

基于统计模型的麦克风阵列语音增强算法

涂井先 冀占江 覃桂荏 蒲保兴

(梧州学院大数据与软件工程学院 广西 梧州 543002)

摘要 提出一种基于统计模型的麦克风阵列语音增强算法。为了估计第一通道语音信号谱幅度平方,假设第一通道语音信号谱的实部和虚部相互独立,并服从方差相同的高斯分布。先利用贝叶斯公式计算第一通道语音信号谱幅度平方的后验概率,再利用数学期望的计算公式得到第一通道语音信号谱幅度平方的最小均方误差估计,最后用第一通道观测语音信号的谱角度来估计第一通道语音信号的谱角度。实验结果表明,所提出的算法优于三种传统的语音增强算法。

关键词 语音增强 最小均方误差估计 去噪性能

中图分类号 TP302.7

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.11.046

MICROPHONE ARRAY SPEECH ENHANCEMENT ALGORITHM BASED ON STATISTICAL MODELS

Tu Jingxian Ji Zhanjiang Qin Guijian Pu Baoxing

(School of Data Science & Software Engineering, Wuzhou University, Wuzhou 543002, Guangxi, China)

Abstract This paper presents a microphone array speech enhancement algorithm based on statistical models. In order to estimate the magnitude-squared spectrum (MSS) of the first channel speech signal, it was assumed that the real part and imaginary part of the first channel speech signal spectrum were independent of each other and obeyed the Gaussian distribution with the same variance. The posterior probability of the MSS of the first channel speech signal was calculated by using the Bayesian formula, and the minimum mean square error (MMSE) estimation of the MSS of the first channel speech signal was obtained by using the mathematical expectation formula, and finally the spectrum angle of the first channel speech signal was estimated by using the spectrum angle of the first channel observed signal. The experimental results show that the proposed algorithm is superior to the three traditional speech enhancement algorithms.

Keywords Speech enhancement Minimum mean square error estimation Denoising performance

0 引言

麦克风阵列语音增强算法是语音信号领域的研究热点^[1-3]。常见的麦克风阵列语音增强算法有维纳滤波算法^[4-5]、子空间算法^[2]、基于概率模型的算法^[7]和卡尔曼滤波算法^[8]等。本文研究基于概率模型的麦克风阵列语音增强算法。

根据麦克风的几何分布,多通道麦克风可划分为麦克风阵列和分布式麦克风^[9]。在麦克风阵列中麦克风之间的间距较小,麦克风之间的相对位置是固定的。而在分布式麦克风中,麦克风之间的间距较大,麦克风

之间的相对位置是可变的。文献[7]提出了基于谱幅度平方估计的分布式麦克风语音增强算法,该算法研究的语音信号模型是无房间混响的。本文对文献[7]的算法进行推广和改进,提出统计模型的麦克风阵列语音增强算法。本文提出的算法研究的语音模型是有房间混响的。实验结果证实了本文提出的算法的有效性。

1 算法设计

1.1 信号模型

在会议室中,演讲者借助 M 个通道的麦克风阵列说话。此时,麦克风阵列接收的语音信号可表示为:

$$y_i(n) = g_i(n) * s(n) + v_i(n) = x_i(n) + v_i(n) \quad (1)$$

$$i = 1, 2, \dots, M$$

式中: * 是卷积, $s(n)$ 是源语音信号, $g_i(n)$ 是源语音信号到第 i 通道麦克风的房间混响, $x_i(n)$ 、 $v_i(n)$ 、 $y_i(n)$ 分别表示第 i 通道语音信号, 噪声信号和接收语音信号。 $v_i(n)$ 和 $s(n)$ 相互独立。利用离散傅里叶变换 (Discrete Fourier Transform, DFT), 式(1)可转化为:

$$Y_i(k, l) = G_i(k, l)S(k, l) + V_i(k, l) =$$

$$X_i(k, l) + V_i(k, l) =$$

$$G_{i1}(k, l)X_1(k, l) + V_i(k, l) \quad (2)$$

式中: k 和 l 分别表示频点标号和帧标号, $Y_i(k, l)$ 、 $G_i(k, l)$ 、 $S(k, l)$ 、 $X_i(k, l)$ 和 $V_i(k, l)$ 分别表示 $y_i(n)$ 、 $g_i(n)$ 、 $s(n)$ 、 $x_i(n)$ 和 $v_i(n)$ 的离散傅里叶变换系数, $G_{i1}(k, l) = \frac{G_i(k, l)}{G_1(k, l)}$ 表示第 i 通道和第 1 通道之间房间混响的比值。利用欧拉公式, 式(2)可写成:

$$\bar{Y}_i(k, l) e^{j\theta_{Y_i}(k, l)} = \bar{G}_{i1}(k, l) \bar{X}_1(k, l) e^{j(\theta_{X_1}(k, l) + \theta_{G_{i1}}(k, l))} +$$

