International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Using Vector Quantization Technique in Recognition of Gurbani Hymns (Japji Sahib): LBG Algorithm (VQ)

Sakshi Dua, Amit Kamra

Abstract- An improved and different variation of Automatic Speech Recognition (ASR) is presented which is based on Vector Quantization (VQ). ASR for different languages and different applications has been introduced so far. In this paper, we have presented a Speech Recognition system to recognize the hymns (paath) of Gurbani (sentences of Japji Sahib) as continuous mode of speech. For this, speech corpus has been generated in which the entire path has been recited by different speakers. The speech mode here can be taken as continuous speech encapsulated with background music and different kinds of additional noises and have been eliminated. The work has been done by using VQ approach of speech recognition and LBG algorithm which design optimal codebooks for the process of recognition. Experimental results are included which show that recognition accuracy for such system was found to be 92.6% and 95.8% for different and same speakers with different and same sentences.

I. INTRODUCTION

Automatic Speech Recognition is an intervention of computer machines to convert the sound and speech signals to some form which can be easily interpretable for human beings. For this reason, ASR is counted under the domain of Human Computer Interface (HCI). Broadly it can be mentioned as a part of Digital Signal Processing (DSP). In order to perform efficient DSP, reducing the noise or background noise effects from signal (speech sample whether isolated, connected or continuous). Various filters have been developed and designed for interpolating the signals, by testing on different signals, different environments with different noises. In early days, DSP aimed to separate the target signals from accompanied noise effects. Nowadays, DSP applications have been expanded to filter the random signals from noise. According to (Dripps, 1999), Signals can be of two types- Deterministic and Stochastic. Filtering the deterministic signals is usually done by using some digital filters. As far as, Speech Communications is concerned, it has played important role for dealing with many languages. It is a process of taking the speech signals by microphone and then converts it into sequence of words/ sentences. Speech sample is processed firstly to extract the required features for recognition. This is the phase known as feature extraction. Several feature extraction techniques are available to perform the same.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Sakshi Dua, Research Scholar, (IKG-PTU, Punjab, India)
Amit Kamra, Assistant Professor, (GNDEC, Punjab, India)

After this, calculating distance score and then classification is made. So, final result is calculated with retrieved scores. These kinds of systems are helpful and successful in various social areas. But the performance of optimal speech recognition is still unattainable. The development in this area is also useful for disabled persons, as they can just speak and their action is completed.

Though automatic speech recognition systems (ASR) have improved their execution and accuracy, still it is not gaining that accuracy and performance especially in some adverse circumstances. Advanced Speech Recognition systems are using different kind of language models which use predefined information to recognize the spoken utterances. Systems with small vocabulary are easy to design and implement. In this paper, we have worked on Gurbani Hymns as a speech which is different kind of speech recognition in system in itself. The research paper has been divided into sections. In section 2, Speech generation in human beings has been discussed. In section 3, speech production has been discussed. In section 4, variations in speech recognition has been presented. In section 5, Year wise trend of techniques, Section 6 containing speech recognition for different languages has been discussed. In section 7 and 8, Methodology of research VQ technique and results has been mentioned.

II. SPEECH GENERATION IN HUMAN BEINGS

Speech is a discrete segment of sound segments often known as phones; each of the phonemes poses certain articulator and acoustic properties. These are the smallest unit of phonemes that comprises of words. By referring to speech we are actually referring to Vocal Tract. The sounds those are possible due to Vocal Tract. The human speech consists of production of different sounds. Vocal tract is the main component of this entire system that acts as a thin passage which is made up muscles and various other tissues. Every phoneme poses certain features depending upon the position of vocal cords, lips, tongue, teeth and jaw. When breathing in and out takes place, lungs generate the power supply for voice.

