International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 4, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

Improvement in Speech Recognition Performance using Beamforming based Speech Enhancement

Milind U. Nemade, Prof. Satish K. Shah

***Abstract* — The speech, being a fundamental way of communication for the humans, has been embedded in various essential applications like speech recognition, voice-distance-talk and other forms of personal communications. There are so many applications of speech still to be far from reality just because of lack of efficient and reliable noise removal mechanism and techniques for preserving or improving the intelligibility for the speech signals. The broad categories of speech enhancement techniques can be listed as speech filtering techniques, beam forming techniques and active noise cancellation methods. In this paper, our work has two-fold objective. First is to improve the speech recognition performance in multi-microphone environment. Second, we attempted to analyze the performance of speech recognition against the filter-bank parameters; filter length and number of subbands. The experiments were performed for 20 words including numbers and commands, 10 words of numbers only and 10 words of commands only for different values of filter bank parameters. The results obtained have proved the speech enhancing capability of the beamforming technique in multi-microphone network where noise and echo-interference can degrade the original speech signal.**

***Keywords* — Beamformer, LMS algorithm, Subband filtering, Mel Frequency Cepstral Coefficient, HMM based classifier.**

I. Introduction

Speech is the fundamental and common medium, hence important for us, to communicate and most effective and reliable means for expressing oneself for personal communication. With advancement in hardware technologies, there are so many electronic and mobile personal communication based devices available, today in market and that too in cheaper cost and with easy availability. The applications like speech recognition, mobile and personal communication, public address system are few of the applications from long list of speech based systems. However, undesired noises in environment like sound from heavy machines, vehicles are also present in one or other form everywhere. These noises cause undesired effects in speech transmission and acquiring systems. Recently, restricted or usable vicinity of applications is moving from one place and close room to more open and multiple locations, leading to several types of undesired signals of mixing with desired speech signal making speech more corrupt with noise. Not only human communications but intelligent machines which trying to automate the things and sometimes also takes decision based on what it receives as a speech, also suffers from the degraded performance.

Since last five decades, various approaches for noise reduction and speech enhancements have been investigated and developed. Among, very early and fundamental approach of noise reduction was introduced to use the theory of the optimum Wiener filter. Given a desired signal and an input signal, the Wiener filter produces an estimate of the desired signal that is optimal, i.e. the squared mean error or difference between the signals is minimized. The Wiener filter can also be adaptively estimated used in an environment where the surrounding noise has time-varying characteristics. Adaptive algorithms such as Least Mean Square (LMS) and Recursive Least Squares (RLS) are well known examples and also widely used.

Recent advances in CPU and multi-core hardware has provided ample amount of computational power and thus, need for today is to design the complex but yet efficient and realistic approach for noise reduction to achieve speech enhancement. The speech enhancement is not only useful for storage and transmission of speech data but it can play vital role in improving much need system based speech recognition where accurate identification of words and sentences can provide automation in most of the human-machine based interface and also be useful in machine-machine interaction based automation. Robotics is a familiar example where speech recognition systems can become boon for today's advanced society at social level in addition to during natural calamities and on war fields.

It is obvious that speech enhancement can boost up the performance of speech recognition systems by keeping low word error rate (WER). There are various types of advanced speech enhancement algorithms in literature and they can be classified in main three categories, namely; filtering/estimation based noise reduction, beam forming and active noise cancellation (ANC) techniques. In this paper, our work has two-fold objective. First is to improve the speech recognition performance in multi-microphone environment. Second, we attempted to analyse the performance of speech recognition against the filter-bank parameters; filter length and number of subbands. The experiments were performed for 20 words including numbers and commands, 10 words of numbers only and 10 words of commands only for different values of filter bank parameters. The results obtained have proved the speech enhancing capability of the beamforming technique in multi-microphone network where noise and echo-interference can degrade the original speech signal.

