

DOI: 10.13718/j.cnki.xdzk.2023.10.013

魏丹丹, 吴军, 周翊. 应用于线上会议系统的改进仿射投影符号算法 [J]. 西南大学学报(自然科学版), 2023, 45(10): 119-128.

# 应用于线上会议系统的改进 仿射投影符号算法

魏丹丹<sup>1</sup>, 吴军<sup>1</sup>, 周翊<sup>2</sup>

1. 遵义师范学院 信息工程学院, 贵州 遵义 563002; 2. 重庆邮电大学 通信与信息工程学院, 重庆 400065

**摘要:** 针对线上会议系统中存在的双端讲话干扰和脉冲噪声问题, 提出一种适用于线上会议系统的改进型仿射投影符号算法。该算法首先用无限脉冲响应(Infinite Impulse Response, IIR)滤波平滑的方式计算信号包络来度量信号能量, 通过比较相关的能量估计值来有选择地更新滤波器权重系数以降低双讲情况对系统的干扰, 然后对噪声建模, 并用新的非线性函数类型替代原算法中的 sign 函数对噪声性能进行提升, 防止算法在脉冲噪声影响下发散。仿真结果表明, 与传统仿射投影算法相比, 该算法可以有效克服线上会议系统中的双端对讲干扰和脉冲噪声干扰, 而且在收敛速度、稳态失调等方面依旧保持良好性能。

**关键词:** 线上会议系统; 脉冲噪声; 双端讲话; 仿射投影算法

中图分类号: TN911.72

文献标志码: A

文章编号: 1673-9868(2023)10-0119-10

开放科学(资源服务)标识码(OSID):



## Improved Affine Projection Algorithm and Its Application in Teleconferencing System

WEI Dandan<sup>1</sup>, WU Jun<sup>1</sup>, ZHOU Yi<sup>2</sup>

1. School of Information Engineering, Zunyi Normal University, Zunyi Guizhou 563002, China;

2. School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

**Abstract:** Aiming at the problems of double-talk interference and impulse noise in the application of online conference system, an improved affine projection symbol algorithm for online conference system was proposed. The new algorithm uses the IIR filter smoothing method to calculate the signal envelope as a measure of signal energy, and by comparing the relevant estimates to selectively update the weight coefficients

收稿日期: 2023-01-02

基金项目: 国家自然科学基金项目(62066049); 贵州省普通高等学校青年科技人才成长项目(黔教合 KY 字[2022]015 号); 贵州省科技厅科技支撑计划(黔科合支撑[2023]257).

作者简介: 魏丹丹, 硕士, 讲师, 主要从事语音信号处理研究.

of the filter, so as to reduce the interference to the system. Then the noise was modeled, and the sigmodal function was used to improve the noise performance to prevent the algorithm from diverging. Finally, the convex combination technique was introduced as the algorithm acceleration scheme. The simulation results show that, compared with the traditional affine projection algorithm, the proposed algorithm can effectively overcome the two-terminal talk interference and impulse noise interference, and still maintain good performance in terms of convergence speed and steady-state.

**Key words:** teleconference system; convex combination; two-terminal speech; affine projection algorithm

得益于 5G 通信技术成熟和业务的蓬勃开展, 宽带语音给用户带来了更优质的接近面对面交流的沟通体验. 目前, 用户在全球线上会议软件的注册量或者使用以每年数 10 倍甚至更多倍数的速度增长. 但是, 目前线上会议系统的研究还存在许多亟待解决的实际应用问题, 例如现有会议系统中搭载的自适应滤波器回声消除算法(Acoustic Echo Cancellation, AEC)在脉冲噪声干扰或双端对讲情况下经常存在发散的问题<sup>[1]</sup>, 实时性要求得不到保证<sup>[2]</sup>. 因此, 创建能够克服这一问题并提升用户体验的回声消除算法是当前的主要任务. 回声消除算法在线上会议系统的应用主要是远端用户将语音信号从发送室传到近端接收室, 远端语音信号通过扬声器播放出来, 经过房间的回波路径不停反射后得到回声信号, 该信号同背景噪声一起被话筒拾取得到期望信号. 而回声消除算法不断迭代更新权重系数实现对回声路径建模, 从而估计出回声信号<sup>[3-4]</sup>. 基于最小均方误差准则(Minimum Mean Square Error, MMSE)的归一化最小均方(Normalized Least Mean Square, NLMS)算法<sup>[5]</sup>、迭代最小均方(Recursive Least Mean, RLS)算法<sup>[6]</sup>及各类变形算法由于计算复杂度低, 结构简单等优点被广泛应用到回声消除中, 但是却无法克服输入语音信号高度相关的影响<sup>[7]</sup>. 对此, Wang 等<sup>[8]</sup>提出一种通过多个输入向量更新权重系数的仿射投影(Affine Projection Algorithm, APA)算法, 但是 APA 收敛速度加快是以计算复杂度增加为代价的. 此外, 在非高斯干扰下, APA 算法性能因为“2-范数准则”特征有所下降, 并利用低阶规范对非高斯噪声的鲁棒性<sup>[9-10]</sup>提出了基于经典“1-范数准则”的一种无矩阵求逆的仿射投影符号算法(Affine Projection Sign Algorithm, APSA)<sup>[11]</sup>. 但是步长因子和正则因子控制的跟踪性能无法满足工程实践的需求, 为了同时满足收敛速度快和稳态误差小的跟踪性能要求<sup>[12]</sup>, 一种基于凸组合方式的算法<sup>[13]</sup>被提出, 所谓凸组合<sup>[14]</sup>就是一种利用两个步长参数进行换挡操作的技术. 小步长参数和大步长参数分别控制误差矢量和收敛速度. 凸组合技术最早是应用在 LMS 算法中, 文献[15]首次将凸组合技术引入 APA 算法中, 接着一系列改进凸组合仿射投影算法甚至凸组合 3 个独立滤波器抽头系数的方案被提出<sup>[16-17]</sup>, 从而提高了算法总体性能.

有学者在对线上会议系统中的双端讲话干扰和脉冲噪声问题深入研究的基础上, 提出了一种改进型凸组合仿射投影符号算法. 该算法首先在 AEC 中加入双端检测(Double Talk Detection, DTD)<sup>[18]</sup>, DTD 方案一般有两种形式: 一种是比较远端语音信号和近端麦克风信号的能量, 如著名的广泛应用于线路回波消除(Line Echo Cancellation)中的 Geigel 法<sup>[19]</sup>; 另一种是通过检测远端和近端麦克风信号之间的相关性来判断 DT(Double Talk). 原理在于回波信号与远端语音信号高度相关, 而近端语音与远端语音的相关性却很低. 相关检测可通过计算远端信号与近端信号的互相关矢量来实现. 然后, 对噪声建模, 用新的非线性函数替代原算法中的 Sign 函数对噪声性能进行提升, 防止算法在脉冲噪声影响下发散.

