

# Development of Standard Speech Database in Marathi Language for Emotion Recognition System

Shinde A.R.<sup>1\*</sup>, Agnihotri P. P.<sup>2</sup> and Khanale P.B.<sup>3</sup>

<sup>1</sup> School of computational sciences, Solapur University, Solapur, India

<sup>2</sup> School of Technology S.R.T.M.U.N. Sub-Campus Latur, India

<sup>3</sup> Dnyanopasak College, Parbhani, India

[www.ijcseonline.org](http://www.ijcseonline.org)

Received: Dec/01/2014

Revised: Dec/13/2014

Accepted: Dec/23/2014

Published: Dec/31/2014

**Abstract** - The biometrics a trading technology to recognize a person based on the physiological or behavioral characteristics. Emotion Recognition system is a part of Human Computer Interaction. The database has been developed to recognize the emotions of human based on speech trait. While recording speech samples different Marathi words and phrases were considered. These words were very useful to identify the emotions of human being. The emotions such as happy, sad, surprise, angry, disgust, fear and neutral were recorded. In the proposed work seven basic emotions were considered such as happy, sad, angry, surprise, fear, disgust, and neutral. According to this, the words like “*Are deva*”, “*Aare re*”, “*kiti chan*”, “*Mast*”, “*Gapre*”, “*khatarnak*” etc. were used to decide what kind of emotion that human being gave through speech. There were total of 50 speakers out of which 25 male and 25 female speakers were recorded. In this paper for each emotion four Marathi emotional words were recorded. In this way there were total 28 words by One speaker and total 1400 words (emotional words/phrases) recorded by 50 speakers including male and female speaker were recorded. This paper discusses about the development of Speech database developed at School of Computational Sciences, Solapur University, Solapur.

**Keywords**—Emotion Recognition, Marathi speech, HCI, Biometric

## I. Introduction

Emotion Recognition is a very recent research topic as a part of Human Computer Interaction (HCI) research domain. Humans can communicate with each other by using natural language. Researchers expect that computer should also be able to interact in the same way as human being [1] [10]. There is need to develop speech database which can effectively express human emotions. There are three methods of data acquisition in emotion research. The first is natural expression, where data is collected from a real-world situation where users are not under obvious observation and are free to express emotions naturally. The second is induced emotional expression, where naïve users are presented with scenarios that induce the required emotional response. The last method of speech acquisition using simulated or portrayed emotional expression makes use of professional or not-professional actors and actresses. For this subjects are instructed to produce emotional expressions for various emotion classes, with varying degrees of intensity [5]. A speech is the most natural form of human communication. Speech is one of the most information-laid signals also speech sounds have a rich and multi-layered temporal-spectral variation that convey words, intention, expression, intonation, accent, speaker identity, gender, age, style of speaking, state of health of the speaker and emotion. There are two broad approaches through which we can able to recognize human emotions. The first is the Facial expression and other is by using speech database (words/phrases).

Speech signal is the fastest and most natural method of communication between human beings. Due to this researchers are attracted to work in speech Emotion Recognition system [1]. Emotion Recognition is very challenging task. Because our aim is recognize human emotions by using computer. An important issue to be considered in the evaluation of an emotional speech recognizes the degree of naturalness of the database used to assess its performance [2].

In this paper the speech database is developed by using speech acquisition device the features like pitch, intensity and some global characteristics were extracted to further recognize human emotions.

## II. Existing Database

In field of emotion recognition research work, some researcher prefers Berlin German database and SJTU Chinese database [6]. Prominent example of acted DB is EMO (Berlin Emotional Speech), the DES (Danish Emotional speech), Polzin in English and Groningen in Dutch [7]. Also Persian database developed by the researcher for emotion recognition. Persian database consists of 2400 wave clips modulated with anger, disgust, fear, sadness, happiness and normal emotions [9]. The best examples for acted databases are the Berlin database of emotional speech and the Danish Emotional speech corpus [4]. In this paper author's aim is to enable a very natural

interaction with the computer by speaking instead of using traditional input devices. The researcher has created a standard database in Persian language; by using Persian speech the features like pitch, intensity and some global characteristics are extracted to recognize human emotions.

