Role of Spectral Peaks in Autocorrelation Domain for Robust Speech Recognition

Poonam Bansal[1], Amita Dev[2] and Shail Bala Jain[3]

[1]Department of Computer Science and Engineering, Amity School of Engineering and Technology, New Delhi, India
[2]Department of Computer Science and Engineering, Ambedkar Institute of Technology, New Delhi, India
[3]Department of Electronics and Communication Engineering, Indira Gandhi Institute of Technology,
Guru Gobind Singh Indraprastha University, New Delhi, India

This paper presents a new front-end for robust speech recognition. This new front-end scenario focuses on the spectral features of the filtered speech signals in the autocorrelation domain. The autocorrelation domain is well known for its pole preserving and noise separation properties. In this paper, a novel method for robust speech extraction is proposed in the autocorrelation domain. The proposed method is based on a novel representation of the speech signal corrupted by an additive noise. Initial filtering stage is used to reduce the additive noise when computing the speech features followed by extraction of the autocorrelation spectrum peaks. Robust features based on these peaks are derived by assuming that the corrupting noise is stationary in nature. A task of speaker-independent isolated-word recognition is used to demonstrate the efficiency of these robust features. The cases of white noise and colored noise such as factory, babble and F16 are tested. Experimental results show significant improvement in comparison to the results obtained using traditional front-end methods. Further enhancement has been done by applying cepstral mean normalization (CMN) to the above extracted features.

Keywords: robust speech recognition, autocorrelation domain, spectral peaks

1. Introduction

Noise robustness is one of the most challenging problems in automatic speech recognition. The performance of automatic speech recognition (ASR) systems, trained with clean speech, may drastically degrade in realistic environments. The main reason for this degradation is the acoustic mismatch between the training and testing environments due to environmental effects. Many techniques have been proposed to overcome this degradation problem [1]. The robustness of the speech recognizer can be accomplished in three ways: (1) Using speech enhancement technique to increase the signal to noise ratio (SNR) [2] (2) extracting the robust parametric representation of speech signal to minimize the effect of noise [3,4] and (3) using model adaptation technique to dynamically adapt the speech recognition models to noisy speech [5,8]. The goal of robust speech recognition is to improve the performance of speech recognition in such adverse environments. The environmental effects are often determined by noise and channel distortion. The noise is usually additive in the spectral domain while channel distortion is multiplicative. One approach to attack the effects of environment is to have a separate training set for each noise type. However, this approach is not practical due to large diversity of noise types encountered in real environments. In order to remove the effect of noise, some methods have been found useful and extensively studied in the literature, such as spectral subtraction (SS), Non-linear spectral subtraction (NSS) [9], Linear-logarithmic RelAtive SpecTrA (*Lin-log RASTA)* [3], Differentiated Power Spectrum (DPS) [10], Parallel Model Combination (PMC) [5] and Minimum Variance Distortionless Response (MVDR) [11]. Furthermore, various techniques have been developed to suppress the channel distortion such as Cepstral Mean Normalization (CMN), RelAtive SpecTrAl (RASTA) [3] and Blind Equalization (BE) [12].

Although these techniques demonstrate the comparable performance, some weaknesses may restrict their practical applications. The SS method suppress the background noises by first estimating non-speech intervals followed by their subtraction from the noisy speech. Recently, the parameters extracted using autocorrelation sequence of the noisy signal have been found useful for robust speech recognition. Some examples include magnitude spectrum of higher lag autocorrelation coefficients [13] whereas other methods are based on Relative Autocorrelation Sequence (RAS) method [14,15]. Furthermore, according to [16], preserving spectral peaks is very important in obtaining a robust set of features in the speech recognition.

In this paper we propose a new approach, utilizing peaks obtained from the autocorrelation spectrum of the speech signal. This approach preserves the autocorrelation spectral peaks. We propose the following front-end description for robust feature extraction. Firstly, we calculate the autocorrelation of the noisy signal. As the temporal autocorrelation of noise is a DC or slowly varying signal, its effect is suppressed by a high-pass filter, such as RAS filter [15]. Then, following the DPS concept, relative autocorrelation sequence spectrum is computed from RAS and differentiated with respect to frequency. This new spectrum is referred to as differentiated relative autocorrelation sequence spectrum (DRASS). Finally, from the magnitude of DRASS, the mel-scale frequency cepstral coefficients (DRASS-MFCC) are derived by passing the DRASS through a mel-frequency filter-bank followed by the discrete cosine transform (DCT).

We can exploit properties of CMN, and combine the proposed robust features DRASS-MFCC with the CMN to further improve the recognition accuracy. The features based on our method are showing remarkable increase in recognition accuracy, especially in noisy environments.

