

文章编号: 1672-2892(2010)04-0431-05

基于频带间相关性的加权噪声功率谱估计

冯 炎^{1,2}, 尼玛扎西¹

(1. 西藏大学 现代教育技术中心, 西藏 拉萨 850012;
2. 西南交通大学 信息科学与技术学院, 四川 成都 610031)

摘要: 噪声功率谱估计是语音增强系统的一个重要组成部分。本文在加权噪声估计的基础上, 考虑了带噪语音在相邻频带间的相关性, 提出了一种新的噪声功率谱估计算法。该算法保留了加权噪声估计算法的优点, 利用频域平滑及时域平滑后的带噪语音来求加权因子, 能够更好地区分弱语音与噪声, 尤其是对强语音后的弱语音与噪声区分更明显, 从而具有更快的跟踪速度及更少的噪声过估计。客观实验和主观实验都证实了本文提出的算法的有效性。

关键词: 语音增强; 噪声功率谱估计; 加权噪声估计

中图分类号: TN912.3

文献标识码: A

Weighted noise power spectrum estimation based on the inter-frequency correlation

FENG Yan^{1,2}, NYIMA Trashi¹

(1. Modern Education Technology Center, Tibet University, Lhasa Tibet 850012, China;
2. School of Information Science & Technology, Southwest Jiaotong University, Chengdu Sichuan 610031, China)

Abstract: Noise power spectrum estimation is a fundamental component of speech enhancement. An improved weighted noise power estimation algorithm is proposed in the paper by taking into account the correlation of inter-frequency in the noisy speech. Because of calculating the weighting factor from the time domain and frequency domain smoothed noisy speech, the new presented algorithm can distinguish noise and weak speech more accurately, especially in the weak speech regions following strong speech. The tracking delay and the overestimates are all considerably reduced compared to competitive noise tracking algorithms. Objective experimental results and a subjective comparison show that the improved noise estimation algorithm is preferred to other noise estimation algorithms when integrated in speech enhancement.

Key words: speech enhancement; noise power spectrum estimation; weighted noise estimation

日常生活中, 移动电话和数字助听器等语音处理设备经常在汽车内和公共场所等嘈杂的噪声环境中使用。语音增强方法能够用来提高这些语音处理设备的性能, 而噪声功率谱估计是语音增强算法的一个关键因素, 先验信噪比的估计也是在噪声功率谱估计的基础上方可进行^[1-3]。较早的方法是通过初始的静音段取平均值获得, 之后便不更新噪声^[4], 这显然不符合噪声的时变特性, 影响了增强后的语音质量。

Cohen 近年提出了最小值控制的递归平滑噪声功率谱估计算法(Minima Controlled Recursive Averaging, MCRA)^[5], 该算法利用了相邻频带间的相关性, 能够更有效地区分噪声与弱语音, 但该算法依赖于最小值的跟踪算法, 无法完全避免在噪声上升区域的噪声欠估计及强语音后面区域出现的噪声过估计的缺陷。

Kato 等人提出的加权噪声估计(Weighted Noise Estimation, WN)^[6]首先由估计的信噪比计算加权因子。通过对带噪语音信号乘以加权因子得到加权值, 进而求平均得到估计的噪声的功率谱。该算法简单有效, 能快速跟踪瞬变噪声, 但在弱语音区域出现噪声过估计, 而在噪声下降时会出现噪声欠估计。本文提出了一种新的噪声功率谱估计算法, 该算法具有更快的跟踪速度及更少的噪声过估计, 因而产生的增强后的语音拥有较好的语音质量及较高的噪声抑制性能。

