

背景噪声下的语音信号分离

云晓花, 景新幸

(桂林电子科技大学信息与通信学院, 广西 桂林 541004)

摘 要: 独立分量分析法在分离含有背景噪声的混合语音时效果不理想。为此, 将独立分量分析算法与卡尔曼滤波相结合, 对语音进行降噪处理, 采用 FastICA 算法对含噪语音进行分离, 分离速率高于 Informax 算法, 能够获得较清晰的语音文件。通过仿真实验验证了该方法的可行性和有效性。

关键词: 独立分量分析; 背景噪声; 卡尔曼滤波; 语音分离; FastICA 算法

Speech Signal Separation Under Background Noise

YUN Xiao-hua, JING Xin-xing

(College of Information & Communication, Guilin University of Electronic Technology, Guilin 541004, China)

【Abstract】 Because of background noisy, there is a problem of directly separating the mixed-signal observed usually can not achieve good results. The basic principle and algorithms of independent component analysis is illustrated. Considering characteristic of Kalman filtering, the speeches under background noisy are denoised; and speeches are separated by the FastICA algorithms. Simulation results indicate the availability and efficiency of this method.

【Key words】 Independent Component Analysis(ICA); background noise; Kalman filtering; speech separation; FastICA algorithm

DOI: 10.3969/j.issn.1000-3428.2011.23.061

1 概述

自动语音识别系统广泛应用于人们的日常生活, 但是往往受到噪声环境的影响, 例如其他说话人的语音、机械噪声, 这些噪声有时很强, 使语音识别系统的性能大幅降低。如果能实现噪声和语音的自动分离, 即在识别前获得较纯净的语音, 就可以彻底解决噪声环境下的识别问题。而盲源分离技术可以将各种本来混合且相对独立的信号分离, 卡尔曼滤波器因其自身高度的灵活性和合理的计算复杂度, 被广泛应用于语音信号的噪声抑制方面。由于背景噪声的存在会导致盲源分离(Blind Source Separation, BSS)分析更复杂, 因此人们通常都在无任何其他背景噪声下对语音信号进行分离, 但在实际中背景噪声的存在是无法避免的。

本文采用一种盲源分离技术即独立分量分析法(Independent Component Analysis, ICA)和卡尔曼滤波相结合的方法, 分离含有背景噪声的混合语音, 用于自动语音识别系统的前端。

2 独立分量分析法

盲源分离的实质是从未知信号的线性混合观测信号中重构出原始的各个源信号。其中, 独立分量分析是一种盲源信号分离方法, 其目的是在统计独立性的假设下, 将观测到的随机矢量分离成各独立分量^[1-3]。图 1 为独立分量分析的原理示意图。

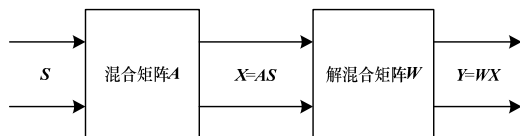


图 1 ICA 原理示意图

假设有 N 个未知的源信号 $s_i(t)$, $i=1 \sim N$, 构成一个向量列 $S(t)=[s_1(t), s_2(t), \dots, s_N(t)]^T$, $X(t)=[x_1(t), x_2(t), \dots, x_M(t)]^T$ 是由

M 个可观测信号 $x_i(t)$, $i=1, 2, \dots, M$ 构成的向量, 且满足下列方程:

$$X(t) = AS(t) \quad (1)$$

其中, A 是一个 $M \times N$ 维矩阵, 一般称为混合矩阵。

设一个 $N \times N$ 的反混合阵 $W=(w_{ij})$, $X(t)$ 经过 W 变换后得到 N 维输出向量 $Y(t)=[y_1(t), y_2(t), \dots, y_N(t)]^T$, 即有:

$$Y(t) = WX(t) = WAS(t) \quad (2)$$

如果实现 $WA=I$ (I 是 $N \times N$ 单位阵), 则 $Y(t)=S(t)$, 从而实现了源信号分离的目的。

目前, ICA 算法基本分为两大类: (1)基于批运算某种相关的判据函数, 这些算法通常需要进行复杂的矩阵或向量运算; (2)基于随机梯度方法的自适应算法, 这些算法的收敛速度慢。在诸多 ICA 算法中, 固定点算法(FastICA)以其收敛速度快、分离效果好被广泛应用于信号处理领域。该算法是一种快速寻优的迭代算法, 它有基于峰值、基于最大似然和负熵最大化等多种形式。FastICA 算法的基本步骤为:

- (1)对观测数据进行适当的预处理, 包括去均值和白化。
- (2)独立分量提取。

3 卡尔曼滤波

卡尔曼滤波^[4]是在均方误差最小准则下的估计, 其不仅适用于平稳系统, 还适用于非平稳信号、时变系统。实际语音信号一般都受到加性噪声干扰, 可以表示为:

$$y(n) = x(n) + v(n) \quad (3)$$

其中, $x(n)$ 为语音信号; $v(n)$ 为加性噪声。实际应用中噪声和信号之间通常是相互独立的。可将叠加噪声近似作为零均

基金项目: 国家自然科学基金资助项目(609661002)

作者简介: 云晓花(1984—), 女, 硕士研究生, 主研方向: 语音信号处理; 景新幸, 教授

收稿日期: 2011-05-11

E-mail: yundaisy@163.com

值、方差为 σ^2 的高斯白噪声。通常用 P 阶的自回归(Auto Regressive, AR)模型描述纯净语音:

$$x(n) = \sum_{i=1}^P a_i x(n-i) + w(n) \quad (4)$$

其中, $w(n)$ 为语音的激励信号。由卡尔曼滤波理论可得系统的状态估计:

$$\hat{x}(n) = A(n-1)\hat{x}(n-1) + K(n)[x(n) - H^T(n)A(n-1)\hat{x}(n-1)] \quad (5)$$

$$K(n) = P(n|n-1)H^T(n)[H(n)P(n|n-1)H^T(n+1)]^{-1} \quad (6)$$

$$P(n|n-1) = A(n-1)P(n-1)A^T(n-1) + \Gamma^T(n-1)N(n-1)\Gamma^T(n-1) \quad (7)$$

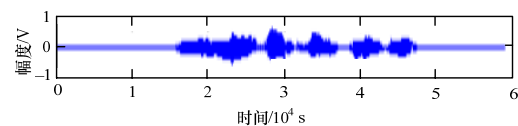
$$P(n) = [1 - K(n)H^T(n)]P^T(n|n-1) \quad (8)$$

其中, $K(n)$ 为卡尔曼增益矩阵; $P(n|n-1)$ 为单步预测误差协方差矩阵; $P(n)$ 为滤波误差协方差矩阵, 则经卡尔曼滤波后的语音信号为 $\hat{s}(n) = [H^T(n), 0]\hat{x}(n)$ 。

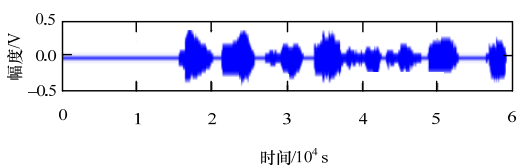
4 实现步骤

本文将 FastICA 算法与卡尔曼滤波法相结合^[2,5], 先利用卡尔曼滤波对分离后的信号进行去噪, 再利用 FastICA 算法对混入的干扰语音进行分离。FastICA 采用基于峰度法推广的通用目标函数 FastICA 改进算法, 语音信号分离方法实现步骤如下:

(1) 读入源语音信号 S , 波形文件如图 2 所示。



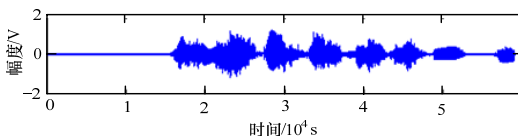
(a) 第 1 个原始语音信号



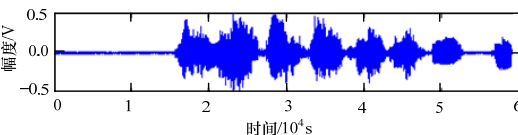
(b) 第 2 个原始语音信号

图 2 原始语音信号

(2) 产生混合语音文件, 并进行加噪处理, 波形文件如图 3 所示。



(a) 混合加噪语音之一



(b) 混合加噪语音之二

图 3 混合加噪语音

(3) 进行预处理(去均值和白化)。

(4) 在进行独立分量提取算法之前先采用卡尔曼滤波进行去噪:

1) 初始化, 置循环变量 $k=1$, 随机初始化权矢量 $w(k)$;

2) $w(k)$ 标准化, $w = w/\|w\|$;

3) 迭代运算: $w(k+1) = E\{v_k(w(k)^T v_k)^3\} - 3w(k)$, 如果 $|w(k+1)^T w(k)|$ 不是足够接近 1, 则继续返回 2) 进行迭代运算,

