

# 基于子带广义旁瓣相消器的麦克风阵列语音增强<sup>\*</sup>

何成林<sup>1,2</sup>, 杜利民<sup>1</sup>, 马 昕<sup>1,2</sup>

(1. 中国科学院 声学研究所 语音交互研究中心, 北京 100080; 2. 中国科学院 研究生院, 北京 100039)

**摘 要:** 为了加快基于广义旁瓣相消器的麦克风阵列语音增强系统的收敛速度, 将其自适应模块的输入信号分解到子带以进行处理, 并将多通道维纳滤波器引入广义旁瓣相消器的非自适应支路, 以更有效地抑制非相干噪声。实际测试结果表明, 相对于基于全带广义旁瓣相消器的麦克风阵列语音增强系统, 采用该子带广义旁瓣相消器结构的语音增强系统具有更快的收敛速度和更高的输出信噪比。

**关键词:** 语音增强; 麦克风阵列; 子带自适应滤波; 广义旁瓣相消器

中图分类号: TN912.3      文献标识码: A      文章编号: 1001-3695(2006)04-0208-03

## Microphone Array Speech Enhancement Based on Subband Generalized Sidelobe Canceller

HE Cheng-lin<sup>1,2</sup>, DU Li-min<sup>1</sup>, MA Xin<sup>1,2</sup>

(1. Speech Interactive Technology Research Center, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100080, China; 2. Graduate School, Chinese Academy of Sciences, Beijing 100039, China)

**Abstract:** In order to accelerate the convergence of the microphone array speech enhancement system based on Generalized Sidelobe Canceller (GSC), the input signals of the adaptive block are decomposed into their subbands to perform the adaptive multi-channel noise cancelling in each subband. And the integration of multi-channel wiener filter into the fixed branch of GSC can suppress the uncorrelated noise efficiently. Experiments indicate that the microphone array speech enhancement system based on this subband GSC has higher converging speed and better output signal-to-noise ratio than its fullband counterpart.

**Key words:** Speech Enhancement; Microphone Array; Subband Adaptive Filter; Generalized Sidelobe Canceller

### 1 引言

对于有些语音应用而言, 如视频会议、车载免提电话等, 由于说话者到麦克风有一定的距离, 麦克风接收到的语音信号受到环境噪声和干扰的影响很大, 严重影响了通话质量。一般的基于单麦克风的语音增强系统难以获得较好的增强效果。麦克风阵列由于利用了目标信号、噪声和干扰的空间信息, 增强从目标信号方向来的信号, 抑制其他方向上的入射信号, 相对于单通道麦克风而言, 基于麦克风阵列的语音增强系统能提供更好的增强效果<sup>[1]</sup>。而一般的基于单通道的语音增强算法, 也可以作为辅助手段结合到麦克风阵列的语音增强系统中, 以获得更高质量的语音。

本文首先介绍了麦克风阵列语音增强中广泛采用的广义旁瓣相消器(Generalized Sidelobe Canceller, GSC), 并介绍了子带自适应滤波的基本原理和子带 GSC 的结构, 最后以实际录音数据对该结构进行测试, 并对测试结果进行了分析和总结。

### 2 广义旁瓣相消器

广义旁瓣相消器作为线性约束最小方差波束形成器(Nearly Constrained Minimum Variance beamformer, LCMV)的另一种实现形式, 由于其结构简单、易于实现, 广泛应用于各种阵列信号处理中。GSC 基本结构如图 1 所示, 各通道接收到的信号经过一段合适的时延  $\tau_m$  补偿后, 使得各路输出信号  $x_m(n)$  在目标信号方向上同步。图 1 中, 上面的非自适应支路为一固定波束形成器(Fixed Beamformer, FBF), 下面阻塞矩阵(Blocking Matrix, BM)的作用相当于一个空间陷波器, 阻塞从阵列指向方向上来的信号, 而通过其他方向上的入射信号。通过调节自适应部分的系数  $W_a$ , 最小化系统干扰和噪声的输出功率, 以实现目标信号的增强。

