

基于参数化波束形成器的 GSC 语音增强方法*

张传营 赵景玉 刘 扬 卜凡亮

(中国人民公安大学信息安全学院 北京 102600)

摘要:针对广义旁瓣相消器(GSC)中固定波束形成器在复杂环境下抑制旁瓣干扰和处理非平稳语音信号时存在局限性,提出了一种基于参数化波束形成器的改进 GSC 语音增强方法。该方法通过动态调节机制,在延迟求和波束形成器与超指向波束形成器之间进行灵活权衡与调节,有效抑制旁瓣干扰,增强了 GSC 在复杂声学环境中的鲁棒性与适应性。此外,引入互相关系数来调节自适应滤波器权重更新步长,有效应对语音信号变化导致的过减问题,提升了在非平稳语音信号中的处理精度。在 MATLAB 环境下开展仿真实验,针对 Babble 噪声、音乐噪声和白噪声环境,对比传统 GSC 和采用均方算法的 GSC,从三维波束方向图、不同背景噪声及参数条件下的降噪效果、互相关系数作用效果等方面进行评估,并利用分段信噪比(SNR)和语音质量感知评估(PESQ)等指标量化分析。结果显示,改进方法在降噪性能和语音清晰度上优势显著。在 Babble 噪声、音乐噪声、白噪声环境中,分段信噪比分别提升至 11.02、6.14 和 10.33 dB, PESQ 值分别提升至 3.65、3.20、3.25,并可通过调节参数实现不同噪声环境下的最佳降噪效果,有力验证了该方法在复杂声学环境中的有效性与优越性。

关键词: 麦克风阵列;波束形成器;广义旁瓣相消器;参数化;互相关系数

中图分类号: TN912.35 **文献标识码:** A **国家标准学科分类代码:** 510.40

GSC improved speech enhancement method based on parameterized beamformer

Zhang Chuanying Zhao Jingyu Liu Yang Bu Fanliang

(School of Information and Network Security, People's Public Security University of China, Beijing 102600, China)

Abstract: To address the limitations of the fixed beamformer in the generalized sidelobe canceller (GSC) in suppressing sidelobe interference and processing non-stationary speech signals in complex environments, this paper proposes an improved GSC-based speech enhancement method utilizing a parameterized beamformer. The proposed method employs a dynamic tuning mechanism to flexibly balance and adjust between the delay-and-sum (DS) beamformer and the super-directive (SD) beamformer, effectively suppressing sidelobe interference and enhancing the robustness and adaptability of the GSC in complex acoustic environments. Furthermore, a cross-correlation coefficient is introduced to regulate the step size of the adaptive filter weight update, mitigating the over-attenuation issue caused by variations in speech signals and improving the processing accuracy for non-stationary speech signals. Simulation experiments were conducted in MATLAB to evaluate the performance of the proposed method under various noise conditions, including Babble noise, music noise, and white noise. The performance of the proposed method was compared with that of the traditional GSC and the GSC with the least mean square (LMS) algorithm. The evaluation was carried out from multiple perspectives, including 3D beam patterns, noise reduction effects under different background noise and parameter conditions, and the effectiveness of the cross-correlation coefficient. Quantitative analysis was performed using performance metrics such as segmental signal-to-noise ratio (SNR) and perceptual evaluation of speech quality (PESQ). The results demonstrate that the proposed method significantly outperforms the traditional GSC in terms of noise reduction performance and speech clarity. In environments with Babble noise, music noise, and white noise, the segmental SNR improved to 11.02, 6.14, and 10.33 dB, respectively, while the PESQ values increased to 3.65, 3.20, and 3.25, respectively. By adjusting the parameters, the proposed method achieves optimal noise reduction effects in different noise environments, validating its

effectiveness and superiority in complex acoustic scenarios.

Keywords: microphone array; beamformer; generalized sidelobe canceller; parameterization; cross-correlation coefficients

0 引言

随着现代通信技术与人工智能的发展,语音作为人类最自然、直接的交流方式,其高效、准确的传递在众多领域中显得至关重要。然而,实际环境中,语音信号不可避免地受到各种噪声干扰,如环境噪声、传输介质噪声以及设备噪声等,这些因素严重影响语音的清晰度、可懂度及后续处理系统的性能。因此,对受噪声污染的语音信号进行降噪,恢复原始纯净语音,成为语音处理技术中不可或缺的一环。

传统的语音增强方法,如谱减法^[1-2]、小波变换^[3-5]、自适应滤波(如卡尔曼滤波^[6]、维纳滤波^[7])等,虽在一定条件下取得了一定成效,但面对复杂、非平稳噪声、声源位置快速变化的场景,往往难以实现精准的噪声抑制与目标语音恢复。谱减法虽然简单易实现,但在处理非平稳噪声时容易产生“音乐噪声”现象;小波变换虽然能够较好地处理瞬态噪声,但在复杂噪声环境下的鲁棒性较差;自适应滤波方法(如卡尔曼滤波和维纳滤波)虽然能够动态调整滤波器参数,但在噪声统计特性快速变化时,性能显著下降。因此,传统方法在复杂环境下的应用受到了较大限制。在这种背景下,麦克风阵列技术^[8-9]因其具备空间滤波特性,能够实现声源定位^[10]、波束形成与噪声抑制的多重功能,日益成为语音增强领域的重要研究对象。

Pradeep 等^[11]提出的延迟求和波束形成(delay-and-sum, DS)方法是麦克风阵列的主要方法。通过适当的延迟补偿,使多通道信号在观察方向上对齐,将对准的信号相加以确保观察方向的增益最大化。Bai 等^[12]于 2022 年提出的广义旁瓣相消器(generalized sidelobe canceller, GSC)与多通道维纳滤波器(multichannel wiener filter, MWF)相结合的方法,利用 GSC 的干扰抑制能力和 MWF 的信号估计能力,实现更高效的语音增强。Kim^[13]提出了一种基于双麦克风信号相位差的目标-非目标方向信号比估计方法,并改进固定波束形成器和控制自适应噪声消除器的权重更新。然而,尽管 GSC 在语音增强领域取得了显著成果,现有研究仍存在一些局限性。首先, GSC 框架中的固定波束形成器在复杂、动态变化环境中的适应能力有限,特别是在指向性、旁瓣水平以及对环境变化的响应方面存在固化的局限性。其次, GSC 方法的误差信号由残留噪声信号和增强的期望信号组成,当语音信号的幅度突然增加时,增强的期望信号部分会增加,导致误差信号的增加,而残留的噪声信号却保持相对不

