Smart Gloves: A novel 3-D Work Space Generation for Compound Two Hand Gestures

Harsh Bhardawaj, Mohit Dhaker, K. Sivani, H. R. Nandi Vardhan
Dept of Electrical and Electronics Engineering
Amrita School of Engineering, Bengaluru
Amrita Vishwa Vidyapeetham
India
+91-080-25183700

harshbhardawaj@outlook.com, mohitdhaker24@gmail.com, sivanikumar3@gmail.com, hr_nandivardhan@blr.amrita.edu

ABSTRACT
Speechlessness is a colossal barrier of communication between the ordinary people and the speech impaired. This paper presents a novel methodology with a working model which is used to convert the sign language to speech in order to help speech impaired people. It uses flex sensor and Inertial Measurement Unit in order to determine the position of finger as well as the position of hand in 3-Dimensional space. These sensors are embedded into the gloves, which when processed, outputs the accurate gesture which has been made by the user, thereby making it smart. The paper also presents a unique division of a matrix in 3-Dimensional space comprising of states. These states have to be estimated in a generalized manner in order to be used by anybody irrespective of their gender or height. The paper also highlights the use of both hands for compound two hand gestures with static and dynamic gesture recognition system. The smart gloves can further be used in a variety of applications such as motion sensing gaming, remote medical diagnosis, and robotics.

CCS Concepts
• Human-centered computing → Human computer interaction (HCI) → Interaction techniques → Gestural input

Keywords
Smart Gloves; Sign language; Flex Sensors; Inertial Measurement Unit; Three-Dimensional State Estimation Method; Gaming; Robotics; Medical field

1. INTRODUCTION
Over 7.68% of the total population of India suffers from the disability of speech impairment [1]. Sign language is a tool that helps the speech impaired people to communicate with the society.

But not every citizen can understand the common sign language used by the speech impaired people which is specially been invented for them. The idea of the project is to design smart gloves which can convert sign language to speech.

The literature indicates a few different approaches for this. However, the two most commonly used techniques in order to recognize the sign language are the vision-based hand gesture recognition system [2] and sensor-based hand gesture recognition system. The camera based or vision-based system is used to track the hand gestures and recognize it [3]. In [4] an idea is proposed about using the Douglas-Peucker algorithm for gesture recognition using image processing. But, it is comparatively less user friendly because of the bulky hardware components which is not easily portable and the processing of the images will take a considerable amount of time [5]. In addition, if the user is in a crowed area multiple gestures will be picked up from the people around which may lead to an incorrect output. Hence, the camera-based system is not feasible [6]. In the work [7] a reverse approach is employed to bridge the communication gap, by converting English words into Indian sign language (ISL).

In the previous method, the gloves that was designed, identifies the gestures of the right hand by using flex sensors. To achieve this, it employs a state estimation method. But, this was an arbitrary method and the states were not generalized. This paper discusses about the design of the gloves for the left hand with a novel 3-D workspace generation. This is achieved by using a simple matrix method to generalize the states and use the identified states to recognize the gesture using two hands. The gesture is converted it into speech using Text -to-speech (TTL) and the output is played through loud speakers.

2. RELATED WORK
In the recent past, there have been various approaches to recognize the hand gestures but they had some limitations of recognition rate, time and generalization. The recent method was developed using flex sensors and Inertial Measurement Unit (IMU). In this method, to detect the finger orientation, flex sensors are embedded behind the finger. As each finger is bent, there will be a corresponding change in radius of curvature of the flex sensor, which in turn changes its resistance. This gives the finger positions. IMU module, placed on the forearm is capable of tracking the movement of the hand and the coordinates in 3D space where the hand moves were named states [6]. Finally, the readings from IMU and flex sensors are combined to obtain the output.

Our research is focused on utilization of flex sensors and IMU. Hand-talk Gloves proposed by Ambika Gujrati et.al, introduces the idea of using flex sensors. To track the motion of hand in 3-D

ICCCV '18, June 15–18, 2018, Singapore, Singapore.

