American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-9, Issue-12, pp-34-39
www.ajer.org
Research Paper Open Access

Approachto Design Acontinous Speech Detection and Identification of Various Individual

B.J. Luckyn[1], TAMUNONENGIYEOFORI. P. DESMOND[2]
[1&2](Electronics/Computer Engineering,Rivers State University, Port Harcourt)

ABSTRACT: *Speech recognition technology is a fast-growing engineering technology in the world and Nigeria. It has a number of applications in different areas and provides potential benefits.Keeping track of the multiple passwords, PINs, memorable dates and other authentication details needed to gain remote access to accounts pose modern lifechallenges. The employment of a speech and voice-based verification as a biometric technology for both children and adults could be a good replacement to the old-fashioned memory dependent procedure. Using speech and voice for authentication could be beneficial in several application areas, including, security, protection, and education, call-based and web-based services. Nearly 20% people of the world are suffering from various disabilities; many of them are blind or unable to use their hands effectively. The speech recognition systems in those particular cases provide a significant help to them, so that they can share information with people by operating computer through speech and voice input. This work attempted to design and implement a speech detector system that would identify different users based on every spoken word uttered. Each user inputs audio samples with a keyword of his or her choice. Language is the most important means of communication and speech is its main medium. In human to machine interface, speech signal is transformed into analog and digital wave form which can be understood by machine. Speech technologies are vastly used and has unlimited practices. These technologies enable machines to respond correctly and reliably to human voices, and provide useful and valuable services. The work gives an overview of the speech recognition process, its basic model, and its application, approaches and implementation.*

KEYWORDS*: Speech and voice recognition, Speech and Voice identification, Security, Cybercrime, Speech and Voice authentication*

Date of Submission: 02-12-2020 Date of acceptance: 17-12-2020

I. INTRODUCTION

In Nigeria, nowadays due to the high rate of cybercrimes and criminality a lot of residential areas and companies or even individuals are using all kinds of security system to make sure their property is secured from cybercriminals such as using passwords and user ID/Pin for protection. Unfortunately, all these security systems are not secured at all because the pin codes can be hacked, the ID card can be stolen and duplicated. Based on these reasons a new form of technology of security system must be introduced to increase back the confidentiality of civilian about the security system. A biometric technology is the one which use the user features parameter as the password. The feature parameters of everyone are unique, even if the users are twins. Therefore, the speech recognition system is safe for administrative user.

Speech is the most natural way to communicate for humans. Speech recognition as a biometric technology is very important and will help in this region to curb the high rate of cybercrime and criminality that is why I will ensure that the aim of this project is carried out so as to help reduce the rate of cybercrimes in the country. Speech recognition is more convenient and accurate this is because the biometrics characteristics of an individual are unique and belongs to the individual until the user dies. It is convenient because with this the fear of carrying confidential belongings such as ID cards and the risk of it being stolen or password being hacked will be reduced.

Speech recognition is classified into two, systems that do not need to be programmed or trained are called "speaker independent"(Kavaler et al., 1987) systems. Systems that need to be programmed of trained are called "speaker dependent". The term speech recognition or speaker identification refers to recognizing the speaker, rather than what they speak (Reynolds et. al., 1995). Recognizing the speaker can simplify the task of translating speech in systems that have been trained on a specific person's voice or it can be used to authenticate or verify the identity of a speaker as part of a security process.

Nearly 50% people of the world are suffering or have suffered from cybercrime such as hacking; many of them are ignorant and lack proper security orientation to secure their important documents, money, etc only relying on passwords as a form of security. The speech recognition systems in those particular cases provide a significant help to them, so that they can protect and always access their properties by operating computer through speech/voice input. This work is designed and developed keeping those factors into mind. Our work is capable to recognize the speech and convert the input audio into text, so as to carry out the given task ofvarious individuals.

II. Aim

The aim of the work is to design approachesof continuous speech detector system that will recognize and indicate the speech sound of different user

III. OBJECTIVES

I. To understand speech recognition and its fundamentals.
II. To differentiate the working and applications of various user.
III. To understand how arduino uno r3 and sound impact sensors work.

IV. RELATED WORKS

Smith et al., 1990,invented and designed an independent speech recognition system with the use of isolated words or phrases programmed or stored into the system. An input word or phrase (the test token) was compared with a stored vocabulary (reference tokens) and was labeled as the closest matching vocabulary entry.

