

基于小波阈值去噪的麦克风阵列语音增强方法

高婉贞, 张玲华

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

摘要: 现有的麦克风阵列语音增强方法中的延迟-求和波束形成算法只对不相干噪声或弱相干噪声有一定的消噪能力, 如果语音中混有较强相干噪声, 则此传统的方法对其没有消除能力。针对这个局限性, 文中把小波阈值去噪的方法与传统的延迟-求和波束形成算法有效结合, 使其对相干噪声也具有很好的消噪能力, 同时减少由于噪声的存在而引起的时延估计误差, 提高时延估计的准确性, 使最终求和结果更好。通过仿真结果表明, 这种改进方法可以改善最终语音效果, 提高语音清晰度, 使人耳更好地接受。

关键词: 麦克风阵列; 小波阈值去噪; 延迟-求和; 时延估计; 时延补偿

中图分类号: TP301

文献标识码: A

文章编号: 1673-629X(2014)10-0144-03

doi: 10.3969/j.issn.1673-629X.2014.10.034

Microphone Array Speech Enhancement Algorithm Based on Wavelet Threshold Denoising

GAO Wan-zhen, ZHANG Ling-hua

(College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: Delay-sum beamforming method as a kind of microphone array speech enhancement algorithms only is effective for irrelevant noise and weak coherent noise. If the speech mixes coherent noise, the conventional method cannot remove it. For this limitation, in this paper wavelet threshold denoising is applied into conventional delay-sum beamforming speech enhancement algorithm, making it to be effective for removing coherent noise. While reducing the time estimation error caused by the noise, improve the accuracy of time delay estimation and make the final sum result better. The simulation results show that the improved method is effective, which improves the clarity of speech and degree of ear acceptance.

Key words: microphone array; wavelet threshold denoising; delay-sum; time delay estimation; time delay compensation

0 引言

生活中由于噪声的存在, 语音的清晰度和可懂度受到影响, 造成了人耳听觉疲劳, 严重影响人们的生活质量。因此采用有效的方法去除语音中的背景噪声以提高信噪比成为语音信号处理中的重点研究对象。目前主要分为单通道语音增强和多通道语音增强技术。单通道方法主要有谱减法^[1]、最小均方误差法^[2]和维纳滤波器法等, 但这些方法已不能满足听力要求, 所以近年来多通道语音增强算法成为人们研究的重点。麦克风阵列在一定的空间域里按某种方式分布排列, 具有空间选择特性, 用“电子瞄准”的方式从所需要的声

源方向提供高质量的语音信号^[3]。现在的麦克风阵列算法主要有三类: 延迟-求和固定波束形成算法^[4-5]、自适应波束形成算法^[6]和具有自适应后置滤波器的波束形成算法^[7]。其中固定波束形成算法首先将阵列中各个传声器接收到的信号进行时延补偿, 以使各通道的语音信号同步, 再对各通道的信号进行加权求和取平均处理, 可以消除不相干噪声, 但语音中混入的相干噪声很难被消除^[8-9], 同时会影响时延估计的准确性, 因此文中设计在时延估计前先对每路信号进行小波阈值去噪处理, 这样处理后再进行时延估计可以增强其准确性并且可以消除相干噪声, 弥补延迟-求和波束

收稿日期: 2013-11-25

修回日期: 2014-03-03

网络出版时间: 2014-07-28

基金项目: 江苏省教育自然科学基金研究重大项目(13KJA510003); 江苏高校优势学科建设工程资助项目(PAPD)

作者简介: 高婉贞(1989-), 女, 安徽蚌埠人, 硕士研究生, 研究方向为语音信号处理; 张玲华, 通讯作者, 教授, 博士研究生导师, 研究方向为语音和说话人的识别和转换技术、各种网络环境中的实时语音通信技术等。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140728.1222.015.html>

形成方法只能消除不相干噪声的缺陷。通过实验仿真表明,文中的方法与传统的延迟-求和固定波束形成算法相比较有更好的效果,并进行听辨测试,得出经改进算法处理后的语音增强信号的自然度和清晰度有所提高。

