International Journal of Modern Trends in Engineering and Research
www.ijmter.com
e-ISSN No.:2349-9745, Date: 2-4 July, 2015

Augmentative Communication for people with severe speech and or physical impairment

Madhuri R. Mahajan[1], Yogita S. Pawar[2]
[1]Department of Computer Engineering, GHRIEM, Jalgaon, pawaryogita04@gmail.com
[2]Department of Computer Engineering, GHRIEM, Jalgaon, madhurimahajansap@gmail.com

ABSTRACT- To make communication easy for people with speech impaired and or physically impaired this system is useful for them. The proposed system recognize disordered speech and give output as possible as in natural form. If speech impaired person is physically impaired and he want to send text message then this proposed system is useful to build text message by using menu oriented text format or by converting text to speech.
A stronger model developed to improve the performance of automatic pronunciation evaluations.
One extended feature is to input text with menu list is optimal binary spelling interfaces is used here.
As well as speech to text are added in system for speech and physical impaired people.
Keywords- speech recognition, discriminative adaptive training (DAT); speaker adaptive training (SAT); minimum phone error (MPE); automatic pronunciation evaluation (APE) ,Binary spelling interface, Speech to Text.

I. INTRODUCTION

1.1 Project Problem Definition

The user is impaired with sever or more speech or and physically handicap. The proposed system gets the disordered speech of the user and build message and then converted into synthetic speech. Some people are less speech impaired as well as handicap and if he wants to send text message then speech to text conversion technique is use here. Proposed System also recognizes the language spoken by user. Menu oriented text is added in proposed system for the people who is more speech impaired and to perform fast communication. To improve the pronunciation level applying more canonical model to the APE.

1.2 Need for New System

Previously available system is voice input voice system. In that system pronunciation problem is occurs and as shown in below figure in user interface entering text is very time consuming that point is covered in proposed system as well as additional pointes added in proposed system is Text-To-Speech, Speech-To-Text. So proposed system is useful for speech and or physically impaired person. components are described below.

1.2.1 Speech Recognition- recognize the speech impaired persons speech as accurate as possible. Is relatively feasible for small words or sentences. For speaker-dependent recognition word recognization accuracy goes down as increasing in words or sentences size. when speech input is highly variable, as is the case with speech impaired voice.

1.2.2 Message Building- This model builds message from the recognized input word. If any word is incorrect in message then it is corrected by using keyboard.

1.2.3 Speech Synthesis-To make sound as natural as possible use synthetic output.

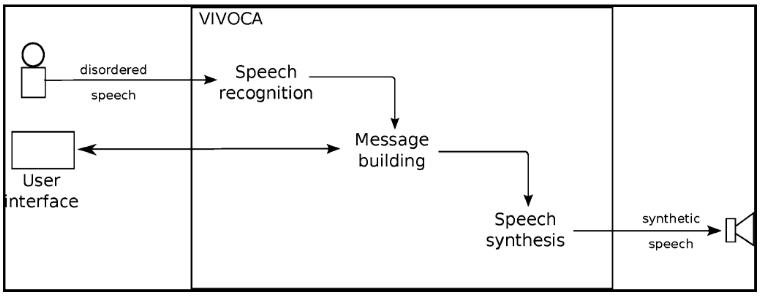

Fig. 1 diagram to convert disordered speech to synthetic speech.

II. ANALYSIS

2.1 Pronunciation

In this proposed work adaptive training is used to eliminate the influence of speaker variations with discriminative training used to easily identify the confused phones. DAT algorithm is a combination of SAT and MPE, referred to here as SAT+MPE. The procedure adaptively trains an ML system as described in Speaker normalization using SAT. Then, the ML-estimated CMLLR linear transforms are fixed and only the model parameters are discriminatively updated.