变,从而在后续帧中可能导致过减现象,进一步影响信号的失真。

针对上述问题,提出了一种基于参数化波束形成器的改进 GSC 语音增强方法。与现有研究不同,通过引入参数化波束形成器设计方法^[14-15],在延迟求和波束形成器与超指向波束形成器之间进行动态调节,有效提升了 GSC 在复杂环境下的适应性和鲁棒性。此外,在估计的噪声和误差信号之间引入互相关系^[16],通过调节自适应滤波器权重更新步长,避免了过减现象的发生,进一步提高了语音增强的效果。与传统的固定波束形成器 GSC 方法相比,不仅能够灵活应对复杂环境中的噪声变化,还能有效抑制非平稳噪声对语音信号的干扰,显著提升了语音的清晰度和可懂度。

1 传统 GSC

GSC 是自适应波束形成器的一种语音增强^[17]模型,旨在优化信号接收 and 处理流程,增强目标信号以及抑制噪声的干扰。该模型主要由固定波束形成器、阻塞矩阵和自适应噪声消除器 3 大核心组件构成。固定波束形成器,通常采用延迟和求和波束形成器^[18-20]的结构,其根本目的是在主瓣方向上增强目标信号的强度,同时削弱非目标方向的信号干扰,提升信号的接收品质。阻塞矩阵的作用是去除期望目标信号的成分,提取出噪声信号,这一步骤为后续的噪声抑制提供了关键的参考信号。自适应噪声消除器根据阻塞矩阵提供的噪声信号,通过自适应算法^[21-22],动态调整权重参数,最大化输出信号中的目标信号,同时最小化噪声干扰信号。

2 设计方法

固定波束形成器是 GSC 的重要组成部分,传统的做法通常采用延迟求和波束形成器。然而,延迟求和波束形成器在应用中存在一些固有的缺陷,比如指向性不强、主瓣波束较宽等问题,这些问题不能应对复杂环境的动态变化,进而影响了降噪效果。因此引入了参数化波束形成器设计方法,使其能够在延迟求和波束形成器和超指向(super directive, SD)波束形成器之间进行灵活的调节。

在设计过程中,通过对零点的转向向量和白噪声增益(white noise gain, WNG)的伪相关矩阵进行酉对角化,得到正交特征向量,特征向量用于构建波束形成器的参数空间,引入参数 N ,构造大小为 $M \times N$ 的特征向量子矩阵,选取前 N 个特征向量的列向量构成新的波束形成器

结构,并定义一个新的向量作用于特征向量子矩阵,得到波束形成器的权重矩阵。

当 $N = 1$ 时,参数化波束形成器为 DS 波束形成器,实现较高的白噪声增益,但指向性较低;当 $N = M$ 时,参数化波束形成器为 SD 波束形成器,表现为较高的指向性因子(directivity factor, DF),但白噪声增益降低;因此可以通过调整参数 N ,在两种波束形成器之间进行调节,实现降噪效果的最优化。

参数化波束形成器能够通过调整波束形成器的权重,实现对波束指向性和宽度的控制。相比于延迟求和波束形成器,参数化波束形成器可以更好地聚焦目标信号,同时抑制来自其他方向的干扰噪声。

3 基于参数化波束形成器的 GSC

3.1 信号模型

改进的 GSC 分为两条路径:主路径和辅助路径。主路径将参数化波束形成器替代传统固定波束形成器,生成具有残留噪声的去噪信号。辅助路径估计主路径输出信号中的噪声分量。 $w_a(n)$ 是使麦克风阵列信号 $x(n)$ 在注视方向上对齐以抑制来自另一个方向的噪声的权值向量。 y_a 为主路径波束形成输出信号的对齐信号之和。辅助路径估计主路径波束形成输出信号中的噪声分量 y_N 作为抵消信号。抵消信号将从波束形成输出信号中去除,将互相关系数引入到权重更新路径中,估计噪声信号和误差信号之间的相关性,来控制自适应滤波器权重更新过程中的步长,通过调整步长,控制滤波器的权重更新过程。增强信号 y 由期望信号组成,得到残差噪声信号,结构如图 1 所示。

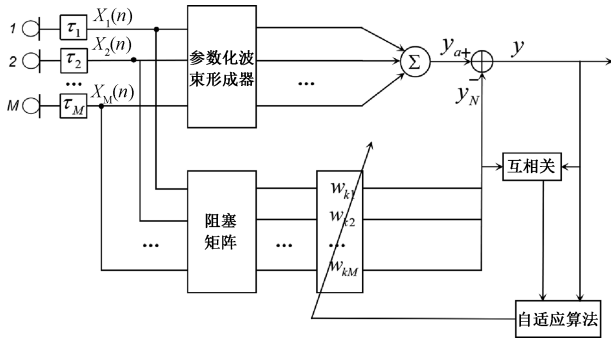


图 1 基于参数化波束形成器的 GSC

Fig. 1 GSC based on parametric beamformer

在当前研究的背景下,考虑一个由 M 个全向麦克风组成的均匀线性阵列,相邻麦克风元素之间的间隔为 δ 。假定在一个无反射,无回声的条件下存在一个远场平面波,该波在声速为 $c = 340$ m/s 的介质中传播,并以入射角 $\theta = 0$ 接触该阵列。在这个场景中,长度为 M 的转向向量可以定义为:

$$\mathbf{d}(\omega) = [1 \quad e^{-j\omega\delta/c} \quad \dots \quad e^{-j(M-1)\omega\delta/c}]^T \quad (1)$$

式中: $\omega = 2\pi f$ 为角频率; f 为时间频率。

可得第 m 个麦克风的观测信号为:

$$x_m(\omega) = d_m(\omega) \times s_m(\omega) + n_m(\omega) \quad (2)$$

式中: s 为干净语音信号; n 为噪声信号,记为:

$$\mathbf{X}(\omega) = [x_1(\omega), x_2(\omega), \dots, x_M(\omega)]^T \quad (3)$$

