

基于相干性滤波器的广义旁瓣抵消器麦克风小阵列语音增强方法

杨立春^{*①②} 钱运涛^①

^①(浙江大学计算机科学与技术学院 杭州 310027)

^②(浙江万里学院智能控制技术研究所 宁波 315101)

摘要: 为了克服传统麦克风小阵列语音增强算法噪声抑制能力有限的问题, 该文提出一种基于相干性滤波器的广义旁瓣抵消器语音增强算法, 该算法基于动态平滑系数噪声谱估计来获得相干性滤波器, 分别对每个阵元接收到的信号进行滤波用以抑制包括混响等噪声信号的干扰, 并把滤波后的信号作为输入信号, 使用基于小阵列的广义旁瓣抵消器波束形成算法抑制残余噪声信号的干扰。模拟和实际试验表明, 该文提出的算法明显优于单独使用小阵列波束形成算法和相干性滤波器算法。

关键词: 语音信号处理; 麦克风小阵列; 语音增强; 相干性滤波器; 广义旁瓣抵消器

中图分类号: TN912.3

文献标识码: A

文章编号: 1009-5896(2012)12-3027-07

DOI: 10.3724/SP.J.1146.2012.00557

Speech Enhancement with Generalized Sidelobe Canceller Based on a Coherence-based Filter for Small Microphone Arrays

Yang Li-chun^{①②} Qian Yun-tao^①

^①(College of Computer Science, Zhejiang University, Hangzhou 310027, China)

^②(Intelligent Control Research Institute, Zhejiang Wanli University, Ningbo 315101, China)

Abstract: In order to overcome the problem that the noise suppression ability of traditional small microphone arrays speech enhancement algorithm is limited, this paper proposes coherence spectral estimation based on the variable coefficients noise filter algorithm to suppress the reverberant interference, and then use the Generalized Sidelobe Canceller (GSC) which based on small microphone arrays canceller to suppress remaining noise and simulations and practical tests show that the proposed algorithm is superior to separate small array beamforming algorithms and coherence filter algorithm.

Key words: Speech signal processing; Small microphone arrays; Speech enhancement; Coherence filter; Generalized Sidelobe Canceller (GSC)

1 引言

麦克风阵列语音增强算法由于具有较多的时域和空域信息, 因而可以获得比单道语音增强算法更高质量的目标语音, 包括更少的目标语音畸变和更好的噪声抑制。而源于窄带信号如雷达、声呐等的波束形成算法是麦克风阵列中常用的语音增强算法, 这些波束形成算法往往要求具有较多的阵元以及较大的阵元间距^[1]。但在空间较小、运算能力有限又对语音质量要求较高的如手机、助听器等设备上, 这些传统的波束形成算法一般难以直接使用。为了

解决这个问题, 基于小阵列的波束形成算法以及根据噪声相关性进行噪声互功率谱密度估计的相干性滤波器方法被广泛研究和应用。

麦克风小阵列波束形成算法包括超指向性算法^[2,3], 一阶差分麦克风阵列^[4]以及基于小阵列的广义旁瓣抵消器算法^[5,6]等。在这几种基于小阵列的波束形成算法中, 广义旁瓣抵消器相对其他算法而言, 不仅可以工作在端射情况, 也可以工作在侧射情况, 并易于实现, 因而被广泛研究和改进。广义旁瓣抵消器算法由文献[7]提出, 主要有 3 部分组成: 固定波束形成器(Fixed BeamFormer, FBF), 阻塞矩阵(Blocking Matrix, BM)以及自适应干扰抵消器(Adaptive Interference Canceller, AIC) 3 部分组成。

固定波束形成器一般由延迟求和波束形成器(Delay-and-Sum Beamformer, DSB)实现, 主要目

2012-05-11 收到, 2012-09-27 改回

国家自然科学基金(61171151), 国家 973 计划项目(2012CB316400), 华为科技项目(YBCB2010059-2), 宁波市创新创业项目(2011B710010)和宁波市重大择优委托项目(2011B1017)资助课题

*通信作者: 杨立春 lichun_y@126.com

的是增强主瓣方向目标信号，而抑制其他方向的噪声信号；阻塞矩阵的目的是产生不含目标语音的信号，为在自适应干扰抵消器中抑制固定波束形成器中的噪声提供参考信号；自适应干扰抵消器一般使用归一化最小均方(Normalized Least Mean Square, NLMS)^[8,9]及其改进算法^[10]实现，用以实现利用参考噪声信号估计出固定波束输出信号中的噪声信号，然后从固定波束形成输出信号中减去该估计信号，以实现目标信号的增强。

