

基于锥形麦克风阵列的语音增强技术研究

孙科伟, 李建海, 杨海东, 宋 博

(空军工程大学 航空航天工程学院, 陕西 西安 710038)

摘 要:针对强噪声环境中小体积应用场合下的语音增强,设计了一种基于锥型麦克风阵列结构的语音增强方法。该方法充分利用锥形结构中锥底麦克风阵列对锥顶麦克风的良好噪声抵消性能,结合端点检测器和两级自适应滤波器设计的多通道抗串扰噪声抵消算法,实现了强噪声环境中的语音增强。理论分析和仿真验证了锥形结构设计的合理性和优越性,仿真结果表明,改进的多路抗串扰噪声抵消算法应用于此锥形结构的麦克风阵列中可提高语音信号的信噪比,对语音的损伤小,在强噪声环境中语音增强效果显著。

关键词:麦克风阵列;多路抗串扰;噪声抵消;语音增强

中图分类号:TN643

文献标识码:A

Study on Speech Enhancement Technique Based on Taper Microphone Array

SUN Kewei, LI Jianhai, YANG Haidong, SONG Bo

(College of Aeronautic and Astronautic, Air Force Engineering University, Xi'an, 710038, China)

Abstract: A speech enhancement technique based on taper microphone array structure is proposed for small volume applications under strong noise environment. The technique takes full advantage of the good noise cancellation performance of the taper bottom microphone array against the taper top microphone in the taper structure and combining with improved multiplex cross-talk resistant noise cancellation algorithm which employed voice activity detector (VAD) and two-stage adaptive filter to realize speech enhancement in strong noise environment. Both the theoretic analysis and simulation results have verified the rationality and advantages of the taper structure. Simulation results indicate that this improved noise cancellation method using for this taper microphone array can significantly increases the SNR, and has small damage to the speech as well as remarkable speech enhancement effect in strong noise environment.

Key words: taper microphone array; multiplex cross-talk resistant; noise cancellation; speech enhancement

0 引言

语音增强是从含噪语音信号中提取尽可能纯净的原始语音,在一些强噪声环境下,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,严重影响了语音通信的质量。麦克风阵列语音增强方法由于利用了信号的空域和时域信息,克服了单通道麦克风语音增强不能同时满足信噪比和语音可懂度的缺点^[1],所以得到了广泛的应用。然而对于一些特殊的场合,如手持送话器、手机中要求麦克风阵列足够小以便嵌入到设备中应用。但是现有许多适用于普通麦克风阵列的语音增强算法运用在小型麦克风阵列中效果不是很好^[2]。针对此,本文在设计麦克

风阵列结构的基础上,结合改进的噪声抵消算法实现了强噪声环境下的语音增强。

1 麦克风阵列结构设计

根据阵列信号处理理论,阵元的空间摆放对阵列系统性能有重要的影响,同样,麦克风阵列的拓扑结构对语音增强系统的性能有重要影响,在常用麦克风阵列结构中,圆形阵列相比一维均匀直线阵列由于其分布在平面内各个离散方向上,因此在不移动阵列的情况下就可对平面内任意方向的语音信号进行增强处理,且二维圆形阵列结构对阵列系统时延估计性能的影响和对阵列输出信噪比的影响都比

收稿日期:2013-06-03

基金项目:航空科学基金资助项目(20095596014)

作者简介:孙科伟(1988-),男,陕西省扶风县人,学士,主要从事强噪声环境下语音降噪技术研究及军事航空通信导航的研究。李建海(1964-),男,陕西人,副教授,主要从事军事航空通信导航的研究。

一维直线阵列优越^[8]。三维阵列对系统的性能提升相比二维阵列更好,但三维阵列结构复杂,成本较高。因此,本文在二维圆形阵列的基础上,设计了锥形麦克风阵列结构,如图1(a)所示。

