

# Hand gesture recognition and voice conversion system for dumb people

V.Padmanabhan, M.Sornalatha

**Abstract**— in our country around 2.78% of peoples are not able to speak (dumb). Their communications with others are only using the motion of their hands and expressions. We proposed a new technique called artificial speaking mouth for dumb people. It will be very helpful to them for conveying their thoughts to others. Some peoples are easily able to get the information from their motions. The remaining is not able to understand their way of conveying the message. In order to overcome the complexity the artificial mouth is introduced for the dumb peoples. This system is based on the motion sensor. According to dumb people, for every motion they have a meaning. That message is kept in a database. Likewise all templates are kept in the database. In the real time the template database is fed into a microcontroller and the motion sensor is fixed in their hand. For every action the motion sensors get accelerated and give the signal to the microcontroller. The microcontroller matches the motion with the database and produces the speech signal. The output of the system is using the speaker. By properly updating the database the dumb will speak like a normal person using the artificial mouth. The system also includes a text to speech conversion (TTS) block that interprets the matched gestures.

**Index Terms**— Gesture, Flex sensor, accelerometer sesor, measuring instrument, microcontroller, TTS, Speaker.

## 1 INTRODUCTION

About nine billion people at intervals the planet unit of measurement dumb. The communication between a dumb and hearing person poses to be an important disadvantage compared to communication between blind and ancient visual people. This creates an extremely little house for them with communication being associate degree elementary aspect of human life [6]. The blind people can speak freely by implies that of ancient language whereas the dumb have their own manual-visual language referred to as language. Language is also a non-verbal form of intercourse that's found among deaf communities at intervals the planet. The languages haven't got a typical origin and thence hard to interpret. A Dumb communication interpreter is also a tool that interprets the hand gestures to sensibility speech. A gesture in associate degree extremely language is also a certain movement of the hands with a particular kind created out of them. Facial expressions collectively count toward the gesture, at constant time. A posture on the other hand is also a static variety of the hand to purpose an emblem. Gesture recognition is classed into a pair of main categories: vision based mostly} and detector based [7] [4]. The disadvantage of vision based totally techniques includes advanced algorithms for process. Another challenge in image and video method includes varied lighting conditions, backgrounds and field of scan constraints and occlusion. The detector based totally technique provides larger quality. The primary aim of this paper is to introduce an issue that will efficiently translate language [3] gestures to every text and sensibility voice. The interpreter makes use of a glove based totally technique comprising of flex detector [9], instrument sensors [8]. For each hand gesture created, a symptom is formed by the sensors appreciate the hand sign [10] [5] the controller matches the gesture with pre-stored inputs. The device not exclusively interprets alphabets but cans even sort words exploitation created gestures. A training mode is gettable on the device therefore it fits every user and accuracy is inflated. The device will even be able to translate larger gestures that require single hand movement. Gesture recognition implies a method by that knowledge is collected from parts of the physical body (usually the hand) and processed to work

out attributes like hand form, direction and speed of gesture being performed. There are presently 2 sorts of answer. Device based mostly techniques involve some variety of guide like a glove or glove-like framework fitted with position trackers and flex sensors to live the condition and position of the hand. Visual based mostly techniques use camera chase technologies, whereby usually the user wears a glove with specific colors or markers indicating individual parts of the hands, specially the fingers. The cameras record the ever-changing image and position of the hand because the user signs and also the pictures are then processed to retrieve the hand form, position and orientation.

