International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

ZIGBEE BASED WIRELESS VOICE TO TEXT TRANSLATOR

S.Monisha[1], T.Monisha[1], R.Mounika[1] Associate Professor; Dr.S.Selvarasu M.E, Ph.D[2]

[1,2]Muthayammal Engineering College,Rasipuram,India.

ABSTRACT

Speech is an important primary need and the most convenient means of communication between people. The communication among - human-computer interaction is called the human-computer interface and it is a very useful means of communication. This project gives an overview of the major technological perspectives of the fundamental progress of speech-to-text conversion and also gives a complete set of speech-to-text conversion based on Zigbee wireless communication. A study of different techniques is done as per stages. This project concludes with the decision on the future direction for developing techniques in human- computer interface systems in the mother tongue and it also discusses the various techniques used in each step of a speech recognition process and attempts an efficient approach for designing an efficient system for speech recognition. However, with modern processes, and methods we can process voice signals easily and recognize the text. In this system, we are going to develop a speech-to-text by using Zigbee. However, the transfer of voice to written text requires special techniques as it must be very fast and accurate. The objective of this paper we use a speech recognition system that recognizes the spoken command by the user and compares it with the already existing database and also Zigbee wireless communication between two systems.

***Keywords**: Speech Recognition, Communication, Algorithm,Zigbee,Based,Wireless,Voice,To,Text,Translator*

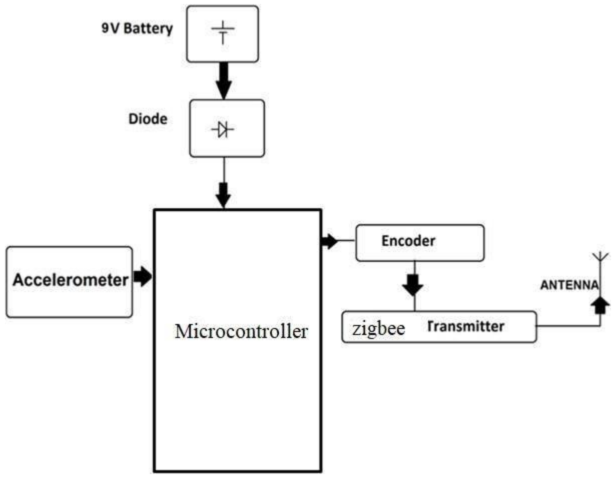
INTRODUCTION

Humans interact with each other in several ways such as facial expression, eye contact, gesture, mainly speech. It is the primary mode of communication among human beings and also a very useful and efficient form of exchanging information among humans in speech. Speech-to-text conversion is widely used in many application areas. Text- to-speech convention transforms information stored as data text into speech. It is widely used in voice reading devices for blind people now a day's. In the last few years, however;the use of text-to-speech conversion technology has grown far beyond the disabled community to become major digital voice storage for voice mail and voice response systems. Developments in Speech technology for various languages have taken place. In this project, we use a voice-recognitionsystem that recognizes the voice command by the user and compares it with the already existing database and also Zigbee wireless communication between two systems. It is the primary and most useful means of communication with people. Whether due to technological curiosity to build machines that mimic humans or desire to automate work with a machine, research in speech recognition is the first step towards human-machine communication. Speech recognition is the process of recognizing the spoken word into a return signal it is very useful for humans.

ZigBee is developed as an open global standard to address the unique needs of low-cost, low-power, wireless sensor networks. Zigbee is the set built around the IEEE802.15.4 wireless protocol. As ZigBee is the technology, we had tried to demonstrate its way of functionality and aspects like kinds, advantages, and disadvantages using a small application of controlling any kind of electronic devices and machines. The Zigbee technology is broadly adopted for fastdata transmission over a dedicated channel.