GSC 的自适应支路系数  $W_a$  的调节可以利用 LMS、RLS 等算法。

(1) LMS(或 NLMS)算法实现简单且易于并行处理, 但在应用于麦克风阵列语音增强中, 如果背景噪声和干扰的特征值分布很广, 利用 LMS 算法收敛速度较慢。

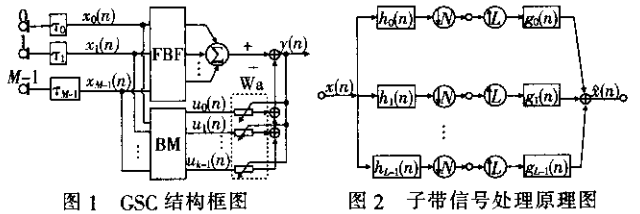
(2) 而 RLS 算法(或其快速横向滤波器实现形式)虽然收敛速度很快, 但运算量大, 且计算中存在数值不稳定性。

为了得到较快的自适应收敛速度和合适的运算量, 本文采用了子带 GSC 的结构<sup>[2]</sup>, 将 GSC 自适应部分的输入信号分解到子带, 再利用 NLMS 算法, 获得了较快的收敛速度。

3 基于子带 GSC 的麦克风阵列语音增强

3.1 子带自适应滤波

为了减少相邻子带间的混叠,子带自适应滤波一般利用过采样滤波器组。如图 2 所示,利用分析滤波器组  $h_l(n)$  将输入的宽带信号分解到  $L$  个子带,并对每个子带的信号以因子  $N$  进行抽取 ( $N < L$ ),再将处理之后的信号以因子  $L$  进行插值。通过合成滤波器组  $g_l(n)$ ,得到处理之后的全带信号。对于一个自适应滤波算法,如果全带滤波器的长度为  $l_p$ ,设完成一次迭代的运算量为  $l_p$  (或  $l_p^2$ ,根据不同的自适应算法而异)数量级,则利用相应的子带自适应滤波。由于以较低的速度更新较少的系数,相应的自适应部分的运算量为原来的  $L/N^2$  (或  $L/N^3$ )<sup>[21]</sup>。



本文中所用的分析、合成滤波器组为过采样 DFT 滤波器组<sup>[31]</sup>。设实值线性相位原型滤波器  $p(k)$  的长度  $l_p$  为偶数,通过广义离散傅里叶变换,得到  $L$  个分析滤波器的冲激响应函数为

$$h_l(k) = p(k) e^{j\frac{2\pi}{L}(l + \frac{1}{2})(k + k_0)}, \quad k = 0, 1, \dots, l_p - 1, l = 1, 2, \dots, L \quad (1)$$

其中的时间偏移因子  $k_0 = -(l_p - 1)/2$ ,保证了如果原型滤波器是线性相位,则各分析、合成滤波器均为线性相位。相应地合成滤波器为

$$g_l(k) = h_l(k), \quad k = 0, 1, \dots, l_p - 1, l = 1, 2, \dots, L \quad (2)$$

式(1)中偏移因子使得各分析滤波器在归一化频率区间  $f \in [0, 2]$  中均匀分布。一般  $L$  为偶数,则对于实值输入信号而言,只用处理其  $[0, 1]$  之间的  $L/2$  个子带的输出,而  $[1, 2]$  之间的  $L/2$  个子带的输出是其复共轭。

图 3 显示了原型滤波器长度  $l_p = 96$ ,子带个数  $K = 8$ ,抽取因子  $N = 6$  的分析(合成)滤波器组的幅频特性。图 4 表示等间隔线性麦克风阵列的子带 GSC 和全带 GSC 的收敛特性的仿真结果。其中阵元个数为 8,阵元间距为 0.04m,目标信号为单位功率的高斯白噪声,入射角为  $0^\circ$ ,宽带干扰信号入射角为  $60^\circ$ ,信号-干扰比  $SIR = -20\text{dB}$ 。各阵元接收到的信号均加了信噪比  $SNR = 20\text{dB}$  的加性高斯白噪声,所有信号的采样率均为  $40\text{kHz}$ ,且目标信号和干扰均位于阵列的远场。两种 GSC 结构中阻塞矩阵取相邻阵元接收到的信号的差值,自适应支路所用的 FIR 滤波器为 60 阶,NLMS 算法的步长为 0.01。其中子带 GSC 的分析、合成滤波器组幅频特性如图 3 所示。图 4 表示 GSC 的输出和目标信号的误差,表征了 GSC 的收敛特性。其中子带 GSC 经过约 1 000 次迭代后基本收敛,稳态误差为  $-27\text{dB}$ ,而全带 GSC 经过约 2 000 次迭代才收敛,稳态误差为  $-21\text{dB}$ 。从中可以看出,子带 GSC 的收敛速度、稳态误差均优

