ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014

Stuttered Speech Recognition For Robotic Control

G. Manjula, Dr. M. Shiva Kumar

Abstract: - Stuttering is a speech disorder which is affecting millions of people in their day to day life. Stuttering is a complex speech problem which affects the verbal communication. A mobile robot is built, which can be controlled by the user by employing the voice commands such as, 'Run', 'Stop', 'Left', 'Right' and 'Backward' issued by the user. The robot will identify the voice commands even if they are stuttered. A novel method employing the DTW algorithm for feature matching has been implemented where the matched features generate the corresponding function key. After processing the stuttered speech, the necessary motion instructions are given to the mobile platform via a RF link. This work can be used to upgrade speech recognition systems.

Keywords: **DTW: Dynamic Time Warping, ED: Euclidean Distance, MFCC: Mel Frequency Cepstral Coefficients, PAA: Piecewise Aggregate Approximation.**

I. INTRODUCTION

Speech is the most effective way of interpersonal communication. Only 5% to 10% of the human population has a completely normal mode of verbal communication with respect to various speech features and healthy voice; and the remaining 90 % to 95 % suffer from one disorder or the other such as stuttering, cluttering, dysarthria, apraxia of speech etc [1], [2]. Stuttering is a speech disorder which is affecting millions of people in their day to day life. Stuttering is a problem that **interferes** with plain speech. A person who stutters may repeat the first part of a word (as in wa wa wa water) or hold a single sound for a long time as in (caaaaaaaaaaaaaaaake). Interjections such as "um" or "like" can occur as well; *particularly when they contain repeated ("u-um-um") or prolonged ("uuuum").* More than 68 million people worldwide stutter and has found to affect males and females in the ratio of 4:1. This disorder is characterized by disruptions in the production of speech sounds, called disfluencies [3], [4].

II. DESIGN METHODOLOGY

As a prototype for analysis of stuttered speech, this work is designed to build a mobile robot which will identify a voice command, even if the command is a stuttered one.

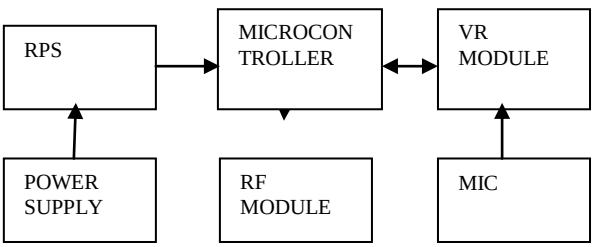

Fig 1: Transmitter of stuttered speech Recognition system

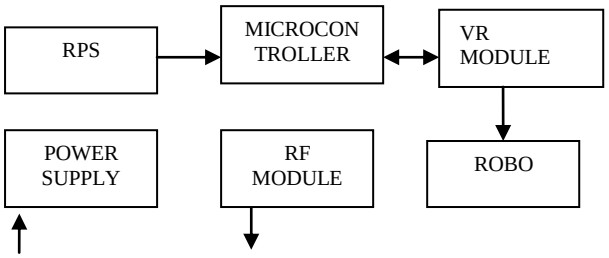

Fig. 2: Receiver of stuttered speech Recognition system to control Robot

The block diagram shown in Fig. 1 indicates the complete overview of transmitter where in stuttered speech is recognized and Fig 2 indicates how the recognized stuttered speech is used to control the mobile robot. The transmitter consists of a Voice Recognition (VR) module. The voice command are sent to the chip using a microphone and then the module converts the voice command to direction command that is predefined and recognizable by the robot. The receiver is the motorized robot. It consists of two DC motors and will make the robot move in forward or in the backward direction. For the recognition of stuttered speech, the test data is converted to templates. The recognition process then consists of matching the incoming speech with stored templates. The template with the lowest distance measure from the input pattern is the recognized word. The best match (lowest distance measure) is based upon dynamic programming. This is called Dynamic Time Warping (DTW) word recognizer. Firstly, the voice command is stored in the Voice recognition module with the help of the function keys available. This acts as a sample voice. Then, the input voice is fed into the microphone. It converts the physical voice signal into electrical signal. The signal received is fed into the voice recognition module. The voice received is processed in the voice recognition system and the characteristics of this voice command are matched with that of the pre-stored sample voice. The voice recognition module sends corresponding string to the microcontroller, which is then transmitted through an RF module. The receiver receives the desired voice signal from the RF transmitter through wireless link. As soon as the RF module recognizes the signal, it is sent to the microcontroller. It must be noted that the microcontroller cannot produce sufficient current to drive the motor directly. The microcontroller output is +5volt and can give a maximum current of 5mA. The D.C. Motor used in this work operates at 12 volt and requires approximately 400mA of current. The motor driver is designed to interface the motor with microcontroller. The driver stage changes the current and voltage level suitably to drive the motor. L293D is a dual

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014

H-bridge motor driver integrated circuit that is used in this work. The 7-segment LED display is interfaced with the microprocessor to see the status of the robot.

