doi:10.6043/j.issn.0438-0479.201612021

一种采用旁瓣增强的麦克风阵列抗混响算法

李剑汶[1],章宇栋[1],童 峰[1*],洪青阳[2]

(1.厦门大学海洋与地球学院,水声通信与海洋信息技术教育部重点实验室,福建 厦门 361102;
2.厦门大学信息科学与技术学院,福建 厦门 361005)

摘要:在音频/视频会议、人机交互、语音识别等领域,严重的混响干扰导致麦克风阵列语音处理性能急剧下降.针对现有逆滤波等抗混响方法需要获得准确的房间传输响应,而波束形成方法抗混响性能有限的问题,基于广义旁瓣抵消器(generalized sidelobe canceller,GSC)结构提出一种采用旁瓣增强的麦克风阵列抗混响算法(Sidelobeenhanceing reverberation mitigation algorithm for microphone array,SERM).该算法首先进行波束形成,初步获得增强的直达语音信号,并对旁瓣获取的混响分量进行自适应迭代增强,再将旁瓣迭代增强的误差信号作为参考噪声进行自适应噪声抵消,最终输出抗混响语音.实验结果表明,在混响环境下该方法能有效改善麦克风阵列的语音信号质量.

关键词:麦克风阵列;波束形成;广义旁瓣抵消器;抗混响处理

中图分类号:TN 912.3 **文献标志码**:A **文章编号**:0438-0479(2017)05-0711-07

近年来麦克风阵列成为智能语音技术的研究热点,被广泛应用于音频/视频会议系统[1]、车载导航[2]、智能家居[3]等领域.但是,在许多实际应用场景中,由于墙体、家具、家电反射等原因存在大量混响,直接影响到麦克风阵列的工作性能.

在目前已有的多种抗混响语音增强方法中:波束形成方法[4]通过对准声源方向进行波束形成,该算法实现简单方便,但其抗混响依赖于波束形成的指向性,而对于语音这类宽带信号而言,实际场合广泛采用的小尺寸阵列指向性往往不够理想[5],影响了抗混响处理性能;谱减法[6]可在频域上对语音信号进行抗混响处理,但在处理过程中会损失语音信号的相位特性,而相位特性对许多麦克风阵列系统而言是至关重要的;盲均衡算法[7]通过使用与输入和输出通道相同数量的有限脉冲响应(finite impulse response,FIR)滤波器来恢复信号的幅度和相位特性以比实现抗混响,然而该算法目前均基于大批量数据的处理,因此很难应用在实时工作的麦克风阵列系统中;逆向滤波法[8]是利用已知条件求得声源与麦克风之间的传输响应,然后通过设计对应的逆向滤波器达到消除混响的目的,但逆向滤波法需要提前获取房间传输响应作为先验知识[8],在实际应用中,麦克风阵列工作环境处于变化之中,很难提前获取准确的房间传输响应.

广义旁瓣抵消器(generalized sidelobe canceller,GSC)算法在麦克风阵列语音信号增强研究领域得到了广泛的研究和应用.李芳兰等[9]在 GSC 中引入可调波束形成器估计声源方位以抑制背景噪声干扰.陈磊等[10]在 GSC 基础上提出了一种可跟踪移动声源方向的麦克风阵列语音增强算法.GSC 算法能很好地抑制非相干背景噪声,但无法抑制相干噪声[11].在混响环境中,由于混响分量和直达分量来自同一声源,具有相关性,所以 GSC 算法无法有效地进行抗混响处理.

考虑到上述问题,本研究基于 GSC 结构提出了一种旁瓣增强的麦克风阵列抗混响算法(Sidelobeenhanceing reverberation mitigation algorithm for microphone array,SERM).该算法在通过声源定位[12]获取声源方向后首先利用波束形成方法初步获取语音直达分量;其次,利用阻塞矩阵输出语音混响分量通过自适应滤波算法以直达分量为预期进行迭代,和旁瓣获得的混响语音进行叠加增强;最后,将旁瓣增强

收稿日期:2016-12-15 **录用日期**:2017-05-27

基金项目:福建省高校产学合作项目(2015H6019);福建省中青年教师教育科研项目(JA14008);厦门市科技计划项目(3502Z20153002)

* **通信作者**:ftong@xmu.edu.cn

引文格式:李剑汶,章宇栋,童峰,等.一种采用旁瓣增强的麦克风阵列抗混响算法[J].厦门大学学报(自然科学版),2017,56(5):711-717.