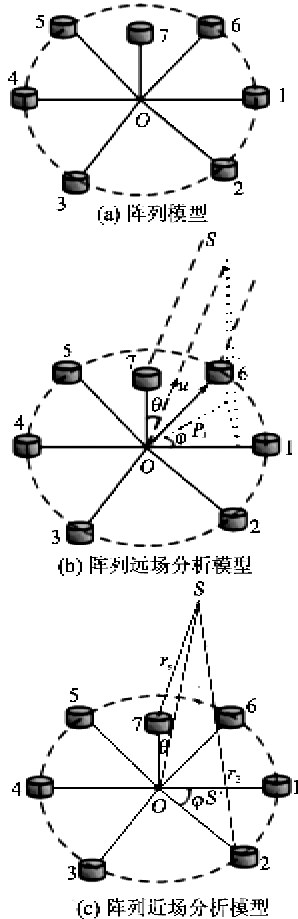


图1 麦克风阵列结构示意图

图1(a)为本文所设计的锥形麦克风阵列的空间摆放位置示意图。6个麦克风均匀的分布在半径为 $r=0.02\text{ m}$ 的圆周上,第7个麦克风分布在垂直于圆周中心点 $d=0.02\text{ m}$ 距离处,选择中心点 O 为整个阵列模型的参考点确定麦克风阵列的位置。 θ 为声源信号与 Z 轴的夹角,即 $90^\circ-\theta$ 为信号的俯仰角, φ 为声源的方位角。

2 锥形麦克风阵列近场和远场信号模型分析

本文所设计的麦克风阵列主要用于抗噪送话器,航空收发电台等小型设备中进行语音增强,因此麦克风阵列接收到语音信号为近场信号,而噪音信号为远场信号,分别针对近场和远场信号模型对阵列性能进行分析。由于这一部分只考虑阵列结构对

阵列增益的影响,所以采用简单的延时求和波束形成方法进行分析。

在远场模型中信号的模型是平面波模型(见图1(b))。忽略各个通道接收到信号的幅度差,只考虑各个通道接收到的信号的相位差。假设信号的方向矢量为

$$\boldsymbol{\mu} = (-\sin\theta\cos\varphi \quad -\sin\theta\sin\varphi \quad \cos\theta)^T \quad (1)$$

设圆形阵列的半径为 r , $\beta=2\pi/M$, $M=6$,圆形阵列上各个阵元的坐标为

$$\boldsymbol{P}_i = [r\cos[(i-1)\beta] \quad r\sin[(i-1)\beta] \quad 0]^T \quad (2)$$

则各个阵元相对于坐标原点的延时为 τ_i ,则

$$\tau_i = \frac{\boldsymbol{\mu}^T \boldsymbol{P}_i}{c} = -\{\sin\theta\cos\varphi r\cos[(i-1)\beta] + \sin\theta\sin\varphi r\sin[(i-1)\beta]\} / c \quad (3)$$

时域的延时对应的是频域信号相位的改变,由此可得在远场模型情况下,阵列的增益函数为

$$B(\omega; \theta, \varphi) = \sum_{i=1}^M \exp(-jk^T \boldsymbol{p}_i) \Big|_{k=\frac{\omega}{c}\boldsymbol{\mu}} \quad (4)$$

在近场模型中,近场波面为球面波,不仅要考虑信号相位的改变,同时要考虑信号幅度的变化(见图1(c))。分析方法和远场分析类似,且各参数不变。可得在近场模型建立下,阵列的增益函数为

$$B(\omega; \theta, \varphi) = \sum_{i=1}^M \frac{r_s}{d_i} \exp[-jk^T (d_i - r_s) \boldsymbol{\mu}] \Big|_{k=\frac{\omega}{c}\boldsymbol{\mu}} \quad (5)$$

式中: r_s 为信号源距离参考点的距离; d_i 为信号源距离第 i 个麦克风的距离。

3 改进的多通道抗串扰噪声抵消算法

由于应用场合的限制使麦克风之间的距离相对较近,用麦克风阵列进行语音增强时,麦克风阵元之间的距离较近,导致语音增强性能受到阵间串扰的影响。由于双通道抗串扰自适应噪声抵消算法在噪声环境较复杂时系统不稳定,且抗噪性能较差^[4],针对此锥形结构提出了多通道抗串扰自适应噪声抵消算法(MCRANC),本文根据前面设计的锥形阵列结构,将锥顶麦克风作为主麦克风,锥底麦克风阵列作为参考麦克风,通过自适应噪声抵消算法实现了系统的噪声抵消性能。

