

DOI: 10.13382/j.jemi.2017.10.018

LCMV 分频的改进维纳滤波后置波束形成算法*

郭业才^{1,2} 陈小燕¹ 王超¹

(1. 南京信息工程大学 江苏省气象探测与信息处理重点实验室 南京 210044;
2. 江苏省大气环境与装备技术协同创新中心 南京 210044)

摘要:针对封闭环境中语音信号受到混响影响,提出了 LCMV 分频的改进维纳滤波后置波束形成算法。该算法通过计算麦克风阵列接收到含混响信号的短时傅里叶变换得到频域阵列信号,对频域阵列信号分频处理,将分频的信号进行线性约束最小方差波束形成滤波处理,该波束滤波根据每个频段上混响时间不同的特性对频域阵列信号进行分频处理后,将波束形成算法分别应用到高低频中,以提高混响抑制的精度;再由频域阵列信号的组合功率谱进行维纳后置滤波以抑制混响,由麦克风阵列接收到混响信号的直达波和反射波之间不相关性及麦克风阵列接收信号的空间信息解决维纳滤波器的精确估计问题;最后由逆短时傅里叶变换恢复出时域信号。仿真结果表明,该算法对混响抑制具有明显的改善;且在混响时间 600 ms 条件下语音增强系统的 PESQ 值提高了 0.26。

关键词: 麦克风阵列;线性约束最小方差;混响;维纳后置波束形成;房间冲击响应

中图分类号: TB52⁺.6 **文献标识码:** A **国家标准学科分类代码:** 510.4040

Improved Wiener post-filter beamforming algorithm based on LCMV divided-frequency

Guo Yecai^{1,2} Chen Xiaoyan¹ Wang Chao¹

(1. Jiangsu Key Laboratory of Meteorological Observation and Information Processing, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing 210044, China)

Abstract: In order to solve the problem that the speech signal is affected by reverberation in the closed environment, the improved Wiener filter based on LCMV is proposed. In this proposed algorithm, the frequency domain array signals can be obtained by the Fourier transform of the signals with the reverberation received by the microphone array and are processed according to frequency division method. The divided-frequency signals also are processed by using linear constrained minimum variance beamforming filter method and according to different frequency reverberation time of each beam, and the beamforming algorithm is applied to the high and low frequency domain to improve the accuracy of reverberation suppression for the frequency domain array signals. The combination power spectrum estimation of the frequency domain array signal is used to treat Wiener post-filter in order to suppress the reverberation. Since the received-reverberation signals of microphone array between the direct wave and reflected wave are not related, so we use space information of the received-reverberation signals of microphone array to solve accurately the estimation problem of the Wiener filter. Finally, the time domain signal is obtained by the inverse Fourier transform. The test results show that the proposed algorithm has a significant improvement on the reverberation suppression, and the PESQ score of the speech enhancement system is improved by 0.26 under the condition of reverberation time 600 ms.

Keywords: microphone array; linearly constrained minimum variance; reverberation; Wiener post-filter beamforming; room impulse response

收稿日期:2017-04 Received Date: 2017-04

* 基金项目:国家自然科学基金(61673222)、江苏省高校自然科学研究重大项目(13KJA510001)、江苏省高校品牌专业建设项目(PPZY2015 B134)资助

0 引言

在封闭环境会议室的免提通信场景下,麦克风阵列采集到不仅有直达路径传播的信号,而且还包含了由于会议室边界以及会议室内物体的反射而产生的延迟衰减信号,这种传播效应称为混响。混响环境下的语音会存在时间和频率模糊,从而带来语音信号时域包络和声纹结构部分失真。因此,在会议室内存在较大混响的情况下,语音将变得难以理解,尤其是听力受损者和老年人、以及机器自动语音识别系统^[1-3]。

迄今为止,一些科研学者相继提出了许多根据混响特征来对含混响的语音信号进行混响抑制的算法。混响抑制的目标是滤除封闭室内含混响效应的语音信号中房间冲击响应的影响。混响抑制算法通常分为逆向滤波器^[2,4]、复倒谱处理技术^[4]、线性预测编码方法^[5-6]、利用统计模型^[7]等类型。逆向滤波处理技术的原理是得知声源与麦克风之间的房间冲击响应或者是已经被估计得到,采用该房间冲击响应对通道传输函数进行逆向滤波来消除混响。该技术在实际环境中,需要保证房间冲击响应^[8,9]是稳定和因果的。此外,传输函数取逆的过程对通道冲击响应的估计误差非常敏感,尤其是在通道传输函数幅度较小的频率部分。复倒谱技术的原理是对接收的混响信号采用对数函数转换方法,根据房间冲击响应和纯净语音信号在时域卷积数学关系,得到复倒谱中的加法处理降低计算难度。但是复倒谱技术处理完后同样要进行传输函数取逆过程运算,考虑到室内声学冲击响应不是最小相位系统的缺点使得通道语音混响抑制方法并不实用。线性预测编码方法对误差的敏感度较低,并且不需要房间冲击响应的先验信息,但是该方法具有较高的复杂度,而且运算时会产生较大的延时,实时性较差。利用统计模型需要估计出各个房间混响特性的参数,进而计算出接收信号中的混响成分。