## 1 APA 及改进算法

本文基于图 1 所示线上会议系统的声学回声消除系统框图进行分析<sup>[20]</sup>, 输入语音信号  $x(n)$  经过回波路径滤波器为  $h(n)$  的发射室产生回波  $y(n)$ , 该回波信号与加性噪声  $v(n)$  一起被麦克风拾取构成期望信号  $d(n)$ . 声学回波路径由一个自适应有限脉冲响应(Finite Impulse Response, FIR)滤波器  $w(n)$  建模, 滤波器输出信号和麦克风输入端信号相比较, 得到误差信号, 通过最小均方误差代价函数对自适应

滤波器的抽头系数进行调整, 最终使包含期望信号的目标函数得到最小化值. 自适应滤波器  $w(n)$  与声学回波路径  $h(n)$  的长度  $L$  相同.

### 1.1 仿射投影算法 APA

APA 是一种数据重用算法, 它不仅利用当前的输入矢量进行滤波器系数更新, 而且重复利用过去的  $K-1$  个输入矢量,  $K$  为算法投影阶数. 假定输入噪声  $v$  服从均值为零, 方差为  $\sigma^2$  的高斯分布, APA 的更新公式可由一个受线性约束的最小化表达式推导, 即

$$\min_{w_i} \| w_i - w_{i-1} \|_2^2 \text{ s.t. } e_{i,p} = 0 \tag{1}$$

式(1)中,  $w_i$  是自适应滤波器权重系数, 后验误差矢量  $e_{i,p} = d_i - U_i^T w_i$ ,  $d_i = (d_i, d_{i-1}, \dots, d_{i-K+1})^T$  是期望信号,  $U_i = (u_i, u_{i-1}, \dots, u_{i-K+1})$  是  $K$  阶输入语音信号矩阵. APA 算法中权重系数的第  $i$  次估计值  $w_i$  依据后验误差  $e_{i,p}$  为 0 的约束条件, 把算法从有约束问题变为无约束的最优化问题<sup>[21]</sup>,  $w_i$  可由式(1)解得

$$w_i = w_{i-1} + \mu U_i (U_i^T U_i + \delta I)^{-1} e_i \tag{2}$$

式(2)中,  $e_i = d_i - U_i^T w_{i-1}$  是误差矢量.  $\delta I$  是为了防止矩阵求逆时数据太大而添加的对角矩阵,  $\delta$  是很小的正数. 引入取值范围在  $0 < \mu < 2$  的步长参数使算法收敛. 值得注意的是, 当投影阶数仿射  $K=1$  时, APA 算法等价于 NLMS 算法.

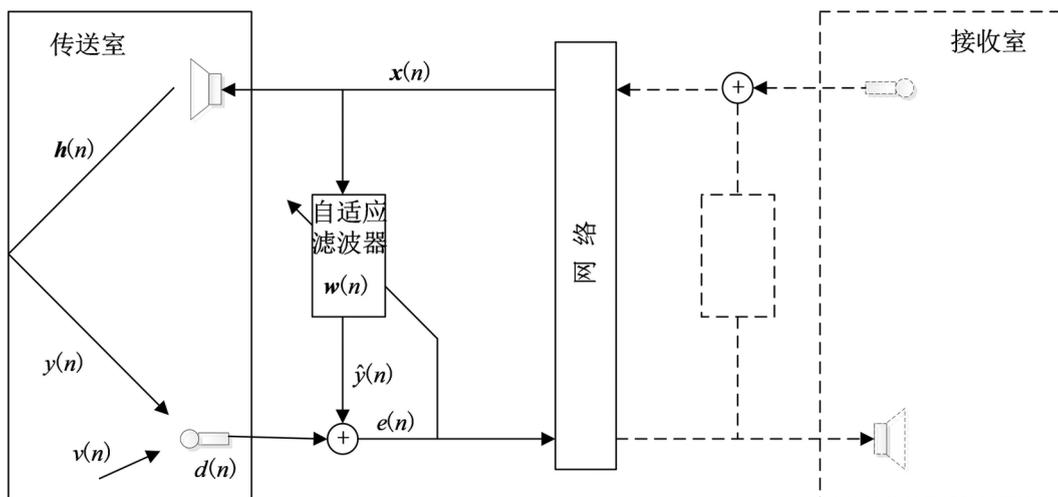


图 1 线上会议系统的 AEC 设置

### 1.2 仿射投影符号算法 APSA

线上会议系统中的输入语音信号具有高度相关性, 传统算法收敛速度明显降低, 而 APA 算法具有降低相关的特性, 输入语音信号严重相关时仍能快速收敛而不引起算法失调. 但是, 算法在外界噪声影响下的鲁棒性能较差, 算法性能严重衰退. 为了克服对自适应滤波算法的不利影响, 文献[11]提出了仿射投影符号算法, 基于  $l_1$  范数的符号函数 APSA 算法在某种程度上可以视为 APA 与符号算法的结合, 算法描述如下所示:

$$\min_{w_i} \| e_{i,p} \|_1 \text{ s.t. } \min_{w_i} \| w_i - w_{i-1} \|_2^2 \leq \mu^2 \tag{3}$$

$$w_i = w_{i-1} + \frac{U_i \text{sign}(e_{i,p})}{2\gamma} \tag{4}$$

$$e_{i,p} \approx e_i \tag{5}$$

$$e_i = d_i - y_i \tag{6}$$

$$y_i = U_i^T w_{i-1} = (y_i, y_{i-1}, \dots, y_{i-K+1})^T \tag{7}$$

式(3)–式(7)中,  $\| \cdot \|_1$  是矢量的  $l_1$  范数,  $\gamma$  是将约束条件改为无约束代价函数时引入的拉格

朗日乘子,  $\text{sign}(\cdot)$  是符号函数. 由于后验误差矢量  $\mathbf{e}_{i,p}$  在实践中无法获得, 将其替代为近似的先验误差矢量  $\mathbf{e}_i$ .

