Development of Emotion Recognition From A Speech In Various Regional Indian Languages: A Review

Vishal Tank[1], Dr. S K Hadia[2]

[1] V T Patel Department of Electronics & Communication Engineering, Chandubhai S Patel Institute of Technology, Charotar University of Science & Technology (CHARUSAT), Changa, Anand, Gujarat, India

[2] V T Patel Department of Electronics & Communication Engineering, Chandubhai S Patel Institute of Technology, Charotar University of Science & Technology (CHARUSAT), Changa, Anand, Gujarat, India

(E-mail: vishaltank.ec@charusat.ac.in)

***Abstract*—** Last few years have given witness of development in the field of emotion recognition from speech. Significance contribution has been provided by the researchers to detect emotions from a speech in various languages across world. India is country where second highest numbers of languages are spoken; in addition to that, same language is spoken with different dialects (flavors) with specified area. By considering this fact review of existing and latest work in emotion recognition and its development in Indian languages is useful for boosting further research. In this review paper enlargement of emotions detection, expansion of various speech databases, kinds of classifiers/ parameters/models used for detection in various Indian languages has been discussed. At last important issues has been presented. The purpose of this paper is to ignite the work in Indian Local languages.

***Keywords*—** *Emotion recognition, Regional Indian Languages, Simulated Indian emotional speech corpus, Prosodic features, Source features, Classifier*

I. INTRODUCTION

There are several ways of communication and speech is the best mode of communication for human beings. Speech is a complex signal which holds information about language, speaker, message, and emotion. Last several decades researcher from psychology and neuroscience has proven that emotions are closely related to decision making and so it plays an important role in the rational actions of human. Emotions can be identified by speech and facial expression but identification from the speech is extremely difficult task. If emotion gets involve in a speech, human robot interaction can be possible in better and effective way [1, 2, 3]. Detail day to day application has been explained further in this paper.

An emotion is a complex psycho-physiological short-time experience resulting from the interaction of biochemical (internal) and environmental (external) factors. Decisions are highly affected by emotions. Recent research emanates that if events occurrence is happened along with emotions, it memorized and recalled with more clarity than without emotions [4].

According to Darwin's theory of evolution, emotions find their provenience in biology. Emotions classification can has been done in two ways approach one is discrete approach and another is continuous approach. According to discrete approach basic seven emotions are defined: happiness, sadness, anger, anxiety, boredom, disgust and neutral [5]. Further complex emotions can be classified by concoction of the basic ones. 40 discrete words for complex emotions were identified and classified in BEEV (Basic English Emotional Vocabulary) for automatic emotion recognition. In continuous approach each emotion can be manifested as a linear amalgam of valence (or evaluation), arousal (or activation) and potency (or power) for example, Thayer's valence-arousal representation and Plutchik's emotional wheel [6].

The paper is methodical as follows: 2. Applications of emotion recognition from speech, 3. Requirement of emotion recognition in India, 4. Available speech corpora for Indian Languages, 5. Development of emotion recognition in various local Indian languages, 6. Conclusion

II. APPLICATIONS OF EMOTION RECOGNITION FROM SPEECH

Emotion recognition from speech has major impact on certain real time applications. Few of them are reviewed and mention as per below [3,7].

Assign priority to customers in various call centres: Call centre is a best real time application where customers with different emotions, different dialogues and different languages exist ultimately, vast variety of emotions based speech conversation can be found. In today's world failure in assignment of priority to customers may lead to failure or loss in business. By using linguistic information and paralinguistic cues, different emotions states can be evaluated at Medical emergency call centre [8]. Emergency cases may be analyzed if priority is assigned to various patients, which may lead to saving a life. By using acoustical variables and feature selection techniques, now a day's real time processing of customer's dialogues is possible through software like ER: Emotion Recognition Software which helps to handle the customer while he/she is speaking. From speech extracting the acoustic features and semantics of the conversation, Call back prioritization can be set effectively by using emotion recognition of customers at various call centers and it can be used in a decision support system for impelling voice messages and delegate a proper agent to respond the message [9].

