

## 基于 Kalman 滤波的 GSC 改进语音增强算法

郭业才<sup>1,2</sup>, 许雪<sup>1</sup>, 刘力玮<sup>1</sup>

(1. 南京信息工程大学电子与信息工程学院, 南京 210044; 2. 南京信息工程大学滨江学院, 无锡 214105)

**摘要:** 针对广义旁瓣相消器 (Generalized sidelobe canceller, GSC) 存在非相干噪声消除性能不佳的缺陷, 提出了采用后置 Kalman 滤波器改进的 GSC 去噪算法。该算法通过归一化最小均方算法校正自适应噪声对消器, 并将滤除方向性干扰噪声后的语音信号输出到 Kalman 滤波器中, 对残余背景噪声进行迭代最小均方误差 (Minimum mean square error, MMSE) 估计, 抑制非相干噪声与麦克风阵元所产生的热噪声。经过在不同信噪比条件下客观语音质量评估 (Perceptual evaluation of speech quality, PESQ) 及语谱图分析后证明, 与传统的 GSC 以及后置谱减法的改进 GSC 相比, 本算法在噪声消除上的表现更为优越, 且增强后信号也更接近目标信号。

**关键词:** 语音增强; 波束形成; 广义旁瓣相消; Kalman 滤波

**中图分类号:** TN911      **文献标志码:** A

### Improved GSC Speech Enhancement Algorithm Based on Kalman Filtering

GUO Yecai<sup>1,2</sup>, XU Xue<sup>1</sup>, LIU Liwei<sup>1</sup>

(1. School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. Binjiang College, Nanjing University of Information Science and Technology, Wuxi 214105, China)

**Abstract:** In view of the poor performance of generalized sidelobe canceller (GSC) of incoherent noise cancellation, an improved GSC denoising algorithm with post-Kalman filter is proposed. The algorithm corrects the adaptive noise canceller through the normalized least mean square algorithm, outputs the speech signal after filtering the directional interference noise to the Kalman filter, and iterates the residual background noise with the minimum mean square error (MMSE) to suppress incoherent noise and thermal noise generated by microphone array elements. After the objective speech quality evaluation, perceptual evaluation of speech quality (PESQ), and spectrogram analysis under different signal-to-noise ratio conditions, it is proved that compared with the traditional GSC and the improved GSC of post-spectral subtraction, this algorithm is better at noise elimination. The performance is better, and the enhanced signal is closer to the target signal.

**Key words:** speech enhancement; beamforming; generalized sidelobe cancellation; Kalman filtering

### 引言

语音增强的主要目的是提高环境中各种噪声干扰下的语音质量与清晰度<sup>[1]</sup>。过去几十年里,

多种单麦克风语音增强方法,如谱减法(Spectral subtraction, SS)<sup>[2]</sup>、Kalman滤波<sup>[3]</sup>等被提出。与单麦克风受方向性噪声影响较大相比,麦克风阵列可利用语音的空间信息来增强目标信号,滤除干扰信号,提高语音的清晰度,因此麦克风阵列语音增强技术近些年来逐渐成为语音增强领域研究热门<sup>[4]</sup>。

1982年由Griffiths等提出的广义旁瓣相消结构<sup>[5]</sup>是线性约束最小方差(Linearly constrained minimum variance, LCMV)波束形成器的等效实现结构,它将LCMV的约束优化问题转化为无约束的优化问题,并且只需要较少的麦克风就能达到较强的噪声抑制性能,因此成为一种主流的语音增强算法。但广义旁瓣相消器(Generalized sidelobe canceller, GSC)同样存在过度抵消、非相干噪声抑制性能不佳等缺陷,学者们已经提出了一些方法来对GSC的性能进行改进<sup>[6-11]</sup>。Zelinski<sup>[12]</sup>设计了一种可以同时抑制非相干噪声以及相干噪声的维纳滤波多通道后置滤波器;Chang等<sup>[13]</sup>通过对GSC输出误差进行估计来估算时变波达方向(Direction of arrival, DOA),可适用于目标源处于移动状态情况;Yu等<sup>[14]</sup>利用SS对GSC输出后的信号进行滤波处理,有效改善了GSC非相干噪声抑制能力差的缺陷,但SS本身的固有缺陷也会产生“音乐噪声”。

