	
9	আমি এ□ দম ভালো নেই। (I am not well.)	
10	বিদায়। (Goodbye!)	

III. RESULT

To evaluate the performance, the system was tasted with 23 disabled people. Every person has repeated each gesture for 20 times. Therefore, each gesture was tested for 460 times. In the case of gestures 7 and 10, the gesture recognition algorithm was having a problem to detect them accurately. The values of gyroscope are considered at this position to resurrect the detection issue, which was successful for both gestures.

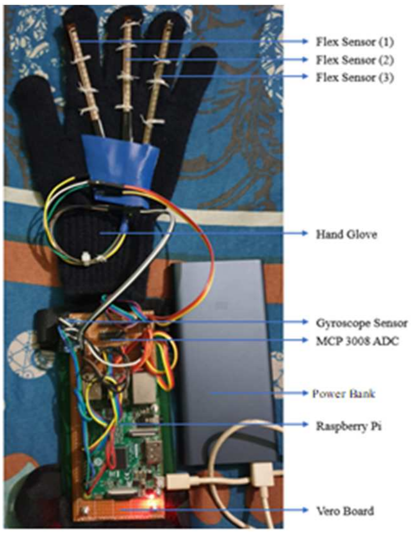


Fig. 4. System Implementation

The Gyroscope data has plotted in Fig 5 to discriminate between the gesture 7 and 10.

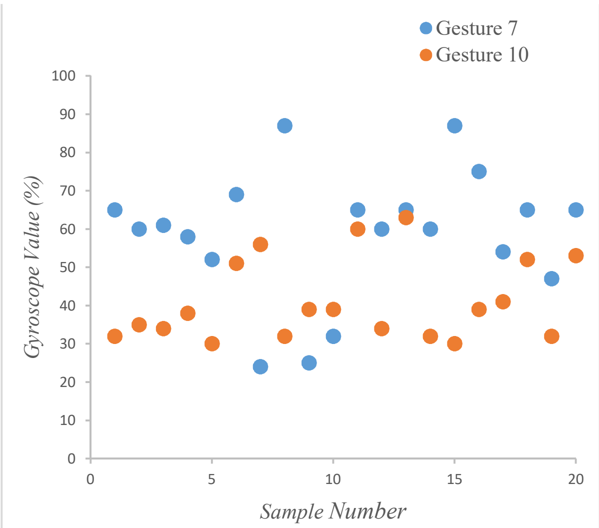


Fig. 5. Gyroscope Datapoints for Gesture 7 and 10

The system has performed overall 92.7% accuracy. The performance can be improved by adding more training datasets.

IV. CONCLUSION

Our proposed hand gloves have made using three flex sensors. Primarily, that limits our capability to track finger gesture for the remaining two fingers. However, due to this purpose, the device's overall cost is less, which is a tremendous advantage from a consumer perspective. On the other hand, the dynamic gesture that has been used here is entirely chosen by us. As these are not universally recognized patterns, the user may need to adapt to the patterns on the first hand. But the output generated using dynamic gestures are considerably faster than static gesture patterns like American Sign Language (ASL).

Mainly we have proposed a method to recognize dynamic gesture patterns using flex sensor and gyroscope which can be used not only in speech but also in drone control, haptics, augmented reality etc. We have converted the data generated from gloves into a two-dimension matrix where in x-axis sampling time and in y-axis flex sensor's resistances are plot. Thus, it will give us a visual idea about the data. This proposed method helped to identify gestures from different user even though the user response wasn't the same as the training data because we have used a form of pattern recognition.

REFERENCES

[1] J. Banks, L. Hayes, L. Smith, and K. Millsap, "Deafness and Hearing Loss," in *Encyclopedia of Diversity in Education*, 2013.

[2] L. Liu, X. Wu, L. Wu, and T. Guo, "Static Human Gesture grading based on Kinect," in *2012 5th International Congress on Image and Signal Processing, CISP 2012*, 2012, pp. 1390–1393.

[3] H. Li, L. Yang, X. Wu, S. Xu, and Y. Wang, "Static hand gesture recognition based on HOG with kinect," in *Proceedings of the 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2012*, 2012, vol. 1, pp. 271–273.

[4] T. Q. Vinh and N. T. Tri, "Hand gesture recognition based on depth image using kinect sensor," in *Proceedings of 2015 2nd National Foundation for Science and Technology Development Conference on Information and Computer Science, NICS 2015*, 2015, pp. 34–39.

[5] S. Bhattacharya, B. Czejdo, and N. Perez, "Gesture classification with machine learning using Kinect sensor data," in *Proceedings - 2012 3rd International Conference on Emerging Applications of Information Technology, EAIT 2012*, 2012, pp. 348–351.

[6] S. Bessa Carneiro, E. D. F. M. De Santos, T. M. A. De Barbosa, J. O. Ferreira, S. G. S. Alcalá, and A. F. Da Rocha, "Static gestures recognition for Brazilian Sign Language with kinect sensor," in *Proceedings of IEEE Sensors*, 2017.

[7] X. Guo, W. Xu, W. Q. Tang, and C. Wen, "Research on optimization of static gesture recognition based on convolution neural network," in *Proceedings - 2019 4th International Conference on Mechanical, Control and Computer Engineering, ICMCCE 2019*, 2019, pp. 398–400.

[8] X. Wang, Z. Chen, X. Wang, Q. Zhao, and B. Liang, "A comprehensive evaluation of moving static gesture recognition with convolutional networks," in *Proceedings of 2018 3rd Asia-Pacific Conference on Intelligent Robot Systems, ACIRS 2018*, 2018, pp. 7–11.

[9] G. Simion and C. D. Caleanu, "Multi-stage 3D segmentation for ToF based gesture recognition system," in *2014 11th International Symposium on Electronics and Telecommunications, ISETC 2014 - Conference Proceedings*, 2015.

[10] G. Plouffe and A. M. Cretu, "Static and dynamic hand gesture recognition in depth data using dynamic time warping," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 2, pp. 305–316, Feb. 2016.

[11] E. Abraham, A. Nayak, and A. Iqbal, "Real-Time Translation of Indian Sign Language using LSTM," in *2019 Global Conference for Advancement in Technology, GCAT 2019*, 2019.

[12] L. T. Phi, H. D. Nguyen, T. T. Q. Bui, and T. T. Vu, "A glove-based gesture recognition system for Vietnamese sign language," in *ICCAS 2015 - 2015 15th International Conference on Control, Automation and Systems, Proceedings*, 2015, pp. 1555–1559.