图2为应用MCRANC实现语音增强的示意图。该语音增强系统主要由语音端点检测(VAD)和有限脉冲响应(FIR)型数字滤波器A、B模块组成。为了阻塞主通道的语音信号 x_0 ,输出纯净的噪

音信号到滤波器 B 模块进行处理,选择其他 N 个麦克风采集的信号 $x_{i_1}, x_{i_2}, \dots, x_{i_N}$, 作为参考信号。通过 VAD 端点检测算法区分出 NonVoice Period (NVP) 和 Voice Period (VP), 在 NVP 阶段, 假定 $s_i(k)=0$, 即各个麦克风接收到的均为纯净的噪音信号, 则由图可得 $x_0 = y_1 + e_1$, 即

$$n_0(k) = w\bar{n}(k) + \text{err}(k) \quad (6)$$

式中: $x_0 = n_0(k)$, $e_1 = \text{err}(k)$ 为预测误差; $y_1 = w\bar{n}(k)$ 为滤波器 A 的输出; w 为 $N(L+1)$ 维系数行向量, 即 $w = (w_1, w_2, \dots, w_N)$, $w_i = (w_{i0}, w_{i1}, \dots, w_{iL})$, $\bar{n}(k)$ 为噪声信号, $\bar{n}(k) = [\bar{n}_1(k), \bar{n}_2(k), \dots, \bar{n}_N(k)]^T$, 其中 $\bar{n}_i(k) = [n_i(k), n_i(k-1), \dots, n_i(k-L)]^T$ 。

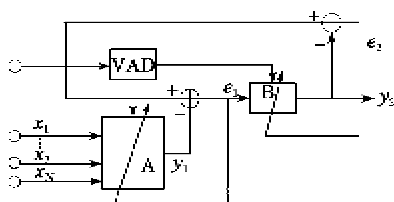


图2 多通道抗串扰噪声抵消语音增强系统

调整权值 w , 使输出误差最小, 此时的 w 为最佳权值 w^0 , 则最佳权值向量

$$w^0 = (w_{10}^0, w_{11}^0, \dots, w_{N0}^0, w_{N1}^0, \dots, w_{NL}^0) \quad (7)$$

此时的最小功率, 即误差信号的功率表示为 $P(\text{err}^0(k))$, 理论上, 最优误差的功率与参考通道数成反比。

在有语音阶段(HP), 假设噪声信号是平稳信号且变化缓慢, 则可利用无语音阶段求得的噪声信号来近似代替有语音阶段的噪声信号, 所以上述求得的滤波器 A 的最优权值向量仍不改变, 则可得

$$y_1(k) = w^0 \bar{x} = w^0 (\bar{s} + \bar{n}) = w^0 \bar{s} + [n_0(k) - e^0(k)] \quad (8)$$

式中: $\bar{x}$ 为接收到的语音信号和噪声信号; $\bar{s}$ 为对纯净语音信号的期望值。其形式与式(8)所示的 $\bar{n}$ 的形式类似。则可得

$$\begin{aligned} e_1(k) &= x_0(k) - y_1(k) = [s_0(k) + n_0(k)] - \\ & y_1(k) = [s_0(k) + n_0(k)] - w^0 \bar{s} - \\ & [n_0(k) - e^0(k)] = s_0(k) - w^0 \bar{s} + \\ & e^0(k) = p(k) + e^0(k) \end{aligned} \quad (9)$$

其中 $p(k) = s_0(k) - w^0 \bar{s}$, 对其求 z 变换可得

$$E_1(z) = P(z) + E^0(z) \quad (10)$$

$$\begin{aligned} P(z) &= S_0(z) - z \left[\sum_{i=1}^N \sum_{j=0}^L w_{ij}^0 s_j(k-j) \right] = \\ & S_0(z) - z \left[\sum_{i=1}^N \sum_{j=0}^L w_{ij}^0 h_{i,s_0}(k-j) \times \right. \\ & \left. s_0(k-j) \right] = \left[1 - \sum_{i=1}^N \sum_{j=0}^L w_{ij}^0 z^{-2j} \cdot \right. \end{aligned}$$

$$H_{i,s_0} \left. \right] S_0(z) = \tilde{H}(z) S_0(z) \quad (11)$$

$$\tilde{H}(z) = 1 - \sum_{i=1}^N \sum_{j=0}^L w_{ij}^0 z^{-2j} H_{i,s_0} \quad (12)$$

