

基于最小控制 GARCH 模型的噪声估计算法

孟宪波, 鲍长春

(北京工业大学电子信息与控制工程学院语音与音频信号处理实验室, 北京 100124)

摘要: MCRA (Minima-Controlled Recursive Averaging) 方法是经典的噪声估计算法, 然而在语音段 MCRA 方法存在不能对噪声功率谱进行有效更新的问题. 针对这一问题, 本文利用广义自回归条件异方差 (Generalized Autoregressive Conditional Heteroskedasticity, GARCH) 模型在时频域对噪声信号建模, 在 MCRA 算法原理的基础上, 提出了基于最小控制 GARCH 模型的噪声估计算法, 实验结果表明, 本文所提的噪声估计算法能够更为准确估计噪声功率谱, 将该算法应用到语音增强中能够获得到较好的语音增强效果.

关键词: 噪声估计; GARCH 模型; MCRA 算法; 语音增强

中图分类号: TN912.3 **文献标识码:** A **文章编号:** 0372-2112 (2016)03-0747-06

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2016.03.037

Noise Estimate Algorithm Based on Minima Controlled GARCH Model

MENG Xian-bo, BAO Chang-chun

(Speech and Audio Signal Processing Laboratory, School of Electronic Information and Control Engineering,
Beijing University of Technology, Beijing 100124, China)

Abstract: Considering the problem that the typical MCRA (Minima-Controlled Recursive Averaging) noise estimate algorithm fails to update the power spectrum of noise effectively when the speech is present, so this paper proposes a noise estimate algorithm based on minima controlled GARCH model. The noise signal is modeled as a GARCH process in time-frequency domain and then the proposed noise estimate algorithm is achieved combined with the basis of the framework of MCRA method. Experimental and testing results indicate that the proposed algorithm can estimate the spectrum of noise more accurately compared with the reference methods. When the proposed algorithm is applied into speech enhancement, a better performance can be achieved as well.

Key words: noise estimate; GARCH model; MCRA algorithm; speech enhancement

1 引言

语音通信过程中不可避免地受到来自周围环境噪声的干扰. 语音增强的目的是尽可能地抑制噪声, 有效地提高信噪比, 同时保证很高的主客观语音质量, 提高语音的整体听觉效果. 经过最近几十年的发展, 学者们提出了许多种单通道的语音增强算法. 在单通道的语音增强算法中, 噪声估计算法的性能直接影响着增强后语音的效果. 因此, 噪声估计算法一直是各国学者广泛研究的重要课题^[1,2].

基于最小统计量 (MS, Minimum Statistics) 的噪声估计算法^[3]可以在搜索窗中较好地搜索噪声功率谱的最小值, 但该方法对于噪声突变情况适应性较差. MCRA

噪声估计算法把 MS 方法与递归平均方法结合到了一起^[4], 比 MS 方法更为平稳和更具适应性. 然而根据 MCRA 算法原理, 在语音存在的假设条件下, 只是简单的利用前一帧的噪声功率谱来更新当前帧信号的噪声功率谱. 因此, 一旦搜索窗内的语音成分较强时, 此算法不能很好地跟踪噪声信号功率谱的变化.

近些年, GARCH 模型广泛应用在金融预测和统计等领域, 它对波动簇信号的预测有很好的表现, 已经有研究者利用 GARCH 模型对语音信号建模, 并在语音增强中取得了很好的效果, 但在噪声估计中鲜有研究. 为了解决 MCRA 算法在语音成分较强时不能很好的跟踪噪声信号功率谱变化的问题, 本文结合噪声信号的功率谱具有波动簇信号的特点, 尝试利用 GARCH 模型在

时频域对噪声信号建模,在 MCRA 算法原理的基础上,提出了一种基于最小控制 GARCH 模型(MC-GARCH, Minima-Controlled GARCH)的噪声估计算法. 根据 MCRA 算法原理,在语音存在的假设条件下,利用针对噪声信号建立的 GARCH(1,1)模型递推公式对噪声功率谱进行计算更新. 这样就可以在语音存在段实现对噪声功率谱的有效跟踪和更新,得到更为准确的噪声功率谱估计. 实验结果表明,本文提出的基于 MC-GARCH 模型的噪声估计算法,能够更为有效的估计噪声信号功率谱,当将所提算法应用到语音增强算法中时,得到了更好的增强效果.