但是波束形成方法包括广义旁瓣抵消器等对抑制混响和非相关噪声能力有限^[11,12]。为了解决波束形成方法对于混响干扰抑制能力有限的问题，基于相干性滤波器方法(coherence based filter)在文献[13]中被提出用以解决混响干扰的问题，其基本思想是假定相邻麦克风间目标语音信号是相关的而噪声信号不相关，这样可以很容易利用信号的互功率谱密度，抑制所有噪声信号包括混响干扰。但实际情况中相邻麦克风，尤其是对小阵列情况，噪声信号往往是相关的。为了解决这个问题，基于噪声谱估计的相干性滤波器^[14,15]成为改进的一个重要方向，其中文献[14]提出基于迭代的噪声谱估计方法，可以估计包括语音段和非语音段噪声谱，且目标语音畸变较小，该方法的不足是由于使用固定的平滑系数来估计噪声谱，导致不能快速跟踪噪声信号的时变特性。

本文提出一种结合基于相干性滤波器和广义旁瓣抵消器的麦克风小阵列语音增强方法，使用基于相干性滤波器作为广义旁瓣抵消器的预处理器，对每个阵元接收到的信号进行滤波，用以抑制非相关性噪声和混响干扰，并把滤波后的信号作为广义旁瓣抵消器的输入信号，实现抑制残余的噪声信号干扰目标。本文的基于相干性滤波器方法中的噪声谱估计是基于文献[14]中的方法并对其进行了改进，本文的方法首先使用语音活动检测(Voice Activity Detector, VAD)^[16]区分语音段和非语音段，然后在语音段和非语音段使用不同的平滑系数对噪声谱进行估计，以便能够快速跟踪噪声谱的变化。

另外本文的算法中广义旁瓣抵消器中使用的是—种扩展的变步长 NLMS 算法，与其他类型 NLMS 权系数在语音段不进行更新相比，本文的 NLMS 算法在语音段使用一个合理值，在保证语音畸变较小的情况下，对权系数进行更新，以便能最大限度跟踪噪声信号的变化。

本文的组织如下：第 2 节介绍了在混响环境中的麦克风阵列信号模型；第 3 节介绍了本文提出的一种基于 VAD 的迭代相干性滤波器方法作为预处

理器结合一种扩展 NLMS 算法的广义旁瓣抵消器的语音增强算法；第 4 节分别在仿真和实际环境中对本文的方法与其他相关的方法进行比较，以证明本文方法的有效性；最后是本文的总结。

2 信号模型

在封闭室内环境中，如房间、礼堂、会议室等场所，声音信号在传播过程中会通过地板、房顶、墙壁以及其他物体的不断反射和衰减而形成一种称为混响的现象，每个信号在封闭环境中遇到物体不断反射而逐渐削弱，这个过程可用脉冲响应函数来描述，麦克风接收到的某一时刻目标信号为目标语音信号与脉冲响应函数的卷积，此时每个麦克风接收到的含噪语音信号可以表示为

$$y_i(n) = a_i * x_i(n) + v_i(n), \quad i = 1, 2, \dots, n \quad (1)$$

式(1)中 a_i 为房间的脉冲响应函数；* 表示卷积操作。对式(1)两边分别进行短时傅里叶变换使其变成频域形式：

$$Y_i(\omega, k) = A_i X_i(\omega, k) + V_i(\omega, k), \quad i = 1, 2, \dots, n \quad (2)$$

式(2)中的 ω 表示频率因子， k 表示帧因子。

对于两个麦克风接收到的相干性函数 $\Gamma_{Y_1 Y_2}(\omega, k)$ 可用式(3)定义：

$$\Gamma_{Y_1 Y_2}(\omega, k) = \frac{\Phi_{Y_1 Y_2}(\omega, k)}{\sqrt{\Phi_{Y_1 Y_1}(\omega, k) \Phi_{Y_2 Y_2}(\omega, k)}} \quad (3)$$