Patricia M.et al., **2006,** developed a system for intelligent pattern recognition to identify the speakers through analyzing signal of sound by use of intelligent technique, like the neural network (ANN) and fuzzy system. At first, they described voice recognition based on a used monolithic NN. Theyimplemented test with 20 different words (Spanish words), the records were from three different speakers. Their achievement on recognition resultedfromthe use of monolithic NN based recognition system that was considerable to about 96% recognition rate with increasing database over 100 words.

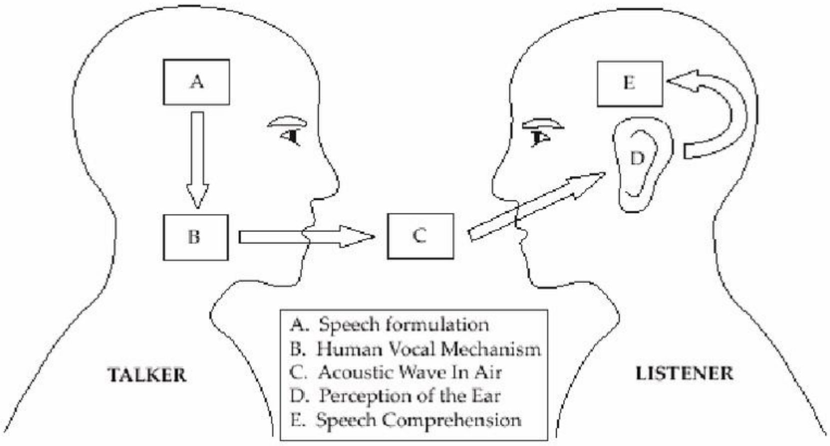
Odette et al., 2007, developed a continuous speech recognition system of polysyllabic word at Radboud University Nijmegen, Netherlands. They used 1463 polysyllabic words selected for that experiment. That experiment showed about 885 utterances made on the system, its accuracy was 64% polysyllabic words at the end of the utterance. Automatic speech recognition system was used for the recognition of the words. They used two types of predictors: first type of predictor was related with absolute and relative values of the word activation and second type of predictor was related to a number of tones that were present till the end of the word. The system gave 81.1% accurate result when local words activation was used to identify word before its last tone was available and the system gave 64.1% of those words which were already recognized as one tone after the uniqueness point.

Marta W. and Jacek D., 2010, developed a voice recognition system with hybridgenetic algorithm with a classifier of K-nearest neighbor. They proposed a satisfactory construction of the model and determined the simulation with parameters that influenced on the classification score. They determined the constructed method of performancefrom results that showed the overall system accuracy to about 94.2% of correct classified patterns in 26 seconds.

Hitesh G. and Deepinder S., 2014, developed an advanced algorithm for speech recognition using combination of FBCC and Genetic algorithm. The genetic algorithm for optimization gives better result for speech recognition. They observed that the level of accuracy using Hidden Markov model (HMM)was strongly influenced by the optimization of extraction process and modeling methods while improved results could be achieved with the help of Genetic algorithm. It was observed that recognition accuracy for feature extraction with FBCC features in comparison with MFCC is better. This algorithm has been tested on samples of various users with and without adding noise, higher degree of accuracy was achieved during recognition.

V. METHOD/IMPLEMENTATION PROCESS

- **SPEECH PRODUCTION IN HUMAN BEINGS**



FigureASchematic diagram of a speech production process in human being's

The figure A shows a schematic diagram of the speech-production/speech-perception process in human beings (L.R.Rabiner and S.E Levinson, 1981). The production (speech-generation) process begins when the speaker formulates a message (in his mind that he wants to transmit to the listener via speech). The machine counterpart to the process of message formulation is the creation of printing text expressing the words of the message. The next step in the process is the conversion of the message into a language code. This roughly corresponds to converting the printing text of the message into a set of phoneme sequence corresponding to the sounds that make up the words, along with prosody markers denoting duration of sounds, loudness of sounds, and pitch accent associated with the sounds. Once the language code is chosen, the talker must execute a series of neuromuscular commands to cause the vocal cords to vibrate when appropriate and to shape the vocal tract such that the proper sequence of speech sounds is created and spoken by the talker, thereby producing an acoustic signal as the final output. The neuromuscular commands must simultaneously control all aspects of articulator motion including control of the lips, jaws, tongue, and velum (a "trapdoor" controlling the acoustic flow to the nasal mechanism).

