Building Real-Time Translation of Arabic Sign Language System using Mechatronic Approach

A.I. El-Alfi, PhD
Dept. of Computer Teacher Preparation
Faculty of Specific Education
Mansoura University, Mansoura, Egypt

R. Adly, PhD
Dept. of Computer Teacher Preparation
Faculty of Specific Education
Mansoura University, Mansoura, Egypt

H.M. Ibrahim
Dept. of Technology Development
Laborers University Mansoura, Egypt

ABSTRACT
Each individual use language to communicate with others. The normal person uses phonetic language but the deaf uses sign language. So, all of them are using different language and having trouble dealing with the other. This paper aims to assist the deaf people to communicate with normal people by building real-time translation Arabic sign language system (ArSL). This system has two parts. The first part translates ArSl to phonetic language by using mechatronic approach. A Data Glove (DG) was built by flex sensors to get the hand gesture of ArSL and convert it to digital data. This data is interrupted to speech via speech knowledge base for Arabic phonetic language. The second part translates Arabic phonetic language to ArSL by using speech recognition for Arabic language to translate speech to text, and then translate text to ArSL via video knowledge base for Arabic language. This system is characterized by mobile, real-time translation, easy in usage and its support for Arabic language.

Keywords
Mechatronic, Data Glove, Flex Sensor, Arabic Sign Language, Glove Print, Hand Gesture, Speech Recognition.

1. INTRODUCTION
A normal people have five senses that enable them to hear and speak with each other. There are some individuals who do not have these important abilities as deaf and dump. So, they depend on communication via sign language to interact with each other. However, communication with the normal is a major concern for them. Each of normal and deaf people are using language not understood by the others. This cause a problem for the deaf and dumb communities to interact with others, particularly when they attempt to involve with educational, social and work environments. So, this paper tends to help deaf in dealing with normal people. Smart agent approach is used to facilitate communication between them.

Mechatronic is the engineering discipline integrating technologies from mechanical engineering, electronics and computing to create more smart devices and machines [1]. Mechanical engineering may be mechanical machines, sensors and actuators [2]. The sensor is a device that detects and responds to some types of input from the physical environment like light, heat, motion, or any one of a great number of other environmental phenomena. The output is generally a signal [3]. Flex sensor is one of the most important sensors. It is used in building DG which is used in such intelligent agents.

Nazrul H. Adnan et. al. used DG in the medical usages. They developed DG to measure the fingers' bending and enhancement the quality of these fingers for the disabled person [4].

Kazuma Iwasako et. al. provided DG as a self-educational tool to learn Japanese sign language. The system helps a learner to recognize motion errors intuitively by himself/herself and provides feedback of errors to the learner [5].

Ahmad Zaki also provided DG to translate Malaysian Sign Language (MSL) to text. He used flex sensors to get the deaf user hand gesture. The result conducted shows that the system successfully detects the alphabet, numbers and words [6].

Norah Alsunaidi et. al. used speech recognition to build a mobile application called Abjad. This application aims to learn children reading which mainly targets children at the age of six years old. It records their voice while reading. Speech recognition converts this voice to text format. Abjad compares text results by the original text then give the children feedback [7].

This paper presents a smart system that depends on obtaining glove print after wearing DG to translate the ArSL into a speech voice. Also, it utilizes speech recognition to translate Arabic speech into ArSL. So that the normal and deaf can communicate with each other.

This paper is organized as following: section 2 presents the proposed system; section 3 presents the proposed system implementation and design; application and results in section 4; and section 5 presents conclusion.

2. THE PROPOSED SYSTEM
The proposed system aims to facilitate communication between deaf and normal users. It is based on translating ArSL to speech and vice versa.

2.1 Translation ArSL to Speech
This part aims to translate Arabic sign language to speech. It includes four main steps:

1-Deaf user wears DG in his hands, and then performs the ArSL (Hand Gesture).

2-DG transfers the hand gesture to digital data (Data Acquisition).

3-A statistical pattern is obtained to each ArSL from digital data. This pattern will define here as ArSL glove print.