本文在前人研究的基础上,提出了一种利用Kalman滤波改进的广义旁瓣相消算法。该算法首先通过归一化最小均方算法改进的GSC波束形成方法对接收到的麦克风阵列信号进行方向性干扰噪声的抑制,输出后的语音信号再利用Kalman滤波器将剩余背景噪声消除以更好地改善语音质量。

## 1 GSC改进算法

广义旁瓣相消波束形成算法是一种经典的麦克风阵列自适应波束形成语音增强方法,该算法主要由3部分组成:固定波束形成(Fixed beamforming, FB)模块、阻塞矩阵(Blocking matrix, BM)模块以及自适应噪声相消(Adaptive noise canceller, ANC)模块。FB模块对时间进行延迟估计后补偿声源到达每个麦克风阵元的延迟时间,并对接收信号校正延迟形成同步,累加生成初级波束信号。BM将阵列信号中的目标语音过滤后生成一个不包含目标语音的带噪参考信号。最后通过归一化最小均方方法进行ANC,噪声一般通过FB输出信号获得的参考噪声进行估计。GSC系统的结构,如图1所示。

假设麦克风阵列有 $M$ 个阵元,首先通过FB模块对每个信道的信号进行过滤,然后将过滤后的信号进行相加,输出为

$$\mathbf{y}_{\text{FBF}} = \mathbf{A}^T \mathbf{X} \quad (1)$$

式中, $\mathbf{X}$ 为 $M$ 个麦克风阵元上接收到的经延迟补偿后具有完整相位的语音信号, $\mathbf{X} = [\mathbf{x}_1(n), \mathbf{x}_2(n), \dots, \mathbf{x}_M(n)]^T$ ;  $\mathbf{A}$ 为FB模块的权系数, $\mathbf{A} = [a_0, a_1, \dots, a_M]^T$ 。

$\mathbf{B}$ 为BM,它将 $\mathbf{X}$ 中的期望信号滤除,只包含干扰信号与噪声。BM输出信号为

$$\mathbf{Z} = \mathbf{B}\mathbf{X} \quad (2)$$

式中 $\mathbf{Z} = [\mathbf{z}_1(n), \mathbf{z}_2(n), \dots, \mathbf{z}_{M-1}(n)]^T$ 。设 $\mathbf{b}_M^T$ 是 $\mathbf{B}$ 中的第一个线性向量,需满足条件

$$\mathbf{b}_M^T \mathbf{I} = 0 \quad (3)$$

在语音信号处理中,BM为

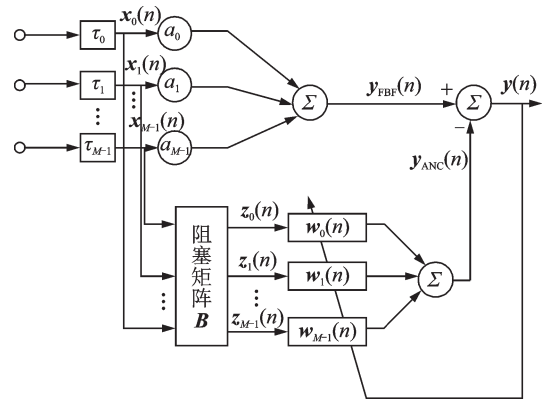


图1 GSC结构图

Fig.1 Structure schematic of GSC

$$B = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix} \quad (4)$$

自适应相消模块将FB输出信号与BM输出信号进行自适应滤波。设自适应FIR滤波器系数表示为 $\mathbf{w}_k(n)$ ,与BM输出噪声相乘后为

$$\mathbf{y}_{\text{ANC}}(n) = \sum_{k=1}^{m-1} \mathbf{w}_k(n) \mathbf{z}_k(n) = \mathbf{W}_K^T \mathbf{Z}(n) \quad (5)$$