设置权向量 $\mathbf{W}_a(\omega) = [w_1(\omega), w_2(\omega), \dots, w_M(\omega)]^T$,在主路径将 $\mathbf{X}(\omega)$ 通过复值线性滤波器 $\mathbf{W}_a(\omega)$ 形成通路,得到接收信号 $y_a(\omega)$:

$$y_a(\omega) = \mathbf{W}_a^H(\omega) \mathbf{X}(\omega) = \mathbf{W}_a^H(\omega) \mathbf{d}_0(\omega) S(\omega) + \mathbf{h}^H(\omega) \mathbf{n}(\omega) \quad (4)$$

假设在所有麦克风上噪声的方差都相同,可以得到 $y_a(\omega)$ 的协方差矩阵为:

$$\Phi_{y_a} = E(y_a y_a^H) = \mathbf{W}_a^H(\omega) \mathbf{d}_0 \mathbf{d}_0^H \mathbf{W}_a(\omega) \phi_s + \mathbf{W}_a^H(\omega) \mathbf{\Gamma} \mathbf{W}_a(\omega) \phi_n \quad (5)$$

式中: $\mathbf{\Gamma}$ 表示为噪声的伪相关矩阵。

在辅助路径将 $\mathbf{X}(\omega)$ 经过阻塞矩阵 $\mathbf{B}$ 滤除目标信号得到噪声估计信号 $\mathbf{U}(\omega)$:

$$\mathbf{U}(\omega) = \mathbf{B}^H \mathbf{X}(\omega) \quad (6)$$

其中,阻塞矩阵 $\mathbf{B}$ 为:

$$\mathbf{B} = \begin{bmatrix} -1 & e^{j\omega\delta/c} & 0 & \dots & 0 \\ 0 & -1 & e^{j2\omega\delta/c} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & e^{j(M-1)\omega\delta/c} \end{bmatrix} \quad (7)$$

将 $\mathbf{U}(\omega, n)$ 通过自适应滤波器抵消 $y_a(\omega, n)$ 中的相应噪声成分,得到增强后的语音信号:

$$y(\omega, n) = y_a(\omega, n) - y_N(\omega, n) \quad (8)$$

3.2 参数化波束形成器

对零点的转向向量和白噪声增益的伪相关矩阵 $\mathbf{d}_0 \mathbf{Id}_0^H$ 进行酉对角化,为了简洁起见,将省略对角频率 ω 的表示为:

$$\mathbf{R}^H \mathbf{d}_0 \mathbf{Id}_0^H \mathbf{R} = \mathbf{\Lambda} \quad (9)$$

式中:特征向量 $\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \dots \quad \mathbf{r}_M]$ 是大小为 $M \times M$ 的满秩矩阵;特征值 $\mathbf{\Lambda}$ 大小为 $M \times M$,且 $\mathbf{\Lambda} = \text{diag}(\lambda_1, 0, \dots, 0)$; λ_1 是矩阵 $\mathbf{d}_0 \mathbf{Id}_0^H$ 的唯一非零实数特征值,并且满足:

$$\lambda_1 = \mathbf{d}_0^H (\mathbf{\Gamma}_d)^{-1} \mathbf{d}_0 \quad (10)$$

λ_1 对应的特征向量为 $\mathbf{r}_1$,可以表示为:

$$q_1 = \frac{(\mathbf{\Gamma}_d)^{-1} \mathbf{d}_0}{\sqrt{\mathbf{d}_0^H (\mathbf{\Gamma}_d)^{-1} \mathbf{d}_0}} \quad (11)$$

由于 $\mathbf{W}_a(\omega)$ 为 $M \times 1$ 的矩阵,所以可由 $\mathbf{R}$ 的线性组合来表示,因此定义一个向量 $\boldsymbol{\alpha}$,使得 $\mathbf{W}_a(\omega) = \mathbf{R} \boldsymbol{\alpha}$,其中 $\boldsymbol{\alpha} = (\alpha_1 \alpha_2, \dots, \alpha_M)^T$ 。代入式(4)可以得到:

$$y_a(\omega) = \mathbf{W}_a^H(\omega) \mathbf{d}_0(\omega) + \mathbf{W}_a^H(\omega) \mathbf{n}(\omega) =$$

$$\alpha^H R^H d_0(\omega) + \alpha^H R^H n(\omega) \quad (12)$$

选取 R 的前 N ($1 \leq N \leq M$) 个特征向量构成新的波束形成器结构:

$$R_{1:N} = [r_1 \ r_2 \ \cdots \ r_N] \quad (13)$$

权向量可以写为:

$$W_a = R_{1:N} \alpha \quad (14)$$

式中: $\alpha = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_N)^T$ 是长度为 N 的向量。

为了实现参数调节,可以利用延迟求和波束形成器可以实现最大化白噪声增益,超指向波束形成器可以实现最大化指向性因子的性质。

将 WNG 和 DF 分别表示为:

$$W(h) = \frac{|W_a^H d_0|^2}{W_a^H I W_a} = \frac{|\alpha^H R_{1:N}^H d_0|^2}{\alpha^H R_{1:N}^H I R_{1:N} \alpha} = \frac{|\alpha^H R_{1:N}^H d_0|^2}{\alpha^H \alpha} \quad (15)$$

$$D(h) = \frac{|W_a^H d_0|^2}{W_a^H \Gamma_d W_a} = \frac{|\alpha^H R_{1:N}^H \Gamma_d R_{1:N} \alpha|}{\alpha^H R_{1:N}^H \Gamma_d R_{1:N} \alpha} \quad (16)$$

最大化 WNG 和 DF:

$$\begin{aligned} \min \alpha^H \alpha & \quad \min \alpha^H R_{1:N}^H \Gamma_d R_{1:N} \alpha \\ \text{s. t. } \alpha^H R_{1:N}^H d_0 &= 1 \quad \text{s. t. } \alpha^H R_{1:N}^H d_0 = 1 \end{aligned} \quad (17)$$

从而可以推导出:

$$W_a = \frac{P d_0}{d_0^H P d_0} \quad (18)$$

其中 $P = R_{1:N} (R_{1:N}^H \Gamma_d R_{1:N})^{-1} R_{1:N}^H$ 。

当 $N = 1$ 时,可以得出:

$$W_a = \frac{1}{d_0^H} = \frac{d_0}{d_0^H d_0} = \frac{d_0}{M} = W_{DS} \quad (19)$$

