# Speech Denoising Using Variant of Minimum Controlled Recursive Average Algorithm and PCA

**Sangita Bavkar & Shashikant Sahare**

M.E Student, Cummins College of Engineering, Pune, India
E-mail : sangitabavkar@gmail.com, Shashikantsahare@rediffmail.com

*Abstract* – **Speech denoising is very important in many applications where noise is unavoidable. The speech accuracy reduces strictly when the systems are operated in noisy environments. There are different Speech enhancement methods, a generalized form of Principal Component Analysis (PCA) is used for speech enhancement.**

**A PCA based algorithm is proposed for denoising of speech degraded by noise interference. Principle Component Analysis (PCA) is a standard tool in modern data analysis because it is simple method for extracting relevant information from complex data matrix using eigenvalues and eigenvectors. By selecting conveniently the numbers of retained principal components, interesting simplified signals can be reconstructed.**

***Keywords – Speech Enhancement, Subspace Approach, Principal Component Analysis (PCA), Noise Estimation, Eigenvalue and Eigenvector.***

## I. INTRODUCTION

The objectives of speech denoising varies according to specific applications, such as to boost the overall speech quality, to increase intelligibility, and to improve the performance of the voice communication device. To improve the performance in noisy environments, modern speech communication systems have integrated speech enhancement as an essential component, generally as a front-end processor. This is an especially important issue for devices used in mobile environments, such as mobile phones and hands-free telephones in cars. Hearing aids can also benefit from speech enhancement techniques to improve speech intelligibility.

There are some of speech enhancement techniques; signal subspace approach has successful results. Using this approach, we obtain a nonparametric linear estimate of the unknown clean-speech signal, based on a decomposition of the observed noisy signal into mutually orthogonal signal and noise subspaces. In this paper enhancement is obtained by removing the noise subspace and optimally weighting the signal subspace to remove noise energies from this 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) [13]. 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. A novel approach for the optimal subspace partition using the Minimum Description Length (MDL) criterion [6]. 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 eigen values and eigenvectors. NOIZOU Database is used for testing.

## II. PROPOSED SYSTEM

### *A. Principal Component Analysis*

Principal components analysis (PCA) is a technique used to reduce a multidimensional data to lower dimensions for analysis. PCA consists of computation of the eigenvalue decomposition or singular value decomposition of a data set, Consider a speech signal s(t) corrupted by an additive stationary background noise n(t). The observed noisy signal can be expressed as follows:

$$x(t)=s(t)+n(t) \quad (2.1)$$

Our noise reduction algorithm operates on a frame by frame basis. A very efficient and robust implementation of the subspace approach is provided by the PCA.

Principal components analysis (PCA) is a technique used to reduce a multidimensional data to lower dimensions for analysis. PCA is based on Eigen analysis; we solve for the eigenvalues and eigenvectors of a square symmetric matrix with sums of squares and cross products. The eigenvector associated with the largest eigenvalue has the same orientation as the first principal component. The eigenvector associated with the second largest eigenvalue determines the orientation of the second principal component. The sum of the eigenvalues equals the trace of the square matrix and the number of eigenvectors equals the number of rows (or columns) of this matrix.

One of the main issues in using principal component analysis (PCA) is the selection of the number of principal components (PCs). There already exist many methods to calculate the number of PCs, but most of them use monotonically increasing or decreasing indices which makes the decision to choose the number of principal components very subjective. Among all, Principal Component Analysis (PCA) showed to be promising. A key issue in developing a PCA model is to choose the adequate number of PCs to represent the system in an optimal way. If fewer PCs are selected than required, a poor model will be obtained and an incomplete representation of the process results. On the contrary, if more PCs than necessary are selected, the model will be over-parameterized and will include noise.