否则, 循环结束。

为了保证每次分离的是不同的独立成分, 需要在上述循环迭代的基础上再增加一个简单的正交投影操作。因此, 将 $w = w - BB^T w$ 添加到 2) 中。

5 实验结果及分析

本文实验读入 2 个语音“wav”文件, 噪声信号为加性高斯白噪声, 噪声信号与语音信号的采样率为 16 kHz, 量化精度为 16 bit; 将它们线性混合成含噪语音信号。用 FastICA 算法进行语音分离, 同时采用卡尔曼滤波进行消噪处理, 最后可以通过“sound”语句试听分离后的语音。结果为先去噪再分离还原的语音比较清晰, 而直接分离还原的语音几乎不能听清。图 4、图 5 分别为利用本文方法分离得到的语音波形图和语谱图, 图 6、图 7 分别为含噪语音直接进行 FastICA 算法分离得到的语音波形图和语谱图。对比图 4、图 6 可见, 先去噪后分离的语音与原输入语音的波形基本一致, 只是在幅值大小上有一定差异; 而直接分离的语音在分离后与原输入语音波形不一致。通过图 5、图 7 的语谱图对比, 显然本文方法比直接进行分离算法的效果好。

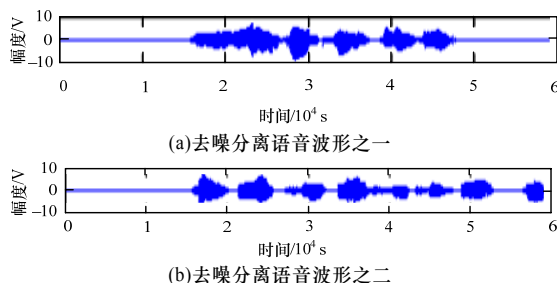


图 4 去噪分离语音波形

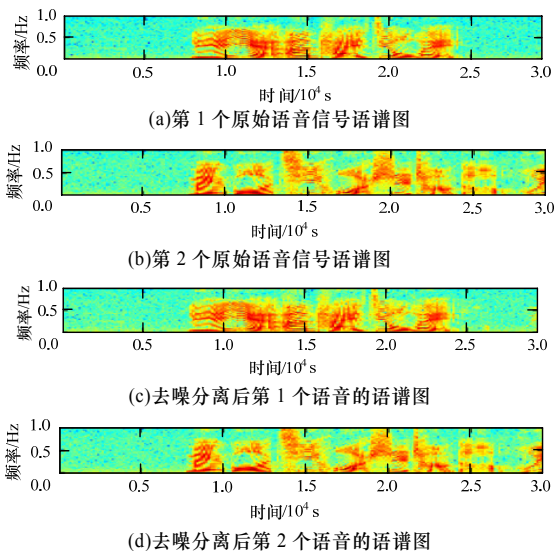


图 5 去噪分离语谱图

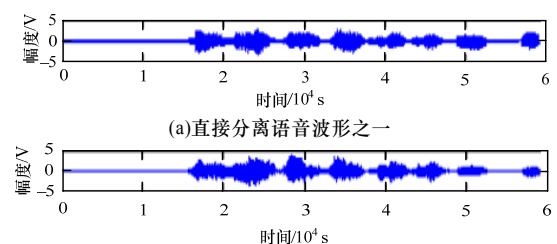


图 6 直接分离语音波形

(下转第 185 页)

大迭代次数设为 500 时的计算机仿真结果, 并与迭代次数为 1 000 的结果进行了比较, 表中斜线后面的数字表示迭代 1 000 次和 500 次的结果之比。

表 1 迭代次数为 1 000 时 5 种算法的仿真结果

算法	适应度(均方误差)			
	最大值	最小值	均值	标准差
QPSO	0.035 72	0.003 13	0.014 54	0.007 89
CPSO	0.013 30	0.002 75	0.005 73	0.002 21
PSO-CA	0.031 38	0.002 29	0.013 91	0.007 05
DE-CA	0.025 09	0.002 81	0.011 23	0.005 59
PSO-DE-CA	0.005 09	0.002 20	0.002 86	0.000 63

表 2 迭代次数为 500 时 5 种算法的仿真结果

算法	适应度(均方误差)			
	最大值	最小值	均值	标准差
QPSO	0.067 98/0.53	0.004 45/0.70	0.028 06/0.52	0.016 29/0.48
CPSO	0.022 84/0.58	0.003 65/0.75	0.008 03/0.71	0.003 72/0.59
PSO-CA	0.086 99/0.36	0.004 95/0.46	0.023 23/0.60	0.017 28/0.41
DE-CA	0.063 28/0.40	0.003 12/0.90	0.024 58/0.46	0.015 65/0.36
PSO-DE-CA	0.005 21/0.98	0.002 22/0.99	0.003 49/0.82	0.000 74/0.85

对比表 1 与表 2 可见, PSO-DE-CA 迭代 500 次的仿真结果接近其迭代 1 000 次的结果, 同时优于其他 4 个算法迭代 1 000 次的结果。

