ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 11 , Issue.3 Sep 2023

Voice Based Prescription Generation System

[1]Minhaj Begum, [2]Nallapu Anushka, [3]Medavarapu Anvitha

Abstract

In India thousands of people dies as a result of wrong medication and normal ailments leads to severe ailments. The scenario is voice recognition. It plays vital role for an effective interaction between human and computer. The scenario of recognition of voice can be used as a trigger to number of virtual application. Although both of above mention scenarios are different they can be unite to address voice based medicine prescription application. This paper proposed a framework that signifies use of voice recognition techniques for medicine prescription. The goal is to avoid wrong medication for normal ailments viz. fever, cough, cold, body pain etc. by designing virtual application on voice based medicine prescription. In this system basic theories for voice recognition such as *pre-emphasis, feature extraction, pattern comparison* is implemented. The MFCC is used as feature extraction technique and DTW as a pattern comparison. Since there is extensive fuzziness in medicine prescription for single disease or symptom, *fuzzy decision logic* is introduced here for appropriate medication. Thus upon spoken symptom the proposed system will prescribe medicine. System performance is analyzed by creating symptom corpus for five people. An accuracy of 90% has been achieved for voice recognition.

Keywords – Automatic Speech Recognition (ASR), Mel Frequency Cepstral Coefficient (MFCC), Dynamic Time Warping (DTW), Fuzzy Logic

1. Introduction

In field of medical science various medicine prescription applications are available. These applications are based on different platforms such as Windows, Android etc. The summary of relevant theories is shown in table no.1 Every application have their unique operation. Some applications are designed for medical administration and some for clinical data management. But very few applications are designed for medicine prescription which also has certain limitations. The motive behind this paper is to design sophisticated virtual application that prescribe medicine for normal ailments and operates on spoken symptom by user. The paper is arranged as fallows. Section 2 describes overall block architecture and system flow. Next two sections discriminates design of ASR through MFCC and DTW.[1][2][3]

[1]Assistant Professor, Department of Information Technology, Bhoj Reddy Engineering College for Women, Hyderabad, India

[2,3] Student, Department of Information Technology, Bhoj Reddy Engineering College for Women, Hyderabad, India

ISSN2321-2152
www.ijmece .com
Vol 11 , Issue.3 Sep 2023

Final section explains use of fuzzy logic for appropriate medication decision. Then results are given for performance of ASR system used.

No.	Applications	Platforms	Features	Cost
1.	Easy Clinic	Windows	Administrative, Informative	$309
2.	Doctors desktop	Windows	Patient registration	$259
3.	Hi-Tech EPR Classic	Windows	Prescription	$450
4.	SalveoRx MD Edition	Windows	Medicine dictionary	Free
5.	HomeoQuest	Windows	Prescription pad	$247
6.	Drug Index	Android	Medicine index	Free
7.	Pill Identifier	Android	Informative	Free
8.	Medicine Rx	Android	Pharmaceutical management	Free
9.	MedRemind	Android	Medicine reminder	Free

Table 1 : Relevance

2. Block Architecture and System Flow

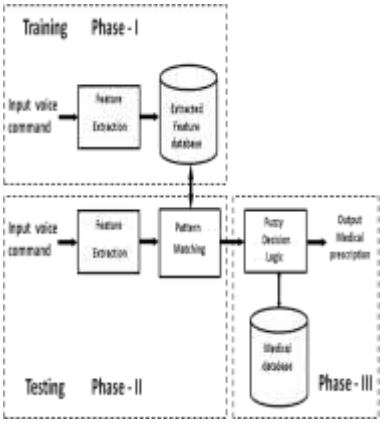

Fig. 1 : Block Architecture

The block architecture for proposed system is as shown in figure 1. It is mainly divided in three phases of implementation. First phase is generally known as training phase. Second phase is testing and third phase is implementation of fuzzy logic. Training phase is accomplished by implementation of feature extraction. Here voiced symptoms to be used are trained using MFCC technique. [4][5][6]
All the extracted features are stored in database for further use. In testing phase feature are extracted for symptom spoken by unknown user. Pattern comparator is nothing but DTW. This block calculates minimum DTW path/score between unknown symptom and stored symptom by using their extracted features. Third phase consist of medical data for reference, and the fuzzy decision logic. The flow of system is shown in figure 2.