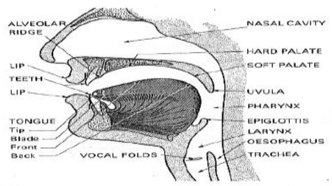

Fig. 1 Sound Production system in human beings [11]

This airflow vibrates the Vocal folds in the larynx to make basic sounds of voice. This process is known as phonation. There is an opening between vocal folds which is known as Glottis. The generated sound is weak enough which is strengthened by the process of resonance. So, we can say there are three significant mechanisms which are involved in production of speech.

a) Energy is generated by the lungs with the help of air that passes through them. This is the process of **Respiration.**
b) Sound is produced by vocal folds inside the larynx. This is the process of **Phonatory System.**
c) The generated sound is then strengthened, filtered and shaped by articulators. This is the process of **Articulation** in mouth.

III. SPEECH PRODUCTION PROCESS AND PRODUCTION MODEL

Speech is the primary input that is required to initiate the process of speech recognition and producing a sound is a complex mechanism. Speech in various languages can be provided to the system to be recognized and understand. Speech engines play significant role for doing so. Dataset plays vital role to do so. This is the domain of HCI, where mechanical recognizer converts the speech signals (input) to some typed transcripts. This mechanism is able to accept the sound waves of speech, process it and generates the output by recoding the information in particular to target language. Speech is an acoustic wave which gives information for the words/sentences which are being spoken by speakers or users of relative language.

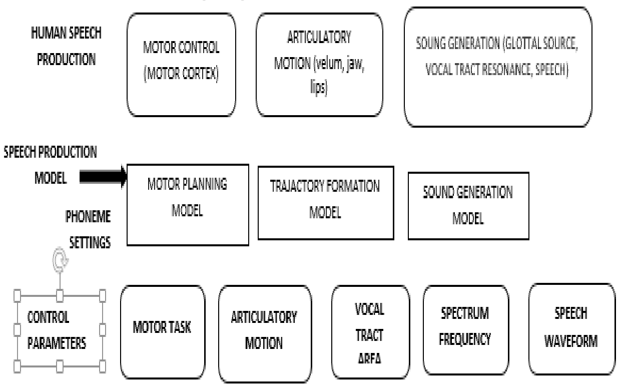

Fig. 2 Speech Production mechanism [30]

IV. VARIATION IN SPEECH RECOGNITION TECHNOLOGY

There are various forms of Speech Recognition systems starting from word to sentence level. At beginning it was performed to recognize the words i.e., Isolated Speech Recognition. Here, only words were able to be recognized [□□□□□], [□□□□□], [□□□□□], [□□□□□], [□□□□], [□□□□]. Then, researchers forwarded towards ***Connected Speech Recognition*** where words with minimal pause were able to be recognized – [□□□ □□□□□], [□□ □□□□□], [□□□ □□□□□]. Nowadays, ***Continuous Speech Recognition***, CSR in which speech of a human being is interpreted and recognized by machine in an efficient manner. The speech with no pause is considered here- [□□□□□□ □□□□ □□ □□□□□ □□], [□□ □□□□ □□□□□ □□□□□ □□□□□ □□ □□□□ □□□□]. Other than approaches, the performance of speech recognition relies upon the design of system and various parameters like mode of speaking, style of speaking, accent of speaker, pronunciation, vocabulary size, variability in transducer and numerous adverse conditions.

