

# Research on Noise Suppression Based on Square Root Wiener Filtering Algorithm

Shuo Zhang, Xue Li, Shun Zhang, Chen Chen, Ruirui Han

School of Opto-Electronic Information, Yantai University, Yantai Shandong  
Email: 98293409@qq.com

Received: Jul. 7<sup>th</sup>, 2018; accepted: Jul. 26<sup>th</sup>, 2018; published: Aug. 2<sup>nd</sup>, 2018

---

## Abstract

A speech enhancement algorithm based on square root Wiener filter is proposed for speech enhancement with noise suppression. Generally speaking, the influence of the two types of speech distortion introduced by the gain function is to enhance the amplification distortion and attenuation distortion of the speech signal amplitude. In fact, when the processed speech contains only attenuation distortion and amplification distortion less than 6 dB, the influence on speech quality is relatively small. Current data show that one of the reasons that existing algorithms cannot improve speech intelligibility is that they allow amplification and distortion more than 6 dB. In this paper, a noise suppression algorithm for amplitude distortion more than 6 dB is proposed. Based on the gain function characteristic of square root Wiener filtering, a noise suppression function based on the constraints of a priori SNR is derived. Experimental results show that the algorithm proposed in this paper can effectively improve speech quality and improve speech intelligibility.

## Keywords

Speech Enhancement, Square Root Wiener Filter, Priori Signal-to-Noise Ratio, Noise Reduction

---

# 基于平方根维纳滤波算法的噪声抑制研究

张 硕, 李 雪, 张 顺, 陈 晨, 韩蕊蕊

烟台大学, 光电信息科学技术学院, 山东 烟台  
Email: 98293409@qq.com

收稿日期: 2018年7月7日; 录用日期: 2018年7月26日; 发布日期: 2018年8月2日

---

## 摘 要

针对噪声抑制的语音增强问题, 本文提出一种基于平方根维纳滤波的语音增强算法。通常增益函数引入

的两种类型的语音失真的影响包括：当增强语音信号振幅的放大失真和衰减失真。事实上，当被处理的语音只包含衰减失真和小于6分贝的放大失真时，对语音质量的影响度较小。目前的数据表明，现有算法不能高质量提高语音清晰度的一个原因是，它们允许放大失真超过6分贝。本文主要提出一种针对幅度失真大于6分贝的噪声抑制算法，根据平方根维纳滤波增益函数特性，推导出基于先验信噪比约束条件的噪声抑制函数。实验结果表明，本文提出的算法能有效的提升语音质量，提高语音的可懂度。

## 关键词

语音增强，平方根维纳滤波，先验信噪比，噪声抑制

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着信息化时代的到来，人工智能(Artificial Intelligence)展现了迅猛的发展势头，因而对增强语音清晰度的要求也越来越高，语音增强显得尤为重要。目前人类已在语音增强算法的研究上取得了很大的进步，主要是谱减法，维纳滤波算法[1]等基于均方误差的统计模型，语音增强的质量得到了提高。但这些语音增强算法中忽略了幅度差值的正负，所以会存在对噪声的过度放大，引入了新的噪声污染，进而影响增强语音的清晰程度。目前，语音增强算法中涉及增益函数主要引入了两种类型的语音失真，一种是处理过后的估计纯净语音信号振幅超过输入纯净语音的放大失真，另一种是经过增强处理的语音信号振幅低于输入纯净语音信号产生的衰减失真。实际研究表明，幅度衰减失真对语音清晰度的影响最小。总的来说，目前数据表明现有算法不能提高语音清晰度的一个主要原因是，它们都允许放大失真超过6分贝。

Loizou 根据增强语音与纯净语音幅度的大小关系，把增强之后的语音失真分为三个部分，分别是幅度衰减部分，幅度放大且放大倍数小于等于2倍(即放大上限为6.02 dB)，幅度放大超过2倍(放大倍数超过6.02 dB)三个部分[2]。现有的语音处理方法，仅对第一和第二部分进行限制，语音质量得到明显的提升，但未对第三部分进行详细有效的约束处理。研究表明，第三部分在增强语音信号中的占比超过35% [3]，具有一定的提升空间。

