

Low-Cost Intelligent Gesture Recognition Engine for Audio-Vocally Impaired Individuals

Rishikanth C[□], Harini Sekar[†], Gautham Rajagopal^μ, Ramesh Rajesh^Ξ, Vineeth Vijayaraghavan^Ω
Dept. of Electrical and Electronics Engg.[□], Dept. of Electronics and Communication Engg.^{†μ}, Department of Research^{ΞΩ}
Sri Venkateswara College of Engineering[□], Easwari College of Engineering^{†μ}, Solarillion Foundation^{ΞΩ}
Chennai, India
(rishikanth.c.in[□], harini.sekar.in[†], gautham.r.in^μ, rajeshramesh90^Ξ, vineethv^Ω)@ieee.org

Abstract — This paper presents a low-cost open source prototype of a gesture recognition and interpretation glove that aids audio-vocally impaired individuals in communication. A sensor based glove for recognition of the gestures in the American Sign Language (ASL) is the input module to the system. It incorporates resistive flex sensors and contact sensors appropriately positioned to obtain a theoretical efficiency of 96% in character recognition. When the sensors are subjected to change in orientation by the user, the variation in electrical resistance of each of the individual sensors is obtained as the input to an Arduino ATmega328 microcontroller that accurately maps it to the letters of the English alphabet through an efficient algorithm. An elegant methodology that is easy to implement using a microcontroller is employed for the mapping process to minimize the complexity of the system. The paper also presents a detailed study of the efficiency characteristics involved in the construction of the custom-designed flex sensors utilized in the system. The paper outlines the calibration and testing results of the sensors. The highlights of the system are its ergonomic and economical design, portability and intelligent gesture recognition and interpretation ability. The implemented system achieves a practical efficiency of up to 80% of theoretical efficiency. The cost of the system is approximately less than USD 5 at laboratory condition using off the shelf components.

Keywords — calibration; character recognition; gesture recognition; microcontrollers; open source hardware; prototype; sensors

I. INTRODUCTION

The World Health Organization (WHO) estimates that about five percent of the total population in the world suffers from hearing loss [1]. The major disadvantage for these individuals lies in the reduced or nullified proximal interactivity with peers having normal hearing and speech. Their interaction with the digital world is also very limited due to the lack of affordable support technologies that aid in their communication. Due to this, deprivation from access to modern technology is also an impediment to their progress in the present scenario.

According to gestural theory, development of the human language originated from simple gestures that were used for communication [2]. For several centuries, sign language that comprises of various gestures has been a major tool for communication among the audio-vocally impaired. An estimation of 13 million people comprising of those with

normal hearing capacity as well as audio-vocally impaired persons use the sign language just in the United States [3]. Considering the third world and developing countries, the previous estimation scales to a much larger number of individuals who need wider access to interventions that help them to interact productively.

The evolution of gesture based recognition systems was initiated with the invention of the Sayre Glove in 1977 which extends to the modern day gesture-based applications in the field of sign language translation [4]. Researchers are now trying to bridge the gap between the advancements in interpersonal communication technology and its adoption by the audio-vocally impaired.

Due to the above, there has been an increase in the innovation of systems consisting of sensor based gloves that act as a recognition engine to interpret various sign languages and translate them to voice output. However the major demerit of these systems is that they are complex and are not cost effective. Developing a low cost, gesture recognition engine along with a detailed study of the construction and calibration of custom-designed open source sensors is the main aim of this paper.

In this paper, a sensor based glove, designed using custom made flex sensors and contact sensors, aid in gesture recognition of the American Sign Language (ASL). The choice of ASL was done based on the fact that it is widely learned as a second language, serving as a bridge language [5].

In our system, the positioning of the flex and contact sensors are done so as to obtain maximum character recognition efficiency of the English alphabet. The sensors connected to the voltage divider circuit generate analog values as input for every unique gesture, through an Arduino ATmega328 microcontroller. The unique values derived are mapped efficiently to the gestures of the ASL.

