

文章编号: 1673-9590(2008)06-0058-04

基于奇异加权的麦克风阵列波束形成语音增强算法

张丽艳¹, 普仕凡², 陈少华¹, 曾洁¹, 郭宇明¹

(1. 大连交通大学 电气信息学院, 辽宁 大连 116028; 2. 大连市 91550 部队 230 所, 辽宁 大连 116023)*

摘 要:提出了一种基于奇异值加权的麦克风阵列波束形成算法. 该算法首先采用相位加权互功率时延估计法对麦克风阵列接收到的信号进行时延补偿. 其次对处理后的矩阵采用雅可比算法进行奇异值分解构造权值并求和, 得到增强的语音信号. 仿真结果表明, 与固定权值波束语音增强算法比较, 该算法对加性非相干噪声有明显的消噪效果.

关键词:语音增强; 麦克风阵列; 波束形成; 奇异值分解

中图分类号: TN912 **文献标识码:** A

A Microphone Array Beamforming Speech Enhancement Algorithm Based on SVD

ZHANG Li-yan¹, PU Shi-fan², CHEN Shao-hua¹, ZENG Jie¹, GUO Yu-ming¹

(1. School of Electrical & Information Engineering, Dalian Jiaotong University, Dalian 116028, China; 2. 230 Research Dalian 91550 Army, Dalian 116023, China)

Abstract: A new microphone array of singular-value weighting beamforming (SWBF) algorithm for speech enhancement is proposed. First, the algorithm adopts the crosspower spectrum phase (CSP) algorithm to compensate time delay of the signal received by the microphone array. Second, to the processed matrix, the algorithm adopts Jacobi algorithm to carry out singular value decomposition (SVD), and then use the singular value as the weighting value to get the weighting sum, which is the output enhancement speech. The simulation shows that in the additive incoherent noise, the output signal to noise ratio (SNR) is improved greatly compared with fixed weighting value beamforming (FWBF) speech enhancement algorithm.

Key words: speech enhancement; microphone array; beamforming; singular value decomposition

智能仪器将在未来的社会中占主导地位, 而用语音实现人与智能仪器交互也将会成为一大趋势. 大多数语音识别系统在无噪的环境下的识别率可达到 95% 以上, 但当环境噪声较大时, 系统的性能迅速降低, 甚至无法识别. 解决这一问题的有效方法是对带噪语音信号进行增强处理, 以消除背景噪声, 提高语音质量. 现有的语音增强方法较多, 按其麦克风的数目, 可以分为单麦克风语音增强方法和麦克风阵列语音增强方法两类. 麦克风阵列由在一定的空间域里按某种方式分布排列的一系列麦克风组成. 麦克风阵列系统具有空间选择性, 它可以用“电子瞄准”的方式从所需要的声源方向提供高品质的声音信号, 因此其语音增强效果要优于单麦克风语音增强效果.

利用麦克风阵列进行语音增强的方法可分为 3 大类: 第一类是传统波束形成技术, 例如延时求和波束形成器^[1]. 加权求和波束形成器、滤波求和波束形成器等. 第二类是自适应波束形成技术, 例如广义

* 收稿日期: 2008-05-13

基金项目: 大连市科技计划资助项目 (2005C11CG208)

作者简介: 张丽艳 (1974 -), 女, 讲师, 博士研究生, 主要从事语音信号处理、嵌入式系统开发等方向的研究

E-mail: zhangly@djtu.edu.cn.

旁瓣语音增强技术^[2]、线性约束最小均方波束形成器^[3]。第三类是子空间语音增强技术,例如基于广义奇异值语音增强算法^[4]。第一类和第三类适用于弱相干噪声消除,第二类适用于消除相干噪声,而对非相干噪声或散射噪声消除效果比较差。且第二类中广义旁瓣语音增强算法通常需以固定加权波束形成增强的语音作为一路参考信号,因此,固定加权波束形成算法的增强效果,会影响广义旁瓣语音增强系统的性能。本文提出的基于奇异值分解的麦克风阵列加权波束形成算法,是属于第一类加权求和波束形成器。该算法对散射噪声有较好的增强效果,并且本文算法比固定加权波束形成语音增强算法输出的信噪比有明显的提高。