The remainder of this paper is organized as follows. Mathematical fundamentals for extracting RAS and DRASS are derived in Section 2. Proposed method with front-end diagram is described in Section 3. In Section 4. experiments conducted in clean and noisy environment with different front-ends are discussed and compared with the proposed front-end method. Finally a conclusion is given in Section 5.

2. Robust Features Extraction

2.1. Calculation of the Autocorrelation for Noisy Signal

Any noisy speech signal $y(m, n)$ can be written as

$$y(m, n) = x(m, n) + u(m, n),$$
$$0 \leq m \leq M - 1, \quad 0 \leq n \leq N - 1 \quad (1)$$

where $u(m, n)$ is the additive noise, $x(m, n)$ noise-free speech signal, N is the frame length (in our case 256 samples), n is the discrete time index in a frame, m is the frame index and M is the number of frames.

If $x(m, n)$ and $u(m, n)$ are assumed uncorrelated, then the autocorrelation of the noisy speech can be expressed as

$$r_{yy}(m, k) = r_{xx}(m, k) + r_{uu}(m, k),$$
$$0 \leq m \leq M - 1, \quad 0 \leq k \leq N - 1 \quad (2)$$

where $r_{yy}(m, k)$, $r_{xx}(m, k)$ and $r_{uu}(m, k)$ are the short-time autocorrelation sequences of the noisy speech, clean speech and noise, respectively, and k is the autocorrelation sequence index within each frame.

The one-sided autocorrelation sequence of each frame can be calculated using a biased estimator as:

$$r_{yy}(m, k) = \sum_{i=0}^{N-1-k} y(m, i) y(m, i + k),$$
$$0 \leq k \leq N - 1 \quad (3)$$

Since additive noise is assumed to be stationary, its autocorrelation sequence can be considered to be identical for all frames. Hence the frame index, m, can be dropped out, and (2) becomes

$$r_{yy}(m, k) = r_{xx}(m, k) + r_{uu}(k),$$
$$0 \leq m \leq M - 1, \quad 0 \leq k \leq N - 1 \quad (4)$$

2.2. Filtering of One-sided Autocorrelation Sequence

As described above, our aim is to reduce the effect of additive noise on the extraction of parameters of the speech signal. If the autocorrelation trajectory of noise is DC or slowly varying,

the noise effect is suppressed by calculating the Relative Autocorrelation Sequence (RAS). The RAS of the noisy speech at the mth index can be found by applying a high pass filtering in the time domain to reduce the effects of additive noises on signal (4) for all indices m and k. Hence consider the differences

$$\Delta r_{yy}(m, k) = \Delta r_{xx}(m, k),$$
$$0 \leq m \leq M - 1, \quad 0 \leq k \leq N - 1 \quad (5)$$

where

$$\Delta r_{yy}(m, k) = r_{yy}(m + 1, k), r_{yy}(m - 1, k)$$

and

$$\Delta r_{xx}(m, k) = r_{xx}(m + 1, k), r_{xx}(m - 1, k)$$

The sequence, $\{\Delta r_{yy}(m, k), 0 \leq k \leq N - 1\}$, is referred to as the RAS of the noisy speech at the mth frame. Denote the output of this temporal

filtering in the time domain as $Z(m, n)$, we can write

$$z(m, n) = x(m, n) + v(m, n),$$
$$0 \leq m \leq M - 1, \quad 0 \leq n \leq N - 1 \quad (6)$$

where $x(m, n)$ and $v(m, n)$ are the clean speech and the remaining noise after filtering, respectively.

2.3. Extraction of DRASS

The autocorrelation function of the filtered speech signal given in (6) can be written as

$$r_{zz}(m, k) = r_{xx}(m, k) + r_{vv}(m, k),$$
$$0 \leq k \leq N - 1 \quad (7)$$