This paper also expounds on the construction and calibration of a flex sensor. The very widely available, industrially manufactured flex sensors are not cost effective. Due to this, the objective to achieve a low cost recognition engine is faltered. Hence open source, low cost, lab-made, bidirectional, sensors were constructed and calibrated to

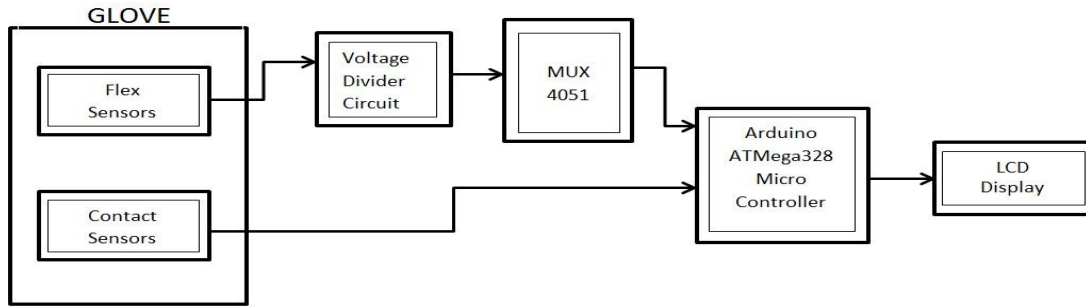


Figure 1. System Block Diagram

produce efficient results at a minimal cost of 0.016 USD per sensor. The system is ergonomic in terms of design, very cost effective and can be afforded by low income individuals and bottom-of-the-pyramid (BoP) societies.

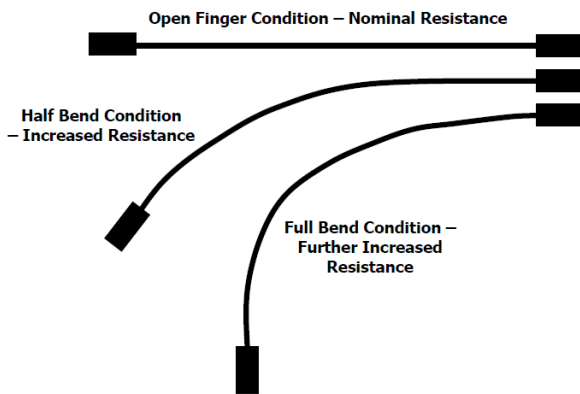


Figure 2. Flex Sensor - Bend Vs. Resistance

II. RELATED WORK

Prior work in the field of gesture recognition using sensor based gloves for character recognition that serves as a useful tool for the audio-vocally impaired individuals has been done.

Yasir Niaz Khan et al [6] developed a sensor based glove that recognizes various gestures of the ASL and translates them into sentences. In the system suggested, the sensor values are recognized using artificial neural networks which involve complex computation. However, the system implemented in our paper uses a simpler method that maps the gestures of the ASL to the English alphabet directly, thus avoiding complex processing. Celestine Preetham et al [7] developed a gesture recognizing glove that generates speech by translating sign language, to aid the speech-impaired. The system suggested made use of low cost piezo-resistive sensors and an accelerometer to recognize the gestures/sign language. Duy Bui et al [8], in their system of gesture recognition of Vietnamese sign language, use a sensing glove made of MEMS accelerometers. The usage of accelerometers in designing the sensing gloves for detecting the gestures proves to be expensive. Hence, in our system, an alternative is designed by positioning the flex sensors and contact sensors to

track the changes in the wrist movement as well, thus making the system economical. Yousuf Khambat et al [9], in their paper presented a GesTALK system that helps the hearing impaired people to interact. Our implementation is a low cost version where character recognition is done through an innovative seven bit mapping whereas the authors use a twenty six bit mapping scheme for the same.

III. PROBLEM STATEMENT

Devices aiding the audio-vocally impaired individuals help translating sign languages/gestures to corresponding audio output. This simplifies the interaction of such individuals. The devices currently made available utilize expensive sensors and complex circuitry. The number of characters recognized and the respective efficiency of such systems is very high. However, such systems are unaffordable for the low income groups.

