



OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

PRACTICE PARTNER FOR STAMMER THERAPY

Prachi Prabhakar Sawant¹, Vaibhav Shivaji Ughade², Yogesh Jorsing Rathod³, Jyotsna Satish Sabne⁴
 Student, Computer Engineering Dept., K. K. Wagh Institute of Engineering, Education and Research, Nashik, India^{1 2 3 4}

sawantprachi840@gmail.com , vaibhavsughade@gmail.com, yrathod426@gmail.com , sabnejyotsna@gmail.com

Abstract: A communication disorder, known as stammering, is a problem in which a person's flow of speech is affected by repetition of a word or syllable or sometimes involuntary pauses leading to missing words. About 1% population in the world suffers from this problem. Even though being a serious concern, the awareness about it and its treatment is far less. In some parts of the world, the idea of implementing such a system has been given a thought, but due to some limitations an effective system is still not developed. The proposed system addresses the issues of repetition and pauses by identifying stammer for literate individuals. Further the scope can be extended for prolongations and other regional languages.

Keywords: Artificial Intelligence, Neural networks, Speech recognition and text analysis, Stammering and treatments.

I INTRODUCTION

Stammering is a communication disorder in which the speech of a person is disrupted by sudden involuntary pauses and/or repetition of letters or words. After consulting some speech language therapists, it was found that a major reason for this could be the improper rate of flow of speech and lack of confidence. According to statistics, around 70 million population in the world and around 5 crore people in India suffer from stammering. Though it is a very major problem, it mostly goes unrecognized due to lack of awareness, unaffordability of treatment cost, fear and embarrassment.

Stammering generally occurs in following manner:

- 1) Repetition: I can can can can speak well.
- 2) Prolongations: I ccc...can speak well.
- 3) Blockages: I ____ speak well.



Speech therapists are specialist doctors who help the patients to overcome communication speech disorders. Following are some common treatments used by them:

- 1) Reading: To control the rate of flow of speech.
- 2) Shadowing: Reading along with the therapist
- 3) Delayed Auditory Feedback: A technique by which the original acoustic speech signal is artificially modified and then fed back after a brief time delay [1]

Nowadays many applications and systems are being developed for handicapped, blind and people facing other kind of serious problems. But stammering goes neglected somewhere and hence this, and all the above-mentioned factors have led to a strong motivation for developing such a system.

In this paper section I provide introduction, idea and motivation for the system. Next section presents survey of work done until now to address this issue. Section III explains the proposed system which aims to provide user friendly and efficient solution. Next section includes conclusion and future work which can be done to increase usability and quality of the system.

II LITERATURE SURVEY

Stammering not only affects physically but also causes emotional and mental effects on the person. In the survey conducted by "I have a voice" stuttering awareness campaign, it was observed that 4 out of 10 people who stutter have been denied a job promotion or school opportunity because of stuttering.[2]

The study of the fact that around 52% people who participated in a survey conducted by "I have a voice", felt

that a digital assistance along with the therapy will definitely be helpful [2]. This laid the foundation of a strong motivational base for the project.

Automatic Speech Recognition is the method to develop audio recognising digital tools. The paper by M. A. Josephine and Dr. E. Chandra discovers the same by targeting young children. The paper includes review of stammering, types of stammering, therapies for stammering and their techniques. The paper also describes computer assisted tools for speech therapy. One of the mentioned tools uses an amplitude plot of the speech signal so that the clients receive an audio visual feedback. A phrase is presented to the user and the client repeats the same phrase. The software compares the user's speech with the given phrase by superimposition in real-time. The software tracks the patient's progress by storing it every time on a disk which the user can carry to consult with the therapist.

The paper also lists some existing tools and devices like Speech Monitor, Fluency Coach, Fluency Master, etc. However, all the existing intelligent systems were found less user friendly and in some way expensive by the software users. Also many of them expect more of input parameters to identify their problems rather than the system autonomously identifying it. [3]

Lim Sin Chee et al. have reviewed techniques that are utilized in automatic stammering recognition for evaluating speech problem for stutterers. They have compared different classification techniques proposed by previous researchers like Artificial Neural Networks (ANNs), Hidden Markov Model (HMM) and Support Vector Machine (SVM).

Method 1: Researchers from many scientific disciplines are designing ANNs to solve problems in pattern recognition, prediction, optimization, associative memory and control. Neural networks play an important role in speech recognition because the repetitions and prolongations are present everywhere in the stammered speech. Each word is judged as fluent, repetitions, prolongations or other stammering categories. The training is achieved by linking the audio input either onto repetitions or prolongations output through a layer of hidden units.

Method 2: HMM is a stochastic model that captures the statistical properties of observed real world data. In speech recognition, speech signal could be viewed as a piecewise stationary signal or a short-time stationary signal. Thus, HMMs are widely used in speech recognition. The paper describes Speech Therapy Assistance Tools (MSTAT) is a system assists SLP to diagnose children for language disorder and to train children with stuttering. The voice patterns of the normal and speech disordered children are used to train the HMM model. If the score is greater than threshold's value, it is diagnosed as normal and vice-versa.

Method 3: SVM is a powerful machine learning tool widely used in the field of pattern recognition. SVM optimization problem attempts to obtain a good separating hyper-plane between two classes in the higher dimensional space. SVM is used as classification tool in stuttering recognition. The paper describes automatic detection method for syllable repetition in reading speech for objective assessment for stuttered dysfluencies. It has four stages comprising of segmentation, feature extraction, score matching and decision logic.