以上混响抑制的方法均不具有指向性,为了解决上述问题,本文提出了混响环境下 LCMV 分频的改进维纳滤波后置波束形成算法。该混响抑制算法首先根据麦克风阵列接收到信号的直达波和反射波之间是不相关,利用阵列接收信号的空间信息解决维纳滤波器的估计问题;其次根据每个频段上混响时间不同对麦克风阵列接收的信号进行分频处理后,将波束形成算法分别应用到高低频中,以提高混响抑制的精度;仿真结果表明,该算法对混响抑制具有明显的改善,且语音增强效果明显提高。

1 阵列数学模型

在封闭环境内,麦克风阵列采集到的语音信号不仅

仅包含直达路径传播的信号,且包含了由于房间反射而产生的延迟衰减信号,这种多径传播效应在接收信号中引入导致谱失真,称为混响。本文考虑 M 个相同的全向性麦克风组成均匀线阵,有 N 个语音信号 ($M > N$),第 m 个麦克风接收到的信号可以表示为:

$$x_m(k) = \sum_{n=1}^M \sum_{l=1}^L H_{nm,l}(k) s_n(k) \quad (1)$$

式中: $H_{nm,l}(k)$ 是第 n 个语音到第 m 个麦克风,长度为 l ,时间序列为 k 的房间冲击响应^[10-15],且 $m = 1, \dots, M$; $n = 1, \dots, N$; $l = 1, \dots, L$ 。根据语音信号的动态非平稳特性,因此对式(1)作短时傅里叶变换,得第 m 个麦克风接收到的频域信号为:

$$x_m(\omega, k) = \sum_{n=1}^M \sum_{l=1}^L H_{nm,l} s_n(\omega, k) \quad (2)$$

式中: $s_n(\omega, k)$ 表示 $s_n(k)$ 第 k 帧信号短时谱。用矩阵表示为

$$\mathbf{x}(\omega) = \mathbf{H}\mathbf{s}(\omega) \quad (3)$$

$$\mathbf{H} = \begin{bmatrix} h_{1n,1} & h_{1n,2} & \cdots & h_{1n,L} \\ h_{2n,1} & h_{2n,2} & \cdots & h_{2n,L} \\ \vdots & \vdots & \cdots & \vdots \\ h_{Mn,1} & h_{Mn,2} & h_{Mn,L} & h_{Mn,L} \end{bmatrix} \quad (4)$$

$$\mathbf{x}(\omega) = [x_1(\omega), x_2(\omega), \dots, x_M(\omega)] \quad (5)$$

$$\mathbf{s}(\omega) = [s_1(\omega), s_2(\omega), \dots, s_N(\omega)] \quad (6)$$

根据第 m 个麦克风所接收的语音信号 $\mathbf{x}_m(\omega)$,麦克风阵列波束形成器输出信号为:

$$\mathbf{Y}(\omega) = \sum_{m=1}^M \mathbf{w}_m(\omega) \mathbf{x}_m(\omega) \quad (7)$$

式中:

$$\mathbf{w}_{m,lcmv}(\omega) = \mathbf{w}_m^{*T}(\omega) \mathbf{d}(\omega) \quad (8)$$

$$\mathbf{d}(\omega) = [e^{-j\pi\omega(-\tau_1)/f}, e^{-j\pi\omega\tau_1(1-\tau_1)/f}, \dots, e^{-j\pi\omega(L-1-\tau_m)/f}, \dots, e^{-j\pi\omega(L-1-\tau_m)/f}]^T \quad (9)$$

式中: $(\cdot)^{*T}$ 为共轭转置, $\mathbf{d}(\omega)$ 是阵列的导向矢量, f 是采样频率, τ_m 是第 m 个麦克风通道和第 1 个麦克风通道的时间延迟。

2 混响环境下麦克风阵列波束形成算法

混响是噪声的一种,称为卷积噪声,且是噪声污染中导致语音质量恶化的重要成分。因此混响环境下的语音会存在时间和频率模糊,从而带来语音信号在时域包络和声纹结构部分失真。

2.1 LCMV 维纳滤波后置波束形成算法

麦克风阵列接收了含混响语音信号的功率谱为:

$$\Phi_{xx}(\omega) = \mathbf{H}\Phi_{ss}(\omega)(\mathbf{H})^{*T} \quad (10)$$

式中: $\Phi_{ss}(\omega)$ 为目标语音信号的功率谱。根据 LCMV 准则,得最小输出功率为:

$$E\{|Y(\omega)|^2\} = \mathbf{w}_{H,lcmv}^{*T}(\omega)\Phi_{xx}(\omega)\mathbf{w}_{H,lcmv}(\omega) \quad (11)$$