于全带 GSC。

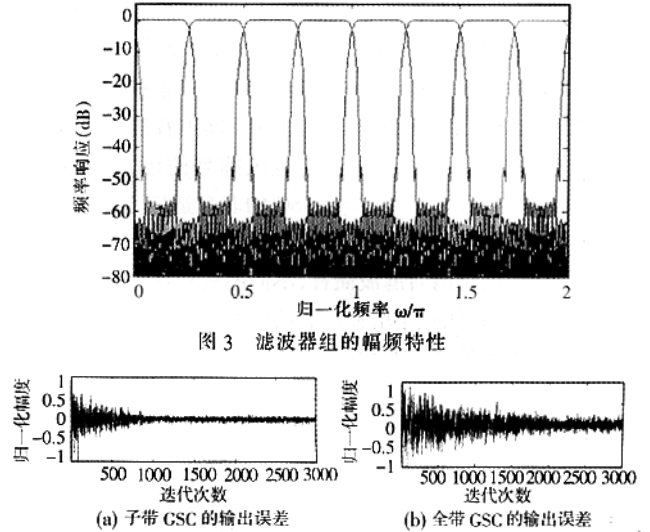


图 3 滤波器组的幅频特性

3.2 子带 GSC 麦克风阵列语音增强的结构和实现

图 5 为基于子带 GSC 的麦克风阵列语音增强系统的结构原理图。各路麦克风接收到的信号经过时延补偿,使得各通道的输出信号在目标信号方向上同步。为了提高时延估计的分辨率,对  $16\text{kHz}$  的输入信号进行 10 倍上采样,再进行基于广义互相关的时延估计,各通道的数据通过时延补偿后,下采样到  $16\text{kHz}$ 。以  $M = 4$  为例,阻塞矩阵为

$$B = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \quad (3)$$

显然阻塞矩阵的输出  $u_l(n)$  表示经过时延补偿后第  $m$  通道和第  $m + 1$  通道接收到的信号的差异。阻塞矩阵的  $M - 1$  路输出和 GSC 上面支路中维纳滤波器的输出分别通过分析滤波器组得到相应的子带信号,再在各个子带进行多参考通道自适应消噪处理,各个子带的输出通过合成滤波器组得到增强之后的语音信号。其中非自适应支路中分析滤波器组后置延时单元  $Z^{-D}$  保证了自适应支路中各子带自适应滤波器的因果可实现性。

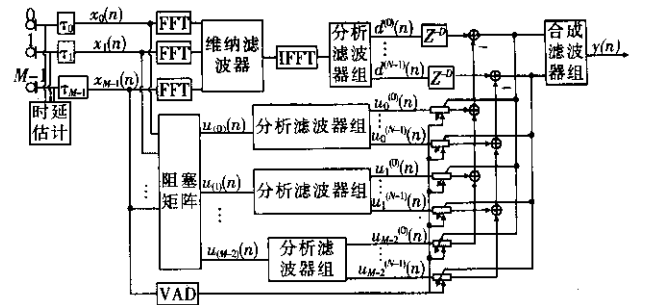


图 5 基于子带 GSC 语音增强系统的结构图

实际中由于时延估计的误差、房间墙壁的反射和各麦克风增益的不一致性等,使得阻塞矩阵的各路输出中含有一定的目标语音信号。如果在所有的时刻自适应滤波器的系数都进行更新,将会导致对目标信号的消除,并在增强之后的语音中有回声现象。实际中为了克服以上缺点,对输入的语音信号进行端点检测(Voice Activity Detection, VAD):在只有噪声期间,

更新自适应滤波器的系数;当有目标语音时,自适应滤波器的系数保持不变。本文中所用的 VAD 检测算法是基于短时零能积(信号短时能量和短时过零率的乘积)的方法<sup>[4]</sup>。