1 基于小波阈值去噪的延迟-求和波束形成算法

传统的延迟-求和波束形成算法首先将阵列接收的信号进行时延估计,然后进行时延补偿,往往由于噪声的存在,时延估计的准确性会受到影响,所以文中先对每路信号进行小波阈值去噪,再做时延估计,这样可以增加时延估计的准确性,并且先消除背景噪声中的相干噪声,再对各路信号进行加权求和,变成一路信号并且消除部分不相干噪声或弱相干噪声。文中设计的系统框图如图 1 所示。

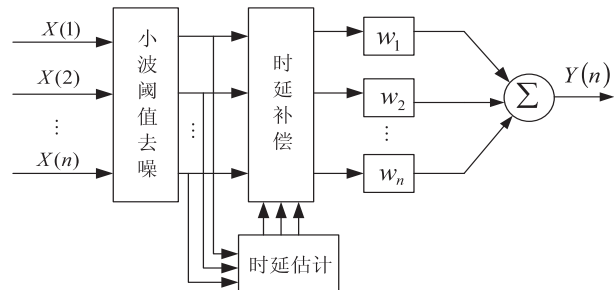


图 1 基于小波阈值去噪的延迟-求和波束形成算法框图

算法流程如下:

- (1) 将麦克风接收到的各路信号先进行小波阈值去噪,去除相干噪声;
- (2) 再将经过小波处理后的各路信号两两相关,计算出各时延,这里选第 1 路为参考信号,其他 $n-1$ 路分别与该路信号通过自相关法求出时延估计,并进行时延补偿;
- (3) 将经过时延补偿的信号加权求和形成 1 路信号。

1.1 传统延迟-求和波束形成算法的原理

传统的延迟-求和波束形成算法主要包括时延估计、时延补偿和加权求和三部分,选用多路麦克风进行接收信号,为了减少算法复杂度,选用了四路线性麦克风阵列模型。图 2 是文中选用的四路线性麦克风阵列图。

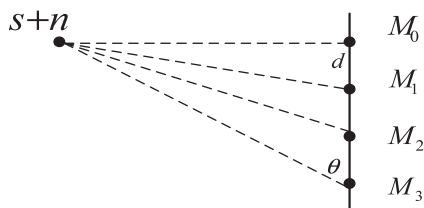


图 2 四路麦克风阵列图

设麦克风阵列的输入信号为 $x_i(t)$, 则对于均匀线性阵列,其原理表达式可以表述如下:

$$x_i(t) = s_i(t - \tau_{0i}) + n_i(t) \quad i = 0, 1, 2, 3 \quad (1)$$

$$y(t) = \sum_{i=1}^N w_i(t) x_i(t - \tau_{0i}) \quad (2)$$

将上式变换到频域上,有

$$Y(\omega) = \sum_{i=1}^N w_i(\omega) X_i(\omega) e^{-j\omega\tau_{0i}} \quad (3)$$

其中, $s_i(t)$ 是纯净的语音信号; $n_i(t)$ 是第 i 路麦克风接收到的噪声信号; $\mathbf{W} = [w_1, w_2, \dots, w_N]^T$ 是麦克风阵列的加权矢量; τ_{0i} 是声源到第 i 路麦克风的时延(由各路信号与第一路信号自相关求出)。

延迟-求和波束形成算法首先对各个麦克风上接收到的信号进行时间补偿,使各个通道语音信号同步,然后对各个通道进行加权求和,权重 w_i 通常取 $1/N$ 。固定波束形成方法的优点是算法简单,易于实现,但通常只对不相干噪声有较好的消噪效果。为了使消噪效果更加理想,将小波阈值去噪法加入其中。

1.2 小波阈值去噪

近年来小波变换方法在语音处理中应用广泛^[10-11],采用小波法去噪的任务是在小波域将信号的小波变换与噪声的小波变换有效地分离。1992 年,Donoho 和 Johnstone 提出了小波阈值去噪的方法,小波阈值去噪的原理是:由于正交小波变换具有很强的去数据相关性,能够使信号的能量在小波域集中在一些大的小波系数中;而噪声的能量却分布在整個小波域内,因此,经小波分解后,信号的小波系数幅值要大于噪声的系数幅值,可以认为幅值比较大的小波系数一般以信号为主,而幅值比较小的系数在很大程度上是噪声。可以设定一个阈值,幅值低于该阈值的小波系数置为 0,高于该阈值的小波系数可以完整保留或做相应的收缩处理^[12]。