A. Programming aspects of Transmitter and Receiver Section

The software implementation of various blocks mentioned in the previous section is done using the Embedded C. In the transmitter side, at first it will initialize the microcontroller and the variables like port. If the logic 1 data is given to recognize, then it will send data 0x05 as move forward command to the receiver. Similarly, the other data will be given to recognise. Likewise, the transmitter will identify the command in the software part and then the corresponding data will be sent to the receiver. At the receiver side, the microcontroller and the RF receiver are initialised to receive the signal. If the data received is 0x05, then the robot will recognize the command as move forward and it will follow the command. Similarly, other data will be recognised.

B. Flow Chart for creation of new database

The creation of the new data base involves the follow steps.

1. The module should be trained with voice instructions.
2. The group will be imported to the database.
3. For each speech signal, features are extracted in the VR Module to generate the templates.

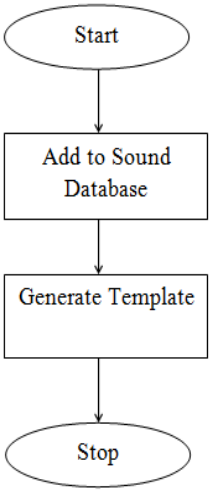

Fig. 3 Flowchart to add to database

Flow Chart for recognition of stuttered speech

1. Input the command to be recognized.
2. Calculate the MFCC for the speech signal.
3. Calculate the Euclidian distance (ED) for the recognition of speech and compare the ED of the stuttered speech signal to be recognized with the stored speech of each signal.
4. Compare each codeword of the signal with the reference patterns and output the corresponding string.
5. If no match is found, score matching is done using DTW.

