

一种改进的子空间语音增强方法

王文杰, 王霞, 王国君, 佟强

(河北工业大学 信息工程学院, 天津 300401)

摘要: 提出一种基于奇异值分解的子空间分解语音增强方法, 该方法是利用最小值统计噪声估计法代替传统的 VAD 方法对噪声进行估计, 并利用所得噪声和带噪语音构造的协方差矩阵得到纯净语音的协方差矩阵, 并将特征值分解的时域约束和频域约束估计方法推广到奇异值分解方法中, 通过奇异值分解、重构得到增强后的语音信息。试验表明: 该方法具有较好的去噪效果。

关键词: 奇异值分解; 噪声估计; 最小统计; 信号子空间

中图分类号: TN912.3

文献标识码: A

文章编号: 1674-6236(2010)06-0127-03

New improved subspace method of speech enhancement

WANG Wen-jie, WANG Xia, WANG Guo-jun, TONG Qiang

(School of Information Engineering, Hebei University of Technology, Tianjin 300401, China)

Abstract: This paper presented a kind of singular value decomposition based on subspace decomposition method of speech enhancement. Using the minimum statistics noise estimation method instead of the traditional VAD method to estimate the noise and use the proceeds of the noise and noisy speech covariance matrix to be constructed pure speech covariance matrix. Eigen value decomposition of the time-domain constraints and frequency domain constraints estimation method are extended to the singular value decomposition method. Through the singular value decomposition, and reconstruction of post-enhanced voice messages, according to tests show that this method has a good de-noising effect.

Key words: singularity value decompose; noise estimate; minimum statistics; signal subspace

随着子空间技术的发展, 子空间方法广泛应用于语音增强领域。子空间法由于其本身具有控制信号失真和残留噪声的平衡机制, 并在消除音乐噪声方面有很好的效果, 因此采用各种改进的子空间方法实现语音增强成为趋势。

Ephraim 和 Van Trees^[1]首先提出白噪声条件下子空间语音增强的算法。Rezayee 和 Gazor^[2]通过对噪声向量的协方差矩阵近似对角化, 将此算法推广到有色噪声; Lev-Ari 和 Ephraim^[3]通过对语音向量和噪声向量的协方差矩阵的联合特征分解, 提出一种基于子空间算法的最优估计; Jabloun 等^[4]又将声学感知模型引入到子空间语音增强中, 利用掩蔽门限在特征域和频域的变换, 对子空间降噪的信号进行噪声滤波。

子空间法在分解的过程中主要包括奇异值分解(SVD)和特征值分解(EVD)两种方法^[5]。奇异值分解主要是对带噪语音的空间进行奇异值分解, 大的奇异值所对应的特征向量包含语音信号信息; 小的奇异值所对应的特征向量包含噪声信息。如果忽略小奇异值及其所对应的特征向量, 那么降维恢复后的信号就是估计的语音信号。特征值分解是利用对含噪子空间的特征值分解, 将分解得到的特征值特征向量分别处

理, 非零特征向量对应信号子空间, 零特征向量对应噪声子空间。只保留非零特征值所对应的特征向量, 由此重构出纯净的语音信号。

1 奇异值分解的基本原理

传统的 SVD 降噪基本原理^[6]:

1) 构造带噪语音的 Hankel 矩阵 Y 。

2) 对 Y 进行奇异值分解, 则 $Y=U\Sigma V^H$, 其中 U 和 V 分别为 $m \times n$, $n \times n$ 维的酉阵, Σ 是一个 $m \times n$ 维的对角阵, 其主对角线上的元素 λ_k 称为 Y 的奇异值。

3) 确定奇异值对角阵 Σ 的有效秩, 即前 P 个最大的奇异值, 然后按式(1)重构 Y 的逼近矩阵 Y_p , 即不受噪声干扰的信号分量所贡献的矩阵。

$$Y_p = [U_p U_{m-p}] \begin{bmatrix} \Sigma_p & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_p^T \\ V_{m-p}^T \end{bmatrix} = U_p \Sigma_p V_p^T \quad (1)$$

式中, U_p, V_p 分别是前 P 个主奇异值对应的左右奇异向量, Σ_p 是主奇异值对应的对角阵, 即

$$\Sigma_p = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p \end{bmatrix} \quad (2)$$