1 算法描述

WN 算法中的信噪比估计是通过当前帧的带噪语音与前一帧的噪声来获得, 而信噪比估计的准确性决定所估计出的噪声谱密度准确性, 本文提出的算法中的信噪比估计是通过频域平滑和时域平滑后的带噪语音与前一帧的噪声来获得。另外, 本文未采用 WN 算法中的局部平均方法, 而采用了 MCRA 算法中的递归平滑方法。本文的方法是首先对带噪语音进行频域平滑获得带噪语音局部能量; 再用时频相关的平滑因子对局部能量进行时域平滑; 通过平滑后的带噪语音局部能量与前一帧的噪声之比来计算信噪比, 进而计算加权因子; 然后在无语音段或弱语音段对加权后的带噪语音用时域递归平滑算法求噪声功率谱, 而在强语音出现时噪声不更新。

1.1 带噪语音的局部能量

用 $x(t)$ 和 $d(t)$ 分别表示纯净语音和不相关的加性噪声, 观测到的带噪语音信号为 $y(t)$, 进行短时离散傅里叶变换后得到:

$$Y(n, k) = X(n, k) + D(n, k) \quad (1)$$

式中 n 和 k 分别表示时间帧序号和频率带序号。

带噪语音的局部能量 $S_f(n, k)$ 通过在频率域对相邻频带的带噪语音功率进行平滑:

$$S_f(n, k) = \sum_{i=-w}^w b(i) |Y(n, k-i)|^2 \quad (2)$$

式中 b 表示长度为 $2w+1$ 的归一化窗函数 ($\sum_{i=-w}^w b(i) = 1$)。

1.2 时域平滑后的局部能量

对局部能量进行平滑来保护强语音后面出现的弱语音, 通过强语音局部能量对弱语音局部能量做递归平滑, 来提高弱语音谱分量的值。

若用 $I(n, k)$ 表示平滑指示函数, $I(n, k)=1$ 表示该区域是强语音后的弱语音区域, 指示需要在该区域保护弱语音, $I(n, k)=0$ 则表示该区域不属于保护的区域。对于初始时刻, 可认为是噪声, 不进行保护, 即 $I(n, k)=0$ 。若前一帧不是所要保护的区域, 当满足 3 个条件时, 即认为当前帧属于保护的区域。这 3 个条件是: 1) 当前帧局部能量比前一帧小; 2) 前一帧出现强语音; 3) 当前帧属于语音帧。若前一帧是所要保护的区域, 仅需满足 2 个条件之一, 便认为当前帧不属于保护的区域。这 2 个条件是: 1) 前一帧平滑后的局部能量值比当前帧小; 2) 当前帧属于噪声帧。指示函数的计算公式描述如下:

当 $I(n-1, k)=0$ 时,

$$I(n, k) = \begin{cases} 1, & \text{if } S_f(n-1, k) > S_f(n, k) \text{ and } \frac{S_f(n-1, k)}{\hat{\lambda}_D(n-1, k)} > \sigma_{\max} \text{ and } \frac{|Y(n, k)|^2}{\hat{\lambda}_D(n-1, k)} > \sigma_{\min} \\ 0, & \text{others} \end{cases} \quad (3)$$

当 $I(n-1, k)=1$ 时,

$$I(n, k) = \begin{cases} 0, & \text{if } \frac{|Y(n, k)|^2}{\hat{\lambda}_D(n-1, k)} \leq \sigma_{\min} \text{ or } S_{pf}(n-1, k) \leq S_f(n-1, k) \\ 1, & \text{others} \end{cases} \quad (4)$$

式中: σ_{\max} 表示强语音阈值, σ_{\min} 表示噪声阈值; $S_{pf}(n, k)$ 表示平滑后的带噪语音局部能量。

根据上面所求平滑指示函数及平滑规则, 对带噪语音的局部能量进行时域平滑, 方法如下:

$$S_{pf}(n, k) = \begin{cases} \alpha_p(n, k) S_{pf}(n-1, k) + [1 - \alpha_p(n, k)] S_f(n, k), & \text{if } I(n, k) = 1 \\ S_f(n, k), & \text{if } I(n, k) = 0 \end{cases} \quad (5)$$