式中： $\Phi_{xx}(\omega)$ 表示为阵列接收信号的功率谱密度，且 $\Phi_{xx}(\omega) = E[\mathbf{x}(\omega)\mathbf{x}^{*T}(\omega)]$ ， $\mathbf{w}_{H,lcmv}(\omega)$ 表示为混响环境下 LCMV 波束形成混响抑制算法权矢量。

麦克风阵列信号经过时延对齐后，为了满足能够在室内达到干扰抑制和降噪的理想性能，在混响环境下的 LCMV 波束形成器设计的问题被建立为：

$$\begin{aligned} \min_{\mathbf{w}(\omega) \in C^M} & \mathbf{w}_{H,lcmv}^{*T}(\omega)\Phi_{xx}(\omega)\mathbf{w}_{H,lcmv}(\omega) \\ \text{s. t} & \mathbf{w}_{H,lcmv}^{*T}(\omega)\mathbf{H} = \mathbf{U}(\omega) \end{aligned} \quad (12)$$

式中： \mathbf{H} 为房间冲击响应矩阵； $\mathbf{U}(\omega) = [U_D(\omega) \ 0]$ 是约束条件响应矢量，用于对目标语音信号去混响。本文采用拉格朗日法求解权矢量，则目标函数 $L(\mathbf{w})$ 为：

$$\begin{aligned} L(\mathbf{w}) &= \mathbf{w}_{H,lcmv}^{*T}(\omega)\Phi_{xx}(\omega)\mathbf{w}_{H,lcmv}(\omega) + \\ \lambda &(\mathbf{w}_{H,lcmv}^{*T}(\omega)\mathbf{H} - \mathbf{U}(\omega)) \end{aligned} \quad (13)$$

式中： λ 为多维拉格朗日矢量。目标函数求对 $\mathbf{w}_{H,lcmv}(\omega)$ 的梯度并令之为 0，即 $\frac{\partial L(\mathbf{w})}{\partial \mathbf{w}} = 0$ ，得：

$$\mathbf{w}_{H,lcmv}(\omega) = (\Phi_{xx}(\omega))^{-1}\mathbf{H}(\mathbf{H}^{*T}(\Phi_{xx}(\omega))^{-1}\mathbf{H})\mathbf{U} \quad (14)$$

假设麦克风阵列中各个麦克风采集的噪声信号之间不相关，则单通道的功率谱为：

$$\Phi_{x_{x_j}}(\omega) = \Phi_{ss}(\omega) \quad (15)$$

式中： $\Phi_{x_{x_j}}(\omega)$ 是信号 x_i 和 x_j 互功率谱。那么通过计算所有可能麦克风组合的互功率谱，维纳滤波器估计为：

$$\mathbf{W}(\omega) = \frac{E\{R\{\sum_{i=0}^{M-2}\sum_{j=i+1}^{M-1}\Phi_{x_{x_j}}(\omega)\}\}}{E\{\sum_{i=0}^{M-1}\Phi_{x_{x_i}}(\omega)\}} \quad (16)$$

式中， $R\{\cdot\}$ 表示实部， $E[\cdot]$ 表示均值。通过计算混响环境下的 LCMV 维纳后置滤波的波束形成算法 (LCMV with Wiener filter beamforming algorithm in reverberation environment, LCMV-WFR) 最优权矢量为：

$$\mathbf{W}_{m,LCMV-WFRopt}(\omega) = \mathbf{w}_{H,lcmv}(\omega)\mathbf{W}(\omega) \quad (17)$$

然而，在现实声场中麦克风阵列的各个通道含混响语音并非不相关，在低频部分其相干系数的接近于 1，直到高频部分才出现较小的值，也就是说基于式 (15) 的假设条件不能完全满足。

2.2 LCMV 的改进维纳后置滤波波束形成算法

基于以上对后置维纳滤波算法的分析，该部分对维纳功率谱进行改进，改进的算法以期弥补前面提到的不足。从房间脉冲响应 $h(k)$ 的角度看，可以把 $h(k)$ 分解成两个部分，第 1 个部分时间很短，可以将其近似看成直达部分语音信号的脉冲响应函数 $h_D(k)$ ，第 2 部分形成混响语音函数 $h_R(k)$ ，表示为：

$$h(k) = \begin{cases} b(k)e^{-\Delta k}, & k > 0 \\ 0, & k < 0 \end{cases}$$

$$h_D(k) = \begin{cases} h(k), & k > k_D \\ 0, & \text{其他} \end{cases} \quad (18)$$

$$h_R(k) = \begin{cases} h(k), & k > k_R \\ 0, & \text{其他} \end{cases} \quad (19)$$