4-Glove print is discriminated to interrupt to speech using Weighted Euclidian Distance.

2.1.1 Hand Gesture Acquisition

1- DG is a device consists of set of sensors. The sensors allow acquiring information to recognize hand gesture [8] as shown in figure 1.

Fig 1: The Data Glove

2- Flex sensor is a sensor that can measure the amount of deflection or bending of the fingers. It is an analog resistor [9]. So, DG converts hand gesture to analog data.

3- Microcontroller converts analog data to digital data via Analog-to-Digital converter (ADC) [9]. So, DG returns values of digital data for each sign as shown in figure 2.
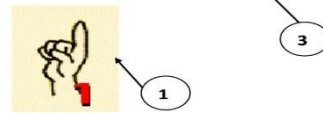

Fig 2: Glove print for number "واحد" by sign language.

2.1.2 Glove Print Acquisition (Learning Phase)

1- The digital data can be reduced by calculating its statistical moments. This yields to obtain the pattern for each ArSL. This pattern will be called here as glove print.

2- Glove print (GP)will consist of 20 features. These features are mean, standard deviation, Kurt and skew as shown in figure 2.

3- Five possible cases for each sign are formulated. Each case will present by GP GP1, GP2, … GPm. (m=5).

4- These five glove prints will be saved in a class. So, each sign has class as shown in table 1. For each class, the appropriate speech record is stored in speech knowledge base.

Table 1. Five glove prints for number "واحد" by sign language

Statistical moments	Class M	Finger1	Finger2	Finger3	Finger4	Finger5
Mean	GP1	0.47	1.12	0.98	1.55	1.12
Std.		0.046	0.009	0.0071	0.052	0.01
Skew		0.048	-0.56	-0.22	-2.13	-0.85
Kurt		-0.48	0.0063	-0.92	5.14	0.60
Mean	GP 2	0.68	1.089	0.90	1.54	1.10
Std.		0.070	0.0152	0.016	0.015	0.0096
Skew		-1.57	-0.564	0.117	-4.0976	-0.0297
Kurt		1.64	-0.390	-1.051	2.361	-1.5405
Mean	GP 3	0.67	1.154	0.964	1.63	1.11
Std.		0.013	0.0106	0.0093	6.76	0.0127
Skew		-0.64	-0.623	-1.467	-1.04	-1.5416
Kurt		-0.73	0.0227	1.017	-2.13	2.2364
Mean	GP 4	0.74	0.995	0.785	1.455	1.1424
Std.		0.014	0.017	0.016	0.010	0.0195
Skew		0.066	-0.62	-1.945	-0.420	-4.783
Kurt		-0.453	-0.86	4.015	0.031	25.251

Mean	GP 5	0.754	0.997	0.791	1.44	1.095
Std.		0.0083	0.025	0.0073	0.033	0.046
Skew		-0.615	-1.450	-0.199	1.623	0.038
Kurt		-0.678	0.950	-1.0851	1.859	-1.247

2.1.3 Glove Print Discrimination (Testing Phase)

1- The user wears the DG, and he/she performs the ArSl.

2- DG sends data to speech knowledge base.

3- Speech knowledge base converts this data to glove print.

4- Through the original and user glove print, speech knowledge base gets Weighted Euclidean Distance (WED) between them throw these following equations:

$$WED = \sqrt{\sum w_i\,(x_i - y_i)^2} \qquad \text{.... (1)}$$

Where

    w: weights = (1 / variance).
    i: data glove sequence
    x: the glove print of original sign.
    y: the glove print of user sign.