Citation:LI J W,ZHANG Y D,TONG F,et al.A Microphone-array de-reverberation method using sidelobe enhancement[J].J Xiamen Univ Nat Sci,2017,56(5):711-717.(in Chinese)

迭代的误差信号作为参考噪声对增强语音进行自适应噪声迭代抵消处理,输出语音增强信号.该方法无需房间传输函数作为先验知识,并可改善常规波束形成方法由于指向性不足导致混响抑制性能不高的问题.实验结果表明,该方法可有效实现混响环境下的语音增强.

1 算 法

1.1 波束形成

目前常用的麦克风阵列波束形成算法有 delay-sum[12]、filter-sum 波束形成方法[12]、自适应波束形成方法[13]、后置滤波方法[14]、子空间方法[15]、子带波束形成方法[16]、近场波束形成方法[17]等.因为 delay-sum 波束形成方法因信号频率的变化会产生不同程度的旁瓣,而 filter-sum 波束形成方法通过滤波器系数调整可以形成宽带的窄波束[12],所以本研究中采用 filter-sum 波束形成方法.即将 M 元麦克风阵列每个通道信号 $x_k(n)(k=1,2,\cdots,M)$ 经过一个 FIR 滤波器,得到输出信号

$$y_k(n)=\sum_{j=-J}^{J}a_{k,j}x_k(n-j), \tag{1}$$

其中,n 表示时间坐标,J 表示 FIR 滤波器的时间跨度,$a_{k,j}$ 表示第 k 通道 FIR 滤波器第 j 个抽头系数.波束形成输出 $y(n)$ 为[12]:

$$y(n)=\sum_{k=1}^{M}y_k(n). \tag{2}$$

1.2 混 响

在典型的家庭、办公环境中,由于墙壁反射等原因,语音信号在室内多径传播产生混响效应.房间传输响应[18]可以分为 3 个部分,如图 1 所示.第一部分为直达声信号的传输响应.第二部分为直达声信号到达之后,在 50 ms 内到达的强反射信号,称为早期混响信号的传输响应,严重影响语音信号处理的性能;同时早期混响信号大小与语音识别性能具有极高的相关度[18],其特性强烈依赖于声源和麦克风的位置.第三部分为后期混响信号的传输响应,其特性为幅度大致呈指数衰减,衰减速率与声源和麦克风的位置无关.

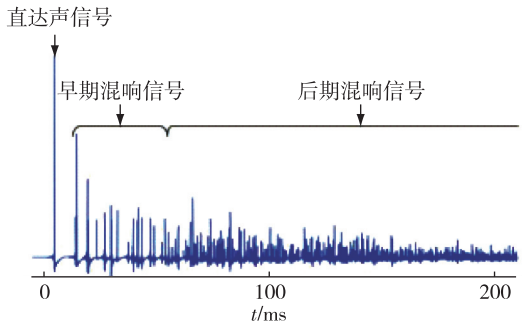
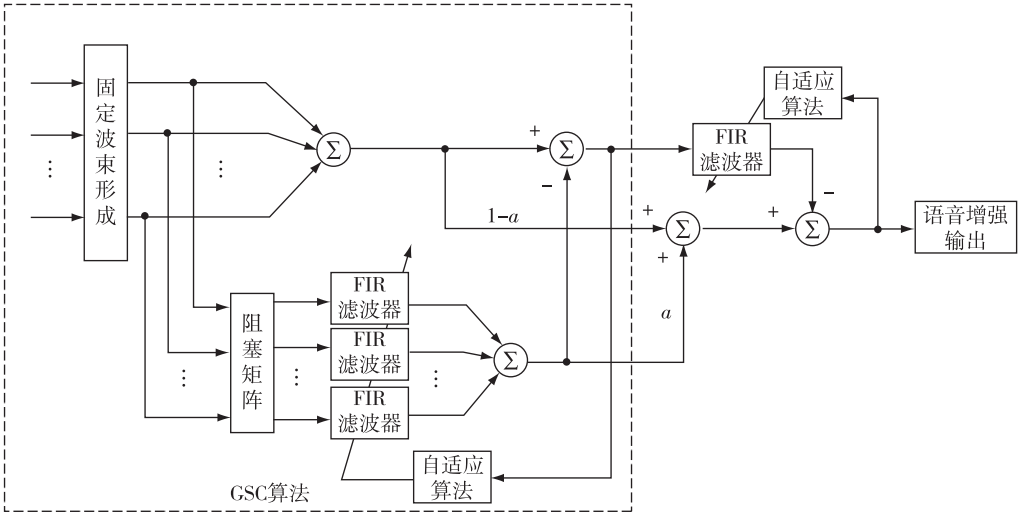