## 2 MATERIALS AND METHODS

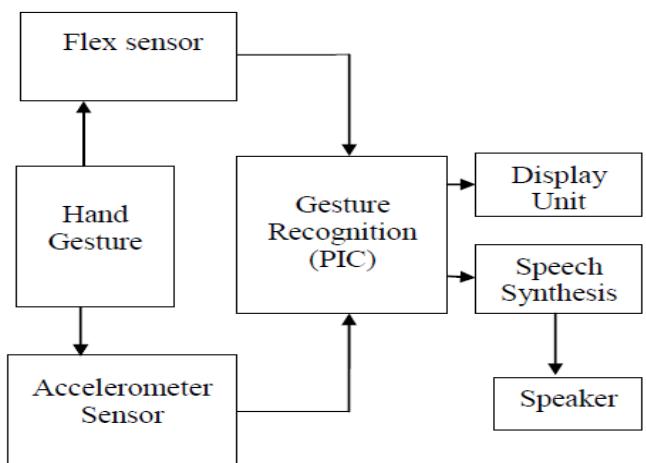


Fig.1. Block diagram of system

The controller employed in the device could be a PIC microcontroller. 5 flex sensors area unit accustomed measure the degree of bending of the fingers. The flex sensors interface with the controller vic-

timization the resistance circuit. The measuring instrument is directly interfaced to the digital ports because it includes the signal acquisition circuit [11]. The device contains these sensors for coaching mode and word formations. This can be interfaced with the digital ports of controller to enclose the digital information. PIC microcontroller processes the information for every explicit gesture created. The controller has 2 modes of operation – coaching mode and operational mode. In coaching mode the gesture is formed by the user and also the voltage levels area units keep in EEPROM. In operational mode the information is being compared with predefined values and also the matched gestures area unit sent with text to speech conversion module. The module consists of TTS block and Speak Jet [1]. The top product is processed and detected via a speaker.

### 3 DATA GLOVE

Data glove consists of 2 detectors; flex sensors and measuring device sensor. The output of the measuring device sensors is detected by the lean detection module, whereas the output of the flex sensors and therefore the overall gesture of the hand square measure detected by the gesture detection module [12]. The gesture detection module provides associate degree 8-bit address for speech synthesis module; 8-bit address is completely different for every gesture. Speech Synthesis module speaks the message severally to deal with received by it.

### 4 FLEX SENSORS

Flex sensors are resistive carbon parts. When bent, the device develops a resistance output correlative to the bend radius. The variation in resistance is [15] just about  $10\text{k}\Omega$  to  $30\text{k}\Omega$ . A global organization flexed device has  $10\text{k}\Omega$  resistance and once bent the resistance will increase to  $30\text{k}\Omega$  at  $90^\circ$ . The device incorporates within the device employing a potential divider network. The potential divider is employed to line the output voltage across 2 resistors connected non-parallel as shown in Figure 2. The electrical device and flex forms a potential divider that divides the input voltage by a quantitative relation determined by the variable and glued resistors.

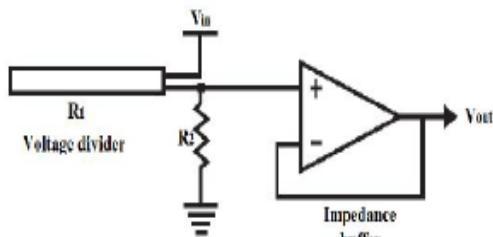


Fig.2. Equivalent circuit of flex sensor

### 5 ACCELEROMETER SENSORS

Accelerometer within the Gesture Vocalized system is employed as a tilt sensing element, which checks the tilting of the hand. ADXL103 measuring system as shown in Figure three. The tip product of the measuring system is provided to 3rd module, which incorporates pipeline structure of 2 ADC. There's a technical issue at this stage of the project, that the analog output of the measuring system as shown

in Figure 3, that ranges from one.5 volts to three.5 volts to a digital 8-bit output systems, becomes terribly sensitive.

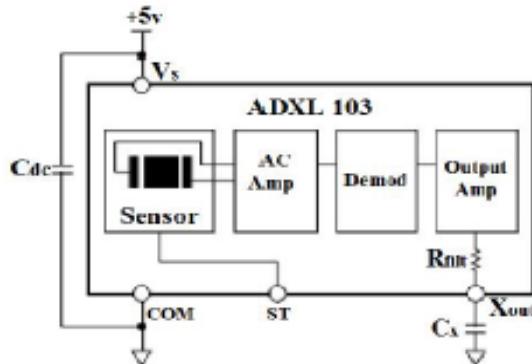


Fig.3. Block diagram of accelerometer sensor

If a lot of sensitive [14] system is employed then there's a large modification within the digital output with the little tilt of the script that is troublesome to be done by.