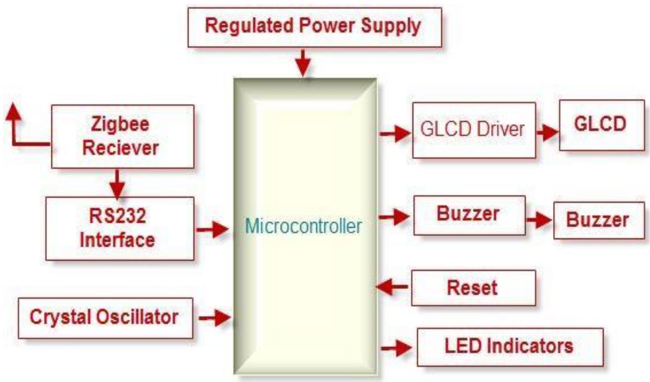
1. EXISTING SYSTEM

In existing system all the sensors data will be stored send send to the doctor using Rf. A Wireless Sensor Network (WSN) for monitoring patient's physiological conditions continuously using Rf. Here the physiological conditions of the patient's are monitored by sensors and the output of these sensors is transmitted via Rf and the same has to be sent to the remote wireless monitor for acquiring the observed patient's physiological signal Infusion pump is a medical device. It is healthcare facilities used in hospitals, and at home. It can deliver fluids both in medicines and nutrients such as pain relievers chemotherapy drugs, hormones or insulin, and antibiotics into a patient's body in any amounts. There are many pumps including insulin pumps, syringe, large volume, elastomeric, patient- controlled analgesia (PCA), and enteral pump widely used. Enteral pump is a pump that is used in medications and deliverliquid nutrients to a patient's digestive tract. Patient-controlled analgesia (PCA) pump is a pump that is used to deliver pain medication. Insulin pump .It is a pump that is used to deliver insulin to patients with diabetes which is frequently used in resident places. These devices are very important in hospitals and nurses used because they can show status of liquid that they give to patients. So, the devices are very popular in hospitals for checking status of patient and medicine..



2. PROPOSED SYSTEM

To propose a system that mainly consists of a transmitter and a receiver section. In the transmitter section (at the patient side), a four-axis accelerometer will be placed on any movable part of the patient. This accelerometer is capable of measuring the acceleration which is static due to gravity and thus finding the angle at which the device is tilted with respect to the earth. Whenever the patient in the hospital needs any help he tilts the accelerometer in different directions. This acts as an input to the accelerometer while the output of it is in the forms of volts that are connected to the controller board which acts as the processing unit. The output of the accelerometer depends onthe angles and is read by the controller. The controller mapsthe input voltages between 0 to 5 volts into integer values between 0 to1023 as analog data from the range of 0-1023 itis very useful. This range provides a lot of sensitivity and a slight shift can lead to a change in value. To reduce the complexity and provide a simple way for the patients, we reduced its sensitivity by mapping it to 0-5 volts and then provided a range for the front, back, forward, and backward. These directions are very easy and used by any person using his/her thumb or any part of the body capable of moving in these directions. A predefined message catering to the basic needs of the patients and those required for an emergency will be stored in the ranges assigned to a particular direction as mentioned above. The accelerometer will be connected to the patient for communication. Then the patient will have a controller board and transmitter for sending his messages. For identification purposes, their name or number is sent to the nurse. All these transmitters can be connected centrally to one ZigBee receiver which works on the same frequency as the transmitter. Thus the proposed system will provide many-to-one communication. On the receiver side, the ZigBee receiver will receive the message and send it to the controller board on the receiver side which will then display the message on the LCD. On reception of the message, the nurse will remotely take the required action to cater to the needs of the message. In case of emergency, the patient has to just press a push button which will signal the processing board to send an emergency alarm to the receiver then the receiver help the patient. The receiver will receive the signal then the controller to activate the buzzer.



