

采用可调波束形成器的 GSC 麦克风阵列语音增强方法

李芳兰¹,周跃海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,

2. 厦门大学信息科学与技术学院, 福建 厦门 361005)

摘要: 基于广义旁瓣抵消器(generalized sidelobe canceller, GSC)算法的麦克风阵列语音增强技术已得到广泛研究,但由于其通常需传统的声源定位方法提供声源方位,语音信号信噪比(SNR)低时声源定位精度将明显下降并影响到语音增强效果. 提出了一种新的麦克风阵列语音增强方法,该方法在 GSC 中引入可调波束形成器估计声源方位以抑制背景噪声影响. 不同类型背景噪声下的实验室语音增强结果表明了该方法的有效性.

关键词: 语音增强; 麦克风阵列; 可调波束形成器; 广义旁瓣抵消器

中图分类号: TN 912. 3

文献标志码: A

文章编号: 0438-0479(2013)02-0186-04

在恶劣的噪声环境下单麦克风接收语音信号的质量将急剧下降,从而影响到语音识别、远程通话、声纹识别等语音信号处理技术的性能^[1]. 近年来,利用麦克风阵列作为语音信号采集前端来改善语音质量越来越受到人们的重视^[2-3]. 麦克风阵列利用了目标信号、噪声和干扰的空间信息,增强了期望方向的信号,抑制其他方向的信号,相对于单通道麦克风而言,能提供更好的增强效果^[4],因而得到了广泛的研究和应用.

广义旁瓣抵消器(generalized sidelobe canceller, GSC)是一种常用的自适应波束形成方法^[5],它能很好地抑制相关干扰. 然而, GSC 算法的信号通路需要借助传统的广义互相关等定位方法^[6-8]提供目标声源方位,当麦克风阵列接收的原始语音信号信噪比(SNR)较低时,会造成定位精度误差,影响了 GSC 算法的消噪性能. 基于此,本文提出了一种采用可调波束形成器的 GSC 麦克风阵列语音增强算法,在 GSC 结构中的固定波束形成器前端引入各通道可调时延补偿,构造可调波束形成器进行声源方位估计,从而在目标声源方位获取阶段即可利用阵列的空间增益来提高方位估计性能.

1 GSC

GSC 是自适应波束形成器的一种通用模型. 线性约束自适应波束形成器可以看作是它的特例,这种模型不但简单,而且更具有一般意义. 有 M 个麦克风的 GSC 由 3 部分组成^[9]: 固定波束形成器、阻塞矩阵和自适应噪声抵消器,如图 1 所示.

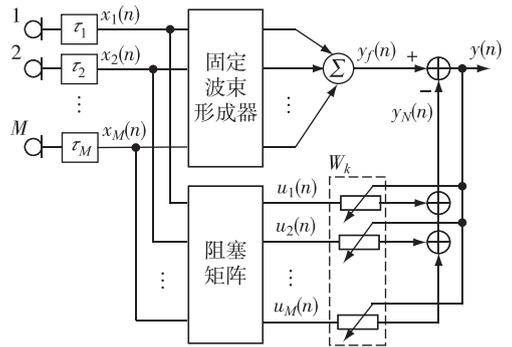


图 1 GSC 结构框图

Fig. 1 Structure of GSC

图 1 中,各通道接收到的信号经过时延 τ_i 补偿后,使得各路输出信号 $x_i(n)$ 在目标信号方向上同步. 固定波束形成器采用系数固定的滤波器对输入信号 $x_i(n)$ 进行滤波,并将滤波后的各路信号相加,实现目标语音的初步增强. 阻塞矩阵的作用相当于一个空间陷波器,阻塞从目标信号方向上来的信号,而通过其他方向上的入射信号. 通过调整自适应部分的系数 W_k ,

收稿日期: 2012-10-16

基金项目: 国家自然科学基金项目(611105026); 福建省高校产学研合作科技重大项目(2010H6022); 厦门市科技计划项目(23502Z20113008)

* 通信作者: ftong@xmu. edu. cn

最小化系统干扰和噪声的输出功率,最终得到增强后的目标语音信号. GSC 自适应支路系数 W_k 的调节可利用实现简单的 LMS(或 NLMS)算法实现.

2 可调波束形成器

对基于 GSC 算法的经典自适应波束形成器,当存在目标语音方位测量误差时,噪声抑制性能受到较大的影响. 要解决这个问题,必须减小背景噪声对传统时延估计算法带来的目标信号方向估计误差. 而传统的利用相关求时延的方法在噪声和混响环境中的估计性能并不理想,直接影响了 GSC 的性能. 因此,本文提出利用可调波束形成器获取目标语音信号方向.