式中: α_p ($0 < \alpha_p < 1$) 表示平滑因子。

将上式合并简化:

$$S_{pf}(n, k) = \tilde{\alpha}_p(n, k) S_{pf}(n-1, k) + [1 - \tilde{\alpha}_p(n, k)] S_f(n, k) \quad (6)$$

式中: $\tilde{\alpha}_p(n, k)$ 为时频相关的平滑因子, 且 $\tilde{\alpha}_p(n, k) = \alpha_p I(n, k)$ 。

图 1 中给出了平滑后的局部能量与未平滑的带噪语音功率对比。图中显示, 平滑后的局部能量值在语音出现区域明显高于带噪语音功率, 而在无语音区域两者差别不明显, 由此可知, 经过时域平滑后的局部能量将弱语音与噪声区分开来, 从而达到保护弱语音的目的。

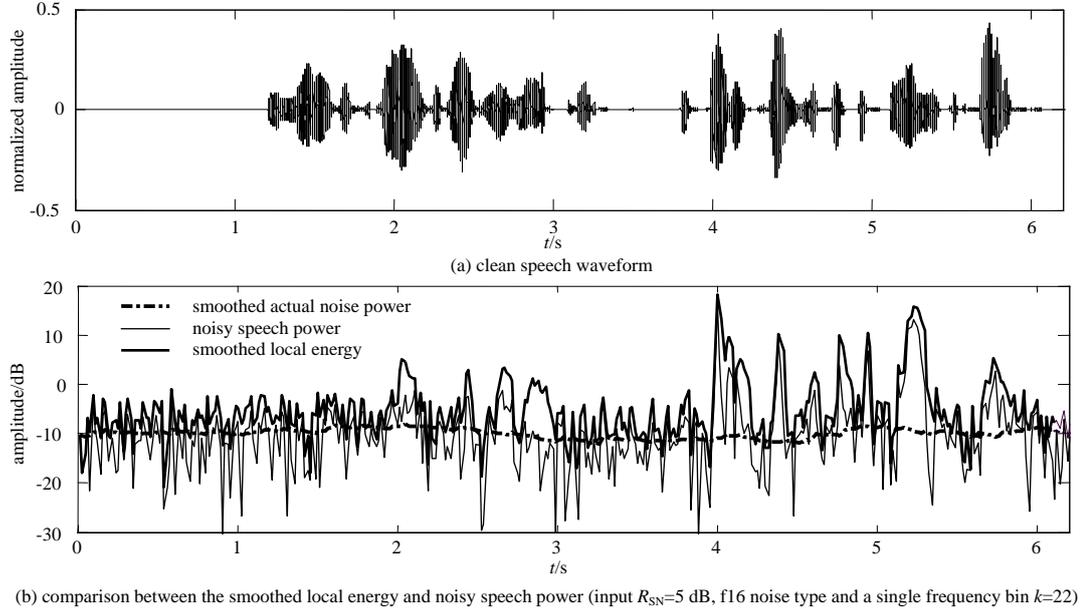


Fig.1 Smoothed local energy
图 1 平滑后的局部能量

1.3 加权因子的计算

根据前面所述, 信噪比的估计是通过时域平滑后带噪语音局部能量与前一帧的噪声功率来获得:

$$\tilde{\gamma}(n,k) = \frac{S_{pf}(n,k)}{\hat{\lambda}_D(n-1,k)} \quad (7)$$

根据上式的信噪比估计, 参考 WN 算法中加权因子的计算方法, 可得加权因子 $W(n,k)$ ^[6]:

$$W(n,k) = \begin{cases} 1, & \text{if } \tilde{\gamma}(n,k) < \tilde{\gamma}_1 \\ \frac{\tilde{\gamma}_2 - \tilde{\gamma}(n,k)}{\tilde{\gamma}_2 - \tilde{\gamma}_1}, & \text{if } \tilde{\gamma}_1 < \tilde{\gamma}(n,k) < \theta_2 \end{cases} \quad (8)$$