Artificial Intelligence and Robotics: Emotion detection has proven its most significance utilization in area of human machine interaction in more natural, effective and expressive way. It is rightly said that human equivalent performance can be reached only when machine can process emotions effectively [10,11]. Some emotion detection softwares are available for robotics and artificial intelligence like nViso, Affectiva.

Improved diagnostic tool: Speech Therapist may use the various emotions states of patient for identification various disorders through software like IcSpeech. Emotion detection has proven its importance in eTherapy domain like scrutinize depression level, early detection of autism spectrum disorder (ASD) [12]. For identification of depression and suicidality various psychomotor disturbances can be identified based on Acoustic analyses of power distribution, voice fundamental frequency (F0), formants and amplitude modulation (AM) [13].

Lie Detection system: Emotion recognition from the speech also helps in area such as Central bureau of Investigation where person's loyalty can be finding through their emotions, stress of speech, and so on. Take an example of X13-VSA software which is unorthodox, advanced and highly developed system and a fully computerized lie detector, operated on voice emotions [14].

Improved Computer Games: Majority of people across the world get entertain through videogames. Indeed, by the nature of what constitutes a game, one cannot dissociate games from emotions [15]. So, games can be utilized/ considered as an arena for eliciting, evaluating, expressing and even synthesising emotions. Existing videogames can be improved and become a lot more entertaining if emotions get involved in it. Apart from that level (Stage level hard, medium and easy) of games can be decided based on emotions of users [16].

Intelligent Teaching/Tutoring System: Intelligent Tutoring Systems are computer-based learning systems that will calibrate delivery of the lesson and the content of materials according to the students' characteristics and needs by analyzing and/or anticipating their behaviours, acknowledgements and feedbacks. So, emotions detection is needed to development of Intelligent Teaching system [17].Recent research has shown that emotions can influence human behavior and learning abilities, as a result developers of tutoring systems have also started to follow these ideas by creating affective tutoring systems [18]. Identification of various emotional states has been made based on a self-assessment method based on Self-Assessment [19]. Various affective teaching systems are also available like WalLis, Primeclimb etc.

Audio/ Video retrieval: For Indexing and better retrieval of audio and video files are also possible based on emotion detection. It can be implemented using support vector machines trained on the extracted features [20].

Smart Car Board System: The aim of in-car emotion recognition should be to increase the driver's safety and comfort zone. In the automatic driving systems, ability of a car to understand human speech and provide a human-like driver assistance system add feathers to existing system [21].

Language Conversion: Language Translation / speech to speech conversion can be more natural if various emotion states of original speaker/ source speech can be added to listener/ target speech [22].

Sorting of Voicemail/ Messages: As of now it is very important to assign a priority to voicemail, which has become an indispensable part of our professional and personal conversation. As of now we have to actively listen to all messages in order to identify the important one. Emotion detection is very useful in such area also [23]. Apart from that to develop naturalness of TTS system can be improved using emotion recognition [24].

III. REQUIREMENT OF EMOTION RECOGNITION IN INDIA

IV. India is Asian country which holds world's second highest numbers of languages are spoken. According to census of India 2001, 122 major languages and 1599 other languages are spoken in India these counts may vary from one sources to another because of dialect and language defines. As on 1st Dec 2007, The Eighth Schedule of the Constitution of India, 22 regional languages were approved based on speaking population as per below Table-1[25]. Major development in the area of emotion recognition from speech is done in English, German, Chinese, Spanish, Japanese, Russian, Swedish, and Italian languages. Reason for development may because of standard emotion speech corpora is available. But unfortunately there is no any registered citation or reference of an emotional speech databases in any of the Indian languages.