式中, $\mathbf{W}_K = [\mathbf{w}_1(n), \mathbf{w}_2(n), \dots, \mathbf{w}_{M-1}(n)]^T (k=1, 2, \dots, m-1)$ ,将初级波束信号 $\mathbf{y}_{\text{FBF}}$ 与 $\mathbf{y}_{\text{ANC}}(n)$ 进行相减后,GSC最终输出为

$$\mathbf{y}(n) = \mathbf{y}_{\text{FBF}}(n) - \mathbf{y}_{\text{ANC}}(n) \quad (6)$$

滤波器权系数的递归公式为

$$\mathbf{W}_K(n+1) = \mathbf{W}_K(n) - \mu \nabla(n) \quad (7)$$

式中, $\mu$ 是自适应收敛系数,决定了算法的收敛速率和收敛稳定性。 $\nabla(n) = \nabla E[\mathbf{y}(n)^2]$ 是难以精确计算的梯度算子,可以采用 $\mathbf{y}(n)^2$ 作为 $E[\mathbf{y}(n)^2]$ 的值估计为

$$\hat{\nabla}(n) = \nabla[\mathbf{y}(n)^2] = 2\mathbf{y}(n)\nabla[\mathbf{y}(n)] \quad (8)$$

$$\nabla[\mathbf{y}(n)] = \nabla[\mathbf{y}_{\text{FBF}}(n) - \mathbf{W}^T(n)\mathbf{Z}(n)] = -\mathbf{Z}(n) \quad (9)$$

将式(9)代入(7),得

$$\mathbf{W}_K(n+1) = \mathbf{W}_K(n) + 2\mu\mathbf{y}(n)\mathbf{Z}(n) \quad (10)$$

式中, $\mu$ 的取值会影响整个GSC系统的降噪能力,其取值越大,滤除噪声越多,但同时也会使语音泄露更严重。现采用归一化最小均方算法来校正算法缺陷,即

$$\mathbf{W}_k(n+1) = \mathbf{W}_k(n) + \frac{\varphi}{\|\mathbf{Z}(n)\|^2} \mathbf{y}(n)\mathbf{Z}(n) \quad (11)$$

式中 $\varphi$ 为调节因子。由于 $\|\mathbf{Z}(n)\|^2$ 处理后非常小,导致 $\varphi/\|\mathbf{Z}(n)\|^2$ 参数太大,进而导致语音信号抵消过于严重。因此引入参数 $\beta > 0$ 来解决这个问题。式(11)改写为

$$\mathbf{W}_k(n+1) = \mathbf{W}_k(n) + \frac{\varphi}{\|\mathbf{Z}(n)\|^2 + \beta} \mathbf{y}(n)\mathbf{Z}(n) \quad (12)$$

## 2 基于Kalman滤波的GSC方法

基于GSC波束形成语音增强算法对方向性干扰噪声具有很好的抑制效果,但在非相干噪声的处理上效果很差,仅依靠FB模块。本文采用Kalman滤波来改善GSC输出后的语音质量。算法整体流程如图2所示。

Kalman滤波器采用递归估计,利用上一时刻状态的

估计值与当前状态的观测值来计算出当前时刻状态的估计值,其计算量和数据存储量都比较小<sup>[15]</sup>。

假设GSC系统的输出信号为

$$\mathbf{y}(n) = \mathbf{x}(n) + \mathbf{v}(n) \quad (13)$$

式中: $\mathbf{x}(n)$ 与 $\mathbf{v}(n)$ 分别表示目标信号和残余噪声,假设两者之间不相关。

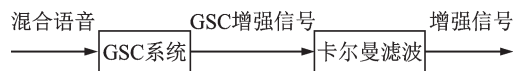


图2 算法流程图

Fig.2 Flow chart of algorithm

语音信号具有非平稳性,因此通常使用短时分析来对其进行处理。即在短时帧(20~30 ms内)内假设语音信号是平稳的,用 $p$ 阶自回归(Autoregressive, AR)模型表示为