5- By using WED, speech knowledge base interrupts Arabic sign language and converts it to its speech record.
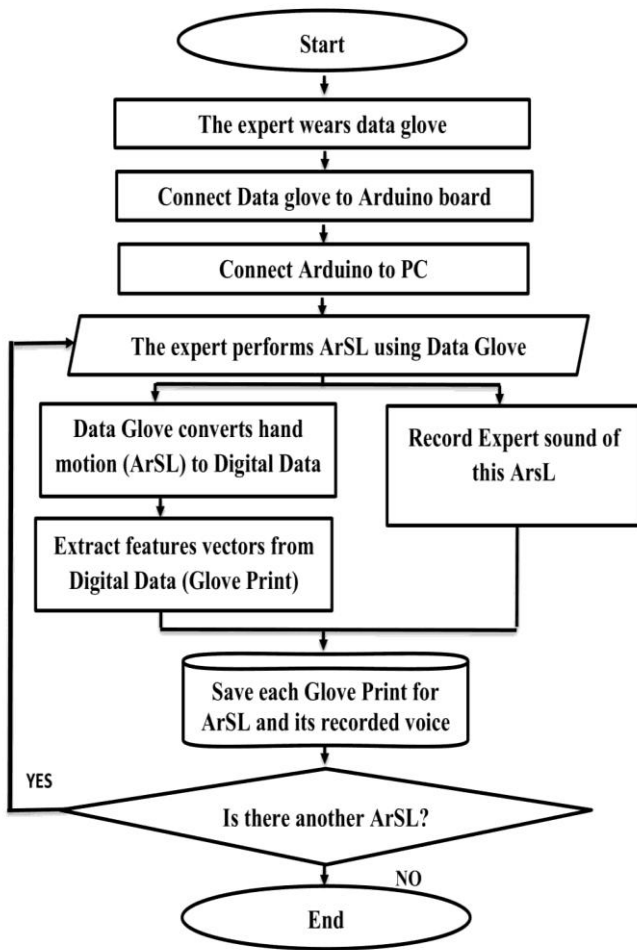

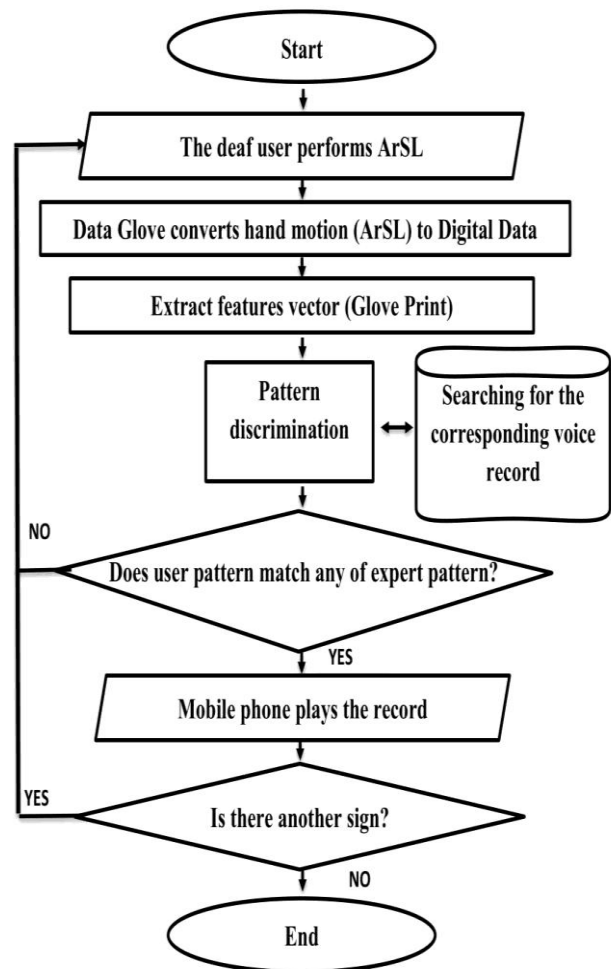**Fig 3: Translation ArSL to speech flow chart Learning phase**

Fig 4: Translation ArSL to speech flow chart Testing phase

2.2 Translation Speech to ArSL Video

The second part of the proposed system aims to translate the speech of normal user to ArSL video. Speech to ArSL video transformation is based on Speech-to-Text then Text-to-ArSL video translating. This process includes four main steps:

1. The normal user speaks on microphone.

2. Speech recognition converts speech to text format [10].

3. The text is processed.

4. The text is translated to ArSL video using video knowledge base.

These processes can be explained as following:

2.2.1 Translation Speech to Text

The proposed system gets the speech from the normal person. It converts these speech words to text format using speech recognition. Speech recognition is included in Application

programming interface (API). It is used to develop the Android OS applications. Speech recognition converts speech to text in real time, supports Arabic language and is available for free [7]. The speech recognition algorithm consists of two steps, first step the audio signal is captured by a microphone and then processed to be recognized, and second step mapping of the signal to words [11]. The proposed system is based on Arabic Language. So, the outcomes of speech recognition maybe one or more probability. The system places these outcomes on array list.

2.2.2 Text Processing

Text processing filters the outcomes placed on the array list to find the correct text. As an example, the normal user said "ذهب احمد الي حديقة الحيوان". Speech recognition converted this speech to text format, saved it in array list consists of M of items, then spilt the text to words as shown in table 2.

Table 2. The most probable text results

M	Word5	Word4	Word3	Word2	Word1
1	الحيوان	حديقة	الي	احمد	ذهب
2	الحيوان	حديقة	الي	احمد	مذهب
3	الحيوان	حديقة	الي	احمد	ذهب

Arabic language has several rules for text processing. The system based on frequent rule to get the correct word. This rule gets the more frequent word in the array. In the previous example, there are two probable results:

- ذهب is frequented two times.
- مذهب is frequented one time.

According to this rule the correct word is ذهب. So, the sentence is "ذهب احمد الي حديقة الحيوان".

2.2.3 Video Processing

Sign language depends on the hands' movement and cannot be represented as single image. So, the system used video to display ArSL. This required two stages. The 1st stage is to build video knowledge base for the ArSL videos (Learning phase) the 2nd stage is to utilize the online conversation and transform it into video.

As mentioned, the system processed the text to get the correct sentence. Thereafter the system splits this sentence into words then transfers each word to its corresponding ArSL video. Knowing that there are two types of words in video knowledge base single words as "ذهب" and compound words as "حديقة الحيوان":

a- The single words are converted to ArSL video knowledge base:

IF (word one == "ذهب") THEN display video "ذهب"

b- The compound words are expression consists of multiple words but have single ArSL video in video knowledge base.

IF (word one == "حديقة")

    IF (word two == "الحيوان")
        THEN   display video "حديقة الحيوان"
    ELSE IF (word two == "الاسماك")
        THEN   display video "حديقة الاسماك"
    ELSE IF (word two == "المنزل")

        THEN   display video "حديقة المنزل"
    END IF
END IF

The system converts the textual words to Arsl videos and displays them in the mobile screen to the deaf user.
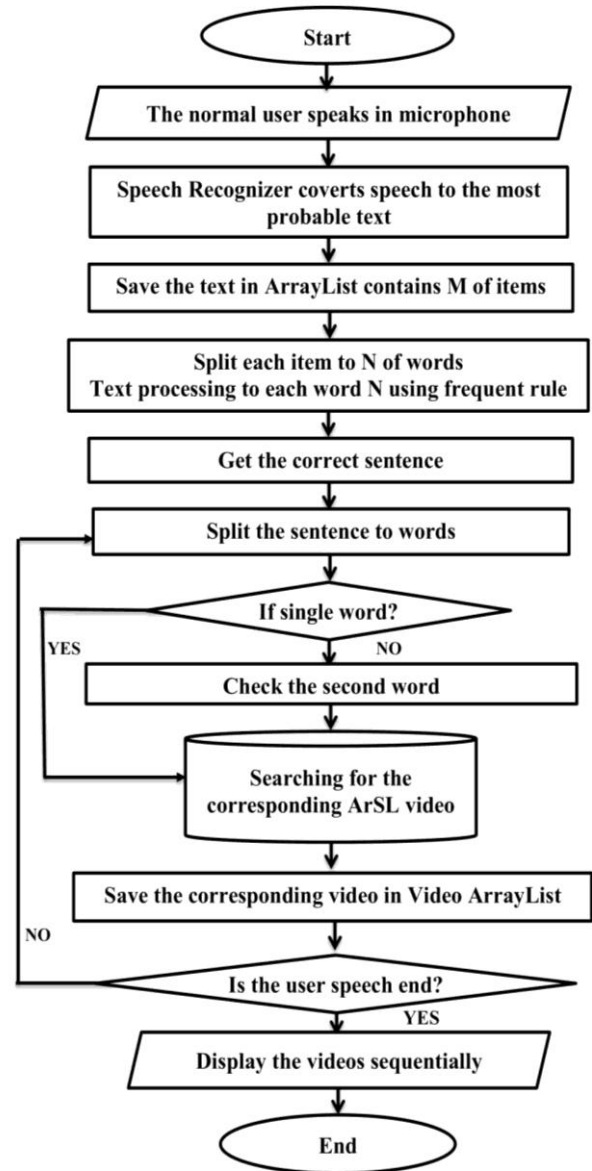