$$\bar{V}_i(k, l) e^{j\theta_{V_i}(k, l)} \quad (3)$$

式中: $\bar{Y}_i(k, l)$ 、 $\bar{G}_{i1}(k, l)$ 、 $\bar{X}_1(k, l)$ 、 $\bar{V}_i(k, l)$ 分别表示 $Y_i(k, l)$ 、 $G_{i1}(k, l)$ 、 $S(k, l)$ 、 $X_i(k, l)$ 的谱幅度, $\theta_{Y_i}(k, l)$ 、 $\theta_{G_{i1}}(k, l)$ 、 $\theta_{X_1}(k, l)$ 、 $\theta_{V_i}(k, l)$ 分别表示 $Y_i(k, l)$ 、 $G_{i1}(k, l)$ 、 $S(k, l)$ 、 $X_i(k, l)$ 的谱角度。为了表达简便, 除了特殊说明外, 在接下来的算式中我们去掉 k 和 l 。本文考虑通过 $\{Y_i\}_{i=1}^M$ 或者 $\{\bar{Y}_i, \theta_{Y_i}\}_{i=1}^M$ 恢复 X_1 或者 $\{\bar{X}_1, \theta_{X_1}\}$ 。若第一通道信号不是目标信号, 需要交换目标信号与第一通道信号。记 $\xi_i = \frac{\sigma_{X_1}^2}{\sigma_{V_i}^2}$ 和 $\gamma_i \triangleq \frac{|Y_i|^2}{\sigma_{V_i}^2}$ 分别为第 i 通道先验信噪比和后验信噪比, 其中 $\sigma_{X_1}^2$ 和 $\sigma_{V_i}^2$ 分别表示 X_1 和 V_i 的方差。

1.2 新估计

为了推导出估计器, 需要做如下假设:

- (1) 各个通道语音信号与噪声信号是相互独立的。
- (2) 各个通道噪声信号是相互独立的。
- (3) 第一通道语音信号谱和噪声信号谱的实部和虚部都服从高斯分布, 它们相互独立, 方差相同。
- (4) 各个通道的观测语音信号的谱幅度平方等于语音成分的谱幅度平方和噪声信号的谱幅度平方之和, 即 $\bar{Y}_i^2 = \bar{G}_{i1}^2 \bar{X}_1^2 + \bar{V}_i^2$ 。

由第 3 个假设, 则第一通道语音信号和噪声信号谱幅度平方分别服从参数为 $1/\sigma_{X_1}^2$ 的指数分布和参数为 $1/\sigma_{V_1}^2$ 的指数分布。因此, \bar{X}_1^2 和 \bar{V}_1^2 的概率密度函数可表示为:

$$f_{\bar{X}_1^2}(\bar{X}_1^2) = \frac{1}{\sigma_{X_1}^2} e^{-\frac{\bar{X}_1^2}{\sigma_{X_1}^2}} \quad (4)$$

$$f_{\bar{V}_i^2}(\bar{V}_i^2) = \frac{1}{\sigma_{V_i}^2} e^{-\frac{\bar{V}_i^2}{\sigma_{V_i}^2}} \quad (5)$$

由第 2 个假设, $\bar{V}_1^2, \bar{V}_2^2, \dots, \bar{V}_M^2$ 的联合密度函数可表示为:

$$f_{\bar{V}_1^2, \bar{V}_2^2, \dots, \bar{V}_M^2}(\bar{V}_1^2, \bar{V}_2^2, \dots, \bar{V}_M^2) = \frac{1}{\sigma_{V_i}^2} e^{-\sum_{i=1}^M \frac{\bar{V}_i^2}{\sigma_{V_i}^2}} \prod_{i=1}^M \sigma_{V_i}^2 \quad (6)$$

由第 4 个假设, $\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2$ 的条件联合密度函数可表示为:

$$f_{\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2}(\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2) = \frac{1}{\sigma_{V_i}^2} e^{-\sum_{i=1}^M \frac{\bar{Y}_i^2 - \bar{G}_{i1}^2 \bar{X}_1^2}{\sigma_{V_i}^2}} \prod_{i=1}^M \sigma_{V_i}^2 \quad (7)$$

定理 1 由式(4)和式(7), 推导出 \bar{X}_1^2 的后验概率密度函数为:

$$f_{\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2}(\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2) = \begin{cases} \bar{\Psi} e^{-\frac{\bar{X}_1^2}{\lambda}} \frac{1}{\sigma_{X_1}^2} \neq \sum_{i=1}^M \frac{\bar{G}_{i1}^2}{\sigma_{V_i}^2} \\ \frac{\bar{G}_{k1}^2}{\bar{Y}_k^2} \quad \text{其他} \end{cases} \quad (8)$$