GSC 对相干噪声有很好的消除性能,对于非相干噪声其消除作用十分有限,而维纳后滤波对非相干噪声有很好的消除作用,已广泛地应用于各种基于麦克风阵列的语音增强系统中<sup>[1]</sup>。本文将维纳滤波结合到 GSC 的非自适应支路中,利用维纳滤波器的输出作为 GSC 自适应支路的目标信号<sup>[5]</sup>。该结构相对于传统的维纳后滤波而言,不但能更有效地消除时变的噪声和干扰,而且能减小增强之后的语音信号的失真。

实际中维纳滤波是在频域实现的:

$$W(f) = \frac{\frac{2}{M(M-1)} \operatorname{Re} \left\{ \sum_{i=0}^{M-2} \sum_{j=i+1}^{M-1} \hat{\Phi}_{x_{pi}x_{pj}}(f) \right\}}{\frac{1}{M} \sum_{i=0}^{M-1} \hat{\Phi}_{x_{pi}x_{pi}}(f)} \quad (4)$$

其中,  $M$  是麦克风的个数,  $\hat{\Phi}_{x_{pi}x_{pj}}(f)$  是经过时延补偿之后通道  $i$  和通道  $j$  之间互功率谱的估计,  $\hat{\Phi}_{x_{pi}x_{pi}}(f)$  是通道  $i$  自功率谱的估计。

$$\hat{\Phi}_{x_{pi}x_{pi}}(f) = |\operatorname{FFT}(x_i(n)/N)|^2/N \quad (5)$$

$$\hat{\Phi}_{x_{pi}x_{pj}}(f) = |\operatorname{FFT}(x_i(n)/N) \times \operatorname{FFT}^*(x_j(n)/N)|^2/N \quad (6)$$

其中  $\operatorname{FFT}(x_i(n)/N)$  表示对通道  $i$  当前帧的数据进行  $N$  点 FFT 运算,  $*$  表示取共轭。实际计算中,式(4)中的互功率谱用前一帧和当前帧的估计结果进行更新:

$$\hat{\Phi}_{x_{pi}x_{pj}}(k, f) = \alpha \hat{\Phi}_{x_{pi}x_{pj}}(k-1, f) + (1-\alpha) \hat{\Phi}_{x_{pi}x_{pj}}(k, f) \quad 0 \leq \alpha \leq 1 \quad (7)$$

其中  $\hat{\Phi}_{x_{pi}x_{pj}}(k, f)$  为通道  $i$  和通道  $j$  当前数据帧互功率谱的估计,  $\hat{\Phi}_{x_{pi}x_{pj}}(k-1, f)$  为前一帧互功率谱的估计,对自功率谱的估计类似。其中功率谱更新因子  $0 \leq \alpha \leq 1$ , 反映了功率谱更新的速度。

#### 4 实验结果及分析

本文用卡内基·梅隆大学实测麦克风阵列数据库<sup>[6]</sup>进行系统测试,该数据库所用的麦克风阵列有 15 个麦克风组成的线性嵌套式阵列,并有一头戴式近距离麦克风作为参考。本文选用其中的第 4~12 号麦克风组成的线性等间隔子阵列所录制的数据进行测试;所用数据为其中 arr3A 组数据,录音环境为一嘈杂的实验室,说话者位于阵列的正前方,离阵列中心的距离为 1m,阵元间距为 0.04m,录音数据的采样率为 16kHz。测试中分析、合成滤波器组选用第 3 节介绍的过采样 DFT 滤波器组,均为 96 阶,语音的帧长为 30ms,帧移为 20ms,式(7)中的功率谱更新因子  $\alpha = 0.8$ ,GSC 中自适应 FIR 滤波器为 40 阶。

图 6 比较了参考信号、麦克风阵列中单个阵元接收到的信号、全带 GSC 的输出、子带 GSC 的输出的时域波形。子带 GSC 结构的系统经过大约 2 000 次迭代后即达到稳态,而全带 GSC 需 5 000 次的迭代,即子带 GSC 有更快的收敛速度。其中单路麦克风接收到的信号的信噪比  $\text{SNR} = -9.6\text{dB}$ ,全带 GSC 增强之后的  $\text{SNR} = 8.1\text{dB}$ ,子带 GSC 增强之后的  $\text{SNR} = 11.0\text{dB}$ 。相对全带 GSC 而言,通过子带 GSC 增强之后的语音的信噪比提高了约 3dB。