图 1 房间传输响应[18]

Fig. 1 The room impulse response[18]

1.3 SERM 算法

本研究提出的基于旁瓣增强的抗混响语音增强 SERM 算法的流程图如图 2 所示.图 2 中虚线框内为 GSC 算法:GSC 算法通过波束形成输出对准方向的带



图中 a 为叠加的权重系数.

图 2 SERM 语音增强算法结构图

Fig. 2 The structure of SERM algorithm

噪语音;再利用阻塞矩阵输出与语音信号不相关的干扰噪声;最后通过自适应迭代滤波处理后,抑制波束形成输出的信号中的干扰噪声,实现语音增强.在混响背景下,SERM 算法在 GSC 算法的基础上通过阻塞矩阵获得旁瓣对应的混响分量;并以直达语音为期望信号进行自适应迭代后将波束形成输出与迭代输出进行旁瓣增强叠加;最后再将迭代误差作为参考噪声以旁瓣增强信号为期望信号进行自适应噪声抵消,实现抗混响语音增强.

由式(1)可知,M 元麦克风阵列各通道信号经波束形成后输出为 $y_k(n)(k=1,2,\cdots,M)$.假设 $\boldsymbol{Y}$ 为阻塞矩阵输入,$\boldsymbol{B}$ 为阻塞矩阵[19],R 为阻塞矩阵输出信号混响分量,则

$$\boldsymbol{Y}=[y_1(n),y_2(n),\cdots,y_M(n)]^{\mathrm{T}}, \tag{3}$$

$$\boldsymbol{B}=\begin{bmatrix}1&-1&0&\cdots&0&0\\0&1&-1&\cdots&0&0\\0&0&1&\cdots&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&\cdots&1&-1\end{bmatrix}_{(M-1)\times M}, \tag{4}$$

$$\boldsymbol{R}=\boldsymbol{BY}=\begin{bmatrix}y_1(n)-y_2(n)\\y_2(n)-y_3(n)\\\vdots\\y_{M-1}(n)-y_M(n)\end{bmatrix}=\begin{bmatrix}r_1(n)\\r_2(n)\\\vdots\\r_{M-1}(n)\end{bmatrix}. \tag{5}$$

将阻塞矩阵输出的旁瓣分量 $r_k(n)(k=1,2,\cdots,M-1)$作为每个 N 阶自适应滤波器的输入信号,波束形成输出 $y(n)$作为期望,$w_{k,n}(i)(i=1,2,\cdots,N_1)$为 n 时刻第 k 通道自适应滤波器的权系数,则第 k 通道自适应滤波器输出 $z_k(n)$为:

$$z_k(n)=\sum_{i=1}^{N_1}w_{k,n}(i)r_k(n-i), \tag{6}$$

$$d(n)=\sum_{k=1}^{M-1}z_k(n), \tag{7}$$

迭代误差为:

$$e(n)=y(n)-d(n), \tag{8}$$

采用最小均方误差(LMS)自适应算法[20],则第 k 个滤波器权重系数更新为:

$$w_{k,n}(i)=w_{k,n-1}(i)+\mu_1\cdot r_k(n)\cdot e(n), \tag{9}$$

其中,μ_1 为旁瓣迭代滤波器的迭代步长.对混响信号以直达信号为期望进行自适应迭代后,将旁瓣迭代滤波输出 $d(n)$与波束形成输出 $y(n)$进行加权叠加得到旁瓣增强输出:

$$s(n)=a\cdot d(n)+(1-a)\cdot y(n), \tag{10}$$

其中 $s(n)$为旁瓣增强后的语音信号抗混响输出.进一步,将旁瓣迭代滤波的迭代误差 $e(n)$作为参考噪声输入,旁瓣增强抗混响输出 $s(n)$为期望信号进行 N_2 阶自适应噪声抵消迭代,获得进一步降噪后的最终语音增强输出 $s_{\mathrm{out}}(n)$:

$$s_{\mathrm{out}}(n)=\sum_{i=1}^{N_2}v_{\mathrm{n}}(i)e(n-i), \tag{11}$$

$$e_2(n)=s(n)-s_{\mathrm{out}}(n), \tag{12}$$

$$v_n(i)=v_{n-1}(i)+\mu_2\cdot e(n)\cdot e_2(n), \tag{13}$$

其中 $v_n(i)(i=1,2,\cdots,N_2)$为 n 时刻噪声抵消迭代自适应滤波器权系数,μ_2 为噪声抵消迭代自适应滤波器的迭代步长.

本研究所提 SERM 算法的自适应迭代采用最小均方差(LMS)自适应算法,在两部分的滤波器阶数均设置为 N 阶时,对 M 元麦克风阵列 LMS 算法每次迭代中需要进行 $M(2N+1)$次乘加法运算[21],运算复杂度不高,可以实现算法的硬件移植.

2　实　验

2.1　实验设置

为了验证 SEMR 算法的有效性,在混响较为严重的室内环境中进行了抗混响语音增强实验.实验房间为典型的办公室,大小为 9.6 m×7.8 m×3 m,墙壁为普通水泥抹灰墙面.房间混响时间 T_{60} 用赛宾公式[22]近似估算为 3.5 s.实验采用 7 元麦克风均匀圆形阵列,阵列直径为 15 cm.实验语音信号由 0.5～4 kHz 的扫频信号和 TIMIT 标准语音信号[23]组合,其中扫频信号用于测量实验环境的房间传输响应,TIMIT 标准语音信号用于评估麦克风阵列输出的语音信号质量,其原始语音信噪比为 28.06 dB.播放测试语音的采样率为 16 kHz,使用 DELL AD211 蓝牙扬声器进行播放.

实验中首先通过声源定位算法来获取目标声源的方向信息,以保证算法的性能不受麦克风阵列和目标声源相对位置变化的影响.同时本研究 SEMR 算法的波束形成方法采用 filter-sum 方法[12],并比较了其与单麦克风采集语音、filter-sum 波束形成算法[12](下文简称波束形成算法)的抗混响性能,其中单麦克风采用的是本文中 7 元阵列中位于 0°方向的麦克风.

实验中用到的各算法参数如表 1 所示:波束形成的滤波器抽头系数由声源相对位置来设定,同时为保证语音信号处理实时输出,波束形成器阶数设置为 120 阶,两部分自适应滤波器阶数 N_1,N_2 均设置为 32 阶.步长因子 μ_1,μ_2 选取 LMS 算法的收敛速度和

表 1 实验参数设置

Tab. 1 The parameters of experiment

信号采样率/kHz	固定波束形成器阶数	N_1,N_2	μ_1	μ_2	0°方向声源相对位置/m	a
16	120	32	0.03	0.01	5	0.5

控制稳态性最优时的值,具体分析见下文.

2.2 实验结果与讨论

将单麦克风和不同抗混响方法处理获得的扫频信号分别进行相关分析,可以得到对应的房间传输响应如图 3 所示.从图 3(a)单麦克风采集语音可以看出,直达声信号在图中约 18.75 ms 处,在 20～40 ms 区域以及 47 ms 处,存在几处较为明显的早期混响成分;从图 3(b)和(c)可以看出,波束形成算法和 SEMR 算法处理后的语音信号对几处明显的混响成分均有较为明显的抑制效果.实验中进一步进行抗混响效果的量化比较、分析.