The remaining part of the paper is organized as follows: In next section II, existing work related to the beamforming have been presented. Section III explains the beamforming structure used in the experiments and how multi-mic environment have been simulated. The beamforming technique and recognition have been explained in brief in section IV. Section V specifies the dataset preparation and discusses on the results obtained with various experiments. Finally, paper is concluded with summary of the work in section V.

International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 4, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

II. Related Work

One of important class of speech enhancement methods is based on the beam-forming, where more than one speech channels (microphones) are used to process the speech. Speech signals are received simultaneously by all microphones and outputs of these sensors are then processed to estimate the clean speech signal. In adaptive beamforming, an array of antennas is exploited to achieve maximum reception in a specified direction by estimating the signal arrival from a desired direction (in the presence of noise) while signals of the same frequency from other directions are rejected. This is achieved by varying the weights of each of the sensors (antennas) used in the array. This kind of speech enhancement techniques can give better performance of the speech applications like automatic speech recognition (ASR) than signal channel processing. Only disadvantage with this class of methods is higher cost of hardware, which can put restriction on using these methods in some speech applications. The basic block diagram of beamformer is shown in figure 1.

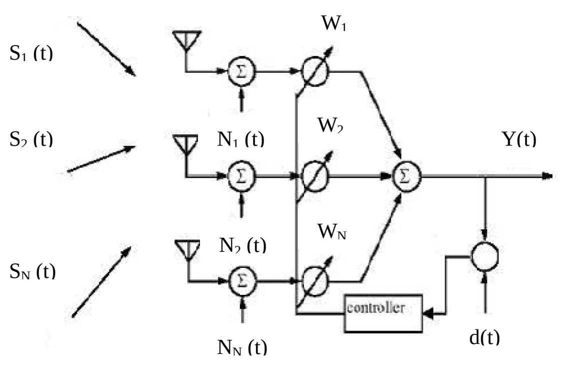


Fig.3. Beamformer: An Adaptive array system

Frost [1] has suggested constrained minimum power adaptive beamforming, which deals with the problem of a broadband signal received by an array, where pure delay relates each pair of source and sensor. Each sensor signal is processed by a tap delay line filter after applying a proper time delay compensation to form delay-and-sum beamformer. The algorithm is capable of satisfying some desired frequency response in the look direction while minimizing the output noise power by using constrained minimization of the total output power. This minimization is realized by adjusting the taps of the filters under the desired constraint using constrained LMS-type algorithm. Griffiths and Jim [2] reconsidered Frost's algorithm and introduced the generalized side-lobe canceller (GSC) solution. The GSC algorithm is comprised of three building blocks. The first is a fixed beamformer, which satisfies the desired constraint. The second is a blocking matrix, which produces noise-only reference signals by blocking the desired signal (e.g., by subtracting pairs of time-aligned signals). The third is an unconstrained LMS-type algorithm that attempts to cancel the noise in the fixed beamformer output. In [2], it is shown that Frost algorithm can be viewed as a special case of the GSC. The main drawback of the GSC algorithm is its delay-only propagation assumption.

In another work [3], switching adaptive filters were used to form the beamformer. This beamformer has two sections and interconnected with switch. The first section determines the adaptive look direction and cues in on the desired speech and is adapted only when speech is present. Second section which adapted during silence-only periods is implemented as multichannel adaptive noise canceller. In [4], authors have proposed the solution to GSC algorithm by estimating ratio of transfer functions (TFs), otherwise it is based on TFs which relates source signal and the sensors. The TF ratios are estimated by exploiting the non-stationarity characteristic of the desired signal. This algorithm can be used normally in reverberating room having acoustic environment. One interesting paper [5], describes how optimal finite-impulse response subband beamforming can be used by including coherent multipath propagation into optimality criterion for speech enhancement in multipath environment.