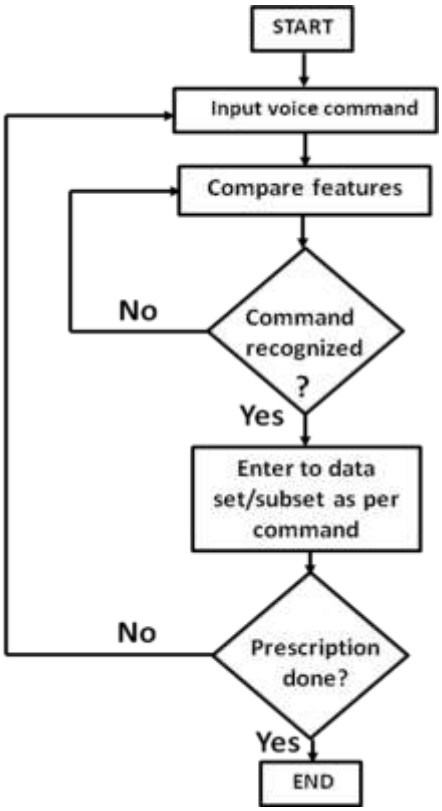

Fig. 2 : System Flow

System starts by taking voiced symptom from user. The MFCC features will be extracted for that input command. These features will be compared by DTW with MFCC features stored in database. Once the input command is recognized its specific value is given to fuzzy decision logic. Depend on recognized symptom fuzzy enters to medical dataset. Thus final medicine prescription is done on simple voice symptoms.

3. Feature Extraction

Feature extraction involves analysis of speech signal. In particular MFCC is the technique used for analysis of mel-frequency component of speech signal. The block diagram for MFCC is shown below in figure 3. This involves pre-emphasis, framing, windowing, DFT calculation, melbank generation and its multiplication with DFT points, log, DCT etc.[7][8][9] Before feature extraction various settings were set on computer. These setting involves, setting microphone volume to high level. Then predefined symptoms were recorded by five persons (2-male, 3-female). Duration for recording was set to 2 seconds. Sampling frequency used were 16000 Hz.

ISSN2321-2152
www.ijmece .com
Vol 11 , Issue.3 Sep 2023

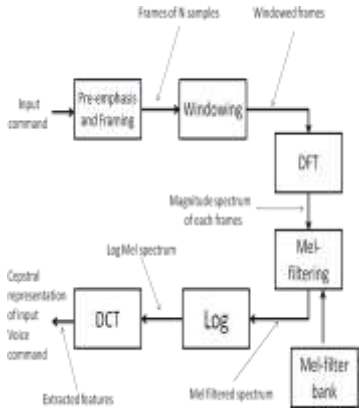

Fig. 3 : MFCC feature extraction

High pass filter is used for pre-emphasis. Framing duration kept was 0.025 seconds i.e. 200 frames/sec. and 80 samples/frame. For windowing Hamming window is used for smoothing end disturbances. 24 melbank filters are used. [10][11][12] In this condition 13 mel frequency Cepstral coefficients were extracted for each recorded symptom.

4. DTW

The key idea behind DTW is to align two non temporal speech sequence and finding minimal DTW score between them. Consider following example shown in figure 4.

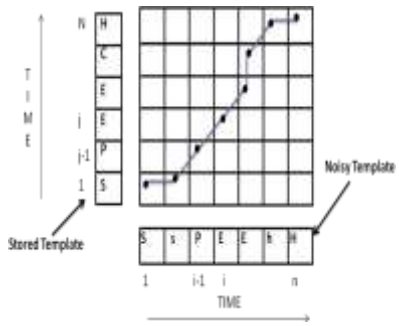

Fig. 4 : DTW aligning

In DTW matrix is computed using MFCC coefficient values. Take all the frames (each frame consist of 13 MFCC coefficient) of stored voice symptom on y-axis and that of testing symptom on x-axis.[8]12][9]15] Calculate local distance between two corresponding frames using Euclidian distance formula [11][14][17] sqrt [$(X_i - Y_j)^2$] Thus cell(n,m) of DTW matrix is result of Euclidian distance between n^{th} frame of test sample and m^{th} frame of reference sample. Then minimum DTW path is constructed through cell(1,1) to cell(r,t) where 'r' and 't' are last frames of reference and test sample respectively. Every next cell is chosen using formula min($cell_{i,j}$ $cell_{i-1,j}$ $cell_{i,j-1}$). Thus by calculating DTW path between test signal and all reference signal, nearest neighborhood approach gives recognized word/symptom.[18]

5. Fuzzy Decision Logic

General approach for designing any fuzzy decision system is as shown in figure 5.