由式(11)可看出, $e_1(k)$ 可用 $s_0(k)$ 线性表示, 因此 $e_1(k)$ 可用来抵消 $s_0(k)$ 。

对于滤波器 B, 只需调整其系数向量, 使 $y_2(k)$ 的功率最小, 则可得语音抵消后噪声信号的估计值。

调整滤波器 B 使其传输函数近似为 $\tilde{H}^{-1}(z)$, 则可得

$$\begin{aligned} Y_2(z) &= H_B E_1(z) = \tilde{H}^{-1}(z) E_1(z) = \\ & \tilde{H}^{-1}(z) [\tilde{H}(z) S_0(z) + E^0(z)] = \\ & S_0(z) + \tilde{H}^{-1}(z) E^0(z) \end{aligned} \quad (13)$$

转换到时域可得

$$y_2(k) = s_0(k) + \tilde{h}^{-1}(z) \times e^0(k) \quad (14)$$

$$\begin{aligned} \| e_2(k) \|^2 &= \| x_0(k) - y_2(k) \|^2 = \\ & \| s_0(k) + n_0(k) - y_2(k) \|^2 = \\ & \| n_0(k) \|^2 + \| s_0(k) - y_2(k) \|^2 + \\ & 2n_0(k)[s_0(k) - y_2(k)] \end{aligned} \quad (15)$$

假定噪声信号 n_0 和语音信号 s_0 不相关, 且 n_0 的均值为 0, 则可得

$$E(e_2^2) = E(n_0^2) + E[(s_0 - y_2)^2] \quad (16)$$

由式(16)可得, 最小化 $E(e_2^2)$ 的值, 相当于使 $E[(s_0 - y_2)^2]$ 值最小, 即等价于 s_0 和 y_2 的误差最小, 亦即从麦克风信号 $x_0(k)$ 中抵消语音信号 $s_0(k)$, 假定主通道信号和参考信号相关, 则可实现 s_0 和 y_2 的误差最小。为适应环境的动态变化, 使最终输出的增强后的语音具有较少的残留噪声, 则可在有语音和无语音阶段对其滤波器 B 系数都作自适应调整。

4 锥形麦克风阵列结构对语音增强性能影响仿真分析

仿真分析锥形结构对阵列增益的影响, 首先分析信号远场模型对阵列增益的影响。将信号的频率固定为 3 000 Hz, 将参考麦克风和主麦克风增益相减, 得到阵列增益, 比较主通道和参考通道的相关程

度。输入信号的俯仰角 $0 \leq \theta \leq 180^\circ$, 方位角 $0 \leq \varphi \leq 360^\circ$, 仿真可得阵列增益变化如图3、4所示。