In application point of view, a constrained switched adaptive beamforming (CSA-BF) [6] was used for speech enhancement and recognition in real moving car environment. This algorithm consists of a speech/noise constraint section, a speech adaptive beamformer and noise adaptive beamformer. The performance obtained with this algorithm was compared with classic delay-and-sum beamforming (DASB) using CU-Move corpus and found decrease in word-error-rate (WER) by 31% in speech recognition. The computational complexity of DASB is very low and can be easily implemented for real-time requirement. It is also effective when direction of desired source is known and can be applied in the car as driver's head position is restricted based on seat position. However, as there is possibility of change in drivers head direction, DASB algorithm could be inconsistent and this inconsistency can be solved by employing CSA-BF algorithm which can improve the SNR by up to +5.5 dB on the average. For the application of hands-free speech recognition, one of the works [7] uses sequence of features to be used for speech recognition itself, to optimize a filter-and-sum beamformer instead of separating the beamformer, to be used for speech enhancement, from speech recognition system. In this work, they used frequency cepstral coefficient (MFCC) and applied to the HMM based classifier for speech recognition.

Optimizing beamformer without knowledge of source or acoustic characteristic of environment is termed as "blind beamforming". One of the papers [8] proposes blind speech enhancement using beamformer which consist of subband soft-constrained adaptive filter using recursive least square (RLS) algorithm, combined with subband weighted time-delay estimator (TDE). Estimation of propagation time difference of arrival of a dominate speech source received by sensor array is based the steered response power with phase transform (SRP-PHAT) algorithm, which was modified to work in subband structure. One recent paper [9] presents phase-based dual-microphone speech enhancement technique based on prior

International Journal of Electronics Communication and Computer Engineering
Volume 3, Issue 4, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209

speech model. In this work, it is claimed that around 23% improvement achieved using this algorithm as compared to the delay-and-sum beamformer, where experiments were conducted on the CARVUI database.

In application point of view, the study presented in [10] addresses the problem of distant speech acquisition in multiparty meeting s using multiple cameras and microphones. The camera, used as a multi-person tracker, was used to give the more precise location of each person to the microphone array beamformer. They evaluated the performance of speech recognition using data recorded in a real meeting room for stationary speaker, moving speaker and overlapping speech scenarios. The result obtained with audio-video speech enhancement was better than that with only audio. In one of the recent work [11], adaptive beamformer based on estimation of power spectral density (PSD) and noise statistics update was proposed. An inactive-source detector based on minimum statistics is developed to detect the speech presence and to acquire the noise statistics. The performances of this beamformers were tested in a real hands-free in-car environment. One of the most recent papers [12] uses GSC based speech enhancement using the location of speaker obtained via localization module. This algorithm relies on time delay compensation, DFT computations, fixed channel compensator, adaptive channel compensator.

III. Beamforming Filter Structure

In order to analyse the performance of the speech recognition in the speech corrupted by noise and echo-interference, the analysis frame work used here is depicted in figure 2. The noisy speech is simulated using multi-microphone speech environment is shown in figure 3. The main section beamforming based speech enhancement, filter bank design is explained in next subsection. In later subsection, multi-microphones speech generation is explained in detail.

A. Beamform Filter Design

Adaptive Filtering is an important technique in the field of speech processing including speech enhancement, echo- and interference cancellation and speech coding. Filter banks have been introduced in order to improve the performance of time domain adaptive filters with additional benefits like faster convergence and the reduction of computational complexity with shorter filters in the subbands being processed at reduced sampling rate [13]. Due to inappropriate structure of filter bank in subband processing and improper design of filters, filter bank may yield degraded performance. The subband FIR filter bank scheme [14] to be used for beamforming is shown in figure 4. The design of filters used here is adapted from and given in detail in references [14,15,16,17]. The design includes the prototype analysis and synthesis filter. The filter bank is obtained by using cosine modulation of prototype filter. The analysis-synthesis filter bank structure is shown in figure 5.