Hence, the need to develop a low-cost, open source, sensor based, gesture recognition engine, that efficiently interprets the ASL and maps it to the English alphabet by using an Arduino ATmega328 microcontroller. Our prototype details on the construction and calibration of the custom-made, low cost flex sensors that were utilized in designing the gesture recognition glove. Our system uses a simple protocol for ultra-fast detection and response, leading to accurate translation of the gestures.

IV. METHODOLOGY

The system comprises of a glove interfaced with an Arduino ATmega328 Microcontroller for gesture recognition of the ASL. The glove is embedded with flex sensors and contact sensors. The flex sensors are connected to a voltage divider circuit to produce the desired range of values for the various gestures. The contact sensors are connected to the microcontroller through resistors. Resistors are included in the circuitry for limiting the flow of current. The various flex sensors connected to the voltage divider circuit are multiplexed using MUX4051. The output of the MUX is given to the Arduino ATmega328 microcontroller which interprets the various gestures made by the user. This is done with the use of a predefined identification table that is discussed in section C of the methodology. The recognized characters are then

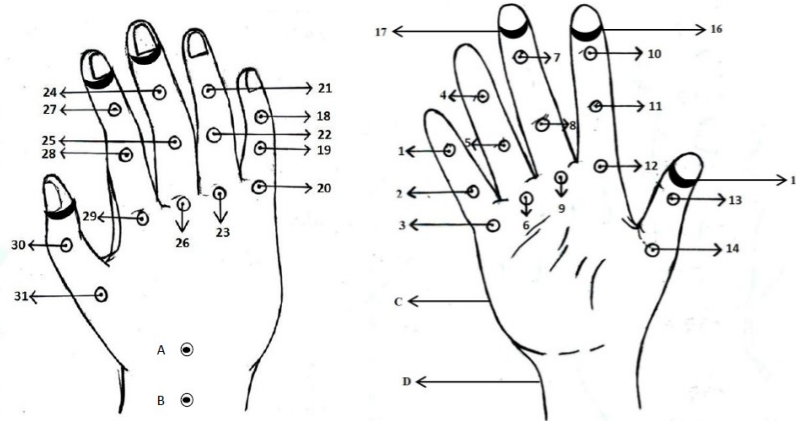


Figure 3. Flex and Contact Sensor Positioning – Physical Sites – (a) Left - Back, (b) Right - Front

displayed in the LCD display by the microcontroller. The system block diagram is illustrated in figure 1.

A. Design Of Glove

The input module to the system is a gesture glove with embedded sensors. The construction of the glove plays a key role in the character recognition efficiency of the system. Initially the glove was designed by positioning the flex sensor one on each of the index finger, middle finger, ring finger and the little finger. However these custom-designed, low cost sensors did not produce traceable difference in range for the open finger condition, half bend and full bend. This is illustrated in figure 2.

The choice of dimensioning and the positioning of the flex sensors are described in detail in the flex sensors subsection. Based on this, the choice was made to position two 5cm flex sensors in each of the index, middle and ring fingers and one flex sensor in the little finger. The system could now accurately recognize 5 characters of the English alphabet and hence the theoretical system efficiency was 19%. To increase the number of characters recognized by the system, Contact Sensors were placed at the tips of the thumb, index finger and middle finger. Now, 14 out of the 26 English letters could be obtained which increased the theoretical efficiency of the system 69%. To further enhance the productivity of the system two more flex sensors were added to track the wrist movement of the user while gesturing and by these, 25 English alphabets could be recognized and thus the percentage theoretical efficiency of the system is maximized to 96%.