本文基于 Loizou 仅对第一和第二区域进行约束，结合平方根维纳滤波增益函数特性，提出了针对第三区域的先验信噪比判定条件，进而对增益幅度进行约束[4]。通过 MATLAB 进行大量实验仿真，将提出的算法与现有算法进行比较，结果显示具有较好的语音增强性能。

## 2. 语音增强基本理论

在语音增强技术中，假设  $t$  时刻的带噪语音信号为  $y(t)$ ，纯净语音信号  $x(t)$  和噪声  $n(t)$  互不相关，其时域表达式如下[1]：

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

对上式进行  $N$  点 DFT 变换后，与时域语音对应的语音谱表示如下：

$$Y_{m,k} = X_{m,k} + N_{m,k}, k = 0, \dots, N-1 \quad (2)$$

其中， $Y_{m,k}$ ， $X_{m,k}$  和  $N_{m,k}$  分别表示带噪语音频谱、纯净语音频谱和噪声谱， $m$ ， $k$  分别表示帧索引和频率。

经过增强处理过后的纯净语音频谱估计可以由增益因子  $G_{k,m}$  表示为:

$$\hat{Y}_{m,k} = G_{m,k} \cdot Y_{m,k} \quad (3)$$

其中,  $\hat{Y}_{m,k}$  是估计的纯净语音谱, 增益因子  $G_{m,k}$  一般来说是先验 SNR 和后验 SNR 的函数。

先验 SNR 的定义[1]为:

$$SNR_{prio} = \frac{E\{|X_{m,k}|^2\}}{E\{|N_{m,k}|^2\}}$$

后验 SNR 的定义为:

$$SNR_{post} = \frac{|Y_{m,k}|^2}{E\{|N_{m,k}|^2\}}$$

其中,  $E\{\cdot\}$  表示数学期望算子。

一般来说, 普通维纳滤波语音增强算法的系统增益因子可以表示仅为先验信噪比参数的函数:

$$G_{m,k} = \frac{SNR_{prio}}{1 + SNR_{prio}}$$

通过式, 可获得维纳滤波语音增强系统的输出  $\hat{Y}_{m,k}$ , 再将其通过  $N$  点 IDFT 变换至时域, 即可得到纯净语音信号的估计。

### 3. 改进算法

为了分析估计的纯净语音信号与输入纯净语音振幅的大小关系对噪声抑制的影响, Loizou 定义了剩余信噪比[5]的概念:

$$SNR_{ESI(m,k)} = \frac{|X_{m,k}|^2}{\left(|X_{m,k}| - |\hat{Y}_{m,k}|\right)^2} = \frac{1}{\left(1 - \frac{|\hat{Y}_{m,k}|}{|X_{m,k}|}\right)^2} \quad (5)$$

由上式  $\hat{Y}_{m,k}$  与  $X_{m,k}$  的比值, 可以把语音信号分为三个区间[3]:

区间 I: 在此区域,  $|\hat{Y}_{m,k}| \leq |X_{m,k}|$ , 表示幅度产生了衰减失真。

区域 II: 在此区域,  $|X_{m,k}| < |\hat{Y}_{m,k}| \leq 2|X_{m,k}|$ , 表示幅度产生了放大失真且放大上限为 6.02 dB。

区域 III: 在此区域,  $2|X_{m,k}| < |\hat{Y}_{m,k}|$ , 表示振幅产生超过 6.02 dB 放大。

根据目前的研究表明, 当对第 I 和第 II 区域进行如下限制约束时语音质量有较大的提高[6]:

$$|\hat{Y}_M| = \begin{cases} |\hat{Y}_{m,k}|, & \text{当 } |\hat{Y}_{m,k}| \leq 2|X_{m,k}| \text{ 时} \\ 0, & \text{其他} \end{cases}$$

由于估计的纯净语音与输入的纯净语音的幅度不能直接进行比较, 本文只能由第 III 区域的定义条件

$$2|X_{m,k}| < |\hat{Y}_{m,k}|$$

并根据  $SNR_{ESI}$  与  $\hat{Y}_{m,k}/X_{m,k}$  的关系图, 发现第 III 区域的具有的普遍特点[7]为:

$$SNR_{ESI} \leq 0$$

同时在单一控制条件下，传统的噪声抑制算法会对输入语音产生一定的噪声污染，典型的如维纳滤波。目前已经证明，平方根维纳滤波相比其他比较复杂的降噪算法，具有计算量小且易于实现的优点，且在语音质量和清晰度上仍然能够达到同样的要求。