ACM ISBN 978-1-4503-6470-6/18/06…$15.00.
https://doi.org/10.1145/3232651.3232670

space, an IMU is used. The IMU is placed on forearm which will help us to trace the motion of hand in any given direction. The end position can be evaluated by integrating the acceleration values obtained from inertial sensors [4]. Angular coordinates (yaw and pitch) used to trace the hand orientation and movement are obtained from Berry IMU and GY87 in this proposed method.

3. METHODOLOGY

3.1 Process Model

The flex sensors are employed on each finger of both the gloves to track the finger orientation. An IMU is placed on the fore arm to track the movement of hands in 3-Dimensional space. Additionally, a flex sensor is fixed on the elbows of both the hands in order to find the depth parameter that will be used to find the states. Compound two hand gestures are the area of focus of the current design. Raspberry Pi and Arduino Uno are used as the processing units. A block diagram of the system is shown in Figure 1 and Figure 2.

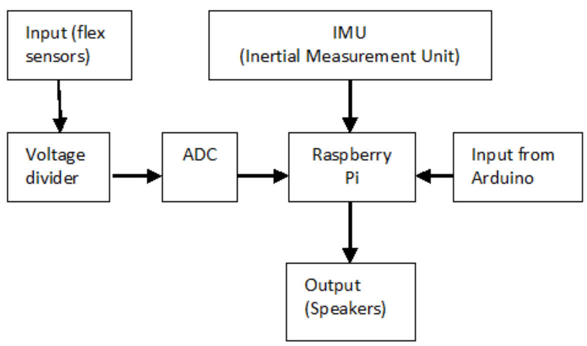

Figure 1. Hardware Block Diagram for the right-hand glove

Figure 1 is the hardware block diagram of the right-hand glove, starting from the flex sensors. The values from the flex sensor is in the form of resistance, which is analog. The changes in resistance is successfully mapped to voltage by employing a voltage divider circuit. The obtained voltage is fed to Raspberry Pi as a digital input. Therefore, the voltage values obtained from the voltage divider circuit is fed to MCP 3008 (ADC) and then given as input to Raspberry Pi, along with the other inputs from IMU and Arduino. This is then processed by the Raspberry Pi and the final output is played through the loud speakers.

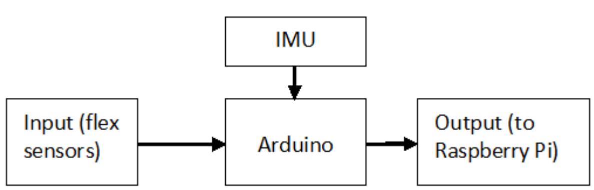

Figure 2. Hardware Block Diagram for the left-hand glove

Figure 2 depicts the hardware block diagram for the construction of the left-hand glove. Here, the Arduino takes the input from the flex sensors and IMU and transmits the data to the Raspberry Pi serially using USB cable. Thus, the output of Arduino is an input to Raspberry Pi. The values obtained from flex sensor and IMU module are used to identify the position of hand in three-dimensional space. The algorithm comprises of two following steps:

- Finger Orientation Detection
- 3-D Orientation Detection

The algorithm binds the data from various sensor on both hands and collates to find out the finger orientation and the three-dimensional position of the hand in space. This data is used to determine the correct gesture. The algorithm for the enhanced model is as shown in Figure 3.