1) *Sensors*: Sensors are transducers that measure a physical quantity and convert it to a signal that can be read by the instrument with which it is interfaced. In the system described in this paper, the sensors used are flex sensors and contact sensors. These sensors measure the physical quantity obtained as an analog value. These are converted to digital values by mapping to a scale of 0-1023 by the Arduino ATmega328 microcontroller through its 10-bit Analog to Digital converter for digital interfacing. The values obtained

while testing and calibration of various sensors belong to this range of values. Flex sensors are made using inexpensive conductive bags, which are readily available off the shelf components.

a) *Flex Sensors*: Popularly known as bend sensors, these produce a variation in resistance with a change in the bend angle. To develop a low cost system, use of industrially available sensors is not feasible due to the high market price of about USD 7.95 for a unidirectional flex sensor of length of 2.2 inches. Hence custom designed flex sensors were built in the lab reducing the cost of a sensor to USD 0.017. The making of the flex sensors was done in two methods.

i) *Development Stage 1*: Two layers of velostat conductive bags that are cut out according to required dimensions are sandwiched between two jump wires of required length, for the two terminals. A layer of masking tape is wound over the conductive bag layer to bind the components together. The custom-designed sensor was then calibrated. However, it was found that between the open finger-condition, half bend and full bend, the sensor produced a narrow difference in range on a scale of 0-1023. To develop on this method's demerits, an enhanced method in Development Stage 2 was adopted.

ii) *Development Stage 2*: To increase the efficiency of the sensors made using Development Stage 1, a third layer of velostat conductive bag was introduced in between the two layers, which increases the range of difference between the various bends. This method produced a remarkable difference in range of 200-400 on a scale of 0-1023 between the various bends after calibration.

The difference in range produced due to the various bends also depends on the length of the flex sensors. 50 Sensors in each of the lengths 15cm, 11cm, 9cm and 5cm were constructed and calibrated.

The calibration was done through a calibration program. Thousand sensor readings for each position were taken, for a

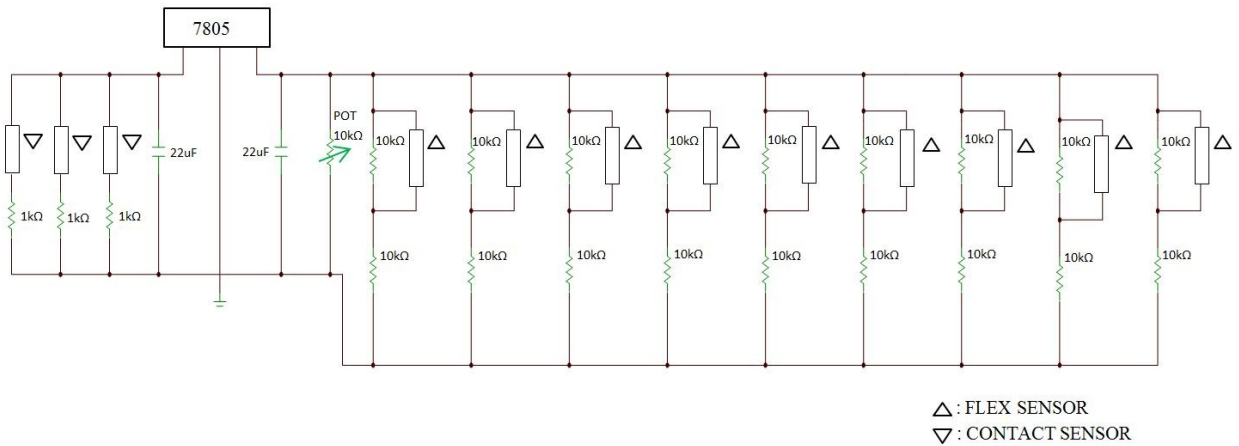


Figure 4. Circuit Diagram

TABLE I. CALIBRATION VALUES - FLEX SENSORS

LENGTH OF SENSOR	STRAIGHT	HALF BEND	FULL BEND
5 cm	504	645	720
9 cm	538	659	743
11 cm	503	681	847
15 cm	595	680	901

single flex sensor positioned on each of the fingers. Each reading was an average of 2000 readings sampled at 0.0625μs (sampling frequency of microcontroller). The result of this calibration is shown in table I.