### 1.3 凸组合仿射投影符号算法 CC-APSA

凸组合技术是一种组合因子的迭代方法, 为工程实践提供了很好的解决方案. 与 APSA 固定步长方案不同的是, CC-APSA 算法应用凸组合技术独立运行两个不同步长来同时满足自适应滤波算法对高收敛速度和低稳态误差的需求. 其中一个用大步长  $\mu_1$  来保障算法的快速收敛, 另一个用小步长  $\mu_2$  来保障算法的低稳态误差. 最终 CC-APSA 算法性能优于其中任何一个步长的自适应滤波算法. 算法的整体输出表示为

$$\mathbf{y}_i = \lambda_i \mathbf{y}_{1,i} + (1 - \lambda_i) \mathbf{y}_{2,i} \quad (8)$$

式(8)中,  $\mathbf{y}_{1,i} = \mathbf{u}_i^T \mathbf{w}_{1,i-1}$ ,  $\mathbf{y}_{2,i} = \mathbf{u}_i^T \mathbf{w}_{2,i-1}$ ,  $\mathbf{w}_{1,i-1}$  和  $\mathbf{w}_{2,i-1}$  分别是大步长和小步长的滤波器权重系数. 组合因子  $\lambda_i$  根据整体误差矢量更新, 即  $\mathbf{e}_i = \mathbf{d}_i - \mathbf{y}_i$ , 采用文献[15]中的方法表示为:

$$\lambda_i = \text{sgm}[\alpha_i] = 1 / (1 + e^{-\alpha_i}) \quad (9)$$

式(9)中,  $\alpha_i$  的自适应更新基于最小误差  $l_1$  范数采用梯度下降方案求出, 即

$$\begin{aligned} \alpha_{i+1} &= \alpha_i - \mu_a \frac{\partial |e_i|}{\partial \alpha_i} = \\ &\alpha_i + \mu_a \lambda_i (1 - \lambda_i) (\mathbf{y}_{1,i} - \mathbf{y}_{2,i}) \text{sign}(e_i) \end{aligned} \quad (10)$$

式(10)中,  $\mu_a$  是  $\alpha_i$  的自适应更新步长, 文献[15]建议  $\alpha_i$  应被限定在对称区间  $[-\alpha^+, \alpha^+]$  内, 并指出将其设定为 4 是合适的选择. 因此, 修改式(4)得到 CC-APSA 算法的更新公式为

$$\mathbf{w}_{2,i} = \beta \mathbf{w}_{2,i-1} + (1 - \beta) \mathbf{w}_{1,i-1} + \frac{\mathbf{U}_i \text{sign}(e_i)}{2\gamma} \quad (11)$$

式(11)中,  $\beta$  是光滑因子.

## 2 双端讲话干扰推导

线上会议系统中的双端讲话检测(DTD)是一个难点, 基于信号包络检测的 DTD 是一种低运算量的算法, 该方法通过比较远端语音信号和近端麦克风信号的能量进而作出相应的 DT 判断<sup>[22]</sup>. 这点和 Geigel 方法性质类似. 但是, Geigel 方法比较的是单个数据点的绝对值, 检测函数的变化非常剧烈. 为了得到更准确稳定的判据, 信号保罗检测法采用的是信号能量而非单个数据点. 通过 IIR 滤波平滑的方式来计算信号包络作为信号能量的度量, 然后比较相关的估计值得出 DT 判断. 其中, 输入语音信号能量为:

$$v_x(n) = \alpha v_x(n-1) + (1 - \alpha) |x(n)| \quad (12)$$

式(12)中,  $\alpha = 0.99$  是遗忘因子, 麦克风信号能量为:

$$v_d(n) = \alpha v_d(n-1) + (1 - \alpha) |d(n)| \quad (13)$$

检测函数因此可以计为

$$\xi(n) = \frac{v_d(n)}{v_x(n-1) + \gamma} \quad (14)$$

式(14)中, 参数  $\gamma$  在只有背景噪声时用来限制检测函数的取值, 一般设为 0.05. 通常情况下, 回波能量相对于远端语音信号能量较弱, 且近端语音信号能量比回波能量高. 因此, 检测函数的值会随着 DT 出现而显著增大, 并随着 DT 结束而减小. 将检测函数值和一个阈值  $T$  相比较, 就可以做出相应的判断. 由于语音信号的能量具有时变特性, 固定阈值  $T$  方案不是好的选择, 必须设计一个能跟踪信号能量变化的时变阈值  $T(n)$ . 当自适应滤波器收敛后, 其输出  $y(n)$  可以近似地认为等同于真实的回波, 因此可以用来作为时变阈值的参考, 于是有:

$$T(n) = \frac{v_y(n)}{v_x(n) + \gamma} + \kappa \quad (15)$$

式(15)中,  $\kappa = 0.02$  是一个小的正数, 用来为检测误差留下一点余地. 比较两者的值可以给出 DT 判

断. 其中,  $v_x(n)$  的作用可看作是为了增加算法准确度引入的一个归一化处理.

### 3 一种新的凸组合仿射投影符号算法 NCC-APSA

本节提出一种新的凸组合仿射投影符号算法 NCC-APSA (Novel Convex Combination Affine Projection Sign Algorithm), 新算法用新的非线性 Sigmoid 函数类型替代原算法中的 sign 函数, 在第 2 部分双端讲话干扰推导基础上定义新的 DT 统计量  $g(n)$ , 有效克服线上会议系统中的双端对讲干扰和脉冲噪声干扰, 而且在收敛速度、稳态失调等方面依旧保持良好性能. 根据滤波器线性约束最小化表达式改写为代价函数形式为:

$$\mathbf{J}(\mathbf{w}_i) = \|\mathbf{e}_{i,p}\|_1 + \gamma(\|\mathbf{w}_i - \mathbf{w}_{i-1}\|_2^2 - \mu^2) \quad (16)$$

利用负梯度最陡下降法, 对式(16)两边关于  $\mathbf{w}_i$  求导:

$$\frac{\partial \mathbf{J}(\mathbf{w}_i)}{\partial \mathbf{w}_i} = -\mathbf{U} \text{sign}(\mathbf{e}_{i,p}) + 2\gamma(\mathbf{w}_i - \mathbf{w}_{i-1}) \quad (17)$$

令

$$-\mathbf{U} \text{sign}(\mathbf{e}_{i,p}) + 2\gamma(\mathbf{w}_i - \mathbf{w}_{i-1}) = 0$$