### 6 SPEECH SYNTHESES

On the idea of this comparison the microcontroller involves acknowledge that that gesture will the hand build. Quickly the microcontroller is aware of that that knowledge is distributed by the bend detection module, and what [13] that means of this knowledge is. That means implies that the microcontroller is aware of, if the hand is creating some outlined gesture and what ought to the system specification. The last step of the system is to convey voice to the every outlined gesture. For this purpose a speech synthesizer is employed. Every word is consisted of some explicit phonemes and just in case of Speech synthesizer every allophone has some explicit addresses as shown in Figure 4. This address is to be sent to the Speech synthesizer at its address lines, to form the speaker, speak that individual word.

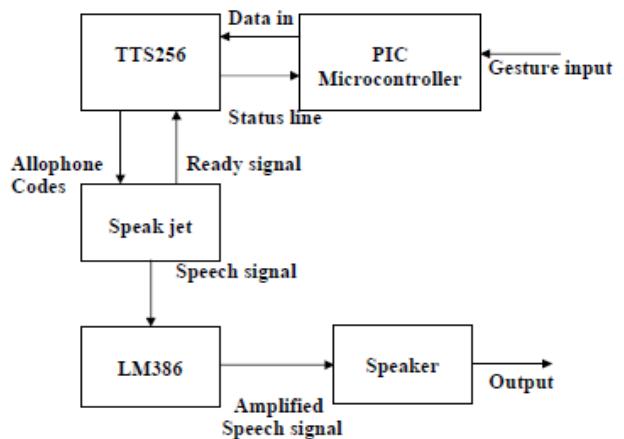


Fig.4. Block diagram of speech synthesizer

Currently these addresses area unit already hold on within the microcontroller [16]. So far, the microcontroller is aware of what the gesture created by the hand, and what ought to be expressed against it. The microcontroller sends the eight-bit address to Speech synthesiz-

er. This eight-bit address is representing the allophones of the word to be spoken. Speech synthesizer offers a sign output. This signal is amplified by victimization the amplifying electronic equipment [2]. The output of the electronic equipment is given to the speech synthesizers.