TABLE 1. APPROVED INDIAN LANGUAGES AS PER CONSTITUENT OF INDIA

Language	Family	Language	Family
Assamese	Indo-Aryan	Konkani	Indo-Aryan
Bengali	Indo-Aryan	Maithili	Indo-Aryan
Bodo	Sino-Tibetan	Malayalam	Dravidian
Dogri	Indo-Aryan	Marathi	Indo-Aryan
Gujarati	Indo-Aryan	Meitei	Sino-Tibetan
Hindi	Indo-Aryan	Nepali	Indo-Aryan
Kannada	Dravidian	Odia	Indo-Aryan
Telugu	Dravidian	Urdu	Indo-Aryan
Punjabi	Indo-Aryan	Kashmiri	Indo-Aryan
Sindhi	Indo-Aryan	Sanskrit	Indo-Aryan
Tamil	Dravidian	Santali	Austroasiatic

In India some of the organization like LDC-IL, CDAC, TIFR, and DIT are contributing toward speech databases in different local Indian languages. There a large scope of the emotion recognition from a speech field across the India and local Indian languages, that's why the paper is based on review of this field in India and Indian languages context. For many of regional Indian languages acoustic–phonetic analysis is not available, so which can further improved [26].

V. AVAILABLE SPEECH CORPORA FOR INDIAN LANGUAGES

Suitable Emotional speech corpora or database is imperative for distinguish emotions either for synthesis or recognition. An important point is that quality of speech corpora and performance of that is highly recommended before selecting it. There are reported speech corpora for English, German, Chinese, Spanish, Japanese, Russian, Swedish and Italian languages. There is speech database available for Indian languages but still at India level some standard speech corpora are available as mention below:

TABLE 2. VARIOUS EMOTION SPEECH CORPORA FOR INDIAN LANGUAGES

S. No	Corpus Name	Language	Types of Emotions
1	IITKGP-SESC : Speech Database for Emotion Analysis	Telugu	anger, disgust, fear, happy, neutral, sad, and sarcastic and surprise
2	Indian Institute of Technology Kharagpur Simulated Emotion Hindi Speech Corpus (IITKGP-SEHSC)	Hindi	anger, disgust, fear, happy, neutral, sad, and sarcastic and surprise
3	Graphic Era University Semi Natural Emotion Speech Corpus (GEU-SNESC)	Hindi	sadness, anger, happiness and neutral.

Koolagudi at al. (2009) have developed the speech corpora named as Indian Institute of Technology Kharagpur Simulated Emotion Speech Corpus (IITKGP-SESC) in Telugu language with the reinforcement of professional maestro from All India Radio (AIR), Vijayawada, India. The speech corpus is collected by simulating eight different emotions using the neutral (emotion free) statements. The subjective listening tests were used to assessed quality of the emotions present in the database [27].

Koolagudi at al. (2011) have developed the developed and simulated emotion speech corpus for Hindi language which is known as Indian Institute of Technology Kharagpur Simulated Emotion Hindi Speech Corpus (IITKGP-SEHSC). The database is recorded using professional maestro from Gyanavani FM radio station, Varanasi, India. During development of speech corpora eight different emotions were considered namely anger, disgust, fear, happy, neutral, sad, and sarcastic and surprise. Values of Mel frequency cepstral coefficients (MFCCs) were used as emotion classifier for spectral information, in addition to that values of Energy, pitch and duration was used for prosody information [28].

Koolagudi at al. (2012) have designed another speech corpora namely Graphic Era University Semi Natural Emotion Speech Corpus (GEU-SNESC). The same authors have extended their work and around 65.3% and 72% average emotion recognition performance was identified and observed for male and female speaker respectively [29]. Summary of speech corpora for regional Indian languages is in Table-2.