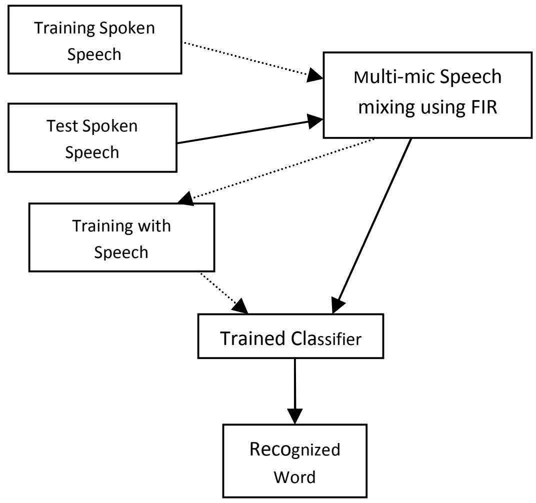


Fig. 2 Analysis Framework for Speech Recognition Performance using Beamforming

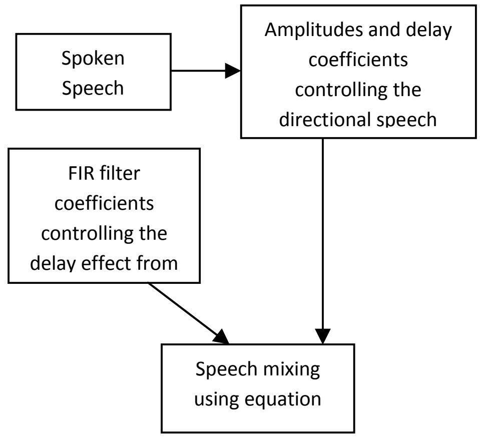


Fig.3 Speech-Splitting Scheme for simulating multi-microphones speech environment

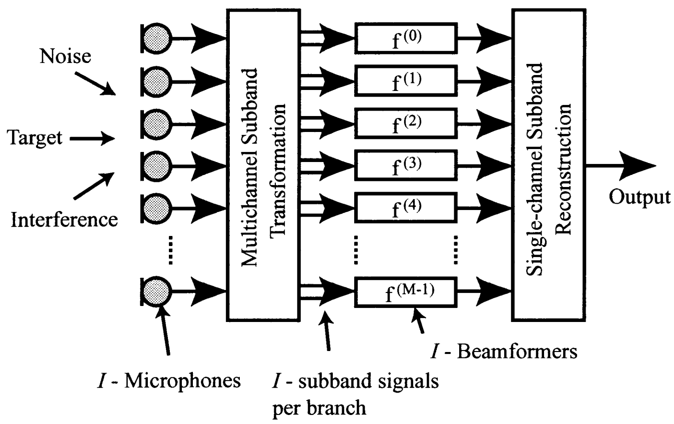


Fig.4. Subband FIR Beamforming Structure

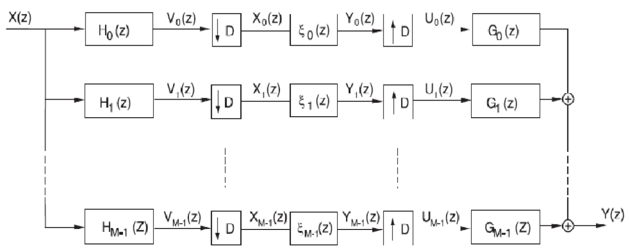


Fig.5. Analysis and Synthesis Filter Banks with Subband Filtering

B. Multi-Microphone Environment

The source of spoken word is from the speaker (person). This speech will travel to all the microphones with different delays and gains depending on the distance between the speaker and microphone. The spoken speech s(n) is simulated to produce N directional sources such that they will be acquired by N different microphones. This is achieved using the amplitude coefficients A(s, k) and filter coefficients b(s, n). The objective of amplitude coefficients is to control the gain of speech sources to be added in speech received by particular microphone. Filter coefficients controls the delay and gain of particular directional source to be mixed with speech being acquired by particular microphone.

N→Number of microphones.