In signal processing applications we use PCA to determine the number of independent source signals from noisy observations by selecting the number of significant principal components. Assuming that the measurement noise corresponds to the smallest equal eigenvalues of the *covariance matrix*, the Akaike Information Criterion and the Minimum Description Length principle have been applied to selecting the number of PCs. However, these criteria apply to PCA based on the *covariance matrix.* The PCA algorithm as follows[3]. Get Speech data Consider some data to apply PCA. Subtract the mean for PCA to work properly. The mean subtracted is the average across each dimension. Calculate the covariance matrix and calculate the eigenvectors and eigenvalues of the covariance matrix since the covariance matrix is square, we can calculate the eigenvectors and eigenvalues for this matrix. These are important, as it tell the useful information about important data.

Choosing components and form the feature vector. The eigenvector with the highest eigenvalue is the principle component of the data set. Once eigenvectors are found from the covariance matrix, the next step is to order them by eigenvalue, highest to lowest. This gives the components in order of importance. I*gnore* the components of lesser significance. If leave out some components, the final data set will have less dimensions than the original. *Feature vector*, which is a matrix of vectors. This is constructed by taking the eigenvectors that you want to keep from the list of eigenvectors

$$\text{FeatureVector}=(eig_1>eig_2>\ldots>eig_n) \quad (2.2)$$

To derive the new dataset chose the highest components(eigenvectors) and formed a feature vector, we simply take the transpose of the vector and multiply it on the left of the original data set, transposed.

$$\text{FinalData}=\text{RowFV}x\text{RowDA} \quad (2.3)$$

RowFV is the matrix with the eigenvectors in the columns *transposed* so that the eigenvectors are now in the rows, with the most significant eigenvector at the top. RowDA is the mean-adjusted data transposed. To recover the old data back, if considered all the eigenvectors in transformation then will get *exactly* the original data back. If reduced the number of eigenvectors in the final transformation, then the retrieved data has missing some information. To get the original data back,

$$\text{RowDA} = \text{RowFV}^{-T}x\text{FinalData} \quad (2.4)$$

Noise reduction can be achieved by reconstructing the initial data using only the p weighted eigenvectors of the signal plus- noise subspace, as proposed by Ephraim et al. in [2]:

$$s(t) = \sum_{n=1}^{p}\left(g_j \text{FinalData}\right) \qquad p < m \quad (2.5)$$

Where $g_j$ is a weighting function given by:

$$g_j = \exp\left\{-\frac{\nu\sigma_n^2}{\lambda \text{j}}\right\} \quad j = 1, \ldots, p \quad (2.6)$$

With $\nu = , \sigma_n^2\,(Noise\ variance)\ and\ \ \lambda j(jth\ eigen\ value)$ .

The parameters m and p are generally chosen in such a way. That the noise is essentially relegated to the residuals of the signal approximation given by Equation (2.5).