2 噪声信号的 GARCH 建模方法

2.1 GARCH 原理简介

GARCH 是常用的统计分析模型工具中时间序列模型的一种,其基本原理是用过去时间的误差和方差信息来预测当前时刻的方差. GARCH 对于方差时变的波动簇时间序列预测有很好的效果,这使得 GARCH 在数据处理方面有重要的应用. 2004 年有学者首次将 GARCH 应用到语音信号处理中^[5]. 其基本原理如下.

假设 $\{y_t\}$ 表示随机实值离散时间序列,使 ψ_t 表示在 t 时刻可利用的信息集合,这样就可以得到在已知信息 ψ_{t-1} 条件下的随机实值离散时间序列 $\{y_t\}$ 在 t 时刻的一个预测值,可以表示为 $y_t | \psi_{t-1}$. 定义 ε_t 表示在 $t-1$ 时刻的条件预测误差,那么条件预测误差 ε_t 可以表示为

$$\varepsilon_t = y_t - E\{y_t | \psi_{t-1}\} \quad (1)$$

定义 y_t 在以 $t-1$ 时刻为信息条件下的条件方差为 λ_t ,那么条件方差 λ_t 可用如下公式表示为

$$\lambda_t = \text{var}\{y_t | \psi_{t-1}\} = E\{\varepsilon_t^2 | \psi_{t-1}\} \quad (2)$$

通过用过去 p 个时刻的条件方差和过去 q 个时刻的条件预测误差建立当前时刻条件方差 λ_t 的如下函数关系,

$$\lambda_t = f(\lambda_{t-1}, \dots, \lambda_{t-p}, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2) \quad (3)$$

我们就完成了一个 (p, q) 阶的广义自回归条件异方差模型,记作 GARCH(p, q)^[5].

GARCH(p, q) 模型通常用线性方程表示为^[6]

$$\lambda_t = \kappa + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \lambda_{t-j} \quad (4)$$

式中 $\kappa, \alpha_i, \beta_j$ 是模型参数,且满足

$$\kappa > 0, \alpha_i \geq 0, \beta_j \geq 0$$

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$$

$$i = 1, \dots, q; j = 1, \dots, p$$

由于最为简单的 GARCH(1,1)模型也隐含地考虑到了过去若干时刻信息对当前时刻的影响,因此在实

际应用中 GARCH(1,1)模型最为常用.

2.2 噪声的 GARCH 建模

通过分析可知,噪声信号的 STFT(Short Time Fourier Transform)系数满足利用 GARCH 建模的基本条件. 即噪声信号的 STFT 系数的方差具有波动性,其时频域在统计分布上具有重尾效应,此外,噪声信号具有一个重要的特点:当在同一频点上观察连续时间的噪声信号的 STFT 系数序列时可以发现,相邻时刻的 STFT 系数的幅度高度相关. 因此我们尝试利用 GARCH 模型在时频域对噪声信号建模.

首先,定义实际噪声信号的 STFT 系数谱为 $D_{t,k}$, Θ_{t-1} 表示已知的前一时刻条件信息,那么就可以得到在信息 Θ_{t-1} 为条件下的当前时刻噪声信号 STFT 谱系数的误差,

$$\sigma_e(t, k) = D_{t,k} - E\{D_{t,k} | \Theta_{t-1}\} \quad (5)$$

式中 $\sigma_e(t, k)$ 是噪声信号 STFT 谱系数的预测误差, t 是帧标号, k 是频点序号.

而噪声信号 STFT 谱系数的方差可以表示为

$$\begin{aligned} \sigma_d^2(t, k) &= \text{var}\{D_{t,k} | \Theta_{t-1}\} \\ &= E\{(D_{t,k} - E\{D_{t,k} | \Theta_{t-1}\})^2\} \end{aligned} \quad (6)$$

这里 $\sigma_d^2(t, k)$ 是噪声信号 STFT 谱系数的条件方差.