式中： $b(k)$ 是零均值的高斯白噪声； Δ 是与混响时间联系在一起衰减因子， $\Delta = 3\ln 10/T_{60}$ ；假设 $s_D(k)$ 与 $s_R(k)$ 分别表示纯净语音信号 $s(k)$ 与 $h_D(k)$ 和 $h_R(k)$ 的卷积，则 $s_D(k)$ 可以看成麦克风阵列接收信号 $x_m(k)$ 的直达信号部分， $s_R(k)$ 可以看成麦克风阵列接收信号 $x_m(k)$ 的混响较强的部分，满足下式：

$$E[|x_i(k)x_j(k)|] = E[|s_D(k)s_D(k)|] + E[|s_R(k)s_R(k)|] \quad (20)$$

$$E[|s_R(k)s_R(k)|] = e^{-2\Delta k}E[|x(k)x(k)|] \quad (21)$$

根据式 (20)、(21) 可以看出，混响是一种声音延迟衰减的叠加效应。基于以上分析，将 T 的取值保持在 T_{60} 的部分，得出直达信号和混响信号的短时能量谱密度之间的关系为：

$$\Phi_{xx}(\omega, k) = \Phi_{DD}(\omega, k) + \Phi_{RR}(\omega, k) \quad (22)$$

$$\Phi_{RR}(\omega, k) = e^{-2\Delta k}\Phi_{xx}(\omega, k) \quad (23)$$

在式 (22)、(23) 的基础上可以求出增益，进行混响抑制处理。改进的最佳滤波器系数为：

$$\mathbf{W}_{DRR}(\omega) = \frac{E\{R\{\sum_{i=0}^{M-2}\sum_{j=i+1}^{M-1}\Phi_{R_{R_i}}(\omega)\}\}}{E\{\sum_{i=0}^{M-1}\Phi_{x_{x_i}}(\omega)\}} \quad (24)$$

那么混响环境下的 LCMV 改进的维纳后置滤波的波束形成算法 (LCMV with improved wiener filter beamforming algorithm in reverberation environment, LCMV-IWFR) 最优权矢量为：

$$\mathbf{W}_{m,LCMV-IWFRopt}(\omega) = \mathbf{w}_{H,lcmv}(\omega)\mathbf{W}_{DRR}(\omega) \quad (25)$$

$$\mathbf{w}_{H,lcmv}(\omega) = (\Phi_{xx})^{-1}\mathbf{H}(\mathbf{H}^{*T}(\Phi_{xx})^{-1}\mathbf{H})\mathbf{U} \quad (26)$$

这样处理有一定效果，但衰减系数与混响时间存在联系，而不同频段的混响时间是不同的，因此各个频率声源信号在封闭空间中产生混响时间是不同的。

2.3 LCMV 分频的改进维纳后置滤波波束形成算法

由于混响是直达声的时间延迟和能量衰减造成的，因此从频谱上看混响部分的语音信号与直达语音信号是混合在一起的，但是不同频率的语音信号产生的混响是有一定差异的，因此采用分频处理方法去混响是很有必要的。由此得到混响环境下的 LCMV 分频改进的维纳后置滤波的波束形成算法 (LCMV frequency division with improved wiener filter beamforming algorithm in reverberation environment, LCMVFD-IWFR)。

将室内环境下的混响抑制作为约束条件，则在混响

环境下的 LCMV 分频改进的维纳滤波后置滤波的波束形成器设计的问题可以被建立为:

$$\begin{aligned} \min_{\mathbf{w}(\omega) \in \mathbb{C}^M} \quad & \mathbf{w}_{FH,low}^{*T}(\omega) \Phi_{xx}(\omega) \mathbf{w}_{FH,low}(\omega) \\ \text{s. t.} \quad & \begin{cases} \sum_{m=1}^M \sum_{l=1}^L \mathbf{w}_{FH,lcmv1}^{*T}(\omega) \mathbf{H}_{1m,l} = \mathbf{U}_{D1}(\omega) \\ \sum_{m=1}^M \sum_{l=1}^L \mathbf{w}_{FH,lcmv2}^{*T}(\omega) \mathbf{H}_{1m,l} = \mathbf{U}_{D2}(\omega) \\ \sum_{m=1}^M \sum_{l=1}^L \mathbf{w}_{FH,lcmv1}^{*T} \mathbf{H}_{nm,l}(\omega) = 0 \\ \sum_{m=1}^M \sum_{l=1}^L \mathbf{w}_{FH,lcmv2}^{*T} \mathbf{H}_{nm,l}(\omega) = 0 \end{cases} \end{aligned} \quad (27)$$

式中: \mathbb{C}^N 表示 N 维复数域, $\mathbf{w}_{FH,low}(\omega)$ 是 LCMV 分频波束形成的权矢量, $\mathbf{w}_{FH,high}(\omega)$ 是 LCMV 高频带波束形成的权矢量, $\mathbf{w}_{FH,low}(\omega)$ 是 LCMV 低频带波束形成的权矢量, $\mathbf{H}_{1m,l}$ 表示目标语音信号方向的第 m 个麦克风、长度为 l 的房间冲击响应矢量, $\mathbf{H}_{nm,l}$ 表示第 n 个语音信号方向, 第 m 个麦克风, 长度为 l 的房间冲击响应矢量, $\mathbf{U}(\omega) = [\mathbf{U}_D(\omega) \quad \mathbf{0}]$ 是约束条件响应矢量, 且 $\mathbf{U}_D(\omega) = [\mathbf{U}_{D1}(\omega) \quad \mathbf{U}_{D2}(\omega)]$ 用于对目标语音信号去混响。采用拉格朗日法, 则目标函数为:

$$L(\mathbf{w}) = \mathbf{w}_{FH,lcmv}^{*T} \Phi_{xx} \mathbf{w}_{FH,lcmv} + \lambda (\mathbf{w}_{FH,lcmv}^{*T} \mathbf{H} - \mathbf{U}) \quad (28)$$