VI. DEVELOPMENT OF EMOTION RECOGNITION IN VARIOUS REGIONAL INDIAN LANGUAGES

The choice of speech features for emotional recognition is a crucial task. Mainly three different types of features namely excitation features, vocal tract system and prosodic features, were considered in majority of languages and same is applicable to Indian languages. This long era of this field reminds that several excitation features like LP residual energy [30], glottal excitation signal [31], excitation source signals [32], were used to find various emotions. Vocal tract features like MFCC [33], LPCC [34], Spectral features using Fourier and Chirp Transformations [35], Prosodic features like duration, intonation and intensity [36], hybrid features (combinations of all mention features) were used to get various emotional distributions [37]. Work carried in Indian languages can be summarized as below:

Hindi Language

Hindi is nation language of India and most spoken language across India. K Rao et al. (2011) have identified six emotions anger, disgust, fear, happy, neutral and sad from Hindi language. With the help of emotion classifier like Auto-associative neural network (AANN) models and Support Vector Machines (SVM) were explored for apprehending the dialect specific and emotion specific information with use of Mel frequency cepstral coefficients (MFCC) and durations of syllables, pitch and energy contours define prosodic features. Hindi dialects were also found along with emotions by using same approach. Recognition performance of the emotions recognition was 78% [38].A Agrwal et al. (2015), have proposed in Hindi language, emotion transfiguration, of neutral speech into emotional speech on a word basis. This emotion transfiguration, is based on the segmentation of the speech and it is done using main prosodic features `pitch' and `intensity', `zero crossing rate' and duration. Speech corpus was limited, better result was expected if strong database was used [39]. Pawar et al. (2015) have developed an algorithm using Sparse DTW which detects five emotions from Hindi Language with an efficiency of 75%. Again strong a standard database has made a question [40].

Marathi Language

Marathi is 4th most spoken language in India. Vishal at al. (2014) have developed a system for Marathi Language to identify five emotions based on temporal and spectral features which are closely related to the pitch, Mel Frequency Cepstral Coefficients (MFCCs) and Formants of speech. Linear Discriminant Analysis (LDA) was used to classify features. This article also describes speech corpora of artificial emotional Marathi speech. Marathi Database was developed by author himself from various Marathi movies (Actors and Actress) [41].Darekar at al. (2016) have proposed to get six emotions by using the concept of fusion for Marathi language. Fusion is the jumbling of results of analysis for getting higher

TABLE 3. REVIEW OF EMOTION RECOGNITION FOR REGIONAL INDIAN LANGUAGES

S. No.	Name of Indian Language	Emotions	Type of speech database	Acoustic Property	Ref.
1	Hindi	Anger, disgust, fear, happy, netural, sad (06)	12000 utterances 10 speakers (5 male & 5 Female) Sampled at 16 KHz	Mel Frequency Cepstral Coefficients (MFCCs), Syllables, Pitch, Energy contours	K Rao (2011)
2	Hindi	Neutral, Happy, Sad, Anger (04)	40 Sentences 3 Speakers (1 male and 2 female) Sampled at 44.1KHz "Praat" Speech Synthesis Software	Stress, Intensity, Pitch, Duration of word	Agarwal (2015)
3	Hindi	Happy, Sad, Surprise, Neutral, Anger (05)	50 People Sampled at 8KHz Audacity Software	Mel Frequency Cepstral Coefficients (MFCCs), SparseDTW	Pawar (2015)
4	Marathi	Happy, Sad, Anger, Afraid and Surprise (05)	Own database using Marathi Movies Sampled at 44 KHz Cool Edit Pro version 2.0 software	Mel Frequency Cepstral Coefficients (MFCCs),	Vishal (2014)
5	Marathi	happy, angry, sad, surprised, fear and neutral (06)	100 speech files Speakers (Male & Female) Sampled at 44.1KHz, Stereo Channel	Pitch, Formants, Energy, MFCC	Darekar (2016)
6	Marathi	happy, angry, sad, surprised, fear and neutral (06)	1200 speech files Speakers (Male & Female) Sampled at 44.1KHz	Energy, MFCCs, pitch values, timbre, vocal tract frequencies	Darekar (2016)
7	Marathi and Kannada	anger, fear, joy, sadness, disgust, surprise, neutral (07)	Not mention	Pitch, Loudness	Nandini (2012)
8	Malayalam	neutral, happy, sad and anger (04)	20 Sentences 16 Speakers (8 Male & 8 Female) Sampled at 44.1KHz Sound Forge Pro 10 Software	Mel Frequency Cepstral Coefficients (MFCCs), Short Time Energy (STE) and Pitch	Rajisha (2015)
9	Malayalam	Neutral, Happy, Sad, Anger, Boredom, newlineFear, Surprise, Calm and Anxiety (09)	5 (3 male and 2 female) Sampled at 8 KHz, Malayalam emotional speech database	Spectral features (LPC and MFCC) Prosodic features (pitch first three formants and the RMS energy) Wavelet features(WPD energy) Hybrid Features	Firoz (2016)
10	Assamese and Bodo	anger, disgust, fear, happiness, sadness, and surprise (06)	16 Speakers (8 Male & 8 Female) Multilingual Emotional Speech Database of North East India	Eigen Values of Autocorrelation Matrix (EVAM) Mel Frequency Cepstral Coefficients (MFCCs),	Aditya (2010)
11	Assamese	anger, disgust, fear, happy, neutral, sad and surprise [07]	4200 Utterances Sampled at 44.1KHz Multilingual Emotional Speech Database of North East India" (MESDNEI)	Prosody Features (pitch, energy, Zero crossing rate), quality features (Formant Frequencies, Spectral features), derived features (i.e. Mel-Frequency Cepstral Coefficient (MFCC), Linear Predictive Coding Coefficients (LPCC)) dynamic feature (Mel-Energy spectrum dynamic Coefficients (MEDC))	Amiya (2015)