式(3)中 $\Phi_{Y_1 Y_2}(\omega, k)$ 是互功率谱密度并定义为

$$\Phi_{Y_1 Y_2}(\omega, k) = E(Y_1(\omega, k) Y_2^*(\omega, k)) \quad (4)$$

式(3)中的 $(\cdot)^*$ 表示共轭操作， $E(\cdot)$ 表示数学期望。另外式(3)中 $\Phi_{Y_i Y_i}(\omega, k)$ 表示信号的功率谱密度，定义如下：

$$\Phi_{Y_i Y_i}(\omega, k) = E(Y_i(\omega, k) Y_i^*(\omega, k)) \quad (5)$$

3 基于相干性滤波器的广义旁瓣抵消器算法

本文的语音增强算法结构如图 1 所示。从图中可以看出，算法主要分为两部分：相干性滤波器和广义旁瓣抵消器，主要思想是利用基于 VAD 的迭代相干性滤波器作为广义旁瓣抵消器的前置预处理滤波器，用以抑制包括混响和非相关性噪声信号在内的干扰，然后使用基于变步长 NLMS 的广义旁瓣抵消器抑制残留的噪声信号。为了简化说明，下面以二元麦克风小阵列为例来介绍本文提出的算法。

3.1 基于 VAD 的迭代相干性滤波器方法

迭代相干性滤波器可以抑制混响干扰，其思想是利用迭代的方法估计出含噪信号中的包含混响在内的噪声谱，进而实现噪声抑制的目的。该方法可

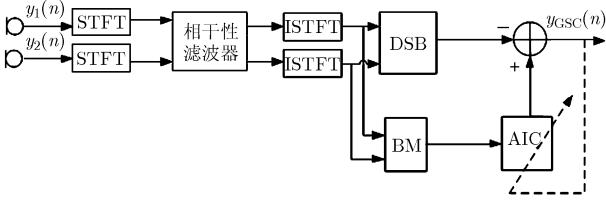


图 1 基于相干性滤波器的广义旁瓣抵消器算法结构图

以在没有使用 VAD 的情况下进行噪声谱估计,但需假定噪声谱的变化较慢而语音谱的变化较快,这样可以使用较小的平滑系数跟踪快速变化的目标语音谱,而使用较大的平滑系数跟踪变化速度较慢的噪声谱,对于第 k 帧噪声谱 $\Phi_{1V_2}(\omega, k)$, 其估计通过式(6)实现:

$$\Phi_{1V_2}(\omega, k) = \lambda_v \Phi_{1V_2}(\omega, k-1) + (1-\lambda_v) Y_1(\omega, k) \cdot Y_2^*(\omega, k) (1-H(\omega, k-1)) \quad (6)$$

式(6)中 λ_v 是噪声谱平滑系数, $H(\omega, k-1)$ 为上一帧的相干性滤波器函数。从式(6)可以看出当前帧噪声谱不但同上一帧噪声谱相关,还与上一帧滤波器相关。第 k 帧相干性滤波器函数 $H(\omega, k)$ 通过式(7)求得

$$H(\omega, k) = \frac{R_{\text{prio}}(\omega, k)}{R_{\text{prio}}(\omega, k) + 1} \Gamma_{Y_1 Y_2}(\omega, k) \quad (7)$$

式(7)中的 $\Gamma_{Y_1 Y_2}(\omega, k)$ 为相干性函数,由式(3)定义; $R_{\text{prio}}(\omega, k)$ 表示先验信噪比,由式(8)定义:

$$R_{\text{prio}}(\omega, k) = \lambda_r H^2(\omega, k-1) \frac{Y_1(\omega, k-1) Y_2^*(\omega, k-1)}{\Phi_{1V_2}(\omega, k-1)} + (1-\lambda_r) R_{\text{post}}(\omega, k) \quad (8)$$

式(8)中的 λ_r 表示先验信噪比平滑系数; $R_{\text{post}}(\omega, k)$ 表示后验信噪比,并通过式(9)定义:

$$R_{\text{post}}(\omega, k) = \max \left[\frac{Y_1(\omega, k) Y_2^*(\omega, k)}{\Phi_{1V_2}(\omega, k)} - 1, 0 \right] \quad (9)$$

另外式(7)中使用的相干性函数中的信号功率谱密度和互功率谱密度也是通过迭代方法估计出来的,其中每个麦克风中的含噪信号功率谱密度迭代估计公式为

$$\Phi_{iY_i}(\omega, k) = \lambda_y \Phi_{iY_i}(\omega, k-1) + (1-\lambda_y) \cdot Y_i(\omega, k-1) Y_i^*(\omega, k-1), \quad i=1,2 \quad (10)$$