式中:

$$k = \arg \min_i \{ \bar{Y}_i^2 / \bar{G}_{i1}^2 \mid i = 1, 2, \dots, M \} \quad (9)$$

$$\bar{\lambda} = 1 / \left(\frac{1}{\sigma_{X_1}^2} - \sum_{i=1}^M \frac{\bar{G}_{i1}^2}{\sigma_{V_i}^2} \right) \quad (10)$$

$$\bar{\Psi} = 1 / (\bar{\lambda} - \bar{\lambda} e^{-\frac{\bar{Y}_k^2}{\bar{G}_{k1}^2 \bar{\lambda}}}) \quad (11)$$

证明: 由式(4)和式(7), 得到:

$$f_{\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2}(\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2) =$$

$$\frac{f_{\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2}(\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2) f_{\bar{X}_1^2}(\bar{X}_1^2)}{\int_0^{\bar{Y}_k^2 / \bar{G}_{k1}^2} f_{\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2}(\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2) f_{\bar{X}_1^2}(\bar{X}_1^2) d\bar{X}_1^2} =$$

$$\frac{f_{\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2}(\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2)}{f_{\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2}(\bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2 | \bar{X}_1^2) f_{\bar{X}_1^2}(\bar{X}_1^2)} =$$

$$\frac{\frac{1}{\sigma_{V_i}^2} e^{-\sum_{i=1}^M \frac{\bar{Y}_i^2 - \bar{G}_{i1}^2 \bar{X}_1^2}{\sigma_{V_i}^2}} \frac{1}{\sigma_{X_1}^2} e^{-\frac{\bar{X}_1^2}{\sigma_{X_1}^2}}}{\int_0^{\bar{Y}_k^2 / \bar{G}_{k1}^2} \frac{1}{\sigma_{V_i}^2} e^{-\sum_{i=1}^M \frac{\bar{Y}_i^2 - \bar{G}_{i1}^2 \bar{X}_1^2}{\sigma_{V_i}^2}} \frac{1}{\sigma_{X_1}^2} e^{-\frac{\bar{X}_1^2}{\sigma_{X_1}^2}} d\bar{X}_1^2} =$$

$$\frac{e^{-\left(\frac{1}{\sigma_{X_1}^2} - \sum_{i=1}^M \frac{\bar{G}_{i1}^2}{\sigma_{V_i}^2}\right) \bar{X}_1^2}}{\int_0^{\bar{Y}_k^2 / \bar{G}_{k1}^2} e^{-\left(\frac{1}{\sigma_{X_1}^2} - \sum_{i=1}^M \frac{\bar{G}_{i1}^2}{\sigma_{V_i}^2}\right) \bar{X}_1^2} d\bar{X}_1^2} = \begin{cases} \bar{\Psi} e^{-\frac{\bar{X}_1^2}{\lambda}} \frac{1}{\sigma_{X_1}^2} \neq \sum_{i=1}^M \frac{\bar{G}_{i1}^2}{\sigma_{V_i}^2} \\ \frac{\bar{G}_{k1}^2}{\bar{Y}_k^2} \quad \text{其他} \end{cases}$$

定理 2 由式(4)和式(8),推导出 \bar{X}_1^2 的 MMSE-MSS 估计:

$$\bar{X}_{1,\text{MMSE}}^2 = E[\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2] = \begin{cases} \left(\frac{1}{w_k} - \frac{1}{e^{w_k} - 1} \right) \frac{\bar{Y}_k^2}{\bar{G}_{i1}^2} \frac{1}{\sigma_{X_1}^2} \neq \sum_{i=1}^M \frac{\bar{G}_{i1}^2}{\sigma_{V_i}^2} \\ \frac{\bar{Y}_k^2}{2 \bar{G}_{i1}^2} & \text{其他} \end{cases} \quad (12)$$

其中:

$$w_k = \left(1 - \sum_{i=1}^M \xi_i \right) \gamma_k / \xi_k \quad (13)$$

证明:当 $1/\sigma_{X_1}^2 \neq \sum_{i=1}^M \bar{G}_{i1}^2/\sigma_{V_i}^2$ 时,由连续随机变量数学期望的定义和式(8),得到:

$$\begin{aligned} \bar{X}_{1,\text{MMSE}}^2 &= E[\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2] = \\ &\int_0^{\bar{Y}_k^2/\bar{G}_{k1}^2} \bar{X}_1^2 f_{\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2}(\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2) d\bar{X}_1^2 = \\ &\int_0^{\bar{Y}_k^2/\bar{G}_{k1}^2} \bar{X}_1^2 \Psi e^{-\frac{\bar{X}_1^2}{\lambda}} d\bar{X}_1^2 = \\ &\int_0^{\bar{Y}_k^2/\bar{G}_{k1}^2} \lambda \Psi e^{-\frac{\bar{X}_1^2}{\lambda}} d\bar{X}_1^2 - \lambda \Psi e^{-\frac{\bar{X}_1^2}{\lambda}} \frac{\bar{Y}_k^2}{\bar{G}_{k1}^2} = \\ &\lambda - 1/(e^{\frac{\bar{Y}_k^2}{\bar{G}_{k1}^2 \lambda}} - 1) \frac{\bar{Y}_k^2}{\bar{G}_{k1}^2} \end{aligned} \quad (14)$$