传统的小波阈值去噪方法^[13]有两种:硬阈值方法和软阈值方法。

(1) 硬阈值化方法:当小波系数的绝对值 $|W_{j,k}|$ 小于给定的阈值 th 时,令其为 0;而大于阈值时,则保持不变,即:

$$\tilde{W}_{j,k} = \begin{cases} W_{j,k}, & |W_{j,k}| \geq th \\ 0, & |W_{j,k}| < th \end{cases} \quad (4)$$

(2) 软阈值化方法:当小波系数的绝对值 $|W_{j,k}|$ 小于给定的阈值 th 时,令其为 0;而大于阈值时,令其都减去阈值,即:

$$\tilde{W}_{j,k} = \begin{cases} \text{sign}(W_{j,k}) (|W_{j,k}| - th), & |W_{j,k}| \geq th \\ 0, & |W_{j,k}| < th \end{cases} \quad (5)$$

式(4)、(5)中,通常采用固定阈值即 $th = \sigma$

$\sqrt{2\ln N}$, σ 是噪声标准差, N 为待处理信号长度。硬阈值方法中, 调整后的小波系数 $\tilde{w}_{j,k}$ 在 th 处是不连续的, 给重构后的信号造成动荡^[14], 噪声消除不干净, 且混有音乐噪声, 而采用软阈值方法调整后的小波系数是连续的, 与硬阈值方法相比效果较好, 所以文中采用软阈值的方法。

1.3 时延估计和补偿

文中采用自相关法进行时延估计, 两信号的相对时延就是两信号广义互相关函数的最大值所对应的位置。由式(1)可得两路信号的互相关表达式为:

$$R_{ij}(\tau) = E[x_i(t)x_j(t+\tau)] = R_{ss}(\tau - \tau_{ij}) + R_{sn_i}(\tau - \tau_{ij}) + R_{sn_j}(\tau) + R_{nn_{ij}}(\tau) \tag{6}$$

又假设信号与噪声、噪声与噪声之间互不相关, 则 $R_{sn_i}(\tau - \tau_{ij}) = R_{sn_j}(\tau) = R_{nn_{ij}}(\tau) = 0$ (7)

式(7)可化为

$$R_{ij}(\tau) = R_{ss}(\tau - \tau_{ij}) \tag{8}$$

由式(8)可知, 当 $\tau = \tau_{ij}$ 时自相关函数取最大值。根据以上分析, τ_{ij} 就是两信号的相对时延估计值。但通常由于各种噪声的存在, 信号与噪声, 噪声与噪声间不是完全不相关的, 因此会造成时延估计误差, 所以文中先对四路信号进行小波阈值去噪, 再做时延补偿。

2 实验与仿真

实验中, 选用语音库中的一段纯净的语音信号, 采样频率为 16 kHz, 每个样值用 16 bit 二进制数进行编码, 并加入信噪比为 -3 dB 的白噪声形成带噪语音。选用线性麦克风阵列模型, 4 路麦克风接收语音信号, 把第 1 路作为参考信号, 其余 3 路做相对的时延处理, 小波阈值处理选软阈值处理方法。图 3 为针对白噪声情况进行处理的结果。

