

文章编号:1001-9081(2006)11-2694-02

盲源分离在单通道语音增强算法中的应用

马建芬,李鸿燕,张雪英,王华奎

(太原理工大学 计算机与软件学院,山西 太原 030024)

(majianfentyut@126.com)

摘要:提出一种单通道语音增强算法。首先由接收到的单声道语音信号的含噪部分构造一个假想噪声源,将这一噪声源和含噪的信号作为多通道自适应去相关(MAD)盲分离算法的输入,得到增强的语音信号。进一步将这一增强的语音作为输入,利用 Daubechies 小波对其进行分解,在小波域中选取合适的阈值函数进行滤波,然后合成时域语音信号。根据以上步骤得到的增强语音有较高的信噪比及可懂度。

关键词:语音增强;盲源分离;小波变换

中图分类号: TN912 **文献标识码:** A

Usage of blind source separation in single channel speech enhancement algorithm

MA Jian-fen, LI Hong-yan, ZHANG Xue-ying, WANG Hua-kui

(College of Computer and Software, Taiyuan University of Technology, Taiyuan Shanxi 030024, China)

Abstract: A new single channel speech enhancement algorithm was put forward. A pseudo noise source was used as the input of the multiple adaptive decorrelation (MAD) algorithm to get an improved speech signal. Moreover, Daubechies wavelet was adapted to transform the signal into wavelet domain. Proper threshold function was used to select wavelet coefficients, and then the signal into time domain was synthesized. Following the scheme provided above, we can get the speech signal with higher SNR and intelligibility.

Key words: speech enhancement; blind source separation; wavelet transform

1 盲源分离^[1]

盲源分离是近年来流行的一种多通道信源分离算法,由于该算法不需要知道被测信号的先验知识,近年来被广泛用于多声道语音增强等应用中。

盲信号分离问题就是求分离矩阵 P ,使得通过它可以仅从观测信号 $x(t)$ 来恢复出源信号 $s(t)$,设 $y(t)$ 为源信号的估计矢量,则分离系统输出可通过如下数学模型表示:

$$y(t) = Ps(t) = WAs(t) \quad (1)$$

文献[1]中提出一种多通道自适应去相关算法(Multiple Adaptive Decorrelation, MAD)用于非平稳信号的分离。该算法是在频域中的去相关算法。

$$X(\omega, t) \approx A(\omega)S(\omega, t) \quad (2)$$

式中 $X(\omega)$ 是 $x(t)$ 以 t 为中心相叠计算的短时傅里叶变换。被测信号的互相关为:

$$\hat{R}_x(\omega, t) = E[X(\omega, t)X^H(\omega, t)] \quad (3)$$

源的互相关为:

$$\hat{\Lambda}_s(\omega, t) = E[S(\omega, t)S^H(\omega, t)] \quad (4)$$

加权矩阵 $W(\omega)$ 可通过下式求得:

$$\hat{W}, \hat{\Lambda}_s = \arg \min_{\hat{W}, \hat{\Lambda}_s} \sum_{\omega=1}^T \|W\hat{R}_x(\omega, t)W^H, \Lambda_s(\omega, t)\|^2 \quad (5)$$

$$W(\tau) = 0, \forall \tau > Q, Q \ll T$$

$$W_{ii}(\omega) = 1$$

由于源的相关为:

$$\hat{\Lambda}_s(\omega, t) = \text{diag} [W(\omega)\hat{R}_x(\omega, t)W^H] \quad (6)$$

为使源的相关的估计逼近式(6),代价函数的选取应使得互相关矩阵 $\hat{R}(\omega, t)$ 的非对角线元素最小化,可以得到以下的学习规则:

$$\Delta W^* (\omega) = 2\mu E(\omega, t) W(\omega) \hat{R}_x(\omega, t) \quad (7)$$

其中 μ 为学习的速率, $E(\omega, t) = W\hat{R}_x(\omega, t)W^H - \hat{\Lambda}_s(\omega, t)$ 。滤波矩阵 $W(\omega)$ 通过解方程(5)得到。在每一个局部最优的计算过程中,混合矩阵的互相关由下式更新:

$$\hat{R}_x(\omega, t) = (1 - \gamma)\hat{R}_x(\omega, t) + X(\omega, t)X^H(\omega, t) \quad (8)$$