Once the speech signal is generated and transmitted to the listener, the speech-perception (or speech recognition) process begins. First the receiver or listener processes the acoustic signal along the basilar membrane in the inner ear, which provides a running spectrum analysis of the incoming signal. A neural transduction process converts the spectral signal at the output of the basilar membrane into activity signals on the auditory nerve, corresponding roughly to a feature extraction process. In a manner that is not well understood, the neural activity along the auditory nerve is converted into a language code at the higher centers of processing within the brain, and finally message comprehension (understanding of meaning) is achieved.

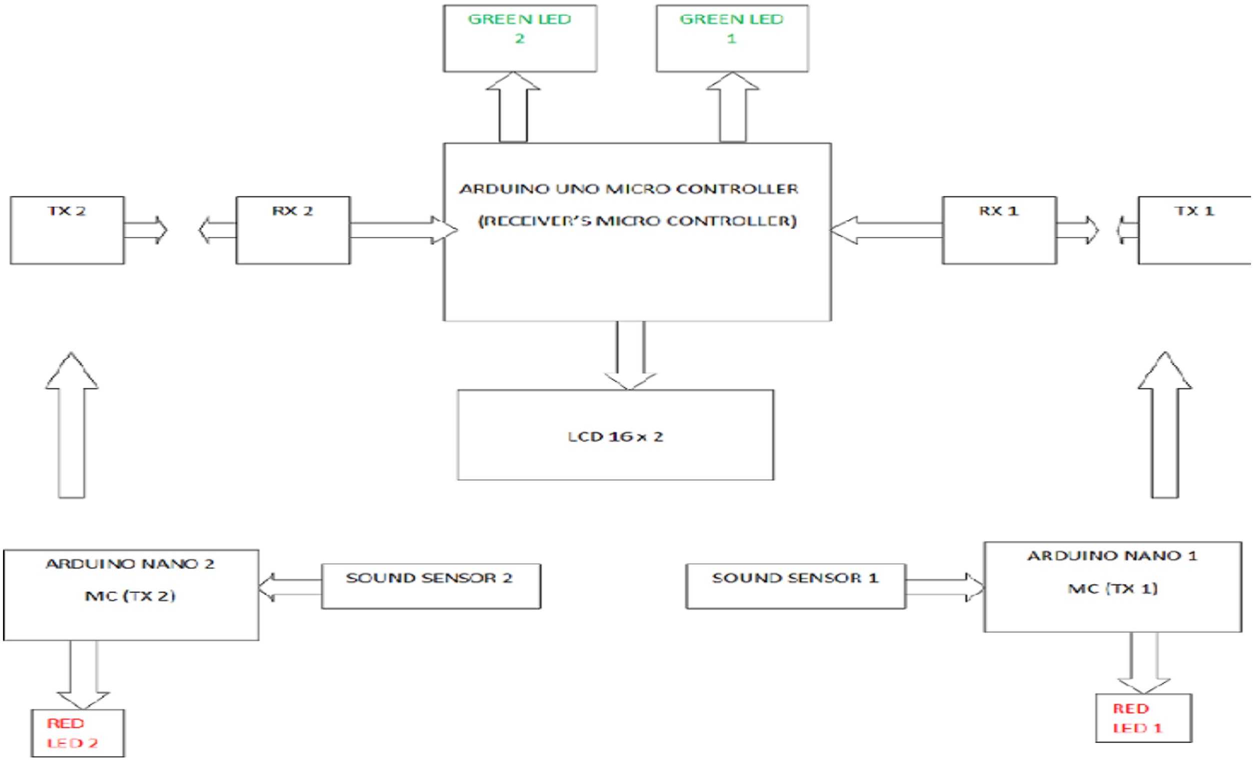


Figure B: Block diagram of a speech detection/recognition system

- **PROPOSED COMPONENTS**
  - Arduino Uno Board
  - Arduino Nano Board x 2
  - Wireless Transmitter x 2
  - Wireless Receiver x 2
  - Sound Impact Sensor (SIS) x 2
  - LCD 16 x 2
  - Green LED x2
  - Red LED x 2
  - Vero Board
  - Jumper Wires
  - Soldering Lead
  - 9V Battery
  - Pin Headers
  - Arduino IDE (Integrated Development Environment)

VI. CONNECTION BETWEEN THE RECEIVING ARDUINO MICRO CONTROLLER AND LCD 16 X 2

i. The receiving micro controller is the Arduino Uno micro Controller and it is connected to an LCD 16 x 2.
ii. The pin A4 of the receiving micro controller was connected to the pin SDA of the LCD.
iii. The pin A5 of the receiving micro controller was connected to the pin SCL of the LCD.
iv. The pin VDD of the LCD (which is 5V) is connected to the VCC pin of the receiving micro controller.
v. The pin VSS of the LCD was connected to the ground pin of the receiving micro controller
vi. This connection alone would transmit and receive signal via the receiving micro controller and finally display it on the LCD.