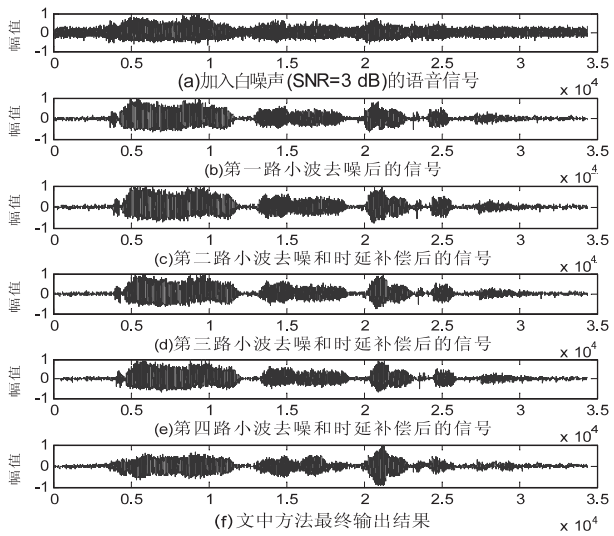


图 3 文中方法仿真实验图

而未做改进的传统方法处理后的语音和原纯语音

如图 4 所示。

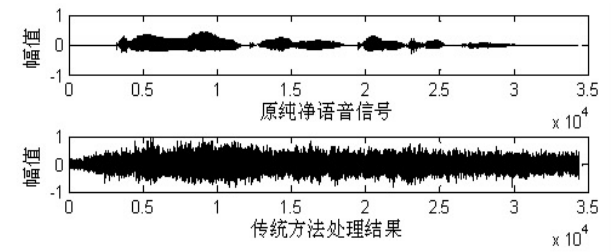


图 4 传统方法仿真实验图

通过图 3 和图 4 的比较可以看出各路信号经过小波阈值法去噪后, 白噪声明显被消除, 同时得到较准确的时延补偿后, 各路信号可以很好地对齐, 使最终处理结果较理想, 接近原语音信号; 而没有加入小波阈值去噪处理的传统延迟-求和波束形成算法处理后的结果, 却没有很好的去噪效果, 时延补偿受到影响。表 1 是加入白噪声形成不同信噪比的带噪信号经过文中模型处理后的信噪比提升结果。

表 1 不同信噪比的带噪信号处理前后的比较

项目	数值		
处理前 SNR/dB	3.0	0.0	-3.0
处理后 SNR/dB	3.704 3	4.758 9	2.160 1

从表 1 可以看出文中的麦克风阵语音增强方法对白噪声有良好的消除效果, 弥补了传统延迟-求和波束形成算法不能消除相干噪声的局限, 增加了可用性。

3 结束语

文中提出了一种小波阈值去噪和麦克风阵语音增强的结合方法, 可以改善传统的延迟-求和波束形成方法的性能, 使其既可以消除不相干噪声也可以消除相干噪声。通过实验仿真和听辨实验的测试, 结果表明被固定白噪声干扰的语音信号通过此方法的处理, 噪声会被消除, 信噪比增加。

参考文献:

[1] Boll S. Suppression of acoustic noise in speech using spectral subtraction[J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1979, 27(2): 113-120.

[2] Ephraim Y, Malah D. Speech enhancement using a minimum mean-square error short-time spectral amplitude estimation [J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1984, 32(6): 1109-1121.

[3] 洪 鸥. 麦克风阵列语音增强技术及其应用[J]. 微计算机信息, 2006, 22(1-1): 142-144.

[4] Kellermann W. A self-steering digital microphone array [C] //Proc of international conference on acoustics, speech and signal processing. Toronto: IEEE, 1991: 3581-3584.

3 测试与分析

测试结果如表 2 所示,其中√表示通过,×表示该功能中有不能使用的部分。项 1 为命令请求,项 2 为信息获取,项 3 为数据回传,项 4 为远程控制。

表 2 测试结果

测试机型	版本号	项 1	项 2	项 3	项 4
SCH-N719	4.1.3	√	√	√	√
HTC Desire HD	2.3.5	√	√	√	√
华为 U8160	2.2.3	√	√	√	√
某山寨机品牌	4.0.3	√	√	√	√

从测试结果可以看出该木马的运行情况良好,各大主流厂商的机器均能够良好运行。测试过程中在一台 ROM 给大幅度修改的山寨机上获取短信时因为短信数据库的格式发生了较大幅度改变导致出现过问题,在对木马进行了一个重新适配后能够正常运行。所以可能在一些大幅度修改的定制 ROM 上使用会出现不可预知的错误。同时从使用效果上来看当受控端在线的时候将请求间隔设置为 10 s 左右就近似可以看成是实时控制了,与短信控制相差无几,甚至在一些情况下能够比短信更快。而且比短信控制更隐蔽不易被发现。

4 结束语

随着移动设备技术的不断发展,其已经在人们的生活中占据了一个极其重要的位置。而 Android 作为其中唯一的一款开源主流操作系统,针对其的攻击显然会非常多。而这些攻击中又以隐私获取与远程控制的危害最大,如何应对这些攻击就成为了当下研究的一个重点。希望文中所提出的木马原型能够对研究起到一定的积极作用。

参考文献:

[1] Developers A. What is android? [J/OL]. 2011. <http://developer.android.com/guide/basics/what-is-android.html>.

[2] 刘仙艳. 移动终端开放平台—Android[J]. 信息通信技术, 2011,5(4):50-53.