level of accuracy. Database was developed by author himself by various dialogues by actors. He also suggested that the emotion generated by a person will be same irrespective of language so and so forth process of emotion detection is independent of language [42]. Darekar at al. (2016) have expanded his work in Marathi language and proved that accuracy can be enlarged by concoction of values of different parameters together [43].

Kannada Language

Kannada is 8th most spoken language in India. Nadini at al. (2016) have proven in two Indian languages namely Marathi and Kannada that although Marathi and Kannada are developed from two different language families, there exist many common emotiphons. Emotiphons helps for identification of emotion in speech processing, adding naturalness to synthesized speech, and in design of dialogue systems [44].

Malayalam Language

Malayalam is 9th most spoken language in India. Rajisha at al. (2010) have developed speech emotional database for Malayalam language and identified four emotions using Mel Frequency Cepstral Coefficients (MFCCs), Short Time Energy (STE) and Pitch as features extraction techniques. For pattern classification Artificial Neural Network (ANN) and Support Vector Machine (SVM) were used and provided accuracy of 88.4 % and 78.2 % respectively [45]. Firoz at al. (2016) have identified the nine emotions states in Malayalam language using Spectral features (LPC and MFCC), Prosodic features (pitch first three formants and the RMS energy), Wavelet features (WPD energy) and Hybrid Features [46].

Assamese and Bodo Language

Assamese is 12th most spoken language in India. Aditya at al. (2010) has extracted six emotions for Assamese and bodo and other three local language of asam (state of India) with the help of Gaussian Mixture Model (GMM) as classifier. A new feature set is evaluated based on Eigen Values of Autocorrelation Matrix (EVAM) of each frame of the speech signal and the performance of the proposed feature set is compared with Mesl Frequency Cepstral Coefficients (MFCCs) [47]. Amiya at al. (2015) has identified seven discrete emotional states in five native Assamese Languages using Multilevel SVM as classifier. The inclusive results of the conducted investigation revealed that the approach of using the combination of features achieved in Multilingual Emotional Speech Corpora of North East India" (MESDNEI) corpus, with an average accuracy rate of 82.26% for speaker independent cases [48].

Summary of all the authors has been mentioned in below Table-3.

