

基于皮尔逊系统的语音信号盲源分离

季 策, 刘梦蝶, 陶奕名

(东北大学 信息科学与工程学院, 辽宁 沈阳 110819)

摘 要: 当利用传统自然梯度算法对所有语音信号都使用同一个激活函数进行分离时, 对语音信号的盲源分离效果都不尽理想. 针对这一问题, 采用基于皮尔逊系统的分段激活函数对传统自然梯度算法进行改进. 通过引入皮尔逊系统, 将皮尔逊函数与传统激活函数相结合, 再利用信号的矩估计方法, 分段选择合适的激活函数代入分离矩阵, 有效克服了传统语音分离算法的缺点和不足. 仿真结果表明, 在对实际的语音信号进行分离时, 改进算法的性能明显优于传统自然梯度算法, 并且在保持了良好收敛速度的同时大大减少了均方误差.

关 键 词: 盲源分离; 自然梯度; 语音信号; 皮尔逊系统; 激活函数

中图分类号: TN 911.6

文献标志码: A

文章编号: 1005-3026(2015)01-0006-04

The Blind Source Separation for Speech Signal Based on Pearson System

Ji Ce, Liu Meng-die, Tao Yi-ming

(School of Information Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: Ji Ce, associate professor, E-mail: jice@ise.neu.edu.cn)

Abstract: The same activation function is always used to separate all speech signals by means of the conventional natural gradient algorithm. Although blind source separation can be achieved, the separation effect is not ideal. To solve this problem, sub-band activation function was used based on Pearson system to improve natural gradient algorithm. By introducing Pearson system, the Pearson function with conventional activation function was combined. The appropriate activation function was selected according to the method of moments-estimating in each part. Then the selected activation function was brought into the separation matrix. This algorithm effectively overcomes the shortcomings and deficiencies of the conventional separation algorithm for speech signal. The simulation results showed that the performance of the improved algorithm is superior to that of the conventional natural gradient algorithm in the actual speech signal separation. In addition, the mean square error is reduced greatly and good convergence rate is maintained at the same time.

Key words: blind source separation; natural gradient; speech signal; Pearson system; activation function

盲信号分离(BSS)是在不知道传输信道特征和真实源信号的情况下,从传感器阵列中分离源信号的过程.盲分离技术已在许多信号处理领域中得到了广泛应用,如语音信号^[1]、图像信号^[2]、生物医学信号处理^[3]、地震信号等.

盲信源分离具有许多有效的算法,其中以自然梯度算法最为著名.该算法以互信息作为代价函数,通过自然梯度最小化该代价函数,从而迭代

得到分离矩阵实现盲分离.影响该算法性能的因素主要有两个:学习步长和激活函数的估计.在基于传统自然梯度算法的语音信号盲分离中,总是对所有语音信号的分离使用同一个激活函数.实际上不同语音信号的分布是不一样的,因此对所有语音信号使用同一个激活函数的分离效果不尽理想.

为了解决不同语音信号的盲分离问题,提高语音信号的分离性能,本文给出了一种新型的分

段激活函数,即将皮尔逊系统作为激活函数引入自然梯度算法,与其他传统的激活函数相结合,构成了一种新的激活函数.以矩估计的方法作为语音信号的判断标准,根据不同的语音信号选择不同的激活函数来实现信号的盲源分离.语音信号的仿真实验证明:本文给出的分段激活函数,在保证收敛速度不变的前提下,大大减少了信号的均方误差,提高了信号的分离性能.

1 盲源分离问题

1.1 语音信号盲源分离的基本模型

设 $S(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$ 为 N 个声音源信号, $X(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T$ 是由 N 个传感器接收到的观察信号,忽略传输延迟和噪声信号, $X(t)$ 与 $S(t)$ 满足如下线性关系:

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

式中, A 是一个 $N \times N$ 的非奇异混合矩阵.盲信号分离的目标就是如何在仅知道观测信号 $X(t)$ 的情况下,找到一个分离矩阵 W ,使输出信号 $Y(t) = WX(t)$ 为源信号 $S(t)$ 的估计.

1.2 自然梯度算法

盲信号分离的关键在于寻求最佳分离矩阵 W_{opt} .为此人们想出了不同的求解方法,如最大熵方法、最大似然估计方法、独立分量分析等.运用自然梯度的概念,将这些方法用统一的迭代公式表示如下:

$$W_{k+1} = W_k + \eta(I - \varphi(y)y^T)W_k. \quad (2)$$

式中: η 为学习步长; $\varphi(y)$ 为激活函数,定义为 $\varphi_i(y_i) = -\frac{p_i(y_i)'}{p_i(y_i)}$, $p_i(\cdot)$ 为第 i 个源信号的概率密度函数^[4].激活函数确定了算法的搜索方向,能否对其进行精确估计是决定算法是否收敛、稳定以及分离效果的重要因素.