其中 γ 是遗忘因子,该算法不需要任何的先验知识,因此称为盲分离算法。

2 基于盲源分离的语音增强算法

在单声道语音增强应用中,为了有效地利用盲源分离算法,基本思想是将包含在单通道中的背景噪声转化为伪源干优信号,以满足盲源分离算法中两路输入的条件。

假设被分离的语音信号为 $\hat{s}_1(t)$, $t \in [t_0, t_f]$, 其中的含噪部分为 $s_n(t) = \hat{s}_1(t)$, $t \in [t_0, t_1]$, $T_r = t_1 - t_0$, $s_n(t)$ 的傅里叶变换由下式构成:

$$S_n(\omega) = |S_n(\omega)| e^{i\Phi_{S_n}(\omega)} \quad (9)$$

收稿日期:2006-05-16;修订日期:2006-07-22

基金项目:国家自然科学基金资助项目(60472094);山西省高校科技开发项目(197-12909083)

作者简介:马建芬(1967-),女,河北容城人,副教授,硕士,主要研究方向:语音信号处理、神经网络、小波理论;李鸿燕(1973-),女,山西运城人,讲师,硕士,主要研究方向:信号处理、神经网络。

计算长度为 T_r 的信号 $\hat{s}_1(t)$ 的傅里叶变换:

$$S_{T_r}(\omega) = |S_{T_r}(\omega)| e^{i\phi_{S_{T_r}}(\omega)} \quad (10)$$

将(10)式中的幅度谱用噪音的幅度谱进行替换,得到:

$$S_{T_r}^a(\omega) = |S_n(\omega)| e^{i\phi_{S_{T_r}}(\omega)} \quad (11)$$

噪音信号由在 T_r 时间内的傅里叶反变换构成,取窗长为 $T_r/2$,可以得到整个噪音的时域信号 $s^a(t)$,上述噪音信号由噪音的谱特性及讲话者的相位特性决定。将 $s^a(t)$ 及 $\hat{s}_1(t)$ 作为 MAD 算法的输入可以进一步去除噪声。

和传统的谱提取算法相比,由于利用人工产生的噪音源进行去噪会增加信号的相关性,此算法中引起的音乐噪声最小。

3 小波滤波用于背景噪声的滤除

经上述算法增强的语音信号有较好的可懂度,为了进一步增强语音,对上述增强语音利用小波变换的方法进一步进行处理。

背景噪声的统计分布具有近似高斯分布的特性,而语音信号具有相对稀疏的分布特性。通过将信号进行空间变换可突出其超高斯性或稀疏性,使语音信号具有较大的值而噪声具有较小的值。从统计的观点来看,经此变换后噪声的值近似为零,可以通过设置合适的阈值滤除。

连续小波变换可以看作通过一组线性滤波将信号映射到一个稀疏空间中。研究表明小波系数具有稀疏特性,该变换

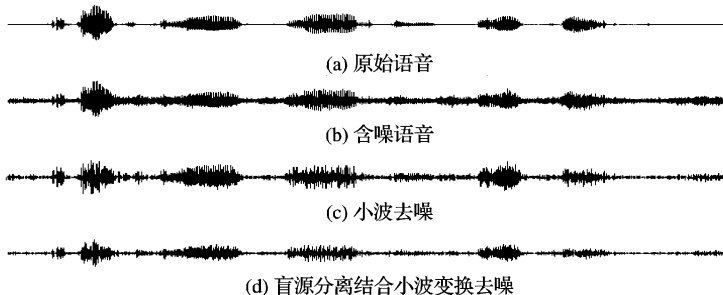


图1 语音时域波形图

5 结语

本文利用小波变换对传统的基于 BSS 的语音增强算法进行了改进,可以看出改进的算法能有效地提高信号的信噪比,可以在不同的背景噪声下进一步测试该算法的性能,对增强的语音用维纳滤波的方法进行增强。

参考文献:

- [1] VISSER E, OTSUKA M, LEE T-W. A Spatio-temporal speech enhancement scheme for robust speech recognition in noisy environments[J]. *Speech Communications*, 2003, 41(2): 393-407.
- [6] 张元亮, 郑南宁, 贾天旭. 自适应聚类滤波[J]. *西安交通大学学报*, 1998, 32(1): 17-20.
- [7] 甘祥超, 武栓虎, 谭正. 基于小波四叉相树的图像去噪方法[J]. *计算机研究与发展*, 2001, 38(3): 341-343.
- [8] 赵红蕊, 阎广建, 邓小炼, 等. 一种简单加入空间关系的实用图像分类方法[J]. *遥感学报*, 2003, 7(5): 358-363.
- [9] ZHAO HR, SHI H, YANG H, *et al*. Separating the radiance contribution of land surface and atmosphere[A]. *IGARSS'03[C]*, 2003.
- [10] SEEMANN T, TISCHER P. Structure preserving noise filtering