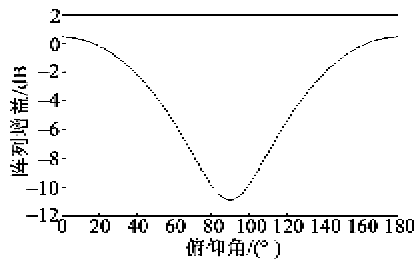


图3 俯仰角 对阵列增益的影响

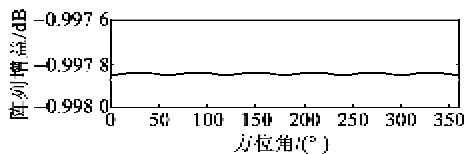


图4 方位角 对阵列增益的影响

分析图3可得,当 $\theta=0^\circ$ 时,阵列的增益最大, $\theta=90^\circ$ 时阵列的增益最小可达 -10.915 dB ,在 $0^\circ \sim 90^\circ$ 范围内呈递减,由于此系统中语音信号在垂直于阵列方向入射,则此时阵列的增益最大,而噪声作为远场信号,在其他方向均有不同程度的衰减,在水平方向的衰减最大。由图4可知,信号的方位角改变对阵列的增益几乎无影响,且由于圆形阵列间隔 60° ,所以阵列增益呈周期性变化。同上仿真方法,分别分析远场和近场模型中信号频率对阵列增益的影响。图5为远场信号频率对阵列增益的影响。图6为近场信号频率对阵列增益的影响。由图可知,在近场和远场模型中信号频率对阵列增益的影响趋势相近。即在垂直于阵列方向入射的信号相比其他方向入射的信号衰减最小。水平方向的衰减最大。随着信号频率的增大阵列增益成周期性衰减变化,对比图5、6可知,在 $50 \sim 3500\text{ Hz}$ 范围内同一频率下远场信号衰减比近场信号更大。在 3000 Hz 时远场信号可衰减到原来信号的 0.282 倍,且频率越小衰减越大。衰减的幅度和增益变化转折点均与阵元之间的距离 r 和 d 有关。

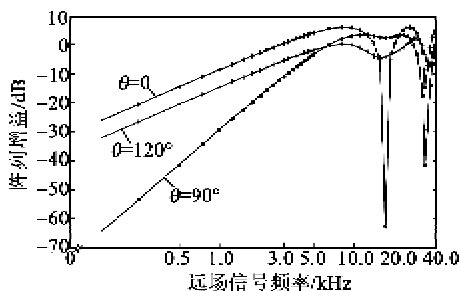


图5 远场信号频率对阵列增益的影响

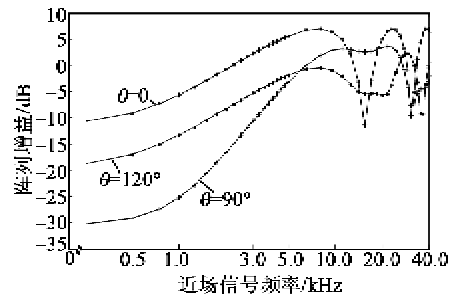


图6 近场信号频率对阵列增益的影响

5 锥形麦克风阵列语音增强性能仿真分析

为验证设计的锥形麦克风阵列结构结合噪声抵消算法的语音增强效果,分别将之前录制好的语音信号和噪音信号保存在计算机中,为满足语音信号近场和噪音信号远场的条件,设置语音信号是在距离麦克风阵列 0.2 m 处采集的人说话音,噪声是之前录制好的机场环境下的噪声信号。不断调整噪声信号和麦克风阵列之间的距离,由于距离和声压成反比,改变距离来控制噪声的声场强度,在语音信号强度不变时,通过调整噪声源的位置来改变输入信号的信噪比。假设噪声源距离麦克风阵列的初始距离为 2 m 。信号的采样频率是 8 kHz ,滤波器A和B中采用的是归一化最小均方(NLMS)算法来改变滤波器的权值向量,其中的端点检测算法(VAD)采用的是文献[5]中的方法。

图7为仿真的结果。调整噪声信号的大小,使其满足强噪音的环境。图7(a)为主麦克风接收到的混合信号,经过测试可得此时信号的信噪比为 -5.012 dB ,图7(b)为经过普通的多通道噪声抵消算法(MANC)后得到的语音增强效果,此时信号的信噪比为 3.8328 dB ,但经此方法增强后的语音信号有很大的失真,增强后的语音中有用的语音信号也被削弱了,通过试听发现语音的可懂度较低。图7(c)为双通道抗串扰自适应噪声抵消算法(DCRANC),此时信号的信噪比为 2.1235 dB ,通过试听增强后的语音可知语音的可懂度相比图7(b)有提升,但残留的噪音信号仍很大。图7(d)为本文所设计的改进的多通道抗串扰自适应噪声抵消算法(MCRANC)增强的结果,此时测试得信号的信噪比为 7.3524 dB ,试听可得此时的语音可懂度较高且残留噪声较小。对比图7可知,MANC对语音信号的破坏最大,MCRANC相比其他两种算法,进一步提高了输入信号的信噪比和语音的可懂度,语