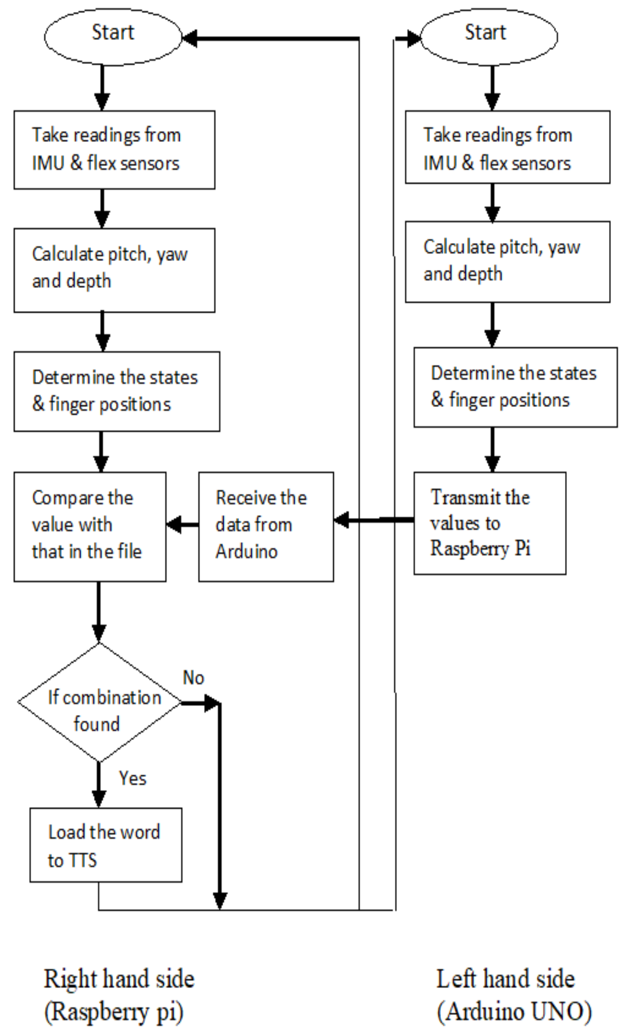

Figure 3. Enhanced Algorithm for compound gestures.

As shown in Figure 3, initially, the algorithm running on Raspberry Pi takes the input from flex sensor and IMU model which are used to calculate depth parameter, pitch and yaw basis. After reading values from the voltage divider, for each finger and the elbow, the position of the fingers and depth parameter is identified. Using the modified method of state estimation as discussed in section 3.3, the states in which the hand moves can be determined.

Similarly, for the left hand, once a start has been initiated by the user, the Arduino takes input from flex sensor and IMU module and calculates the pitch and yaw. Here, the usage of a voltage divider is not necessary as the Arduino has an inbuilt ADC. After reading values for each finger and the elbow from the Arduino, the position of the fingers and depth parameter is identified. Using the novel method of state estimation as explained in section 3.3, the states in which the hand moves can be determined.

The data from left hand is used along with the data from right hand to form the search string for the gesture. This string is matched with the database to identify the gesture. The text corresponding to the gesture identified in the database is appended into the text file. The TTS engine is used to voice out the text for this file. If there is no corresponding match found the algorithm iterates again.

3.2 Finger Orientation Detection

To identify the position of fingers, the flex sensors are fixed along the length of fingers. Flex sensors consists of a thin flexible substrate of carbon resistive elements. When the fingers bend, the flex sensors also bend along with it. During the process of bending, the area of the carbon substrate varies and so does the resistance value correspond to that degree of bend. Basically, as the bend increases, the resistance also increases.

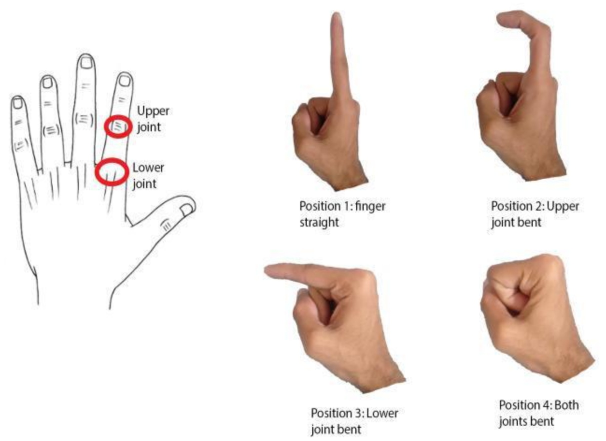

Figure 4. Finger joints under consideration and possible finger positions [2]

The various position of the fingers with respect to finger joints are shown in Figure 4. Two 2.2" flex resistor is used for the thumb of each hand and ten 4.5" flex resistors are used for the other eight fingers and two elbows [6]. The flex sensors are biased with the voltage of 5V and the output of theses is connected to the ADC which converts the value into a corresponding digital value. The position of the finger is chosen based on the set voltage ranges. For finger position 1 of the straight finger then If θ1 is the range for position 1 viz. a straight finger

Example:
if (digital_voltage_value> θ1):
position="1"

The voltage ranges with respect to various finger positions for the left hand, including that of the elbow are tabulated in Table 1 below.

