RESEARCH ARTICLE

Design and Development of Speech Database for Travel Purpose in Marathi

Pooja V. Janse [1], Ratnadeep R. Deshmukh [2], Smita B. Magre [3]

[1] Dept. of CS & IT, Dr. B. A. M. University, Aurangabad-431004, India
[2] Dept. of CS & IT, Dr. B. A. M. University, Aurangabad-431004, India
[3] Dept. of CS & IT, Dr. B. A. M. University, Aurangabad-431004, India

[1] puja.janse@hotmail.com; [2] rrdeshmukh.csit@bamu.ac.in; [3] smit.magre@gmail.com

Abstract— The paper represents the brief information about developing speech database in Marathi language for Travel purpose in Aurangabad District. Development of speech database is very primary requirement for developing an Automatic Speech Recognition System. The accuracy of speech recognition depends on the quality of the speech data recorded and the algorithms implemented for the development of ASR. The data collection procedure from various speakers from Aurangabad district is described in the paper for developing ASR system in Marathi language for travel domain.

Keywords— Speech recognition, Marathi Language, Travel domain, Isolated words

I. INTRODUCTION

Speech is the way of Communication between human being. Speech has the capability to be used as an interface for computer system [1]. Estimated number of languages in the world varies between 6,000 and 7,000. Human being has long been motivated to develop the computer that can understand and talk like human. Since 1960, computer researcher has trying ways and means to make computer record, interpret and understand human speech.

The computer system which can understand the spoken language are very useful in various domain like education sector, domestic sector, military sector, medical sector, agriculture sector, artificial intelligence sector etc. So to perform any type of research, researcher requires some previous data. Generally databases are fundamental for research.

This paper will describe about Aurangabad District, need of development of speech database for travel purpose, Marathi Language, methodology used for collecting text corpora, steps for data collection, recording procedure and detail in various section. The conclusion and future work is stated.

A. About Aurangabad city

Aurangabad is an important city in the state of Maharashtra, India. It sees a lot of international tourist traffic due to world heritage sites, Ajanta & Ellora caves. Aurangabad, a dream destination of Mughal Emperor Aurangzeb, founded in 1610, is known today, as one of the four important cities in Maharashtra. The city is heading towards rapid industrial growth, but retains its past glory and charms, heritage and traditions.

The entire city of Aurangabad was fortified and huge entrance gates were erected. Except for Bharkal gate all the other gates are associated with the period of Aurangzeb. The gates with prime importance were the ones facing the four directions, Delhi Gate facing the North, Paithan gate facing South, Makai Gate (Mecca Gate) facing East and the Khas Gate facing the West. In all there were 54 gates in Aurangabad. Out of 54 gates only 13 have survived over the period of time. The important and architecturally most appealing gates include Delhi Gate, kala Darwaza, Makia Gate, Paitahn gate and Rangeen Darwaza.

There are various tourist places in Aurangabad. The Ajanta Caves depict the stories of Buddhism spanning from the period from 200 B.C. and 650 A.D. These 29 caves were built by Buddhist monks using simple tools like hammer & chisel.

The Ellora (Verul) Caves, Ghrishneshwar temple, Bibi ka Maqbara, Panchakki, Aurangabad   Caves, Soneri Mahal, Gautala Wildlife Sanctuary, Pitalkhora, Siddharth Garden are attractions in Aurangabad city. So lots of tourist visits to Aurangabad.

B. Research Problem

Automation of travel related application over the web has been a successful business model for many internet companies. This is a step towards replacing human travel agents with intelligent machines. The next step towards a more human like experience in automated travel reservations is to use automatic speech recognition and natural language understanding to interact with a machine and use the travel related application using a more natural communication medium, namely speech.

The problems posed by spontaneous speech can be divided into four categories
- User noise - breath noise, filled pauses and other user generated noise
- Environment noise - door slams, phone rings, etc.
- Out-of-vocabulary words - The subject says words that the system doesn't know.
- Grammatical coverage - Subjects often uses grammatically ill-formed utterances and restarts and repeat phrases.