由于通常情况下假设 $D_{t,k}$ 具有零均值,所以进一步可以得到

$$\begin{aligned} \sigma_e^2(t, k) &= |D_{t,k}|^2 \\ \sigma_d^2(t, k) &= E\{\sigma_e^2(t, k) | \Theta_{t-1}\} \end{aligned} \quad (7)$$

根据 GARCH 建模原理,可以得到关于噪声信号 STFT 谱系数的 GARCH(p, q)模型为

$$\begin{aligned} \sigma_d^2(t, k) &= f(\sigma_d^2(t-1, k), \dots, \sigma_d^2(t-p, k), \\ &\quad \sigma_e^2(t-1, k), \dots, \sigma_e^2(t-q, k)) \end{aligned} \quad (8)$$

为实际应用,定义噪声信号功率谱预测的 GARCH(1,1)模型为

$$\begin{aligned} \sigma_d^2(t, k | t-1) &= \alpha + \rho \sigma_e^2(t-1, k | t-1) \\ &\quad + \eta \sigma_d^2(t-1, k | t-2) \end{aligned} \quad (9)$$

式中式中 α, ρ, η 为模型参数. 满足

$$\alpha > 0, \rho > 0, \eta > 0, \rho + \eta < 1$$

这样,式(9)就是我们所要建立的关于噪声信号 STFT 谱系数的 GARCH(1,1)模型的递推表达式.

3 基于 MC-GARCH 模型的噪声估计

3.1 估计原理

在基于语音存在和语音不存在的两种假设基础上,MCRA 噪声估计算法针对语音存在与否,给出如下表达式^[4]:

$$V_0^{t,k} : \hat{\sigma}_d^2(t, k) = \alpha \hat{\sigma}_d^2(t-1, k) + (1-\alpha) |Y_{t,k}|^2$$

$$V_1^{t,k} : \hat{\sigma}_d^2(t, k) = \hat{\sigma}_d^2(t-1, k)$$

(10)

其中, $V_1^{t,k}$ 和 $V_0^{t,k}$ 分别表示第 t 帧、第 k 个频点语音存在和不存在的两种假设条件, $Y_{t,k}$ 是含噪语音信号的幅度谱, α 为常数因子.

分析式(10)可知, MCRA 噪声估计算法只在语音不存在时更新噪声信号的功率谱估计, 而当语音存在时, 噪声信号每个频点的功率谱估计用上一帧该频点的噪声估计结果近似代替. 因此, MCRA 噪声估计算法不能在语音存在段对噪声进行有效的估计更新, 因此会影响噪声估计算法的准确性.

为解决这一问题, 本文对式(10)进行改进, 在语音存在即 $V_1^{t,k}$ 的假设条件下, 引入噪声信号的 GARCH(1, 1) 模型递推关系来更新当前帧信号的噪声估计.

$$\begin{aligned} V_0^{t,k} : \hat{\sigma}_d^2(t, k | t-1) &= \beta \hat{\sigma}_d^2(t-1, k) + (1-\beta) |Y_{t,k}|^2 \\ V_1^{t,k} : \hat{\sigma}_d^2(t, k | t-1) &= \alpha + \rho \hat{\sigma}_c^2(t-1, k | t-1) \\ &\quad + \eta \hat{\sigma}_d^2(t-1, k | t-2) \end{aligned} \quad (11)$$

式中 β 为常数因子, α, ρ, η 是 GARCH(1, 1) 模型参数, 其满足的约束条件在上节中已给出, $\hat{\sigma}_d^2(t, k | t-1)$ 是在第 $t-1$ 帧噪声功率谱估计条件下的第 t 帧噪声的功率谱估计, $\hat{\sigma}_c^2(t-1, k | t-1)$ 是噪声功率谱估计的条件误差.

从式(10)可以看出, 其得到的噪声功率谱估计实质是以第 $t-1$ 帧噪声功率谱估计为条件下的第 t 帧噪声功率谱估计. 因此式(11)就可以表示为

$$\begin{aligned} V_0^{t,k} : \hat{\sigma}_d^2(t, k) &= \beta \hat{\sigma}_d^2(t-1, k) + (1-\beta) |Y_{t,k}|^2 \\ V_1^{t,k} : \hat{\sigma}_d^2(t, k) &= \alpha + \rho \hat{\sigma}_c^2(t-1, k) + \eta \hat{\sigma}_d^2(t-1, k) \end{aligned} \quad (12)$$

对比式(10)和式(11)可以发现, MCRA 噪声估计方法在语音存在时刻只是利用前一帧的噪声功率谱来更新得到当前帧的噪声功率谱, 这就使得 MCRA 对语音存在段的噪声估计能力较弱, 而本文利用 GARCH 模型在语音存在段对噪声信号建模, 利用了 GARCH 模型对波动信号的预测的性能优势, 能够较好跟踪噪声信号的波动趋势. 从而使得本文所提的噪声估计方法可以更为准确的估计噪声, 尤其是在语音存在段, 本文提出方法的效果更为明显.