上式中 λ_y 是含噪语音谱平滑系数。两个麦克风之间的含噪语音互功率谱密度估计公式为

$$\Phi_{Y_1 Y_2}(\omega, k) = \lambda_y \Phi_{Y_1 Y_2}(\omega, k-1) + (1-\lambda_y) Y_1(\omega, k) Y_2^*(\omega, k) \quad (11)$$

由于上述未用 VAD 的迭代相干性滤波器方法要求噪声信号比语言信号变化慢,导致噪声抑制能力有限。

为了克服这个问题同时加快噪声谱和含噪语音谱的估计,我们在上述方法基础上引入了 VAD^[16] 算法,然后分别对语音段和非语音段采用不同的平滑系数用以加快对噪声谱和含噪信号谱的估计。由于较小的平滑系数可以比较大的平滑系数能更快地跟踪信号变化,在语音段使用较小的含噪语音谱平滑系数和较大的噪声谱估计平滑系数,这样可以较快地跟踪到语音信号的变化,并可以防止由于噪声谱估计错误而导致的语音信号的畸变,在非语音段则使用较小的噪声谱估计平滑系数,以便加快噪声谱估计。

式(12)和式(13)从直观的角度说明了本文采用的含噪信号互功率谱密度估计的平滑系数 λ_y 和噪声信号互功率谱密度估计的平滑系数 λ_v 的取值方法:

$$\lambda_y = \begin{cases} \lambda_1, & \text{语音帧} \\ \lambda_2, & \text{非语音帧} \end{cases} \quad (12)$$

$$\lambda_v = \begin{cases} \lambda_2, & \text{语音帧} \\ \lambda_1, & \text{非语音帧} \end{cases} \quad (13)$$

由前面描述可知,式(12)和式(13)中的 λ_1 与 λ_2 关系为 $\lambda_1 < \lambda_2$ 。

为了能够利用相干性滤波输出结果进行 GSC 语音增强,首先需要估计目标信号相对两个麦克风的位置关系^[17,18],然后在进行功率谱估计前首先需对两个麦克风接收到的信号进行相位补偿。假定目标信号传播方向与阵列的麦克风 1 一侧方向的夹角为 θ ,在小阵列中,相邻麦克风之间的脉冲响应函数可以近似地认为是相等的或者认为其区别可以忽略不计,则第 1 个麦克风接收到的信号乘以它们的相位差 $e^{-j\omega d \cos \theta / c}$ 后,其目标信号与第 2 个麦克风接收到的目标信号基本相同,然后分别对这两个信号使用相干性滤波器进行去混响处理,得到的经过相干性滤波器后获得的信号 $Y_{c1}(\omega, k)$ 和 $Y_{c2}(\omega, k)$:

$$Y_{c1}(\omega, k) = Y_1(\omega, k) e^{-j\omega d \cos \theta / c} H(\omega, k) \quad (14)$$

$$Y_{c2}(\omega, k) = Y_2(\omega, k) H(\omega, k) \quad (15)$$

注意到经过相位补偿后的两个信号所含目标语音信号基本相同,因而经过相同滤波器滤波后的信号 $Y_{c1}(\omega, k)$ 和 $Y_{c2}(\omega, k)$ 中的目标语音信号也基本相等。

3.2 广义旁瓣抵消器改进算法

广义旁瓣抵消器噪声抑制的方法是利用参考噪声信号估计 DSB 输出中的噪声信号。麦克风接收到的信号经过相干性滤波器后的信号 $Y_{c1}(\omega, k)$ 和 $Y_{c2}(\omega, k)$ 已对混响和部分噪声进行了抑制,且它们所包含的目标信号基本相同,因而可以得到 DSB 输出为

$$Y_{\text{DSB}}(\omega, k) = (Y_{c1}(\omega, k) + Y_{c2}(\omega, k)) / 2 \quad (16)$$