K→Number of speech sources to be mixed with speech, being acquired by microphone, where speech sources are target speech, echo (interference) and noise.

s(n)→Spoken speech by speaker

A(d, k)→k^{th} directional source, to be added with s^{th} microphone speech

b(s, n)→Speech to be filtered with coefficients set b, n=1: L coefficients to produce s^{th} directional speech.

Thus, the speech Y_s received by s^{th} microphone is given by

$$Y_s = f(s(n), A(d, k), b(s, n))$$

IV. Methodology

A. Beamforming

The signal obtained in each of the microphone is passed through the subband filter bank. The beamformers are formed by using the FIR adaptive filters, whose coefficients are determined by using the LMS algorithm. The beamformer filter is placed between each of analysis subband filter bank and each of microphone branch. This control the gain of each of the subband output from each microphone branch to be passed through the synthesis filter bank for each of the microphone line. The output of entire synthesis filter bank from each of the microphone line is added to form the reconstructed speech output.

B. Recognition

First of all, the features are extracted from the speech of spoken words. The feature Mel frequency cepstral coefficients (MFCC) have been proved to give better performance in case of speech recognition and hence widely used in speech recognition applications [18,19,20].

In speech processing, the mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of a speech, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency. The recognition process consists of training the classifier and testing the spoken words with trained classifier. The classier used here is nearest neighbour classifier (NN) based on Euclidian distance metric.

IV. Experimental Results

A. Dataset Preparation

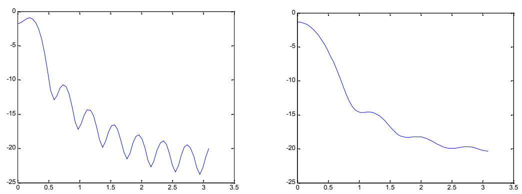
For analyzing the performance of speech recognition, we have considered here four speaker's 20 number of spoken words. These words are listed below and can be categorized on the basis of their use, as numbers and commands. The spoken word from speaker has a length of 2 sec in time. The speech to be used in the experiment is created using multi-microphone mixing environment as described earlier.

TABLE I
List of Spoken Words

Spoken words (each for 2 sec)	
Numbers	**Commands**
one	yes
two	no
three	hello
four	open
five	close
six	start
seven	stop
eight	dial
nine	on
ten	off

B. Experiments

The prototype filter was designed to construct the filter bank. The frequency spectrums of prototype filter for different length of filters and different numbers of sub-bands are shown in figure 6. For training classifier, 2 speaker's spoken words used and for testing we used 4 speakers, wherein 2 speakers are unknown and 2 speakers are same as they were in training phase. Each person (speaker) has 20 spoken words, which includes 10 words for numbers and 10 words for commands as listed in table I.



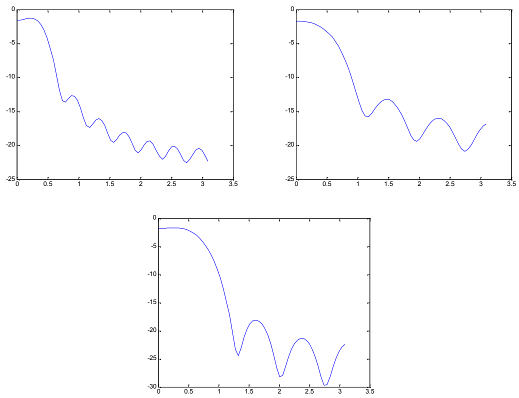


Fig.6. Frequency Response of Prototype filter with different specifications (no of subbands – filter length): Row-1-Col 1) 16 -16; Row-1-Col-2) 16-8; Row-2-Col-1) 8-16; Row-2-Col-2) 8-8 ; Row-3) 4-8

The experiments are performed separately with following class of words:

- Numbers and Commands together (20 words)
- Numbers only (10 words)
- Commands only (10words)

The recognition accuracy is calculated as the ratio of correctly recognized words and total words used for recognition test experiment. We have used MATLAB® environment for performing all experiments.