平方根维纳滤波增益函数  $G_{m,k}$  为[8]:

$$G_{m,k} = \sqrt{\frac{SNR_{prio}}{1 + SNR_{prio}}} \quad (4)$$

其中  $SNR_{prio}$  是先验  $SNR$ ，根据以下方程得[9]:

$$SNR_{prio}(m,k) = \alpha \frac{|\hat{Y}_{m-1,k}|^2}{\lambda_{N(m-1,k)}} + (1-\alpha) \cdot \max \left[ \frac{|\hat{Y}_{m,k}|^2}{\lambda_{N(m,k)}} - 1, 0 \right] \quad (5)$$

其中， $\lambda_{N(m,k)}$  是估计的背景噪声功率谱， $\alpha$  是平滑因子，通常取值为 0.98。

根据平方根维纳滤波增益函数  $G_{m,k}$  特性可知，其增益函数  $G_{m,k}$  在信噪比  $SNR$  为负值时，斜率大约为 1 [8]，即可以估计在第 III 区域时  $G_{m,k} = 1$ 。

$$\begin{aligned} |\hat{Y}_{m,k}| &\geq 2|X_{m,k}| \\ \Rightarrow |\hat{Y}_{m,k}|^2 &\geq 4|X_{m,k}|^2 \\ \Rightarrow G_{m,k}^2 \cdot Y_{m,k}^2 &\geq 4|X_{m,k}|^2 \\ \Rightarrow Y_{m,k}^2 &\geq 4|X_{m,k}|^2 \end{aligned}$$

由目前幅度平方估计:

$$Y_{m,k}^2 \approx X_{m,k}^2 + N_{m,k}^2$$

可直接得到第 III 区域的判定条件为:

$$SNR_{prio} \leq 1/3$$

得到对第 III 区域的幅度约束条件:

$$|\hat{Y}_M| = \beta |\hat{Y}_{m,k}|, \text{ 当 } |\hat{Y}_{m,k}| \geq 2|X_{m,k}| \text{ 时}$$

经过大量实验可知，由在系数  $\beta = 0.68$  时得到语音质量较好。

#### 4. 仿真实验结果分析

为了检验改进算法的性能，本文运用 MATLAB 软件进行实验仿真，选取多种来自语音库的纯净语音 sp01.wav-sp10.wav，噪声来自于噪声库的 white, babble, pink, F16 座舱噪声，输入信噪比分别取 -5 dB, 0 dB, 5 dB, 10 dB。所有信号的采样频率均为 8 kHz/s，仿真实验中语音信号帧长  $K = 256$ ，重叠率为 50%，系数  $\beta = 0.68$ ，以下通过主观语音质量评估(PESQ)，分段信噪比(SegSNR) [10]来对比分析两种语音增强性能。

从表 1 中可以看出在输入信噪比分别为 -5 dB, 0 dB, 5 dB, 10 dB 时输入噪声为的 White, Pink, Babble, F16 座舱四种噪声环境下，改进算法的 PESQ 值明显大于现有的算法，表明本文提出的算法输出的语音与纯净语音更接近，语音的增强效果更好。从表 2 中可以看出在以上四种噪声环境下，改进算法的分段信噪比要优于目前的 I + II 约束算法，具有较好的语音增强性能。因此改进算法在一定噪声环境下，相比目前的算法在语音增强性能上有了一定的提升。

**Table 1.** PESQ data contrast table  
**表 1.** PESQ 数据对比表

噪声类型	算法	输入SNR			
		-5 dB	0dB	5dB	10dB
White	I + II约束	2.8867	3.1922	3.3645	3.5102
	改进算法	<b>2.9188</b>	<b>3.3122</b>	<b>3.5607</b>	<b>3.7298</b>
Babble	I + II约束	2.6560	2.9581	3.1599	3.4051
	改进算法	<b>2.7026</b>	<b>3.0239</b>	<b>3.2612</b>	<b>3.4951</b>
Pink	I + II约束	2.7094	3.0130	3.4486	3.4943
	改进算法	<b>2.7650</b>	<b>3.1123</b>	<b>3.5417</b>	<b>3.6378</b>
F16	I + II约束	2.7882	2.9855	3.3219	3.5265
	改进算法	<b>2.8134</b>	<b>3.0806</b>	<b>3.4749</b>	<b>3.6218</b>