可调波束形成器结构如图 2 所示,它利用阵列形成不同波束角度时 GSC 的输出能量作为代价函数来估计波束对准声源时各通道语音的时延补偿值,以实现目标语音信号方向的定位. 其基本思想是:假定目标信号入射角为 α ,利用不同的 α 求出对应的通道补偿时延值 τ 对各通道接收的语音信号进行时延补偿,然后通过传统 GSC 算法得到各自对应的增强语音,并比较不同 α 对应的输出信号能量,最大能量对应的 α 即为实际目标信号方向,相应的 τ 即为目标语音到达相邻麦克风的时延补偿值.

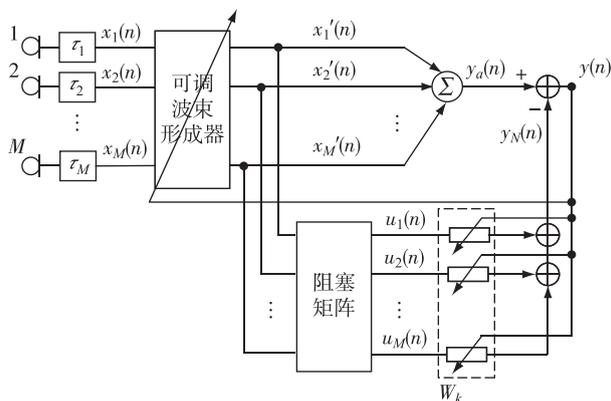


图 2 采用可调波束形成器的 GSC 结构图
Fig. 2 Structure of adjustable beamformer GSC

考虑图 3 中的均匀线列阵,阵列孔径为 d ,目标语音信号源为 $s_1(n)$,其入射角估计值为 α_k . 设声场为远场传播模型,声波以平行波方式传播. 则 α_k 可通过式 (1) 来确定:

$$\alpha_k = \frac{k}{N} \times \pi, k = 0, 1, \dots, N, \quad (1)$$

其中 N 为整数, N 值越大,求出的入射角越精确. 文中取 $N=63$.

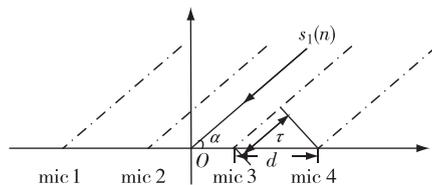


图 3 均匀直线阵
Fig. 3 Uniform linear microphone array

利用 α_k 可求得相应的时延估计值 $\tau(\alpha_k)$:

$$\tau(\alpha_k) = \frac{d \cos \alpha_k}{c} \times f_s, \quad (2)$$

其中 c 为空气中声速.

针对以 f_s 采样频率采样后的数字信号,各通道时延补偿值应为采样点数,则需对 $\tau(\alpha_k)$ 取整,即:

$$\tau'(\alpha_k) = \lceil \tau(\alpha_k) \rceil = \left\lceil \frac{d \cos \alpha_k}{c} \times f_s \right\rceil, \quad (3)$$

式中符号 $\lceil \cdot \rceil$ 表示取整.

对于偶数个麦克风组成的阵列,以阵列中心为坐标原点,对阵列接收信号进行端点检测后截取长为 m 的语音段记为 $x_i(n), i=1, 2, \dots, M$,对 $x_i(n)$ 进行时延补偿得:

$$x'_i(\alpha_k, n) = x_i \left[n - \frac{(M+1-2i)}{2} \times \tau'(\alpha_k) \right]. \quad (4)$$

记 $\mathbf{X}(\alpha_k, n) = [x'_1(\alpha_k, n), x'_2(\alpha_k, n), \dots, x'_M(\alpha_k, n)]^T$,将 $\mathbf{X}(\alpha_k, n)$ 通过固定波束形成通路,可得到入射角 α_k 波束形成的接收信号 $y_a(\alpha_k, n)$:

$$y_a(\alpha_k, n) = \mathbf{W}_a^T \mathbf{X}(\alpha_k, n), \quad (5)$$

其中权向量 $\mathbf{W}_a(\alpha_k, n) = [\omega_1(\alpha_k, n), \omega_2(\alpha_k, n), \dots, \omega_M(\alpha_k, n)]^T$, \mathbf{W}_a 满足 $\mathbf{W}_a^T \vec{1} = 1$.