TABLE II. SENSOR NOTATIONS

Sensor Notations	Sensor Specifications
S1	Flex sensor of length 5 cm
S2	Flex sensor of length 9 cm
S3	Flex Sensor of length 11 cm
S4	Flex Sensor of length 15 cm
C1	Contact Sensor placed at the tip of the thumb
C2	Contact Sensor placed at the tip of the Index Finger
C3	Contact Sensor placed at the tip of the Middle Finger

Selection of flex sensors was done based on the comparison of calibration table of all the sensors. Table II denotes the physical sensor specifications of various sensor notations. Based on the tables, decision was made to use flex sensors of length S2 and S1, as these sensors produced a remarkable difference in range between the open finger-condition, half bend and full bend.

The positioning of flex sensors on the glove played a critical role in gesture recognition. The positioning was done in three methods, with method three producing optimal result.

- Method 1: Initially, on the index finger S4 was placed covering all the joints to sense the Open finger condition, half and full bends of the finger. The positioning of sensors in the middle, ring and little fingers were done in a similar manner. However, this method was not feasible as the difference in the range of values between the various bend angles were not sufficient and led to generation of false positive signals. Hence, the values could not be mapped accurately.

- Method 2: To overcome these false positive signals, a new method of positioning was devised. In this method, on the index finger, S2 was placed on the point 29, as shown in figure 3 (a), to sense the full bend of the finger and S1 was placed on the inner side of the index finger between the points 10 and 12, as illustrated in figure 3 (b), to sense the half bend of the finger.

The positioning of sensors on the middle, ring and little fingers were done in a similar manner. The positioning of flex sensors on the inner side of the fingers was done for better detection of the various degrees of bends. The shortcoming of this method was that the S2 positioned on the index, ring and middle fingers showed values for half as well as full bends owing to the length and this could not be disabled.

- Method 3: This was the most efficient method of positioning, wherein the demerits of the above two methods were overcome and a wide difference in range of values for the various bends were obtained. In this method, two S1's were utilized for each of the index, middle and ring fingers.

TABLE III. CHARACTERS RECOGNIZED - FLEX ONLY

CHARACTERS RECOGNIZED BY EMBEDDING FLEX SENSORS IN THE GLOVE
B, F, K, M, W

In the index finger one S1 was positioned on point 29 as shown in figure 3 (a) to track the full bend of the finger and another S1 was placed between points 10 and 11, shown in

figure 3(b), on the inner side of the index finger to track the half bend of the finger. Similar positioning of flex sensors was done for the ring and middle fingers. For the little finger only one S1 was placed on point 20 to track the various bends. This method produced best results and was adopted for designing the glove. Characters recognized by embedding flex sensors as described in this section are illustrated in the table III.

b) Contact Sensors: Contact sensors are those which produce output only when they are in contact with an object. These sensors are introduced in the design of the glove to increase the number of characters recognized by the system. In the design implemented, contact sensors are rings made of aluminium foil placed at the tip of the thumb, index and middle fingers. These are indicated as 15, 16, 17 in figure 3(b). The support circuitry are resistors connected in series with each of the three sensors further connected to a 5V source. Characters recognized after introduction of contact sensors in the design of the glove as described in this section are illustrated in table IV (a & b).

TABLE IV. (a) CHARACTERS RECOGNIZED – INCLUSION OF CONTACT SENSORS (b) FLEX AND CONTACT

CHARACTERS RECOGNIZED AFTER INCLUSION OF CONTACT SENSORS
A, C, D, E, L, O, S, U, V, Y, Z, T, X

CHARACTERS RECOGNIZED AFTER INCLUSION OF FLEX AND CONTACT SENSORS - CUMULATIVE
B, F, K, M, W, A, C, D, E, L, O, S, U, V, Y, Z, T, X

c) Flex Sensors for Tracking Wrist Movement : For further improving the number of characters recognized by the system, two additional flex sensors are employed. The positioning of the flex sensors between points A,B and C,D respectively as highlighted in figure 3 (a & b) will help in tracing the wrist movement thus increasing the output of the system. The characters recognized due to introduction of these sensors in the design of the glove in illustrated in the table V (a & b).