[3] 网秦安全公司. 2013 上半年网秦全球手机安全报告[R]. 北京:网秦安全公司,2013.

[4] Enck W,Ongtang M, McDaniel P. Understanding android security[J]. IEEE Security & Privacy,2009,7(1):50-57.

[5] Enck W,Octeau D,McDaniel P,et al. A study of android application security[C]//Proceedings of the 20th USENIX conference on security. [s. l.]:[s. n.],2011:21-44.

[6] Ongtang M,McLaughlin S E,Enck W,et al. Semantically rich application-centric security in Android[J]. Security and Communication Networks,2011,5(6):658-673.

[7] 宋 杰,党李成,郭振朝,等. Android OS 手机平台的安全机制分析和应用研究[J]. 计算机技术与发展,2010,20(6):152-155.

[8] 董 蕾,黄淑华,尹浩然,等. 基于 Android 平台的手机木马关键技术分析[J]. 信息网络安全,2012(11):63-65.

[9] 耿东久,索 岳,陈 渝,等. 基于 Android 手机的远程访问和控制系统[J]. 计算机应用,2011,31(2):559-561.

[10] 蔡罗成. Android 后台监听实现机制浅析[J]. 信息安全与通信保密,2010(6):39-41.

[11] 赵建勋. 基于 Android 平台的移动位置服务的开发与实现[J]. 现代商贸工业,2010,22(20):271-272.

[12] 黄志勇,赵 霁. 基于 Android 平台的移动位置信息服务开发研究[J]. 自动化技术与应用,2011,30(12):22-26.

[13] 温 敏,艾丽蓉,王志国. Android 智能手机系统中文件实时监控的研究与实现[J]. 科学技术与工程,2009,9(7):1716-1719.

[14] 黄锦川,金炜东. 基于 Android 平台 Web 服务的应用研究[J]. 铁路计算机应用,2010,19(11):24-27.

(上接第 146 页)

[5] Flanagan J L, Johnston J, Zahn R. Computer-steered microphone arrays for sound transduction in large rooms[J]. Journal of Acoustical Society of American,1985,78(5):1508-1518.

[6] Griffiths L J, Jim C W. An alternative approach to linearly constrained adaptive beamforming[J]. IEEE Transactions on Antennas and Propagation,1982,30(1):27-34.

[7] Zelinski R. A microphone array with adaptive post-filtering for noise reduction in reverberant rooms[C]//Proc of international conference on acoustics,speech and signal processing. New York:IEEE,1988:2578-2581.

[8] 马晓红,殷福亮,陆晓燕,等. 基于小波变换的传声器阵列语音增强方法[J]. 大连理工大学学报,2003,43(4):511-515.

[9] 王冬霞,殷福亮. 联合波束形成与谱减法的麦克风阵列语

音增强算法[J]. 大连理工大学学报,2006,46(1):121-126.

[10] 张宝琳,张玲华,林志敏. 基于多分辨率小波阈值去噪的响度补偿方法[J]. 计算机技术与发展,2012,22(12):83-86.

[11] 徐耀华,王 刚,郭 英. 基于时频阈值的小波包语音增强算法[J]. 电子与信息学报,2008,30(6):1363-1366.

[12] 孙延奎. 小波分析及其应用[M]. 北京:机械工业出版社,2005.

[13] 李海东,李 青. 基于阈值法的小波去噪算法研究[J]. 计算机技术与发展,2009,19(7):56-58.

[14] 李如玮,鲍长春,窦慧晶. 基于双正交小波包分解的自适应阈值语音增强[J]. 仪器仪表学报,2008,29(10):2135-2140.

基于小波阈值去噪的麦克风阵列语音增强方法

作者：[高婉贞](#)，[张玲华](#)，[GAO Wan-zhen](#)，[ZHANG Ling-hua](#)
作者单位：[南京邮电大学 通信与信息工程学院, 江苏 南京, 210003](#)
刊名：[计算机技术与发展](#)[ISTIC](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2014(10)

引用本文格式：[高婉贞](#), [张玲华](#), [GAO Wan-zhen](#), [ZHANG Ling-hua](#) [基于小波阈值去噪的麦克风阵列语音增强方法](#)[期刊论文]-[计算机技术与发展](#) 2014(10)