当 $N = M$ 时,可以得出:

$$W_a = \frac{\Gamma_d^{-1}(\omega) d_0}{d_0^H \Gamma_d^{-1}(\omega) d_0} = W_{SD} \quad (20)$$

通过调整正整数 N ,可以得到性能介于 W_{DS} 和 W_{SD} 的之间不同的波束形成器。

3.3 互相关系数

互相关系数是一种衡量两个信号之间相关性的指标。在许多情况下,语音信号与噪声信号是不相关的。因此,可以利用估计的噪声信号和误差信号之间的互相关系数,来控制自适应滤波器权重更新过程中的步长。通过调整步长,可以更好地控制滤波器的权重更新过程,从而避免发生过减现象,提高信号处理的效果。

$$w_k(n+1) = w_k(n) + \rho(k-1) \mu x(n) e(n) \quad (21)$$

$$\rho(k) = E(y_N(k) \times y(k)) = E[y_N(k) \times (s(k) + r(k))] = E(y_N(k) \times s(k)) + E(y_N(k) \times r(k)) \quad (22)$$

式中: $\rho(k)$ 是估计噪声信号 $y_N(k)$ 和误差信号 $y(k)$ 之间的第 k 帧的互相关系数; $s(k)$ 表示干净的期望语音信号; $r(k)$ 是输出信号的残留噪声信号。因为语音信号通常被假设为与噪声信号不相关,所以估计的噪声和语音

信号之间的互相关系数接近于 0。

$$E(y_N(k) \times s(k)) \approx 0 \quad (23)$$

因此, $\rho(k)$ 等于估计噪声信号 $y_N(\omega, n)$ 和残留噪声信号 $r(k)$ 之间的互相关系数:

$$\rho(k) \approx E(y_N(k) \times r(k)) \quad (24)$$

当 $\rho(k)$ 很小时,残余噪声与估计噪声之间的相关性几乎为 0,此时权重系数变化不大。反之,权重系数则会显著调整,以便迅速减少残余噪声的影响。

自适应噪声抵消模块的输出 $y_N(\omega, n)$ 为:

$$y_N(\omega, n) = \sum_{l=0}^{L-1} [W_k(\omega, n)]^T U(\omega, n-l) \quad (25)$$

增强后的语音信号为:

$$y(\omega, n) = y_a(\omega, n) - y_N(\omega, n) \quad (26)$$

4 仿真实验

4.1 实验环境

仿真实验是在 MATLAB 环境中进行的。为了验证该方法的有效性和优越性,与传统 GSC 语音增强方法进行了比较,并对比了不同参数条件下的降噪效果。实验使用的均匀线性麦克风阵列由 8 个麦克风组成,相邻两个麦克风之间的间距为 $\delta = 1$ cm。实验使用的干净信号是一段清晰、标准的语音,为了模拟真实环境中的噪声情况,分别加入了有色噪声信号和白噪声,得到两种噪声环境下的降噪效果分别如图 2 和 3 所示。

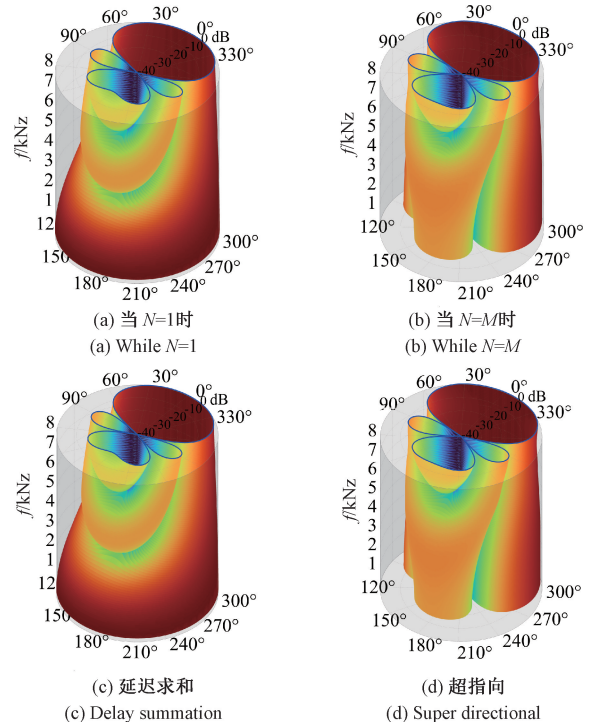


图 2 三维波束方向图

Fig. 2 3D beam pattern

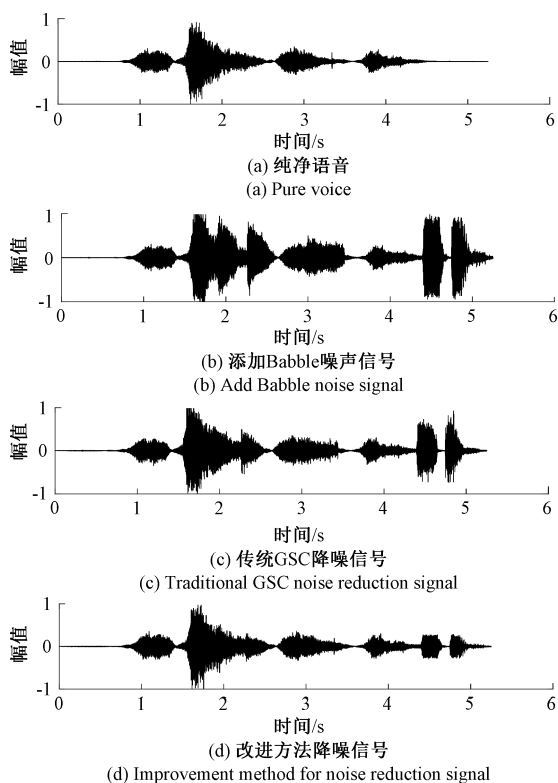


图3 添加 Babble 噪声降噪对比

Fig. 3 Adding Babble noise for noise reduction comparison

4.2 实验结果

1) 三维波束方向图

波束形成器三维波束方向图如图 2(a) 和 (b) 所示, 根据式 (17) 和 (18) 所得到的延迟求和波束形成器和超指向波束形成器的波束方向图如图 2(c) 和 (d) 所示。根据图 2 显示的实验结果, 可以看出当 $N = 1$ 时与 DS 的波束方向图几乎一致, 表明当 $N = 1$ 时所得到的波束形成器可以退化为延迟求和波束形成器, 此时该波束形成器白噪声增益最好, 而在低频时指向性能较差; 当 $N = M$ 时与 SD 的波束方向图一致, 表明当 $N = M$ 时所得到的波束形成器可以退化为超指向波束形成器, 此时可以看出该波束形成器在低频时也能表现出良好的指向性, 但白噪声增益相对较弱。