Table1. Voltage range corresponding to all the fingers for right hand

Finger	Voltage Range (Digital Value)	Finger Position
Thumb Finger	> 41000	1
	< 41000 and > 39400	2
	< 39400	3
Index Finger	> 42600	1
	< 42600 and > 41600	2
	< 41600 and > 40600	3
	< 40600	4
Middle Finger	> 42000	1
	< 42000 and > 40000	2
	< 40000 and > 38500	3
	< 38500	4
Ring Finger	> 39600	1
	< 39600 and > 37600	2
	< 37600 and > 35500	3
	< 35500	4
Little finger	> 36500	1
	< 36500 and > 33500	2
	< 33500 and > 30200	3
	< 30200	4
Elbow	> 41500	3
	< 41500 and > 39000	2
	< 39000	1

3.3 Enhanced State Estimation Algorithm

A state is a 3-D space which is identified by a range of pitch, yaw and depth values. They are used as the coordinates to find the coordinates of the position of hands in an imaginary 3D workspace which is geometrically cuboidal in shape. In the previous work [6], the states were identified on repetitions without any concrete geometrical shape. The enhanced state estimation algorithm uses a standard shape which is a generalized mechanism. The previous algorithm doesn't cater to portability as it has to be restructured based on the user as height and the length of the hand hampers the idealized states. The three sides of the workspace are identified as the depth, yaw and pitch as shown in Figure 5.

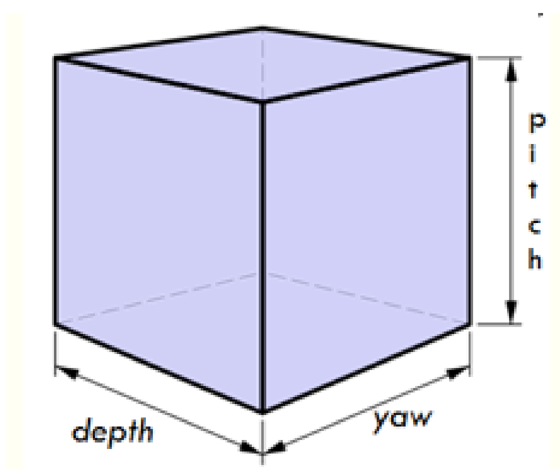

Figure 5. 3-D Representation of Pitch, Yaw and Depth

The work area has been divided into a three-dimensional matrix of the order $4 \times 3 \times 3$ (Length=3 units, Breadth=3 units, Height = 4 units) inspired by the antenna shaping in [8], comprising of 36 states. When the hand is moving along a horizontal plane, a maximum deviation of 360 (-180 to 180) can be observed. This is the yaw basis value calculated from the IMU reading, strategically placed on the fore arm. When the hand moves horizontally, the yaw basis value changes with respect to the starting position of the hands. The ranges of these yaw values are grouped into three categories resulting into either '1, 2, 3' as the final reading. The hand can reach up to a maximum deviation of 180 °(-90 to 90) in the vertical plane. These values are the pitch values which is calculated from the accelerometer. The depth ranges from 1 to 3. A three-dimensional representation of states in the modified state estimation method is shown in Table 2.

Table 2. 3-D state values for different depth

Depth=1			Depth=2			Depth=3		
1	2	3	13	14	15	25	26	27
4	5	6	16	17	18	28	29	30
7	8	9	19	20	21	31	32	33
10	11	12	22	23	24	34	35	36

The depth parameter is one of the dimension used for the consruction of a 3D matrix for the modified state estimation method. It is recognized as the breadth of the imaginary cuboid which is depicted in Figure 5. Similary, the pitch and yaw are the height and length of the cuboid respectively.These cubiods are stacked on the basis of depth values as seen in Table 2. Thus, to identify the position of hand in the specific distance away from body, we have to find the depth using the flex sensor placed on the elbow as shown in Figure 6.