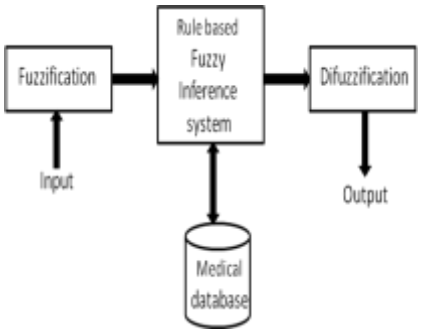

Fig. 5 : Fuzzy Decision Logic

Fuzzification is converting crisp value of recognized symptom to some value depend on degree of membership function of corresponding recognized symptom. Fuzzy inference system is rule based block that maps the output with input using predefined rules. These rules are extracted from clinical data for normal ailments. Fuzzy decision system executes through following steps.[19][20][21]

1. Fuzzyfication of input variable
2. Application of fuzzy operator on antecedent (input variable)
3. Implication from antecedent
4. Aggregation of consequent across the rules
5. Difuzzification

Symptom, *Symptom Details* and *Age* are the three input variable for Fuzzification. Rules were designed upon these input variables. Using combination of these three inputs a particular group of drug is assigned for each group. Gaussian function is used as membership function for input variables *Symptom* and *Age* while triangular function is used for *Symptom Details.* Separate Gaussian function is used for each drug in output.

6. Results

All of the three techniques explained above are implemented in MATLAB v7.10 and performance is verified and enhanced. The ASR duo MFCC-DTW gives accuracy up to 90% and speed up to 2.60 seconds. These results are summarized in table no. 2 Fuzzy decision logic is designed in FIS editor of Fuzzy toolbox in MATLAB. FIS appearance is shown in figure 6 while GUI for user interface is shown in figure 7.

ISSN2321-2152
www.ijmece .com
Vol 11 , Issue.3 Sep 2023

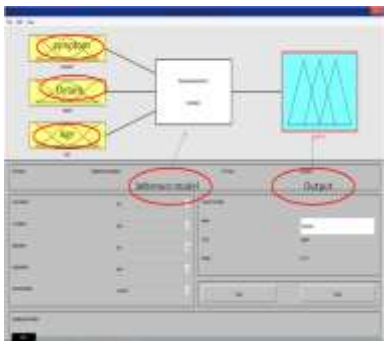

Fig. 6 : FIS

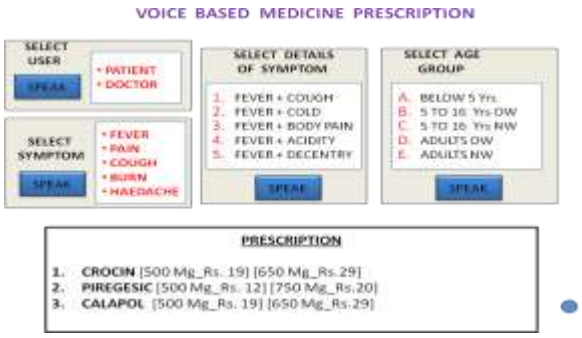

Fig. 7 : GUI

Sr. No.	Commands	Accuracy	Speed
1	Acidity	100 %	2.90
2	Burn	80 %	2.84
3	Cough	100 %	2.91
4	Cold	100 %	2.94
5	Diarrhea	80 %	2.96
6	Fever	100 %	2.95
7	Headache	60 %	2.92
8	Vomiting	80 %	2.96
9	Pain	80 %	2.96

Overall system accuracy is 90% and speed is 2.6

REFERENCES

[1] Sanjib Das, "Speech Recognition Technique: A Review," *International Journal of Engineering Research and Applications*, Vol.2, Issue 3, pp.2071-2083, May-Jun 2012.

[2] Aarati V. Jadhav and Rupali V. Pawar, "Review of Various Approaches towards Speech Recognition," *International Conference on Biomedical Engineering* (ACoBE),pp.99-103,27-28 February 2012.