直达混响比[24] R_{DR}是一个衡量房间混响程度的声学参数,定义为直达声信号与混响的能量比值,其计算公式为:

$$R_{\mathrm{DR}}=10\lg\left(\frac{h^2(t)}{\sum\limits_{n=0}^{t/\Delta t-1}h^2(n)+\sum\limits_{n=t/\Delta t+1}^{L/\Delta t-1}h^2(n)}\right), \quad (15)$$

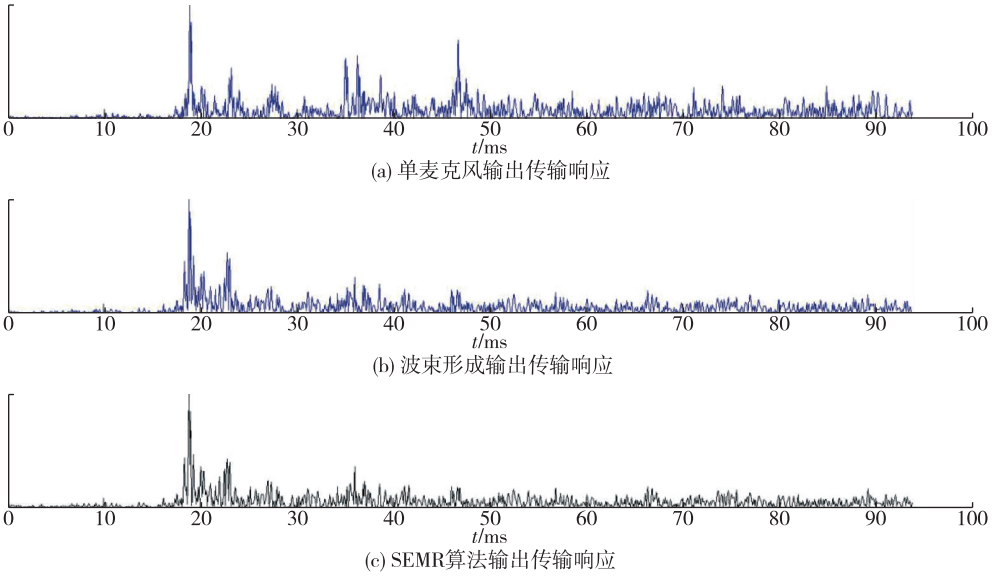
其中:$h(n)$为语音信号在 n 时刻的房间传输响应值;t 为直达声信号对应的时刻;Δt 为采样时间间隔;L 为计算时间长度,取值为 50 ms,即为早期混响时间段.实验计算 3 种输出信号在早期混响时间长度下的 R_{DR} 如表 2 所示.从表中可以看出,相对于单麦克风输出语音信号的 R_{DR},波束形成算法输出语音信号的 R_{DR} 提高了 1.51 dB,SEMR 算法输出语音信号的 R_{DR} 提高了 2.15 dB.作为参考,表 2 同时列出了测试语音的信噪比,相对于单麦克风输出语音信号的信噪比,波束形成算法输出语音信号的信噪比提高了 0.14 dB,SEMR 算法输出语音信号的信噪比提高了 0.64 dB.可见,SEMR 算法的抗混响性能优于波束形成方法.

表 2 实验语音信号输出 R_{DR} 和信噪比

Tab. 2 The R_{DR} and SNR of the testing speech signal

算法类型	R_{DR}/dB	信噪比/dB
单麦克风	−11.33	18.75
波束形成	−9.82	18.89
SEMR	−9.18	19.39

图 4 为语音信号的波形频谱图.可以看出,与原始语音相比,噪声和混响干扰造成单麦克风、波束形成及 SEMR 算法语音信号不同程度的质量下降;同时,从图 4(a)、(c)、(e)、(g)方框内波形可以看出,SEMR 算法输出语音时域波形得到了较好的增强.为了更好地进行性能比较,进一步对实验语音进行语音音质量