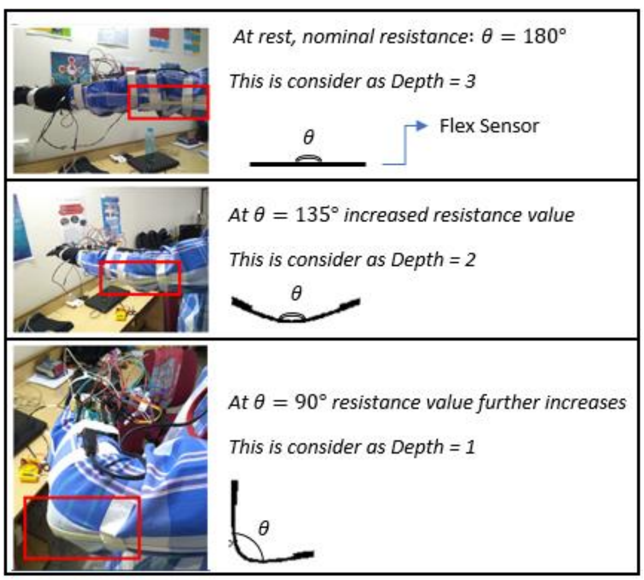

Figure 6. Detection of depth parameter using flex sensor at the elbow

3.4 RESULT

The experiment was setup and the modified state estimation algorithm was used to determine compound two hand gestures. The string below shows the word corresponding to string values of the gestures that were watched out along with the equivalent action word.

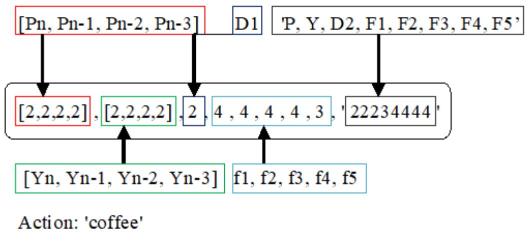

Figure 7. Output of the result on the screen

In Figure 7, 'Pn to Pn-3' and 'Yn to Yn-3'represents the four consecutive pitch and yaw values respectively for the right hand, where 'Pn' and 'Yn' are the latest pitch and yaw values. D1 is the depth corresponding to the right hand when a particular gesture is being performed and f1, f2, f3, f4 and f5 represents the five fingers of the right hand starting from little finger to the thumb. For the left hand, 'P' represents the current pitch value, 'Y' represents the current yaw value, 'D2' represents the depth value and the remaining five values are for the finger positions, starting from the thumb. The corresponding gesture "coffee" for the read values are shown in the Figure 8.

Figure 8. Gesture for ' coffee'

The string so obtained is compared with the database. Once a matching pattern is found, the respective action word is fetched and appended to the text file. The TTS engine will give a voice output. The sensor values for "sorry", "welcome", "good" are tabulated in Table 3.

Table 3. Outputs with their equivalent sensor values and gestures

[Pn, Pn-1, Pn-2, Pn-3]	[Yn, Yn-1, Yn-2, Yn-3]	[D1, f1, f2, f3, f4, f5]	['P, Y, D2, F1, F2, F3, F4, F5]	Voiced output
[3,3,3,3]	[2,2,2,2]	[1,4,4,4,4,3]	[41311111]	sorry
[2,2,2,2]	[2,2,2,2]	[2,1,1,1,1,1]	[22211111]	welcome
[3,3,3,3]	[3,3,3,3]	[3,4,4,4,4,1]	[41311111]	good

Thus, the motion of a human hand in 3D space can be successfully traced with the help of flex sensors and the IMU modules. This will help the speech impaired people to voice out the gestures, thereby enabling an efficient way of communication with the external world.

3.5 CONCLUSION

The values from the flex sensors from both the hands are compared along with the values from the IMU in order to identify the coordinates of the position of hand in the newly constructed 3-D matrix. This is then compared with our database of values to identify the exact gesture.