Provided that the noise and the clean speech signal are assumed mutually uncorrelated, we
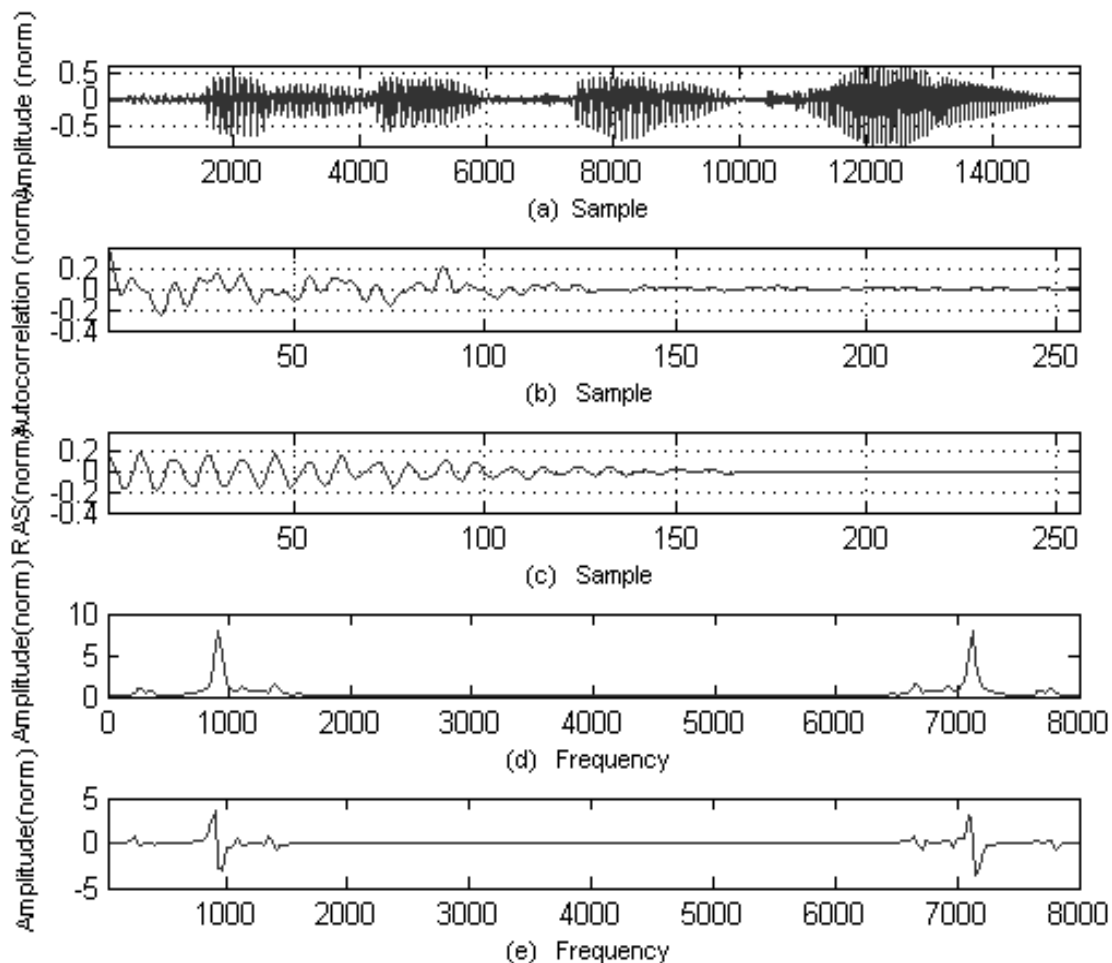

Figure 1. (a) A sample speech signal (b) Autocorrelation magnitude of a 16 msec. frame (c) Relative autocorrelation magnitude of a 16 msec. frame (d) Relative autocorrelation power spectrum of 16 msec. frame (The power spectrum is estimated through 256- point FFT) and (e) Differentiated relative autocorrelation spectrum.
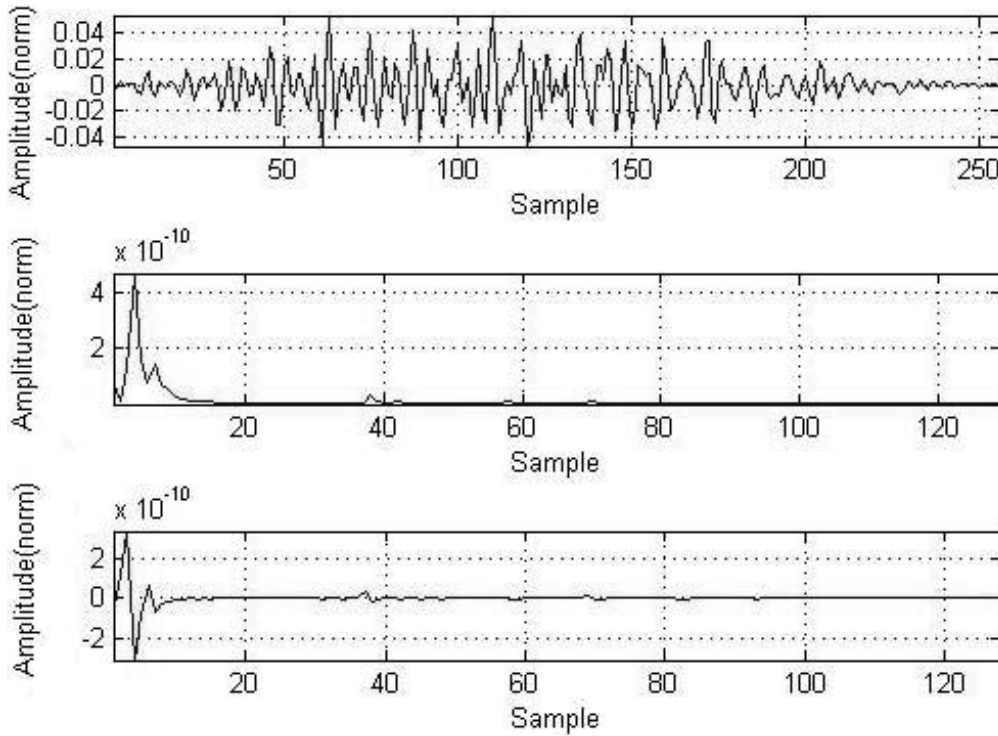
Figure 2. (a) A sample speech signal, and (b) the autocorrelation spectrum magnitude and the differentiated autocorrelation spectrum magnitude of the same signal with a 256-point FFT. For clarity, only 128 points of the spectrum are shown.