与 DSB 输出中的残余噪声信号相关的参考信号为

$$V_{\text{ref}}(\omega, k) = Y_{c1}(\omega, k) - Y_{c2}(\omega, k) \quad (17)$$

然后分别把式(16)和式(17)转换到时域,再利用变步长 NLMS 算法对残余噪声进行降噪处理。

传统的 NLMS 算法在语音段滤波器权系数固定不变,这虽然不会导致目标信号的畸变,但同时也很难跟踪时变的非平稳信号,当语音段连续很长时,这种问题变得更加严重。本文采用了两种步长分别估计在噪声段和语音段的噪声信号,用以平衡语音畸变和噪声估计。具体思路是在噪声段采用相对较大的步长来更新滤波器的权系数,以加快噪声的估计,而在语音段则使用较小的步长,对权系数进行更新,同时其更新需考虑到噪声抑制和语音畸变之间的平衡,以语音没有明显畸变为前提,这样可以最大限度地保证能够在较长语音段内跟踪噪声信号的变化。也就是说滤波器步长 μ 满足式(18)要求:

$$\mu = \begin{cases} \mu_1, & \text{语音段} \\ \mu_2, & \text{非语音段} \end{cases} \quad (18)$$

式(18)中的滤波器步长 μ_1 和 μ_2 均为常数,且前者小于后者。

经过本文算法处理后的信号,一般还会含有少量噪声干扰,可以使用单道降噪算法进一步进行处理。

4 实验分析

实验分为仿真实验和实际录音环境实验,采样频率均为 16 kHz,另外为了验证目标信号相对阵列不同位置时噪声抑制的效果,在仿真实验中分别进

行了端射和侧射位置的实验。我们选取 GSC, GSC-OMLSA^[19]方法以及 coherence-based^[14]作为比较。VAD 采用文献[16]提出的方法,短时傅里叶变换采用 hamming 窗,每帧长度为 1024 个采样点,短时傅里叶反变换采用 50% 的重叠相加法,式(12)和式(13)中的 λ_1, λ_2 分别取 0.1 和 0.9。

4.1 仿真实验

仿真实验中,两个麦克风间距为 20 mm。第 1 个实验目标声源为端射方向并位于第 1 个麦克风一侧,距离第 1 个麦克风为 15 cm。噪声来自 Noise-92 数据库。混响仿真环境通过 Habets^[20]算法实现,该算法基于 Allen 和 Berkley 的图像方法^[21]。麦克风阵列位于大小为宽 3 m、长 6 m 以及高为 2.8 m 的房间内,混响时间 $RT_{60}=0.4$ s。

另外两个麦克风的坐标(单位:m。下同)分别为 (1.92 2.5 1.6)和(1.94 2.5 1.6)。目标声源位于(1.77 2.5 1.6),两个 babble 干扰噪声源分别位于(2 3 1.6)和(2.5 1 1.7)。

图 2(a)和图 2(b)分别为在上述仿真条件下麦克风 1 接收到的纯净语音信号和含噪语音信号时域和频域图;图 2(c)-2(f)为 4 种不同基于麦克风小阵列算法对噪声处理后的时域和频域图,可以很明显看出,本文的算法在端射情况下对噪声抑制效果非常突出。

第 2 个实验的噪声源位置同第 1 个实验,唯一不同的是目标声源位于侧射方向(1.93 2.65 1.6)。处理后的时域和频域效果如图 3 所示,从中可以看出本文算法表现较优。