代入约束条件和基于  $l_1$  范数的凸组合技术, 然后代入式(10)组合因子推导结果, 更新公式为:

$$\mathbf{w}_{2,i} = \beta \mathbf{w}_{2,i-1} + (1 - \beta) \mathbf{w}_{1,i-1} + \frac{\mu_2 \mathbf{U}_i \text{sign}(\mathbf{e}_i)}{\sqrt{\text{sign}(\mathbf{e}_i^T) \mathbf{U}_i^T \mathbf{U}_i \text{sign}(\mathbf{e}_i)}} \quad (18)$$

为了保证算法在脉冲噪声下的鲁棒性, 将式(18)改进为更加平滑的非线性函数 sigmoid 形式, 在算法实施时, 为了克服双端对讲干扰, 本文做了一点处理. 定义新的 DT 统计量  $g(n)$ , 采用  $\zeta(n) - T(n)$  作为检测函数, 通过试验设计固定阈值和控制环节来实现更稳定的 DTD, 改进后的权重系数更新为:

$$\mathbf{w}_{2,i} = \beta \mathbf{w}_{2,i-1} + (1 - \beta) \mathbf{w}_{1,i-1} + \frac{\mu_2 \mathbf{U}_i \text{sigm}(\mathbf{e}_i) g(n)}{\sqrt{\text{sigm}(\mathbf{e}_i^T) \mathbf{U}_i^T \mathbf{U}_i \text{sigm}(\mathbf{e}_i)}} \quad (19)$$

式(19)中,  $g(n)$  由如下表达式获得:

$$g(n) = \begin{cases} 1 & \text{if threshold} > \{\zeta(n) - T(n)\} \\ 0 & \end{cases} \quad (20)$$

当单端讲话时, 近端语音能量值非常小, 即 DT 统计量  $g(n) = 1$ . 在算法迭代初期, 误差信号能量值较大, 而滤波器输出信号能量和麦克风信号能量值较小, 算法以大步长参数更新, 收敛速度快的同时伴随着稳态误差不理想. 待算法迭代趋于稳定后, 误差信号能量值变小, 而滤波器输出信号能量和麦克风信号能量值逐渐增大, 算法以小步长因子更新, 降低稳态误差. 随着迭代次数不断增加, 算法收敛程度增加, 麦克风信号能量和滤波器输出信号能量的差值始终小于误差信号能量, 由此可以保证式(19)推导的改进型仿射投影符号算法的收敛性. 当发生双端对讲时, 近端语音能量快速增加到一个较大的值, 即 DT 统计量  $g(n) = 0$ , 此时更新过程近乎停止, 从而防止了算法发散. 当脉冲噪声干扰时由式(19)可知, 会导致误差信号大幅度增加, 从而造成  $\mu_2 \mathbf{U}_i \text{sigm}(\mathbf{e}_i) g(n) / \sqrt{\text{sigm}(\mathbf{e}_i^T) \mathbf{U}_i^T \mathbf{U}_i \text{sigm}(\mathbf{e}_i)} \approx 0$ , 权系数的增加量接近于零, 算法的迭代几乎停止, 因此基本消除了脉冲噪声对算法性能的影响, 而回声抑制能力不受影响. 步长衡量因子满足文献[13]中提出的抗脉冲干扰条件. 步长调节因子在脉冲噪声发生时自动调节步长, 算法性能趋于均方误差自适应算法, 消除了权重矢量由于大幅度误差信号引起更新的可能. 以上分析表明, 改进算法 NCC-APSA 在脉冲噪声和双端对讲的三重干扰下仍能保持算法收敛, 表现出了复杂环境下良好的适应性.

### 4 仿真结果和分析

本文基于 MATLAB 仿真平台对算法性能进行模拟, 第 1 部分采用不同类型输入信号和系统脉冲响应分别分析 APA 算法、APSA 算法、CC-APSA 算法及 NCC-APSA 算法在收敛速度、均方误差和跟踪能力等方面

的性能. 第 2 部分则验证改进型凸组合仿射投影符号算法在有双端讲话干扰的线上会议系统中的 AEC 性能.

#### 4.1 算法性能对比

本节采用两种不同类型的远端输入信号进行仿真, 8 kHz 采样频率, 16 bit 量化. 输入信号如图 2 是由零均值高斯随机信号通过 AR 模型  $T(z)=1/(1-0.7z)$  产生, 信噪比  $SNR=30$  dB. 图 3 是仿真中系统辨识采用的两种不同单位脉冲响应, 左边是截断长度为 256 阶的汽车单位脉冲响应, 右边是会议系统中常见的 1 024 阶单位脉冲响应. 两种脉冲都不属于典型的稀疏信道, 稀疏度分别为 0.52 和 0.32. 衡量算法性能指标为归一化均方偏差(Normalized Mean Square Deviation, NMSD)<sup>[23]</sup>, 用于度量自适应滤波器与目标系统的逼近程度, 即回声抑制能力, 其定义如下:

$$NMSD = 10 \log_{10} \left( \frac{\| \mathbf{w}_i - \mathbf{w}_o \|_2^2}{\| \mathbf{w}_o \|_2^2} \right) \quad (21)$$

归一化均方偏差包括稳态误差和收敛速度两个跟踪性能, 稳态误差越小, 代表回声抑制效果越好, 而收敛速度代表算法需要多少次迭代才能达到理想的回声抑制效果. 本文所有仿真曲线均是通过 50 次独立实验求平均值获得.