(a) 单麦克风输出传输响应

(b) 波束形成输出传输响应

(c) SEMR算法输出传输响应

图 3 房间归一化传输响应

Fig. 3 The normalized transmission response of the room

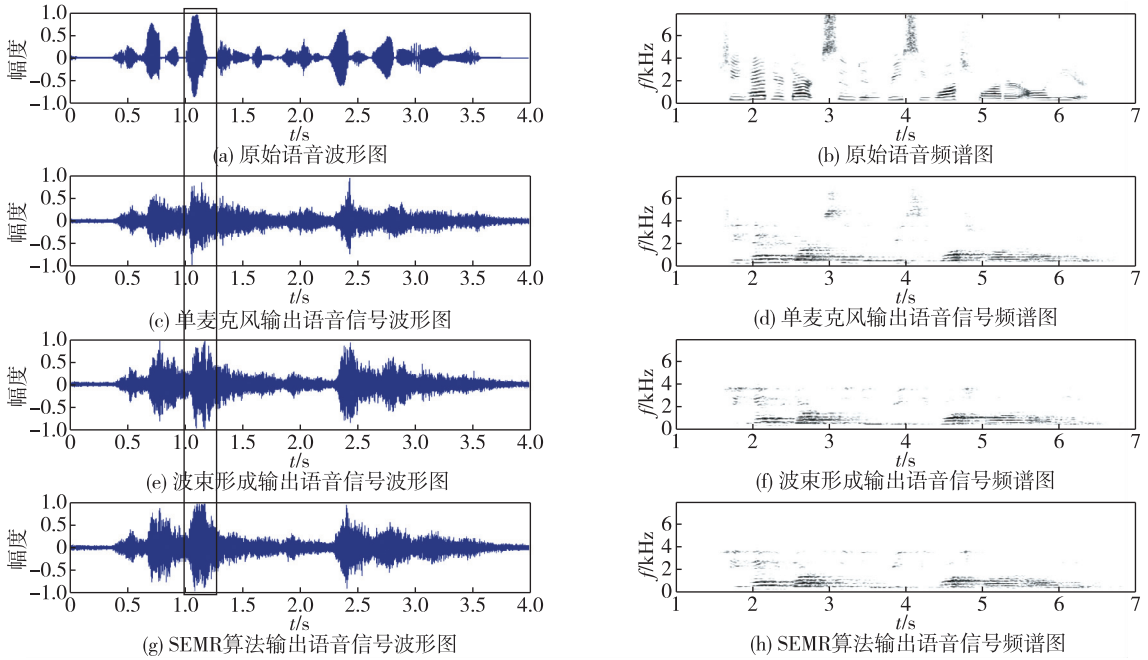


图 4　语音信号波形图和频谱图

Fig. 4　The waveform and spectrogram of the speech signal

化评估.

感知语音质量评价(PESQ)是一种采用改进型认知模型技术和听觉模型技术的语音质量客观评价算法[25].其满分为 4.50 分,得分越高则说明信号质量越好.本文中实验 PESQ 评估测试的语音信号为国际 TIMIT 标准语音库的语音片段.3 种算法输出语音的 PESQ 量化评估结果如表 3 所示.从表 3 可以看出,波束形成输出的语音信号质量与 SEMR 算法输出的语音质量均优于单麦克风语音,但 SEMR 算法输出语音的 PESQ 得分值更高.可见,SEMR 算法输出的语音质量优于波束形成方法输出.

表 3　实验语音信号输出 PESQ 评估结果

Tab. 3　The PESQ evaluation results of the testing speech signal

算法类型	PESQ 值
单麦克风	1.24
波束形成	1.33
SEMR	1.64

语音识别是目前麦克风阵列的热门研究方向之一,语音识别效果是衡量麦克风阵列性能的一个重要标准,也是本研究算法主要的实际应用背景.本文中对波束形成输出语音和本研究算法输出语音输出进行了初步识别测试.实验采用 HTK 语音识别工具箱[26]作为识别端,将麦克风阵列输出语音通过声卡以 16 kHz 采样率输入 PC,由 PC 上运行的 HTK 识别端进行语音识别.识别实验分别由测试者 A 和 B 进行,测试者均位于麦克风阵列 0°方向,3 m 距离处,测试指令词共 40 个,实验测试结果如表 4 所示.从表 4 可以看出,测试者 A 和 B 基于 SEMR 算法输出识别率相对于固定波束形成语音输出识别率分别提高了 10 和 5 个百分点.

