ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE)
Vol 3, Issue 5, May 2016

Automatic Recognition for Instances of Repetitions in Stuttered Speech using Sequence Training Algorithm

[1] G. Manjula, [2] Dr. M. Shiva Kumar
[1] Associate Proffesor Dept. of ECE [2] Prof & Head, Dept. of EIE,
[1][2]GSSS Institute of Engineering & Technology for Women, Mysore

Abstract: ---- **Stuttering is a speech disorder in which normal flow of speech is messed out by occurrences of dysfluencies. There are high proportion of prolongations and repetitions in stuttered speech. Stuttered speech recognition technology is a great aid to admit the challenges and it is a prominent technology for Human-computer Interaction . Conventionally, stuttering assessment is done by counting number of dysfluent words as a proportion of total words in a passage. Stuttering assessment is also done by measuring the time of dysfluencies and comparing with the duration of the entire passage. However, it is time-consuming and it results in poor agreement with different individuals assessing the same speech data. Therefore, automatic stuttered speech recognition system is used to automate the dysfluency count and type of dysfluency classification, thus providing an objective and consistent assessment of stuttered speech. Such approach can support Speech Language Pathology (SLP) by doing tedious routine works and allowing more time for therapeutic session between SLP and stutterers. Instances of repetitions in stuttered speech are recognized by using back propogation based sequence training method. Artificial neural network is used to make intelligence through the learning process. In this work Multi-layer Feed forward Network and error back propagation learning algorithm is used to train the network. The combination of the artificial neural network and the back propagation algorithm is the sequence training technique. The experimental investigations reveal that the proposed method shows promising results in identifying the instances of repetitions in stuttered speech**

Keywords: **-Stuttered speech recognition, Sequence training, Back Propagation, PRAAT tool**

I. INTRODUCTION

Speech is the most common way of communication for people. We are born with the skill of speaking and learn it easily during our early childhood. By the developments of communication technologies in the last era, speech starts to be an important interface for many systems. Instead of using complex different interfaces, speech is easier to communicate with computers and robot. People are comfortable with speech therefore persons would also like to interact with computers or other type of machine via speech, rather than using primitive interfaces such as keyboards and pointing devices. This can be accomplished by developing an Automatic Speech Recognition (ASR) system which allows robot to identify the words that a person speaks into a microphone and converts it into written text or commands. 1% of world population suffer from stuttering [1-3]. Stuttering is known as dysphonia, which is a disorder that affects the fluency of the speech. It can cause a person to become fearful of human interaction and meeting new people, causing considerable anxiety and low self-confidence. The term stuttering is most commonly associated with involuntary sound repetition, but it also encompasses the abnormal hesitation or pausing before speech, referred to by people who stutter as blocks, and the prolongation of certain sounds, usually vowels or semi-vowels.

Speech recognition system can be separated in different classes by describing what type of utterances they can recognize such as Isolated Word, Connected Word, Continuous speech and Spontaneous speech. Stuttered speech is a type of continuous speech. Recognizer with continues speech capabilities are most difficult to create because they utilize special method to determine utterance boundaries. Classification methods such as, Hidden

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE)
Vol 3, Issue 5, May 2016

Markov Model, and Dynamic time warping for stuttered speech recognition are usually limited by computer memory space and computational time. We propose a new method using the sequence training and back propagation techniques and this paper provides an overview of stuttered speech recognition system using back propagation based sequence training technique. Back-propagation based sequence training method is attractive due to its mathematical simplicity and it can be efficiently implemented. In this method each node computes and stores shortest path that it has to reach from and finally tracing back in shortest path to recover the sequence is back propagation.

This work aims to analyze the stuttered speech and to control an unmanned vehicle even if the commands are stuttered. The phases are requirement analysis, designing, coding, testing, implementation and finally maintenance. Requirement analysis involves feasibility study and requirement of the needs to accomplish the desired results. Design involves hardware and software design for this work. Hardware design includes designing of various circuits such as preamplifier circuit for amplifying the stuttered voice, Motor driven circuit for the movement of motor. Software design includes analysis of library functions and coding for the purpose of training and testing. Coding involves software development for the designed work. Testing involves debugging of all individual modules and testing functionalities. During the testing or recognition phase, the feature of test pattern (test speech data) is matched with the trained model of each and every class if it is matched then the function is performed else error occurs.