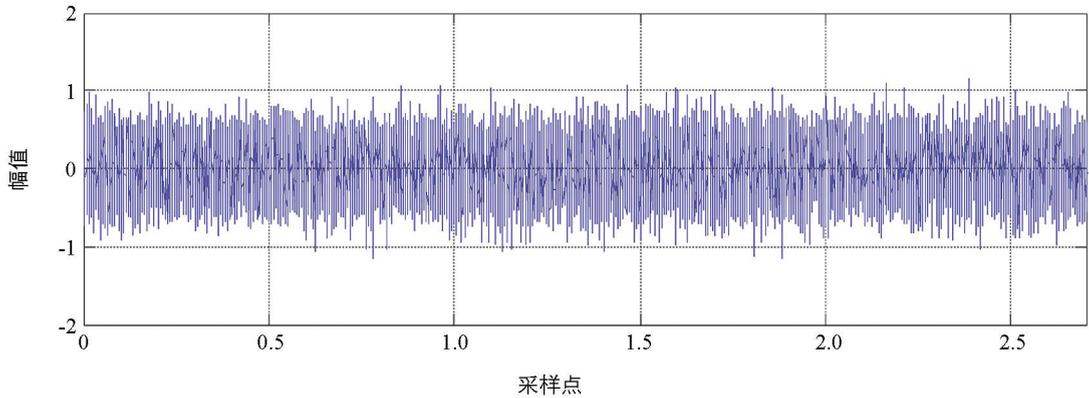


图 2 输入高斯白噪声信号

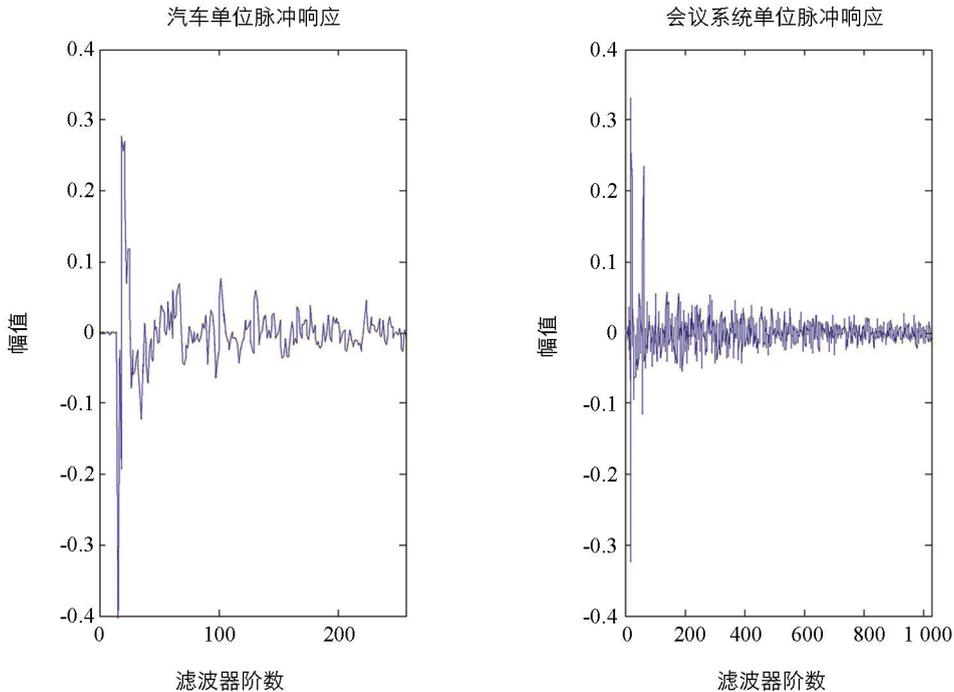


图 3 系统辨识

##### 4.1.1 稳态输入信号的算法性能对比

输入图 2 中自回归模型下的高斯白噪声, 期望信号由图 3 的真实单位脉冲响应产生, 无输出噪声.

作为第一步, 本文仿真了 APA 算法、APSA 算法、CC-APSA 算法和 NCC-APSA 算法在自回归模型下的跟踪性能, 图 4 和图 5 分别展示了  $L=256$  和  $L=1024$  两种系统辨识的情况. 实验结果表明, 所提出的 NCC-APSA 算法具有优秀的收敛速度和稳态误差, 当滤波器大小  $L=256$  时, CC-APSA 算法和 NCC-APSA 算法无论是收敛速度还是稳态误差都好于 APA 算法、APSA 算法. CC-APSA 算法和 NCC-APSA 算法收敛速度相似, 有几乎相同的稳态 NMSD, 都是大约 900 次迭代后开始收敛. 当  $L=1024$  时, 各算法性能均有不同程度的下降, 算法无法达到 900 次迭代的稳态, 但是可以看出 CC-APSA 算法跟踪性能略微优于 NCC-APSA 算法.

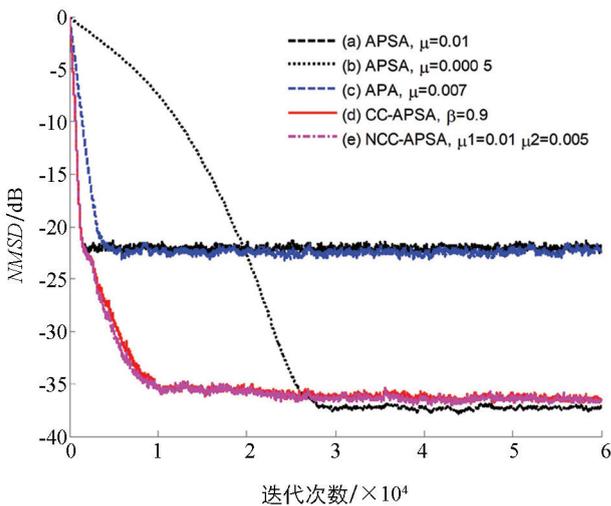


图 4 相关自适应算法性能比较(汽车脉冲响应)

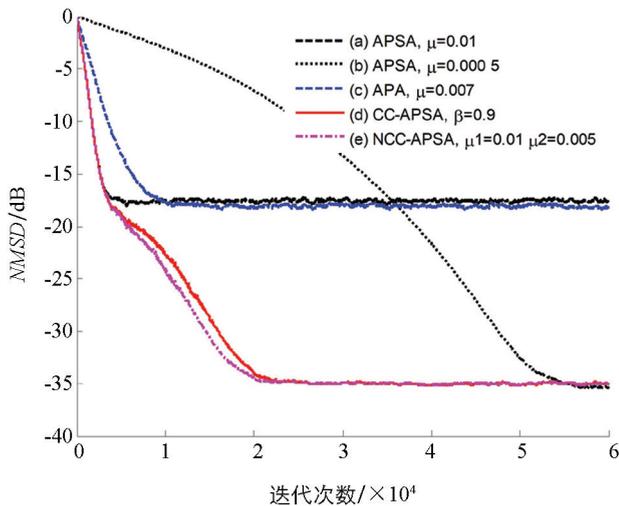


图 5 相关自适应算法性能比较(会议系统脉冲响应)

### 4.1.2 脉冲噪声模型下的算法性能对比

图 6 和图 7 是算法在期望信号中加入两种不同类型脉冲噪声后的跟踪性能曲线, 利用混合高斯模型对脉冲噪声进行仿真<sup>[24]</sup>. 模型为  $k_i A_i$ ,  $k_i$  为伯努利过程  $P[k_i=1]=praba$ , 其中  $praba$  表示该过程中冲激噪声的出现概率.  $A_i$  为零均值, 功率为  $\sigma_A^2=1000\sigma_y^2$  的高斯随机过程, 从图 6 可以看出, NCC-APSA 算法在冲激噪声逐步增大的情况下, 虽然误差增大, 但是仍然保持了稳定性, 因而验证了 NCC-APSA 算法对一般冲激噪声具有很好的通用性和鲁棒性. 图 7 中冲激响应在第 50000 次迭代时从  $w$  突变为  $-w$ , 本文主要观测算法的收敛和跟踪性能, 在冲激噪声下 NCC-APSA 算法不仅收敛, 而且在未知系统发生截断误差后算法很快跟上了系统变化, 可见系统突变对 NCC-APSA 算法未损失收敛速度和稳态误差性能.