2.1.1 Accent adaptation using MLLR- Model was trained using native speakers, whose speech data acoustic feature space might differ from those of non-native speakers. Since the aim is to assess the quality of speakers, the non-native English speakers are not expected to pronounce these words the same as the native speakers.. Although there may be some mismatch between the acoustic features of native and good non-native speakers, this mismatch does not affect normal communication. To get a more model score an adapted MLLR was used . The adaptation was based on the pronunciation of the good speakers. The accent adaptation in APE differs from automatic speech recognition (ASR). For APE, accent adaptation is only used to reduce mismatches caused by the different characteristics of the speakers.

III. BINARY SPELLING INTERFACE

Any binary SI can be associated with a corresponding binary tree with exactly one letter (or menu item) in each of its leaves. This implies that the writing process for one character starts from the origin of the tree and proceeds by selecting the left branch or the right branch depending on the single binary signal at each internal node, which lays on the path from the origin to a leaf where this character is located. Letters that are sequentially selected in this manner make up a final text. it is possible that a false leaf will be selected and a false letter. As a means of correction of the written text a so-called *delete-option* or *backspace-option* can be used . This delete-option is a special character, which erases the last character of the text similar to the action of a backspace-key on a computer keyboard

3.1 Optimization Algorithm For Strong Inhomogeneous Binary Spelling Interface

As it was already noted above, when selection probabilities and essentially distinct one from the other ("strong inhomogeneous" SI) and the number of letters is large, the globally optimal solution for SI optimization problem could not be derived by making a locally optimal (greedy) choice. It means that algorithms, which build the full binary tree in a bottom-up manner can not be applied to this case. The globally optimal solution could not be arrived at by making a greedy choice just because the two least-frequent objects can belong to different parent nodes at different levels of the optimal binary tree. In this case the deepest leaves of the tree do not correspond to minimal values of probabilities to reach these leaves without any error.

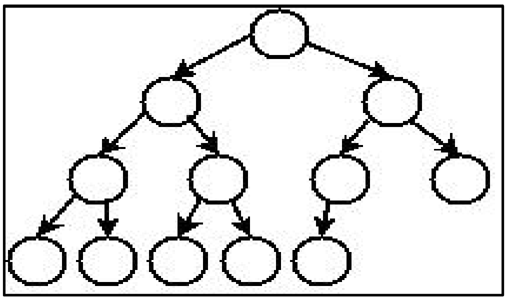

Fig. 2 Full binary tree with 6 leaves.

For our purposes we need to compute values of the optimization criteria and comparing these values in order to choose the minimum. To calculate these values in the effective way, we need a decoding algorithm, which can compute pairs of left and right steps(xi,yi) for all leaves from the given i€[0,n+1] P-sequence (p1,p2,…pn). Such an algorithm can be written as shown below. Note that a stack , which implements a LIFO (last-in, first-out) policy for a set of leaves' "power" coordinates {(x,y)} is used. Here, represents the number of left steps and y—the number of right steps required to achieve a particular letter:

P0 <- 0 Pn+1 <- pn X <- 0
Y <- 0 Push(S,(x,y))
For I <- 1 to n+1 do (x,y) <- Pop(s)
For j <- 0 to pi-pi-1-1 do xR <- x
yR <- y+1 Push(S,(xR,yR)) X <- x+1 Enddo
Xi <- x
Yi <-y
Enddo.

This algorithm can be illustrated with the P-sequence (2, 4, 4, 4, 5) for the tree from Fig. 1. We will follow the algorithm step by step and describe the values of all variables and the stand of the stack , which will be filled (with the standard Pushprocedure) and popped (with the standard Pop-procedure) from the left to the right.

for i=1 to n do pi = i
enddo
while I = max{k|pk < n} exists do pi <- pi+1
for j = i+1 to n do pj <- max{pi,j} enddo
enddo.

IV. SPEECH TO TEXT CONVERSION

4.1 Bakground

4.1.1 Speech to Text for people with Special Needs- Pure audio recordings can be effective for nonnative English speakers and user with learning disabilities or handicap who may simply need more time to comprehend speech or not able to write. Thus, SR-mVA is providing text is important for people who is not recognize voice. user with hearing impairments or handicap cannot process the audio speech. The SR--mVA approaches caption user's speech into text. Handicap person is not able to edit text message so this system is useful for such people.