式中: $\tilde{\gamma}_1$, $\tilde{\gamma}_2$ 和 θ_2 是加权因子参数。

1.4 噪声功率谱估计

用加权因子对带噪语音进行加权来求噪声估计。当 $\tilde{\gamma}(n,k) \geq \theta_2$ 语音出现的可能性很大, 不更新噪声; 而 $\tilde{\gamma}(n,k) < \theta_2$ 时, 出现的可能是噪声或弱语音, 对带噪语音加权后再用时域递归平滑求噪声功率谱:

$$\hat{\lambda}_D(n,k) = \begin{cases} \hat{\lambda}_D(n-1,k), & \text{if } \tilde{\gamma}(n,k) \geq \theta_2 \\ a_d \hat{\lambda}_D(n-1,k) + (1-a_d)[W(n,k)|Y(n,k)|^2], & \text{others} \end{cases} \quad (9)$$

式中 a_d ($0 < a_d < 1$) 是平滑参数。初始几帧可认为是噪声, T_{init} 表示初始帧的大小, 对初始帧进行平均^[6]:

$$n < T_{\text{init}}, \hat{\lambda}_D(n,k) = \frac{1}{n} \sum_{i=1}^n |Y(i,k)|^2 \quad (10)$$

为了直观显示本文新提出的基于频带间相关性的加权噪声功率谱估计算法的性能优势, 图 2 给出了本文提出的算法与 MCRA, WN 算法的噪声功率谱估计对比。图 2 使用的带噪语音与图 1 相同。图 2 中显示新提出的算法在弱语音区域未出现噪声过估计, 尤其是强语音后的弱语音。从图可以看出, 在初始的纯噪声区域本文的算法与 MCRA 算法均接近于真实噪声, 而 WN 算法欠估计严重; 在 2 s~3 s 之间的弱语音区域, 本文算法未出现过估计, 而 MCRA 算法与 WN 算法都出现过估计; 4 s~6 s 之间的强语音频繁出现区域, MCRA 算法出现了严重过估计, 而 WN 算法出现严重欠估计, 相比之下, 本文算法更接近于真实噪声。由此可以看出, 本文算法总体性能要优于 WN 算法和 MCRA 算法。

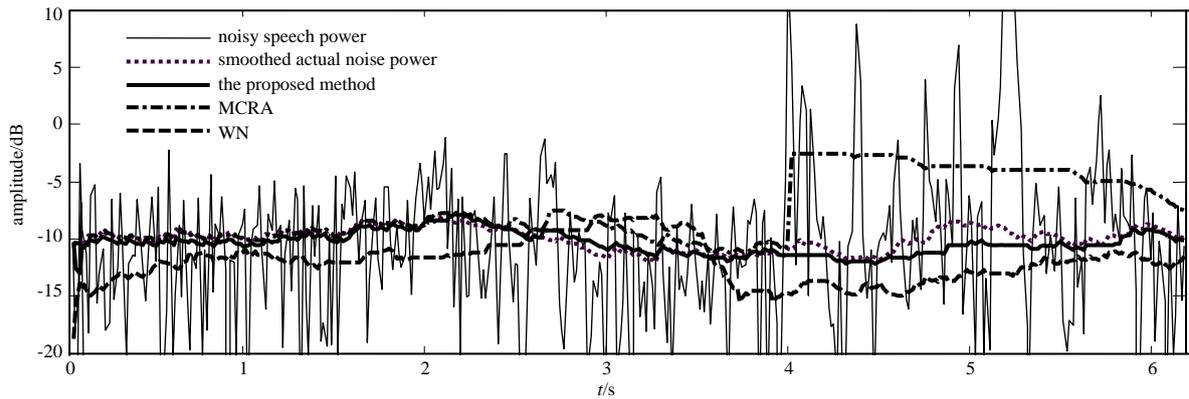


Fig.2 Noise power estimation obtained by the proposed method, MCRA and WN (input $R_{SN}=5$ dB, f16 noise type and a single frequency bin $k=22$)