V. YEAR WISE TREND OF TECHNIQUES

In early times, ASR system was the concept of acoustic-phonetic process, where one of the aim was to describe the phonetic features of speech signal and then explaining how they are recognized acoustically in given speech signal. These features were indicating number of phonemes, articulation way. LPC is one of the techniques which was used by Atal and Itakura in 1960’s that significantly estimated the vocal tract response of human speech. By the mid of 1970, pattern recognition technique with the application of LPC, was proposed and designed by Itakura, Rabiner and Levinson. In this time, Tom Martin also founded the Speech Recognition Commercial Company known as Threshold Technology Inc. this company introduced real ASR product known as VIP-100 system which was able to influence Advanced Research Projects Agency (ARPA) of U.S.A. Between 1980 and 1990 speech recognition systems posed another improvement by using template-based approach. The concept of HMM came into existence and understood by few labs. (IBM and IDA). After the publication of its theory, HMM became the esteemed method for ASR. The hidden Markov model, which is a doubly stochastic process, models the intrinsic variability of the speech signal (and the resulting spectral features) as well as the structure of spoken language in an integrated and consistent statistical modeling framework **(Rabiner 2004)**. In 1980’s, another technology named as Artificial Neural Network (ANN) was introduced but unfortunately failed to yield optimal results. Error Propagation Method was used as a training method to mimic the human like reasoning. In early times, Neural Network (NN) was made to recognize simple speech of few words (Isolated Speech Recognition) and yielded optimal results. Towards the end of this decade, discriminative training and kernel based technique known as Support Vector Machines (SVM) came into trend, which was based on classification. Then, several new and different speech recognition systems were introduced like- SPHINX, BYBLOS and DECIPHER from CMU, BBN and SRI respectively. CMU Sphinx successfully trained the context-dependent phonemes and showed significant results for LVSRS (Large Vocabulary Speech Recognition Systems). ATIS, process involving spontaneous speech conversation, WSJ system for the transcription of speech (spoken paragraphs) from wall street journal and size of dataset was 60K words. In late 1990’s remarkable progress was made towards the development of Software tools and processing of tens of thousands of phone units was possible. Cambridge University team under the guidance of Steve Young, introduced new system known as Hidden Markov Model Tool Kit (HTK), and it was widely accepted tool for ASR. Year wise utilization of techniques has been shown in Table 1.

International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Table 1. Year wise utilization of techniques

Year	Technique used
2018	PLC(Packet Loss Concealment)
2018	HMM&MFCC, LPC, PLP
2018	HMM, K-Means Clustering, MFCC
2018	Pattern Matching, MATLAB CODE
2018	LPC, SVM (SM-SVM & LS-SVM)
2018	Spectral-Domain Speech Enhancement Technique
2017	Voice Activity Detection, Speech Enhancement Algorithm, HMM, PC-Based MATLAB Programs
2017	Concluded HMM Optimum for NLP
2017	GMM-HMM, MFCC, PLP Features
2017	HMM, MLP & MLE
2017	HMM, HTK, GMM-HMM
2016	HMM & OOV
2016	HMM + DNN, HMM + GMM+DNN
2016	HMM + New Improved MFCC
2015	MFCC, HMM, MLP, RBPNN, SVM, LVQ
2015	MFCC, LPCC, PLP, LPC
2015	MFCC, LPC, VQ, HMM
2015	HMM + VQ (LBG)

VI. SPEECH RECOGNITION FOR DIFFERENT LANGUAGES

ASR has performed significant role in different languages and has achieved remarkable performance. Different recognition techniques with different feature extraction methods have shown ASR a successful domain under HCI. Different languages for which ASR has been used, discussed below-

A. Automatic Speech Recognition in Assamese Language

Assamese is an official language of Assam state. Its phonetic character has been derived from Sanskrit. It has 11 vowels and 41 consonants in its script. Assam is a semi-literate state (Ghai, 2012). Sarma have worked for Assamese Speech Recognition, by developing numeral recognition system. Here, two variations- Gender and Mood were taken in consideration while recording and 10 numeral digits were spoken at KHz via mono channel microphone. LPC and PCA were used as feature extraction techniques to extract speech sounds and other irrelevant features. LVQ is comprised of Self Organised Map and Multi Layer Perceptron network used as classifiers. Success rate of this recognition rate was found to be 95%. Singh et. al. developed speech recognition system for Assamese language by using Deep Neural Network (DNN) by using Long Short Term Memory Network, which is a special kind of RNN (RECURRENT NEURAL NETWORK). MFCC was used as feature extraction technique in this work and 100% accuracy rate was achieved in this development.