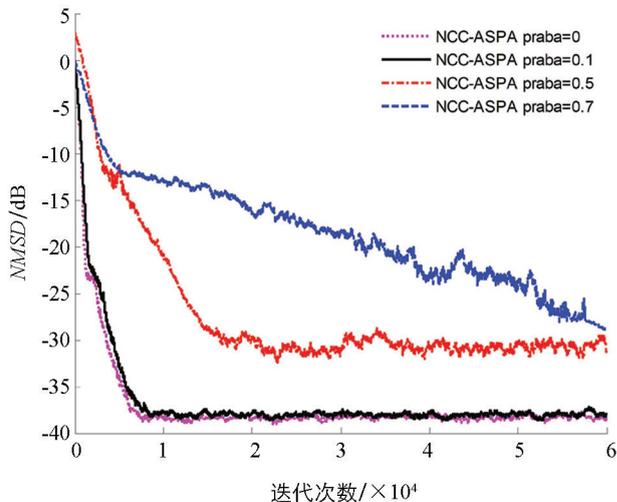


图 6 不同脉冲噪声水平下算法跟踪性能

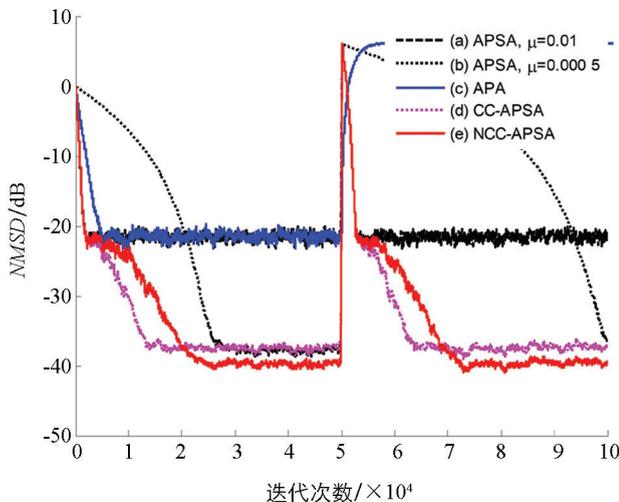


图 7 相关自适应算法待辨识系统突变的性能比较

### 4.1.3 语音输入信号的算法性能对比

语音由于具有非平稳特性和频谱范围大的特点而非理想的脉冲响应激励信号<sup>[25]</sup>, 在实际应用场景中, 输入信号往往是语音信号而非高斯白噪声, 所以本文采用真实语音信号进行仿真实验, 材料取自一段中文男声, 具有 10~30 ms 的准稳态特性及高相关性. 本文对 NCC-APSA 算法使用语音信号作为输入信号进行了对比实验, 分别采取图 4 所示的两个单位回声脉冲响应进行滤波. 对于这两个实验, 本文都添加一个信噪比为 50 dB 的高斯白噪声成分. 图 8 和图 9 分别显示了长度为  $L=256$  的汽车系统和长度为  $L=1\ 024$  的线上会议系统的结果. 与 APA、APSA 和 CC-APSA 算法相比, NCC-APSA 算法在暂态和稳态状态下都具有良好的 AEC 性能.

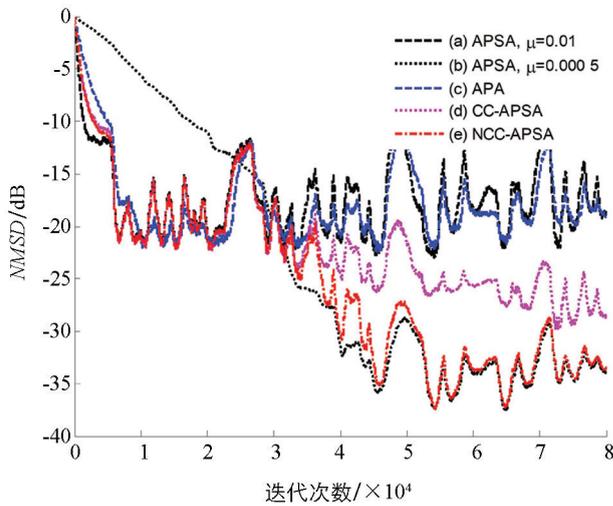


图 8 语音输入下相关自适应算法  
跟踪性能(汽车脉冲响应)

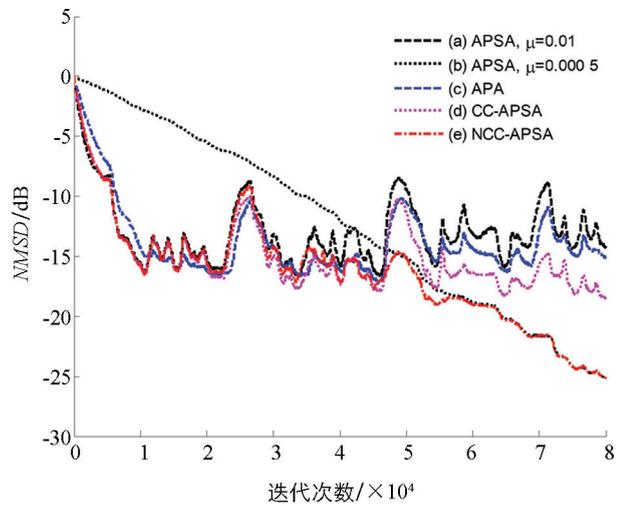


图 9 语音输入下相关自适应算法  
跟踪性能(会议系统脉冲响应)

## 4.2 基于线上会议系统的算法验证实验

为了进一步验证 NCC-APSA 算法的性能, 本文采用真实语音信号进行线上会议系统的回声消除实验, 分别在两个不同房间进行音频数据采集. 两个房间分别配置了不同的普通家具家电, 发送室大小约为  $4.35\text{ m} \times 3.2\text{ m} \times 2.8\text{ m}$  ( $L \times W \times H$ ), 接收室大小约为  $3.8\text{ m} \times 2.6\text{ m} \times 2.5\text{ m}$  ( $L \times W \times H$ ). 用于语料采集的器材包括电脑、麦克风和扬声器等, 扬声器和麦克风分别摆放在电脑两侧, 且两者距离 0.45 m 左右. 电脑摆放在中间负责数据处理, 麦克风录制远端语音信号, 远端信号经过房间的回声信道形成回声后被麦克风拾取, 同时被拾取的还有近端信号和环境噪声(图 10).