图2 本文方法与MCRA,WN算法的噪声功率谱估计对比(输入信噪比为5 dB的f16噪声,频率点 $k=22$)

2 仿真实验与结果分析

将本文提出的噪声谱估计算法应用于最小均方误差语音增强系统^[1]。试验中分别采用3种类型的噪声:白噪声(White)、汽车噪声(volvo)以及飞机驾驶舱内噪声(f16),这些噪声取自NOISEX92数据库,噪声信号与语音信号的采样频率均为8 kHz,分别使用不同的信噪比:0.5 dB,10 dB,15 dB。纯净语音取自NOIZEUS数据库。分帧的帧长为256,帧间重叠128,所使用的窗函数为汉明窗。先验信噪比估计中的参数设定^[7]: $\alpha=0.98$, $\xi_{\min}=-25$ dB。MCRA中的参数设定^[5]: $a_d=0.95$, $a_s=0.8$, $a_p=0.2$, $w=1$, $\delta=5$, $L=125$, h 为汉明窗。WN算法的参数设定为^[4]: $T_{\text{init}}=4$, $L_z=20$, $\tilde{\gamma}_1=0$ dB, $\tilde{\gamma}_2=10$ dB, $\theta_z=7$ dB。

本文算法中, $\alpha_d=0.95$ ^[5], $w=1$ ^[5], h 为汉明窗^[5], $T_{\text{init}}=4$ ^[4],因本文采用了 $w=1$ 的汉明窗对带噪声语音进行频域平滑, θ_z , $\tilde{\gamma}_1$, $\tilde{\gamma}_2$ 的取值约为WN算法中的2倍,即: $\theta_z=10$, $\tilde{\gamma}_1=3$, $\tilde{\gamma}_2=18$;其它参数由经验值确定: $\alpha_p=0.5$, $\sigma_{\min}=1$, $\sigma_{\max}=12$ 。

表1给出了本文算法与MCRA,WN的相对估计误差对比。从数据可看出,不同噪声环境和不同输入信噪比的试验中,本文算法的相对估计误差都比WN算法和MCRA算法要小。说明本文算法估计出的噪声功率谱更接近于真实噪声功率谱。表2给出了分别用本文算法、MCRA及WN算法增强后的语音PESQ得分对比实验数据。PESQ是ITU-T P.862建议的语音质量感性评价标准^[8]。本文算法的PESQ得分都优于MCRA及WN算法。充分证实了本文算法的有效性。

表1 本文算法与MCRA,WN的相对估计误差对比

Table1 Average relative estimation error in the proposed method, MCRA and WN

noise type	input R_{SN}/dB	proposed method	WN	MCRA
white	0	0.046 3	0.091 3	0.124 9
	5	0.054 5	0.102 5	0.312 4
	10	0.054 6	0.124 2	1.706 5
	15	0.070 5	0.145 0	8.926 8
volvo	0	0.080 4	0.261 7	0.102 3
	5	0.079 9	0.259 2	0.102 5
	10	0.079 9	0.259 1	0.102 2
	15	0.079 9	0.259 7	0.105 8
f16	0	0.058 7	0.124 0	0.089 2
	5	0.070 2	0.139 3	0.141 0
	10	0.088 7	0.177 6	0.442 4
	15	0.074 0	0.158 7	2.624 3

表2 本文算法与MCRA,WN算法的PESQ得分

Table2 PESQ scores by our proposed approach, MCRA and WN

noise type	input R_{SN}/dB	proposed method	WN	MCRA
white	0	1.905	1.858	1.854
	5	2.330	2.270	2.282
	10	2.659	2.622	2.609
	15	2.924	2.898	2.846
volvo	0	3.472	3.387	3.324
	5	3.812	3.733	3.601
	10	4.081	4.020	3.790
	15	4.253	4.223	3.911
f16	0	2.110	2.069	2.098
	5	2.474	2.435	2.450
	10	2.764	2.741	2.722
	15	3.022	3.015	2.966