can calculate the relationship between relative autocorrelation power spectrum of noisy speech signal, clean speech signal and noise by using short time DFT to both sides of equation (7), as follows:

$$\text{DTF}\{r_{zz}(m,k)\} = \text{DTF}\{r_{xx}(m,k)\} \\ + \text{DTF}\{r_{vv}(k)\}, \quad \text{or} \quad (8) \\ Z(\omega) = X(\omega) + V(\omega)$$

Where $\text{DTF}[\ ]$ denotes the discrete Fourier transform, and ω is the frequency in radians. Further enhancement can be obtained by computing differentiation of the relative autocorrelation power spectrum for extracting spectral peaks. The differential power spectrum is defined as:

$$D_Z(\omega) = \frac{dZ(\omega)}{d\omega} = \frac{dX(\omega)}{d\omega} + \frac{dV(\omega)}{d\omega} \\ = D_X(\omega) + D_V(\omega) \quad (9)$$

Where $D_X(\omega)$ and $D_V(\omega)$ are the differential relative autocorrelation power spectra of the clean speech signal and the noise respectively. $D_Z(\omega)$ is referred to as the Differentiated Relative Autocorrelation Sequence Spectrum (DRASS) of noisy speech.

This approach combines advantages of RAS and DPS. Figure 1(a) depicts a sample speech signal, its autocorrelation magnitude (b) and relative autocorrelation magnitude (c) for a 16 msec. frame. Figure 1 clearly shows that initial filtering in terms of RAS retains the signal information and filters out the additive noise at primary level. Figure 1(d) and 1(e) show relative autocorrelation spectrum and the differentiated relative autocorrelation spectrum estimated by the 256-point DFT. As shown in the Figure 1 and Figure 2, by differentiation the flat parts of the spectrum are transformed to values that are approximately zero. The spectral peaks are preserved, except each peak is split into the positive and negative parts. The spectral peaks that convey the most important information in the speech signal are preserved, and thus, no important information is lost in the process.

3. Description of the Proposed Method

In this section, we describe the proposed method for robust feature extraction for the speech recognition. First, the speech signal is divided into

frames of 256 samples each, and a pre-emphasis filter is applied on each frame. A Hamming window is used, and then, the autocorrelation sequence of the frame signal is obtained using a biased estimator, as given by eq. (4). A temporal filtering is then applied to the autocorrelation sequence to obtain the relative autocorrelation sequence (RAS) in order to suppress the additive noise. When speech is corrupted by an additive noise, the noise component is additive to the speech not only in the autocorrelation domain, but also in the power spectrum domain. In the next step, differentiated relative autocorrelation sequence spectrum (DRASS) is calculated by taking the differentiation of the power spectrum of the relative autocorrelation coefficients (RAS). Differentiation of the autocorrelation spectrum of the noisy speech signal preserves the spectral peaks, however, each peak has now positive and negative parts. In addition, the flat part of the power spectrum is approximately zero. Figure 2 depicts a sample speech signal, its short-time autocorrelation spectrum and the differentiated short-time autocorrelation spectrum. This sample signal corresponds to one frame of sample speech. In order to simplify the spectrum representation, only the significant lower-frequency parts of the spectrum have been shown and the non-significant parts have been omitted. As shown in Figure 2 and as mentioned above, the flat parts of the filtered autocorrelation spectrum have been transformed to zero by differentiation and each peak has two

parts, positive and negative. Since the spectral peaks convey the most important information in the speech signal, this fact that the differential power spectrum retains spectral peaks means that we will not lose the important information of the speech signal. Furthermore, since the noise spectrum is often flat and the differentiation either reduces or omits the relatively flat parts of the spectrum, it will lead to omission of the effect of the noise on the signal leading to more robust features. A set of robust mel-frequency cepstral coefficients (DRASS-MFCC) can be derived from the magnitude of the differentiated relative autocorrelation power spectrum by applying it to a conventional mel-frequency filter-bank and finally passing its logarithm to the DCT block. MFCC feature vector set of dimension 39 is formed by concatenating energy feature, Delta MFCC and Delta-Delta MFCC. CMN (Cepstral mean normalization) is used for removing the effect of channel distortion on robust feature vector set. Figure 3 displays the proposed front-end diagram of our method.