Fig.5.Flow chart of system execution

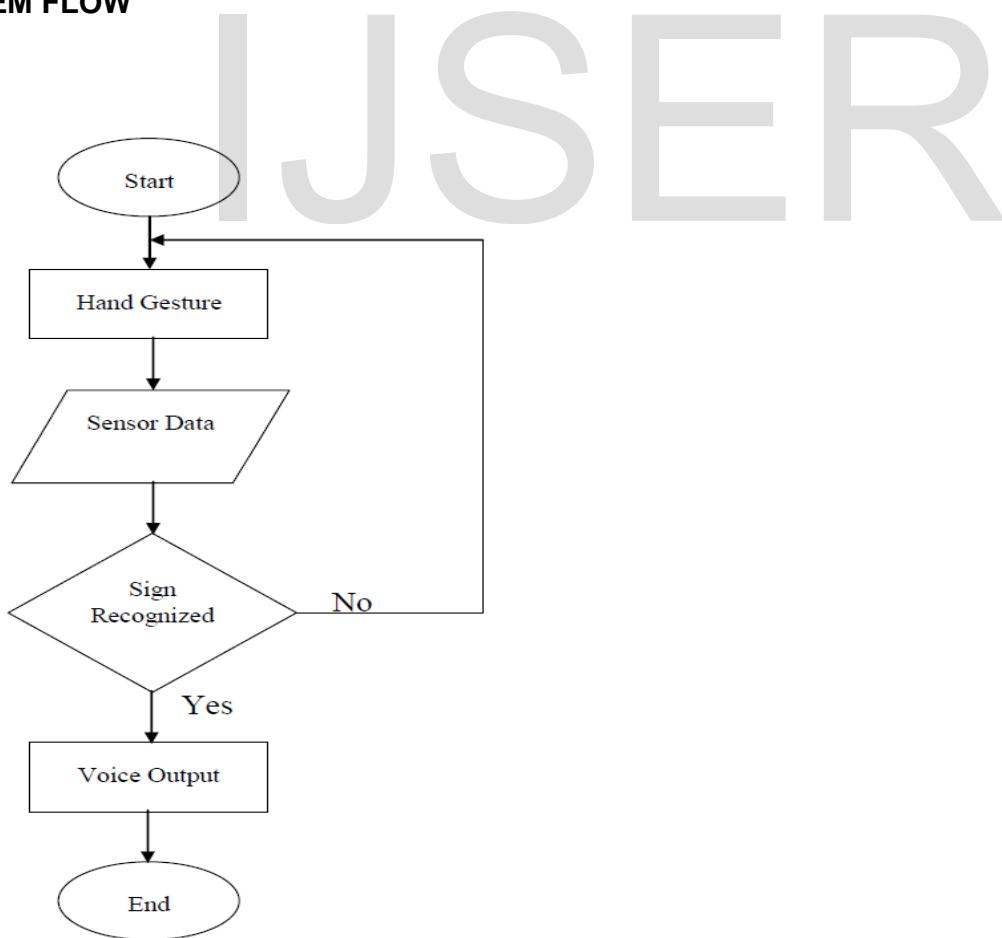
## 9 RESULTS AND ANALYSIS

## 7 SYSTEM FUNCTIONALITY

Table 1

When power is ON, the position and orientation of hand is obtained from the data glove that consists of five Flex sensors on fingers (Thumb, index, middle, ring, and pinky) and one accelerometer of PIC microcontroller LM386 Speaker Speakjet TTS256 Three outputs (X, Y, and Z positions). Tilting of the palm can be captured by the accelerometer where Flex sensors can measure the bend of the five fingers when making a sign. When the user performs a gesture/letter and press a button, signals coming from the sensors are amplified via a dedicated amplification circuit to each signal, and then captured by the microcontroller which convert the analog signals to digital values through its 8-channel ADC. These values are formatted into a simple state matrix: five values for the Flex sensors, one for each axis of the accelerometer. As a result, each letter in the ASL will have a specific digital level for the five fingers and the three axis of the accelerometer. Each level is represented by a value between 0 and 255; an interval of  $\pm 3$  levels should be taken into consideration in case the user could not keep his hand steady.

## 8 SYSTEM FLOW



Letter	Thumb	Index	Middle	Ring	Pinky	X,Y,Z positions
A	12	15	18	12	12	18
B	23	23	28	23	23	23
C	35	38	35	33	35	35
D	46	48	41	41	46	43
E	51	53	53	56	56	53
F	64	64	61	69	69	64
G	79	74	76	79	74	76
H	84	81	87	87	81	84
I	94	92	94	97	92	94
J	104	102	104	107	102	104
K	112	112	115	115	117	117
L	122	125	127	122	127	125
M	133	135	133	138	135	133
N	143	148	143	148	143	145
O	151	153	157	155	158	151
P	163	168	165	168	163	163
Q	174	171	176	177	177	174
R	184	183	189	186	189	182
S	191	194	193	195	197	199
T	201	206	209	209	202	204
U	215	217	219	212	215	217
V	225	227	223	221	223	225
W	237	232	237	235	235	237
X	241	243	241	243	244	243
Y	247	246	248	247	248	248
Z	253	253	252	252	254	254

people. This project aims to lower the communication gap between the mute community and additionally the standard world. The projected methodology interprets language into speech. The system overcomes the necessary time difficulties of dumb people and improves their manner. Compared with existing system the projected arrangement is compact and is possible to carry to any places. This system converts the language in associate passing voice that's well explicable by blind and ancient people. The language interprets into some text kind displayed on the digital display screen, to facilitate the deaf people likewise. In world applications, this system is helpful for deaf and dumb of us those cannot communicate with ancient person. The foremost characteristic of this project is that the gesture recognizer may be a standalone system, that's applied in common-place of living. It's in addition useful for speech impaired and paralysed patient means those do not speak properly and in addition used for Intelligent Home Applications and industrial applications.