1 固定波束麦克风阵列语音增强算法

1.1 信号模型

给定一个由 3 个麦克风 M_0, M_1 和 M_2 组成的均匀线阵,相邻麦克风的间距为 d ,声源位置为 (x, y) ,则三个麦克风接收到的信号可以表示为

$$X_i(t) = \alpha_i r(t - t_i) + n_i(t) \quad (i = 0, 1, 2) \quad (1)$$

式中: $r(t)$ 为声源信号, α_i 为声波传播到第 i 个麦克风的衰减因子; t_i 为声波传播到第 i 个麦克风所需的时间; $n_i(t)$ 第 i 个麦克风的干扰成分, $r(t)$ 与 $n_i(t)$ 不相关。

1.2 固定权值波束形成算法

为使某一方向上的信号获得最大增益,需要在各个麦克风上加一时间延迟补偿 τ_{ij} (参考麦克风的标号), $j = 0, \dots, 2$ 为阵列中麦克风的标号,以补偿不同麦克风在接收该方向声波的时间延迟。以 M_0 为参考麦克风,其时间延迟为 τ_0 ,则麦克风阵列接收到的信号可表示为^[5]

$$X_i(t) = \alpha_i r(t - t_i - \tau_{0i} - \tau_0) + n_i(t) \quad (i = 0, 1, 2) \quad (2)$$

对式 (2) 的信号进行加窗傅立叶变换,得到

$$X_i(t, f) = \alpha_i R_w(t, f) e^{-j2\pi f(t_i + \tau_{0i} + \tau_0)} + N_i(t, f) \quad (i = 0, 1, 2) \quad (3)$$

式中, $R_w(t, f)$ 表示信号 $r(t)$ 的加窗傅立叶变换, $N_i(t, f)$ 表示第 i 个麦克风噪声的加窗傅立叶变换。对 $X_i(t, f)$ ($i = 0, \dots, 2$) 进行加权求和,利用 $\tau_0 + \tau_{00} = \tau_1 + \tau_{01} = \tau_2 + \tau_{02}$,并将噪声进行分解,可得到

$$X(t, f) = \frac{1}{3} \sum_{i=0}^2 X_i(t, f) \alpha_i R_w(t, f) e^{-j2\pi f(\tau_{0i} + \tau_0)} + \frac{1}{3} \sum_{i=0}^2 N_{in}(t, f) + \frac{1}{3} \sum_{i=0}^2 N_{ic}(t, f) \quad (4)$$

式中, $N_{in}(t, f)$ ($i = 0, \dots, 3$) 为噪声中的不相干成分, $N_{ic}(t, f)$ ($i = 0, \dots, 3$) 为噪声中的相干成分。

由式 (4) 可看出,声源信号部分基本没有变化,噪声部分分两种情况: 各麦克风的非相干噪声由于彼此之间不相干,噪声强度变为原来的 $1/3$,得到了一定程度的抑制; 各麦克风的相干噪声部分由于彼此之间具有一定的相关性,对其抑制的程度取决于它们之间的相关性,相关性越弱,抑制效果越好。因此,固定波束形成语音增强方法可抑制不相干或若相干噪声。

1.3 基于互功率相位加权的时延估计

参考麦克风与阵列中其他麦克风之间的归一化互功率谱为^[6]

$$C_i(t, f) = \frac{X_0(t, f) X_i(t, f)}{|X_0(t, f)| |X_i(t, f)|} \quad (5)$$

式中, $X_i(t, f)$ ($i = 0, \dots, 2$) 表示信号 $X_i(t)$ ($i = 0, 1, 2$) 的加窗傅立叶变换。对 $C_i(t, f)$ 进行傅立叶反变换,则有

$$C_i(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} C_i(t, f) e^{j2\pi f\tau} df \quad (6)$$

其中, $C_i(t, \tau)$ 是相关函数,在延迟 $\tau_{0i} = \tau_0 - \tau_i$ 处有极大值点。

2 基于奇异值加权的波束形成语音增强方法

将每个麦克风接收到的带噪语音信号分帧作为输入矩阵的列向量。对输入矩阵作奇异值分解,奇异

值的不同,说明各个麦克风输出信号的能量分布的不同,其原因在于各个麦克风受到的噪声干扰不同.因此可在奇异值分解的基础上,估计出麦克风测量数据的可信度,对于可信度高的高的麦克风测量数据赋予大的权值,对于可信度低的麦克风赋予小的权值,从而得到增强的语音信号.