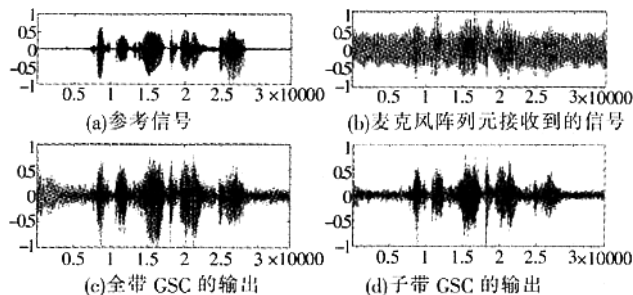


图 6 时域波形比较

#### 5 结论

本文介绍了基于子带广义旁瓣相消器的麦克风阵列语音增强结构,通过将自适应部分的宽带输入信号分解到各个子带的“预白化”处理,加快了自适应部分的收敛速度,GSC 非自适应支路中多通道维纳滤波器有效地抑制了非相干噪声。实验结果表明,相对于基于全带 GSC 的麦克风阵列语音增强系统,基于该子带 GSC 结构的系统有更快的收敛速度,且增强后的语音信噪比约有 3dB 的提高。

但如果 VAD 检测出现错误,如将语音段判别为静音(噪声)段,导致在语音期间更新自适应滤波器的系数,则在输出结果中将会存在明显的回声。当信噪比较高时,基于信号短时零能积的 VAD 检测算法能给出较准确的结果;当信噪比很低时,必须选择更稳健的 VAD 检测算法(如基于最小二乘周期估计的方法)。此外,当应用环境混响很大时,基于广义互相关的时延估计会产生较大的误差,此时必须选取更稳健的结构<sup>[1]</sup>,但均可以利用子带信号处理的方法,加快其自适应部分的收敛速度,提高系统对时变噪声的处理性能。

#### 参考文献:

- [1] M Brandstein, D Ward. Microphone Arrays [M]. New York: Springer-Verlag 2001.
- [2] Wei Liu, Stephan Weiss, Lajos Hanzo. Subband Adaptive Generalized Sidelobe Canceller for Broadband Beamforming [C]. Proceedings of the 11th IEEE Workshop on Statistical Signal Processing, 2001. 591-594.
- [3] M Harteneck, S Weiss, R W Stewart. Design of Near Perfect Reconstruction Oversampled Filter Banks for Subband Adaptive Filters [J]. IEEE Transactions on Circuits and Systems, Part II: Analog and Digital Signal Processing, 1999, 46(8): 1081-1086.
- [4] H G Hirsch, C Ehrlicher. Noise Estimation Techniques for Robust Speech Recognition [C]. IEEE Proc. of the ICASSP 95, 1995. 153-156.
- [5] A Abad, J Hernando. Speech Enhancement and Recognition by Integrating Adaptive Beamforming and Wiener Filtering [C]. IEEE Sensor Array and Multichannel Signal Processing Workshop, SAM, Sitges 2004.
- [6] Tom Sullivan. CMU Microphone Array Database [DB/OL]. [http://www.speech.cs.cmu.edu/databases/micarray/cmu\\_tms\\_multimic.tar.gz](http://www.speech.cs.cmu.edu/databases/micarray/cmu_tms_multimic.tar.gz), 2004-04.