Fig 5: Translation speech to ArSL flow chart

3. THE PROPOSED SYSTEM IMPLEMENTATION AND DESIGN

The proposed system depends on mobile application and DG. The mobile application is designed and developed by Android studio.

DG proposed consists of Arduino board (UNO), Flex Sensor 2.2" (SEN-10264), 10K Ohm resistors, Jumper Wires, Bluetooth (HC05), Battery, and Breadboard. To build DG follow the following steps:

3.1 Build the Voltage Divider Flex Sensor Circuit

As mentioned earlier the flex sensor is analog. So, voltage divider of flex sensor circuit was built and connect it with resistor [12] as shown in figure 6.
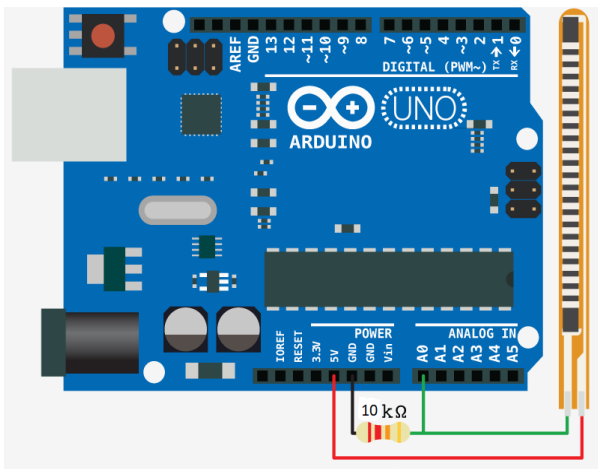

Fig 6: Voltage divider flex sensor circuit

3.2 Build Breadboard Circuit

Breadboard is a thin plastic board used to hold electronic components that are wired together [13]. The proposed system contains breadboard with fixed jumper wires on its back as shown in figure 7.
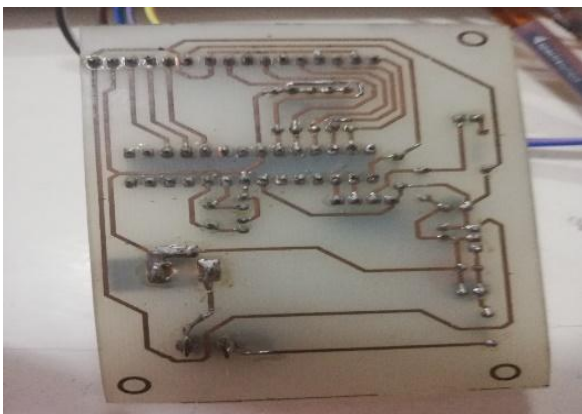