### III. Device and Software Used

We have used Linovo P950 gaming headset with microphone having features like maximum power input is 30 mw, sensitivity is -38dB/mw, impedance 2200 ohms and frequency response 100-10000 Hz. Also this Headset which itself provide the facility of noise cancellation. While recording the seven basic emotions are considered like happy, sad, surprise, angry, disgust, fear, and neutral. In this we have selected some Marathi words like “*Aare Deva*”, “*Kitti Chan*”, “*Are Bapre*” etc. For each emotion total four words are selected. Praat speech analysis software is used for recording these Marathi words by keeping standard frequency 44100 Hz. We have recorded total 700 words/phrases in Marathi language by 25 female speakers (25 X 28 emotional words for each speaker) and 700 ((25 X 28) by male speakers for seven basic emotions.

### IV. About Database

In this paper, we have developed a speech database in Marathi language. This is the regional language (Mother tongue) of Maharashtra state. To express different emotions by using Marathi words (Verbal communication i.e. speech), we have considered most common Marathi words which is strongly used to express the exact emotion of human being. For development of these database total 50 speakers we have considered (age between 20 to 45 years). Total 25 Male speaker and 25 Female speakers has recorded their Marathi speech samples in closed computer Lab to reduce noise. For noise free recording

In the proposed work, we have developed speech database in Marathi language. In Marathi language there are various words and phrases are available which can be useful for recognizing human emotions. The emotions such as happy, angry, sad, surprise, fear, disgust and neutral are the basic emotions. The words in Marathi such as “*Aare Deva*” (Oh God!), “*Kiti Chan*” (How Good!) are used to express human emotions. While developing this database, we have considered four words/phrases of each emotion as shown in table 1.

Table 1: Four Marathi Emotional Words/Phrases (Representing **Happy** Emotion)

Sr. No.	Words/phrases in Marathi	Meaning in English	Written in Devanagari Word-net
1	Aare Wa	Great !	†, ēā १०० I
2	Kitti Chan	How	× āṛōḥ āṛō I

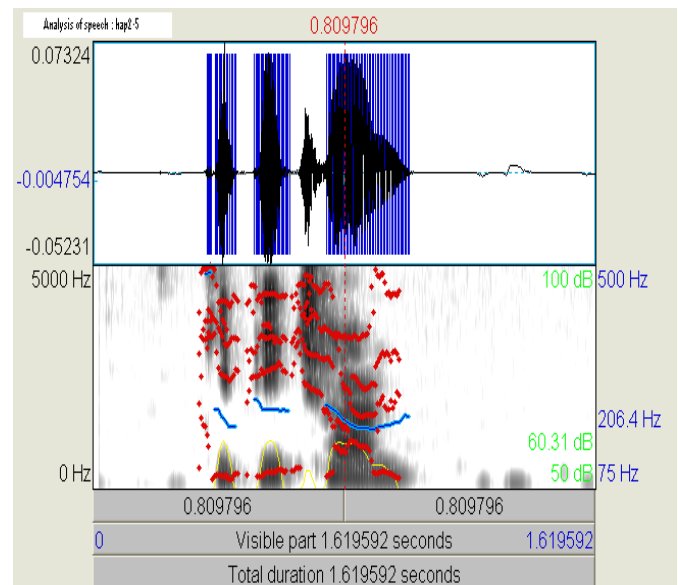
		good I	
3	Kitti God	How nice !	× āṛōḥ āṛō
4	Mast	Superb !	†āṛō I

Table 1: Four Marathi Emotional Words/Phrases (Representing **Angry** Emotion)

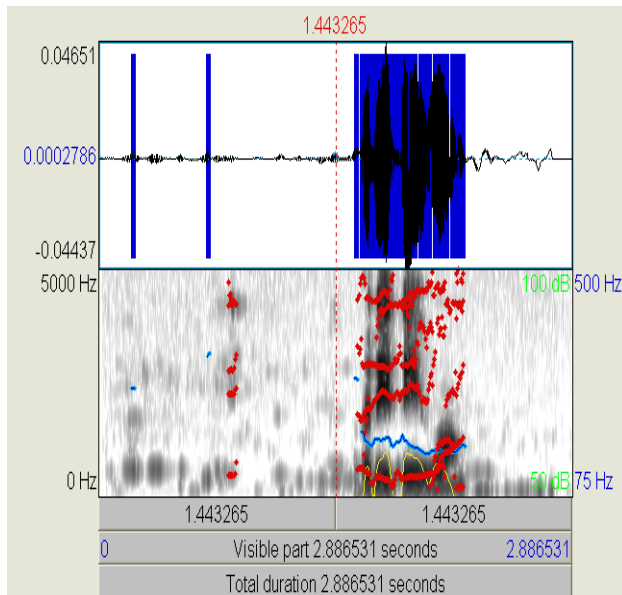
Sr. No.	Words/phrases in Marathi	Meaning in English	Written in Devanagari Word-net
1	Gap Re	Don't talk	ṠṠ, ēā I
2	Hat	Shut up	Āṛō
3	Nakoch mala	Don't need!	Ṡ āṛōḥ ṠṠ
4	Chal Nigh	Gate out	ṠṠ × ṠṠ

### V. Methodology

To recognize the Human emotion by using speech, the following general structure is used. To analysis the Recorded speech we have used **Praat** software. By using this software we can extract some features of speech such as pitch, intensity, formant, frequency etc. Also we can use **MATLAB** to extract these features and according to this we can recognize the human emotion.



