Singular Value Decomposition Based Dual Channel Spectral Subtraction

Sandhya N.K

Correspondence: sandhya.nerlige@gmail.com

Abstract: SVD based Dual channel Spectral subtraction algorithm [SVD-DSS], is a new approach to enhance the speech using spectral subtraction [SS] method and low-rank approximation. This method mainly uses new class of earphones/headphones consisting of two microphones at different positions. The SVD decomposition is used for low-rank approximation of noise[7] and spectrally subtract from the speech to get enhanced speech. For input signal modelling a speech of 10 seconds was recorded in the presence of traffic noise at different Signal to Noise ratio (SNR) positions. For evaluating the performance of SVD-DSS algorithm, SNR, Spectrogram analysis and subjective listening tests were performed. This analysis showed the effective improvement in the speech signal SNR.

Keywords: Singular Value Decomposition - SVD, Spectral Subtraction, Low-rank approximation, Dual Channel Spectral Subtraction, Multi-Source Communication, Speech Reconstruction, Noise reduction.

1. Introduction:

Spectral subtraction is one of the earliest speech enhancement methods. It is a simple algorithm to suppress background noise in the speech. This algorithm faces difficulties in detecting true or approximate noise level to subtract it from the speech signal spectrally, in real-time. This work mainly deals with the Dual-microphone devices. This is because, usually the noise source we considered are mainly vehicle traffic noise. These noisy sound waves will generate from a point object and will be planar when they reach the speaker/microphones. When two microphones are placed with a considerable distance between them, on the same plane, parallel to the planar waves of noisy sound, each one receives approximately same level of noise. But also, with that, when one microphone is nearer to the speaker mouth, and another a few inches away, the speech signal level each microphone receives is very different, i.e. one with speech dominance [SD] and another with noise dominance [ND].

From this process, we can easily estimate the noise level at every sample and use it for spectral subtraction. When we use SVD and low rank approximation, ND signal gets more dominance towards noise. This noise estimation can be used to subtract noise from speech dominant signal.

2. Spectral subtraction:

Given a clean speech signal 's', distorted with additive noise 'n' resulting noisy speech 'y' can be expressed as:

In time domain it is given by:

$$y(n) = s(n) + z(n) \qquad (1)$$

In Frequency domain it is given by:

$$Y(\omega) = S(\omega) + Z(\omega)$$

or $$Ye^{j\emptyset_{y,k}} = S_k e^{j\emptyset_{s,k}} + Z_k e^{j\emptyset_{d,k}} \qquad (2)$$

where $Y(\omega)$, $S(\omega)$, $Z(\omega)$ are the Fourier transform coefficients of noisy speech, clean speech and noise.

ISSN 2229-5518

Then the clean speech estimation/speech enhancement of 'y' using spectral subtraction will be expressed as follow,

$$\hat{S}_{k,p}^{p} = \hat{Y}_{k,p}^{p} - E[Z_k^p] \qquad (3)$$

Here $E[Z_k^p]$ is the expectation of noise, estimated from the speech absent segments of noisy speech. For magnitude spectral subtraction 'p' [2].

2.1 Low-Rank approximation:

For matrix 'A' of dimension (nxm), The low-rank approximation [21] of that signal is given by

$$\widehat{A_k} = {}_{\mathrm{i=1}}^{\mathrm{k}}\Sigma\Omega_{\mathrm{i}}\mathrm{U_i}\mathrm{V_i^T} \qquad (4)$$

Where Ω_i are the largest k singular values of '*A*', U_i *and* V_i^T are respectively, left and right singular vectors, resulted by the SVD decomposition of 'A'. SVD based Dual Channel Spectral subtraction [SVD-DSS]

The main necessity of SVD here is to estimate much proper noise level. With the help of low-rank approximation, using Toeplitz structure, both noise estimation and weighted averaging of noise samples can be done at a time. Subtracting this better estimation of noise, from absolute value gives enhanced speech signal. 'μ' is the oversubtraction parameter, mainly to control spectral peaks by over subtraction. When these Here 'α' is the spectral flooring parameter used to fill up the gaps between peak values of power or magnitude spectrum, and also to remove negative co-efficient. These gaps and negative coefficients, together, creates musical noise after converting them to time domain.

3. Algorithm:

Input: Noise dominant signal('d'), Speech dominant signal('s')

Procedure:

1) Using SVD decomposition, find the Low-rank estimation of 'n':

$$\widehat{d_k} = {}_{\mathrm{i=1}}^{\mathrm{k}}\Sigma\Omega_{\mathrm{i}}\mathrm{U_i}\mathrm{V_i^T} \qquad (5)$$

Where Ω_i are the largest k singular values of '*d*', U_i and V_i^T are unitary matrices and rectangular diagonal matrices.