式中: λ 为多维拉格朗日矢量。目标函数求对 $\mathbf{w}_{FH,lcmv}$ 的梯度并令之为 0, 即 $\frac{\partial L(\mathbf{w})}{\partial \mathbf{w}} = 0$, 得:

$$\mathbf{w}_{FH,lcmv} = (\Phi_{xx})^{-1} \mathbf{H} (\mathbf{H}^{*T} (\Phi_{xx})^{-1} \mathbf{H})^{-1} \mathbf{U} \quad (29)$$

式中: 根据 LCMV 准则, 由维纳滤波器输出语音信号的频域信息分别计算出高频段 $\mathbf{w}_{FH,lcmv1}$ 和低频段 $\mathbf{w}_{FH,lcmv2}$ 的最优权值为:

$$\mathbf{w}_{FH,lcmv1} = LCMV(\mathbf{x}(\omega), \mathbf{C}, \mathbf{U}_{D1}(\omega)) \quad (30)$$

$$\mathbf{w}_{FH,lcmv2} = LCMV(\mathbf{x}(\omega), \mathbf{C}, \mathbf{U}_{D2}(\omega)) \quad (31)$$

通过对比全频带的波束形成器的响应, 则:

$$\mathbf{w}_{FH,lcmv}(\omega) = \alpha \mathbf{w}_{FH,lcmv1} + (1 - \alpha) \mathbf{w}_{FH,lcmv2} \quad (32)$$

式中: α 是矩阵加权系数, 是正常数; $\mathbf{w}_{FH,lcmv}(\omega)$ 是 LCMV 分频波束形成的最优权矢量。那么全频带的处理权值为:

$$\mathbf{W}_{m,LCMV-IWFRopt}(\omega) = \mathbf{w}_{H,lcmv}(\omega) \mathbf{W}_{DRR}(\omega) \quad (33)$$

那么通过对比全频带的处理权值, 则分频维纳滤波器的 LCMV 波束形成的最优权矢量为:

$$\mathbf{W}_{m,LCMVFD-IWFRopt}(\omega) = \mathbf{w}_{FH,lcmv}(\omega) \mathbf{W}_{DRR}(\omega)$$

$$\mathbf{W}_{m,LCMVFD-IWFRopt}(\omega) =$$

$$\frac{(\alpha \mathbf{w}_{FH,lcmv1} + (1 - \alpha) \mathbf{w}_{FH,lcmv2}) \mathbf{W}_{DRR}(\omega) E[R\{\sum_{i=0}^{M-2} \sum_{j=i+1}^{M-1} \Phi_{RR_{ij}}(\omega)\}]}{E[\sum_{i=0}^{M-1} \Phi_{xx_i}(\omega)]} \quad (34)$$

3 性能评价

本文采用分段信噪比^[13] (SNRseg) 和语音质量评估^[14] (Perceptual evaluation of speech quality, PESQ), 用于评估语音去混响的性能。分段信噪比定义为:

$$SegSNR = \frac{10}{N_s} \sum_{n=1}^{N_s} \lg \left\{ \frac{\sum_{k=2n_{offset}N_s}^{2n_{offset}N_s+2n_{offset}-1} s^2(k)}{\sum_{k=2n_{offset}N_s}^{2n_{offset}N_s+2n_{offset}-1} [s(k) - \hat{s}(k)]^2} \right\} \quad (35)$$

式中: N_s 是语音帧个数, n_{offset} 是时间序列偏移量, $2n_{offset}$ 是语音帧长度, $s(k)$ 为原信号, $\hat{s}(k)$ 为估计信号。

对于 PESQ 分数^[14], 它是由 ITU-T 为 3.2 GHz 的移动电话和窄带语音编解码器的语音质量评估的建议。该语音质量评分范围 -0.5 ~ 4.5 分。当待测试的语音失真越严重, 得分越低, 特别是 -0.5 最低分。PESQ 定义为:

$$PESQ = 4.5 - 0.1D_{ind} - 0.0309A_{ind} \quad (36)$$

式中: D_{ind} 是平均干扰值, A_{ind} 是对称干扰值。

4 仿真结果和性能评价

本文实例通过 Image 镜像源函数来模拟封闭房间混响环境。首先设置仿真房间长宽高为 5.5 m × 3.3 m × 2.3 m, 其次选择线性阵列的麦克风位置 (2.00, 1.00, 1.20), (2.00, 1.05, 1.20), (2.00, 1.10, 1.20), (2.00, 1.15, 1.20), (2.00, 1.20, 1.20), (2.00, 1.25, 1.20), (2.00, 1.30, 1.20) (单位为 m), 且采用声源到麦克风阵列的角度为 90°, 该声源位置为 (3.10, 1.00, 1.20); 最后房间吸声系数为 (0.6, 0.9, 0.5, 0.6, 1.0, 0.8), 混响时间为 600 ms。