近端讲话的语音内容是一段中文男声“园林采用新中式建造手法……”, 语音数据的采样率设置为 8 000 Hz, 根据房间实际脉冲来选择截断长度, 确定自适应 FIR 滤波器阶数  $L=1\ 024$  阶. NCC-APSA 算法处理结果如图 11 所示, 左边是近端语音信号(长度设置为 400 000)和双端讲话语音的时域波形图, 语音信号在迭代次数为 156 800 时加入双讲语音, 其余设置为 0, 无双端对讲影响. 图 11 右侧从上往下分别是麦克风信号和回声消除后的输出信号时域波形图. 从图 11 中可以看出, NCC-APSA 算法可以有效地滤除回声干扰噪声, 听感方面的回声抑制效果也很好, 工程实践中具有很高的实用价值.

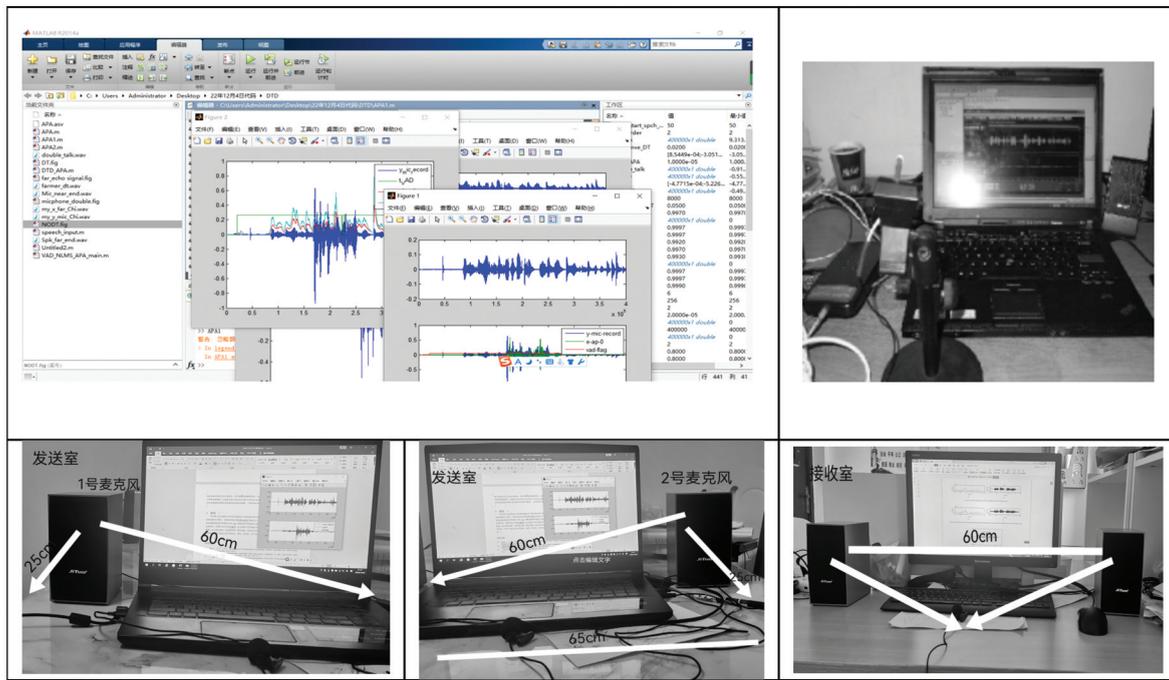


图 10 单通道 AEC 软、硬件实验设置

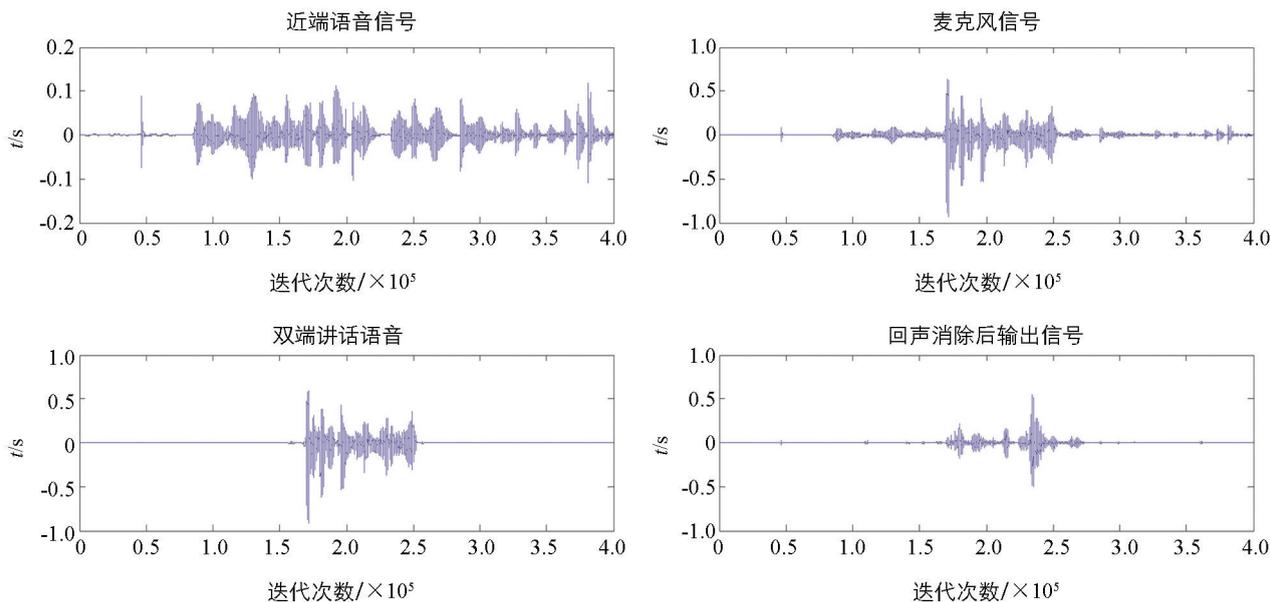


图 11 信号时域波形图

### 5 结论

针对线上会议系统中存在的双端讲话干扰和脉冲噪声问题, 本文提出一种适用于线上会议系统的改进型仿射投影符号算法. 仿真结果表明, 该算法可以有效克服线上会议系统中的双端对讲干扰和脉冲噪声干扰, 而且依旧保持良好的跟踪性能. 尽管本文算法在一定程度上改善了算法性能, 但是没有考虑稀疏信道下的性能优化, 而且计算复杂度依然较高, 不利于实际应用中推广. 因此, 如何利用信道的稀疏先验特性, 进一步开发适应稀疏信道的低复杂度算法是下一步研究的重点工作.

### 参考文献:

[1] DUTTWEILER D. Proportionate Normalized Least-Mean-Squares Adaptation in Echo Cancelers [J]. IEEE/ACM

- Transactions on Audio Speech and Language Processing, 2000, 8(5): 508-518.
- [2] HANSLER E, SCHMIDT G. Acoustic Echo and Noise Control: A Practical Approach [M]. Hoboken: John Wiley & Sons, 2004.
- [3] NI J G, LI F. Adaptive Combination of Subband Adaptive Filters for Acoustic Echo Cancellation [J]. IEEE Transactions on Consumer Electronics, 2010, 56(3): 1549-1555.
- [4] 王文博, 姚英彪, 刘兆霆. 二值传感器网络的分布式稀疏 LMS 算法 [J]. 信号处理, 2019, 35(1): 86-92.
- [5] HASSANI I, AREZKI M, BENALLAL A. A Novel Set Membership Fast NLMS Algorithm for Acoustic Echo Cancellation [J]. Applied Acoustics, 2020, 163(C): 107210.
- [6] 赵益波, 严涛, 李春彪, 等.  $\alpha$ -稳定分布噪声环境下的非线性回声消除研究 [J]. 电子学报, 2020, 48(1): 59-65.
- [7] 吉利鹏, 倪锦根. 偏差补偿符号子带自适应滤波器 [J]. 电子学报, 2020, 48(11): 2220-2225.
- [8] WANG B, LI H, WU C, et al. Variable Step-Size Correntropy-Based Affine Projection Algorithm with Compound Inverse Proportional Function for Sparse System Identification [J]. IEEE Transactions on Electrical and Electronic Engineering, 2022, 17(3): 416-424.
- [9] LIU W F, POKHAREL P. Principle Correntropy: Properties, and Applications in Non-Gaussian Signal Processing [J]. IEEE Transactions Signal Process, 2007, 55(11): 5286-5298.
- [10] AL-SAYED S, ZOUBIR A M, SAYED A H. Robust Adaptation in Impulsive Noise [J]. IEEE Transactions on Signal Processing, 2016, 64(11): 2851-2865.
- [11] SHAO T G, Zheng Y R, Benesty J. An Affine Projection Sign Algorithm Robust Against Impulsive Interferences [J]. IEEE Signal Processing Letters, 2010, 17(4): 327-330.