Table II
Recognition Accuracy With And Without Beamforming For Numbers And Commands Together

Filter length	Number of subbands	Recognition Accuracy in %		
		Pure speech	Multi-Microphone speech	Multi-mic Beamformed speech
16	16	75	27.5	30
8	16	75	27.5	32.5
16	8	75	27.5	31.25
8	8	75	27.5	21.25
8	4	75	27.5	25

TABLE III
Recognition Accuracy with and without Beamforming for Numbers only

Filter length	Number of subbands	Recognition Accuracy in %		
		Pure speech	Multi-Microphone speech	Multi-mic Beamformed speech
16	16	72.5	35	40
8	16	72.5	35	45
16	8	72.5	35	52.5
8	8	72.5	35	27.5
8	4	72.5	35	40

For each class of experiment, the recognition accuracy is calculated in three scenarios. First when pure speech is feed to the recognition experiment without any noise and interference. Secondly, speech was prepared with multi-mic environment with an inclusion of noise and interference (echo). Finally, using beamforming multi-mic speech is enhanced with beamforming-filter bank structure and then fed to the recognition experiment. The last three columns of each of the following tables showing recognition accuracy represents the performance obtained in these three situations. In order to analyze the speech recognition performance against the parameters of filter-bank, we selected the two parameters: filter length and number of subbands in filter bank. The experiments were repeated for different values of these two parameters as mentioned in the first two columns of following tables.

Table IV
recognition accuracy with and without beamforming for commands only

Filter length	Number of subbands	Recognition Accuracy in %		
		Pure speech	Multi-Microphone speech	Multi-mic Beamformed speech
16	16	82.5	50	52.5
8	16	82.5	50	50
16	8	82.5	50	37.5
8	8	82.5	50	62.5
8	4	82.5	50	35

A. Discussion

The recognition accuracy obtained in various experiments is shown in tables II, III and IV. The main objective of the speech enhancement is to bring up the speech recognition performance in the presence of noise and echo-interference to the performance obtained with pure speech signals, which is the ideal case. Thus our aim was to boost up the performance of beamforming based speech enhancement to that in the case of ideal signal. It can be observed that from table II, the speech recognition performance can be improved using the beam forming based speech enhancement. This is also visible in other two experiments, table III and IV, where only numbers and only commands were used for speech recognition. These three cases observations are depicted as below;

1. Numbers+Commands: In the case of numbers and commands together in recognition experiment, accuracy reduces to 27.5% from ideal value 75% due to noise and echo mixing. Using beam forming, the degraded performance can be improved to the optimized performance 32.5%. This is a significant improvement in the recognition accuracy.
2. Numbers: In this case, interference due to noise and echo mixing causes decrease in recognition performance to 35% from ideal value 72.5%. Using beam forming, the degraded performance can be boost up to the optimized performance of 52.5%. This is a much significant improvement in the recognition accuracy.
3. Commands: In this case, due to noise and echo mixing, recognition performance degrades to 50% from ideal value 82.5%. Using beam forming, the degraded performance can be brought up to the optimized performance 62.5%. This is a much significant improvement in the recognition accuracy.

These results are in consistent with the fact that proper beamforming can improve the recognition performance. Since, this improvement can be achieved with less computational parameters of sub-band filtering; this technique is suitable for real-time application of speech recognition.

The performance is also dependent on the parameters of filter bank used for sub-band filtering. Another objective of this work is to analyse the effect of the filter-bank parameters on the speech recognition. It can be seen easily that there is an undesired effect of improper selection of the parameters like filter length and number of subbands on the recognition accuracy. In general, more the number of subbands, more parallelism can be achieved by the system and larger the filter length, better the frequency response but with higher computational complexity. Thus, it is important to design the system with proper selection of these parameters so that system yield can be improved. For the all word experiments, it can be seen that best performance we get, more precisely, is with filter length 8 and number of subbands 16. However, roughly, it can be seen that the filter length and number of subbands with values 8 or 16 giving better results. The same conclusion can be inferred from other two experiments also.