3 结论

本文在考虑了相邻频带间的相关性后结合加权噪声估计提出了一种新的噪声功率谱估计算法,该算法能更准确地区分弱语音与噪声,尤其是对强语音后的弱语音与噪声区分更明显,进一步减少了过估计,从而提供更准确的噪声估计。仿真结果表明,本文算法的相对噪声估计误差小于MCRA及WN算法,而且本文算法应用于语音增强系统时,在多种噪声背景下都具有更为出色的语音增强效果。

参考文献:

- [1] Ephraim Y, Malah D. Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator[J]. IEEE Trans. Acoust., Speech, Signal Processing, 1984,32(6):1109-1121.
- [2] 冯炎. 基于直接判决估计和预测估计的语音增强算法[J]. 信息与电子工程, 2010,8(1):76-79. (FENG Yan. Speech enhancement based on Decision-Directed incorporating predicted estimation[J]. Information and electronic engineering, 2010,8(1):76-79.)
- [3] Ephraim Y, Cohen I. Recent advancements in speech enhancement[R/OL]. [2010-03-29]. <http://teal.gmu.edu/%7EYephraim/Papers/crc%202004%20enhan%20ee.pdf>.
- [4] Boll S F. Suppression of acoustic noise in speech using spectral subtraction[J]. IEEE Trans. Acoust., Speech, Signal Process., 1979,ASSP-27(2):113-120.
- [5] Cohen I, Berdugo B. Noise estimation by minima controlled recursive averaging for robust speech enhancement[J]. IEEE Signal Process. Lett., 2002,9(1):12-15.
- [6] Kato M, Sugiyama A, Serizawa M. Noise suppression with high speech quality based on weighted noise estimation and MMSE STSA[C]// Proc. 2001, IWAENC2001:183-186.
- [7] Cappé O. Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor[J]. IEEE Trans. Speech and Audio Proc., 1994,2(2):345-349.
- [8] ITU-T P.862: Perceptual Evaluation of Speech Quality (PESQ), an Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs[S], 2001.

作者简介:



冯炎(1979-),男,山西省吕梁市人,在读硕士研究生,主要研究方向为语音信号处理。email:fy4528@163.com.

尼玛扎西(1964-),男,西藏察隅县人,博士,教授,主要研究方向包括计算机网络、藏文信息处理等。

中国电子学会电路与系统分会第22届学术年会征文通知

中国电子学会第二十二届电路与系统学术年会将于 **2010年10月15-18日** 在上海召开。会议由中国电子学会电路与系统分会主办,复旦大学和中国科学院电子学研究所承办,同济大学和上海交通大学协办。本次年会包括大会邀请报告和学术专场报告。我们真诚地邀请电路与系统及相关领域的科研工作者参加本次年会!现将会议有关事项通知如下:

一、大会特邀报告

- 1.刘瑞文 教授,美国 IEEE/CAS 前主席(题目待定)
- 2.Prof. Wang Wei, State University of New York (王慰 教授,纽约州立大学)
题目: Circuits and Systems Based on 3D IC Integration
- 3.郑立荣 教授,复旦大学
题目: 物联网—从芯片到系统

二、征文范围

本次年会诚征电路与系统领域未发表过的文章。主要包括以下所列主题(但不限于以下主题):混合技术系统、模拟电路与信号处理、多媒体与通讯系统、移动通信技术、数字信号处理、电子设计自动化、一般及非线性电路与系统、VLSI 电路与系统、神经网络、智能系统、图论与系统优化。

会议通知全文可参看 <http://www.iaeej.com> “会讯”。有关会议的最新消息请查看: <http://jeit.ie.ac.cn> “公告栏”。