4. Experiments

A digital database of 200 Hindi words spoken by 30 speakers (Table 1) has been used for the experiment of speaker- independent isolated word recognition system.
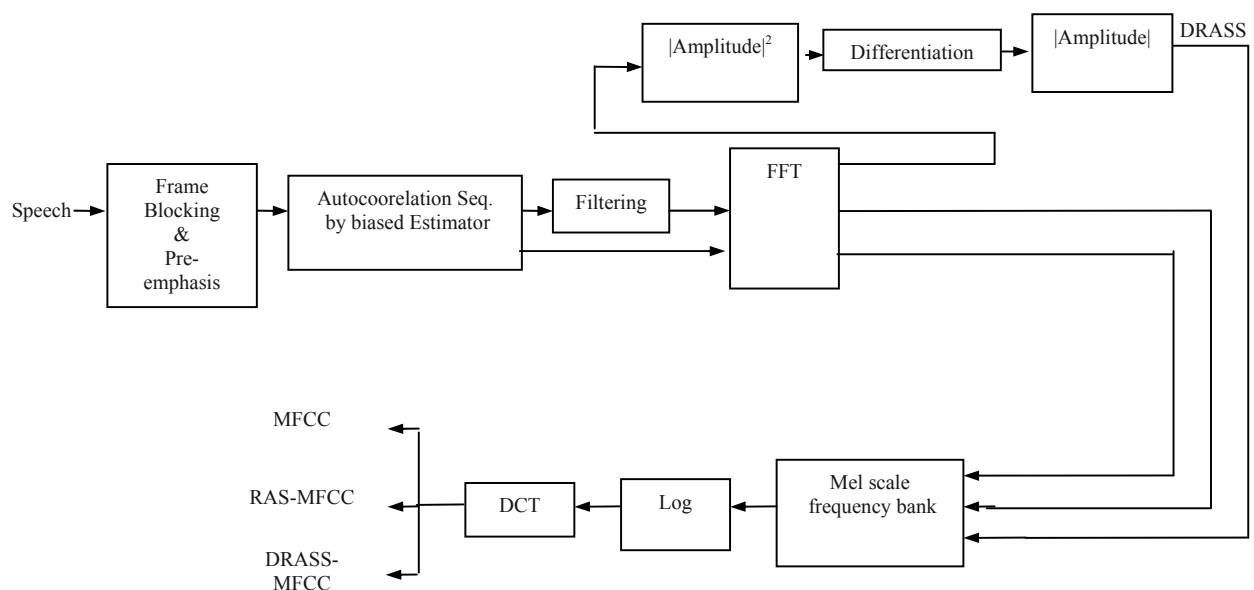

Figure 3. A Front end diagram for feature extraction.

The spoken samples are recorded by 15 male, 10 female and 5 child speakers in the studio environment using the Sennheiser microphone model MD421 and a tape recorder model Philips AF6121. Each speaker pronounced 5 repetitions of words. The resulting database was partitioned for the use of training and testing. 2500 utterances from 20 speakers were used for training of the HMM model. The test set also contains similar data from 10 speakers that were not included in the training set.

1. Language	: Standard Hindi (Khari Boli)
2. Vocabulary Size	: A set of 200 most frequently occurring Hindi words
3. Speakers	: 30 speakers
4. Utterances	: (15 male, 15 female and 5 children) 5 repetitions each
4. Audio Recording	: Recording on a cassette tape in studio SNR>40dB
5. Digitization	: 16KHz. Sampling, 16 bit quantization.

Table 1. Hindi speech database for a vocabulary of 200 words used in the experiment.

The features in all the cases of training and testing are computed using 16 millisecond frames with 8 millisecond frame shifts. Pre-emphasis coefficient used is 0.9375. For each speech frame, a 20-channel mel-scale filter bank is used. Optimum seven state Hidden Markov Model [17] has been used for training and testing. Features vector sets of size 39 are extracted using different front-ends: MFCC (for comparison purposes), RAS-MFCC and our method DRASS-MFCC. With these features vector sets, the word models of training database for different front-ends are created. Afterwards, word recognition rates for testing database are computed with all the above front-ends and compared with the traditional MFCC. Another comparison is made by enhancing the above features with cepstral mean normalization (CMN)

Testing on clean speech

The purpose of this experiment is to evaluate the performance of MFCC, RAS-MFCC and DRASS-MFCC, when training data and the testing data are in a clean environment, i.e., assuming 40 dB signal to noise ratio (SNR). The results in Table 2 are the baseline results for comparison purposes. We observe that

the recognition rates are approximately identical for MFCC and RAS-MFCC. While using DRASS-MFCC, the recognition rate is improved to 99.6% when compared to 98.24% of RAS-MFCC. As the mel frequency cepstral coefficients are enhanced by CMN, the word recognition rates always show an improvement. It is true for all the cases (MFCC, RAS-MFCC and DRASS-MFCC). This shows that the cepstral mean normalization is useful in overcoming the other distortions in speech signal effectively.