C. Need of Development of speech database

Major population in Aurangabad speaks Marathi- the regional language of Maharashtra. Hindi, Urdu and English are also spoken widely. Generally people who come in close contacts with people do understand and speak fairly good Marathi.

As little work is done for travel domain in Marathi, it leads to develop ASR system in Marathi.

II. MARATHI LANGUAGE

Marathi first appeared in writing during the 11th century in the form of inscription on stone and copper plates. From the 13th century until mid of 20th century, it was written with the Modi alphabet. Since 1950 it has been written with Devanagari alphabet.

Marathi language is Indo-Aryan language. It is official language of Maharashtra state of India and it is one of 23 official languages spoken by about 71 million people mainly in Indian state of Maharashtra state and neighboring state [2].

Many government and semi-government organizations exist which work for the regulation, promotion and enrichment of the Marathi language. These are either initiated or funded by Government of Maharashtra. The major dialects of Marathi are called standard Marathi and warhdi Marathi. Standard Marathi language is based on based on dialects used by academics and print media. The work for Marathi language is majorly being carried out in IIT Bombay and TIFR (Mumbai). Few universities have also started working for development of system in Marathi Language in Maharashtra state [3, 4].

III. SELECTING TEXT CORPORA

For developing a speech database the basic requirement is of correct Text corpus which would be recorded from various speakers [5]. The text corpus designed should be grammatically correct. The isolated words are selected from various categories.

The text corpus was generated using information about Aurangabad District. Total 124 words are selected and they are grouped according to their category i.e. Malls, Cinema halls, Markets, Temples, Playgrounds, Station and Airport, Cultural halls, Hotels, Tourist Places and Restaurants.

IV. DATA COLLECTION

In this stage, the steps followed for developing speech corpora are described. The recording media is chosen first and then the data has been recorded using high quality microphones and laptop using PRAAT for recording speech signal.

A. Speaker Selection

The speech data will be collected from the native speakers of Marathi Language. The selected speakers will be from different regions of Aurangabad District. They would be comfortable with reading and speaking the Marathi Language. The speakers are classified on the basis of gender [6].

B. Speech Data collection

The speaker will be asked to speak 124 words with 3 utterances of every word. The speech data will be collected by visiting 9 Taluca places of Aurangabad District. At every Taluca place speech data from 10 speakers will be collected. Among this 10 speaker 5 will be male and 5 will be female. The Speech Data was collected in a noisy environment with the sampling frequency 22050 Hz.

C. Data Collection Statistics

The speech data is collected from 100 speakers. Each speaker was asked to speak 124 words with 3 utterances. 372 utterances of words will be collected from every speaker. Total 37200 utterances of words are recorded. Till date we have collected 37200 utterances from 100 speakers in which 50 male and 50 female speakers. Table shows process details which are followed during data collection work.

TABLE I: Process details followed in work

Process	Description
Total Number of Words Selected	124
Utterances Recorded	Three utterance of each word
Total Utterance per Speaker	372
Total Speaker	100
• Male Speaker	50
• Female Speaker	50
Total Male Speaker Utterances	18600
Total Female Speaker Utterances	18600
Total Utterances	37200
Total Size of Database	3.34 GB
• Male Database Size	1.64 GB
• Female Database Size	1.69 GB
Software Used for Recording	PRAAT
Tools	Microphone Sennheiser PC360 and Sennheiser PC350
Recording Frequency	22050 Hz

V. RECORDING ENVIRONMENT

The speech data will be recorded using high quality microphones like Sennheiser PC 350 and Sennheiser PC 360 with the help of open source PRAAT speech software. The data will be recorded in Noisy environment. The purpose of recording in noisy environment is to develop robust ASR System.