VII. CONNECTION BETWEEN THE RECEIVING ARDUINO MICROCONTROLLER AND TWO RECEIVER CIRCUITS

i. There are two receivers circuit under this session which are the RX 1 and RX 2 connected to the main receiving micro controller for the main circuit.
ii. Each RX has a definite LED which are LED 3 and LED 4 respectively.

iii. The digital OUT pin of the RX 1 was connected to the pin 11 of the receiving micro controller and the ANT pin of the RX 1 was connected to the VCC of the receiving micro controller.
iv. The digital OUT pin of the RX 2 was connected to the pin 7 of the receiving micro controller and the ANT pin of the RX2 was connected to the VCC of the receiving micro controller
v. The GND pin of the RX 1 and 2 are all connected individually to the ground pin of the receiving micro controller.
vi. The VCC of the RX 1 and 2 are all connected to the VCC of the receiving micro controller.
vii. The LED 3 for RX 1 has its positive pin connected to the pin 12 of the receiving micro controller and its negative pin directly into the GND of the receiving micro controller.
viii. The LED 4 for RX 2 has its positive pin connected to the pin 4 of the receiving micro controller and its negative pin directly into the GND of the receiving micro controller.
ix. This connection alone would receive signal transmitted from the respective transmitters into the respective RX via the receiving micro controller and then display on the LCD.

VIII. SOUND IMPACT SENSOR 1 AND 2, TRASNMITTER 1 AND 2 AND TRANSMITTING MICRO CONTROLLER 2 AND 3PROCESS OF CONNECTION AND WORKING METHOD

i. There are two transmitting micro controllers under this session which are Arduino Nano 1 and 2 known as micro controller 2 and 3 respectively.
ii. There two transmitters circuit TX 1 and TX 2 under this session.
iii. There are two sound impact sensors which are sound impact sensor 1 and 2.
iv. The digital OUT pin of the sound impact sensor 1 was connected to the pin 8 of the micro controller 2 and the digital OUT of the sound impact sensor 2 was connected to the pin 11 of the micro controller 3.
v. The VCC of both sound impact sensors are connected to the VCC of the individual micro controllers 2 and 3 respectively.
vi. The GND of both sound impact sensors are connected to the ground pin of the individual micro controllers 2 and 3 respectively.
vii. The DATA pin of TX1 was connected to the pin 6 of the micro controller 2 and the DATA pin of TX 2 was connected to the pin 7 of the micro controller 3.
viii. The ANT pin of TX 1 and 2 were connected to the VCC of the individual micro controllers 2 and 3.
ix. The VCC pins of the TX 1 and 2 were connected to the VCC of the individual micro controllers 2 and 3 respectively.
x. The positive pin of LED 1 was connected to the pin 4 of the micro controller 2 and the positive pin of LED 2 was connected to the pin 5 of the micro controller 3.
xi. The negative pins of LED 1and 2 were connected to the ground pins of micro controllers 2 and 3 respectively.
xii. This entire connections will send sound signals from the individual sound impact sensors 1 and 2 transmitting directly through the individual transmitters TX 1 and 2, which passes these signals from the respective micro controllers 2 and 3 and eventually being received by the receiver's circuits (RX 1 and 2) interpreted by the receiving micro controller 3.
xiii. The LED 1 and 2 were turned ON once transmitting sound signals passes through them from the sound impacts sensors (transmitted by the TX 1 and 2) through the micro controllers 2 and 3.
xiv. The LED 3 and 4 were turned ON once these sound signals were received by the respective receivers circuit RX 1and 2 and interpreted by the receiving micro controller 3 as voices.
xv. This is how the entire circuit works.

IX. SOFTWARE REQUIREMENT

i. Arduino IDE (1.8.13): This is the open source Arduino software which makes it easy to write code and upload it to the board. This software can be used with any arduino board, it is used to program the arduino board in the manner in which you want it to work or function. It is also used for running sketches.
ii. Proteus (8.7): this software is used for electronic circuit designs, and creation of schematics, simulation and electronic prints for manufacturing printed circuit boards. This software was used in other to execute this project.