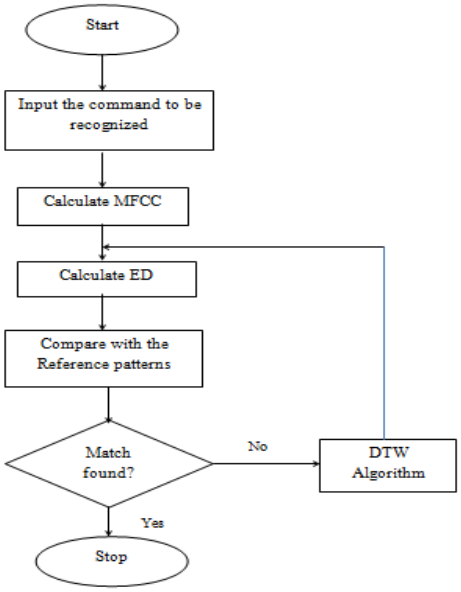

Fig.4 Flowchart for Stuttered Speech Recognition

III. SIMULATION AND RESULT

A. Transmitter and Receiver sections

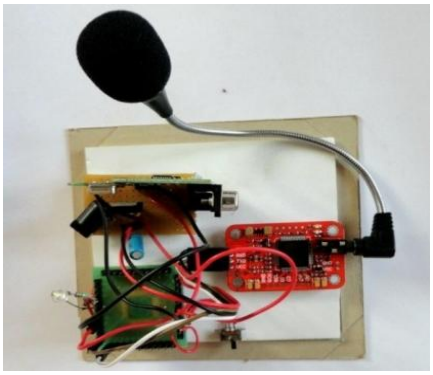

Fig. 5 Transmitter section of stuttered speech recognition module

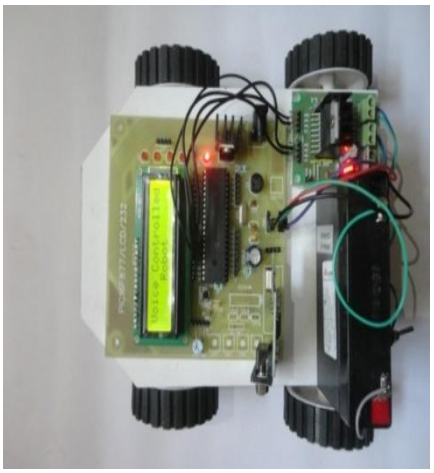

Fig. 6 Receiver section of stuttered speech recognition module

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014

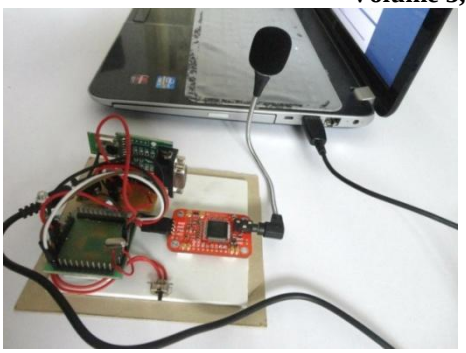

Fig. 7 loading the samples to transmitter module

The robot moves forward, left, right, backward, stop depending on the user, what he speaks into the microphone for example say "r r r r r run, "lef f f f f t", "r r r r r right", b b b backward, s s s s stop respectively as shown in Fig 8, wherein the robot is moving forward.

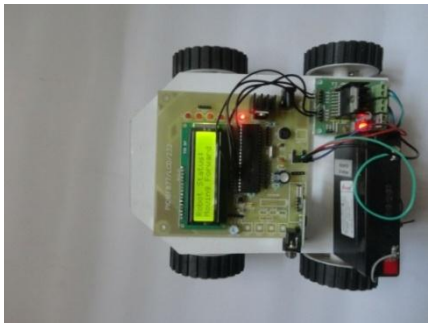

Fig 8 Robot moving forward/backward/left/right/stop

IV. CONCLUSION

In this work, a voice controlled robot is developed which is also able to recognise stuttered speech. This is fuelled by the need to improve the speech quality by identification and elimination of stuttering and also help Speech Language Pathologists (SLP) to access stuttering patients. The system has practical coverage up to a few meters. It uses the DTW algorithm for feature matching and the matched features generate the corresponding function key. After processing the speech, the necessary motion instructions are given to the mobile platform via a RF link. As the training data was small as compared to the average needed data for building a stuttered speech recognition system, the error rates were pessimistic. The error rates could be bifurcated into two groups depending upon speaker independency. DTW classifier was trained using the utterances of a single person. Speaker dependant: The error rates were approximately 10% to 20% Speaker independent: The error rates were approximately 30% to 40%

REFERENCES

[1] Young, M. A., "Predicting Ratings of Severity of Stuttering" [Monograph], Journal of Speech and Hearing Disorders, Pp. 31-54, 1961.

[2] Sherman.D, "Clinical and Experimental use of the Iowa Scale of Severity of Stuttering", Journal of Speech and Hearing Disorders, Pp. 316-320, 1952.

[3] Cullinan. W.L, Prather. E.M & Williams. D, "Comparison of Procedures for Scaling Severity of Stuttering", Journal of Speech and Hearing Research, Pp. 187-194, 1963.

[4] Oliver Bloodstein, "A Handbook on Stuttering", 5th Edition, Singular Publishing Group, Inc., San-Diego and London, 1995.

[5] M N Hegde, Deborah Davis, Text book on "Clinical Methods and Practicum in Speech Language Pathology", 5th Edition, Cengage learning publisher, 2005.

[6] Andrzej Czyzewski, Andrzej Kaczmarek, Bozena Kostek, "Intelligent Processing of Stuttered Speech", Journal of Intelligent Information Systems archive, Volume 21, Issue 2, Pp. 143-171, September 2003.

[7] K. Ravikumar, B. Reddy, R. Rajagopal, and H. Nagara, "Automatic Detection of Syllable Repetition in Read Speech for Objective Assessment of Stuttered Disfluencies", Proceedings of World Academy Science, Engineering and Technology, Pp. 270-273, 2008.

[8] 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) 2012, Pg no:240-245, ISBN :978-1-4673-0960-8.

[9] Jiang. J, Lu C, Peng D, Zhu C, Howell P (2012) Classification of Types of Stuttering Symptoms Based on Brain Activity PLoS ONE. 7(6): e39747. doi: 10.1371/journal.pone.0039747

[10] Arghavan Bahodorinejad, Farshad Almasganj. "Delayed Auditory Feedback for Speech Disorders". 2012 International Conference on Biomedical Engineering, 978-1-4577-2, 2011 IEEE.

[11] Marius Cristian, Adrian Graur, "Developing a Logopaedic Mobile Device Using FPGA", 1-4244-1234-X, 2007, IEEE.