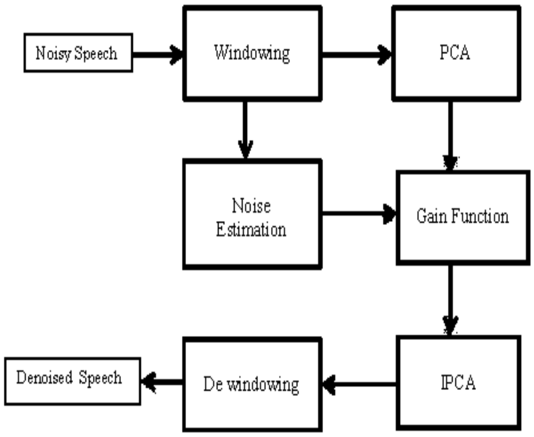


Fig.1: Block Diagram of Proposed speech denoising Using PCA

### B. *Noise Estimation*

In most speech-enhancement algorithms, it is assumed that an estimate of the noise spectrum is available. Such an estimate is critical for the performance of speech-enhancement algorithms as it is needed, for instance, to evaluate the Wiener filter in the Wiener algorithms (Lim and Oppenheim,1978) or to estimate the a priori SNR in the MMSE algorithms (Ephraim and Malah, 1984) or to estimate the noise covariance matrix in the subspace algorithms (Ephraim and Van Trees, 1993). The noise estimate can have a major impact on the quality of the enhanced signal. If the noise estimate is too low, annoying residual noise will be audible, while if the noise estimate is too high, speech will be distorted resulting possibly in intelligibility loss. The simplest approach is to estimate and update the noise spectrum during the silent (e.g., during pauses) segments of the signal using a voice-activity detection (VAD) algorithm (e.g., Sohn and Kim,1999). Although such an approach might work satisfactorily in stationary noise (e.g., white noise), it will not work well in more realistic environments(e.g., in a restaurant) where the spectral characteristics of the noise might be changing constantly. Hence there is a need to update the noise spectrum continuously over time and this can be done using noise-estimation algorithms[5].

### C. *Subjective and Objective Evaluation*

The main evaluation methods for speech enhancement systems look at the effect of the system on the intelligibility of the speech signal and the improvement in the overall quality of the signal. There are two methods of assessing the Intelligibility or quality of speech: subjective and objective

Now a day, the most accurate method for evaluating speech quality is through subjective listening tests. Although subjective evaluation of speech enhancement algorithms is often accurate and reliable provided it is performed under toughest conditions ,it is costly and time consuming. For that reason, much effort has been placed on developing objective measures that would predict speech quality with high correlation. Many objective speech qualities measures have been proposed in the past to predict the subjective quality of speech [5]. Most of these measures, however, were developed for the purpose of evaluating the distortions introduced by speech codecs and communication channels. There are different from those introduced by speech enhancement algorithms. As a result, it is not clear whether the objective measures originally developed for predicting speech coding distortions [1] are suitable for evaluating the quality of speech enhanced by noise suppression algorithms.

To reduce the length and cost of the subjective evaluations, only a subset of the NOIZEUS corpus was processed . Several objective speech quality measures were evaluated: segmental SNR (segSNR), perceptual evaluation of speech quality (PESQ), including the log-likelihood ratio (LLR), Itakura-Saito distance measure (IS).

## III. ALGORITHM

1. Take input noisy speech signal x (t).
2. Convert x (t) into frame using windowing.
3. Separate the speech and noise for estimate the noise variance.
4. Find the weighting function.
5. Perform PCA.
6. Using equation (5) reconstruct the speech signal with maximum noise reduction and minimum speech distortion.
7. Perform IPCA.
8. Using Overlap/Add reconstruct the enhanced speech signal.

## IV. EXPERIMENTAL RESULTS

### A . *Denoising in White Gaussian Noise*

For the performance evaluation the NOIZOU database is used. The evaluations of speech on the global SNR, Itakura saito are listed in Table I. From the informal listening tests increased the intelligibility of speech.

TABLE I. EVALUATION IN THE CASE OF WHITE GAUSSIAN NOISE.

| Noisy | | PCA | |
|---|---|---|---|
| SNR | IS | SNR | IS |
| 0 dB | 5.10 | 8.8 | 4.77 |
| 5dB | 4.18 | 11.38 | 3.86 |
| 10dB | 3.29 | 14.77 | 3.01 |
| 15dB | 2.5 | 17.73 | 2.1 |

### *B. Comparison of Objective Parameters in White Gaussian Noise*

The noisy signal with 5dB SNR is evaluated with this algorithm. The denoised and noisy speech signal is compared using objective parameter like global SNR, segmental SNR, Itakura saito, PESQ, Log Likelihood Ratio (LLR). The fig 2. Shows the comparison of noisy and denoised speech signal with respect to objective parameter.

