

Enhancement of Speech Signal Using Improved Minimum Controlled Recursive Average & Principal Component Analysis

Sangita Bavkar, Shashikant Sahare

Abstract— In this paper we present a speech enhancement algorithm for noisy speech signal. A worldwide subspace approach is used for enhancement of speech corrupted by noise. The proposed approach is based on the instantaneous diagonalisation of the clean speech and noise covariance matrices. The system designed in this paper takes the PCA method as its basis and used a strong and accurate noise estimation algorithm that can update the noise variance. Objective measures demonstrated significant improvements over other subspace-based methods when tested with sentences corrupted with white Gaussian Noise.

Index Terms— Eigenvalue Analysis, Improved Minimum Controlled Recursive Average (IMCRA), Noise Estimation, Objective Evaluation , Principal Component Analysis(PCA),Speech Enhancement, Subspace Signal.

1 INTRODUCTION

THE Communication can be greatly slowed down by noise. In different speech processing system speech enhancement is used as preprocessor block. Speech enhancement seeks to eliminate noise in a variety of environments, the most famous of which are telecommunications applications. After the lots of research work there is no perfect solution exists to speech enhancement problem.

All speech enhancement systems suffer from distortion or residual noise due to imperfect noise estimation. The system designed in this paper takes the subspace method as its basis and develops a robust and accurate noise estimation algorithm.

One particular class of speech enhancement techniques that has gained a lot of attention is signal subspace filtering. In this approach, a nonparametric linear estimate of the unknown clean-speech signal is obtained based on a decomposition of the observed noisy signal into mutually orthogonal signal and noise subspaces. This decomposition is possible under the assumption of a low rank linear model for speech and uncorrelated additive (white) noise interference. Under these conditions, the energy of less correlated noise spreads over the whole observation space while the energy of the correlated speech components is concentrated in a subspace there of. Also, the signal subspace can be recovered consistently from the noisy data. Generally speaking, noise reduction is obtained by

removing the noise subspace and by removing the noise contribution in the signal subspace.

In this paper we propose a subspace approach for single channel speech enhancement in noisy environments based on the KLT, and implemented via Principal Component Analysis (PCA) [1]. The KLT provides an optimum compression of information, while the DFT and the DCT are suboptimal. The main problem in subspace approaches is the optimal choice of the different parameters [6]. Use of the non-negative eigenvalue for signal estimation [3] [4] [5]. This criterion provides consistent parameter estimates and allows us to implement an automatic noise reduction algorithm that can be applied almost blindly to the observed data. The goal of speech enhancement varies according to specific applications, such as to reduce listener fatigue, to boost the overall speech quality, to increase intelligibility, and to improve the performance of the voice communication device. We try to improve the overall speech quality while minimizing any speech intelligibility loss.

The corrupted speech therefore needs modern data analysis because it is simple method for extracting relevant information from complex data matrix using eigenvalues and eigenvectors. NOIZEUS Database is used for testing.

2 PRINCIPAL COMPONENT ANALYSIS

Principal components analysis (PCA) is a technique used to reduce a multidimensional data to lower dimensions for analysis [2]. PCA consists of computation of the eigenvalue decomposition or singular value decomposition of a data set, Usually after mean centering the data for each attribute. In general PCA methods are based on the Karhunen-Loève Transform (KLT). PCA involves a mathematical procedure

- Sangita K. Bavkar is currently working with PVG's COET Pune as Assistant Professor and pursuing ME from Cummins College of Engineering for Women Pune, Maharashtra, India ,E-mail: sangitabavkar@gmail.com
- Shashikant L. Sahare is Associate with Cummins College of Engineering for Women Pune, Maharashtra, India and pursuing Ph.D. from Sathyabama University, Chennai, India, Email: shashikantsahare@rediffmail.com

that transforms a number of correlated variables into a number of uncorrelated variables called *principal components*. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. However, with more than three dimensions, we usually need a little help. What PCA does is that it takes your cloud of data points, and rotates it such that the maximum variability is visible. Another way of saying this is that it identifies your most important gradients. It is very easy method to implement based on Eigen analysis [9].

3 NOISE ESTIMATION

Noise spectrum estimation is a fundamental component of speech enhancement systems. Traditional noise estimation methods, which are based on voice activity detectors (VADs), restrict the update of the estimate to periods of speech absence. Additionally, VADs are generally difficult to tune and their reliability severely deteriorates for weak speech components and low input SNR.