表 4　麦克风阵列识别性能测试结果

Tab. 4　The results of the recognition performance of the microphone array　%

测试者	波束形成方法识别率	SEMR 算法识别率
A	82.5	92.5
B	85	90

在 SEMR 算法中,步长因子 μ_1,μ_2 是控制 LMS 自适应算法的收敛性能和稳态失调的重要参数[27],步长因子需适当设置以保证算法性能的充分发挥.实验中测试了 SEMR 算法中不同步长因子 μ_1,μ_2 的语音信号输出 R_{DR} 如图 5 所示,其中曲线 1 表示当步长因子 μ_1 为 0.03 时不同步长因子 μ_2 的语音信号输出 R_{DR},曲线 2 表示当步长因子 μ_2 为 0.01 时不同步长

因子 μ_1 的语音信号输出 R_{DR}.从图 5 可以看出,在一定范围内 SEMR 算法性能对于步长因子 μ_1 具有较好的稳健性,而相对步长因子 μ_2 则较为敏感;同时从图 5 可以看出,SEMR 算法中步长因子 $\mu_1=0.03$,$\mu_2=0.01$ 时,算法输出的语音信号 R_{DR} 较高,抗混响性能较好.

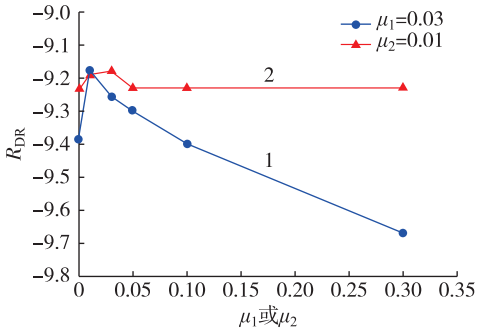


图 5 不同迭代步长的语音信号输出的 R_{DR}

Fig. 5 The R_{DR} values of the output signal for different iterations

3 结 论

针对严重混响环境情况下麦克风阵列处理性能降低的问题,本研究在 GSC 结构基础上提出了一种利用旁瓣获得的混响语音与波束形成获得的直达语音进行自适应迭代增强,并结合迭代噪声进行自适应噪声抵消的抗混响方法 SEMR.实验结果表明,本研究提出的 SEMR 算法相对于传统波束形成算法,在抗混响能力、输出语音信号质量和识别性能上都有一定程度的提高.

参考文献:

[1] NISHIURA T, GRUHN R, NAKAMURA S. Collaborative steering of microphone array and video camera toward multi-lingual tele-conference through speech-to-speech translation[C]// IEEE Workshop on Automatic Speech Recognition and Understanding. Trento:IEEE,2001:119-122.

[2] MEYER J,SIMMER K U.Multi-channel speech enhancement in a car environment using wiener filtering and spectral subtraction[J].IEEE Transactions on Acoustics, Speech,and Signal Processing,1997,21(2):1167-1170.

[3] 陈成斌.针对于家居环境的语音增强系统的研究与开发[D].广州:华南理工大学,2014.

[4] 王冬霞,殷福亮.基于近场波束形成的麦克风阵列语音增强方法[J].电子与信息学报,2007,1(29):67-70.

[5] 何赛娟,陈华伟,尹明婕,等.基于差分麦克风阵列和语音稀疏性的多源方位估计方法[J].数据采集与处理,2015,2(30):372-382.

[6] 陈欢,邱晓辉.改进谱减法语音增强算法的研究[J].计算机技术与发展,2014,4(24):69-71.

[7] YOSHIOKA T,NAKATANI T.Generalization of multi-channel linear prediction methods for blind MIMO impulse response shortening[J]. IEEE Trans. Audio, Speech,Language Precess,2012,10(20):2707-2720.

[8] WANG H,ITAKURA F.An approach of dereverberation using multi-microphone subband envelope estimation[C]// IEEE International Conference on Acoustics, Speech and Signal Processing.Toronto:IEEE,1991:953-956.

[9] 李芳兰,周跃海,童峰,等.采用可调波束形成器的 GSC 麦克风阵列语音增强方法[J].厦门大学学报(自然科学版),2013,52(2):186-189.

[10] 陈磊,江伟华,童峰.一种可跟踪移动声源方向的麦克风阵列语音增强算法[J].厦门大学学报(自然科学版),2015,54(4):551-555.

[11] REUVEN G, GANNOT S, COHEN I. Dual-source transfer-function generalized sidelobe canceller[J].IEEE Transactions on Audio, Speech, and Language Processing,2008,4(16):711-727.