由一个过采样移不变的多分辨滤波带构成。该滤波带将信号映射到一个正交的子空间中,该子空间具有固定的带宽与中心频率之比,与语音识别中的梅尔尺度是一致的,这样通过小波变换可同时得到信号的稀疏表示与生理感知。

首先计算信号 $y(t)$ 在每个子带内的含噪小波系数 y_i ,任务是计算每个子带内的去噪小波系数 y_i^s 以使噪音信号和语音信号的联合概率分布 $P(y_i - y_i^s | \sigma_i)P(y_i^s)$ 最大,其中 σ_i 为噪音的方差。已知噪音和语音信号的分布分别为高斯分布及拉普拉斯分布:

$$P(y_i - y_i^s | \sigma_i) \approx e^{-((y_i - y_i^s)^2 / 2\sigma_i^2)}, P(y_i^s) \approx e^{-|y_i^s|} \quad (12)$$

考虑到对数似然度, y_i^s 可由以下公式近似计算:

$$y_i^s = k(y_i) = \text{sign}(y_i) \max(|y_i| - \sqrt{2}\sigma_i, 0) \quad (13)$$

式中 σ_i 必须用传统的方法进行精确的估计。

4 实验结果

为了评价改进算法的性能,我们对 8kHz 抽样的语音数据分别按信噪比为 5dB, 10dB, 15dB 的要求加入汽车背景噪声,实验中我们分别对信号按上述方法进行语音增强。定义增强语音的信噪比为: $SNR = 10 \log (\|X\|^2 / \|Y - X\|^2)$, 其中 X 是原始信号, Y 是含噪信号。

表 1 为一男声语音和一女声语音在不同信噪比的情况下用上述算法进行增强及直接按小波增强的算法所得到的信噪比。

表1 直接利用小波与利用盲源分离进行预处理

测试语音	信噪比(单位: dB)	
	直接利用小波的方法	盲源分离结合小波变换
5	10.1	12.1
10	12.3	14.5
15	16.5	17.7
5	10.2	11.6
10	12.4	14.0
15	16.2	17.1

- [2] PARRA L, SPENCE C. Convolutional blind separation of non-stationary sources[J]. *IEEE Transactions on Speech Audio Processing*, 2000, 8(3): 320-327.
- [3] MISEP Linear and Nonlinear ICA Toolbox[EB/OL]. <http://neural.inesc.pt/~lba/>, 2006.
- [4] 863 语料库试用数据[EB/OL]. <http://www.chineseldc.org/resource.asp>, 2006.
- [5] JUTTEN C, HERAULT J. Blind separation of sources, part 1: an adaptive algorithm based on neuromimetic architecture[J]. *Signal Processing*, 1991, 24(1): 1-10.

(上接第 2693 页)

- [6] 张元亮, 郑南宁, 贾天旭. 自适应聚类滤波[J]. *西安交通大学学报*, 1998, 32(1): 17-20.
- [7] 甘祥超, 武栓虎, 谭正. 基于小波四叉相树的图像去噪方法[J]. *计算机研究与发展*, 2001, 38(3): 341-343.
- [8] 赵红蕊, 阎广建, 邓小炼, 等. 一种简单加入空间关系的实用图像分类方法[J]. *遥感学报*, 2003, 7(5): 358-363.
- [9] ZHAO HR, SHI H, YANG H, *et al*. Separating the radiance contribution of land surface and atmosphere[A]. *IGARSS'03[C]*, 2003.
- [10] SEEMANN T, TISCHER P. Structure preserving noise filtering

using explicit local segmentation [A]. *Proceedings IAPR Conference on Pattern Recognition[C]*, 1998.

- [11] 郭桂蓉. 模糊模式识别[M]. 长沙: 国防科技大学出版社, 1992. 178-187.
- [12] ARCE GR, FOSTER RE. Detail-preserving ranked order based filters for image processing[J]. *IEEE Transactions on ASSP*, 1987, 37(1): 83-98.
- [13] THONG T. Digital image processing test patterns [J]. *IEEE Transactions on ASSP*, 1983, 31(3): 675-690.