4.1.2 OBJECTIVE- This proposed system automatically convert user's speech to text. Two different methods of SR--mVA were compared during different course subjects as a technical feasibility and case study. Our specific objectives were to: identify the issues regarding the use of SR--mVA as a standard tool in capturing spoken information,. Check the word reorganization accuracy by comparing RTC and PLA.

4.1.3 Utilities and Tools- The SR--mVA is beneficial to record sound, generate speech transcripts.

4.1.3.1 Recording Speech Audio- The users' voice is recorded using an Audio-Technica 700 Series Freeway 8-channels UHF wireless microphone system. Audacity provided various configuration options for voice recording in SRcompatible format settings.

4.1.3.2 Display TEXT- Synote(www.synote.org/synote), a web-based application, created text and tags to synchronize audio or video recordings, SR transcripts. Registered users can save their own remarks and notifications for each set of Speech.

4.1.3.3 Error Correction- A Word Error Rate (WER) evaluation tool developed by IBM is a command-line utility that compared a raw "decoded" SR text transcript to an edited text transcript. WER is percentage of total errors for the total number of words n speech. Accuracy is the percentage of correct to total words transcribed.

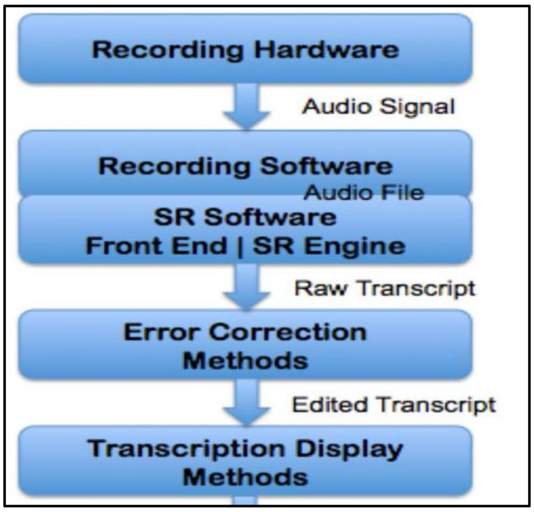

Fig. 3 General SR- mVA methodology.

CONCLUSION

This Proposed work described the development of portable, voice output communication aid controllable by automatic speech recognition. The device can be configured to enable the user to create either simple or complex messages using a combination of a relatively small set of input, with menu oriented list of word and produce intelligible speech output.

We try to solve device and pronunciation related problem by using different model.

If in case speech impaired person is also impaired physically and user want to send text message then speech to text conversion is done here. As well as user stored this message in his database.

REFERENCES

[1] Mikhail Tregoubov and Niels Birbaumer, "On the Building of Binary Spelling Interfaces for Augmentative Communication", VOL.52, NO. 2, FEBRUARY 2005.

[2] SONG Yin, LIANG Weiqian, "Experimental Study of Discriminative Adaptive Training and MLLR for Automatic Pronunciation", Evaluation Institute of Microelectronics, singhua University, Beijing 100084, China; Department of Electronic Engineering, Tsinghua University,Beijing 100084, China

[3] Rohit Ranchal, Member, Yiren Guo, Keith Bain, Heather Martin, J. Paul Robinson, Member, IEEE, and Bradley S. Duerstock, Using Speech Recognition for Real-Time Captioning and Lecture Transcription in the Classroom,IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, OCTOBER-DECEMBER 2013.

[4] Mark S. Hawley, Stuart P. Cunningham, Phil D. Green, Pam Enderby, Rebecca Palmer, Siddharth Sehgal, and Peter OâAZNeill,A Voice-Input Voice-Output Communication Aid for People With Severe Speech Impairment, IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 21, NO. 1, JANUARY 2013.