B. Automatic Speech Recognition in Hindi Language

HHindi is an official language in Himachal Pradesh, Bihar, Uttar Pradesh, Chhattisgarh and Jharkhand. This language was first used with Brahmi Script. In 11th century AD onwards, it was written with Devanagri Script. Hindi language possesses 40 consonants, 10 vowels and 2 modifiers (Ghai, 2012). Samudravijaya developed a speaker independent/ continuous speech recognition system for spoken queries of railways reservation in Hindi Language. The dataset was containing 320 sentences and 161 words. The system used 10 speakers both male and female and developed by using HMM technique. Kumar and Aggarwal worked for developing an ASR for recognizing isolated words of Hindi Language. HMM was used to train and test the system by using MFCC as feature extraction. Dataset of 30 words was used with accuracy of 94.63%.

C. Automatic Speech Recognition in Bengali Language

Bengali is official Language in Bengal and is second most commonly spoken language in Language. The main aspect about this language is that National Anthem of our nation has been written in Bengali Language and the name of our national animal is also named as The Royal Bengal Tiger. For Bangladesh, Bengali is an official language. Phonemic inventory of Bengali contains 29 consonants and 14 vowels including 7 nasalised vowels (Ghai, 2012). Chowdhury has worked to develop ASR for Bangla by using CMU-SPHINX tool. Two different decoders were used which gave two different accuracy rates- 90.65% and 86.79% respectively. Hasnat et. al. developed Isolated and Continuous ASR for Bangla Language by using HTK tool. To eliminate noise, Adaptive filter was used. MFCC for extracting the features and HMM training model was used for training the system. For isolated recognition 90% and Continuous speech 80% accuracy was achieved.

D. Automatic Speech Recognition in Punjabi Language

Kumar has developed a speaker dependent, real time, isolated and connected words recognition ASR system. This system used medium sized vocabulary. For this system they made use of VQ and DWT techniques. LPC analysis was used for feature extraction process. The system was trained for 1500 isolated words and 500b isolated words were used for evaluating the performance. The entire performance/accuracy that was observed for isolated words was 61% and lesser for connected words. Kumar and Singh developed Automatic Spontaneous Punjabi Speech Recognition System. For the execution, they have used Sphinx toolkit and Java programming. The system generated 98.6% accuracy for Punjabi sentences.

E. Automatic Speech Recognition in Urdu Language

Urdu, being national language of Pakistan, is one of 22 scheduled languages in India and official language of 5 states. Urdu is mutually intelligible with standard Hindi spoken in India. Urdu has recognised dialects such as Dakhni, Rekhta and Modern Vernacular Urdu. It has 28 consonants and 10 vowels (Puncase_study). Raza et. al. developed a Spontaneous ASR for Large Vocabulary dataset. For this system they have used HMM and Sphinx 3.

As a speech, spontaneous conversation from Urdu speakers as a interview data has been recorded. Sarfraz et. al. developed Large Vocabulary Continuous Speech Recognition system for Urdu Language. CMU Sphinx Open Source Toolkit was made use for this system. The training data for system was recorded in noisy environment.

VII. PROBLEM STATEMENT

In this work we have worked for the new concept of speech recognition. Recognition of recited hymns in Gurudwara Sahib has been done so far. Japji Sahib Ji has been recognized by machine by using MFCC and LBG (VQ) algorithm.

VIII. WHY VECTOR QUANTIZATION FOR SPEECH RECOGNITION?

Vector quantization is the quantization of data in contiguous blocks referred as vectors. Quantization is the mapping process of infinite scalar/vector quantities to finite scalar/vector quantities. Quantization perform significant role in signal processing, speech processing and image processing. VQ has spread wings in these areas with development of algorithms like (LBG) Linde Buzo Gray algorithm.