[12] ZITOUNI M S,HAMMADIH M L,ALSHEHHI A,et al.Simulated analysis and hardware implementation of voiceband circular microphone array[C]// IEEE 20th International Conference on Electronics,Circuits and Systems(ICECS).Abu Dhabi:IEEE,2013:508-511.

[13] FROST O L.An algorithm for linearly constrained adaptive array processing[J].Proceedings of the IEEE,1972,60(8):926-935.

[14] MVCOWAN A I, BOURLARD H. Microphone array post-filter for diffuse noise field[C]// IEEE International Conference on Acoustics,Speech and Signal Processing. Florida:IEEE,2002,1:905-908.

[15] ASANO F,HAYAMIZU S,YAMADA T,et al.Speech enhancement based on the subspace method[J]. IEEE Transactions on Speech and Audio Processing,2000,8(5):497-507.

[16] GRBIC N,NORDHOLM S,CANTONI A.Optimal FIR subband beamforming for speech enhancement in multipath environments[J].IEEE Signal Processing Letters,2003,10(11):335-338.

[17] ZHENG Y R,GOUBRAN R A,EI-TANANY M,et al. A microphone array system for multimedia applications with near-field signal targets[J].IEEE Sensors Journal,2005,5(6):1395-1406.

[18] YOSHIOKA T,SEHR A,DELCROIX M,et al.Making machines understand us in reverberant rooms:robustness against reverberation for automatic speech recognition[J].IEEE Signal Processing Magazine,2012,6(29):114-126.

[19] 罗丁利,徐伟.一种阻塞矩阵的构建方法[J].火控雷达技术,2009,4(38):54-56.

[20] 曹亚丽.自适应滤波器中 LMS 算法的应用[J].仪器仪表学报,2005,26(s2):452-454.

[21] 马伟富.自适应 LMS 算法的研究及应用[D].成都:四川大学,2005:4-7.

[22] 张武威.关于室内混响时间的计算问题[J].电声技术,2005,29(3):17-20.

[23] GAROFOLO J,LAMEL L,FISHER W,et al.TIMIT acoustic-phonetic continuous speech (ms-wav version)[J].Journal of the Acoustical Society of America,1993,88(88):210-221.

[24] THOMAS M R P,TASHEV I J,LIM F,et al.Optimal beamforming as a time domain equalization problem with application to room acoustics[C]// IEEE 14th International Workshop on Acoustic Signal Enhancement (IWAENC).Antibes:IEEE,2014:75-79.

[25] ITU.Perceptual evaluation of speech quality (PESQ):An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codes:ITU-T Recommendation P.862[S].Geneva:ITU 2001:1-21.

[26] YOUNG S, EVERMANN G, GALES M, et al. The HTK book[M]Cambridge:Cambridge University Engineering Department,2006.

[27] 曾伟,吴国辉,代冀阳,等.LMS 自适应滤波算法的改进及性能分析[J].计算机仿真,2011,28(4):111-114.

A Microphone-array De-reverberation Method Using Sidelobe Enhancement

LI Jianwen[1],ZHANG Yudong[1],TONG Feng[1*],HONG Qingyang[2]

(1.Key Laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, College of Ocean and Earth Sciences,Xiamen University,Xiamen 361102,China;2.School of Information Science and Engineering,Xiamen University,Xiamen 361005,China)

Abstract:In many practical applications such as audio/video conferences,man-machine interfaces,speech recognitions,the reverberation leads to significant degradation of microphone array speech processing.However,while the conventional de-reverberation methods,such as inverse-filtering method,need to obtain accurate room transmission responses,the classic beam-forming method only yields limited performance.Based on the Generalized Sidelobe Canceller (GSC) structure,this paper proposes a novel microphone-array de-reverberation method to achieve sidelobe exploitation.First,we use the contemporary beam-forming method to initially obtain the direct speech signals.Then we adopt the adaptive iterative algorithm to explore reverberation components in sidelobes for further enhancements.Finally,the error of the sidelobe enhancement is adopted as the reference noise of adaptive noise cancellation to achieve the de-reverberation.Experimental results show that the proposed method effectively improves the quality of speech signal under the reverberation environment.

Key words:microphone-array;beam-forming;generalized sidelobe canceller (GSC);de-reverberation method