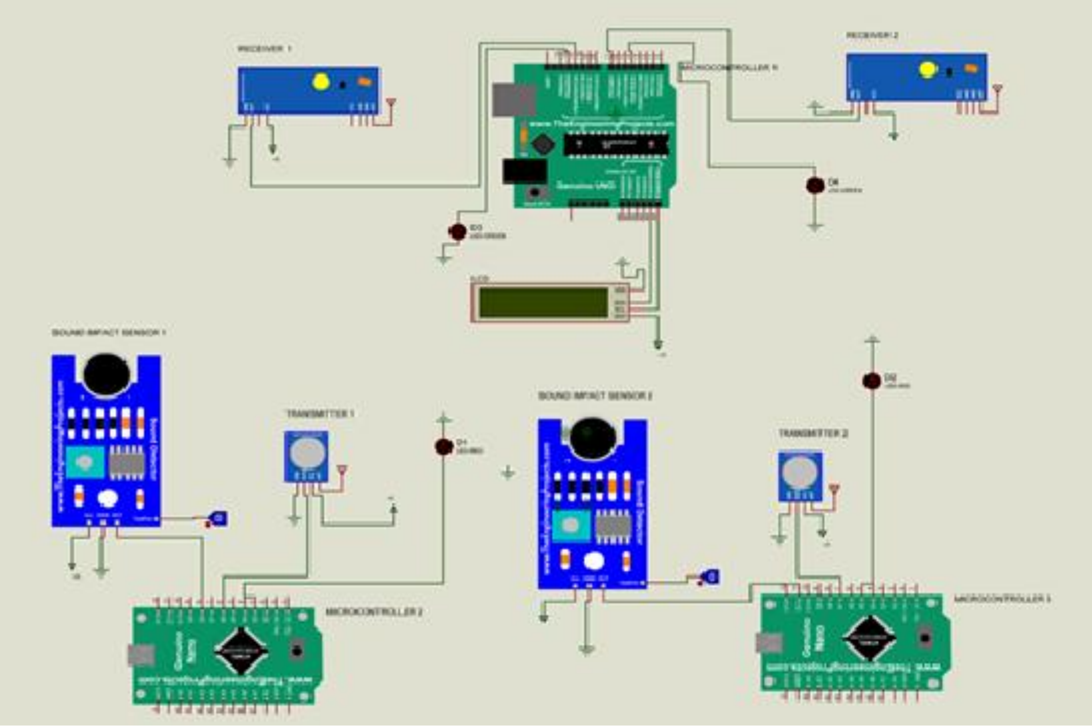


Figure C: The circuit design of a speech recognition system

X. CONCLUSION

The work shall avoid providing sensitive and risky personal information, institutions may opt to use speech detection/recognition to authenticate identities of different clients. This will help to curb fraud and phone crimes by use of speech biometrics in various institutions within the pandemic era.Finally, this work when constructed and applied on the systems such as phones, computers, indoor and outdoor units will curb the widespread pandemic.

REFERENCES

[1]. L.R.Rabiner & S.E Levinson., (1981) "Isolated and connected Word Recognition- Theory and Selected Applications,'' IEEE Trans. Communications, COM-29(25):621-659.

[2]. Patricia M., Jerica U., Daniel S., Miguel S., Miguel L., and Oscar C., (2006)"Voice Recognition with Neural Networks,Type-2 Fuzzy Logic and Genetic Algorithms", Engineering Letters, 13:2, EL_13_2_9 (Advance online publication: 4 August 2006).

[3]. Marta W., and Jacek D.,(2010) "Voice Command Recognition Using Hybrid Genetic Algorithm", Faculty of Technical Physics and Applied Mathematics, Gdansk University of Technology.

[4]. Odette S, Lou Boves and Louis ten Bosch., (2007) "On-line Early Recognition of Polysyllabic Words in Continuous Speech'', Roudboud University Nijmegen, Netherland.

[5]. Hitesh G., and Deepinder Singh W.,(2014) "Speech Feature Extraction and Recognition Using Genetic Algorithm", International Journal of Emerging Technology and Advanced Engineering (IJETAE), Vol. 4, Issue: 1, ISSN 2250-2459, pp. 363-369.

[6]. Kavaler, R. et al., (1987) A Dynamic Time Warp Integrated circuit for a 1000-Word Recognition System. IEEE Journal of Solid-State Circuits, vol SC-22, NO .p. 3-14.

[7]. Reynolds, Douglas, Rose & Richard., (1995) ''A Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models'', IEEE Journal on Speech and Audio Processing, VOL 3,.p 1-5

B.J. Luckyn, et. al. "Approachto Design Acontinous Speech Detection and Identification of Various Individual." *American Journal of Engineering Research (AJER)*, vol. 9(12), 2020, pp. 34-39.