Some important Points to be discussed: 1) As per Indian context more research is needed to develop a standard emotional speech database on all local languages and more work should be carried out. 2) Speech corpora needed to be more expand in terms of more speakers and generality. Majority of database should be based on real emotions. 3) Design of appropriate accurate predication models is required. Hybrid classifier can be used for emotion classification to improve result. 4) More research needed in Cross lingual emotion recognition to make field more interesting because emotions is universal, independent of languages, gender and independent of speaker. 5) Cross lingual emotion may be possible for Indian regional languages because 16 languages out of 22 languages are coming same family (Indo-Aryan). 6) Expression of emotions are multimodal activity so, facial expression can be added with speech to get robust emotion recognition systems. 7) Standard and registered emotional speech corpora are available in very few languages as mention in paper. More emotional speech corpora are needed in other languages of world also. 8) Out of 22 Indian languages only 2-3 languages (13-15%) contain standard Emotional Speech corpora. This can be improved further. 9) Out of 22 Indian languages research work has been done in only 6-7 Indian languages (25-30%) which is an serious issue. 10) Predominance part of Indian languages are obtained from Sanskrit language so, more attention towards Sanskrit language may provide great impact.

VII. CONCLUSION

Emotion recognition from speech is blooming area of research. Researchers across the world have put lots effort and considerable amount of that, but as far as Indian local languages are concern it requires more improvement. Out of 22 Indian local languages in only 6-7 languages, some amount of work in the area of emotion recognition and emotional corpora has been accomplished. So there is a scope of research in other regional Indian languages also. This paper contains review of recent work in Indian languages and some of important issues are also discussed in the paper.

REFERENCES

[1] Schuller, B., Rigoll, G., & Lang, M," Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture". In Proc. IEEE int. conf. acoust., speech, signal processing ,pp. 577– 580, 2004.

[2] Dellert F., Polzin T., and Waibel A., "Recognizing emotion in speech", 4th international conference on spoken language processing, Philadelphia, PA, USA, pp. 1970–1973, 1996.

[3] Koolagudi and K Rao, "Emotion recognition from speech: a review". International journal of speech Technology, Springer, Vol. 15, pp. 99-117, 2012.

[4] Mélanie Fernández Pradier," Emotion Recognition from Speech Signals and Perception of Music" ,Ph.D. Dissertation, University of Stuttguart, 2011.

[5] P. Ekman.(1992). An argument for basic emotions. Cognition and Emotion, Vol. 6, pp. 169–200, 1992.

[6] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, K. Kollias, W. Fellenz, and J.G. Taylor. ,"Emotion recognition in Human-Computer Interaction", IEEE Signal Processing Magazine, Vol. 22(1),pp. 32-80, 2005.

[7] S. Ramakrishnan," Recognition of Emotion from Speech: A Review", Book chapter Speech Enhancement, Modeling and Recognition- Algorithms and Applications. InTech, DOI: 10.5772/39246, pp. 121-138 2012.

[8] Laurence Devillers and Laurence,"Real-Life Emotion Recognition in Speech", Speaker classification-II, Springer, pp. 34-42, 2007.

[9] Valery Petrushin. ," Emotion in Speech: Recognition and Application to Call Centers", Proceedings of Artificial Neural Networks in Engineering, pp. 1-8, 1999.

[10] O'Shaughnessy, "Speech communication human and machine", .Reading: Addison-Wesley,1987.

[11] Affectiva software. Available: https://www.affectiva.com/

[12] Pascal Ackermann, Christian K, Jo A B, Laus W & Sabina J. ,"EEG-based automatic emotion recognition: Feature extraction, selection and classification methods", IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), 2016.

[13] France DJ, Richard G. Shiavi, Stephen Silverman, Marilyn Silverman, and D. Mitchell Wilkes, "Acoustical properties of speech as indicators of depression and suicidal risk", IEEE Transactions On Biomedical Engineering, Vol.-4(7), pp. 829-837, 2000.

[14] Suresh A. and sahseen ,"Speech Stress Analysis based on Lie Detector for Loyalty Test", International Journal of Printing, Packaging & Allied Sciences, Vol. 4(1), pp. 631-638, 2016.

[15] Yannakakis G.N., Karpouzis K., Paiva A.and Hudlicka E,"Emotion in Games". Affective Computing and Intelligent Interaction, 2011.