Alternative techniques, based on histograms in the power spectral domain, are computationally expensive, require much memory resources, and do not perform well in low SNR conditions. Furthermore, the signal segments used for building the histograms are typically of several hundred milliseconds, and thus the update rate of the noise estimate is essentially moderate [10].

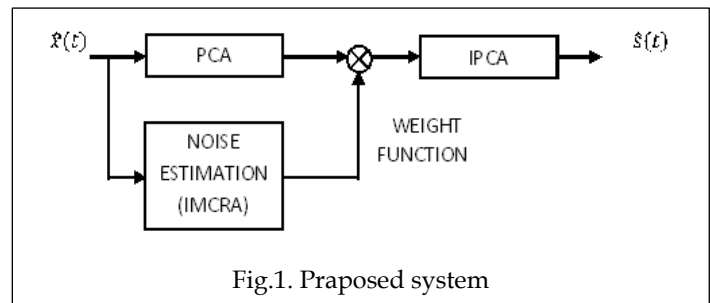
The improved minima controlled recursive averaging (IMCRA) approach [7], for noise estimation in adverse environments involving nonstationary noise, weak speech components, and low input signal-to-noise ratio (SNR). The noise estimate is obtained by averaging past spectral power values, using a time-varying frequency-dependent smoothing parameter that is adjusted by the signal presence probability. The speech presence probability is controlled by the minima values of a smoothed periodogram. The IMCRA includes two iterations of smoothing and minimum tracking. The first iteration provides rough voice activity detection in each frequency band. Then, smoothing in the second iteration excludes relatively strong speech components, which makes the minimum tracking during speech activity robust.

Cohen shown that in nonstationary noise environments and under low SNR conditions, the IMCRA approach is very effective. In particular, compared to a competitive method, it obtains a lower estimation error, and when integrated into a speech enhancement system achieves improved speech quality and lower residual noise.

Noise estimation is very important part of any speech enhancement algorithm because the estimated speech quality and intelligibility is depend on it.

4 PROPOSED SYSTEM

The fig 1 shows the proposed system which is based on PCA and IMCRA noise estimation method.



4.1 PCA Algorithm

Consider a speech signal $s(t)$ corrupted by an additive stationary white Gaussian noise $n(t)$. The observed noisy signal can be expressed as follows:

$$x(t) = s(t) + n(t) \quad (1)$$

- Step 1: Get noisy speech data
 - Step 2: Subtract the mean and get data adjusted.
 - Step 3: Calculate the covariance matrix
 - Step 4: Calculate the eigenvectors and eigenvalues of the covariance matrix
 - Step 5: Choosing components and forming a feature vector in descending order.

$$\text{FeatureVector} = (\text{eig}_1 > \text{eig}_2 > \text{eig}_3 \dots \text{eig}_n > 0)$$
 - Step 6: Deriving the new data set

$$\text{FinalData} = (\text{FeatureVector})^T \times (\text{DataAdjust})$$
 - Step 7: Getting Old Data Back

$$(\text{DataAdjust}) = (\text{FeatureVector})^T \times \text{FinalData}$$
 - Step 8: Calculate weight function as wf

$$(\text{DataAdjust}) = (\text{FeatureVector})^T \times \text{FinalData} \times \text{wf}$$
- But, to get the actual original data back,

$$(\text{DataAdjust}) = (\text{FeatureVector})^T \times \text{FinalData} + \text{OriginalMean}$$
This formula also applies to when you do not have all the eigenvectors in the featurevector. So even when you leave out some eigenvectors, the above equation still makes the correct transform.

5 PERFORMANE EVALUATION

For the performance evaluation the NOIZEUS database is used. The resulting files were evaluated perceptually and objectively by quality measures. From the informal listening tests, the increased intelligibility of speech is verified. Sampling frequency is 8 kHz. The frame size is $N=320$ and we have used the Hamming window with 50% overlapping.

5.1 Performance evaluation in White Gaussian Noise

The objective evaluation [8] of speech on the global SNR, Itakura-Saito (IS) and perceptual evaluation of speech quality (PESQ) and Log Likelihood Ratio (LLR) is listed in Table I. The following test has been taken with 0dB, 6dB, and 11dB and 18dB SNR of the noisy signal. The fig 2 shows the plot of

clean, noisy and enhanced speech signals. The performance for different SNR levels depends on the variance of the Gaussian noise.