4) 通常 Y_p 不再是一个 Hankel 矩阵, 可由 Y_p 按式(3)构

收稿日期: 2009-10-13

稿件编号: 200910038

作者简介: 王文杰(1984—), 男, 河北宣化人, 硕士研究生。研究方向: 信号与信息处理。

造 Hankel 矩阵 \hat{S} 。

$$\hat{S} = \begin{bmatrix} \hat{s}[0] & \hat{s}[1] & \cdots & \hat{s}[M-1] \\ \hat{s}[1] & \hat{s}[2] & \cdots & \hat{s}[M] \\ \vdots & \vdots & \ddots & \vdots \\ \hat{s}[L-1] & \hat{s}[L] & \cdots & \hat{s}[N-1] \end{bmatrix} \quad (3)$$

$$\text{式中, } \hat{s}[k] = \frac{1}{\beta - \alpha - 1} \sum_{i=\alpha}^{\beta} Y_P(k-i+2, i) \quad (k=0, 1, \cdots, N-1), \quad (4)$$

$$\alpha = \max(1, k-L+2), \beta = \min(M, k+1) \quad (5)$$

同时 $\hat{s}[k]$ 就作为真实信号的一种估计, 即降噪后信号。

2 最小统计噪声估计法

Martin^[7-8]提出一种基于带噪语音能量的最小统计量进行噪声功率谱估计的算法, 该方法不需要 VAD, 在一段连续的时间窗内, 通过跟踪带噪语音的特征值的最小值来得到噪声的估计和更新。

$$R_Y = U \Lambda_Y U^T \quad (6)$$

$$\Lambda_N = E(U^T N N^T U) = U^T R_N U \quad (7)$$

由式(6)和式(7)以及语音和噪声不相关的假设, 即 $R_Y = R_X + R_N$, 则有

$$R_Y = U(\Lambda_X + \Lambda_N)U^T = U \Lambda_Y U^T \quad (8)$$

式中, $\Lambda_Y = \text{diag}(\lambda_Y^1, \lambda_Y^2, \cdots, \lambda_Y^K)$ 为 R_Y 的特征值组成的对角阵。

式(8)表明, 带噪语音的特征值 Λ_Y 等于纯净语音 Λ_X 和 Λ_N 的和。因此, 具体到第 n 帧第 i 个特征向量对应的特征值上, 当不含语音时, 带噪语音 $\lambda_Y^i(n)$ 即衰减到 $\lambda_N^i(n)$ 的水平, 即有

$$\lambda_Y^i(n) = \begin{cases} \lambda_X^i(n) + \lambda_N^i(n) & \text{含有语音时} \\ \lambda_N^i(n) & \text{不含语音时} \end{cases} \quad (9)$$

因此可沿着每一个特征向量跟踪 $\lambda_Y^i(n)$ 的最小值来获得 $\lambda_N^i(n)$ 的估计。

3 改进方法

为了更好的去噪效果, 将最小值统计噪声估计法运用到基于奇异值分解的子空间法中, 同时将特征值分解法中的时域约束和频域约束方法推广到奇异值分解中, 从而在得到语音信号矩阵时更加精确。方法步骤如下:

1) 定义并设置阈值等相关参数。

2) 利用最小统计噪声估计法求每一帧噪声的协方差矩阵, 其原理步骤如下:

①特征值平滑

首先, 在每个特征向量上对带噪语音 R_Y 的特征值 $\lambda_Y^i(n)$ 进行一阶平滑

$$\lambda(i, n) = \alpha(i, n)\lambda(i, n-1) + (1-\alpha(i, n))\lambda_Y^i(n) \quad (10)$$

式中, $\lambda(i, n)$ 为第 i 个特征向量上的平滑特征值, n 为当前帧; 参数 $\alpha(i, n)$ 为平滑因子, 取值在 0~1 之间。

根据噪声在时间帧上的前后相关性, 定义

$$\alpha(i, n) = 1 - \min\left\{1, \left[\frac{\lambda_Y^i(n)}{\lambda_N^i(n)(n-1)}\right]^Q\right\} \quad (11)$$

式中, Q 为整数, $\lambda_N^i(n)(n-1)$ 表示此前 10~20 帧噪声估计的平均值

$$\bar{\lambda}_N^i(n)(n-1) = \frac{1}{10} \sum_{p=1}^{10} \lambda_N^i(n-p) \quad (12)$$

Q 控制 $\frac{\lambda_Y^i(n)}{\lambda_N^i(n)(n-1)}$ 变化的快慢, 实验值选取 $3 \leq Q \leq 5$,