TABLE V. CHARACTERS RECOGNIZED – (a) INCLUSION OF WRIST TRACK FLEX SENSORS (b) CUMULATIVE

CHARACTERS RECOGNIZED AFTER INCLUSION OF WRIST TRACK FLEX SENSORS
G, H, I, J, P, Q, N

CHARACTERS RECOGNIZED CUMULATIVELY BY GLOVE
B, F, K, M, N, W, A, C, D, E, L, O, S, U, V, Y, Z, G, H, I, J, P, Q, T, X

B. Hardware Implementation

Figure 4 illustrates the circuitry implemented for designing the glove. When the circuit was implemented, due to loading effect the tail end sensors received a voltage, less than the ideal value of 3.3V. This led to misrepresentation of signals for the various bends. To overcome this, the battery supply is

routed through a voltage regulator which provides a constant supply of 5V. The voltage regulator used in the circuit for this purpose is IC 7805. The constant 5V supply is further scaled down using a potentiometer to obtain a stable supply of 3.3V.

C. Software Implementation

The analog values obtained from the glove are converted into digital values using a software comparator. Each sensor value is represented by a bit. The bits obtained from the each of the sensors are shifted and then converted into their respective decimal values through computation. The decimal and binary values corresponding to each alphabet is illustrated in table VI. The various characters mapped to their corresponding decimal place values stored in an array.

Each letter of the English alphabet is represented by a train of bits obtained from the various sensors while gesturing, by the user. The train of bits obtained is converted to the respected decimal values. The conflicts due to identical decimal values are recognized and the correct letter is identified. This is done by implementing a conditional logic that is capable of differentiating the gestures based on inputs from the sensors used for tracking wrist movement and the contact sensors. A predefined array of size 127 is initialized where the letters of the English alphabet are stored. The locations of the stored letters are indicated by their respective previously computed decimal values.

TABLE VI. CHARACTER RECOGNITION DECIMAL - BINARY

Alphabet	Decimal	Binary
A	127	1111111
B	0	0000000
C	42	0101010
D	31	0011111
E	127	1111111
F	96	1100000
G	95	1011111
H	87	1010111
I	126	1111110
J	126	1111110
K	23	0010111
L	31	0011111
M	85	1010101
N	87	1010111
O	42	0101010
P	87	1010111
Q	95	1011111
R	-	-
S	127	1111111
T	63	0111111
U	7	0000111
V	7	0000111
W	1	0000001
X	63	0111111
Y	126	1111110
Z	95	1011111

Table VII illustrates the pseudocode for the software implementation of the gesture recognition engine.

TABLE VII. SOFTWARE IMPLEMENTATION PSEUDOCODE

<ol style="list-style-type: none"> 1. Initialize No. Of Samples N as required. 2. Declare a character array of size 127 and store all the letters in positions defined by the decimal equivalents to which the letters are mapped. 3. For No. Of Samples N=500, read Analog Sensor values on a scale of 0-1023, average it over N=500 and store it in separate variables for each of the 7 sensors. 4. Calculate gesture recognition time. 5. Call a2d(). <ol style="list-style-type: none"> a. This function will convert the analog values to digital as per set reference value and store the same in separate variables for each sensor. b. Each sensor value is represented by 1 bit. 6. The overall decimal equivalent is obtained using left shift and logical OR operations. 7. IF, the decimal value equals any one of the conflicting values. <ol style="list-style-type: none"> a. The statuses of the wrist and contact sensors are read. b. Letter matching with the status is displayed on the LCD. 8. ELSE, refer character stored in array whose position is defined by the decimal equivalent. <ol style="list-style-type: none"> a. IF, the character in the array position is a valid letter of the English alphabet, push the letter stored in the defined array position to the LCD.