当定义第 t 帧、第 k 个频点的语音存在概率 $p_{t,k} = P(V_1^{t,k} | Y_{t,k})$ 时, 噪声功率谱估计可以表示为

$$\begin{aligned} \hat{\sigma}_d^2(t, k) &= (\beta \hat{\sigma}_d^2(t-1, k) + (1-\beta) |Y_{t,k}|^2) (1-p_{t,k}) \\ &\quad + (\alpha + \rho \hat{\sigma}_c^2(t-1, k) + \eta \hat{\sigma}_d^2(t-1, k)) p_{t,k} \end{aligned} \quad (13)$$

当令模型参数 $\alpha = (1-\beta) |Y_{t,k}|^2$, 而参数 $\eta = \beta$ 时, 上式就可以简化为

$$\begin{aligned} \hat{\sigma}_d^2(t, k) &= (1-\beta) |Y_{t,k}|^2 + P_{t,k} \rho \hat{\sigma}_c^2(t-1, k) \\ &\quad + \eta \hat{\sigma}_d^2(t-1, k) \end{aligned} \quad (14)$$

进一步令 $\gamma = 1 - \beta$, 可得

$$\begin{aligned} \hat{\sigma}_d^2(t, k) &= \gamma |Y_{t,k}|^2 + P_{t,k} \rho \hat{\sigma}_c^2(t-1, k) \\ &\quad + \eta \hat{\sigma}_d^2(t-1, k) \end{aligned} \quad (15)$$

式中 γ, ρ, η 为参数因子, 根据实验经验取得, 满足 $\gamma > 0, \rho > 0, \eta > 0, \rho + \eta < 1$. 这里分别取 $\gamma = 0.005, \rho = 0.3, \eta = 0.6$.

从式(15)可以看出本文所提噪声估计算法的递推关系式同样为一个 GARCH(1, 1) 模型的形式, 与 MCRA 噪声算法相似, 本文所提噪声估计算法的递推关系式同样受到 $p_{t,k}$ 的影响, 因此将本文提出的噪声估计算法称为最小控制 GARCH 模型噪声估计算法, 记为 MC-GARCH. 而语音存在概率也是同通过在时间窗内搜索含噪信号的功率谱最小值并与含噪信号平滑功率谱比较判定的方法得到的, 文献[7]中给出了求解语音存在概率 $p_{t,k}$ 的详尽算法, 这里不再赘述.

3.2 噪声估计误差项的更新方法

式(15)中 $\hat{\sigma}_c^2(t-1, k)$ 被定义为噪声信号功率谱估计的误差项, $\hat{\sigma}_c^2(t-1, k)$ 的有效更新对 MC-GARCH 模型噪声估计算法的准确性有直接的影响. 本文结合多带谱减法^[8]的原理实现对其更新, 对于 8kHz 采样的窄带语音信号, 子带划分如表 1 所示. 由图 1 可知, $t-1$ 时刻噪声信号的误差项是利用 $t-1$ 时刻增强语音的功率谱信息得到的.

表 1 子带划分表

子带标号	实际频率 (Hz)	频点
1	0 - 1000	1 - 32
2	1000 - 2000	33 - 64
3	2000 - 4000	65 - 128

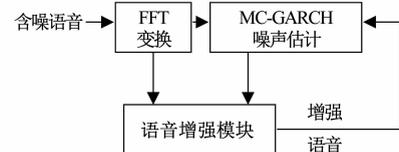


图1 噪声功率谱估计误差更新原理框图

因此, 噪声功率谱的误差 $\hat{\sigma}_c^2(t, k)$ 可表示为

$$\hat{\sigma}_c^2(t, k) = |Y_{t,k}|^2 - \varepsilon_n \varphi_n |\hat{X}_{t,k}|^2 \quad (16)$$

式中, ε_n 为带减系数, φ_n 是过减系数, n 是子带标号, $\hat{X}_{t,k}$ 为语音信号幅度谱估计.