$$x(n) = \sum_{i=1}^p a_i x(n-i) + g(n) \tag{14}$$

式中: $p$ 为模型阶数, $\{a_i\}$ 为线性预测系数<sup>[16]</sup>。

令 $X(n)=[x(n-p+1), x(n-p+2), \dots, x(n)]^T$ ,在状态空间中式(13,14)分别转化为状态方程与测量方程,即

$$y(n) = HX(n) + v(n) \tag{15}$$

$$X(n) = AX(n-1) + Gg(n) \tag{16}$$

式中

$$H = G^T = [0, 0, \dots, 0, 1]_{1 \times p} \tag{17}$$

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 1 \\ a_p & a_{p-1} & \dots & \dots & a_1 \end{bmatrix}_{p \times p} \tag{18}$$

式中: $g(n)$ 为过程噪声,这里假设 $g(n)$ 与 $v(n)$ 是均值为0、方差分别为 $\sigma_g^2$ 与 $\sigma_v^2$ 的高斯白噪声,并且两者互不相关; $A$ 是状态转移矩阵。现用Kalman滤波器的循环迭代对每帧带噪语音进行估计:

(1) 初始化

$$\hat{X}(0|0) = 0, P(0|0) = 0$$

(2) 迭代过程:对于 $n = 1, 2, \dots$

$$\hat{X}(n|n-1) = A\hat{X}(n-1|n-1) \tag{19}$$

$$P(n|n-1) = AP(n-1|n-1)A^T + \sigma_g GG^T \tag{20}$$

$$K(n) = \frac{P(n|n-1)H^T}{HP(n|n-1)H^T + \sigma_v^2} \tag{21}$$

$$\hat{X}(n|n) = \hat{X}(n|n-1) + K(n)(y(n) - H\hat{X}(n|n-1)) \tag{22}$$

$$P(n|n) = [I - K(n)H]P(n|n-1) \tag{23}$$

式中: $\hat{X}(n|n-1)$ 为 $n-1$ 时刻对 $n$ 时刻状态的估计值; $\hat{X}(n|n)$ 为 $n$ 时刻与观测值结合估计出的真实状态; $y(n)$ 为 $n$ 时刻观测值; $P(n|n-1)$ 与 $P(n|n)$ 分别表示预测和估计的误差协方差矩阵; $K(n)$ 为Kalman增益矢量。图3为Kalman算法流程图。

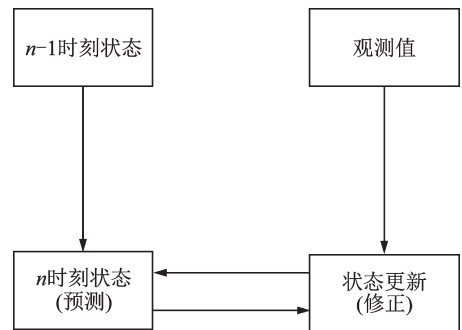


图3 Kalman滤波流程图

Fig.3 Flow chart of Kalman filtering

### 3 仿真实验

#### 3.1 仿真环境

为验证本算法的噪声抑制能力,本文使用MATLAB平台搭建麦克风阵列进行实验仿真,并与传统GSC语音增强方法以及文献[14]所提出的GSC-SS算法进行分析比较。实验设置安排如下:仿真采用6阵元均匀线性麦克风阵列,阵元间距为5 cm,目标声源距离2 m入射角为30°,干扰声源距离2.5 m入

射角为 $-20^\circ$ ,采样频率为16 000,如图4所示。Kalman滤波器迭代次数为12次,自适应滤波器为20阶。其余参数均与文献[14]一致:GSC在32阶自适应滤波器中,正实数 $\varphi=1, \beta=40$ 。

下面采用客观语音质量评估PESQ分别对GSC增强信号、GSC-SS增强信号及所提算法增强信号进行评估分析,并通过不同算法增强信号的波形图与语谱图更为直观地展现本文算法的性能。