Speech tone for happy emotion  
Fig.1 speech Analysis: For Happy Emotion



Speech tone for sad emotion  
Fig.2 Speech Analysis: For Sad Emotion

## VI. Conclusion and Future Work

In this proposed work, we have developed a standard database in Marathi Regional language from Maharashtra state of India. This will very useful in the speech emotion recognition system. By using this database we can extract various speech features. The features like intensity, pitch, formant, frequency etc. are responsible parameters to recognize emotion of human being. By analyzing these features we may easily identify the particular emotion of that human being. Also in future we can add different languages like Hindi (National language of India), English (International language), and other local languages such as Kannada, Telugu, etc (Regional languages). Definitely it will support for all researchers whose interest is in the field of Emotion Recognition System.

## Acknowledgment

We are thankful to University Department of Computer Science, School of Computational Sciences of Solapur University, Solapur and Department of Computer Science, College, Parbhani. We express our sincere thanks towards research group of Japan to make available their image database JAFFA for our research and development.

## REFERENCES

- [1] Mina Hamidi, Maharram Mansoorizade, "Emotion Recognition from Persian Speech with neural network", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.5, Sept.2012.
- [2] Bhoomika Panda, Debananda Padhi, Kashamamayee Dash, Prof. Sanghamitra Mohanty, "Use of SVM Classifier & MFCC in Speech Emotion Recognition System", International Journal of Advanced Research in

Computer Science and Software Engineering, Vol.2, Issue 3, March 2012, ISSN: 2277-128X.

- [3] Mina Hamidi, Maharram Mansoorizade, "Emotion Recognition from Persian Speech with neural network", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.5, Sept.2012.
- [4] Thirid Vogt, Elisabeth Andr'e, and Johanne Wagner, "Automatic Recognition of Emotions from Speech: A Review of the Literature and Recommendations for Practical Realisation", Affect and Emotion in HCI, LNCS 4868, Pp. 75-91, 2008. Springer-Verlag Berlin Heidelberg 2008
- [5] Donn Morrison, Ruili Wang, Liyanage C. De Silva, "Ensemble methods for spoken emotion recognition in call-centres", Speech Communication 49 (2007) 98-112.
- [6] Yixiong Pan, Peipei Shen and Liping Shen, "Speech Emotion Recognition Using Support Vector Machine", International Journal of Smart Home Vol. 6, No. 2, April, 2012.
- [7] Giovanni Costantini, Iacopo Iadarola, Andrea Paoloni, Massimiliano Todisco, "EMOVO Corpus: an Italian Emotional Speech Database"
- [8] Julien Epps, Roddy Cowie, Shrikanth Narayanan, Björn Schuller and Jianhua Tao, "Emotion and mental state recognition from speech", Epps et al. EURASIP Journal on Advances in Signal Processing 2012, 2012:15.
- [9] Mina Hamidi and Maharram Mansoorizade, "emotion recognition from persian speech with neural network", "International Journal of Artificial Intelligence & Applications (IJAIA)", Vol.3, No.5, September 2012.
- [10] Sharad B. Jadhav, Shriram D. Raut, Rajivkumar S. Mente, Ashok R. Shinde, "A Research on Contactless Palm print Recognition System", International Journal of Research in Computer Science and Information Technology, Vol 02, issue 2, Pp. 45-48, ISSN: 2319-5010, March 2014.

**Ashok R. Shinde** has completed M.Sc.in Computer Science in 1997 and qualified UGC-SET in Computer Science and Applications in 2006. He is perusing his Ph.D. in Computer Science. He is currently working as Assistant Professor in Solapur University, Solapur. He is author of more than 10 research articles published in national and international journals and conference proceedings. His area of interest is Data Mining and Warehousing, Network Security, Digital Image Processing etc. He has 5 years of research experience and 14 years of teaching experience at UG and PG level.