TABLE 1
EVALUATION IN WHITE GAUSSIAN NOISE

SNR(dB)		PESQ		IS		LLR	
Noisy	PCA-IMCRA	Noisy	PCA-IMCRA	Noisy	PCA-IMCRA	Noisy	PCA-IMCRA
0	8	1.4	1.8	5.5	3.1	2.1	1.7
6	12	1.8	2.4	4.17	2.1	1.7	1.3
11	16	2.1	2.8	3.3	2.4	1.5	1.1
18	21	2.5	3.2	2.2	2.0	1.1	0.7

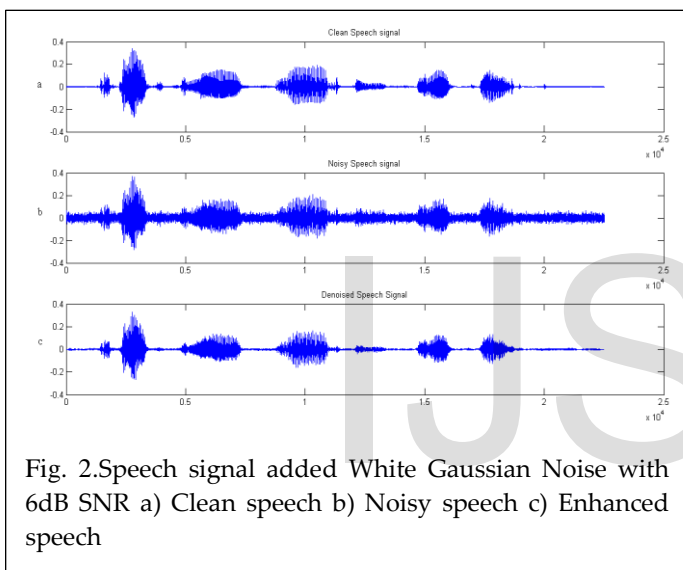


Fig. 2.Speech signal added White Gaussian Noise with 6dB SNR a) Clean speech b) Noisy speech c) Enhanced speech

6 CONCLUSION

We have proposed in this paper a subspace approach for speech enhancement in highly noisy environments. This approach is based on PCA. The performance evaluation based on SNR, Itakura-Saito distortion measure, Perceptual Evaluation of Speech Quality and Log Likelihood Ratio (LLR) , as well as in formal listening tests, shows clearly that our algorithm provides some signal distortion and a higher noise reduction using improved minimum controlled recursive average noise estimation algorithm.

ACKNOWLEDGMENT

We would like to thanks to all of those who provided us with valuable guidance. We would like to specially thanks to staff members of PVG's College of Engineering Technology and Cummins College of Engineering Pune for their guidance and interest.

REFERENCES

- [1] Jonathon Shlens, "A Tutorial On Principal Component Analysis", Center for Neural Science, New York University, New York City, NY 10003-6603 and Systems Neurobiology Laboratory, Salk Institute for Biological Studies La Jolla, CA 92037 (Dated: April 22, 2009; Version 3.01)
- [2] Lindsay I Smith, "A tutorial on Principal Components Analysis", February 26, 2002.
- [3] Y. Ephraim and H.L. Van Trees, "A signal subspace approach for speech enhancement", IEEE Trans. on Speech and Audio Proc., Vol. 3, No. 4, pp. 251-266, July 1995.
- [4] R. Vetter, N. Virag, P. Renevey, and J.-M. Vesin, "Single channel speech enhancement using principal component analysis and MDL subspace selection", in *Eurospeech'99, Budapest, Hungary*, 1999.
- [5] Amin Haji Abolhassani, Sid-Ahmed Selouani, Douglas O'Shaughnessy, "Speech Enhancement Using PCA and Variance Of The Reconstruction Error In Distributed Speech Recognition", *Automatic Speech Recognition & Understanding*, 2007. ASRU. IEEE workshop on Digital Object Identifier: 10.1109/ASRU.2007.4430077, Publication year: 2007, Page 19-23.
- [6] C.J. James and D. Lowe, "Extracting multisource brain activity from a single electromagnetic channel," *Artificial Intelligence in Medicine*, vol. 28, issue.1, pp. 89-104, 2003.
- [7] Cohen, I. (2003). Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging. *IEEE Transactions on Speech and Audio Processing*, 11(5), 466-475.
- [8] Yi Hu and Philipos C. Loizou, Senior Member, IEEE "Evaluation of Objective Quality Measures for Speech", *IEEE, Speech, and Language Processing*, vol. 16, no. 1, January 2008.
- [9] Saeed V. Vaseghi, Advanced "Digital Signal Processing and noise Reduction", Chapter 9, Wiley, Fourth Edition
- [10] P. C. Loizou, "Speech Enhancement: Theory and Practice", 2nd ed. Boca Raton, FL.: CRC, 2007.