Design and Development of Telugu Diphone Database

D. Nagaraju, RJ Ramasree, M. Sri Poojitha, K. Nikhitha, M. Swetha

Abstract: Diphone database plays a vital role in text to speech system. In this paper, we present the design and development of Telugu diphone database which supports the MBROLA. In the previous article of the same author generation of diphone database for Telugu presents generation of Telugu diphone database [01]. This is the second Telugu diphone database developed by the same author. This database is the upgraded version of first database and few problems are rectified in this version.

Key words: diphone database – Telugu - MBROLA

I INTRODUCTION

Diphone database is the database of sounds. In designing and develop a database of Telugu diphone database there are several steps to follow, first selection of sample words or wav files, which contains diphones, second phone selection, third is diphone selection and extraction of diphone, finally segmenting or equalizing the diphones to make diphone database. In previous article few diphones are not working properly, because their energy levels are not smoothened such few diphones are ui, ai, ea etc. Number of diphones is less in previous diphone database. Above two are major drawbacks in previous database. Theses drawbacks are the motion to make new diphone database in Telugu language.

II RELATED STUDY

A diphone database for Arabic language was designed and developed. Arabic language has 28 consonants, 3 long vowels and 3 short vowels. 102 sentences and 654 words are used to extract 35 phones including pause and 1190 diphones [02]. Romanian language has 1156 possible diphones including 30 allophones [03]. MBROLA supported Sinhala diphone database is designed with 1004 diphones, without considering the allophones [04]. Polish diphone database is created with 1444 diphones and polish extended with 1681 diphones [05]. A Korean diphone database covers 1853 out of the 3977; it covers only 47% of all possible diphones [06]. Lithuanian has 58 phones and two diphone databases are builds with possible 5003 diphones [07]. Kurdish language has 37 phonemes, thus it has 37X37=1369 diphones. Diphone database designed with 1300 diphones [08]. Bengali has 44 alphabets and the database designed with 1936 diphones [09].

III TELUGU PHONE LIST AND AUDIO COLLECTION

Telugu language is one of the familiar languages in south India and it is the primary language of the states Andhra Pradesh and Telangana. Telugu people are living across the India and nearly 50 countries in the world. Telugu have 16 vowels and 36 consonants total 52 alphabets, among these there alphabets (ఏప్పుడు ఇప్పుడు) are obsolete now [10]. The remaining 49 Telugu phones are considered to build diphone database. Telugu phone set is ఎందా ఏం హేచ్ ఫ్యాం డా నా బా భా మా యా రా లా వా సా శా సా లా క్షా and its transcriptions are a aa i ii u uu e ea o oe ou aM a@h ka kha ga gha ~ma ca cha ja Ja ~n Ta Tha Da Dha Na ta tha da dha na pa pha bha ma ya ra la va Sa sha sa ha La ksha.

Telugu speech corpus taken from IIT Chennai Telugu mono database wav files [11], these wav files are used to extract diphones. It have 2840 sentence in wav form, among these we used 940 sentences for extraction of diphones.

IV DIPHONE SELECTION AND EXTRACTION

Diphone is an adjacent pair of phones. Diphones are useful in speech synthesis. Pre-recorded diphones are combined to create synthesized speech; the resulting sounds are much more natural than combining just simple phones, because the pronunciation of each phone varies based on surrounding phones. The number of phones in a language is P, and then the number of possible diphones is P². In previous article 106 sentences/622 words are used to extract 1356 diphones and generated the Telugu diphone database. The projected Telugu diphone database is designed with 49 phones and 1806 diphones, these diphones are extracted from 940 sentences. Praat tool is used to extract the diphones from wav files. The General frame work for generation of diphone database is as shown in figure 1.

V EQUALIZATION OF DIPHONES

The energy levels at the beginning and at the end of a segment are modified in order to eliminate amplitude mismatches the energy of all the phones of a given phonemes are set to phones’ average value [12].

Revised Manuscript Received on November 06, 2019.

D. Nagaraju, Associate professor in Audisankara college of engineering and technology, gadur, Nellore, Andhra Pradesh, India -524101

Dr. RJ Ramasree, Professor, dept of computer science, Rashtriya Sanskrit vidyaapeetha, tirupati, Andhra Pradesh, India

M. Sri Poojitha, K. Nikhitha, M. Swetha, Students in Audisankara college of engineering and technology, gadur, Nellore, Andhra Pradesh, India
Design and Development of Telugu Diphone Database

Fig 1. General frame work for generation of diphone database

The extracted diphones are then put forward to the evaluation stage in which both database files and the diphone files are evaluated by manually. The annotation diphone files allow manually comparison of the annotations with signals in the diphone files. Mbrolator generates diphone database by taking Database SEG file and all available diphones in the given language.

VI STEP BY STEP PROCEDURE TOWARDS DIPHONE DATABASE GENERATION

Step 1. Wav file reading by using Praat tool

Step 2. Click on view and edit command

Step 3. Create annotation wav file

Step 4. View & Edit both sound and text grid files
Step 5. Phone selection

Step 6. Diphone selection

Step 7. Diphone extraction and save selected file as wav

VII CONCLUSION

In August 2013, the first Telugu diphone database was published. Having properly generated the parameters of the duration and fundamental frequency curves, it is possible to synthesize a sound of a sufficiently high quality. This can be testified using the examples presented with the database. Therefore, the database alongside the MBROLA software can be used as a module for signal generation in the Telugu speech synthesizer of a high quality.

Previous database have 1356 diphones, it is upgraded up to 1806 diphones and properly not working diphones are upgraded, thus they are working properly.

REFERENCES


4. N. G. J. Dias: MBROLA Formatted Diphone Database for Sinhala Language, University of Kelaniya 2015

5. Jolanta Bachan: Efficient Diphone Database Creation for MBROLA, IEEE 2010


7. Pijus Kasparatis: Diphone Databases for Lithuanian Text-to-Speech synthesis, Informatics 2013