#### 作者简介:

何成林(1977-),男,博士研究生,主要研究方向为数字信号处理、语音信号处理;杜利民(1957-),男,研究员,教授,博士生导师,主要从事信号处理和语音交互技术的研究;冯昕(1969-),男,山东泰安人,博士研究生,主要研究方向为语音识别。

# 基于子带广义旁瓣相消器的麦克风阵列语音增强

作者: [何成林](#), [杜利民](#), [马昕](#), [HE Cheng-lin](#), [DU Li-min](#), [MA Xin](#)  
作者单位: [何成林, 马昕, HE Cheng-lin, MA Xin\(中国科学院, 声学研究所, 语音交互研究中心, 北京, 100080; 中国科学院, 研究生院, 北京, 100039\)](#), [杜利民, DU Li-min\(中国科学院, 声学研究所, 语音交互研究中心, 北京, 100080\)](#)  
刊名: [计算机应用研究](#) **ISTIC** **PKU**  
英文刊名: [APPLICATION RESEARCH OF COMPUTERS](#)  
年, 卷(期): 2006, 23 (4)  
被引用次数: 3次

## 参考文献(6条)

1. [M Brandstein;D Ward](#) [Microphone Arrays](#) 2001
2. [Wei Liu;Stephan Weiss;Lajos Hanzo](#) [Subband Adaptive Generalized Sidelobe Canceller for Broadband Beamforming](#)[外文会议] 2001
3. [M Harteneck;S Weiss;R W Stewart](#) [Design of Near Perfect Reconstruction Oversampled Filter Banks for Subband Adaptive Filters](#)[外文期刊] 1999 (08)
4. [H G Hirsch;C Ehrlicher](#) [Noise Estimation Techniques for Robust Speech Recognition](#)[外文会议] 1995
5. [A Abad;J Hernando](#) [Speech Enhancement and Recognition by Integrating Adaptive Beamforming and Wiener Filtering](#) 2004
6. [Tom Sullivan](#) [CMU Microphone Array Database](#) 2004

## 本文读者也读过(10条)

1. [何成林](#), [杜利民](#), [马昕](#), [He Chenglin](#), [Du Limin](#), [Ma Xin](#) [麦克风阵列语音增强的研究](#)[期刊论文]-[计算机工程与应用](#) 2005, 41 (24)
2. [栗晓丽](#), [傅丰林](#), [Li Xiaoli](#), [Fu Fenglin](#) [基于子带TF-GSC麦克风阵列语音增强](#)[期刊论文]-[电子科技](#) 2008, 21 (2)
3. [栗晓丽](#) [基于GSC结构的麦克风阵列语音增强算法研究](#)[学位论文] 2008
4. [邱作春](#) [麦克风阵列语音增强用于抗噪说话人识别](#)[期刊论文]-[大众科技](#) 2008 (12)
5. [林章儒](#) [麦克风阵列语音增强系统的设计与实现](#)[学位论文] 2007
6. [洪鸥](#), [Hong, Ou](#) [麦克风阵列语音增强技术及其应用](#)[期刊论文]-[微计算机信息](#) 2006, 22 (1)
7. [李金涛](#) [麦克风阵列语音增强算法研究](#)[学位论文] 2010
8. [王冬霞](#), [殷福亮](#), [Wang Dongxia](#), [Yin Fuliang](#) [一种近场环境下的麦克风阵列语音增强方法](#)[期刊论文]-[信号处理](#) 2007, 23 (2)
9. [张骞](#), [刘志红](#), [仪垂杰](#) [麦克风阵列语音增强技术及其消噪性能研究](#)[期刊论文]-[大众科技](#) 2010 (3)
10. [杜军](#), [桑胜举](#), [Du Jun](#), [Sang Shengju](#) [基于麦克风阵列的语音增强技术及应用](#)[期刊论文]-[计算机应用与软件](#) 2009, 26 (10)

## 引证文献(3条)

1. [王晓荣](#), [赵劲劼](#), [赵治栋](#) [基于改进广义旁瓣抵消器的语音增强](#)[期刊论文]-[杭州电子科技大学学报](#) 2007 (5)
2. [林静然](#), [彭启琮](#), [邵怀宗](#), [居太亮](#) [基于麦克风阵列的宽带健壮自适应波束形成算法](#)[期刊论文]-[通信学报](#) 2006 (12)
3. [郭威](#), [申利民](#), [谭云福](#), [陶凤莉](#) [嵌入式语音识别在混响环境中的信号增强方法](#)[期刊论文]-[计算机应用研究](#) 2010 (12)

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_jsjyyyj200604071.aspx](http://d.g.wanfangdata.com.cn/Periodical_jsjyyyj200604071.aspx)