第 3 个实验为了验证在不同信噪比(Signal-to-Noise Ratio, SNR)情况下处理结果比较,仿真环境

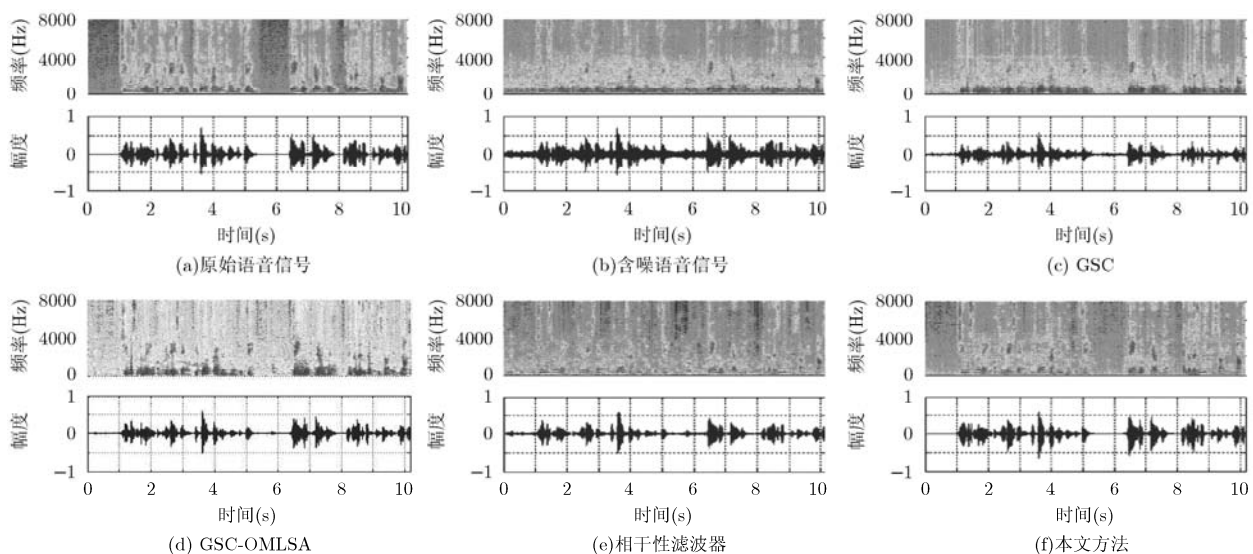


图2 端射方向、背景噪声为 babble、 $RT_{60}=0.4$ s 环境下处理前后时频图

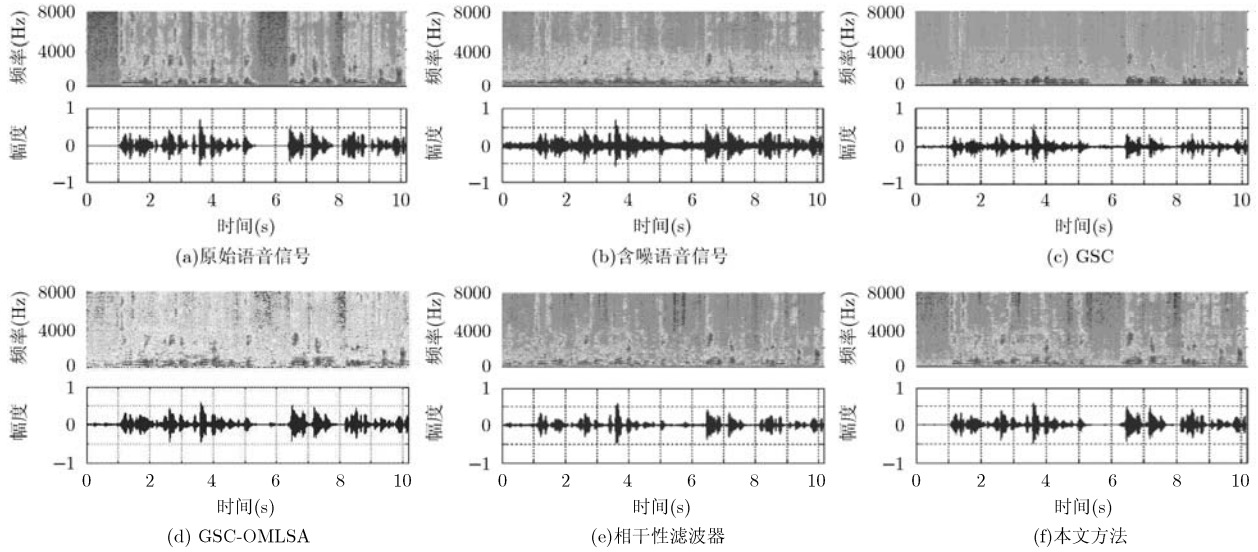


图 3 侧射方向、背景噪声为 babble、 $RT_{60}=0.4$ s 环境下处理前后时频图

设置同第 2 个实验，比较结果如图 4 所示。从图 4 中可以看出，本文的算法在不同信噪比环境下噪声抑制效果都优于同类语音增强算法。

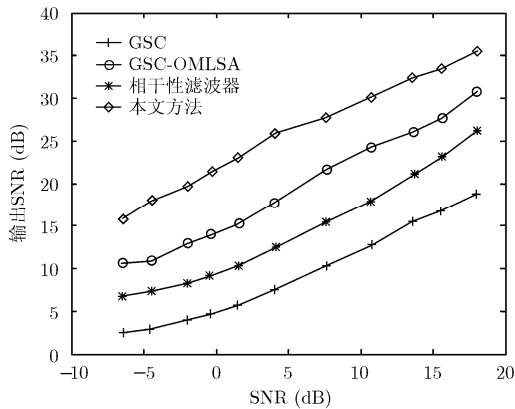


图 4 侧射方向不同信噪比处理效果比较图， $RT_{60}=0.4$ s

第 4 个实验为主观评测，其目标是通过不同人员主观测试给出降噪后输出信号的效果好坏。由于客观评价不能准确反映人们对语音的感受，现代通信语音质量评价中，由 ITU-TP.800 定义的 MOS (Mean Opinion Score) 可以作为主观评价语音质量的标准，其方法是用人耳主观感受语音质量，并用记分方式做出判定，最低分为 1，最高分为 5，且得分越高表示质量越好。测试者打分前先听原始语音信号，然后再听处理后的语音信号，对处理后的信号从失真度、可理解性、残余噪声大小、舒适度等方面进行全面判定，并给出分数。

实验环境同第 3 个实验， $RT_{60}=0.4$ s，分别在 babble、音乐、汽车、办公室、工厂 5 种噪声环境

下分别使 GSC, GSC-OMLSA, coherence-based 以及本文方法进行了降噪，初始信噪比均为 0 dB 左右，然后请 16 个与本项目不相关人员分别给出分数，最后求出每种方法处理后语音评测的平均分，其结果如表 1 所示。可以看出本文算法主观测试成绩也优于另外 3 种语音增强算法。

表1 MOS主观测试得分结果比较

噪声	得分(分)			
	GSC	GSC-OMLSA	相干性滤波器	本文方法
babble	2.6	3.7	3.4	4.1