同时, $\mathbf{X}(\alpha_k, n)$ 经过阻塞矩阵 \mathbf{B} 滤除目标信号得到噪声估计信号:

$$\mathbf{U}(\alpha_k, n) = \mathbf{B} \mathbf{X}(\alpha_k, n). \quad (6)$$

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

$$y(\alpha_k, n) = y_a(\alpha_k, n) - y_N(\alpha_k, n), \quad (7)$$

其中, $y_N(\alpha_k, n)$ 为自适应噪声抵消模块的输出,

$$y_N(\alpha_k, n) = \sum_{l=0}^{L-1} [\mathbf{W}_k(\alpha_k, n)]^T \mathbf{U}(\alpha_k, n-l); \mathbf{W}_k(\alpha_k, n)$$

为滤波系数,且 $\mathbf{W}_k(\alpha_k, n+1) = \mathbf{W}_k(\alpha_k, n) +$

$$\frac{\mu}{\|\mathbf{U}(\alpha_k, n)\|^2} y(\alpha_k, n) \mathbf{U}(\alpha_k, n).$$

求出 α_k 对应的输出语音段能量 $e(\alpha_k)$ 作为可调波束形成器的代价函数:

$$e(\alpha_k) = \sum_{n=0}^m y(\alpha_k, n)^2, \quad (8)$$

则目标语音信号实际方向 α 可通过式(9)来确定:

$$\alpha = \operatorname{argmax}[e(\alpha_k)], k = 0, 1, \dots, N, \quad (9)$$

相应地, α 对应的时延值 $\tau'(\alpha)$ 即为目标语音信号方向上的各通道时延补偿值.

3 实验结果及分析

本文在实验室中利用实测数据对该语音增强系统进行了测试,并将本文算法测试结果与采用广义互相关^[6]进行定位的传统 GSC 算法输出进行了比较.实验室长约为 7 m,宽约为 3 m,麦克风阵列为由 4 个麦克风组成的间距 $d=14$ cm 的均匀线阵列.目标语音到麦克风阵列中心的距离为 2.50 m,通过位于 68° 方向的音箱播放.房间中空调、房间外来往的车声等构成了环境背景噪声(背景噪声级为 57.9 dB(A)),以位于 135° 方向的音箱播放的 3 种噪声为典型的干扰噪声源,包括白噪声、音乐噪声和语音噪声,音箱距麦克风阵列中心的距离为 2.20 m.麦克风阵列和音箱距地面的高度为 1.10 m.

实验中采用的麦克风为单指向性的驻极体麦克风,阵列语音采样频率为 16 kHz,采样时间为 3 s.实验算法中相关参数设置如下:可调波束形成器模块语音段长度 $m=1\ 500$,滤波器长度为 $L=3$,步长为 $\mu=0.2$;2 种算法中语音增强部分滤波器长度均为 $L=20$,步长为 $\mu=0.2$.文中语音段的端点检测采用文献^[10]的检测算法.

表 1 给出了 2 种比较算法的语音增强结果.从表中可以看出,在 3 种不同噪声、不同 SNR 下,本文算法的消噪性能均优于传统的 GSC 算法.

实验计算了不同背景噪声、不同 SNR 下阵列的波束指向图.图 4 以可调波束形成器代价函数 $e(\alpha_k)$ 曲线的形式给出了语音噪声背景下 3 种不同 SNR 条件下

表 1 2 种算法语音增强结果比较

Tab. 1 Comparison of the speech enhancement

	results by two algorithms		dB
	背景噪声	带噪语音	
白噪声	传统 GSC 算法	20.22	29.91
	本文算法	17.55	26.54
		13.49	20.26
音乐噪声	传统 GSC 算法	18.07	28.46
	本文算法	16.22	24.73
		11.95	19.94
语音噪声	传统 GSC 算法	22.49	30.75
	本文算法	19.05	28.35
		14.81	25.07

阵列的波束指向图,图中实线箭头表示目标语音信号方向,虚线箭头表示干扰噪声方向.从图 4 可以看出,可调波束形成器通过角度 α 扫描形成的对准目标语音信号源的波束与实际方向一致,同时可调波束形成器也形成了对准噪声方向的波束,在不同 SNR 条件下语音信号源方向与噪声源方向波束呈现出不同的强度比.

图 5 比较了参考信号、语音噪声背景下麦克风阵列中单个阵元接收到的信号、传统 GSC 输出和本文算法的输出时域波形.从图 5 也可以看出,本文算法对语音背景噪声的抵消性能明显优于传统的 GSC 算法.

4 结 论

本文提出了采用基于可调波束形成器的 GSC 麦克风阵列语音增强方法,通过可调波束形成器对目标语音信号的声源方向进行估计,然后利用 GSC 进行噪声的抑制.在不同的背景噪声条件下,本文方法通过对可调波束形成器各通道目标语音信号进行的时延补