麦克风阵列接收信号是由声源与各个阵元间的脉冲响应采用镜像 Image 模型产生。根据语音的动态非平稳特性, 将麦克风阵列接收信号分割成互不重叠的信号帧, 每帧信号先加 Hamming 窗, 再进行傅里叶变换。

采用 Image 模型仿真条件: 以 mic1 为基准麦克风仿真麦克风阵列含混响语音信号。根据声源位置和麦克风阵列位置的先验信息, 分别计算 LCMV-WFR 的权值、LCMV-IWFR 的权值以及 LCMVFD-IWFR 的权值得到波束指向性; 然后, 采用不同混响抑制算法进行求和运算, 分别得到由每帧信号叠加重构的输入语音信号, 并对它们进行比较。原始语音和由 Image 模型产生的混响语音的时域波形和语谱如图 1、2 所示。

由图 1 和图 2 的时域波形知: 与原始语音信号相比, 镜像 Image 模型产生的参考麦克风 mic1 的混响语音信

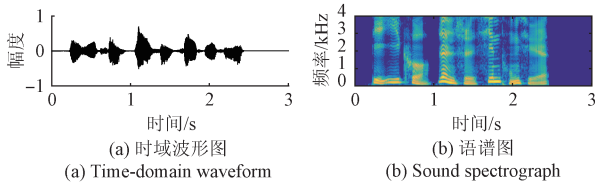


图1 原始语音信号
Fig.1 Original speech signal

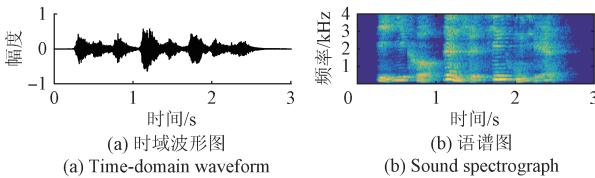


图2 参考麦克风 mic1 的混响语音
Fig.2 Reverberation speech of reference mike mic1

号在每个音节后面产生幅度衰减以及时间上的拖尾现象。图2比图1的时域波形多出了很多部分。根据封闭房间存在的混响效应定义,该多出的部分是叠加在原始语音上的混响部分。

由图1和2语谱图中可以看出,镜像 Image 模型产生的混响语音相对于原始语音的共振峰发生了频率的偏移,声纹能量不集中,呈现发散状态且颜色变浅。根据语谱图的定义,能量值的大小是通过颜色来表示的,以颜色深表示该点的语音能量越强。

混响环境下 LCMV 后置滤波的波束形成算法的输出信号的时域波形和语谱图分别如图3所示。

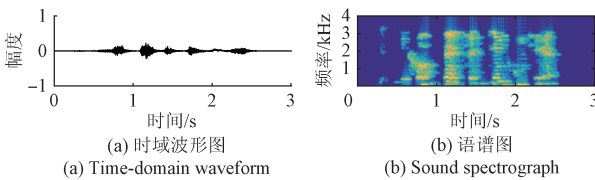


图3 混响环境下 LCMV 的维纳滤波后置波束形成算法输出信号(LCMV-WFR)
Fig.3 Output signal of Wiener post-filter beamforming algorithm of LCMV under reverberation environment (LCMV-WFR)

图3的时域波形表明,混响环境下 LCMV 的维纳后置滤波的波束形成算法(LCMV-WFR),LCMV-WFR 算法输出信号的时域波形相对于图2的时域波形,在幅度上波峰波谷界限分明,但是数据丢失严重。图3的语谱图表明,LCMV-WFR 算法输出信号相对于图2的谱图,共振峰之间干扰消除了,但谱信息丢失严重,抑制混响效果不好。

从图2的语谱图看出不同频率的混响时间不同,特

别是低频 1 000 Hz 以下的混响较为严重,以及高频信息的丢失。因此,需提高低频混响抑制效果,减少高频信息的丢失。本文两种改进的种不同混响抑制算法,在这两个方面具有较好效果,其输出信号的时域波形图和语谱图,如图4、5所示。

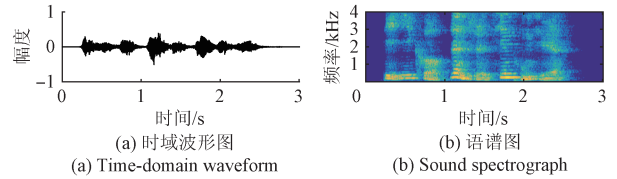


图4 混响环境下 LCMV 的改进维纳后置滤波波束形成算法输出信号
Fig.4 Output signal of improved Wiener post-filter beamforming algorithm of LCMV under reverberation environment