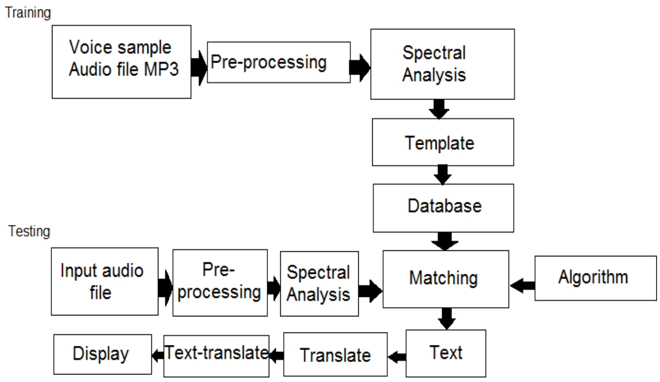
3. LITERATURE SURVEY

The conversion of speech signal to words in a orderly manner by means of algorithm applied as a machine program is termed speech recognition. The objective of speech recognition area is to make changes and develop a system for speech input to a machine based on progress of statistic modeling of speech.[1] the most extensive and prominent way to extract spectral features is determining Mel-frequency cepstral coefficients (MFCC).MFCC used in speech recognition depend on frequency domain using Mel scale that intern depends on human ear scale. MFCC exhibits real cepstral of a windowed short time signal which is evolved from fast Fourier transform (FFT). Audio feature extraction technique based on MFCC extracts characteristics from identical speech to once that are preowned by human for hearing speech whereas parallelly all other information's are neglected.[2]

Herman sky developed a model called perceptual linear prediction (PLP).In PLP the concept of psychophysics of hearing is used that models human speech. PLP enhances speech recognition rate by removing unjustified information of the speech to equate characteristics of human auditory system changes have been brought in spectral characteristics which is the only difference between PLP and LPC technique.[2][3] PNCC, a front end technique which is slightly different than MFCC, which uses gamma tone filters instead of Mel- scale transformation [5] Imitating the performance of cochlea. To increase robustness, the term called medium time power bias removal is included further. The evaluate the quality reduction of speech due to noise the arithmetic to geometric mean ratio is calculated by bias vector.[4] Speech to text engine transforms speech to text from an instant voice input, complementing users giving better ideas for a different choice of data entry. HMM are used to perform speech to text conversion. HMM generates stochastic models from a known statement and in contrast with the possibility that the unknown statement was generated by each model. can be observed as a piece wise stationary signal or a short time stationary signal, hence HMM are thereby used in speech to text conversion.[6] A parametric density function that indicates weighted sum of Gaussian components densities is referred as Gaussian Mixture Model (GMM).To compare the feature extracted from the model with stored template, Gaussian Mixture Model is used. Representation of Gaussian Mixture model is done with the help of Gaussian distribution which is thereby calculated by its mean, variance and weight of the same.[7]

Machine translation or MT comes under computational linguistic that examines use of software to translate text from source language to target language .A translation from an intermediate work then it is useful to demonstrate therepresentation that imitates the meaning of actual sentence is used in transfer based machine translation which is similar to interlingual machine translation. In this knowledge of the source and target languages are used to evaluate its grammatical and actual structure of language, transferring that to a structure appropriate for developing text in a target language, and thereby obtaining the desired text.[8]Dictionary uses dictionary entries, similar to that of a normal dictionary-word by word, generally with not much correlation of meaning amongst the language. Morphological analysis of lemmatization may or may not be used in lookup. Whereas this of translation is least polished, but to translate long lists of phrases on a significant level, dictionary based MT is Ideal.

4. SYSTEM IMPLEMENTATION



Speech Recognition is the process of recognizinga certain word spoken by a particular speaker based on individual information included in speech waves. This technique is used to identify a speaker's voice to verify his/her identity and provide controlled access to services like voice-based biometrics, database access services, voice-based dialing, voice mail, and remote access to computers. Voice recognition basically means talking toa computer, then recognizing what one is saying. There are many types of features, which are derived differently and have a good impact on the recognition rate. This project presents the techniques to extract the feature set from a voice signal, which can be used in voice recognition systems. The voice recognition system performs two important operations: signal modeling and pattern matching.

5. PRE-PROCESSING

Before recognition, first the speech signal is preprocessed.It consists of pre-emphasis, endpoint detection, framing, and windowing. Speech recognition is has 4 stages:

- Analysis
- Feature extraction
- Modeling
- Matching

Analysis:

It deals with the stage with a suitable size for segmentingvoice signals for further analysis.