而带减系数 ε_n 不同子带对应不同的值, 即

$$\varepsilon_n = \begin{cases} \varepsilon_1, & f \leq 1\text{kHz} \\ \varepsilon_2, & 1\text{kHz} < f \leq 2\text{kHz} \\ \varepsilon_3, & f > 2\text{kHz} \end{cases} \quad (17)$$

其中 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 为经验因子, 这里取 $\varepsilon_1 = 1.5, \varepsilon_2 = 2.5, \varepsilon_3 = 1.2$.

然后根据含噪信号的功率与增强语音功率谱的比值的方式找到对应的过减系数 φ_n , 定义比值为 SSR_n

$$SSR_n = 10 \log_{10} \left(\frac{\sum_l^h |\bar{Y}_{t,k}|^2}{\sum_l^h |\hat{X}_{t,k}|^2} \right) \quad (18)$$

式中 l 和 h 分别为子带的下边界和上边界, $|\bar{Y}_{t,k}|^2$ 为含噪信号的时域平滑功率谱. 对于不同的 SSR_n 的范围, 可定义过减系数 φ_n 的取值为:

$$\varphi_n = \begin{cases} \varphi_1, & SSR_n < -3 \\ \varphi_2, & -3 \leq SSR_n \leq 15 \\ \varphi_3, & SSR_n > 15 \end{cases} \quad (19)$$

这里 $\varphi_1, \varphi_2, \varphi_3$ 为对应不同频带的实验经验常数因子. 根据实验经验这里取 $\varphi_1 = 3.75, \varphi_2 = 3 - \frac{3}{20} SSR_2, \varphi_3 = 1$.

为了保证噪声功率谱误差 $\hat{\sigma}_e^2(t, k)$ 为正值, 本文做如下处理

$$\hat{\sigma}_e^2(t, k) = \begin{cases} \hat{\sigma}_e^2(t, k), & \text{if } |Y_{t,k}|^2 > \varepsilon_n \varphi_n |\hat{X}_{t,k}|^2 \\ \zeta |\bar{Y}_{t,k}|^2, & \text{else} \end{cases} \quad (20)$$

式中常数因子 ζ 取 0.001

通过以上过程就完成了对噪声信号功率谱估计的误差项更新过程.

4 实验及测试结果分析

本文实验采用的纯净语音信号选自 NTT 标准语音库中的中文字库, 采样率为 8kHz. 噪声信号选自 ITU-T 噪声库.

下面将给出本文所提算法和参考算法得到的在特定频率处的噪声功率谱估计曲线.

图 2 和图 3 分别是在白噪声和有色噪声 (street) 条件下的含噪语音信号频谱图和在频率约为 680Hz 处 MC-GARCH 模型法、MCRA 方法和 MS 方法所对应的噪声功率谱估计曲线比较示例. 从中可以看出, 本文提出的 MC-GARCH 模型算法对应的噪声功率谱估计能够较好的跟踪噪声信号功率谱的趋势. 特别是在有语音成分存在的时间段, MC-GARCH 模型算法能够较好地更新噪声功率谱估计, 这是因为本文在语音存在时段引入了 GARCH 对噪声信号建模, 而 GARCH 模型对于在时频域上具有波动簇特性的噪声信号有很好的估计效果. 因此, 相对 MCRA 方法, MC-GARCH 模型法可以在语音存在时段对噪声功率谱估计进行更为有效的更新.

针对噪声估计算法在噪声估计方面的性能表现, 本文采用计算均方误差 (MSE) 的方法对所提出的噪声估计算法和参考方法进行测评. 它能反映估计噪声与

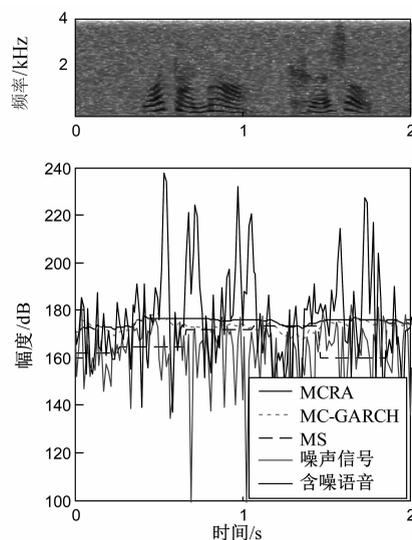


图2 白噪声功率谱估计比较示例 (SNR=10dB, 频点 $k=25$, 约为680Hz)