2) 不同背景噪声条件下的降噪效果对比

在实验中, 添加了 Babble 噪声、音乐噪声和白噪声, 研究不同噪声背景下降噪对比, 设置参数 $N = 2$, 初始信噪比分别设定为 $SNR = -4.9, -9.42, -4.27$ dB, 得到了处理前后的语音时域波形, 如图 3~5 所示, 并将该方法与传统 GSC 方法的降噪效果进行比较。图 3~5 横轴为信号序列的采样点数, 纵轴为信号幅度。从图 3~5 可以看出, 3 种不同类型的噪声环境下, 传统 GSC 降噪方法虽有一定的降噪效果, 但相比于改进方法效果较差, 该方法能

够有效地降低噪声水平, 并且保留了语音信号的清晰度和可懂度。

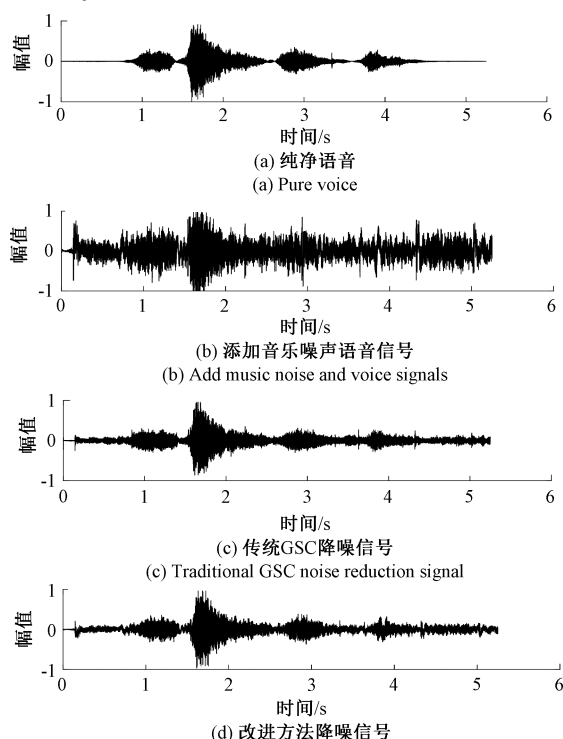


图4 添加音乐噪声降噪对比

Fig. 4 Adding music noise for noise reduction comparison

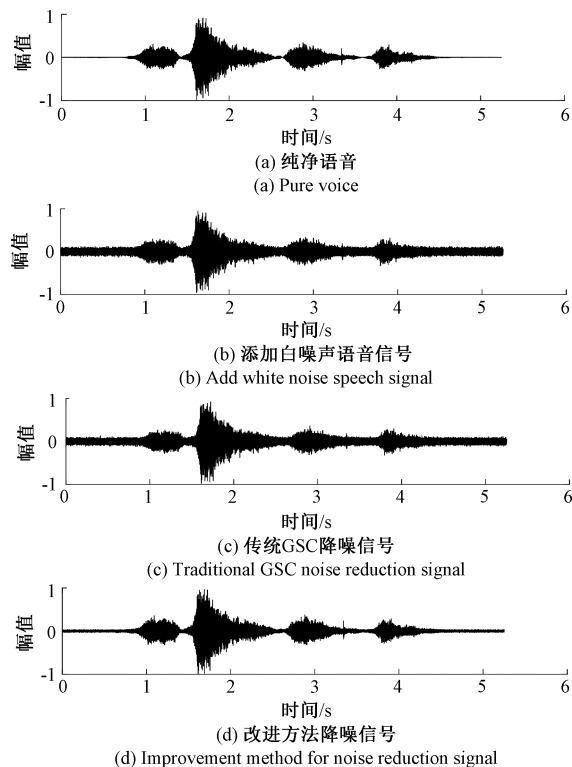


图5 添加白噪声降噪对比

Fig. 5 Adding white noise for noise reduction comparison

3) 不同参数条件下的降噪效果对比

通过调整参数 N ($N=1,2,5,8$),得到处理后的语音时域波形,如图 6 所示。从图 6 可以看出当 $N=1$ 时,改进的降噪方法与传统 GSC 语音降噪方法有相似的特性,因为当 $N=1$ 时,参数化波束形成器为延迟求和波束形成器,另外引入了自相关系数,相比较传统 GSC 方法,更接近于原始干净信号,降噪效果更好。通过进一步比较,当 $N=2$ 时,降噪效果相较于 $N=1$ 有了显著的提升,而当 $N=5$ 时,降噪效果又明显下降,因此可以通过调节参数 N ,找到最佳的参数配置。