2) Take Fourier Transform of noise estimate $\widehat{d_k}, \widehat{D_k}$ and speech dominant signal *s*, S_f

3) By spectral subtraction method, estimate the enhanced signal Fourier coefficients

$$\widehat{S} = \begin{Bmatrix} {S_f}^p - \mu\widehat{D^p}_k & if\ S_f > \mu\widehat{N_k} \\ \alpha\widehat{M} & else \end{Bmatrix} \qquad (6)$$

Here $\widehat{M}$ is the average value of preceding 3 consecutive values of $\widehat{S}$. 'α' ($0<\alpha<1$) is the spectral flooring parameter, μ is oversubtraction parameter.

Using phase of original speech-dominant signal 's' and by taking inverse Fourier transform on $\widehat{S}$,reconstruct the enhanced speech signal.

Output: Speech enhanced signal ($\hat{s}$)

4. Results:

The performance of the algorithm is checked on the signal, recorded from dual microphone placed at different signal level and noise level. The two recordings were sampled at rate of 16000Hz, hence 20ms (320 samples) window was used for spectral subtraction, which considers even transition phases in speech properly [1]. For low-rank approximation, 3-6 largest singular values, and tested for better estimation of noise.

ISSN 2229-5518

Experimentally it shows that, better noise estimation can be achieved with largest 4 singular values and over-subtraction factor between 5-3 [3]. To reduce background musical noise [4], spectral flooring parameter varying between 0.4-0.7 can be used. Spectrogram of original Speech dominant signal and reconstructed speech dominant signal can be observed for the enhancement.

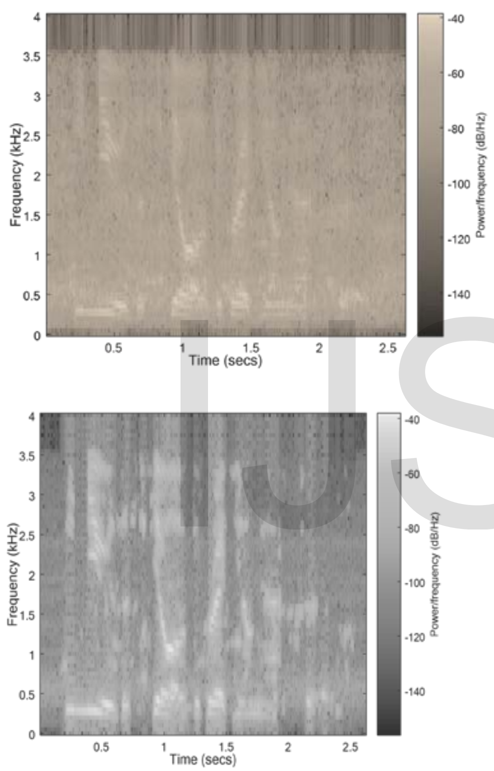

Fig1: Spectrogram of original speech dominant signal and reconstructed speech signal shows the speech enhancement and the formants can be observed clearly in the reconstructed signal.

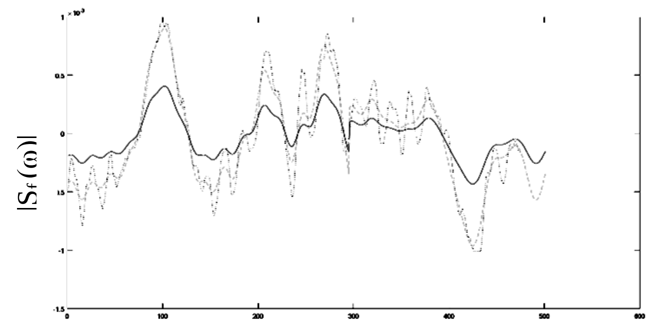

Fig2: shows the original noise level of a 20ms noise dominant signal segment (dotted line) and its estimate using low-rank approximation(in thick line) .