[16] Alexandru P., Joost B., Marteen V S, "GAMYGDALA: An Emotion Engine for Games", IEEE Transactions on Affective Computing, Vol. 5(1), pp. 32-44, 2014.

[17] Roger Nkambou,"A Framework for Affective Intelligent Tutoring Systems", 7th International Conference on Information Technology Based Higher Education and Training, 2006.

[18] Sintija Petrovica, AllaAnohina-Naumeca and Hazım Kemal Ekenel,"Emotion Recognition in Affective Tutoring Systems: Collection of Ground-truth Data", Procedia Computer Science, Vol.104, pp. 437-444, 2017.

[19] Jacob Whitehill, Marian B and Javier M., "Automatic facial expression recognition for intelligent tutoring systems", IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008.

[20] Tao Li, "Content-based music similarity search and emotion detection", IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings, 2004.

[21] Florian Eyben, Martin Wöllmer, Tony Poitschke, Björn Schuller, Christoph Blaschke, Berthold Färber, Nhu Nguyen-Thien and lorian Eyben, "Emotion on the Road—Necessity, Acceptance, and Feasibility of Affective Computing in the Car", Advances in Human-Computer Interaction, HindawiPublications, 2010.

[22] Moataz El Ayadi, Mohamed K and Fakhri K,"Survey on speech emotion recognition: Features, classification schemes, and databases", Elsevier journal on Pattern Recognition. Vol. 44(3), pp. 572-587, 2011.

[23] Zeynep Inanoglu & Ron C. ," Emotive alert: HMM-based emotion detection in voicemail messages", Proceedings of the 2005 International Conference on Intelligent User Interfaces, January 10-13, 2005 San Diego, California, USA, 2005.

[24] Bidisha Sharma and S R Prasanna , " Polyglot Speech Synthesis: A Review", IETE Technical Review, Vol. 34 (4), pp. 366-389,2017.

[25] Public Document on "Constitution of India 2008" online available: https://web.archive.org/web/20140909230437/http://lawmin.nic.in/coi/coiason29july08.pdf

[26] Biswajit Sharma and S R Prasanna, "Acoustic–Phonetic Analysis for Speech Recognition: A Review", IETE Technical Review, 2017.

[27] Koolagudi, S. Maity, V. A. Kumar, S. Chakrabarti and K. S. Rao, "IITKGP-SESC : Speech Database for Emotion Analysis", Communications in Computer and Information Science, JIIT University, Noida, India: Springer,2009.

[28] Koolagudi, Ramu Reddy, Yadav and K. S. Rao. , " IITKGP-SEHSC : Hindi Speech Corpus for Emotion Analysis", IEEE proceedings on International Conference on Devices and Communications (ICDeCom),2011.

[29] Koolagudi, Swati, Bhavna, Anurag & K. S. Rao,"Recognition of emotions from speech using excitation source features", Elsevier proceedings for international conference on modeling, optimization and computing,2012.

[30] Rao, K. S., Prasanna S. R. M., and Yegnanarayana B ," Determination of instants of significant excitation in speech using Hilbert envelope and group delay function", IEEE Signal Processing Letters, 14, 762–765, 2007.

[31] Cummings K. E., and Clements M. A., "Analysis of the glottal excitation of emotionally styled and stressed speech", The Journal of the Acoustical Society of America, pp. 88–98, 1995.

[32] Hua L. Z., Yu H. and Hua W. R., "A novel source analysis method by matching spectral characters of LF model with STRAIGHT spectrum", Berlin: Springer, 2005.

[33] Mubarak, O. M., Ambikairajah, E., and Epps, J., "Analysis of an MFCC-based audio indexing system for efficient coding of multimedia sources", 8th international symposium on signal processing and its applications, Sydney, Australia, 2005.

[34] Pao, T. L., Chen Y. T., Yeh J. H., Cheng Y. M., and Chien, C. S. ,"Feature combination for better differentiating anger from neutral in Mandarin emotional speech". Berlin: Springer. 4738, 2007.