**Table 2.** SegSNR data contrast table  
**表 2.** SegSNR 数据对比表

噪声类型	算法	输入SNR			
		-5 dB	0dB	5dB	10dB
White	I + II约束	6.2465	8.0932	9.7660	11.4656
	改进算法	<b>6.5243</b>	<b>8.2917</b>	<b>9.9304</b>	<b>11.5940</b>
Babble	I + II约束	4.2135	6.1960	8.4592	10.2078
	改进算法	<b>4.5964</b>	<b>6.4487</b>	<b>8.6108</b>	<b>10.3343</b>
Pink	I + II约束	5.7400	7.6298	9.2991	11.2238
	改进算法	<b>6.0056</b>	<b>7.8953</b>	<b>9.4282</b>	<b>11.3494</b>
F16	I + II约束	5.3564	7.2823	9.3195	11.0511
	改进算法	<b>5.6058</b>	<b>7.4841</b>	<b>9.4812</b>	<b>11.1919</b>

## 5. 总结

本文主要研究了基于平方根维纳滤波, 针对噪声被过度放大对语音信号的影响, 提出了对增强语音信号第 III 区幅度限制的算法。根据平方根维纳滤波增益函数的特性, 幅度平方谱估计, 推导出与先验信噪比有直接联系的约束条件, 并且实施起来简单易行。实验结果表明, 新算法在 PESQ, SegSNR 评价标准下具有较好的性能, 能有效地提升语音质量, 提高语音的可懂度。

## 参考文献

- [1] Loizou, P. (2007) Speech Enhancement: Theory and Practice. CRC Press LLC, Boca Raton, Florida.
- [2] Kim, G. and Loizou, P.C. (2011) Gain-Induced Speech Distortions and the Absence of Intelligibility Benefit with Existing Noise-Reduction Algorithms. *Journal of the Acoustical Society of America*, **130**, 1581-1596. <https://doi.org/10.1121/1.3619790>
- [3] Scalart, P. and Filho, J. (1996) Speech Enhancement Based on a Priori Signal to Noise Estimation. *Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing*, Atlanta, 9-9 May 1996, 629-632. <https://doi.org/10.1109/ICASSP.1996.543199>
- [4] Chung, K. (2004) Challenges and Recent Developments in Hearing Aids: Part, I. Speech Understanding in Noise, Microphone Technologies and Noise Reduction Algorithms. *Trends in Amplification*, **8**, 83-124. <https://doi.org/10.1177/108471380400800302>

- 
- [5] Ma, J., Hu, Y. and Loizou, P.C. (2009) Objective Measures for Predicting Speech Intelligibility in Noisy Conditions Based on New Band-Importance Functions. *Journal of the Acoustical Society of America*, **125**, 3387-3405. <https://doi.org/10.1121/1.3097493>
- [6] Loizou, P.C. and Kim, G. (2011) Reasons Why Speech Enhancement Algorithms Do Not Improve Speech Intelligibility and Suggested Solutions. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **19**, 47-56. <https://doi.org/10.1109/TASL.2010.2045180>
- [7] Rangachari, S. and Loizou, P.C. (2006) A Noise-Estimation Algorithm for Highly Non-Stationary Environments. *Speech Communication*, **48**, 220-231. <https://doi.org/10.1016/j.specom.2005.08.005>
- [8] Hu, Y. and Loizou, P.C. (2007) A Comparative Intelligibility Study of Single-Microphone Noise Reduction Algorithms. *Journal of the Acoustical Society of America*, **122**, 1777-1786. <https://doi.org/10.1121/1.2766778>
- [9] Hu, Y. and Loizou, P.C. (2007) Subjective Comparison and Evaluation of Speech Enhancement Algorithms. *Speech Communication*, **49**, 588-601. <https://doi.org/10.1016/j.specom.2006.12.006>
- [10] Lu, Y. and Loizou, P.C. (2011) Estimators of the Magnitude-Squared Spectrum and Methods for Incorporating SNR Uncertainty. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **19**, 1123-1137. <https://doi.org/10.1109/TASL.2010.2082531>

#### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2163-3983, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [hjwc@hanspub.org](mailto:hjwc@hanspub.org)