由式(10)和式(13), $\bar{\lambda}$ 可写成:

$$\bar{\lambda} = \frac{\bar{Y}_k^2}{\bar{G}_{k1}^2 w_k} \quad (15)$$

将式(15)代入式(14),推出:

$$\bar{X}_{1,\text{MMSE}}^2 = \left(\frac{1}{w_k} - \frac{1}{e^{w_k} - 1} \right) \frac{\bar{Y}_k^2}{\bar{G}_{k1}^2} \quad (16)$$

当 $1/\sigma_{X_1}^2 = \sum_{i=1}^M \bar{G}_{i1}^2/\sigma_{V_i}^2$ 时,

$$\begin{aligned} \bar{X}_{1,\text{MMSE}}^2 &= E[\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2] = \\ &\int_0^{\bar{Y}_k^2/\bar{G}_{k1}^2} \bar{X}_1^2 f_{\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2}(\bar{X}_1^2 | \bar{Y}_1^2, \bar{Y}_2^2, \dots, \bar{Y}_M^2) d\bar{X}_1^2 = \\ &\int_0^{\bar{Y}_k^2/\bar{G}_{k1}^2} \bar{X}_1^2 \frac{\bar{G}_{k1}^2}{\bar{Y}_k^2} d\bar{X}_1^2 = \frac{\bar{Y}_k^2}{2 \bar{G}_{k1}^2} \end{aligned} \quad (17)$$

综上所述,定理 2 成立。

本文用第一通道观测语音信号的谱角度来估计第一通道语音信号的谱角度,即:

$$\hat{\theta}_{X_1} = \theta_{Y_1} \quad (18)$$

2 参数估计

本节将介绍相关参数的估计方法。第 i 通道噪声

信号谱方差 $\sigma_{V_i}^2$ 是采用最小值控制递归平均法 (Minima Controlled Recursive Averaging, MCRA)^[10] 来更新,第 i 通道先验信噪比 ξ_i 是采用直接决策 (Decision-Directed, DD) 方法来更新。

$$\hat{\xi}_i = \alpha \hat{G}_{i1}^2 \frac{\bar{X}_1^2}{\sigma_{V_i}^2} + (1 - \alpha) \max(\gamma_i - 1, 0) \quad (19)$$

式中: $\alpha \in (0, 1)$ 是平滑因子。第 i 通道和第 1 通道之间房间混响的比值的谱幅度 \bar{G}_{i1} 通过式(20)来更新。

$$\hat{G}_{i1} = \sqrt{\frac{\sigma_{X_i}^2}{\sigma_{X_1}^2}} = \sqrt{\frac{\sigma_{Y_i}^2 - \sigma_{V_i}^2}{\sigma_{Y_1}^2 - \sigma_{V_1}^2}} \quad (20)$$

式中: $\beta \in (0, 1)$ 是平滑因子, $\sigma_{Y_i}^2$ 表示 Y_i 的方差。在只含有噪声的信号帧, $\sigma_{V_i}^2$ 保持不变。而在含有语音成分的信号帧中, $\sigma_{V_i}^2$ 通过式(21)来更新。

$$\sigma_{V_i}^2 = \gamma \sigma_{V_i}^2 + (1 - \gamma) \bar{Y}_i^2 \quad (21)$$

式中: $\gamma \in (0, 1)$ 是平滑因子。

所提出的算法采用传统的基于短时能量的方法进行语音活动检测 (Voice Activity Detection, VAD)。

3 算法描述

算法的输入是 M 个通道的观测信号,算法输出是估计的第 1 个通道语音信号。下面给出所提出的算法在处理每帧信号的主要步骤:

步骤 1 采用离散傅里叶变换将时域的观测信号转化为频域的信号。

步骤 2 采用最小值控制递归平均法计算 $\sigma_{V_i}^2$ 。

步骤 3 用传统的基于短时能量的方法进行语音活动检测。

步骤 4 分别通过式(19)、式(20)、式(21)计算 ξ_i 、 \bar{G}_{i1} 、 $\sigma_{V_i}^2$ 。

步骤 5 通过式(12)计算 \bar{X}_1^2 。

步骤 6 通过式(18)计算 θ_{X_1} 。

步骤 7 计算 $X_1 = \bar{X}_1 e^{j\theta_{X_1}}$ 。

步骤 8 对频域的信号进行离散傅里叶逆变换 (Inverse Discrete Fourier Transform, IDFT)。

4 实验

在这节,我们通过一些仿真例子来比较本文提出的算法、DM-MMSE-MSS 算法^[7]、SDW-MWF 算法^[4]和 R1-SDW-MWF 算法^[5]。仿真实验在 MATLAB R2016a 平台上进行。