Another important point that can be observed here is that for numbers type speech and commands type of speech, different parameters of sub-band filtering are required. This is due to the fact that numbers are normally pronounced with short duration support and commands are comparatively long duration speech words. Hence, having optimized parameters in general application is required. The highest performance in both individual experiments (table III and IV), has different lengths, 16 and 8. However, second highest performance with beamforming is with filter length 8 and number of sub-bands 16, in both types of experiments. Thus these parameters can be selected as optimized parameters for general application.

V. Conclusion

The main objective of the speech enhancement is to bring up the speech recognition performance in the presence of noise and echo-interference to the performance obtained with pure speech signals, which is the ideal case. Thus, our aim was to approach the performance of beamforming based speech enhancement to that in the case of ideal signal. Another objective of this work is to analyze the effect of the filter-bank parameters on the speech recognition. It can be seen easily that there is an undesired effect of improper selection of the parameters like filter length and number of subbands on the recognition accuracy. The experiments were performed for 20 words including numbers and commands, 10 words of numbers only and 10 words of commands only for different values of filter bank parameters. The results obtained have proved the speech enhancing capability of the beamforming technique in multi-microphone network where noise and echo-interference can degrade the original speech signal.

Acknowledgement

Milind U. Nemade thanks to Dr. S. G. Kirloskar, Principal and Prof. ArchanaBhise, Vice Principal K. J. Somaiya Institute of Engineering and Information Technology Sion, Mumbai for their valuable guidance and support.

References

[1] O. L. Frost, III, “An algorithm for linearly constrained adaptive array processing,” *Proc. IEEE*, vol. 60, pp. 926–935, Jan. 1972.

[2] Griffiths, L.; Jim, C., "An alternative approach to linearly constrained adaptive beamforming,", *IEEE Transactions on Antennas and Propagation*, vol.30, no.1, pp. 27- 34, Jan 1982.

[3] Van Compernolle, D. "Switching adaptive filters for enhancing noisy and reverberant speech from microphone array recordings," *International Conference on Acoustics, Speech, and Signal Processing*, ICASSP-90., vol., no., pp.833-836 vol.2, 3-6 Apr 1990.

[4] Gannot, S.; Burshtein, D.; Weinstein, E., "Signal enhancement using beamforming and nonstationarity with applications to speech," *IEEE Transactions on Signal Processing*, vol.49, no.8, pp.1614-1626, Aug 2001.

[5] Grbic, N.; Nordholm, S.; Cantoni, A., "Optimal FIR subband beamforming for speech enhancement in multipath environments," *IEEE Signal Processing Letters*, vol.10, no.11, pp. 335- 338, Nov. 2003.

[6] Xianxian Zhang; Hansen, J.H.L., "CSA-BF: a constrained switched adaptive beamformer for speech enhancement and recognition in real car environments," *IEEE Transactions on Speech and Audio Processing*, vol.11, no.6, pp. 733- 745, Nov. 2003.

[7] Seltzer, M.L.; Raj, B.; Stern, R.M., "Likelihood-maximizing beamforming for robust hands-free speech recognition," *IEEE Transactions on Speech and Audio Processing*, vol.12, no.5, pp. 489- 498, Sept. 2004.

[8] Yermeche, Z.; Grbic, N.; Claesson, I., "Blind Subband Beamforming with Time-Delay Constraints for Moving Source Speech Enhancement," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.15, no.8, pp.2360-2372, Nov. 2007.

[9] Guangji Shi; Parham Aarabi; Hui Jiang, "Phase-Based Dual-Microphone Speech Enhancement Using A Prior Speech Model," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.15, no.1, pp.109-118, Jan. 2007.