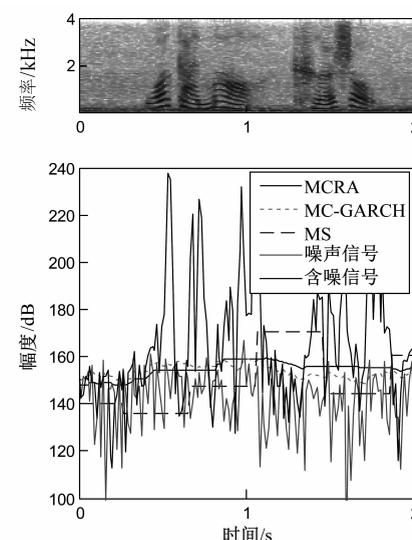


图3 有色噪声功率谱估计比较示例 (SNR=10dB, 频点 $k=25$, 约为680Hz)

实际噪声之间的接近程度, 相对估计误差值越小则表明估计噪声功率谱越接近实际噪声的功率谱^[9].

相对估计误差 (MSE) 定义如下:

$$MSE = \frac{1}{L} \sum_{t=0}^{L-1} \frac{\sum_k [\hat{\sigma}_d^2(t, k) - \sigma_d^2(t, k)]^2}{\sum_k (\sigma_d^2(t, k))^2} \quad (21)$$

式中 L 为总帧数, K 是频点个数, t 是帧号, k 为频点序号, MSE 是相对估计误差, $\hat{\sigma}_d^2(t, k)$ 为噪声功率谱估计, $\sigma_d^2(t, k)$ 实际噪声功率谱.

在信噪比为 0dB、5dB、10 dB 和 15dB 下, 本文分别对白噪声、street 噪声、volvo 噪声和 babble 噪声功率谱估计进行了 MSE 测评, 结果如表 2~5 所示.

分析各表可以看出在不同信噪比和不同噪声环境中,本文提出的 MC-GARCH 模型噪声估计方法 MSE 均明显小于参考算法,这说明本文提出的算法对噪声的估计更为准确.原因是 MC-GARCH 模型噪声估计方法可以在语音存在段有效的更新噪声估计,因此本文所提算法比参考算法更为准确有效.

为进一步说明本文所提方法的优势,图 4 给出了一段基于统计模型^[7]的增强语音的语谱比较示例,噪声为街道噪声,输入信噪比为 10dB,使用的噪声估计方法分别为 MS 法、MCRA 法和 MC-GARCH 模型法,图 4(a)~图 4(d)分别对应含噪语音的语谱、用 MS 增强的语谱、用 MCRA 增强的语谱和用 MC-GARCH 模型增强的语谱.从频谱图可以看出本文提出的 MC-GARCH 模型噪声估计算法对应的图 4(d)中的噪声残留明显少于参考方法,并且语音成分也得到了较好的保留.尤其是在语音存在时间段,与参考方法相比,MC-GARCH 模型方法对应的图 4(d)中的噪声得到了较好的消除.

表 2 白噪声功率谱估计的 MSE 比较

信噪比	噪声估计方法		
	MS	MCRA	MC-GARCH
0dB	0.730213	0.609153	0.570238
5dB	0.740615	0.618945	0.591536
10dB	0.751392	0.631205	0.606321
15dB	0.761438	0.650141	0.621538

表 3 street 噪声功率谱估计的 MSE 比较

信噪比	噪声估计方法		
	MS	MCRA	MC-GARCH
0dB	0.778213	0.638126	0.615468
5dB	0.795214	0.657317	0.620032
10dB	0.801536	0.670346	0.642857
15dB	0.810635	0.686391	0.665971

表 4 volvo 噪声功率谱估计的 MSE 比较

信噪比	噪声估计方法		
	MS	MCRA	MC-GARCH
0dB	0.983510	0.867143	0.844516
5dB	1.098424	0.892517	0.859618
10dB	1.132685	0.912551	0.880674
15dB	1.270539	0.936427	0.907201

表 5 babble 噪声功率谱估计的 MSE 比较

信噪比	噪声估计方法		
	MS	MCRA	MC-GARCH
0dB	0.780125	0.639217	0.617385
5dB	0.791270	0.660135	0.636094
10dB	0.802583	0.682350	0.660237
15dB	0.818634	0.700215	0.678901