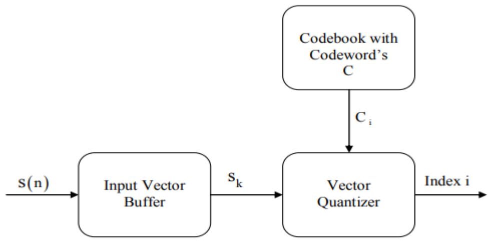

Fig. 3 Block Diagram of Vector Quantizer [31]

Vector quantization poses some limitations also like computational complexity and memory requirements, required for searching/storing codebooks. VQ in the field of speech coding was involved by Dudley and Smith in 1960. There are numerous techniques of vector quantization with own limitations and advantages. Each technique is supposed to decrease spectral distortion, computational complexity and requirements for memory. Several techniques of VQ are- Split Vector Quantization (SVQ) technique, Multistage Vector Quantization (MSVQ) technique, Split-Multistage Vector Quantization (S-MSVQ) technique and Switched Split Vector Quantization (SSVQ) technique. The performance of each and every technique depends upon efficient codebooks and codebooks are generated by using wide training set. It has been noticed that by decreasing the number of bits for generating codebooks also decreases the computational complexity and memory requirements but increases spectral distortion.

IX. WHAT IS LBG ALGORITHM

For implementing our work, LBG algorithm has been deployed. Vector Quantization technique generates the codebooks and codebooks here are generated by using iterative algorithm called LBG (Linde-Buzo-Gray) algorithm. This algorithm uses training sequence as input. Training sequence is a set of LSF vectors which are obtained from speakers of different groups and ages. Speech signals can be recorded in open environments and computer rooms. In this research work, speech signals have been recorded in room with different backgrounds and allowed music.

LBG algorithm works in following manner-

Obtaining the training sequence from set of speech samples which are recorded in different kind of environments.

Finding a Cendroid i.e., initial codebook obtained from training sequence.

Splitting of Cendroids into relevant set of codeword's.

Computing D, the difference between training sequence and codeword's.

Obtain new Cendroids by replacing old ones.

Repeating above process until required number of codewords in codebook are introduced.

X. STEPS OF IMPLEMENTATION, DESIGNING DATABASE AND RESULT EVALUATION

The proposed system of Gurbani Recognition has been implemented by using VQ (Vector Quantization) and LBG algorithm where optimal codebooks have been generated so far. For extracting features MFCC technique has been involvedthis system is able to recognize recited Gurbani hymns. The system is speaker independent continuous speech recognition using dataset which is developed exclusively for proposed system. The recognition of this kind, not only faces the barrier of ordinary background noises but background music also. So this was challenging fact of our system.

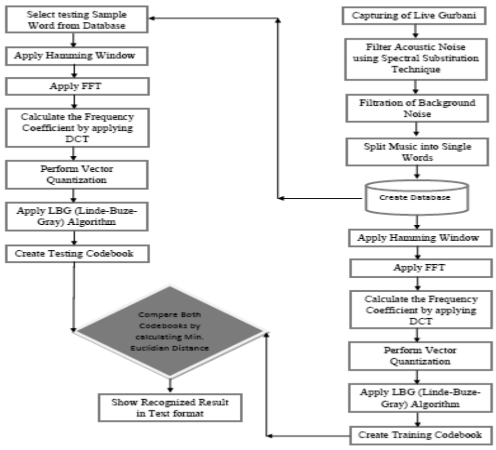

Fig. 4 Proposed Speech Recognition System

1) *Designing of Database*

In this research work, the database has been designed for recognizing the Gurbani Hymns.

Table 2. Details of Utterances and Speakers

Speech Corpus	Accent and Language	No. of Males	No. of Females
Training Corpus	Gurbani Hymn (Japji Sahib) Continuous	5	21

The size of the database contains 1092 sample files, recorded from 26 different speakers twice. The sentences have been recited in continuous way along with background music. The speech samples have been recorded by using optimal quality microphones and by using NCH Suite software. Speakers were in the age of 18-65 years. While training of system 1006 sentences were used (26 speakers, 21 Females and 5 male speakers). Sampling rate for samples was 44.1 kHZ and are recorded by using unidirectional microphone.