4.1 实验环境

在仿真环境中,房间的尺寸为长 6.7 m、宽 6.1 m、

高 2.9 m。考虑带有 10 个全方向麦克风和相邻麦克风间距为 0.1 m 的线性均匀麦克风阵列。为了方便描述,以房间的长、宽、高为 x 、 y 、 z 轴建立空间直角坐标系。图 1 给出了麦克风阵列和源语音信号的位置的平面图。源语音信号位于 $(3.35, 1.0, 1.4)$, 10 个麦克风分别位于 $(2.85 + 0.1(i-1), 1.1, 1.4)$ 处, 其中 $i = 1, 2, \dots, 10$ 。根据各个麦克风与源语音信号的距离, 从小到大对各个通道的语音信号进行排序。第一通道观测语音信号的麦克风离源语音信号是最近的, 本文的目标是利用多个通道的观测语音信号来恢复第一通道语音成分。随机从 NOIZEUS 语音库^[11] 挑出 20 个语音句子拼接成时长约为 50 秒的源语音信号。从 NOISEX 噪声库^[12] 选取噪声。通过镜像方法^[13] 生成房间混响。将房间混响卷积作用于源语音信号生成每个通道的语音成分, 然后根据输入信噪比来确定噪声信号的权重, 最后每个通道的语音信号成分加上乘以权重后的噪声信号, 生成每个通道的观测语音信号。本文考虑两种噪声 (factory 噪声和 airport 噪声), 两个输入信噪比水平 (-5 dB 和 5 dB) 和混响时间为 480 ms 各种组合情形。

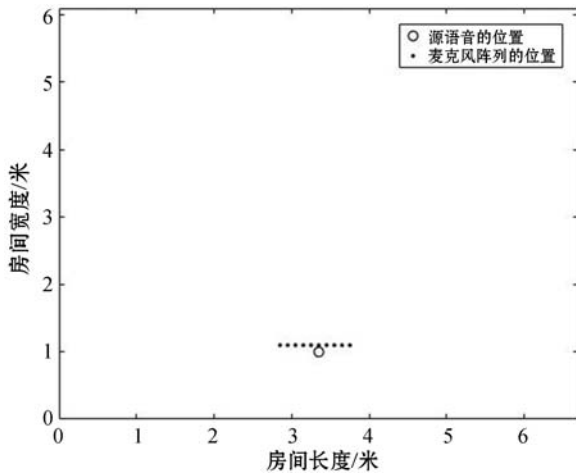


图1 麦克风阵列和源语音信号的位置的平面图

为有效地评价算法, 四种算法部分实验参数统一设定为帧长为 256 (32 ms), 重叠率为 50%, 傅里叶变换点数为 512, 平滑因子为 0.98。语音活动检测方法采用传统的基于短时能量的方法。

4.2 实验结果

采用分段信噪比 (Segmental Signal to Noise Ratio, SSNR) 增量, 对数似然比 (Log-likelihood Ratio, LLR) 和语音质量感知评价 (Perceptual Evaluation of Speech Quality, PESQ) 作为算法去噪性能的评价指标^[14]。SSNR 越大, 表示去噪量较多。LLR 越小, 表示语音失真较少。PESQ 是衡量噪声减少和语音失真综合的指标, 它越大, 表示去噪性能越好。

图 2 - 图 13 给出了在 factory 噪声和 airport 噪声, 输入信噪比为 -5 dB 和 5 dB 以及混响时间为 480 ms 各种组合情形下 4 种算法的 SSNR 增量、LLR 和 PESQ 的实验结果。可以看出, 4 种算法的去噪性能从好到差的排序为本文提出的算法、DM-MMSE-MSS 算法、R1-SDW-MWF 算法和 SDW-MWF 算法。本文提出的算法和 DM-MMSE-MSS 算法的三个评价指标明显优于 R1-SDW-MWF 算法和 SDW-MWF 算法。在 airport 噪声和输入信噪比为 5 dB 情形下, 本文提出的算法在 SSNR 增量、LLR 和 PESQ 方面都优于其他三种算法。在 factory 噪声和各种输入信噪比组合情形下, 本文提出的算法在 SSNR 增量和 PESQ 方面优于其他三种算法, 在 LLR 方面劣于 DM-MMSE-MSS 算法, 但优于 R1-SDW-MWF 算法和 SDW-MWF 算法。在 airport 噪声和输入信噪比为 -5 dB 情形下, 本文提出的算法在 LLR 方面优于其他三种算法, 在 SSNR 方面跟 DM-MMSE-MSS 算法差不多, 但优于 R1-SDW-MWF 算法和 SDW-MWF 算法, 在 PESQ 方面略劣于 DM-MMSE-MSS 算法, 但优于 R1-SDW-MWF 算法和 SDW-MWF 算法。