## REFERENCES

1. Amarjot, S., Devinder, K., Phani, S., Srikrishna, K., and Niraj, A. "An intelligent multi-gesture spotting robot to assist persons with disabilities", *International journal of computer theory and engineering*, Vol. 4, No. 6, pp. 1290-1420,2012.
2. Alois, F., Stefan, R., Clemens, H., and Martin, R. "Orientation sensing for gesture-based interaction with smart artefacts", *IEEE Transactions on audio, speech, and language processing*, Vol.28, No. 8, pp. 1434-1520, 2007.
3. Donglin, W., Ajesh, P., Hye, K., and Hosub, P. "A Deconvolutive neural network for speech classification with applications to home service robot," *IEEE Transactions on instrumentation and measurement*, Vol. 59, No. 12, pp. 2334-4620,2010.
4. Istvan, I., Zoran, M., and Nikola, D. "Hands-free voice communication with TV", *IEEE Transactions on consumer electronics*, Vol. 17, No. 2, pp. 1378-1695, 2011.
5. Ibrahim, P., and Srinivasa, R. "Automated speech recognition approach to continuous cue symbols generation", *International journal of power control signal and computation*, Vol. 18, No. 8,pp. 434-520,2007.
6. Jean, C., and Peter, B., "Recognition of Arm Gestures Using Multiple Orientation Sensors: Gesture Classification", *IEEE Intelligent transportation systems conference on electronics*, Vol. 13, No. 1, pp. 334-520,2004.
7. Joyeeta, S., and Karen, D. "Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique", *International journal of advanced computer science and applications*, Vol. 4, No. 2, pp. 434-820, 2013.
8. Ravikiran, J., Kavi, M., Suhas, M., Sudheender, and S., Nitin, V. "Finger Detection for Sign Language Recognition", *Proceedings of the International multi conference of engineers and computer scientists*, Vol. 2, No. 1, pp. 234-520,2009.
9. Kekre, M., and Vaishali, K. "Speaker identification by using vector quantization," *IEEE Transactions on mechatronics*, Vol. 15, No. 1, pp. 1034-1620, 2010.
10. Kouji, M., Kazuya, M., Tsutomu, H., and Ryo, K. "A decision method for placement of tactile elements on a sensor glove for the recognition

## 10 CONCLUSIONS

Sign language may be a helpful gizmo to ease the communication between the deaf or mute community and additionally the standard

of grasp types”, IEEE *Transactions on image processing*, Vol. 28, No. 8,  
pp. 654-1834, 2004.

11. Mark, H., Rebecca, P., Peter, O. “A voice-input voice-output communication aid for people with severe speech impairment”, IEEE *Transactions on neural systems and rehabilitation engineering*, Vol. 21, No. 1, pp. 1534-4320, 2013.
12. Oyaaran, B. “Vision based sign language recognition modelling and recognizing signs with manual and non-manual components”, *International journal of computer science and engineering*, Vol. 12, No. Paper-back, 03 pp. 378-1695, 2008.
13. Otiniano, R., and Amara, C. “Finger spelling recognition from rgb-d information using kernel descriptor”, IEEE *Transactions on neural systems and rehabilitation engineering*, Vol.28, No. 8, pp. 124-184, 2006.
14. Ravikiran, J., Kavi, M., Suhas, M., Dheeraj, R., Sudheender, S., and Nitin V. “Finger detection for sign language recognition”, *Proceedings of the international multi conference of engineers and computer scientists*, Vol. 39, No. 12, pp. 334-620, 2009.
15. Zhengmao, Z., Prashan, R., Monaragala, N., Malin, P. “Dynamic hand gesture recognition system using moment invariants”, IEEE *Transactions on neural networking and computing*, Vol. 21, No. 1, pp. 1034-1320, 2010.