- [12] ZHOU Y, LIU H, CHAN S C. New Partial Update Robust Kernel Least Mean Square Adaptive Filtering Algorithm [C] // 2014 19th International Conference on Digital Signal Processing. Hong Kong: IEEE, 2014.
- [13] 郭莹, 于和芳, 赵璐, 等. 非高斯噪声下基于 Wilcoxon 范数的变步长符号扩散式仿射投影算法 [J]. 电子与信息学报, 2021, 43(2): 303-309.
- [14] 黄梓桐, 阿里甫·库尔班, 韩文轩, 等. 基于最大相关熵的簇稀疏仿射投影算法 [J]. 新疆大学学报(自然科学版), 2020, 37(2): 177-182.
- [15] CHAN S C, ZHOU Y, HO K L. A New Sequential Block Partial Update Normalized Least Mean M-Estimate Algorithm and its Convergence Performance Analysis [J]. Journal of Signal Processing Systems, 2010, 58(2): 173-191.
- [16] LOGANATHAN P, HABETS E AP, NAYLOR P A. A Proportionate Adaptive Algorithm with Variable Partitioned Block Length for Acoustic echo Cancellation [C] // Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Prague: IEEE, 2011.
- [17] CHOI J H, CHO H, JEONG J J, et al. Combination of Step Sizes For Affine Projection Algorithm with Variable Mixing Parameter [J]. Electronics Letters, 2013, 49(18): 1149-1150.
- [18] ALDAJANI M A. Adaptive Step-Size Sign Least Mean Squares [C] // 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Montreal: IEEE Press, 2004.
- [19] WEI Y, ZHANG Y G, WANG C C. Block-Sparsity-Aware LMS Algorithm for Network Echo Cancellation [J]. Electronics Letters, 2018, 54(15): 951-953.
- [20] JIANG S, GU Y. Block-Sparsity-Induced Adaptive Filter for Multi-Clustering System Identification [J]. IEEE Transactions on Signal Processing, 2014, 63(20): 5318-5330.
- [21] 魏丹丹, 周翊, 师黎明, 等. 基于反双曲正弦函数的抗冲激块稀疏自适应滤波算法 [J]. 计算机应用, 2017, 37(1): 197-199.
- [22] Iqbal M A, Stokes J W, Grant S L. Normalized Double-Talk Detection Based on Microphone and AEC Error Cross-Correlation [C] // IEEE International Conference on Multimedia & Expo. Beijing: IEEE, 2007.
- [23] ROTILI R, CIFANI S, PRINCIPI E, et al. A Robust Iterative Inverse Filtering Approach for Speech Dereverberation in Presence of Disturbances [C] // IEEE Asia Pacific Conference on Circuits & Systems. Macao: IEEE, 2008.
- [24] PRINCIPI E, VESPERINI F, SQUARTINI S, et al. Acoustic Novelty Detection with Adversarial Autoencoders [C] // International Joint Conference on Neural Networks. Copenhagen: IEEE, 2017.
- [25] COHEN I. Optimal Speech Enhancement Under Signal Presence Uncertainty Using Log-Spectral Amplitude Estimator [J]. IEEE Signal Processing Letters, 2002, 9(4): 113-116.