EXPERIMENTAL RESULTS

The intelligence engine theoretically recognizes 25 out of 26 characters of the English alphabet, producing a theoretical efficiency of 96%. The practical efficiency obtained using custom-designed hardware is 80% as the engine recognizes 20 out of 25 theoretically viable characters of the ASL.

Every unique character of the ASL was gestured repeatedly for 100 times, for varying number of samples and their corresponding gesture recognition time were observed, yielding accurate results. The characters that were not fully recognized by the system and their corresponding gesture recognition times were noted as shown in table VIII.

TABLE VIII. GESTURE RECOGNITION TIME - CHARACTERS

Gesture Recognition Time (milliseconds)	Unrecognized Characters
10	D, G, K, Q, T, X, Z
20	D, G, K, Q, T, X, Z
40	D, G, Q, T, X, Z
79	G, Q, T, X, Z
160	G, Q, T, X, Z
399	G, Q, T, X, Z

Gesture recognition time denotes the predetermined program time for identifying a unique gesture corresponding to a particular character of the ASL.

For 25 samples at a gesture recognition time of 20 milliseconds, characters 'D, G, K, Q, T, X, Z' of the ASL

could not be recognized by the engine. Further on increasing the number of samples the efficiency of the engine increased as the character 'K' was recognized for 50 samples at a gesture recognition time of 40 milliseconds. When the number of samples were increased further to 100, Character 'D' of the ASL could also be recognized hence attributing to a greater efficiency of the system and this persisted even on increasing the number of samples further to 500. Thus the practical efficiency of the engine was increased from 72% to 80% of the theoretical efficiency for an increase in gesture recognition time from 20 to 399 milliseconds. However as no change in efficiency was obtained in increasing the gesture recognition time from 79 to 399 milliseconds, the ideal gesture recognition time for this system was benchmarked at 79 milliseconds. This enables the system to identify 12 characters per second which is far higher than the practical speed of gesturing for an average human.

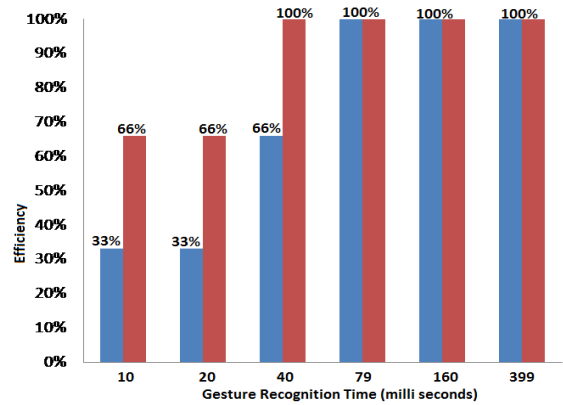


Figure 5. Graph - Efficiency Vs. Recognition Time

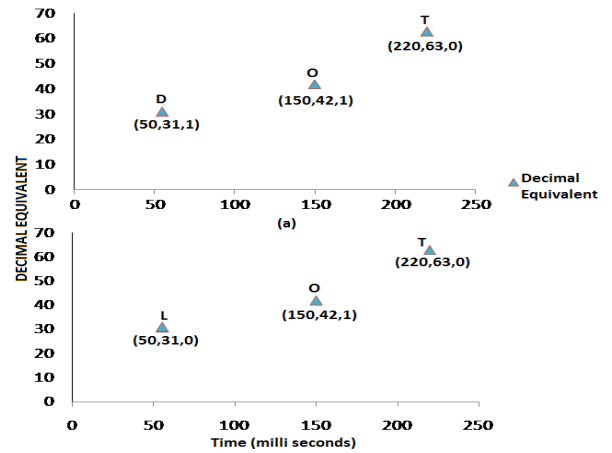


Figure 6. Graph - Recognition Sampling

Also, character identification efficiency of the characters D and K were found to be directly proportional to the number of samples i.e., efficiency increases with increase in the number of samples.

$$\eta \propto n_s \propto t_g$$

Where, η is the character identification efficiency, n_s represents the number of samples and tg represents the corresponding gesture recognition time.