[3] M. A. Anusaya and S. K. Katti, "Speech Recognition by Machine: A Review," *International Journal of Computer Science and Information Security*, Vol.6, No.3, pp.181-199, 2009.

[4] Vimala. C. and Dr. V. Radha, "A Review on Speech Recognition Callenges and Approaches," *World of Computer Science and Information Technology Journal* (WCSIT), ISSN: 2221-0741, Vol. 2, No. 1, pp.1-7, 2012.

[5] Bharati W. Gawali, Santosh Gaikwad ,Pravin Yannawar, and Suresh C. Mehrotra, "Marathi Isolated Word Recognition System using MFCC and DTW Features," *ACEEE International Journal on Information Technology*, Vol. 01, No. 01, pp.21-24, Mar.2011.

[6] Ingyin khaing, "Mynmar Continuous speech Recognition System Based on DTW and HMM," *International Journal of Innovations in Engineering and Technology* (IJIET), Vol.2, ISSUE. 1, 78-83, February 2013.

[7] Palden Lama and Mounika Namburu, "Speech recognition with Dynamic Time Warping using MATLAB,"CS 525, SPRING 2010-PROJECT REPORT.

[8] Zbyni k Tychtl and Josef Psutka, "Speech Prroduction Based on the Mel-frequency Cepstral Coefficients,".

[9] Chadawan Ittichaichareon,Siwat Suksri, and Thaweesak Yingthawornusuk, " Speech Recognition Using MFCC," *International Conference on Computer Graphics, Simulation and Modeling*(ICGSM'2012),pp.135-138, July 28-29, 2012.

[10] Fang Zheng, Guoliang Zhang and Zhanjiang Song, "COMPARISON OF DIFFERENT IMPLEMENTATIONS OF MFCC," *Journal Computer science & Technology*, 16(6): 582-589, Sept. 2001.

[11] Yan-Sheng Lin and Chang-Peng Ji, "Research on Improved Algorithm of DTW in Speech Recognition," *International Conference on Computer Application and System Modeling* (ICCASM 2010), pp.418-421, 2010 *IEEE*.

[12] Li Yang, Le Jing, Yang Yuxiang, and Wang Jian, "Improvement Algorithm of DTW on Isolated-Word Recognition,"*International Conference* 2011 *IEEE.*

[13] Shivanker Dev dhingra, Geeta Nijhawan, and Poonam Pandit, "ISOLATED SPEECH RECOGNITION USING MFCC AND DTW," *International Journal of advanced Research in Electrical,Electronics and Instrumentation Engineering*,Vol.2, Issue 8, August 2013.

[14] Shicheng Li, and Haipeng Ren, " An isolated word recognition system based on DSp and improved dynamic time warping algorithm,"pp.136-139, *International Conference* 2010 *IEEE.*

[15] Lindasalwa Muda, Mumtaj Begam and I.Elamvazuthi, "Voice Recognition Algorithm using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time warping (DTW)Technique,"JOURNAL OF COMPUTING.VOLUME 2.ISSUE 3,MARCH 2010,*ISSN 2151-9617*.

[16] Zhang Jing and Zhang Min, "Speech Recognition system Based Improved DTW Algorithm," *International Conference on Computer, Mechatronics, Control and Electronics Engineering* (CMCE), 2010 IEEE.

[17] Talal Bin Amin and Iftekhar mahmood, "Speech recognition Using Dyanamic Time Warping,"*2nd International Cionference on Advances in Space*

Technology, pp.74-79,29th -30th November, 2008 IEEE.

[18] Ran Yaniv and David Burshtein, "An Enhanced Dyanamic Time Warping model for improved estimation of DTW Parameters,"*IEEE TRANSACTION ON SPEECH AND AUDIO PROCESSING*, VOL.11, NO.3, MAY 2003.

[19] Lawren Rabiner, B. Yagnanarayna, B. H. juang, "Fundamentals of Speech Recognition"Pearson Publication,2009.

[20] D Driankov, H Hellendoorn, M reinfrank, "An Introduction to Fuzzy Control" Narosa Publication. 2001.

[21] Timoyht J. Ross, "Fuzzy Logic With Engineering Application" 3rd edition Wiley India Publication, 2011.