将时延补偿后的麦克风接收数据分段作为输入矩阵的列向量,则 M 个麦克的全部测量值构成输入矩阵 X ,

$$X = [X_1 \ X_2 \ \dots \ X_M] = \begin{matrix} X_1(1) & X_2(1) & \dots & X_M(1) \\ X_1(2) & X_2(2) & \dots & X_M(2) \\ \dots & \dots & \dots & \dots \\ X_1(N) & X_2(N) & \dots & X_M(N) \end{matrix} \quad (7)$$

对式 (7) 采用了单边 Jacobi 算法^[7-8]进行奇异值分解,从而得到未经排序的 M 个奇异值 $\lambda_i (i = 1, \dots, M)$,也就确定了每个奇异值与输入矩阵的列向量的对应关系.奇异值反映了每个麦克风受到噪声影响的程度.根据误差理论,若将较大的权值赋予具有较小误差的测量值,为将较小的权值赋予具有较大误差的测量值,这样的估计结果将优于由算术平均得到的估计结果.采用归一化权值

$$W_i = \frac{1}{\lambda_i} \frac{1}{\sum_{i=1}^M \lambda_i^{-1}}, \quad i = 1, \dots, M \quad (8)$$

则增强的语音信号为

$$Y = \sum_{i=1}^M W_i X_i \quad (9)$$

式中, $Y = Y_i, (i = 1, \dots, N)$.

3 算法的实现步骤

(1)时延补偿: 用式 (5) 计算参考麦克风和其余麦克风之间的归一互功率函数 $\hat{c}_i(t, f)$; 用式 (6) 计算参考麦克风和其余麦克风之间的相关函数 $C_i(t)$; 估计出参考麦克风和其余麦克风之间的时间延迟 $\hat{\alpha}_{0i} (i = 0, 1, 2)$; 用式 (2) 对每个麦克风接收到的信号进行时间延迟补偿.

(2)奇异值分解构造权值: 用式 (7) 构造输入矩阵; 根据单边 Jacobi 算法计算输入矩阵的奇异值,得到 M 各奇异值; 用式 (8) 计算权值.

(3)用式 (9) 加权求和得到增强的语音信号.

4 实验结果

实验中,纯净的语音采用计算机录制的语音,噪声选自 NOISEX-92 数据库,语音信号和噪声信号的采样率均为 8 kHz,均匀线性阵列中的麦克风的间距 d 为 4 cm,入射角约为 60° ,分帧长度为 80 点/帧,在每一帧分四段同时处理.时延补偿时,因采用的是均匀线阵,因此,只需估计出 M_0 与 M_1 的时延 $\hat{\alpha}_{01}$,而 $\hat{\alpha}_{02}$ 可由式 (10) 计算的到

$$\hat{\alpha}_{02} = 2 \hat{\alpha}_{01} \quad (10)$$

语音增强前后的波形比较,如图 1 所示.在图 1 中,(a)是纯净的语音;(b)是一路输入信噪比为 10 dB 的带白噪声的语音;(c)是固定波束形成增强得到的语音;(d)是本文算法增强得到的语音.各图中横轴为语音采样点数,纵轴为信号的幅度.(b)和 (d)的语音时域波形相比较所反映出的本文算法的增强效果.(c)和 (d)反映了固定波束形成算法与本文算法的增强比较效果.

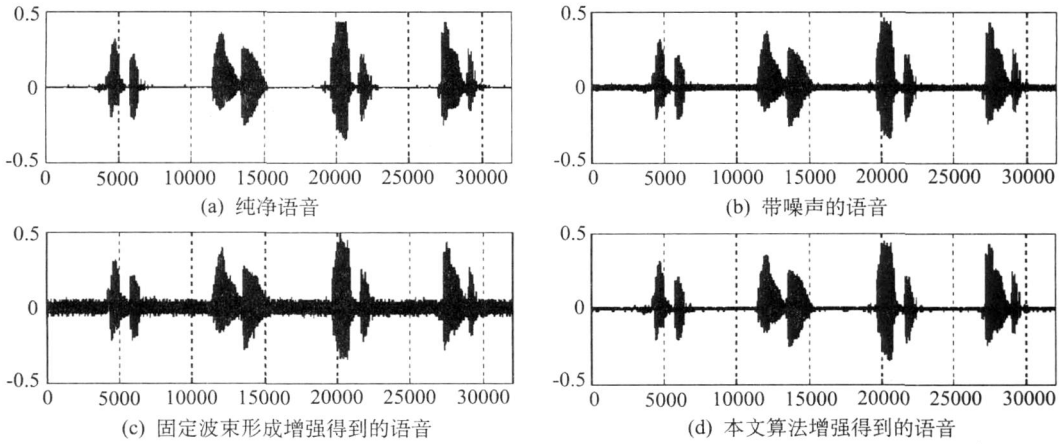


图 1 语音波形图

本文还采用如下分段输入信噪比公式和分段输出信噪比公式来度量算法的语音增强效果.