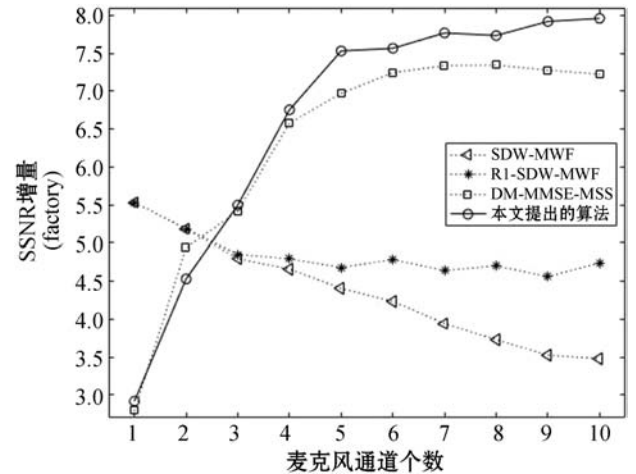


图2 factory 噪声和 -5 dB 输入信噪比情形下 4 种算法的 SSNR 增量

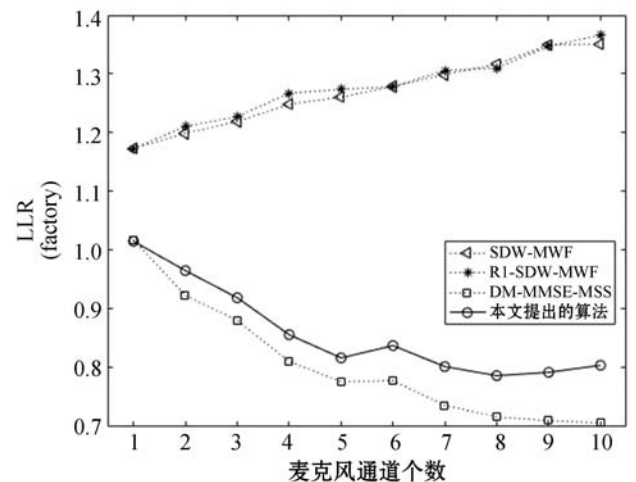


图3 factory 噪声和 -5 dB 输入信噪比情形下 4 种算法的 LLR

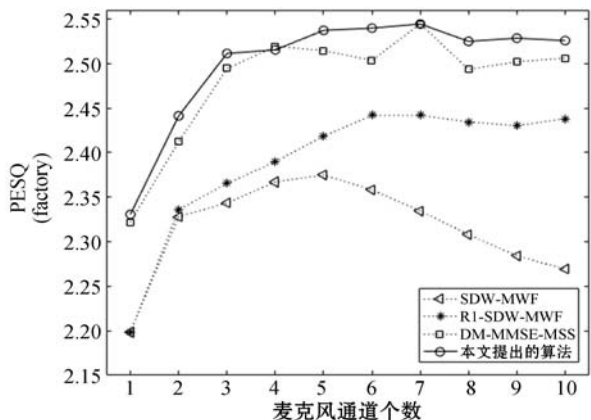


图 4 factory 噪声和 -5 dB 输入信噪比情形下 4 种算法的 PESQ

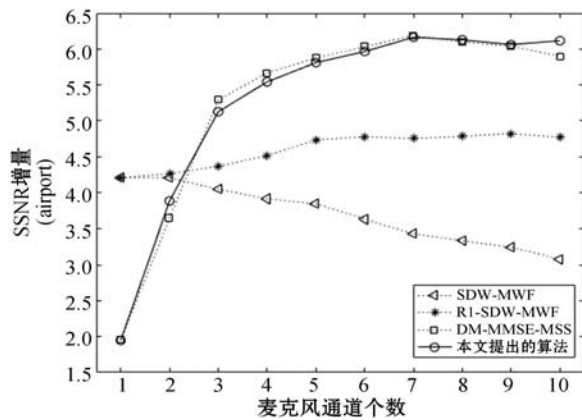


图 8 airport 噪声和 -5 dB 输入信噪比情形下 4 种算法的 SSNR 增量

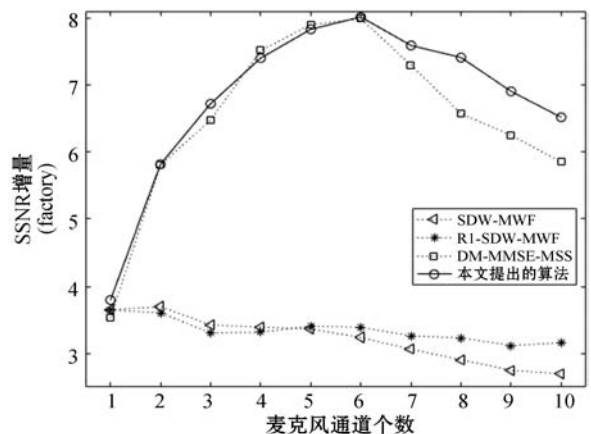


图 5 factory 噪声和 5 dB 输入信噪比情形下 4 种算法的 SSNR 增量

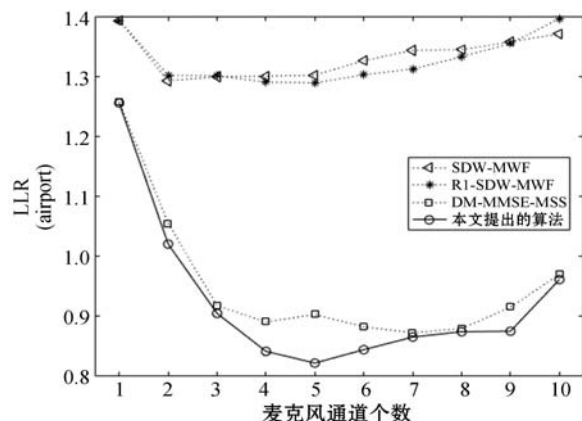


图 9 airport 噪声和 -5 dB 输入信噪比情形下 4 种算法的 LLR

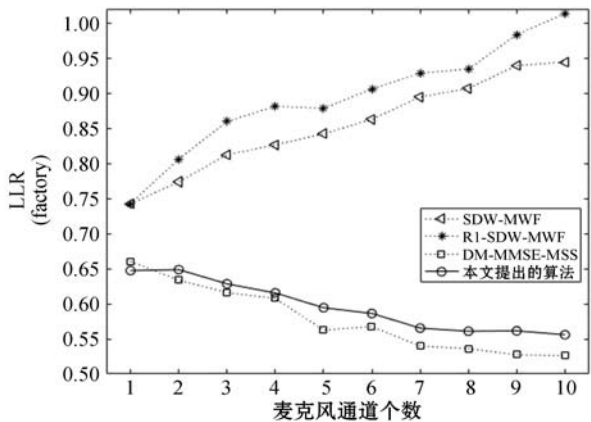


图 6 factory 噪声和 5 dB 输入信噪比情形下 4 种算法的 LLR

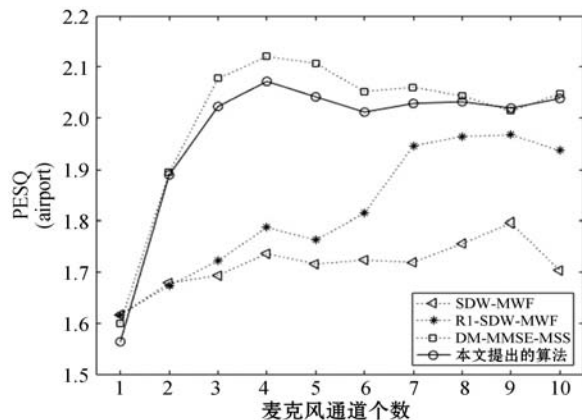


图 10 airport 噪声和 -5 dB 输入信噪比情形下 4 种算法的 PESQ

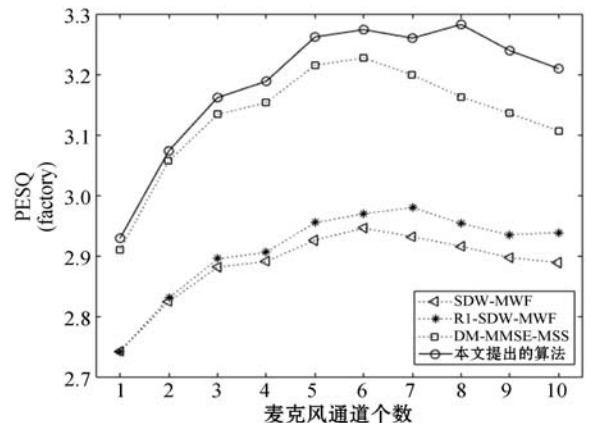


图 7 factory 噪声和 5 dB 输入信噪比情形下 4 种算法的 PESQ

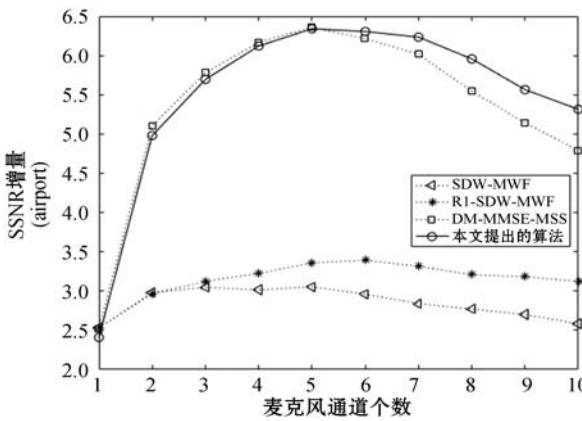


图 11 airport 噪声和 5 dB 输入信噪比情形下 4 种算法的 SSNR 增量

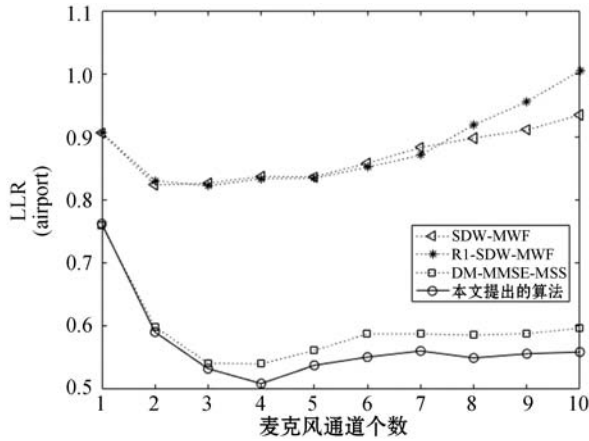


图 12 airport 噪声和 5 dB 输入信噪比情形下 4 种算法的 LLR

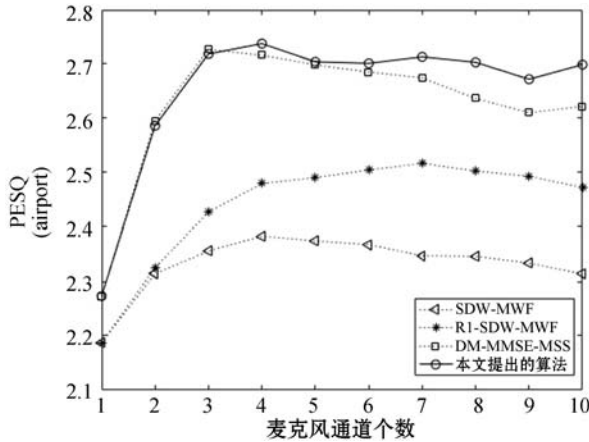


图 13 airport 噪声和 5 dB 输入信噪比情形下 4 种算法的 PESQ

5 结 语

本文提出了基于统计模型的麦克风阵列语音增强算法,算法考虑有混响的语音信号模型。在各个通道的观测语音信号的谱幅度平方等于语音成分的谱幅度平方和噪声信号的谱幅度平方之和的假设下,推导出语音成分谱幅度平方的最小均方误差估计。跟文献[7]提出的算法相比,本文算法适用范围更广,更符合实际环境。仿真实验结果表明,本文提出的算法在三种评价指标下优于传统的多通道语音增强算法。

参 考 文 献