有助于避免噪声的过估计。平滑因子 $\alpha(i, n)$ 的作用方式如下: 如果当前帧 $\lambda_N^i(n)$ 远远大于此前帧噪声的平均值 $\bar{\lambda}_N^i(n)(n-1)$, 则此特征向量上很可能含有语音信息, $\alpha(i, n) \approx 1$, 因此噪声几乎不更新, 前一帧的估计值继续用作当前的估计; 如果 $\lambda_Y^i(n)$ 的值接近于 $\bar{\lambda}_N^i(n)(n-1)$, 则很可能不含语音, $\alpha(i, n) \approx 1$ 因此噪声的估计值用当前帧来更新。

②最小值跟踪与更新

当得到平滑的特征值后 $\lambda(i, n)$, 在连续 L 帧的时间窗内, 沿着每个特征向量找出 $\lambda(i, n)$ 的最小值 $\lambda_{\min}(i, n)$, 每输入一帧语音, 都更新此最小值。由式(12)可得

$$\lambda_{\min}(i, n) = \min\{\lambda(i, n-L-1), \cdots, \lambda(i, n-1), \lambda(i, n)\} \quad (13)$$

窗长 L 决定搜索精度。 L 的选取要足够长, 以保证跟踪的最小值与噪声相关, 而 L 的上限应考虑避免噪声的估计更新过慢, 实验表明, 0.2~1.0 s 的窗长满足上述条件。

③偏差补偿

考虑到带噪语音特征值的最小值 $\lambda_{\min}(i, n)$ 与真实噪声之间的偏差, 采用文献[7]的方法, 对 $\lambda_{\min}(i, n)$ 进行补偿, 这可以通过乘以一个补偿因子 B_{\min} 来实现。

$$\lambda_N^i(n) = B_{\min} \lambda_{\min}(i, n) \quad (14)$$

B_{\min} 的取值由实验确定, 其范围为 1.1~1.4。 $\lambda_N^i(n)$ 即为当前帧第 i 个特征向量上的噪声的特征值估计。

3) 分帧, 对每一帧带噪语音构造协方差矩阵。

4) 对相减得到带噪语音的协方差矩阵求奇异值分解。

5) 根据奇异值贡献率修改奇异值谱。忽略小奇异值以及其所对应的特征向量, 降维恢复后的信号就是估计的语音信号。

6) 重构当前帧的协方差矩阵, 得到增强后的语音。

7) 将各帧语音连接得到新的增强后的语音。

实验结果: 对一段语音加入粉红噪声, 使信噪比为 0 dB, 分别用传统的奇异值方法及改进的奇异值分解法对这段语音进行去噪。图 1~图 5 分别为使用各种方法去噪后语音信号的波形图。

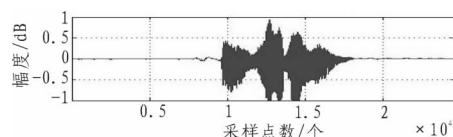


图1 原始语音信号

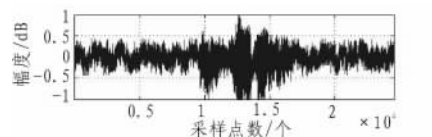


图2 加噪后的信号

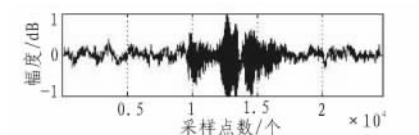


图3 传统奇异值方法

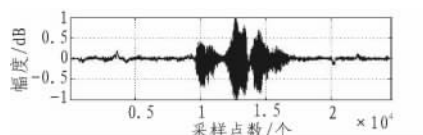


图4 奇异值时域约束估计法

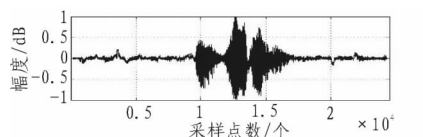


图5 奇异值频域约束估计法

由实验波形图可以看到,通过最小值统计噪声估计法和时域约束、频域约束改进过的奇异值方法比传统的奇异值方法在去噪效果上有明显提高。

4 结 论

通过实验表明:相比其他传统的子空间分解方法,改进的奇异值分解法,在有色噪声情况下,通过最小值估计法对噪声估计的改进和在分解过程中时域约束和频域约束对奇异值法的改进实现了很好的语音增强效果。

参考文献:

[1] Ephraim Y, Van Trees H L. A signal subspace approach for

speech enhancement [J]. IEEE Trans Speech and Audio Processing, 1995, 3(4): 251-266.