$$SegSNR_{input} = 10 \lg [(\sum_{k=1}^N s^2(k)) (\sum_{k=1}^N n^2(k))^{-1}] \tag{11}$$

式中, $s(k)$ 为输入的纯净语音信号, $n(k)$ 为输入的加性噪声信号.

$$SegSNR_{output} = 10 \lg [(\sum_{k=1}^N \tilde{s}^2(k) - \tilde{n}^2(k)) (\sum_{k=1}^N \tilde{n}^2(k))^{-1}] \tag{12}$$

式中, $\tilde{s}(k)$ 和 $\tilde{n}(k)$ 分别表示算法增强输出信号中语音部分和噪声部分的成分.

图 2 为固定波束形成语音增强算法与本文算法输出信噪比的对比图, 信噪比由式 (12) 计算得到, 其中 N 值取 50. 由图可知, 本文算法比固定波束形成算法的信噪比约提高 5 dB.

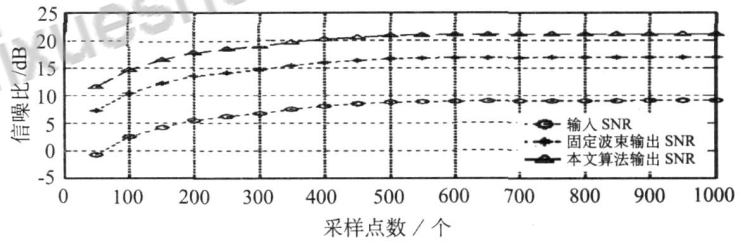


图 2 固定波束形成语音增强算法与本文算法输出信噪比的对比图

总之本文算法对加性非相干噪声有明显的消噪效果. 无论从信噪比的提高, 还是从实际视听语音增强效果来看, 该方法都明显的消除了带噪语音中的噪声成分, 这一点具有实际应用价值.

参考文献:

[1] FLANAGAN J L, JOHN SHON J D, ZAHN R, et al Computer-steered microphone array for sound transduction in large rooms [J]. J. Acoust Soc. Am., 1985, 78(5): 1508-1518

[2] FUDGE G L, NEBARGE D A. A calibrated generalized side-lobe canceller for wideband beamforming [J]. IEEE Trans Signal Processing, 1994, 42: 2871-2875.

[3] GRIFFITHS J, JIM C W. An alternative approach to linearly constrained adaptive beamforming [J]. IEEE Trans AP, 1982, 30: 27-34.

[4] DOCLO S, MOONEN M. GSVD-Based Optimal Filtering for Single and Multimicrophone Speech Enhancement [J]. IEEE Trans on Signal Processing, 2002, 50(9): 2231-2244.

[5] 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.

[6] OMOLOGO M, SVAIZER P. Use of the cross-power-spectrum phase in acoustic event localization [J]. IEEE Trans On Speech and Audio Processing, 1997, 5(3): 288-292.

[7] 周长发. 科学与工程数值算法 [M]. 北京: 清华大学出版社, 2002: 60-79, 95-103.

[8] 张宝琳, 谷同祥, 莫则绕. 数值并行计算原理与方法 [M]. 北京: 国防工业出版社, 1999: 100-106.

论文降重、修改、代写请扫码



免费论文查重, 传递门 >> <http://free.paperyy.com>



阅读此文的还阅读了:

1. [子带MCRASC-MGSC微型麦克风阵语音增强算法](#)
2. [Quattro3 MT303 / DuetEXE: 麦克风](#)
3. [基于多统计模型和人耳听觉特性的麦克风阵列后滤波语音增强算法](#)
4. [脸颊上的诱惑——清晰语音从下颚骨开始](#)
5. [A noise cross PSD estimator for dual-microphone speech enhancement based on minimum stat](#)
6. [基于麦克风阵列的语音增强算法概述](#)
7. [基于听觉感知特性的信号子空间麦克风阵列语音增强算法](#)
8. [UTalk中快速禁用麦克风](#)
9. [基于听觉感知特性的信号子空间麦克风阵列语音增强算法](#)
10. [数字滤波的滑动平均法和低通滤波法](#)