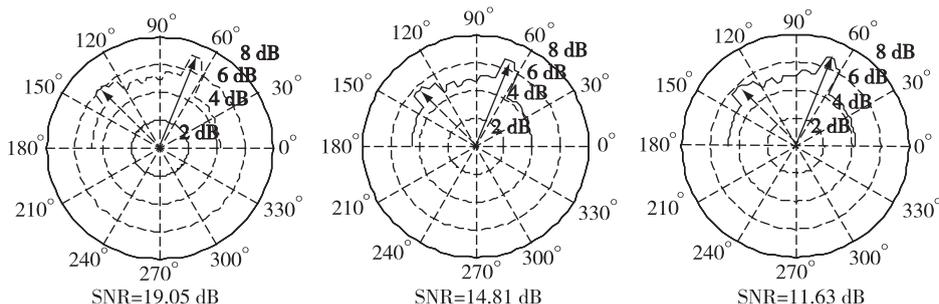


图 4 语音噪声背景下的可调波束形成器获取的语音信号指向图

Fig. 4 Beampattern of the linear microphone array under voice noise environment

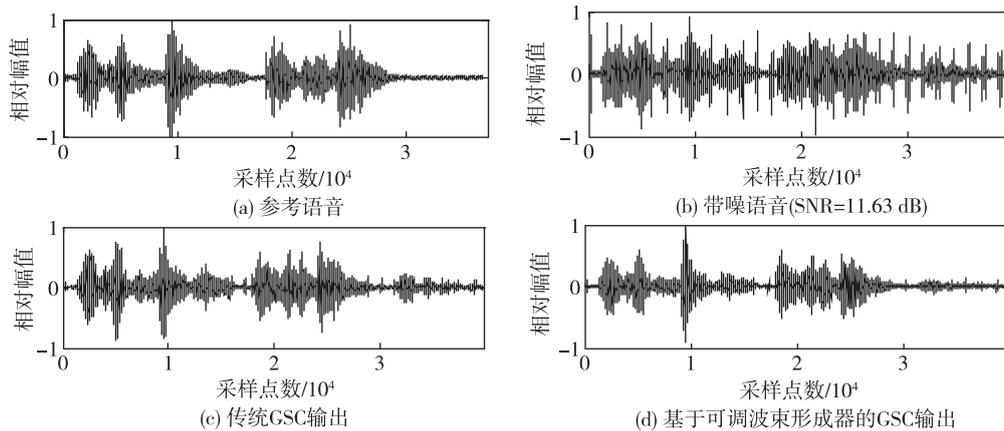


图 5 语音噪声背景下时域波形比较

Fig. 5 Speech waveforms of clean, noisy and enhanced under voice noise environment

偿,可获得较好的声源定向性能,减少了背景噪声对声源方向估计的影响.实验室语音增强实验结果表明了本文算法的有效性.

参考文献:

- [1] 周跃海,童峰,洪青阳.采用 DTW 算法和语音增强的嵌入式声纹识别系统[J].厦门大学学报:自然科学版,2012,51(2):174-178.
- [2] 罗金玉,刘建平,张一闻.麦克风阵列信号处理的研究现状与应用[J].现代电子技术,2010(23):80-84.
- [3] Flanagan J L, Berkley D A, Elko G W, et al. Autodirective microphone systems[J]. Acoustica, 1991, 73(2):58-71.
- [4] 栗晓丽,傅丰林.基于子带 TF-GSC 麦克风阵列语音增强[J].电子科技,2008,21(2):33-36.

- [5] Griffiths L J. An alternative approach to linearly constrained adaptive beamforming[J]. IEEE Transactions on Antennas Propagation, 1982, 30(1):27-34.
- [6] Knapp C H, Carter G C. The generalized correlation method for estimation of time delay[J]. IEEE Trans Acoustics, Speech, and Signal Processing, 1976, 24(4):123-128.
- [7] 夏阳,张元元.基于矩形麦克风阵列的改进的 GCC-PHAT 语音定位算法[J].山东科学,2011,24(6):75-79.
- [8] 陶巍,刘建平,张一闻.基于麦克风阵列的声源定位系统[J].计算机应用,2012,32(5):1457-1459.
- [9] 刘鹏鹏.麦克风阵列语音增强技术的研究与实现[D].大连:大连理工大学,2007.
- [10] 路青起,白燕燕.基于双门限两级判决的语音端点检测方法[J].电子科技,2012,25(1):13-19.

Microphone Array Speech Enhancement Based on Adjustable Beamformer Generalized Sidelobe Canceller

LI Fang-lan¹, ZHOU Yue-hai¹, TONG Feng^{1*}, HONG Qing-yang²

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

Abstract: Generalized sidelobe canceller (GSC) has been widely investigated in microphone array speech enhancement. However, the traditional source localization methods are usually adopted to obtain the direction of source, which is needed for the GSC algorithm. Under relatively low signal noise ratio (SNR) the performance of the GSC will degrade due to the decreasing precision of the source localization. This paper presents a new method for the GSC based microphone array speech enhancement, which uses an adjustable beamformer (ABF) to estimating the direction of the speech source to suppress the background noises. Experimental speech enhancement results under different type of background noises validated the effectiveness of the proposed method.

Key words: speech enhancement; microphone array; adjustable beamformer; generalized sidelobe canceller