The main strength of PRAAT is its graphical user interface. PRAAT also provides the functionality of General analysis (waveform, intensity, spectrogram, pitch, duration) Spectral analysis, pitch analysis, voice analysis, format analysis, intensity analysis, PCA and many facilities [7].
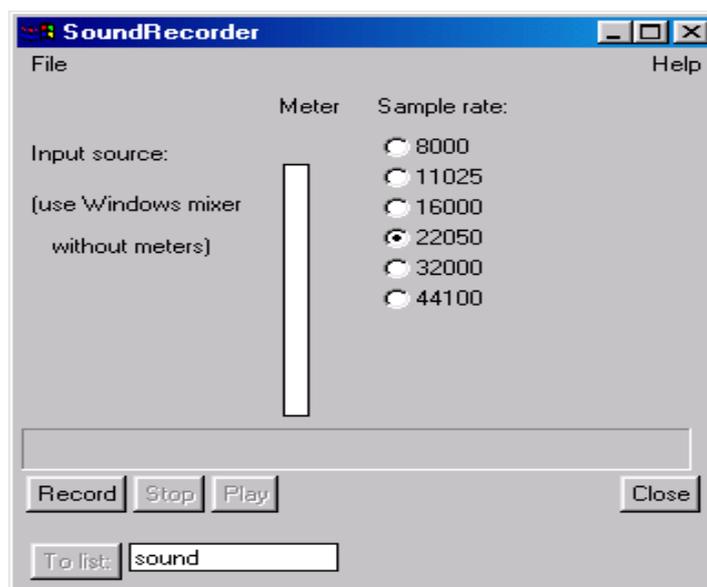
Following figure 1 shows Sound Recorder window of PRAAT.

Fig 1: Sound Recorder Window of PRAAT

VI. CONCLUSION AND FUTURE WORK

We have faced many problems while data collection as getting information related to all tourist places, Hotels, Restaurants, Malls, Markets and other Tourist information, to check the grammar and syntactic error in the collected text corpus, to convince people to collect speech sample, pronunciation of correct word while recording the speech data.

We have represented the procedure that is been followed by us to develop speech database of Isolated Marathi Words for travel domain from Aurangabad District of Maharashtra. It will be used for the development of robust Automatic Speech Recognition System. The developed system can also be used as guide for tourists from Maharashtra who visits to Aurangabad.

ACKNOWLEDGMENTS

REFERENCES

[1] P. P. Shrishrimal, R. R. Deshmukh, Vishal B. Waghmare, "Development Of Isolated Words Speech Database Of Marathi Words For Agriculture Purpose", Asian Journal of Computer Science And Information Technology 2: 7 (2012) 217 – 218.

[2] Pukhraj P. Shrishrimal, Ratnadeep R. Deshmukh, Vishal B. Waghmare, "Indian Language Speech Database: A Review", International Journal of Computer Applications (0975 – 888) Volume 47– No.5, June 2012.

[3] S. S. Agrawal, K.K. Arora, S Arora, Samudravijaya K., "Recent Advances of Speech Databases development activity for Indian Language", Proc. ISCS LP 2006, Singapore, pp. 771-776.

[4] Samudravijaya K., P. V. S. Rao and S. S. Agarwal. 2000, "Hindi Speech Database", Sixth International Conference on Spoken Language Processing (ICSLP 2000), Beijing, China.

[5] Agrawal S. S., K.K. Arora, S Arora, Samudravijaya K, "Text and Speech Corpus Development in Indian Languages", ibid, pp. 94-97

[6] Gopalakrishna Anumanchipalli, Rahul Chitturi, Sachin Joshi, Rohit Kumar, Satinder Pal Singh, R.N.V. Sitaram, S P Kishore, "Development of Indian Language Speech Databases for Large Vocabulary Speech Recognition Systems".

[7] http://www.fon.hum.uva.nl/praat/ cited on 09/03/2012.