### 3.2 PESQ指标

PESQ是评价语音质量的一项客观指标,同时也是国际电信联盟ITU-T的P.862建议评价算法。PESQ以 $-0.5\sim 4.5$ 之间的数值表示增强后语音的质量高低,通过数学模型量化语音的响度、频率等物理与人类心理的感知特性。评估分数越高则输出语音质量越高,反之则质量越低。PESQ的结构如图5所示。

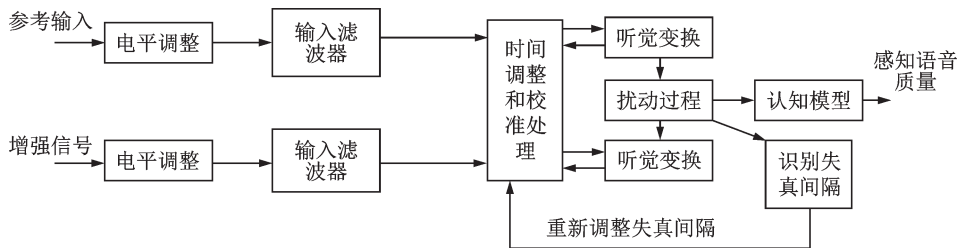


图5 PESQ结构框图

Fig.5 Structure diagram of PESQ

### 3.3 PESQ实验

为验证所提算法的性能,在 $-22\sim 5$  dB不同输入信噪比条件下分别计算GSC输出信号、GSC-SS输出信号与所提算法的PESQ值,如图6所示。图6表明,与GSC及GSC-SS算法相比,本文所提的算法的PESQ值在不同信噪比环境下均较高,这说明所提算法增强后的语音信号失真度更低,更符合人耳听觉主观感受。尤其在信噪比较低时,本文所提算法依旧有较高的PESQ值,相比其他两种算法优越性明显。

### 3.4 语谱图分析

图7(a,b)分别为原始目标语音信号与阵列接收到带噪信号的波形图和语谱图。图7(c)是经GSC增强后的语音信号的波形图和语谱图,可见在对干扰性噪声的抑制上,GSC效果显著,但残留的非相干噪声仍然很多。图7(d,e)分别是采用文献[14]所提出GSC-SS和本文所提出的算法增强后的波形图与语谱图,由比较可知,本文所提出算法在对背景噪声的抑制性能上比GSC-SS更优越,同时有效保留目标信号减少了语音失真。