Feature type	Recognition rate (%)
MFCC	98.241
MFCC_CMN	99.270
RAS-MFCC	98.246
RAS-MFCC_CMN	99.324
DRASS-MFCC	99.642
DRASS-MFCC_CMN	99.842

Table 2. Comparison of clean-train and clean test recognition rates for various features.

Testing on noisy speech

The polluted testing utterances are generated by adding the artificial noises at five SNR levels. The white noise is generated by a random number generator program, and other colored noises such as factory noise, F16 noise, and babble noise, are extracted from the NATO RSG-10 corpus [18]. The noises are added to the clean speech signal at 20, 15, 10, 5 and 0 dB of SNR. Both RAS-MFCC and DRASS-MFCC are evaluated and the word recognition rates are compared with the traditional MFCC front-end. Another comparison is made by taking into consideration the effect of cepstral mean normalization. Figures 4(a)-(d) show the results obtained using MFCC, RAS and DRASS front-ends respectively. For the case of white noise in Figure 4(a), the performance of MFCC degrades most significantly among all features. Although the MFCC with cepstral mean normalization can make some improvement, its performance is still worse than RAS-MFCC and DRASS-MFCC. It is obvious that RAS-MFCC and DRASS-MFCC are quite robust to the additive noises. Their performance has been further improved by combining it with the cepstral mean normalization.
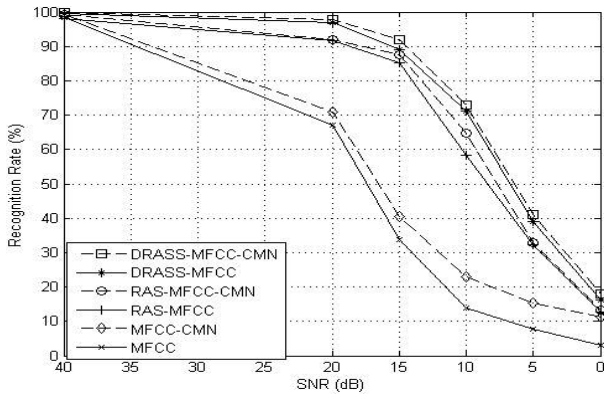
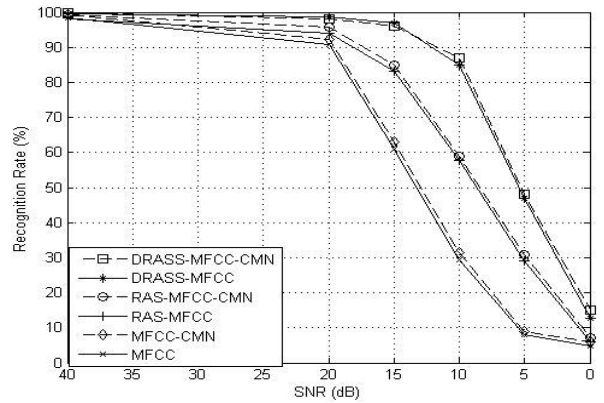*Figure 4a.* Recognition rate (%) for testing speech corrupted by a white noise.

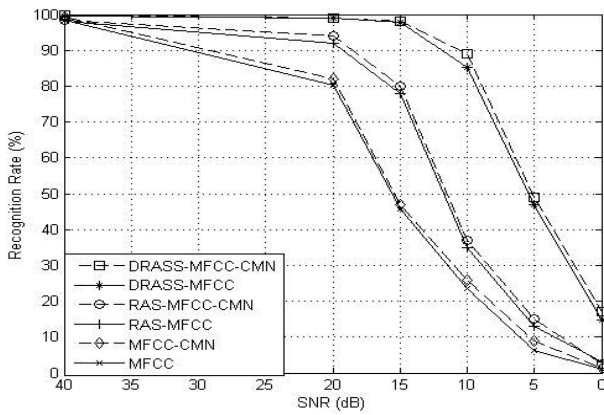*Figure 4b.* Recognition rate (%) for testing speech corrupted by a factory noise.