Fig 7: The fixed jumper wire in breadboard

The resistance and Bluetooth can be fixed on it to become save. But Microcontroller have to remove from Arduino and fixed it in breadboard to become possible to separate the Arduino board from DG as shown in figure 8.
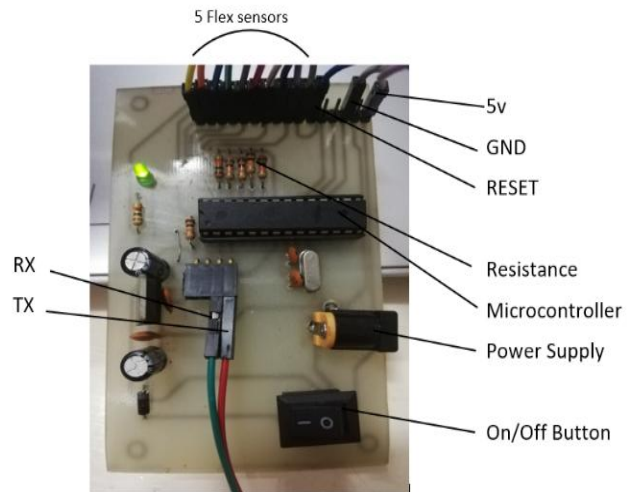

Fig 8: Breadboard component

At first, breadboard has to be connected with Arduino to can programming the sensors. Just connect 5v by 5v, GND by GND, RESET by RESET, finally RX by RX and TX by TX using jumper wires as shown in figure 6 and 8.

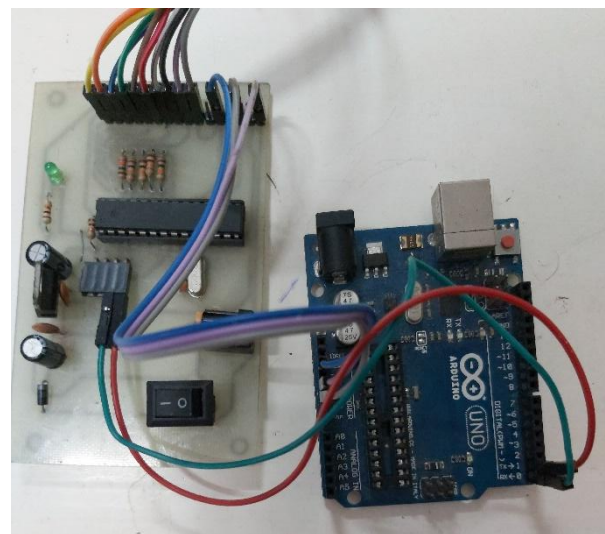Finally, the breadboard and Arduino become as shown in figure 9.

Fig 9: Connect Arduino and breadboard

Now DG become ready to run on laptop and programming it. After programming, Arduino board has no longer function and can be separated from DG and connect Bluetooth and battery. The DG finely become as shown in figure 10.

Fig 10: The data glove

4. APPLICATION AND RESULTS

The proposed system is applied on a sample of teachers and students in deaf and dumb School. The application aims to test the system's ability to convert the ArSL to speech and vice versa. So, there are two applications and two results.

4.1 The Results of Translation ArSL to Speech

The initial experiment done was to test the ability of the system to translate ArSL to speech. This experiment applied on ten deaf students. Each deaf student wears the DG, then performed each alphabet/number/word sign saved in speech knowledge base. Speech knowledge base contains ArSL for numbers from one to ten, alphabet, and simple words in Arabic language.

The system translated each ArSL for student to speech and also written words which appeared on the mobile screen. The results of these attempts are shown in table 3.

Table 3. The results of translating from ArSL to speech

Sign	Num. of attempts	True	false	The success percentage
Numbers	100	96	4	96%
Alphabet	280	259	21	92.5%
Words	50	44	6	88%

4.2 The Results of Translation Speech to ArSL

The second part of the experience done was to test the ability of the system to translate speech to ArSL. This part was tested on ten teachers to translate their speech to ArSL video via video knowledge base. Video knowledge base contains ArSL video for numbers from one to ten, alphabet, and simple words in Arabic language. The system translated the teacher speech to ArSL video displayed on mobile screen. The results of these attempts are shown in table 4.