2) Feature Extraction phase

The main role of feature extraction is to extract the set of features or properties from input speech signal. Such features carry significant characteristics of speech information. Various feature extraction techniques are – Principal Component analysis (PCA, which is non linear and faster method). Linear Discriminative Analysis (LDA, which is linear and best classification method). Linear Predictive Coding (LPC, which is static feature extraction method). Cepstral Analysis, a static feature extraction method. Mel-Frequency Scale Analysis where Mel scale is computed. Filter Bank Analysis method which is filter based feature extraction method. (MFCC) Mel-frequency cestrum coefficient method is robust and dynamic method. Wavelet technique or spectral subtraction method is also robust extraction method. RASTA (Relative Spectra Processing, method which is used in noisy speech recognition.

In this work, MFCC has been used to perform the feature extraction method. The working pattern of MFCC can be represented as-

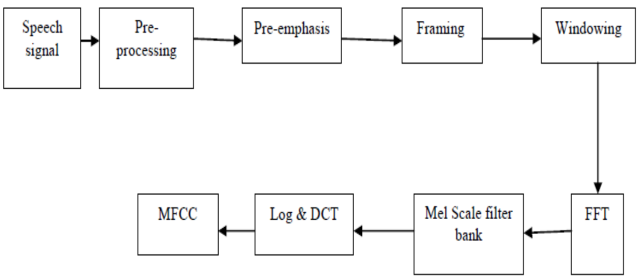

Fig. 5 Block Diagram representation for extracting MFCC features [24]

The process of MFCC consists of preprocessing phase where framing of speech samples and windowing is carried out. FFT is then calculated for every frame of input speech to obtain the frequency components. Thirdly, logarithmic filter bank is applied. DCT is calculated against the Mel Scaled filtered bank. In this work, the input speech is sampled at 44.1 kHz, and processed at 10 ms frame rate with a Hamming window of 30 ms.

3) Training Phase

1006 samples have been involved in the process of Training where 28 speakers were used (22 Female and 6 Male).
Firstly, input of spoken sentences of Japji Sahib Ji.
Segmentation of words was done so far.
At last, training file will be created for all the input sentences.

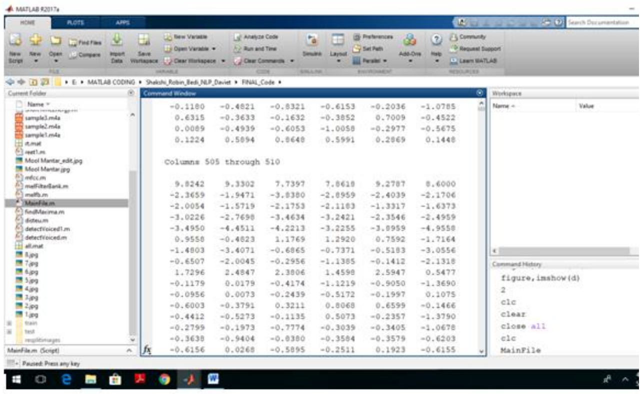

Fig. 6 Training phase of words in proposed system

Sampling rate was 44.1 kHZ and recording was done by using NCH Suite software, by using unidirectional microphone and under different background noises. During training, each word of speech is segmented which is shown in figure-

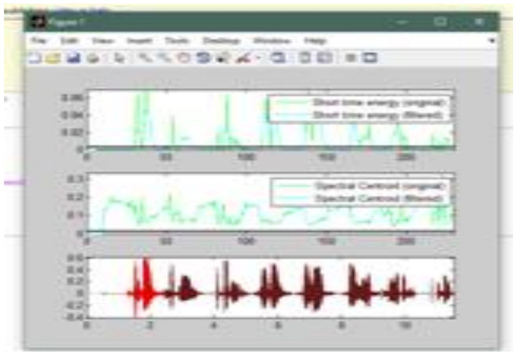

Fig. 7 Segmentation of words in proposed system

XI. RESULTS

Based on following stated experiments, the system has achieved 95.8% accuracy with same speakers as of training but on different sentences with 12.67% WER. With different speakers and same sentences the designed system is yielding 92.6% and 13.89% WER, showed in Table 3.