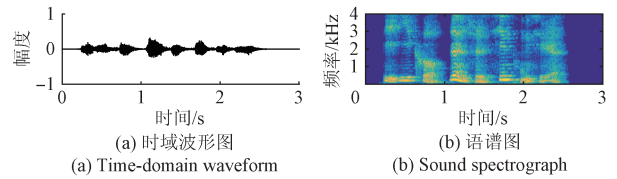


图5 混响环境下 LCMV 分频的改进维纳后置滤波波束形成算法输出信号
Fig.5 Output signal of improved Wiener post-filter beamforming algorithm of LCMV frequency under reverberation environment

图4、5的时域波形表明,LCMV-IWFR 波束输出比图2的时域波形图波峰波谷更加清晰,混响抑制效果有一定程度的提高,但是与图1的时域波形图相似度较低,因此混响抑制效果并不是很明显。与图4的时域波形图相比,本文提出的 LCMVFD-IWFR 算法信号输出较波峰波谷更加清晰,语音信息保留较好,混响抑制效果有改善。

图4、5的语谱图表明,LCMV-IWFR 波束输出的语谱图比图2的语谱图的共振峰前后重叠少,且共振峰频率时频域的偏移少,混响抑制效果有一定程度的提高,但是与图1的语谱图相似度不高,颜色变浅、能量削减严重,混响抑制效果并不是很明显。与 LCMV-IWFR 波束算法相比,而本文提出的 LCMVFD-IWFR 算法输出语谱图的共振峰彼此之间重叠少,且共振峰频率时频域的偏移少,混响抑制效果有改善。与图1的语谱图相似度较高,颜色变浅。因此 LCMVFD-IWFR 混响抑制效果较好。

采用两者评价标准,即分段信噪比和语音质量评估,来定量的评价目标语音去混响的性能,如表1所示。

表 1 混响抑制不同算法性能比较

Table 1 Performance comparison of different reverberation suppress algorithms

混响时间 600 ms	分段信噪比	语音质量评估
LCMV-WFR	2.76	1.937 5
LCMV-IWFR	3.38	2.109 4
LCMVFD-IWFR	3.67	2.204 4
均值	3.27	2.083 7

从表 1 中可以得出,应用这两种评价指标的评测结果显示混响抑制性能有一定的改善。综合 SNRseg 以及 PESQ 值,从表 1 知 LCMVFD-IWFR 的分段信噪比为 3.67 dB,语音质量评估为 2.2044,高于均值。这说明本文提出的 LCMVFD-IWFR 算法不仅能消除加性噪声的影响,而且对目标期望的方向上混响抑制具有一定程度改善。

5 结 论

本文主要研究提出了 LCMV 分频改进的维纳后置滤波的波束形成算法。分析混响环境下,延时对齐后的麦克风阵列接收信号各个通道间语音信号是相关的,而噪声是不相干的。因此通过计算各个通道之间的互相关的功率谱计算后置维纳滤波的系数来抑制噪声,但是这样的处理时时频域的处理针对封闭空间中存在的混响效果并不好,本文的直达声在冲击响应上称为前期混响,在一定的时间范围内,其余的称为后期混响,通过计算前期混响和后期混响的功率谱,给出改进的维纳后置滤波系数,但是这样的处理并没有完全考虑空域滤波,因此本文采用 LCMV 处理空间信息,给出相关约束条件计算整体权矢量,通过拉格朗日计算该算法的权矢量,在空域上兼顾频域信息,通过计算低频和低频两个频带的权矢量,加权为最优权矢量。最后,在混响时间的条件下,利用语音增强系统的评价标准,对该算法的性能进行了验证。

参考文献

- [1] ZOU Y X, WANG P. Speech enhancement with an acoustic vector sensor: An effective adaptive beamforming and post-filtering approach [J]. EURASIP Journal on Audio, Speech, and Music Processing, 2014, 14(1):17.
- [2] KOKKINAKIS K. Selective-tap blind dereverberation for two-microphone enhancement of reverberant speech[J]. IEEE Signal Processing Letters, 2009, 16(11): 961-964.
- [3] HIKICHI T, DELCROIX M. Inverse filtering for speech dereverberation less sensitive to noise and room transfer function fluctuations[J]. EURASIP Journal on Advances in Signal Processing, 2007, 7(1):1-12.
- [4] GOETZE S, KALLINGER M, MERTINS A, et al.

Estimation of the optimum system delay for speech dereverberation by inverse filtering [C]. International Conference on Acoustics, 2012.