The paper is organized as follows. There have been many research works done in the past for developing objective methods to facilitate the SLP during stuttering assessments. section 2 presents an overview of significant works that have been conducted for the objective assessment of stuttered speech, how they design the experiments and discuss their findings.
The methodology of the system is presented in Section 3. Furthermore, this section covers database that used in the experiment, feature extraction algorithm and classification technique. Experimental results and discussion are presented in Section 4. Finally, conclusions and future works are discussed in Section 5.

II. LITERATURE SURVEY

Speech is the most important mode of communication among human beings. Stuttering is a speech disorder which is affecting millions of people. The literature survey of stuttered speech analysis and recognition is carried out. Over the past 2 decades, there are a lot researchers being done on analysis and classification of stuttered speech recognition. It is observed that there are certain deviations in stuttered speech when compared to normal speech. The most common deviations identified are repetition of a sound, repetition of a syllable, repetition of a phrase or a sentence. The less common stuttering features include prolongations, blocking, and interjections. The research till now is not been able to identify any single cause for stuttered speech. Stuttering is currently thought to be a problem with the neural processing (brain activity) that underlies speech production. From the analysis carried out, it is evident there are modifications in speech production mechanism in the production of stuttered speech with respect to normal speech [16].

This analysis features of stuttered speech helps in identifying the best suitable feature extraction and classifier for the automatic identification of instances in stuttered speech. The most commonly used classifiers and feature extraction techniques to classify types of dysfluencies and also between stutterers and non stutterers are ANNs, HMM, SVM, Mel Frequency Cepstral Co efficient (MFCC), Dynamic Time Wrapping (DTW) . Each classifier and feature extraction technique provides different accuracies where in Mel Frequency Cepstral Co efficient (MFCC) feature extraction technique and classifier Hidden Markov Model (HMM) is proven to give highest accuracy 96% [20]. Table 1.2 gives the comparison between different classifiers and feature extraction techniques for an analysis and classification of stuttered speech.

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE)
Vol 3, Issue 5, May 2016

Table 1: Comparison between different feature extraction and speech recognition algorithms

Sl.No	Feature extraction /Classifier	Suitable For
1	Mel Frequency Cepstral Co efficient (MFCC)	Short time power spectrum
2	Dynamic Time Wrapping (DTW)	Requires less storage space but completely dependent on time and efficient for isolated
3	Hidden Markov Model (HMM)	Computationally more complex and needs more storage space, needs more training
4	Artificial Neural Networks (ANN)	Solves complex problems, it is an irreplaceable tool which is widely used in stuttering recognition.

III.DESIGN METHODOLOGY

In this study, sequence training technique derived from three different steps which are used for the recognition of stuttering events. Sequence training technique involves 3 steps: Split, Predict, Update as shown in Fig 3.1. ANN is employed to evaluate the effectiveness of the feature in the recognition of repetition and prolongation in stuttered speech. Fig. 3.2 illustrates general block diagram of stuttering recognition.