[10] Maganti, H.K.; Gatica-Perez, D.; McCowan, I., "Speech Enhancement and Recognition in Meetings with an Audio–Visual Sensor Array," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.15, no.8, pp.2257-2269, Nov. 2007.

[11] Hai Huyen Dam; Hai Quang Dam; Nordholm, S., "Noise Statistics Update Adaptive Beamformer with PSD Estimation for Speech Extraction in Noisy Environment," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.16, no.8, pp.1633-1641, Nov. 2008.

[12] Han, S.; Hong, J.; Jeong, S.; Hahn, M., "Robust GSC-based speech enhancement for human machine interface," *IEEE Transactions on Consumer Electronics*, vol.56, no.2, pp.965-970, May 2010.

[13] John J. Shynk, “Frequency-domain and multirate adaptive filtering,” *IEEE Signal Processing Magazine*, vol. 9, pp. 14–37, 1992.

[14] Jan Mark de Haan, Nedelko Grbic, Ingvar Claesson, and Sven Erik Nordholm, “Filter bank design for subband adaptive microphone arrays,” *IEEE Trans. Speech and Audio Proc.*, vol. 11, no. 1, pp. 14–23, Jan. 2003.

[15] Kumatani, K.; McDonough, J.; Schachl, S.; Klakow, D.; Garner, P.N.; Weifeng Li, "Filter bank design based on minimization of individual aliasing terms for minimum mutual information subband adaptive beamforming," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008. ICASSP 2008, vol., no., pp.1609-1612, March 31 2008-April 4 2008.

[16] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, 1993.

[17] Kenichi Kumatani, Tobias Gehrig, Uwe Mayer, Emilian Stoimenov, John McDonough, and MatthiasW¨olfel, “Adaptive beamforming with a minimum mutual information criterion,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 8, pp. 2527–2541, 2007.

[18] L. Rabiner and Biing-Hwang Juang, *Fundamentals of Speech Recognition*, Prentice Hall PTR, 1993.

[19] Joseph W. Picone, “Signal Modeling Techniques in Speech Recognition”, *Proceedings of the IEEE*, vol. 81, No. 9, pages 1215--1247, 1993.

[20] Steven B. Davis and Paul Mermelstein, “Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-28, No. 4, August 1980.

AUTHOR’S PROFILE

Milind U. Nemade

was born in Maharashtra, India 1974. He graduated from the Amaravati University, Maharashtra, India in 1995. He received M.E (Electrical) degree with specialization in Microprocessor Applications from M.S. University of Baroda, Gujrat, India in 1999. Now he is Associate Professor and Head at Department of Electronics and Telecommunication, K.J. Somaiya Institute of Engineering and Information Technology Sion, Mumbai, University of Mumbai, India. He started PhD study at Electrical Department, Faculty of Technology and Engineering, M.S. University of Baroda, Gujrat, India. He presented and published four papers in national conferences and three papers in the proceedings of international conferences. His research interest includes speech and audio processing.

Email_id : mnemade@tech.somaiya.edu,munemade@gmail.com

Prof. Satish Shah

is a professor in the Electrical Engineering Department at Faculty of Technology, M S University of Baroda for **last Twenty Five** years. He is a fellow of IE(I) for past Fourteen Years and has also served as the member of Committee of Vadodara local center for more than SIX years in past.

He has guided more than hundred projects at UG/PG level and completed a research project on DSP based Active Power filter sponsored by AICTE, New Delhi. He has written three books on Embedded System design/ Microprocessors/ Microcontrollers and presented/ published more than 25 research papers in national/international conferences/Journals. He has attended and organized several seminars, workshops, and symposiums for UGC, AICTE, IETE, and MSU.

He is a fellow of other technical associations such as: IETE, ISA and IEEE (NY) & ISTE. He has served as the member, Hon Secretary and Treasurer of their local executive committees for a span of six-eight years.

Email_id : satishkshah_2005@yahoo.com