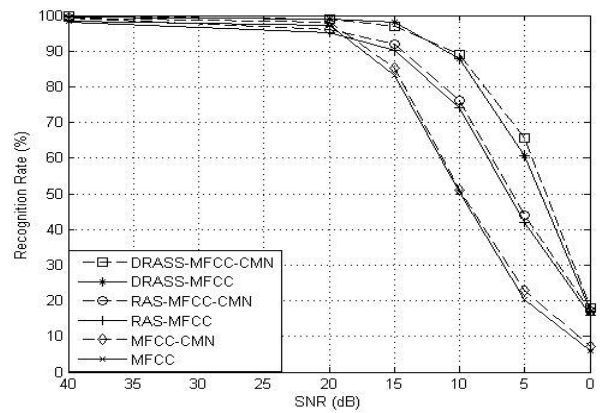*Figure 4c.* Recognition rate (%) for testing speech corrupted by F16 noise.

Figure 4d. Recognition rate (%) for testing speech corrupted by a babble noise.

Figures 4(b), (c) and (d), show the performance when the testing speech is corrupted by factory, F16, and babble noises, respectively. These figures indicate that the performance of MFCC degrades significantly. The best performance comes from DRASS-MFCC combined with cepstral mean normalization. Generally, the DRASS-MFCC with CMN is better than RAS-MFCC with CMN in all the cases. This is due to the peak preserving property in the power spectrum which helps the speech recognition in noisy environment. The normalization in the frequency domain also provides compensation of the additive noise, when the testing speech is corrupted by a colored noise. The average rates of word recognition are computed with different types of front-ends (MFCC, RAS-MFCC, DRASS- MFCC) at different levels of noises and are displayed in Table 3.

Feature type	Noise levels (dB)					
	40	20	15	10	5	0
MFCC	98.24%	83.88%	55.87%	29.38%	10.5%	3.7%
MFCC_CMN	99.27%	85.81%	58.97%	32.81%	14.05%	6.40%
RAS-MFCC	98.24%	93.19%	84.13%	56.31%	29.11%	9.35%
RAS-MFCC_CMN	99.32%	94.52%	86.11%	59.20%	30.71%	9.97%
DRASS-MFCC	99.64%	98.39%	95.43%	82.26%	48.42%	15.08%
DRASS-MFCC_CMN	99.84%	98.01%	95.19%	84.48%	51.02%	16.96%

Table 3. Average recognition rates (%) for various feature types with different SNR levels .

5. Conclusion

In this paper, cepstral features derived from autocorrelation spectral domain are proposed in order to improve the robustness of the speech recognition systems. The concept of DRASS introduces a new set of cepstral features, exploiting the benefits of RAS and DPS, for improving the robustness of speech recognition. In the case of noisy environment, CMN contributes to better performance in terms of speech recognition. Experimental results show that the proposed approach DRASS-MFCC with CMN is more effective in overcoming additive noises which are stationary in nature as well as it allows to reduce the effects of channel distortion at low SNR's. Since the proposed method eliminates additive noise in two stages, it outperforms the well known RAS approach. The use of DRASS can significantly increase the word recognition rates especially at low SNR, for example, by 9.08% at 15dB, 3.49% at 20 dB compared to RAS. Furthermore, this proposed method works well for different types of noises including white, babble and factory noise. Further improvement of speech recognition by feature extraction will be considered in future work since it is likely to have great influence on the ASR system performance.

References

[1] Y. GONG, Speech recognition in noisy environments: A survey. *Speech Communication*, Vol. 16 (1995), pp. 261–291.

[2] S. F. BOLL, Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustic Speech and Signal Processing*, 27 (2), (1979), pp. 113–120.

[3] HERMANSKY AND MORGAN, RASTA processing of speech. *Speech Communication*, Vol. 41, (2003), pp. 469–484.

[4] J. HERNANDO AND C. NADEU, Linear prediction of the one-sided autocorrelation sequence for noisy speech recognition. *IEEE Trans. Speech Audio Processing*, Vol. 2, No. 5, (1994) pp. 578-586.

[5] M. J. F. GALES AND S. J. YOUNG, Robust speech recognition in additive and convolutional noise using parallel model combination, *Comput. Speech Lang.*, Vol. 9, (1995), pp. 289–307.

[6] M. J. F. GALES AND S. J. YOUNG, Robust speech recognition in additive and convolutional noise using parallel model speech recognition. *Proceedings of ICASSP*, (2003), pp. 644–647.

[7] D. MANSOUR AND B. H. JUANG, A family of distortion measures based upon projection operation for robust speech recognition. *IEEE Trans. Acoust. Speech and Signal Processing*, Vol. 37, No. 11, (1989) pp. 1695–1671.