5 结束语

本文提出双种群文化算法。该算法的双种群空间分别按照 PSO 和 DE 算法独立进化。信仰空间提炼并保存来自 PSO

和 DE 种群空间的文化知识, 反过来指导 PSO 和 DE 种群空间的进化, 使 PSO 和 DE 种群空间可以相互协作、共同进化。因此, 该算法能充分发挥 PSO 和 DE 算法的优势, 有效克服 2 个算法独立运行的缺陷。本文将该算法应用于 FIR 滤波器的优化设计并进行了仿真, 结果表明其优于其他算法。

参考文献

[1] 周国昌, 沈绪榜. 基于 LS CSIMD 的可变阶 FIR 并行算法研究[J]. 计算机工程, 2007, 33(1): 7-9.

[2] 方伟, 孙俊, 须文波. 基于自适应量子粒子群算法的 FIR 滤波器设计[J]. 系统工程与电子技术, 2008, 30(7): 1378-1381.

[3] Zhang Jidong, Jia Dongli, Li Kui. FIR Digital Filters Design Based on Chaotic Mutation Particle Swarm Optimization[C]//Proc. of International Conference on Audio, Language and Image Processing. Shanghai, China: [s. n.], 2008: 418-422.

[4] Zhao Zhongkai, Gao Hongyuan. FIR Digital Filters Based on Cultural Particle Swarm Optimization[C]//Proc. of International Workshop on Information Security and Application. Qingdao, China: [s. n.], 2009: 252-255.

[5] Gao Hongyuan, Diao Ming. Differential Cultural Algorithm for Digital Filters Design[C]//Proc. of the 2nd International Conference on Computer Modeling and Simulation. Sanya, China: [s. n.], 2010: 459-462.

编辑 张正兴

(上接第 182 页)

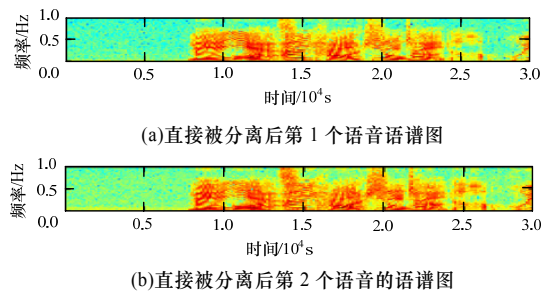


图 7 直接分离语谱图

同时为了说明该算法的有效性, 进行了定性分析。本文通过信噪比 (SIR) 性能指标来进行度量, 即 $SIR(y_i) = 10 \lg(\sum_{i=1}^n s_i^2 / \sum_{i=1}^n (y_i - s_i)^2)$, 其中, s_i 表示的是源信号; y_i 是分离后的信号, SIR 越大, 分离性能越好。3 种算法的实验结果见表 1。

表 1 算法性能比较

方法	SIR/dB	时间/s
直接分离混合信号	4.192 3	0.417 757
去噪再采用 FastICA 算法分离	16.099 7	2.613 064
去噪再采用 Informax 算法分离	15.940 8	2.915 247

由表 1 结果可以得出, 先去噪再进行分离的效果比直接对混合信号分离的方法效果好。同时, FastICA 算法的分离速率比 Informax 算法高。

6 结束语

独立分量分析作为盲信号分离技术的一种新技术, 在语音识别方面的应用越来越广泛。本文采用卡尔曼滤波与独立分量分析结合的方法处理噪声背景下的语音信号。大量实验结果显示, 无论在信号波形还是声音的效果上, 本文方法对背景噪声下的语音信号都有明显的改善效果, 验证了其在噪声环境下语音识别系统前端应用的有效性。

参考文献

[1] 张启发. 盲信号处理及应用[M]. 西安: 西安电子科技大学出版社, 2006.

[2] Hyvarinen A. Fast and Robust Fixed-point Algorithms for Independent Component Analysis[J]. IEEE Transactions on Neural Networks, 1999, 10(3): 626-634.

[3] Kokkinakis K, Loizou P C. Using Blind Source Separation Techniques to Improve Speech Recognition in Bilateral Cochlear Implant Patients[J]. Journal of Acoustical Society of America, 2008, 123(4): 2379-2390.

[4] Ludeman L C. 随机过程——滤波、估计与检测[M]. 邱天爽, 译. 北京: 电子工业出版社, 2005.

[5] 李鸿燕, 郝润芳, 马建芬, 等. 基于维纳滤波和快速独立分量分析的有噪混合图像盲分离[J]. 计算机应用研究, 2007, 24(10): 161-162.

编辑 张正兴