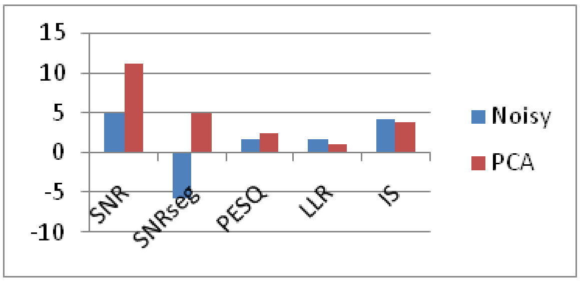


Fig.2: Comparisons of Objective Parameter for noisy signal with 5dB SNR

### *C. Denoising of Speech Signal using PCA*

The clean signal is taken from the NOIZOU database, the white Gaussian noise is added with 5dB SNR, and after applying the PCA we got the improved speech signal. The fig.3 shows the first clean speech, second noisy speech and third enhanced speech signal

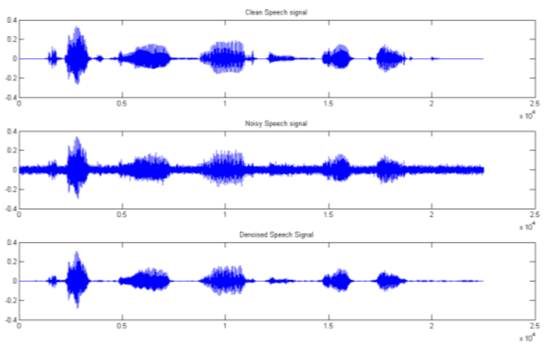

Fig.3: Denoising of Speech Signal using PCA

## V. CONCLUSIONS

We have proposed in this paper a subspace approach for single channel speech enhancement in highly noisy environments. This approach is based on PCA. The performance evaluation based on segmental SNR, Itakura-Saito distortion measure, observation of the spectrograms, as well as in formal listening tests, shows clearly that our algorithm provides some signal distortion and a higher noise reduction that existing enhancement methods based on traditional subspace approaches.

## VI. ACKNOWLEDGMENT

We would like to thanks to all of those who provided us with useful and helpful 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.

## VII. REFERENCES

[1] Rolf Vetter ,Nathalie Virag,Lausanne,Switzerland "Single Channel Speech Enhancement Using Principal Component Analysis and MDL Subspace Selection"

[2] 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.

[3] Lindsay I Smith, February 26, 2002 "A tutorial on Principal Components Analysis"

[4] Yi Hu and Philipos C. Loizou, Senior Member, IEEE "Evaluation of Objective Quality Measures for Speech Enhancement " IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 16, NO. 1, JANUARY 2008,

[5] Loizou, P., Sundarajan R. "A Noise estimation Algorithm for highly non–stationary Environments", speech communication 48 (2006) 220-231 Science direct.

[6] S. Valle,W. Li, and S.J. Qin, "Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods," vol. 38, pp. 4389–4401, Industrial and Engineering Chemistry Research, 1999.

[7] J. Rissanen, “Modeling by shortest data description,” vol. 14, pp. 465–471, Automatica, 1978.

[8] G. Schwarz, “Estimating the dimension of a model,” vol. 6, pp. 461–464, Ann. Statist., 1978.

[9] S.J. Qin and R. Dunia, “Determining the number of principal components for best reconstruction,” Journal of Process Control, vol. 10, pp. 245–250, 2000.

[10] Saeed V. Vaseghi,Advanced Digital Signal Processing and noise Reduction, WILEY Fourth Edition

[11] “A Tutorial On Principal Component Analysis” By Jonathon Shlens

[12] P. C. Loizou, Speech Enhancement: Theory and Practice, 1st ed. Boca Raton, FL.: CRC, 2007.

[13] S. Haykin, “Adaptive filter theory,” Printice Hall.

[14] Gerkmann, T. &Hendriks, R. C. “Unbiased MMSE-Based Noise Power Estimation With Low Complexity and Low Tracking Delay”IEEE Trans Audio, Speech, Language Processing, 2012, 20, 1383-1393

[15] K. Judd and A. Mees, \On selecting models for nonlinear time series", Physica D, Vol. 82, pp. 426-444, 1995.

◈◈◈