- [1] Benesty J, Chen J D, Huang Y T. Microphone array signal processing[M]. Berlin Heidelberg: Springer-Verlag, 2008.
- [2] Huang Y T, Benesty J, Chen J D. Analysis and comparison of multichannel noise reduction methods in a common framework[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2008, 16(5): 957-968.
- [3] 倪峰, 周翊, 刘宏清. 利用信号功率谱密度的麦克风阵列噪声消除方法[J]. 信号处理, 2020, 36(3): 373-381.
- [4] Doclo S, Spriet A, Wouters J, et al. Frequency-domain criterion for the speech distortion weighted multichannel wiener

filter for robust noise reduction[J]. Speech Communication, 2007, 49(7/8): 636-656.

- [5] Serizel R, Moonen B, Dijk M, et al. Rank-1 approximation based multichannel wiener filter algorithms for noise reduction in cochlear implants[C]//38th IEEE International Conference on Acoustics, Speech and Signal Processing, 2013: 8634-8938.
- [6] Serizel R, Moonen M, Dijk B V, et al. Low-rank approximation based multichannel wiener filter algorithms for noise reduction with application in cochlear implants[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2014, 22(4): 785-798.
- [7] Tu J X, Xia Y S. Fast distributed multichannel speech enhancement using novel frequency domain estimators of magnitude-squared spectrum[J]. Speech Communication, 2015, 72: 96-108.