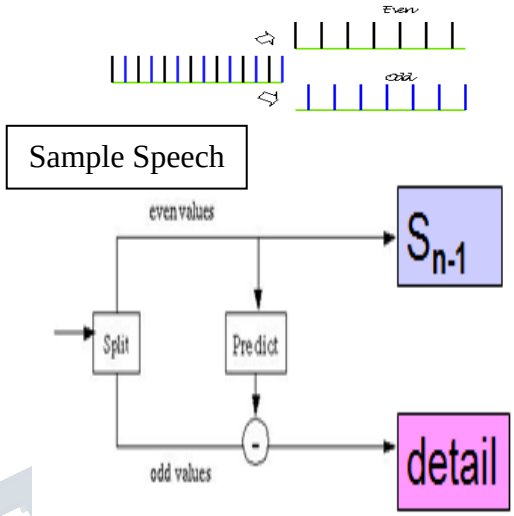

Fig 3.1: Sequence Training Technique

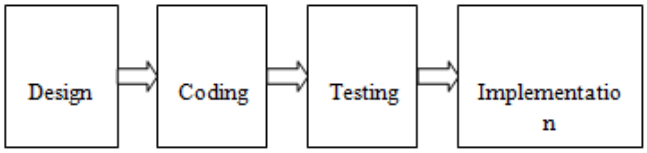

Figure 3.2 : Design Flow

Designing the prototype for this project includes both hardware and software units. First part of the project deals with recording stuttered voice signal by microphone, while recording there is lot of noises, using high quality microphone it is able to record human voice and helps to reduce noise. This stuttered speech signal is then analyzed to produce a representation consisting of salient features of the speech. The most prevalent feature of speech is derived from its short-time spectrum each short-time spectrum is transformed into a feature vector, and the temporal sequence of such feature vectors thus forms a speech pattern. Voice module is previously burned with recorded stuttered voice commands. The LEDs on the VR module glow indicating the various stages like waiting for record, recording stage, termination of record stage. If the given stuttered voice command is matched with recorded voice command then related motor action is performed. Microcontroller controls the action on Arduino platform with the power supply either from the battery or from the laptop.

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE)
Vol 3, Issue 5, May 2016

The second part is the software, the software must be equipped with vocabulary and voice commands. This words or vocabulary must be trained with different stuttered voices,and then software will have many different voices and pitches to use them as source to recognize the voices. In the last part the unmanned vehicle will receive the signal. Robot navigation is done using ATmega328 microcontroller.

IV. IMPLIMENTATION

The below block shows the important requirements for the project are VR module, DSP processor, Arduino board, ATmega328 microcontroller, motor driver circuit, motors, PRAAT and Arduino software.

VR module: It could recognize the voice. It receives configuration commands or responds through serial port interface. This module can store up to 15 pieces of voice instruction. We should train the module with voice instructions. This module is speaker dependent. It has DSP processor and PRAAT software package embedded in it and microphone is connected to the module. The code is burned to the processor of VR module for the functioning of the motor for the given stuttered voice commands.

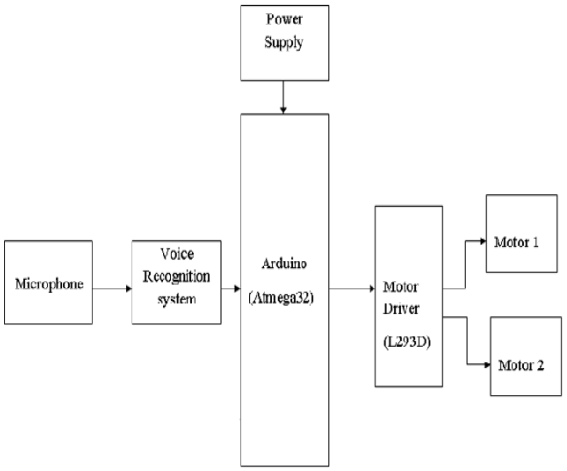

Figure 4.1: Block Diagram

ATmega328 Microcontroller: The true computer on a chip is nothing but a microcontroller. It is a powerful microcontroller that provides a highly-flexible and cost-effective solution to many embedded control applications.

Arduino Uno: It is a microcontroller board based on the ATmega328. It has 14 digital I/O pins, 6 analog inputs, a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button.The Arduino Uno can be powered via the USB connection or with an external power supply. External power can come either from an AC-to-DC adapter or battery.

Motor Driver: L293D is a dual H-bridge motor driver integrated circuit. Motor driver act as current amplifiers since they take a low-current control signal and provide a higher-current signal. This higher current signal is used to drive the motors.

Microphone and DC motor: It is an acoustic-to-electric transducer or sensor that converts sound in air into an electrical signal. It takes the analog voice commands and sends it to voice recognition chip in the form of electrical signal.Geared DC motor used is of 300 RPM, 12V with metal gearbox and metal gears.

Arduino software: The easiest way to program the board is with the Arduino software which is free and open source, and available for Windows, Mac OS X and Linux users. The programming language that the Arduino platform uses embedded C. Arduino calls these programs sketches. Then a utility called compiler reads the sketch and converts it to machine instructions that the Arduino understands. The most challenging part of the entire system is designing and interfacing various stages together.