[2] Gazor S, Rezayee A. An adaptive KLT approach for speech enhancement [J]. IEEE Trans on Speech and Audio Processing, 2001, 9(2): 95-97.

[3] Lev-Ari H, Ephraim Y. Extension of the signal subspace speech enhancement approach to colored noise [J]. IEEE Signal Processing Lett, 2003, 10(4): 104-106.

[4] Jabloun F, Champagne B. Incorporating the human hearing properties in the signal subspace approach for speech enhancement [J]. IEEE Transactions on Speech and Audio Processing, 2003, 11(6): 700-708.

[5] Gazor S, Zhang W. Speech enhancement employing Laplacian-gaussian mixture [J]. IEEE Transactions on Speech and Audio Processing, 2005, 13(5): 896-904.

[6] Doclo S, Dologlou I, M Moonen. A novel iterative signal enhancement algorithm for noise reduction in speech [C]. Pros. of the 5th International Conference on Spoken Language Processing (ICSLP98), Sydney, Australia, December, 1998: 1435-1438.

[7] Martin R. Noise power spectral density estimation based on optimal smoothing and minimum statistic [J]. IEEE Trans on Speech and Audio Processing, 2001, 9(5): 504-512.

[8] 赵胜跃, 戴蓓蓓. 基于最小统计噪声估计的信号子空间语音增强 [J]. 数据采集与处理, 2007(12): 453-457.

适合高绝缘电压应用 10MBd 数字光电耦合器

Avago Technologies(安华高科技)推出具备高绝缘电压规格的 10MBd 数字光电耦合器, 适合电源和高功率电机控制应用。Avago 是为通信、工业和消费类等应用领域提供模拟接口零组件的领导厂商。

ACNV2601 单通道 10MBd 数字光电耦合器可以满足需要高数据率和高绝缘电压性能的新市场需求, Avago 的 ACNV2601 是业内第一款结合 10MBd 数据率并维持甚至提高绝缘电压规格的光电耦合器产品。

ACNV2601 在设计上可以承受 7 500 Vrms/一分钟隔离电压、1 768 Vpeak 连续工作绝缘电压和 12 000 V 的瞬变电压, ACNV2601 目前正在申请可以承受电气冲击应用的强化绝缘 IEC/UL/CSA 规范认证。

Avago 特有的 AlGaAs 发光二极管和光检测器间内部屏蔽可以保证共模电压 1,500 V 时 20,000V/μs 的共模瞬变抑制能力。

Avago 的 ACNV2601 拥有的高绝缘电压使它非常适合 690 V 电源电压或使用高电压 IGBT 的高速功率开关设备应用, 例如风力涡轮、智能功率模块以及如 RS-232/RS-485 和 CANBus 等通信接口都可以由 ACNV2601 的高速和高绝缘电压特性获益。

在传统线路接收器无法运作的场合, ACNV2601 可以在超高接地或感应噪声下提供可靠的逻辑接口和输入输出缓冲。

ACNV2601 功能特点如下: 10MBd 典型数据率和高绝缘电压; ACNV2601 全球安全规范认证申请中: IEC/EN/DIN 60747-5-5 1 768 V 工作绝缘电压和 12 000 Vpeak 瞬变电压, UL1577 7 500 Vrms 一分钟输入输出绝缘电压, CSA: 20 kV/μs 共模瞬变抑制能力, Vcm = 1 500 V 时; 2 mm 内部封装绝缘穿透距离 (DTI); 13 mm 最低封装爬电距离和电气间隙; 5 V 工作电源; 10-pin 宽体 DIP 封装, 并提供鸥翼式 SMD 引脚选择; -40 °C 到 +105 °C 宽广工作温度范围。

ACNV2601 提供有 10-pin 500mil 宽体 DIP 封装和适合表面贴装程序鸥翼式引脚选择, ACNV2601 10MBd 光电耦合器可以在 -40 °C 到 +105 °C 的宽广温度范围下工作, 并保证交直流电气性能。

除了高绝缘能力外, ACNV2601 的封装也提供大于 13 mm 的爬电距离和电气间隙以及 2 mm 的绝缘穿透距离 (DTI)。

咨询编号: 2010061013

-129-