[35] Sigmund, M ," Spectral analysis of speech under stress", International Journal of Computer Science and Network Security,Vol. 7, pp.170–172, 2007.

[36] Lee C. M., and Narayanan S. S., "Toward detecting emotions in spoken dialogs. IEEE Transactions on Audio, Speech, and Language Processing, Vol. 13, pp. 293–303, 2005.

[37] Nicholson, J., Takahashi, K., & Nakatsu, R, "Emotion recognition in speech using neural networks", In 6th international conference on neural information processing (ICONIP-99), Perth, WA, Australia, pp. 495–450, 1999.

[38] K Sreenivasa Rao and Koolagudi,"Identification of Hindi Dialects and Emotions using Spectral and Prosodic features of Speech", Journal Of Systemics Cybernetics And Informatics, Vo. 9(4), pp. 24-33, 2011.

[39] A Agarwal and Amita Dev. "Emotion Recognition and conversion based on segmentation of speech in Hindi Language",2nd IEEE International Conference on Computing for Sustainable Global Development (INDIACom),pp.1843-1847, 2015.

[40] V. Pawar. and Nupur Patel ,"Emotion recognition from Hindi Speech using MFCC and sparse DTW", International journal of Engineering & Technology, Vol. 4(6), pp.1-5, 2015.

[41] Vishal Wagmare, Ratandeep D, Pukhraj S and Ganesh J," Emotion Recognition System from Artificial Marathi Speech using MFCC and LDA Techniques", Elsevier 5th International Conference on Advances in Communication, Network, and Computing,2014.

[42] R. V. Darekar & A P Dhande," Enhancing effectiveness of emotion detection by multimodal fusion of speech parameters", IEEE International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp.3242-3246, 2016.

[43] R. V. Darekar & A P Dhande," Improving emotion detection with speech by enhanced approach", IEEE International Conference on 3rd International Conference on Signal

Processing and Integrated Networks (SPIN), pp. 364-369, 2016.

[44] Nandini Bondale & Thipuur S., "Emotiphons: Emotion Markers in Conversational Speech - Comparison across Indian Languages", Proceedings of the 2nd Workshop on Sentiment Analysis where AI meets Psychology.73-79, 2012.

[45] T. M. Rajisha, A P Sunija & K Riyas ,"Performance Analysis of Malayalam Language Speech Emotion Recognition System using ANN/SVM", Elsevier International Conference on Emerging Trends in Engineering, Science and Technology ,2015.

[46] Firoz Shah, "Study and analysis of speech emotion recognition", Ph.D. dissertation, India, 2016. Available: http://shodhganga.inflibnet.ac.in/handle/10603/122185.

[47] Aditya K, Aurobinda R and Tapan B, "Vocal emotion recognition in five languages of Assam using features based on MFCCs and Eigen Values of Autocorrelation Matrix in presence of babble noise", IEEE National Conference on Communications (NCC), 2010.

[48] Amiya K, Kamalakanta M, Bibek K & Aurobindo R , " A novel approach of Speech Emotion Recognition with prosody, quality and derived features using SVM classifier for a class of North-Eastern Languages", IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS), pp. 372-377, 2015.

Mr. V P Tank pursed Bachelor of Engineering from Dharmsinh Desai University, Nadiad in 2009 and Master of Engineering from Gujarat Technological University in year 2011. He is currently pursuing Ph.D. and currently working as Assistant Professor in V T Patel Department of Electronics & Communication Engineering, Charotar University of Science & Technology, since 2012. He is a life time member of IETE since 2014. He has published more than 10 research papers in reputed international journals and conferences and it's also available online. His main research work focuses on Digital Speech processing, Bioelectronics, Digital signal processing. He has 7 years of teaching experience and 1 years of Research Experience.

Dr S K Hadia is an associate professor *at* V T Patel Department of Electronics & Communication Engineering, Charotar University of Science & Technology. His main research work focuses on Optical communication computer network, Image processing.. He has more than 15 years of teaching experience.