[8] B. A. CARLSON AND M. A. CLEMENTS, A projection based likelihood measure for speech recognition in noise. *IEEE Trans. Acoust. Speech and Signal Processing*, Vol. 2, No. 1, part 2, (1994), pp. 97–102.

[9] MICHAEL J. CAREY, Robust speech recognition using non-linear spectral smoothing. Eurospeech'03, (2003), pp. 3045–3048.

[10] J. CHEN, K. K. PALIWAL AND S. MAKAMURA, Cepstrum derived from differentiated power spectrum for robust speech recognition. *Speech Communication*, Vol. 41, (2003), pp. 469–484.

[11] U. H. YAPANEL AND S. DHARANIPRAGADA, Perceptual MVDR-based cepstral coefficients (PMCCs) for noise robust speech recognition. *Proceedings of ICASSP*, (2003), pp. 644–647.

[12] L. MAUUARY, Blind equalization for robust telephone-based speech recognition, *Proc. of European Signal Processing Conference*, (1998).

[13] B. SHANNON AND K. K. PALIWAL, Feature extraction from higher-lag autocorrelation coefficients for robust speech recognition. *Speech Communication*, Vol. 48, No. 11, (2006), pp. 1458–1485.

[14] K.-H. YOU, T.-H. HWANG AND H.-C. WANG, Robust features for noisy speech recognition based on temporal trajectory filtering of short-time autocorrelation sequences. *Speech Communication*, Vol. 28, (1999), pp. 13–24.

[15] K. H. YUO AND H. C. WANG, Robust features derived from temporal trajectory filtering for speech recognition under the corruption of additive and convolutional noises, in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP 98)*, Seattle, Washington, pp. I-577–580.

[16] B. STROPE AND A. ALWAN, Robust word recognition using threaded spectral peaks, in *Proc. of ICASSP'98*, pp. 625–628.

[17] P. BANSAL, A. DEV AND S. B. JAIN, Optimum HMM combined with vector quantization for Hindi speech word recognition. *IETE Journal of research*, Vol. 54, 4, (2008), pp. 239–243.

[18] A. VARGA AND H. J. M. STEENEKEN, Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems. *Speech Communication*, Vol. 12, (1993), pp. 247–251.

Contact addresses:
Poonam Bansal
Amity School of Engineering and Technology
Guru Gobind Singh Indraprastha University
New Delhi, India
e-mail: pbansal89@yahoo.co.in

Amita Dev
Ambedkar Institute of Technology
Shakarpur, New Delhi, India
e-mail: amita_dev@hotmail.com

Shail Bala Jain
Indira Gandhi Institute of Technology
Guru Gobind Singh Indraprastha University
New Delhi, India
e-mail: shailbala.jain@gmail.com

POONAM BANSAL obtained her B. Tech. degree form Delhi College of Engineering and completed her M. Tech. in Computer Science and Engineering from Delhi College of Engineering, Delhi University, in 2001. Presently she is assistant professor in the Computer Science and Engineering department of Amity school of Engineering and Technology, a premier Institute under GGSIP University. She is pursuing her Ph.D. in the area of speech recognition from Guru Gobind Singh Inderprastha University, New Delhi, India. She has published various papers in renowned national and international journals and in the proceedings of leading conferences. Her research interests include speech recognition and speaker identification. She is a member of IEEE, IETE, IEICE and Acoustic society of India.

AMITA DEV has obtained her B. Tech. degree from Punjab University and completed her post graduation at BITS, Pilani. She has completed her PhD. at Delhi University in the area of speech recognition. She has more than twenty two years of working experience and presently she is working as Principal of Ambedkar Polytechnic, New Delhi and Bhai Parmanand Institute of Business Studies, New Delhi. She has been awarded "National Level Best Engineering Teacher Award" in 2001 by ISTE, for significant contribution in the field of engineering and technology. She has also been awarded "State Level Best Teacher Award" by Department of Training and Technical Education. Govt. of Delhi. She is recipient of "National Level AICTE Young Teacher Career Award" for pursuing advance research in the field of speech recognition. She has written several text and reference books in the area of Computer Science and Engineering and published more than 20 papers in the renowned national and international journals and in the proceedings of leading conferences.

SHAIL BALA JAIN obtained M. Tech. in 1970 in electronics and communication from Indian Institute of Technology, Delhi and Ph.D. in the field of digital signal processing from Delhi University, New Delhi India. She has thirty one years of teaching and research experience at Delhi College of Engineering, New Delhi, India. Since 2002, she is Head of Electronics and Communication Engineering Department of "Indira Gandhi Institute of Technology", a premier Institute under Guru Gobind Singh Indraprastha University. She has several research publications and is co-author of two books. Her fields of interest are digital signal processing and linear integrated circuits. She is senior member of IEEE and fellow member of IETE.