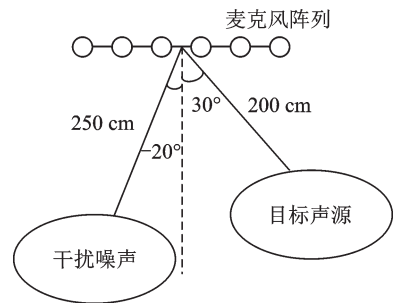


图4 麦克风阵列示意图

Fig.4 Schematic of microphone array

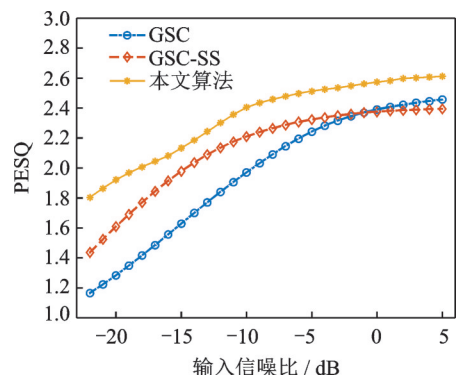


图6 不同信噪比下PESQ评估

Fig.6 PESQ evaluation at different signal-to-noise ratios



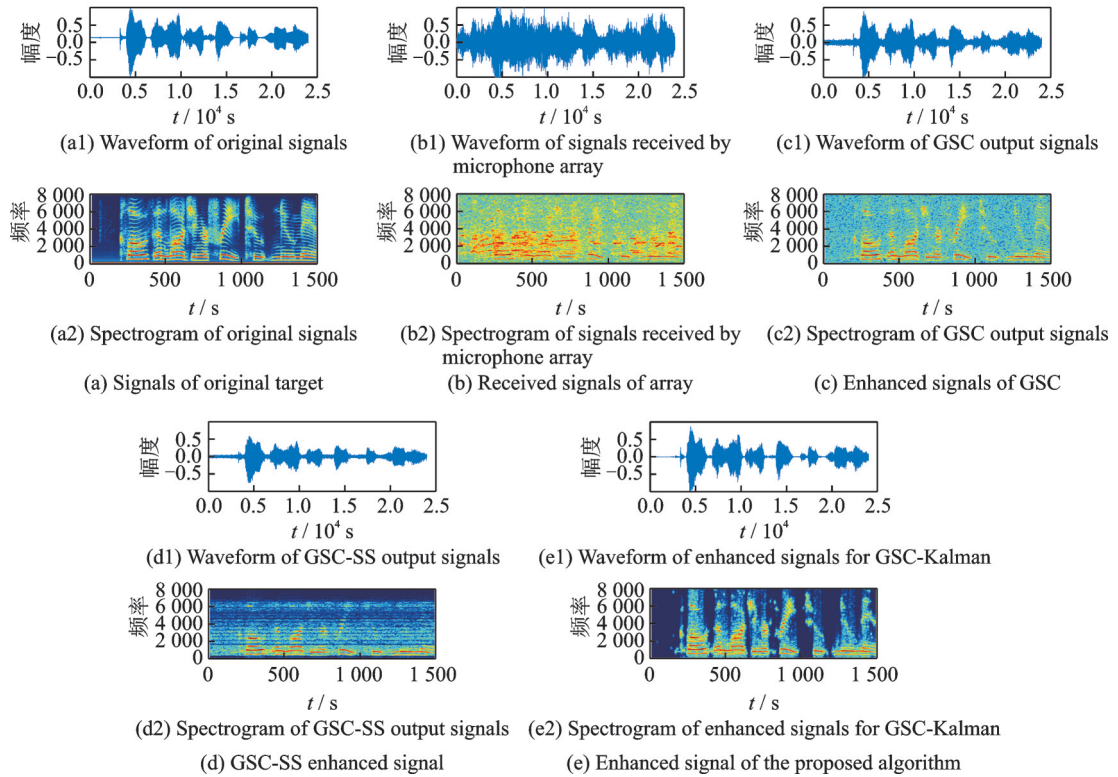


图7 不同算法语谱图对比

Fig.7 Comparison of spectrogram of different algorithms

#### 4 结束语

本文在传统GSC波束形成语音增强方法的基础上加以改进,采用归一化最小均方算法校正自适应对消模块语音泄露的缺陷。输出后的GSC增强信号仍存在热噪声及非相干噪声残留过多的问题,提出后置Kalman滤波器进行迭代MMSE估计,利用前刻状态与当前状态观测值估计当前值。仿真结果验证,本文所提算法在噪声抑制上与传统的GSC算法及后置谱减滤波算法相比具有明显的优越性,增强后的信号可保留更多有用信息并增强语音可懂度。但在声学环境较为复杂无法准确获取方向角信息时性能表现较差,值得加以改进。

#### 参考文献:

- [1] PRIYANKA S S, KUMAR T K. GSC adaptive beamforming using fast NLMS algorithm for speech enhancement [C]// Proceedings of International Conference on Imaging, Signal Processing and Communication (ICISPC). Singapore: [s.n.], 2019: 160-165.
- [2] ROY S K, PALIWAL K K. A non-iterative Kalman filter for single channel speech enhancement in non-stationary noise condition [C]// Proceedings of International Conference on Signal Processing and Communication Systems (ICSPCS). Cairns, Australia: [s.n.], 2018: 1-7.
- [3] MEYER P, ELSHAMY S, FINGSCHIEDT T. A multichannel Kalman-based Wiener filter approach for speaker interference reduction in meetings [C]// ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain: [s.n.], 2020: 451-455.
- [4] 闫姝, 权建军. 基于麦克风阵列的语音增强算法研究[J]. 自动化仪表, 2019, 40(9): 59-62.