This relation between efficiency and the gesture recognition time is depicted in the graph shown in figure 5.

The glove was tested for the identification of a series of characters to ensure proper functioning of the flex and contact sensors and one such sample tested using the glove is illustrated below.

- D, O, T
- L, O, T

This process was plotted as a graph shown in figure 6 where the x, y, z co-ordinates represent the decimal equivalent of characters, time and the contact sensor value respectively. The letters D and L were uniquely identified based on the contact sensor value thus ensuring high accuracy of character identification.

CONCLUSION

Thus a low-cost, custom-made flex and contact sensor driven gesture recognition glove was prototyped and tested using an inexpensive Arduino ATmega328 microcontroller. The cost of this glove built under laboratory conditions using off-the-shelf components at retail price was realized at USD 5. Further, the practical efficiency of 80% (20 out of 25 letters) was achieved over a theoretical efficiency of 96% (25 out of 26 letters). This indicates the robustness and high efficiency of the implemented prototype. Given that the ideal gesture recognition time to operate the system is recommended at 79 milliseconds, the system does not impose any efficiency constraints on the speed of human gesturing as it is capable of recognizing up to 12 characters per second.

ACKNOWLEDGMENT

The authors would like to acknowledge Solarillion Foundation for its support and funding of the research work carried out for the experimentation and implementation of the gesture recognition glove prototype.

REFERENCES

[1] Fact Sheet No.300, "Deafness and Hearing Loss", World Health Organization (WHO), February 2014, [Online] Available : <http://www.who.int/mediacentre/factsheets/fs300/en/>

[2] Gestural Theory, "Origin of Language", Wikipedia.Org, [Online] Available: http://en.wikipedia.org/wiki/Origin_of_language#Gestural_theory

[3] Hernandez-Rebollar, J.L.; Kyriakopoulos, N.; Lindeman, R.W., "A new instrumented approach for translating American Sign Language into sound and text," Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on , vol., no., pp.547,552, 17-19 May 2004, doi: 10.1109/AFGR.2004.1301590

[4] Dipietro, L.; Sabatini, A.M.; Dario, P., "A Survey of Glove-Based Systems and Their Applications," Systems, Man, and Cybernetics, Part C:

Applications and Reviews, IEEE Transactions on , vol.38, no.4, pp.461,482, July 2008, doi: 10.1109/TSMCC.2008.923862

[5] "American Sign Language", Wikipedia.Org, [Online] Available : http://en.wikipedia.org/wiki/American_Sign_Language

[6] Mehdi, S.A.; Khan, Y.N., "Sign language recognition using sensor gloves," Neural Information Processing, 2002. ICONIP '02. Proceedings of the 9th International Conference on , vol.5, no., pp.2204,2206 vol.5, 18-22 Nov. 2002, doi: 10.1109/ICONIP.2002.1201884

[7] Preetham, C.; Ramakrishnan, G.; Kumar, S.; Tamse, A.; Krishnapura, N., "Hand Talk-Implementation of a Gesture Recognizing Glove," India Educators' Conference (THIEC), 2013 Texas Instruments , vol., no., pp.328,331, 4-6 April 2013, doi: 10.1109/THIEC.2013.65

[8] Bui, T.D.; Nguyen, L.T., "Recognizing Postures in Vietnamese Sign Language With MEMS Accelerometers," Sensors Journal, IEEE , vol.7, no.5, pp.707,712, May 2007, doi: 10.1109/JSEN.2007.894132

[9] Khambaty, Y.; Quintana, R.; Shadaram, M.; Nehal, S.; Virk, M.A.; Ahmed, W.; Ahmedani, G., "Cost effective portable system for sign language gesture recognition," System of Systems Engineering, 2008. SoSE '08. IEEE International Conference on , vol., no., pp.1,6, 2-4 June 2008, doi: 10.1109/SYSE.2008.4724149