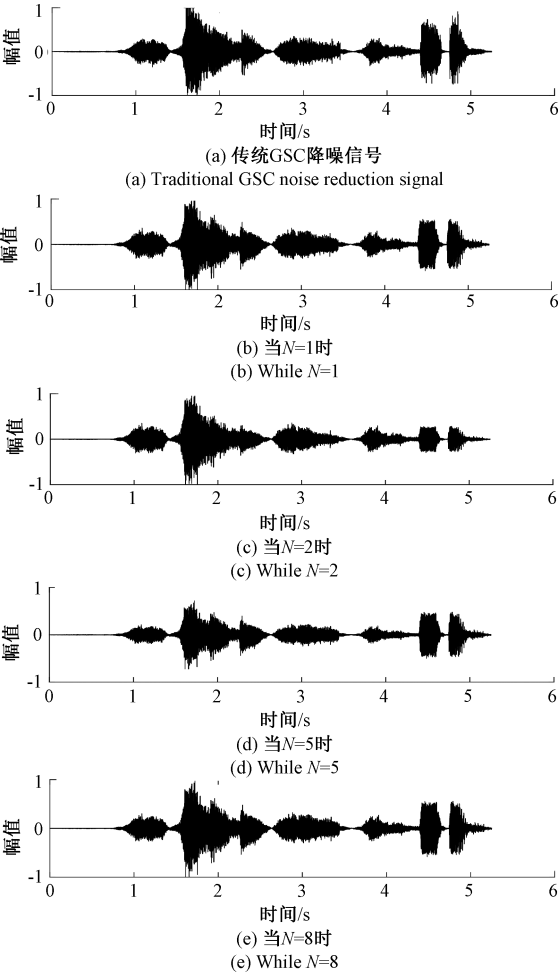


图 6 不同参数降噪对比
Fig. 6 Comparison of noise reduction with different parameters

4) 互相关系数作用效果

为了验证互相关系数在非平稳信号处理中的有效性,可以通过比较估计噪声能量与实际噪声能量来进行评估。如果估计噪声能量大于实际噪声能量,说明出现了过减现象,会导致额外的失真,从而影响语音质量。相反,如果估计噪声能量小于实际噪声能量,意味着自适应

滤波器在估计和消除噪声方面表现良好,能够有效地减少或消除噪声的影响,而不会引入过多的失真^[24]。未引入互相关系数时噪声帧能量的比较结果如图 7 所示。引入互相关系数后的噪声帧能量比较结果如图 8 所示。通过比较两个实验的结果,在引入互相关系数后,估计噪声能量明显小于实际噪声能量,表明引入互相关系数有助于避免过减现象的发生,提高了自适应滤波器在非平稳信号中对噪声的估计能力,从而提升了语音处理的整体性能。

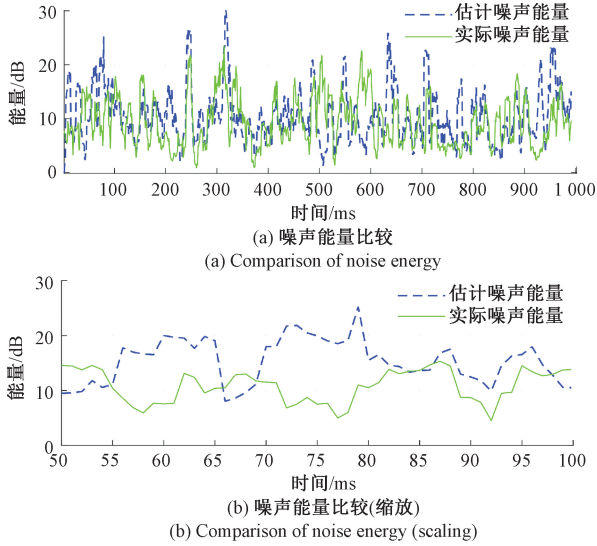


图 7 未引入互相关系数噪声能量比较
Fig. 7 Comparison of noise energy without introducing cross-correlation numbers

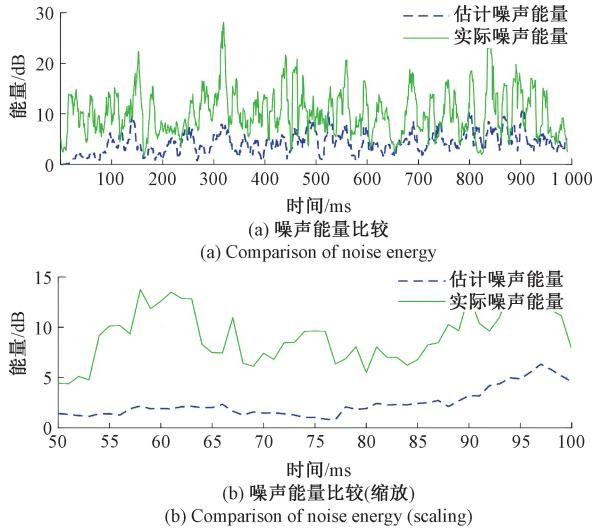


图 8 引入互相关系数后噪声能量比较
Fig. 8 Comparison of noise energy after introducing cross-correlation coefficients

4.3 客观评价标准

为了进一步比较改进的方法与传统 GSC 方法、采用最小均方 (least mean square, LMS) 的 GSC 波束形成器的性能,采用了分段信噪比 (segmented signal-to-noise ratio, SNR), 语音质量感知评估 (perceptual evaluation of speech quality, PESQ) 等指标,以全面评估改进方法与传统 GSC 方法、采用 LMS 的 GSC 波束形成器在不同噪声环境和不同信噪比水平下的性能。

在不同噪声类型和 SNR 水平下,通过调整参数设置所得到的分段信噪比结果如表 1 所示。与传统 GSC 和采用 LMS 的 GSC 相比,改进方法在各个测试噪声条件下均表现出更高的 SNR 值。此外可以调节参数设置,找到针对不同噪声类型和初始 SNR 的最优配置。可以看出,在 3 种噪声类型下, N 的最优选择与初始信噪比有关。

在 5.16 dB 的音乐噪声背景下, $N = 2$ 时降噪效果最好,而在相同信噪比的白噪声背景下,此时 $N = 3$ 时的降噪效果最好,这一结果说明, N 的最优选择不仅与初始信噪比有关而且还与噪声类型有关。表 2 为 PESQ 值在未引入互相关系数情况下的结果,一方面可以看出在 3 种噪声类型下,未引入互相关系数的 PESQ 值均高于传统 GSC 方法和采用 LMS 的 GSC 方法,并且可以通过选择最优参数配置达到最佳降噪效果,说明提出的参数化波束形成器,性能表现更优异。另一方面,对比表 3 引入互相关系数的参数化波束形成器 GSC 降噪方法, PESQ 的值相对较低,说明引入相互相关系数可以进一步提高 PSEQ 的值,由于互相关系数可以调节权重更新步长,可以有效解决非平稳信号处理过程中的过减问题,减少了信号失真,进一步提高了语音质量。

表 1 不同噪声类型下的分段信噪比 (有互相关系数)