- YAN Shu, QUAN Jianjun. Research on speech enhancement algorithm based on microphone array[J]. *Process Automation Instrumentation*, 2019, 40(9): 59-62.
- [5] GRIFFITHS L J, JIM C W. An alternative approach to linearly constrained adaptive beamforming[J]. *Antennas and Propagation, IEEE Transactions on*, 1982, 30(1): 27-34.
- [6] XU Y, NIU L, GAO Q, et al. Research of speech signal dereverberation method based on TF-GSC and minimum phase decomposition[C]//*Proceedings of Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. Chengdu, China: [s.n.], 2017: 1273-1277.
- [7] 马子骥, 余旭, 倪忠. 基于 CTF-GSC 和后置滤波的麦克风阵列语音增强[J]. *西北大学学报(自然科学版)*, 2017, 47(6): 829-837.
- MA Ziji, YU Xu, NI Zhong. Speech enhancement based on CTF-GSC and postfiltering for microphone array[J]. *Journal of Northwest University (Natural Science Edition)*, 2017, 47(6): 829-837.
- [8] YU K, FERNANDEZ M F. LCMV and GSC algorithms with low-sidelobe pattern control[C]//*Proceedings of 2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. Xiamen, China: IEEE, 2017: 1-6.
- [9] PRIYANKA S S, KUMAR T K. GSC beamforming using different adaptive algorithms for speech enhancement[C]//*Proceedings of 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. Kanpur, India: [s.n.], 2019: 1-6.
- [10] NI F, ZHOU Y, LIU H. A robust GSC beamforming method for speech enhancement using linear microphone array[C]//*Proceedings of 2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSp)*. Kuala Lumpur, Malaysia: IEEE, 2019: 1-5.
- [11] 李斌, 张玲华. 一种广义旁瓣抵消器结构的语音增强改进算法[J]. *数据采集与处理*, 2017, 32(2): 307-313.
- LI Bin, ZHANG Linghua. Improved speech enhancement algorithm with generalizated sidelobe canceller[J]. *Journal of Data Acquisition and Processing*, 2017, 32(2): 307-313.
- [12] ZELINSKI R. A microphone array with adaptive post-filtering for noise reduction in reverberant rooms[J]. *Proc of ICASSP88*, 1988(5): 2578-2581.
- [13] CHANG D, ZHENG B. Adaptive generalized sidelobe canceler beamforming with time-varying direction-of-arrival estimation for arrayed sensors[J]. *IEEE Sensors Journal*, 2020, 20(8): 4403-4412.
- [14] YU C, SU L. Speech enhancement based on the generalized sidelobe cancellation and spectral subtraction for a microphone array[C]//*Proceedings of 2015 8th International Congress on Image and Signal Processing (CISP)*. Shenyang: [s.n.], 2015: 1318-1322.
- [15] 梅铁民. 基于反幂法和卡尔曼滤波的自适应语音去混响方法[J]. *信号处理*, 2018, 34(7): 776-786.
- MEI Tiemin. Adaptive speech dereverberation based on inverse power method and Kalman filter[J]. *Journal of Signal Processing*, 2018, 34(7): 776-786.
- [16] GOH Z, TAN K C, TAN B T G. Kalman-filtering speech enhancement method based on a voiced-unvoiced speech model[J]. *IEEE Transactions on Speech and Audio Processing*, 1999, 7(5): 510-524.

## 作者简介:



郭业才(1962-), 通信作者, 男, 博士, 教授, 博士生导师, 研究方向: 通信信号处理、水声信号处理等, E-mail: guo-yecai@163.com。



许雪(1997-), 女, 硕士研究生, 研究方向: 信号处理, E-mail: xu55xue@163.com。



刘力玮(1996-), 男, 硕士研究生, 研究方向: 阵列信号处理。

(编辑: 张彤)