car	2.8	4.0	3.9	4.3
factory	2.7	3.6	3.3	4.1
music	2.5	3.6	3.4	4.0
office	2.8	3.9	3.8	4.3

4.2 实际实验

实际实验是在一个大小为 4 m 宽，7 m 长和 3 m 高的房间内录音和处理；二元麦克风小阵列是由楼氏电子提供的直径为 4 mm 的两个全指向性硅微麦克风组成，阵列大约位于房子的中间位置；目标声源是一个女生在两个麦克风中心线附近的侧射方向进行诗词朗读，距离阵列约为 15 cm 左右；干扰噪声为一个音箱播放的各种类型的背景噪声，它位于麦克风 1 一侧，与目标声源的夹角约为 60° ，距离阵列约为 1 m 左右；音频采集卡使用的是福建泉州恒通数码科技的 DAR-2000 进行信号采集，采样率为 16 kHz。我们选取 5 种不同类型的噪声作为

干扰源, 处理前以及基于小阵列不同算法处理后的信噪比见表2。

表2 不同类型背景噪声环境处理信噪比比较

噪声	信噪比(dB)				
	输入信号	GSC	GSC-OMLSA	相干性滤波器	本文方法
babble	-0.56	5.92	11.95	10.01	14.31
car	3.53	14.69	17.97	16.23	21.83
factory	-7.90	-0.85	8.11	6.34	11.94
music	-3.10	3.97	10.04	7.36	13.93
office	6.86	15.10	19.03	17.21	21.80

从表2可以看出, 本文的语音增强算法对各种类型的噪声干扰, 其处理效果均优于同类型的其他算法。

5 结束语

本文以二元麦克风小阵列为例, 提出一种基于相干性滤波器和广义旁瓣抵消器组合的麦克风小阵列语音增强算法, 在语音段和非语音段采用不同的平滑系数分别进行噪声谱估计, 可以加快跟踪噪声谱的变化, 并能保证较小的目标语音畸变; 而对于采用的NLMS算法, 在语音段和非语音段采用不同的步长对滤波器权系数进行更新, 可以最大限度避免在长语音段对噪声信号跟踪失效的问题。从仿真实验可以看出, 本文提出的算法可以有效工作在端射和侧射方向, 而实际录音实验也进一步证明了本算法在真实环境中的有效性。