Table 1 Segmented signal-to-noise ratio under different types of noise (cross-correlation coefficient)

背景噪声	SNR/dB	传统 GSC	GSC with LMS	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$	$N = 7$	$N = 8$
Babble 噪声	-4.89	2.84	9.80	9.34	11.02	10.85	10.68	9.61	8.30	6.67	6.06
	-1.37	4.37	14.03	13.01	14.23	14.24	13.64	11.28	11.42	7.89	9.34
	4.65	6.14	21.23	21.32	22.23	22.57	21.40	18.66	19.41	15.20	18.30
音乐噪声	-9.42	-3.32	4.62	5.33	6.14	5.83	5.52	4.78	4.10	4.01	4.00
	-0.17	5.16	11.68	12.75	13.50	13.62	11.22	9.60	9.76	6.43	7.16
	5.16	9.45	18.32	19.77	20.31	20.18	19.34	17.82	17.76	15.57	15.68
白噪声	-4.27	4.42	10.21	10.33	9.88	9.94	8.02	8.56	8.66	8.61	8.42
	-0.67	7.43	14.56	14.59	14.73	14.75	14.70	14.58	14.61	14.16	14.40
	5.16	12.10	17.38	18.08	18.97	19.27	18.50	16.96	16.60	12.32	12.76

表 2 不同噪声类型下的 PESQ 值 (无互相关系数)

Table 2 PESQ values under different types of noise (no cross-correlation coefficient)

背景噪声	SNR/dB	传统 GSC	GSC with LMS	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$	$N = 7$	$N = 8$
Babble 噪声	-4.89	3.142	3.568	3.368	3.609	3.539	3.367	3.250	3.200	3.093	3.329
	-1.37	3.217	3.691	3.518	3.727	3.666	3.751	3.400	3.357	3.258	3.475
	4.65	3.284	3.782	3.684	3.816	3.781	3.827	3.582	3.542	3.447	3.643
音乐噪声	-9.42	2.643	2.921	2.813	3.115	3.015	3.145	2.637	2.578	2.469	2.721
	-0.17	2.868	3.432	3.237	3.584	3.460	3.619	3.058	2.997	2.887	3.138
	5.16	2.969	3.668	3.427	3.759	3.656	3.787	3.217	3.142	3.025	3.319
白噪声	-4.27	2.612	3.021	2.903	3.201	3.107	3.236	2.732	2.681	2.598	2.829
	-0.67	2.783	3.267	3.076	3.303	3.268	3.398	2.911	2.851	2.735	3.012
	5.16	3.103	3.496	3.353	3.619	3.541	3.653	3.212	3.158	3.040	3.305

表 3 不同噪声类型下的 PESQ 值 (有互相关系数)

Table 3 PESQ values under different types of noise (cross-correlation coefficient)

背景噪声	SNR/dB	传统 GSC	GSC with LMS	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$	$N = 6$	$N = 7$	$N = 8$
Babble 噪声	-4.89	3.142	3.593	3.435	3.655	3.592	3.675	3.301	3.251	3.158	3.368
	-1.37	3.217	3.721	3.584	3.762	3.706	3.776	3.441	3.396	3.314	3.512
	4.65	3.284	3.823	3.723	3.835	3.803	3.844	3.619	3.579	3.501	3.673
音乐噪声	-9.42	2.643	2.960	2.902	3.201	3.099	3.232	2.723	2.662	2.558	2.808
	-0.17	2.868	3.526	3.335	3.683	3.563	3.716	3.141	3.082	2.977	3.230
	5.06	2.969	3.782	3.535	3.838	3.745	3.864	3.324	3.251	3.122	3.422
白噪声	-4.27	2.612	3.086	2.965	3.253	3.156	3.281	2.778	2.719	2.634	2.867
	-0.67	2.783	3.312	3.152	3.429	3.334	3.455	2.981	2.915	2.804	3.067
	5.16	3.103	3.536	3.438	3.681	3.601	3.703	3.276	3.220	3.122	3.356

基于参数化波束形成器的 GSC 语音增强方法在两种客观评价指标上均超越了传统 GSC 方法和采用 LMS 的 GSC 方法,不仅在 SNR 上实现了显著提升,保障了信号的纯净度,而且在 PESQ 指标上的表现也证明了该方法在降低噪声、提高语音感知质量方面有着显著的效果,验证了改进方法在复杂声学环境下的有效性和优越性。

5 结 论

针对 GSC 中固定波束形成器在复杂环境下适应性不足、旁瓣干扰抑制效果有限以及非平稳语音信号处理中容易出现过减现象的问题,提出基于参数化波束形成器的改进 GSC 语音增强方法。该方法通过动态调节机制,在延迟求和波束形成器与超指向波束形成器之间进行灵活权衡与调节,有效抑制旁瓣干扰,增强了 GSC 在复杂声学环境中的鲁棒性与适应性。引入互相关系数来调节自适应滤波器权重更新步长,有效应对语音信号变化导致的过减问题,提升了在非平稳语音信号中的处理精度。通过参数化波束形成器的设计,本文方法能够在不同噪声环境下动态调整波束形成器的指向性和增益,显著提升了语音增强的效果。该方法不仅展示了在语音增强领域的先进性和必要性,也为未来进一步提高增强效果、优化算法以适应实时应用、以及在复杂声学环境中的应用提供了新的思路和方向。下一步,可以考虑在实际场景中结合深度学习算法,可以实现参数的自动调节和优化,进一步提升系统的实时性和适应性。针对多噪声源和复杂声场环境,未来的研究可以探索多维阵列和分布式麦克风网络的应用,进一步提升系统的空间分辨率和噪声抑制能力。

参考文献

[1] 吴礼福,申浩.掩蔽法减少谱减法去混响中的音乐噪声[J].电子测量与仪器学报,2017,31(11):1855-1859.
WU L F, SHEN H. Masking method to reduce musical noise in reverberation reduction using spectral subtraction[J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(11): 1855-1859.

[2] YAN X, YANG Z, WANG T, et al. An iterative graph spectral subtraction method for speech enhancement[J]. Speech Communication, 2020, 123: 35-42.