Table 3. Recognition Results

Exp. No.	Same Speakers, Different Sentences.		Different Speakers Same Sentences	
	Recog. Rate	WER	Recog. Rate	WER
Exp 1	94.87	11.57	91.33	10.67
Exp 2	94.42	10.75	93.56	11.88
Exp 3	95.02	13.88	92.90	11.09
Exp 4	94.45	11.97	94.06	10.98
Exp 5	95.28	12.05	92.55	9.55
Exp 6	94.78	9.95	91.34	12.45

Exp 7	95.67	10.89	91.09	13.22
Exp 8	94.32	11.99	90.01	10.12
Exp 9	95.09	12.45	93.89	9.43
Exp 10	94.06	13.77	94.66	10.99
Exp 11	95.11	14.55	92.44	11.89
Exp 12	94.78	12.99	93.76	10.67
Average Results	**95.8 %**	**12.67**	**92.6%**	**13.89**

XII. CONCLUSION AND FUTURE WORK

The main aim of this research paper was to develop a speech recognition system for recognizing gurbani hymns (sentences of recited Japji Sahib Ji) as a continous mode of speech. The consideration factor about proposed system was not only that it was able to recognize recited sentences of Gurbani but also was elimination of additional music in recitation. Removal of additional music which was present in sentences was complex task in itself. The process is carried out by using Vector quantization approach for recognition and LBG algorithm. MFCC feature extraction method has been used to deploy the features. Sentence recognition accuracy for three phases of our experiment is found to be in the range of 91-94.6% and word recognition rate was 93-95.4%. The work has been limited to sentences of Japji Sahib (hymns). In future, we will try to work for more speedy recitation and announcements with optimal recognition rate and other techniques.