There are 3 techniques of performing analysis of speechsignals:

Segmentation analysis, Sub segmental analysis, Supra segmental analysisSupra segmental analysis:

In this analysis voice signal is analyzed using the behavior of the speaker. This paper refers to this technique for analyzing the voice signal as impulse voice signal is provided as input. The analysis takes place using the analysis of two parameters:

I. Amplitude

II. FrequencyExtraction:

Converting the noise waves into a parametric representation is a major part of any voice recognition approach. Here both static and dynamic features of voice used for voice recognition task because the vocal track is not completely characterized only by static parameters. For this various algorithm are available such as MFCC, PLP, and PNCC.This paper refers PNCC algorithm of feature extraction

PNCC:

Power Normalized Cepstral Coefficient is a front end technique that is slightly different than MFCC. It uses gamma filters instead of Mel-scale transformation. This provides Substantial improvement in accuracy as compared to MFCC and PLP.PNCC processing requires only about 33 percent more computation compared to MFCC.

Modeling:

Modeling refers to generating a speaker model using speaker-specific feature vector. This includes various types of approaches such as Acoustic phonetic Patternrecognition, Template-based, Dynamic time wrapping, Knowledge based, Statistical based, Learning-based. In this paper template-based approach for modeling is used In Template-based approach Unknown speech is comparedagainst a set of pre-recorded words (Templates) to find the best Match. This has the advantage of using accurate word models.

Matching:

For quick and automatic voice recognition technology, the digital processing of speech and voice recognition algorithms is considered to be very essential. A study of speech recognition is categorized as: speaker recognition and speaker identification.

Consider a sequence of feature vector {x1, x2.... xi} that represents the voice of an unknown speaker in the speaker recognition phase.These sequences are compared with codesfrom the pre-defined database. Euclidean distance is there

distortion distance measured between two vector set which aids identifying the unknown speaker.The formula used tocalculate the Euclidean distance can be defined as following:

The Euclidean distance between two points

[9] Uwe Mnegge(2006),"AN Excellent Application for Machine Translation;Automatic Translation of a large Database.in Elizabeth ,Proceedings of the annual conferance of the German society of Technical Commutator,Stuttgart;Telkom,1

$P = (p_1, p_2 \ldots p_n)$ and $Q = (q_1, q_2 \ldots q_n)$,

$$= \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$$

The speaker is identified by choosing the speakerwith lowest distortion distance.

6. CONCLUSION

This paper basically voice to text conversion and then text translation is done. Speech-to-Text conversion system is implemented with the help of PNCC for feature extraction asit provides substantial improvement in accuracy and for recognition Template-based approach is adopted which gives easiness. In this voice database, few audio files are recorded and these are analyzed to get feature vectors. These features are initially modeled in template-based after that the test speaken word is compared with the forward algorithm of template. Here for wireless communication purpose ZIGBEE technology is referred as it is broadly adopted for huge and fast transmission of information.

7. REFERENCE

[1] Santosh K Gaikwad,Bharti W Gawali,Pravin yannawar,"A Review on Speech Recognition Technique;University Aurangabad,Maharashtra."

[2] Namrata Dave,"Feature Extraction Methods LPC,PLP,MFCC in Speech Recognition; G.H. Patel College of Enginnering,Gujrat".

[3] h.Hermansky,"Perceptuallinear analysis of speech," Acoustical Society of America Journal,Vol 87,PP 1788-1752,Apr 1990

[4] Gellert Sarosi ,Mihaly Mozsary,Peter Mihajlik,and Tibor Fegyo, "Comparison of Feature Extraction Methods for Speech Recognition in noise free and traffic noise environment,Dept of Telecommunication and Media Informatics Budapest University of Technology and Economics Budapest ,Hungary

[5] ,THINK Research centre nonprofit LLC,Aitia International Inc".

[6] R.D.Patterson, K.Robinson, J.Holdsworth, D.MCKrown, C.Zhang & M.H.Allerhand, "Complex Sounds and auditoryimages ," in Pergamon Press,Oxford PP.429-446,1992

[7] D.B.Paul,"Speech Recogniton using HMM",The Lincoln Laboratory Journal,Vol 3,Number1 (1990)

[8] Virendra Chauhan,Shobhana

[9] Dwivedi, Karale,Prof.S.M.Potdar;"Speech To Text Converter using Gaussian Mixture Model; Sinhgad Academy of Enginnnering - Pune "

[10] Jurafsky Danie;Martin,James H(2009)."Speech and Language Processing ,PP.906-908"