- [8] Tu J X, Xia Y S. Effective Kalman filtering algorithm for distributed multichannel speech enhancement[J]. Neurocomputing, 2018, 275: 144-154.
- [9] Trawicki M B, Johnson M T. Distributed multichannel speech enhancement with minimum mean-square error short-time spectral amplitude, log-spectral amplitude, and spectral phase estimation[J]. Signal Processing, 2012, 92(2): 345-356.
- [10] Cohen I, Berdugo B. Noise estimation by minima controlled recursive averaging for robust speech enhancement[J]. IEEE Signal Processing Letters, 2002, 9(1): 12-15.
- [11] Loizou P C. Speech enhancement: Theory and practice[M]. Boca: CRC Press, 2007.
- [12] Varga A, Steeneken H J M. Assessment for automatic speech recognition: II, NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems[J]. Speech Communication, 1993, 12: 247-251.
- [13] Allen J B, Berkley D A. Image method for efficiently simulating small-room acoustic[J]. The Journal of the Acoustical Society of America, 1979, 65: 943-950.
- [14] Hu Y, Loizou P C. Evaluation of objective quality measures for speech enhancement[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2008, 16(1): 229-238.

(上接第 272 页)

- [12] 余波, 台亮青, 马治杰. 云计算环境下基于属性和信任的 RBAC 模型研究[J]. 计算机工程与应用, 2020, 56(9): 84-92.
- [13] 吴绍涵, 胡卫. 基于 SM2 算法和区块链的移动端身份认证协议设计[J]. 网络与信息安全学报, 2018, 4(9): 60-65.