REFERENCES

1. Svoboda Marcus, Mihaela Frigura-Iliasa, DoruVatau " New Interpolation Tools for Digital Signal Processing", IEEE, The International Conference on Information and Digital Technologies, 2016
2. J H Dripps "AN INTRODUCTION TO MODEL BASED DIGITAL SIGNAL PROCESSING", 'Printed and published by the IEEE, 1997
3. M. Unser, "Splines: A perfect fit for signal and image processing," IEEE Signal Processing Magazine, vol. 16, no. 6, pp. 22-38, 1999.
4. Bhiksha Raj, Vipul N. Parikh, and Richard M. Stern "The Effects Of Background Music On Speech Recognition Accuracy", IEEE 1997
5. Jab, U., Siegler, M. A., Doh, S.-J., Gouvea, E., Mmno, P. J.,B Raj, B., and Stern, R. M., "Recognition of continuous broadcast News with multiple unknown speakers and environments", Proc. ARF'A Speech Recognition Workshop, pages 61-66, February, 1996.
6. P.S.Gopalakrishnanetal, "Suppressing background music from corrupted data of the ARPA Hub4 task", proceedings of ARPA speech recognition workshop. Pages 81-84, February, 1996.
7. Acero, A., "Acoustical and Environmental Robustness in Automatic Speech Recognition", Kluwer Academic Publiisbers, Boston,MA, 1993.
8. Nakadai, Y., &Sugamura, N., " A speech recognition method for noise environments using dual inputs". In Proceedings of The international conference on spoken language processing (pp. 1141–1144), Kobe, Japan, November, 1990
9. [9] O'Shaughnessy, D. (1987). Speech communication: Human and machine. New York: Addison-Wesley.
10. J. Allen. "How do humans process and recognize speech?" IEEE Transactions on Speech and Audio Processing, 2(4):567–577, Oct. 1994.
11. X. L. Aubert, "An overview of decoding techniques for large vocabulary continuous speech recognition", Computer Speech and Language, 16(1):89–114, Jan. 2002
12. S. F. Chen and J. Goodman, "An empirical study of smoothing techniques for language modelling", Computer Speech and Language, 13(4):359–393, Oct. 1999.
13. A. P. Dempster, L. N. M., and R. D. B, "Maximum likelihood from incomplete data via the EM algorithm", Journal of the Royal Statistical Society. Series B, 39(1): 1–38, 1977.
14. H. Hermansky, "Should recognizers have ears?" Speech Communication, 25(1–3): 3–27, 1998.
15. T. Hirsimaki, "Advances in Unlimited-Vocabulary Speech Recognition for Morpholy Rich Languages", PhD thesis, Helsinki University of Technology, 2009.
16. F. Jelinek, "Continuous speech recognition by statistical methods", Proceedings of the IEEE, 64(4):532–556, Apr. 1976.
17. M. Kurimo, "Using Self-Organizing Maps and Learning Vector Quantization for Mixture Density Hidden Markov Models", PhD thesis, Helsinki University of Technology, 1997.
18. S. Levinson, "Structural methods in automatic speech recognition". Proceedings of the IEEE, 73(11):1625–1650, Nov. 1985.
19. Saini P., Kaur P., Dua M. "Hindi Speech Recognition Using HTK", International Journal of Enginering Trends and Technology (IJETT), Volume Issue 6-June 2013
20. Boruah S., Basishtha S., "A Study on HMM based Speech Recognition System" IEEE, 2013
21. Ghai W, Singh N., "Analysis of Automatic Speech Recognition Systems for Indi –Aryan Languages: Punjabi A Case Study", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, March 2012
22. Ghai W., Singh N., "Continuous Speech Recognition for Punjabi Language", International Journal of Computer Applications (0975 – 8887) Volume 72– No.14, May 2013
23. Mishra et. al., "Study of robust extraction techniques for speech recognition system", IEEE, 2015
24. R. P. Lippmann, "Speech recognition by machines and humans", Speech Communication, 22(1):1–15, Jul. 1997.
25. J. J. Odell, "The Use of Context In Large Vocabulary Speech Recognition", PhD thesis, Queens' College, 1995.
26. D. O'Shaughnessy, "Invited paper: Automatic speech recognition: History, methods and challenges", Pattern Recognition, 41(10) :2965–2979, Oct. 2008.
27. D. Povey, "Discriminative Training for Large Vocabulary Speech Recognition", PhD thesis, University of Cambridge, 2003.
28. C. E. Shannon, "A mathematical theory of communication", The Bell System Technical Journal, 27:379–423, 1948.
29. Rabiner R. L & Schafer W. R, "Introduction to Digital Speech Processing", Foundations and Trends R in Signal Processing Vol. 1, Nos. 1–2 (2007) 1–194 © 2007
30. Makhoul J, Roucos S, Gish H., "Vector Quantization in Speech Coding", proceedings of the IEEE, VOL. 73. NO 11. November 1985

AUTHORS PROFILE

Sakshi Dua is working as Assistant Professor in the field of Computer Science Dept. (Lyallpur Khalsa College for Women, Jalandhar, Punjab, INDIA). She is having 11years of experience in this field. and pursuing her PhD in the field of Natural Language Processing (NLP) from Inder Kumar Gujral Punjab Technical University (IKG-PTU, Kapurthala).

Dr. Amit Karma is working as Assistant Professor in Guru Nanak Dev Engineering College, Ludhiana, Punjab, INDIA. He is area of Interest are Medical Image Processing, Detection of Cancer, Biometrics. He is member of Indian Institution of Industrial Engineering. He is having numerous publications in National and International Journals.