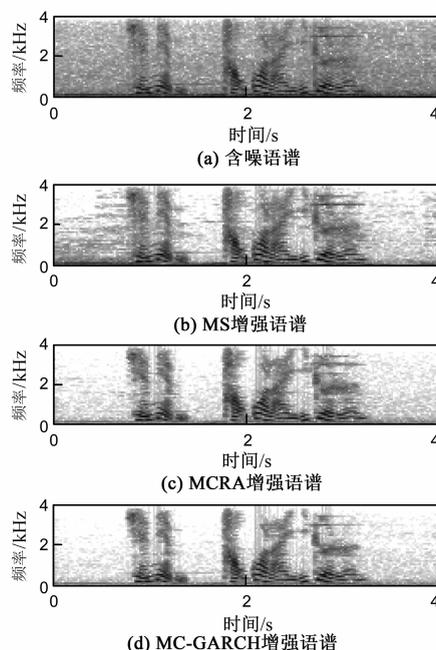


图 4 10dB 街道噪声环境下增强语音语谱比较示例

5 结论

针对经典 MCRA 噪声估计方法在语音段不能对噪声信号进行有效更新的问题.本文利用 GARCH 模型在时频域对噪声信号建模,在 MCRA 算法原理的基础上,提出了基于 MC-GARCH 模型的噪声估计算法,实验及测试结果表明,与参考算法相比,本文提出的基于 MC-GARCH 噪声估计算法可以更为有效地跟踪噪声信号的变化,能够较为准确的进行噪声信号功率谱估计,当应用到语音增强后算中可以获得较好的增强效果.

参考文献

- [1] Hendriks R C, Heusdens R, Jensen J. MMSE based noise PSD tracking with low complexity [A]. 2010 International Conference on Acoustics, Speech, and Signal Processing [C]. Dallas, Texas, USA: IEEE Press, 2010. 4266 - 4269.
- [2] 张东方, 蒋建中, 张连海. 一种改进型 IMCRA 非平稳噪声估计算法 [J]. 计算机工程, 2012, 38(13): 270 - 272. Zhang Dong-fang, Jiang Jian-zhong, Zhang Lian-hai. Improved IMCRA non-stationary noise estimation algorithm [J]. Computing Engineering, 2012, 38(13): 270 - 272
- [3] Martin R. Spectral subtraction based on minimum statistics [J]. European Signal Processing Conference, 1994, 1182 - 1185.
- [4] Cohen I, Berdugo B. Noise estimation by minima controlled recursive averaging for robust speech enhancement [J]. Signal Processing Letters, IEEE, 2002, 9(1): 12 - 15.
- [5] Cohen I. Modeling speech signals in the time-frequency do-

- main using GARCH model[J]. European Signal Processing, 2004, 84(12):2453-2459.
- [6] Cohen I, Speech spectral modeling and enhancement based on generalized autoregressive conditional hetero-skedasticity models[J]. European Signal Processing, 2006, 86(4):698-709.
- [7] Loizou P. Speech Enhancement: Theory and Practice[M]. Boca Raton, FL, USA; CRC Press, 2007. 435-439.
- [8] Kamath S, Loizou P. A multi-band spectral subtraction method for enhancing speech corrupted by colored noise [A]. IEEE International Conference on Acoustics Speech and Signal Processing[C]. Orlando, Florida, USA; IEEE, 2002. 4164-4164.
- [9] 周旋. 基于统计模型和经验模式分解的宽带语音增强技术研究[D]. 北京:北京工业大学, 2012.

作者简介



孟宪波 男, 1987 年出生, 河北承德人, 北京工业大学硕士研究生, 主要研究方向为语音与音频信号处理.

E-mail: mengxianbo@emails.bjut.edu.cn



鲍长春 (通信作者) 男, 1965 年出生, 内蒙古赤峰人, 博士, 北京工业大学教授、博士生导师, IEEE 高级会员, 国际语音通信学会 (ISCA) 会员, 亚太信号与信息处理学会 (APSIPA) 会员, 中国电子学会理事, 中国声学学会理事, 信号处理专业委员会委员. 主要研究方向为语音与音频信号处理.

E-mail: chchbao@bjut.edu.cn