[3] YANG Y, LIU P, ZHOU H, et al. A speech enhancement algorithm combining spectral subtraction and wavelet transform[C]. 2021 IEEE 4th International Conference on Automation, Electronics and Electrical

Engineering (AUTEEE). IEEE, 2021: 268-273.

[4] TALBI M, BOUHLEL M S. A new speech enhancement technique based on stationary bionic wavelet transform and MMSE estimate of spectral amplitude[J]. Security and Communication Networks, 2021: 1-11.

[5] 李红延,周云龙,田峰,等.一种新的小波自适应阈值函数振动信号去噪算法[J].仪器仪表学报,2015,36(10):2200-2206.
LI H Y, ZHOU Y L, TIAN F, et al. A new wavelet adaptive threshold function vibration signal denoising algorithm[J]. Chinese Journal of Scientific Instrument, 2015, 36(10): 2200-2206.

[6] KHODARAHMI M, MAIHAMI V. A review on Kalman filter models[J]. Archives of Computational Methods in Engineering, 2023, 30(1): 727-747.

[7] JAISWAL R K, YEDURI S R, CENKERAMADDI L R. Single-channel speech enhancement using implicit wiener filter for high-quality speech communication[J]. International Journal of Speech Technology, 2022, 25(3): 745-758.

[8] 潘超,黄公平,陈景东.面向语音通信与交互的麦克风阵列波束形成方法[J].信号处理,2020,36(6):804-815.
PAN CH, HUANG G P, CHEN J D. Microphone array beamforming methods for speech communication and interaction[J]. Journal of Signal Processing, 2020, 36(6): 804-815.

[9] HUANG G, CHEN J, BENESTEY J. A flexible high directivity beamformer with spherical microphone arrays[J]. The Journal of the Acoustical Society of America, 2018, 143(5): 3024-3035.

[10] CHUNG M A, CHOU H C, LIN C W. Sound localization based on acoustic source using multiple microphone array in an indoor environment[J]. Electronics, 2022, 11(6): 890-901.

[11] PRADEEP M N, SURESH M. Speech enhancement using delay and sum beamforming approach[J]. 2020, 3(2):42-54.

[12] BAI M R, KUNG F J. Speech enhancement by denoising and dereverberation using a generalized sidelobe canceller-based multichannel wiener filter[J]. Journal of the Audio Engineering Society, 2022, 70(3): 140-155.

[13] KIM S M. Hearing aid speech enhancement using phase difference-controlled dual-microphone generalized

- sidelobe canceller [J]. IEEE Access, 2019, DOI: 10.1109/ACCESS.2019.2940047.
- [14] GONGPING H, JACOB B, ISRAEL C, et al. A simple theory and new method of differential beamforming with uniform linear microphone arrays [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020, 28: 1079-1093.
- [15] JIN J, BENESTY J, HUANG G, et al. On differential beamforming with nonuniform linear microphone arrays[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2022, 30: 1840-1852.
- [16] SU H, LEE C M. Modified GSC method to reduce the distortion of the enhanced speech signal using cross-correlation and sidelobe neutralization [J]. Applied Sciences, 2021, 11(14): 6288-6304.
- [17] 罗庆予, 张天骐, 方蓉, 等. 联合频谱映射与掩蔽估计的协作式语音增强方法[J]. 电子测量与仪器学报 2023, 37(10): 14-23.
- LUO Q Y, ZHANG T Q, FANG R, et al. A cooperative speech enhancement method combining spectrum mapping and mask estimation [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(10): 14-23.
- [18] PAUL S, THOMAS A, SINGH M S. Delay-and-sum-to-delay-standard-deviation factor: A promising adaptive beamformer [J]. Optics Letters, 2021, 46(18): 4662-4665.
- [19] ZIKSARI M S, ASL B M. Fast beamforming method for plane wave compounding based on beamspace adaptive beamformer and delay-multiply-and-sum[J]. Ultrasound in Medicine & Biology, 2023, 49(5): 1164-1172.
- [20] HUANG G, WANG Y, BENESTY J, et al. Combined differential beamforming with uniform linear microphone arrays [C]. ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021: 781-785.
- [21] WU C, YU P. A study on active noise reduction of automobile engine compartment based on adaptive LMS algorithm[J]. Acoustics Australia, 2020, 48: 431-440.
- [22] KAPOOR J, PATHAK A, RAI M, et al. Speech quality enhancement through noise cancellation using an adaptive algorithm[J]. IAENG International Journal of Computer Science, 2022, 49(3): 653-665.

作者简介



张传营, 2022 年于河北科技大学获得学士学位, 现为中国人民公安大学硕士研究生, 主要研究方向为阵列及语音信号处理。
E-mail: 962409974@qq.com

Zhang Chuanying received his B. Sc. degree from Hebei University of Science and Technology in 2022. Now he is a M. Sc. candidate at People's Public Security University of China. His main research interests include array and speech signal processing.



赵景玉, 2021 年于东北大学获得学士学位, 现为中国人民公安大学硕士研究生, 主要研究方向为阵列信号处理与自动化控制。
E-mail: 1754633076@qq.com

Zhao Jingyu received his B. Sc. degree from Northeastern University in 2021. Now he is a M. Sc. candidate at People's Public Security University of China. His main research interests include array signal processing and automatic control.



刘扬, 2022 年于山东科技大学获得学士学位, 现为中国人民公安大学硕士研究生, 主要研究方向为阵列信号处理与安全防范技术。
E-mail: 1419311746@qq.com

Liu Yang received his B. Sc. degree from Shandong University of Science and Technology in 2022. Now he is a M. Sc. candidate at People's Public Security University of China. His main research interests include array signal processing and security and prevention technology.



卜凡亮 (通信作者), 1993 年于成都电子科技大学获得学士学位, 1996 年于西安交通大学获得硕士学位, 2000 年于西安交通大学获得博士学位, 现为中国人民公安大学安全防范系教授, 主要研究方向为信息与信号处理。
E-mail: bufanliang@sina.com

Bu Fanliang (Corresponding author) received his B. Sc. degree from Chengdu University of Electronic Science and Technology in 1993, M. Sc. degree from Xi'an Jiao tong University in 1996, and Ph. D. degree from Xi'an Jiao tong University in 2000, respectively. Now he is a professor in the Department of Security Prevention at People's Public Security University of China. His main research interests include the signal and information processing.