参 考 文 献

- [1] Brandstein M and Ward D. *Microphone Arrays: Signal Processing Techniques and Applications*[M]. Berlin: Springer Verlag, 2001, Chapter 1.
- [2] Doclo S and Moonen M. Superdirective beamforming robust against microphone mismatch[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2007, 15(2): 617-631.
- [3] Wang L, Ding H, Yin F, *et al.*. Combining superdirective beamforming and frequency-domain blind source separation for highly reverberant signals[J]. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010, 2010: 1-13.
- [4] Elko G W and Pong A N. A simple adaptive first-order differential microphone[C]. *IEEE International Conference on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, USA, Oct. 1995: 169-172.
- [5] Wolff Buck M, Haulick T, and Schmidt G T. A compact microphone array system with spatial post-filtering for automotive applications[C]. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, April. 2009: 221-224.
- [6] Chen J, Phua K, Shue L, *et al.*. Performance evaluation of adaptive dual microphone systems[J]. *Speech Communication*, 2009, 51(12): 1180-1193.
- [7] Griffiths L and Jim C W. An alternative approach to linearly constrained adaptive beamforming[J]. *IEEE Transactions on Antennas and Propagation*, 1982, 30(1): 27-34.
- [8] Chan S C and Zhou Y. On the performance analysis of a class of transform-domain NLMS algorithms with gaussian inputs and mixture gaussian additive noise environment[J]. *Journal of Signal Processing Systems*, 2011, 64(3): 429-445.
- [9] Huang H and Lee J. A new variable step-size NLMS algorithm and its performance analysis[J]. *IEEE Transactions on Signal Processing*, 2012, 60(4): 2055-2060.
- [10] 李涛, 李国林, 廖辉荣. 基于广义旁瓣相消器的滤波器组旁瓣干扰抑制[J]. *电子与信息学报*, 2011, 33(1): 85-89.
- [11] Li Tao, Li Guo-lin, and Liao Hui-rong. Side-lobe interference suppression for filter banks based on GSC[J]. *Journal of Electronics & Information Technology*, 2011, 33(1): 85-89.
- [12] Joerg Bitzery, Klaus Uwe Simmerz, and Karl Dirk Kammery. Theoretical noise reduction limits of the general-ized sidelobe canceller(GSC) for speech enhancement[C]. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, AZ, USA, March. 1999: 2965-2968.
- [13] Huang Y, Chen J, and Benesty J. Immersive audio schemes[J]. *IEEE Signal Processing Magazine*, 2011, 28(1): 20-32.
- [14] Allen J B, Berkley D A, and Blauert J. Multimicrophone signal-processing technique to remove room reverberation from speech signals[J]. *The Journal of the Acoustical Society of America*, 1977, 62(4): 912-915.
- [15] Rahmani M, Akbari A, and Ayad B. An iterative noise cross-psd estimation for two-microphone speech enhancement[J]. *Applied Acoustics*, 2009, 70(3): 514-521.
- [16] Kalle F, Ghorbel M, Frikha M, *et al.*. A noise cross PSD estimator based on improved minimum statistics method for two-microphone speech enhancement dedicated to a bilateral cochlear implant[J]. *Applied Acoustics*, 2012, 73(3): 256-264.
- [17] Ghosh P K, Tsiartas A, and Narayanan S. Robust voice activity detection using long-term signal variability[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2011, 19(3): 600-613.
- [18] Ribeiro F, Zhang Cha, Florencio D A, *et al.*. Using

- reverberation to improve range and elevation discrimination for small array sound source localization[J]. *IEEE Transactions on Speech and Audio Processing*, 2010, 18(7): 1781-1792.
- [18] Blandin C, Ozerov A, and Vincent E. Multi-source tdoa estimation in reverberant audio using angular spectra and clustering[J]. *Signal Processing*, 2012, 92(8): 1950-1962.
- [19] Abutalebi H R, Dashtbozorg B, and Zare T. Speech enhancement using hybrid generalized sidelobe canceller and spectral estimator[C]. *International Symposium on Telecommunications*, Tehran, Aug. 2008: 564-569.
- [20] Habets EAP. Room impulse response generator[OL]. http://home.tiscali.nl/ehabets/rir_generator.html, 2010, 9.
- [21] Allen J B and Berkley D A. Image method for efficiently simulating small-room acoustics[J]. *Journal of the Acoustical Society of America*, 1979, 65(4): 943-950.
- 杨立春: 男, 1975年生, 博士生, 讲师, 研究方向为麦克风阵列语音增强处理.
- 钱运涛: 男, 1968年生, 博士, 博士生导师, 教授, 主要研究方向为信号处理、模式识别、机器学习等.