- [5] YEGNANARAYANA B, MURTHY P S. Enhancement of reverberant speech using LP residual signal [J]. IEEE Transactions on Speech and Audio Processing, 2000, 8(3):267-281.
- [6] WU M. A two-stage algorithm for one-microphone reverberant speech enhancement [J]. IEEE Transactions on Audio, Speech, Lang Process, 2006, 14(3): 774-784.
- [7] HABETS E A P. Dual-microphone speech dereverberation using a reference signal [C]. International Conference on Acoustics Speech and Signal Processing (ICASSP), 2007, 8(4): 901-904.
- [8] LÖLLMANN H W, VARY P. Low delay noise reduction and dereverberation for hearing aids [J]. EURASIP Journal on Advances in Signal Processing, 2009, 1(1): 1-9.
- [9] HABETS E A P, BENESTY J, COHEN I, et al. On a tradeoff between dereverberation and noise reduction using the MVDR beamformer [C]. International Conference on Acoustics Speech and Signal Processing (ICASSP), 2009:3741-3744.
- [10] HABETS E A P, BENESTY J, COHEN I, et al. New insights into the MVDR beamformer in room acoustics [J]. IEEE Transactions on Audio Speech & Language Processing, 2010, 18(1):158-170.
- [11] GUSTAFSSON T, RAO B D, TRIVEDI M. Source localization in reverberant environments - modeling and statistical analysis [J]. IEEE Transactions on Speech and Audio Processing, 2003, 11(6): 791-803.
- [12] LEHMANN E A, JOHANSSON A M. Prediction of energy decay in room impulse responses simulated with an image-source model [J]. Journal of the Acoustical Society of America, 2008, 124(1): 269-277.
- [13] HU Y, LOIZOU P C. Subjective evaluation and comparison of speech enhancement algorithms [J]. Speech Communication, 2007, 49(7): 588-601.
- [14] FUJITA K, KATO T, YAMADA H, et al. SNR-dependent background noise compensation of PESQ values for cellular phone speech [C]. INTERSPEECH 2005 - Eurospeech, European Conference on Speech Communication and Technology, 2005:3165.
- [15] 毕杨, 王英民, 王奇. 宽带聚焦波束形成算法优化 [J]. 电子测量与仪器学报, 2016, 30(6):951-958.
- BI Y, WANG Y M, WANG Q. Research on broad-band focused beamforming algorithm optimization [J]. Journal of Electronic Measurement and Instrumentation, 2016, 30(6):951-958.

作者简介



郭业才,1986年于安庆师范学院获得学士学位,2003年于西北工业大学获得博士学位,现为南京信息工程大学教授、博导,主要研究方向为通信信道盲均衡理论与算法,气象信息技术与安全,水声信号处理等。

E-mail: guo-yecai@163.com

Guo Yecai, received B. Sc. from Anqing Normal College in 1986 and Ph. D. from Northwest University of Technology in 2003, respectively. And now, he is professor and Ph. D. supervisor in Nanjing University of Information Science & Technology. His main research direction is communication channel blind equalization theory and algorithm, information technique and safety, etc.

罗德与施瓦茨公司推出全新高端 GNSS 模拟器,高度还原现实测试场景

R&S SMW200A 高端 GNSS 模拟器是罗德与施瓦茨公司卫星导航系统模拟器产品家族的一颗耀眼的新成员。它可以扩展至多达 4 个 RF 输出,可同时模拟多天线和多频段环境下的 GNSS 信号,是市场上唯一一款能够内部并行模拟 GNSS 信号和复杂干扰环境的仪器。R&S SMW200A 的问世为罗德与施瓦茨公司开启了 GNSS 模拟的革新时代。

越来越多的 GNSS 接收机需要处理来自多个卫星导航系统不同频段的信号,例如 GPS、Glonass、Galileo 或者 BeiDou,有时甚至使用多根天线并行处理以提高定位精度,此外差分 GNSS (DGNSS) 定位技术也可以进一步提高定位精度。这些技术主要应用于自动驾驶以及飞机着陆等场景,定位的精确性和可靠性是不可或缺的。应用于这类场景的 GNSS 接收机必须经过大量的测试才能部署在车辆或飞机上。

借助复杂干扰功能,高度还原现实测试场景

R&S SMW200A GNSS 模拟器提供了一个革新的测试解决方案,可轻松仿真复杂且高度现实的测试场景,满足了多种多样的 GNSS 应用。为了测试多频段和多天线系统,用户可将 72 个 GNSS 信道分配多达 4 个 RF 输出。R&S SMW200A 不仅可以产生 GPS、Glonass、Galileo 和 BeiDou 信号,也可产生 QZSS 和 SBAS 信号。该解决方案使用户能够轻松快速地验证真实条件下接收机的定位精度。

R&S SMW200A 高端 GNSS 模拟器,是市场上第一个也是唯一一个具有内部噪声发生器同时可以产生具有多个干扰源的复杂干扰场景。所有信号(GNSS,噪声和干扰)都直接在仪器中生成,无需额外的干扰源,极大地简化了测试设置。

用户友好,面向未来

R&S SMW200A 内部集成直观的 GUI 人机交互界面,用户可快速轻松地生成各种 GNSS 场景,无需外部计算机进行配置和操作。得益于多种仪器选件,该方案可以根据用户需求进行最佳配置。R&S SMW200A 是一个可扩展的、面向未来的平台,可满足未来的各种测试要求,如测试新的 GNSS 信号。

罗德与施瓦茨公司已经正式发布全新高端的 R&S SMW200A GNSS 模拟器,该仪器将在俄勒冈州波特兰 ION GNSS +2017 贸易展上正式亮相。更多信息,请访问官方网站: www.rohde-schwarz.com/ad/press/gnss-solutions