According to the conclusion presented in the "Overview of automatic stuttering recognition system", it was concluded that HMM provided 96% accuracy, SVM yielded 94.35% accuracy while ANNs achieved 94.9% accuracy.[4]

There are 2 probable ways for implementation of such a system: Either speech signal processing or speech to text conversion and then processing as suggested by De Silva, 2016. [1]

Chen et al. implemented a project which included speech signal processing. The objective of this research was to develop an algorithm that allows for identification and removal of a specified word within an audio file through waveform analysis. The application of this objective was to reduce time for manual editing of audio streams. A specific application for this program was the identification of excessive stutter words in a voice recording. But, the recognition of stammered words using waveform analysis however provided less accuracy and couldn't be used for real time sound processing. [5]

Hence the proposed system is based on De Silva's "Smart speech therapist for stammer". The author applied NLG (Natural Language Generation) with Hidden Markov model for self-learning and automated sentence generation therapy methodology for stammer patients. SSTS application is expected to mimic some of the treatment activities performed by the Speech and Language Therapist in the real life like:

- Provide a random sentence to the child. (This undoubtedly shall be simple and appropriate).
- Recognize what is being pronounced by the child.
- Compare the original sentence and the voiced sentence. It is important to differentiate whether the sentence is correctly pronounced, and to identify whether there are any repeated (stuttered), missed or additional words or sentences in the speech.
 - Calculate a score for the speech. (According to a scoring criterion prepared using the repeated, missed or additionally pronounced words, etc.)
 - Speech results (Score, repeated words) shall be recorded and should be used in generating new user specific sentences.

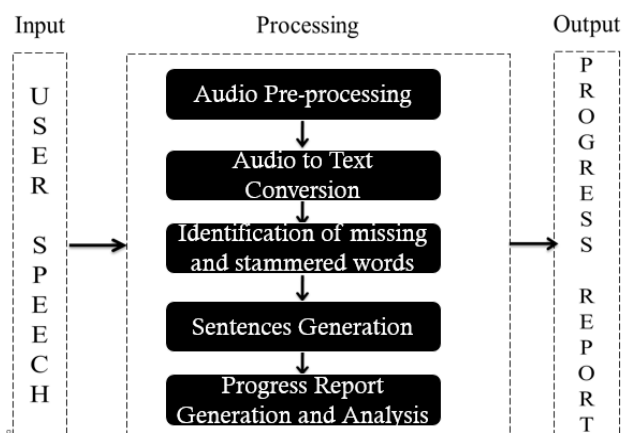
- User should be able to measure their progress (via progress charts).

The author could implement all the five main modules by the end of the project as it was proposed. But unexpectedly the accuracy rate of the Adaptive Paragraph Generation module was limited to 50%, which was not enough. Hence he suggested using neural networks or vectors for sentence generation instead of Hidden Markov model. [1]

Hence the proposed system intends to use artificial neural networks and implement a similar system with an expectation of achieving higher accuracy. The user of the system can track his progress by periodic progress report generation. A detailed description of the system is provided in the next section.

III PROPOSED SYSTEM

The proposed system consists of 5 important modules as shown in the below figure.



- **Audio Pre-processing**

Before providing the input to the API for speech recognition, noise removal algorithms are applied so as to provide improved speech input and improve efficiency of speech recognition.

- **Audio-to-text conversion or speech recognition**

After pre-processing, the Google's voice recognition API, which is already implemented and available for usage, is used for speech recognition. Since it is currently the best available API, it is a preferred choice.

- **Identification of stammered words**

The text file generated from previous process contains stammered words which will be detected using neural networks algorithms and some inbuilt language constructs.

- **Sentences Generation**

Based on the identified word, some sentences are provided to the user to read, that contain those same stammered words. The user has to read those paragraphs and again an audio file will be generated and the entire previous process is repeated in a cyclic way.

- **Progress report generation and analysis**

Ultimately, the user audio inputs will be stored in the backend and whenever the user wants to generate a progress report, the current input will be compared with previous ones to provide the improvement analysis.

IV CONCLUSION AND FUTURE WORKS

The proposed system aims to provide a digital system to stammering people to overcome their stammering problem. By making the use of proposed system, the speech therapist can treat their patients using a digital assistant and evaluate the performance using the progress report. It aims to help stammering people by providing a user friendly and a simple digital software which is easy to use and access and is inexpensive.

In future, the scope of the system can be expanded to identify prolongation type of stammer. Also currently the systems available are for English only. The scope can be extended to other regional languages also.

ACKNOWLEDGMENT

We are grateful to our guide, Prof. Neha Sharma for her valuable guidance in this research work. We would also like to show gratitude towards our institution K. K. Wagh Institute of Engineering Education & Research, Nashik and Head of Computer Engineering Department Prof. Dr. S. S. Sane. We would also like to thank our colleagues for their useful insight towards this work.

REFERENCES

- [1] Yasas De Silva, "Automated and Self-learning sentence generation methodology for stammer patients by means of Natural Language Generation", in, 2016
- [2] <http://www.ihaveavoice.info/survey.html>, Stuttering Facts and Stats (Online)
- [3] M A. Josephine Sathya, Dr. E. Chandra, "Types and Tools Available for Fluency Disorder – Speech Therapy", in International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 2, February 2015, ISSN: 2277 128X./
- [4] Lim Sin Chee, Ooi Chia Ai, Sazali Yaacob, "Overview of Automatic Stuttering Recognition System," in proceedings of International Conference on Man-Machine Systems (ICoMMSS) 11-13 October 2009, Batu Ferringhi, Penang, Malaysia
- [5] Recognition of Sound Sample Sequences Using Waveform Analysis: Detecting Stutter Words in an Audio Stream