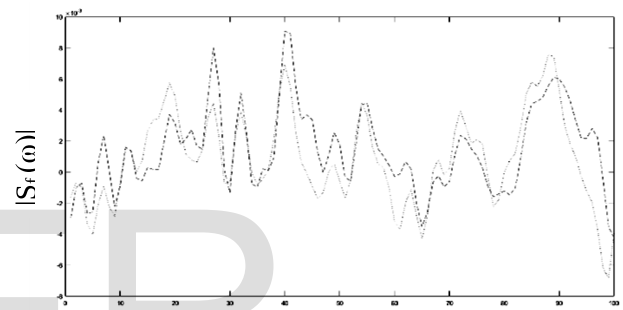

Fig3: The original and enhanced speech dominant signal. The very useful advantage with this SVD-DSS is, speech can be enhanced effectively in both voiced and unvoiced speech frames [1], as the noise and speech signal levels can be separately estimated and analysed for spectral subtraction.

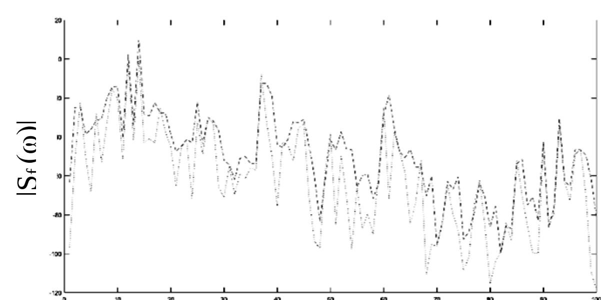

Fig4: SNR improvement of original and enhanced speech segment.

5. Conclusion:

Spectral subtraction is one of the earliest methods for speech enhancement. The performance of this method would be great and faster, when this method is accompanied with better noise estimating methods. The SVD-DSS method is one such method for estimating very precise noise level for spectral subtraction. By using two microphones for recording same speech signal with same background noise with different spectral levels of noise and speech, solves this noise estimating problem, even in real-time scenario. As low-rank approximation averages [8] the diagonal values of the resulting matrix make the noise estimate better and applicable to spectral subtraction. Since Noise dominant speech signal would be having higher values in corresponding noise frequency components, one can identify which frequency band got affected more and accordingly can control the spectral subtraction with the help of over subtraction and spectral flooring parameters [5]. Over-subtraction and spectral flooring parameters effectively control the musical tone noise created by the inverse transform of negative or zero Fourier coefficients in random frequency positions. With this simple complexity algorithm one can achieve better noise cancellation and speech enhancement and can be easily implemented in earphones/headphones with low cost effectiveness.

6. Acknowledgement:

The author thanks BTL Institute of Technology and Management and Visvesvaraya Technological University for the support towards the research.

7. References:

1. Philipos C. Loizou: "SPEECH ENHANCEMENT Theory and Practice".
2. S.F Boll: "Suppression of acoustic noise in speech using spectral subtraction", *IEEE Trans. Acoust. Speech and Sig. proc., 27:113-120,1979.*
3. M. Berouti, r. Schwartz, J. Makhoul, "Enhancement of speech corrupted by acoustic noise", *Proc, IEEE ICASSP, 208-211, 1979.*
4. Boh Lim Sim, Yit Chow Tong, Joseph S Chang, and Chin Tuan Tan, "A parametric formulation of the generalized spectral subtraction method", IEEE Trans. Speech Audio Process.
5. Sunil D. Kamath and Philipos C. Loizou, "A Multi-Band Spectral Subtraction Method For Enhancing Speech Corrupted By Colored Noise" - IEEE international conference May 2002.
6. Ephraim, Y. and Van Trees, H.L. (1995), A signal subspace approach for speech enhancement, IEEE Trans. Speech Audio Process., 3(4), 251–266.
7. Tufts, D. and Shah, A. (1993), Estimation of a signal waveform from noisy data using low-rank approximation to a data matrix, IEEE Trans. Signal Process., 41(4), 1716–1721.
8. Lawerence R. Rabiner and Ronald W. Schafer, "Introduction to Digital Speech Processing".
9. Kaladharan N, "Speech Enhancement by Spectral Subtraction Method", International Journal of Computer Applications (0975 – 8887), June 2014.

10. Tobias Goehring, Federico Bolner, Jessica J.M. Monaghan, Bas van Dijk, Andrzej Zarowski, Stefan Bleeck, "Speech enhancement based on neural networks improves speech intelligibility in noise for cochlear implant users", Hearing Research, IEEE international conference – November 2016.
11. S. China Venkateswarlu, A. Subba Rami Reddy & K. Satya Prasad "Speech Enhancement using Boll's Spectral Subtraction Method based on Gaussian Window", Global Journal of Researches in Engineering: FElectrical and Electronics Engineering, 2014.
12. Navneet Upadhyay, and Abhijit Karmakar, "Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study" Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015).