[12] Walter Dosch, Nontasak Jacchum, "Stuttering Removal-Developing Mealy and Moore Style Implementations of an Interactive Component"

[13] Juray Palfy, Jirt Pospichal, "Pattern Search in Disfluent Speech", 2012 IEEE International Workshop on Machine Learning for Signal Processing, Sept 23-26.2012, Santander, Spain.

[14] Ravi Kumar. K. M, Ganesan. S, "Comparison of Multidimensional MFCC Feature Vectors for Objective Assessment of Stuttered Disfluencies". Advance Networking and Applications, Volume: 02 Issue: 05, Pages: 854-860(2011).

[15] W. reichl and G. Ruske, "Syllable Segmentation of Continuous Speech with Artifical Neutral Networks." In processing of Euro speech, Berlin, Vol.3,pp. 1771-1774,1993.

[16] 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.p.1085,1997

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

[18] Coneliu Octavian DUMITRU, Inge GAVAT,"A Comparative Study of Feature Extraction Methods Applied to Continuous

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014

Speech Recognition Language", 48th International Symposium EL MAR-2006, 07-09 June 2006, Zadar, Croatia.

[19] Vimala.C, Dr.V.Radha, "A Review on Speech Recognition Challenges and Approaches", World of Computer Science and Information Technology Journal, ISSN 2221-0741, Vol 2.No 1, 1-7, 2012.

[20] N. Maheswari, A. P. Kabilan, R. Venkatesh, "A Hybrid Model of Neural Network Approach for Speaker Independent Word Recognition", International Journal of Computer Theory and Engineering. Vol. 2, No 6, December, 2010 1793-8201.

[21] Tomi Kinnunen and Pasi Franti, "Speaker Discriminative Weighting Method for VQ based Speaker Identification".

[22] Jui-Feng Yeh, Ming -chi Yen, "Speech Recognition with word Fragment Detection Using Prosody Features for Spontaneous Speech", Applied Mathematics & Information Sciences, 669s-675s (2012).

[23] Ravi Kumar. K. M, Ganesan. S, "Comparison of Multidimensional MFCC Feature Vectors for Objective Assessment of Stuttered Did fluencies," Advance Networking and Applications, Volume:02 Issue:05,Pages:854-860(2011).

[24] W.reichl and G. Ruske, "Syllable Segmentation of Continuous Speech with Artificial Neutral Networks," In processing of Euro speech, Berlin, Vol.3,pp. 1771-1774,1993.

[25] E. Keogh, "Exact indexing of dynamic time warping", In VLDB, pp. 406-417, Hong Kong, China, 2002.

[26] H. Silverman and D. Morgan, "The application of dynamic programming to connected speech Segmentation," IEEE ASSP Mag. 7,no. 3,pp. 7-25,1990

[27] Ravi Kumar, K. M, Ganesan. S, "Comparison of Multidimensional MFCC Feature Vectors for Objective Assessment of Stuttered Disfluencies," Advance Networking and Applications, Volume: 02 Issue: 05, Pages: 854-860(2011).

[28] I. Swietlicka W. Kuniszyk - Jozkowiak and E. Smolka, "Artificial Neutral Networks in the disabled Speech Analysis," in the computer Recognition System 3. Vol. 57 /2009: Springer Berlin /Heidelberg, May 12, 2009, pp 347-354.

[29] 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: ANN recognition of repetition and prolongation with supplied word segment markers," Journal of speech Language and Hearing Research Vol 40.p.1085,1997

[30] Ooi Chia Ai, Hariharan.M. "MFCC based recognition of repetition and prolongation in Stuttered speech using k-NN and LDA", Research and Development (SCOReD), 2009 IEEE Students Conference on 16-18 Nov. 2009.

AUTHOR'S BIOGRAPHY

Dr M Shivakumar did his M.Tech in 1998 and Ph.D. in 2011 from University of Mysore. He has a teaching experience of more than 20 years. At present, he is working as Professor in the department of Electronics and Communication Engineering at GSSS Institute of Engineering and Technology for Women, Mysore. He has published several papers in international conferences and journals. His areas of interest include Speech Signal Processing and Mobile Robotics.

Mrs. G. Manjula has done her h M.Tech in 2007 and pursuing Ph.D. in the area of Signal processing from VTU under the guidance of Dr. M. Shiva Kumar. At present, she is working as Assoc. Professor in the department of Electronics and Communication Engineering at GSSS Institute of Engineering and Technology for Women, Mysore. She has published several papers in international conferences and journals. Her areas of interest include VLSI and Signal Processing.