
# Variance Normalized Perceptual Subspace Speech Enhancement

Sudeep Surendran, T. Kishore Kumar

PII: S1434-8411(17)30236-4
DOI: http://dx.doi.org/10.1016/j.aeue.2017.01.023
Reference: AEUE 51790

To appear in: *International Journal of Electronics and Communications*

# Variance Normalized Perceptual Subspace Speech Enhancement

Sudeep Surendran[a], T. Kishore Kumar[b]

[a] *Research scholar, Department of ECE, National Institute of Technology Warangal, Warangal-506004, Telangana, India*
[b] *Associate Professor, Department of ECE, National Institute of Technology Warangal, Warangal-506004, Telangana, India*

**Abstract**

In this paper, a perceptual subspace speech enhancement method using masking property of the human auditory system and variance normalization is presented. The simultaneous masking property of the human auditory system is used while deciding the gain parameters of the algorithm. Spectral Domain Constrained estimator was employed in determining the filter coefficients, and colored noise was handled by replacing the noise variance by Rayleigh quotient. Normalized variance was used as control parameter which adaptively decided the amount of filtering to be performed. Adaptively varying the control parameter enabled the proposed algorithm to perform better in non-stationary colored noise environments, compared to some of the other existing algorithms. The waveforms and spectrograms indicate that the proposed algorithm removes most of the noise without distorting the speech signal. The results obtained in terms of objective parameters ($wcep$, $WSS$, $SNR_{LOSS}$ and $SNR_{LESC}$) and informal listening test support the claim.

*Keywords:* Speech Enhancement, Signal Subspace Approach, Masking Property, Variance Normalization
*2010 MSC:* 15Axx, 15Bxx, 65Fxx

## 1. Introduction

Speech enhancement refers to the improvement in quality and intelligibility of noise-corrupted speech signals by using supervised or unsupervised speech enhancement methods. It is commonly used as a pre-processing block in a lot of applications like automatic speech recognizer and other communication systems.

Supervised methods achieve noise reduction by considering a model for both the speech and noise signals, whose parameters are estimated from the training samples of that signal and then defining an interaction model. Hidden Markov model (HMM) based methods [1] and Gaussian Mixture Models(GMM) [2, 3], belong to this type of algorithm. The performance of the supervised approaches depends on the prior information fed to the system which limits its performance in non-stationary noise environments.

Unsupervised methods assume a statistical model for the speech and noise signals. Clean speech is

*Corresponding author
*Email address:* sudipsuren@nitw.ac.in (Sudeep Surendran)