Table 4. The results of translating from ArSL to speech

Sign	Num. of attempts	True	false	The success percentage
Numbers	100	98	2	98%
Alphabet	280	267	13	95.36%
Words	50	45	5	90%

Based on the results shown in Table 3 and 4, the accuracy for all the tests is quite high. In the number of 10 attempts tested for each sign, at least 8 attempts were successful. The alphabet and numbers have higher accuracy due to the usage of flex sensors. The words have lower accuracy because the words involve motion which needs to be detected by more sensors. Overall, the accuracies range from 88% to 98%, which is reasonably high. So, the results of the tests show that the proposed system successfully transfer ArSL to speech and speech to ArSL.

5. CONCLUSION

The proposed system aims to translate real time ArSL to speech and vice versa. Translation of ArSL to speech based on DG. DG consists of flex sensors. Flex sensor measures the bending of the finger and converts this bending to digital data. So, DG returns for each ArSL five digital data. The system calculates the statistical moments to obtain sign language glove print. By using WED the system compares this glove print with the original glove prints saved in speech knowledge base. Speech knowledge base returns the record which its glove print matches with the user glove print. Accordingly, the system translated ArSL into speech.

On the other side, translation of speech to ArSL is based on speech recognition. Speech recognition converts speech to text. The system processes the text then converts it to ArSL video using video knowledge base. Accordingly, the system translated speech into ArSL.

This system translates numbers from one to ten, alphabet, and some words in Arabic language. It is applied on a sample of teachers and students in deaf and dumb School. It has achieved 98% success.

6. REFERENCES

[1] Stefan Kozsk 2016 "Multidisciplinary Approach and Dual Education in Control Engineering for Mechatronic", faculty of electronic Engineering and Information Technology, Slovak University of Technology in Bratislava, Slovakia, IFAC-Papers OnLine 49-6 (2016) 052–056

[2] Dr. Kevin Craig 2002 "ACTUATORS & SENSORS IN MECHATRONICS", Associate Professor of Mechanical Engineering.

[3] M. Anjanappa, et al 2002 "Sensors and Actuators", University of Maryland Baltimore County.

[4] Nazrul H. Adnan et al 2012 "Measurement of the Flexible Bending Force of the Index and Middle Fingers for Virtual Interaction", International Symposium on Robotics and Intelligent Sensors, 388 – 394

[5] Kazuma Iwasako et al 2014 "Development of finger motion skill learning support system based on data gloves", International Conference on Knowledge-Based and Intelligent Information & Engineering Systems.

[6] Ahmad Zaki Shukor and et al 2015 "A New Data Glove Approach for Malaysian Sign Language Detection",

IEEE International Symposium on Robotics and Intelligent Sensors (IRIS 2015)

[7] Norah Alsunaidi et al 2018 "Abjad: Towards Interactive Learning Approach to Arabic Reading Based on Speech Recognition"

[8] Julien Maitre et al 2019 "Basic Daily Activity Recognition with a Data Glove", The 10th International Conference on Ambient Systems, Networks and Technologies (ANT) April 29 - May 2, 2019, Leuven, Belgium

[9] Kunal R. Jambhulkar 2017 "Review on Sensor based Hand Gesture Recognition System", Department of Embedded System and Computing, Nagpur Institute of Technology, India.

[10] K. saksamudre et al 2015 " A Review on Different Approaches for Speech Recognition System."

+International Journal of Computer Applications 115(22): 23-28.

[11] Prerana Das et al 2015 "VOICE RECOGNITION SYSTEM: SPEECH-TO-TEXT", Department of Computer Science & Engineering and Information Technology, School of Technology, Assam Don Bosco University, Assam, India.

[12] Abidhusain Syed et al 2012 "Flex Sensor Based Robotic Arm Controller Using Micro Controller", Department of Electronics and Communication, BLDEA College, India.

[13] Mark Keanu et al 2019 "Materials Selection of 3D Printing Filament and Utilization of Recycled Polyethylene Terephthalate (PET) in a Redesigned Breadboard", Department of Mathematics, Ateneo de Manila University Philippines.