Vision based hand gesture recognition system uses camera to capture image and then the captured image is processed and the desired feature (hand gesture) is extracted. This extracted information is used for the recognition of gesture [2]. The disadvantage of using vision-based gesture recognition system is that the background should be clear and the person whose gesture has to be recognized should be in focus in order to avoid the ambiguity of unwanted gesture recognition. The sign language to speech conversion using flex sensors studied in [9] is limited in its static and dynamic recognition as it doesn't track the entire motion of the hand. Our previous work [2] has a main disadvantage that the evaluation of position of hand is in 2-D space. As all the possible combination cannot be realized by only be restricting the system in 2-Dimension space.

In this paper, the newly constructed gloves work efficiently as these problems have been rectified by the generation of a novel 3D workspace along with the algorithm associated with it. The glove can be used to navigate a wheel-chair just by changing the TTS system with a suitable system as proposed in [10]. The gloves can also find its applications in gaming industry, defense and production industries and in various medical fields:

Gaming industry: Hand gestures are very useful in virtual reality games, where the user can control the players in the game by moving his hands.

Defense and production industries: The gloves could be used to control a robotic arm which is capable of building products in industries in order to reduce time and human efforts. It could also be used in bomb diffusion by incorporating haptic feedback, avoiding the risk of hazards caused to humans during the process.

Medical field: To aid people with Cerebral palsy which is a permanent movement disorder that appears in early childhood. It can also be used in remote medical surgery where the operation can be performed by the robots itself as they will be remotely controlled by the doctor without being physically present in that location.

4. REFERENCES

[1] Census of India 2011, 2011. *http://www.censusindia.gov.in /2011census/ population_enumeration.html*

[2] Y. Zhu, Z. Yang and B. Yuan 2013.Vision Based Hand Gesture Recognition.In *Proceedings of the International Conference on Service Sciences (ICSS).* Shenzhen, 260-265.

[3] P. Vijayalakshmi, M. Aarthi 2016. Sign language to speech conversion. In *Proceedings of the International Conference on Recent Trends in Information Technology (ICRTIT).*

[4] M. Geetha, Menon, R., Jayan, S., James, R., and Janardhan, G. V. V. 2011.Gesture recognition for American sign language with polygon approximation. In *Proceedings - IEEE International Conference on Technology for Education, T4E.* Chennai, Tamil Nadu, 241-245.

[5] K. K. Dutta, Satheesh Kumar Raju K., Anil Kumar G. S. and Sunny Arokia Swamy B. 2015. Double handed Indian Sign Language to speech and text. In *Proceedings of the Third International Conference on Image Information Processing (ICIIP).* Waknaghat, 374-377.

[6] Abhijith Bhaskaran K., Anoop G Nair, Deepak Ram K., Krishnan Ananthanarayanan, H. R. Nandi Vardhan 2016. Smart Gloves for Hand Gesture Recognition. In *Proceedings of the International Conference on Robotics and Automation for Humanitarian Applications (RAHA 2016).*

[7] R. J. Raghavan, Prasad, K. A., Muraleedharan, R., and M. Geetha, 2013. Animation system for Indian Sign Language communication using LOTS notation. In *Proceedings of the Emerging Trends in Communication, Control, Signal Processing Computing Applications (C2SPCA).*

[8] Chia-Hsu Kuo, Chun-Cheng Lin, Jow-Shiun Sun 2017. Modified Microstrip Franklin Array Antenna for Automotive Short-Range Radar Application in Blind Spot Information System. In *IEEE Antennas and Wireless Propagation Letters.*

[9] Ambika Gujrati, Kartigya Singh, Khushboo, Lovika Sora and Mrs. Ambikapathy 2013. Hand-talk Gloves with flex Sensors: A Review. In *International Journal of Engineering Science Invention ISSN (Online).* vol. 2, no. 4, 43-46.

[10] Rajesh Kannan Megalingam, A. Thulasi, A., and R. Krishna, R.2012. Methods of wheelchair navigation: Novel gesture recognition method. In *International Journal of Applied Engineering Research*. vol. 7, 1654-1658.