音增强效果较好^[6]。

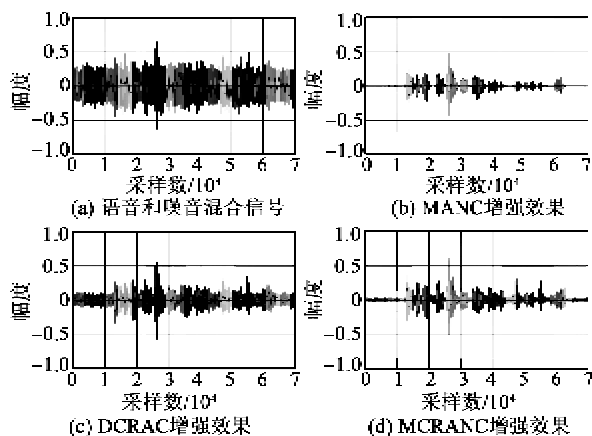


图 7 语音增强结果

6 结束语

本文在研究阵列信号处理理论和自适应噪声抵消方法的基础上设计了一种锥形麦克风阵列语音增强系统,通过对麦克风阵列结构优化和算法的改进实现复杂噪声环境下纯净语音信号的提取。仿真分析了锥形阵列结构对阵列增益的影响。提出了改进的多通道抗串扰自适应噪声抵消算法,仿真结果表明将此算法应用在此锥形结构中可取得很好的语音增强效果。在同等条件下相比其他自适应噪声抵消方法信号的信噪比有很大的提升,且语音的可懂度较高。本文所设计的小型麦克风阵列语音增强系统可广泛的应用于抗噪送话器等小型设备的前端进行

语音信号的采集。

参考文献:

- [1] 袁坤,梁维谦,张国先. 基于等边三角形麦克风阵列的语音增强[J]. 清华大学学报:自然科学版, 2009, 49(7): 987-990.
- [2] ZENG Qingning, OUYANG Shan. Speech enhancement by array crosstalk resistant ANC and spectrum subtraction[J]. Chinese Journal of Acoustics, 2008, 27(1): 85-96.
- [3] 王冬霞,赵光,郑家超. 麦克风阵列拓扑结构对语音增强系统性能影响的理论分析[J]. 辽宁工业大学学报:自然科学版, 2010, 30(1): 1-4.
- [4] 曾庆宁,欧阳缮. 子带 MCRASC-MGSC 微型麦克风阵列语音增强算法[J]. 西安电子科技大学学报:自然科学版, 2010, 37(6): 1011-1016.
- [5] ZHAO H, ZHAO L X, ZHAO K, et al. Voice activity detection based on distance entropy in noise environment[C]//Seoul, Korea; Fifth International Joint Conference on INC, IMS and IDC, 2009: 1364-1367.
- [6] 李银国,薛雯,徐洋. 基于噪声短时谱动态估计的语音增强谱减算法[J]. 重庆邮电大学学报:自然科学版, 2010, 22(2): 127-130.

LI Yinguo, XUE Wen, XU Yan. Spectral subtraction algorithm for speech enhancement based on dynamic estimation of noise short-time spectrum[J]. Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition, 2010, 22(2): 127-130.

(上接第 305 页)

参考文献:

- [1] SAROIU S, GUMMADI P K, GRIBBLE S D. A measurement study of peer-to-peer file sharing systems [C]//Washington; Proc of the Multimedia Computing and Networking 2002, MMCN 2002, 2002: 156-170.
- [2] MU T, NANDI A, RANGAYYAN R M. Screening of knee-joint vibroarthrographic signals using the strict 2-surface proximal classifier and genetic algorithm[J]. Computers in Biology and Medicine, 2008, 38(2): 1103-1111.
- [3] MASCARO B, PRIOR J, SHARK L K, et al. Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints [J]. Medical Engineering and Physics, 2009, 31(3): 1013-1022.
- [4] 史峰,王辉,郁磊. matlab 智能算法 30 个案例分析 [J]. 北京航空航天大学:自然科学版, 2010(12): 27-37.
- [5] 文福拴,韩祯祥,田磊,等. 基于遗传算法的电力系统故障诊断的解析模型与方法(三)[J]. 电力系统及其自动化学报, 1999, 11(3): 8-13.
- [6] 何正嘉,岑艳阳,孟庆丰,等. 机械设备非平稳信号的故障诊断原理及应用[M]. 北京:高等教育出版社, 2001.
- [7] 李舜酩,许庆余. 微弱振动信号的谐波小波频域提取 [J]. 西安交通大学学报, 2004, 38(1): 51-55.