V.RESULTS

Ideal Condition:

Table 2: Results obtained under ideal condition

Commands	No. of times command given	No. of times command recognized	Efficiency (%)
Left	10	10	100
Right	10	10	100
Forward	10	10	100
Back	10	10	100

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE)
Vol 3, Issue 5, May 2016

Normal Condition:

Table 3: results obtained under normal condition

Commands	No. of times command given	No. of times command recognized	Efficiency (%)
Left	10	9	90
Right	10	9	90
Forward	10	10	100
Back	10	9	90

VI. CONCLUSION AND FUTURE SCOPE

In this project we have designed a user dependent continuous word recognition system with the recognition accuracy of about 92.5% under normal conditions using 4 stuttered speech commands.

The accuracy can be improved further and the system can be used for more number of words during training of the system. Noise conditions affect recognition accuracy of the system. To overcome disadvantage in this system Adaptive Noise Modeling has to be done to improve the system's performance.

REFERENCES

[1] Andrzej Czyzewski, Andrzej Kaczmarek1and Bozena Kostek, "Intelligent processing of stuttered speech", *Journal of Intelligent Information Systems*, 10.1.1.109.715, 2003

[2] Hariharan.M, Vijean.V, Fook.C.Y, Yaacob.S, "Speech stuttering assessment using sample entropy and Least Square Support Vector Machine", *Signal Processing and its Applications (CSPA)"*, ISBN: 978-1-4673-0960-8, pp. 240-245, 2012

[3] Michael D. McClean, D.C. Stephen M. Tasko,MI Charles M. Runyan James "Orofacial Movements Associated With Fluent Speech in Persons Who Stutter" *Journal of Speech, Language, and Hearing Research*,Vol. 47, pp 294–303, April 2004 ©American Speech-Language-Hearing Association 1092-4388/04/4702-0294

[4] Jiang J, Lu C, Peng D, Zhu C, Howell P, "Classification of Types of Stuttering Symptoms Based on Brain Activity*",* *PLoS ONE, e39747. doi:10.1371/journal*.pone.0039747,2012

[5] Arghavan Bahadorinejad, Farshad Almasganj, " Delayed Auditory Feedback for Speech Disorders", *International Conference on Biomedical Engineering*, 978-1-4577-2, 2011.

[6] Marius Cristian, Adrian Graur, "Developing a Logopaedic Mobile Device Using FPGA", *Applied Computational Intelligence and Informatics*, 1-4244-1234-X, 2007

[7] Walter Dosch, Nontasak Jacchum, " Stuttering Removal-Developing Mealy and Moore Style Implementations of an Interactive Component", *Computer Systems and Applications*, ISBN 978-1-4244-3807-5, pp. 302-308, 2009

[8] Juray Palfy, Jirt Pospichal, " Pattern Search in Disfluent Speech", *IEEE International Workshop on Machoine Learning for Signal Processing*, 2012

[9] K.M Ravikumar, R.Rajagopal, H.C.Nagaraj "An Approach for Objective Assessment of Stuttered Speech Using MFCC Features" *ICGST, DSP Journal*, Volume 9, Issue 1, June, 2009 pp. 19-24, june, 2009.

[10] W. reichl and G. Ruske, " Syllable Segmentation of Continuous Speech with Artificial Neural Networks," *In processing of Eurospeech*, Berlin, Vol.3,pp.1771-1774,1993.

[11] P.Howell, S. Sackin and K. Glenn,"Development of a two stage procedure for the automatic recognition of disfluencies in the speech of children who stutter: II. ANN recognition of repetitions and prolongations with supplied word segment markers," *Journal of speech Language and Hearing Research*, Vol 40, pp. 1085, 1997

[12] Y.V. Geetha, K. Pratibha, R.Ashok, and S.K.Ravindra" Classification of childhood disfluencies using neural networks" *Journal of Fluency disorders*, vol.25, pp.99-117, 2000.

[13] Coneliu Octavian, "A Comparative Study of Feature Extraction Methods Applied to Continous Speech Recognition in Romanian Language", *48th International Symposium ELMAR-2006*, 07-09 june 2006,

[14] Vimala.C, Dr.V.Radha,"A Review on Speech Recognition Challenges and Approaches", *World of Computer Science and Information Technology Journal*, Vol 2 No1, ISSN 2221-0741, pp. 1-7, 2012.