传统的利用自然梯度算法中,总是选用双曲正切(tanh)函数作为激活函数^[5].这是因为语音信号在大部分情况下是服从超高斯分布的,所以在对分离性能要求不是很高的情况下,tanh 完全可以满足分离需要.但是,如果要得到更好的分离效果,必须考虑语音信号的各种分布情况,然后选择相对应的激活函数,代入分离矩阵进行迭代分离.

2 基于皮尔逊系统的盲源分离

2.1 皮尔逊系统

Karvanen 第一次将皮尔逊系统引入盲分离

系统,用来分离简单的超高斯和亚高斯信号^[6].皮尔逊系统基本的数学模型为

$$f'(x) = \frac{(x-a)f(x)}{b_0 + b_1x + b_2x^2}. \quad (3)$$

其中, a, b_0, b_1 和 b_2 分别为皮尔逊系统的参数.将式(3)转换成标准的激活函数的形式为

$$\varphi(x) = -\frac{f'(x)}{f(x)} = -\frac{(x-a)}{b_0 + b_1x + b_2x^2}. \quad (4)$$

其中的参数 a, b_0, b_1 和 b_2 , 可通过矩估计方法求出:

$$b_1 = a = -\mu_3(\mu_4 + 3\mu_2^2)/C, \quad (5)$$

$$b_0 = -\mu_2(4\mu_2\mu_4 - 3\mu_3^2)/C, \quad (6)$$

$$b_2 = -(2\mu_2\mu_4 - 3\mu_2^2 - 6\mu_3^2)/C. \quad (7)$$

其中, $C = 10\mu_4\mu_2 - 12\mu_3^2 - 18\mu_2^3$, μ_2, μ_3, μ_4 分别为信号的二阶矩、三阶矩和四阶矩.通过这些参数的计算,可以确定皮尔逊激活函数.

理想情况下,矩估计应由源信号来计算.但在实际情况下,源信号是未知的.因此,可以用分离出来的估计信号去代替源信号进行计算,即用样本矩来代替理论矩.具体的矩估计公式如下:

$$\left. \begin{aligned} \hat{\alpha}_1 &= \bar{y} = \sum_{i=1}^n y_i/n, \\ \hat{\alpha}_2 &= \hat{\sigma}^2 = \sum_{i=1}^n (y_i - \bar{y})^2/n, \\ \hat{\alpha}_3 &= \sum_{i=1}^n (y_i - \bar{y})^3/(n\hat{\sigma}^3), \\ \hat{\alpha}_4 &= \sum_{i=1}^n (y_i - \bar{y})^4/(n\hat{\sigma}^4). \end{aligned} \right\} \quad (8)$$

式中, $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3$ 和 $\hat{\alpha}_4$ 分别为估计出来的一阶矩、二阶矩、三阶矩和四阶矩^[7].从式(8)可以看出,一阶矩和二阶矩实际上就是均值和方差.若用 $\hat{\alpha}_2, \hat{\alpha}_3$ 和 $\hat{\alpha}_4$ 来代替 μ_2, μ_3 和 μ_4 ,就可计算出皮尔逊系统的参数.

2.2 基于皮尔逊系统的分段激活函数

皮尔逊系统可以模拟多种信号的分布,可以将皮尔逊系统作为一种激活函数引入语音信号盲分离.但是,由于皮尔逊系统分离的信号是有一定范围的^[6],超过这个范围后,用皮尔逊系统作为激活函数便不能实现盲分离或盲分离效果很差.基于这一点,本文提出将皮尔逊系统作为一种激活函数,与传统语音分离的激活函数相结合,构成一种新型的分段激活函数.然后,根据矩估计的方法,为信号选择合适的激活函数进行盲源分离.本文提出的分段激活函数如图1所示.

图1中:当 $\alpha_4 < \alpha_3^2 + 1$ 时,语音信号不能分离,而实际上,语音信号不会达到这个区域;当 $\alpha_3^2 + 1 < \alpha_4 < 2.6$ 时,语音信号服从亚高斯分布,

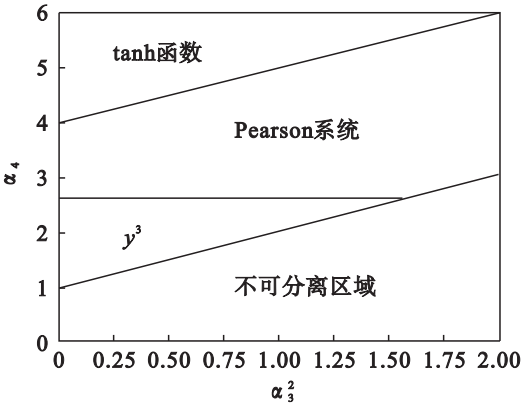


图 1 基于皮尔逊系统的分段激活函数
Fig. 1 Sub-band activation function based on Pearson system

选用 y^3 函数作为分离矩阵的激活函数;当 $2.6 < \alpha_4 < \alpha_3^2 + 4$ 时,选用皮尔逊系统作为激活函数;当 $\alpha_4 > \alpha_3^2 + 4$ 时,语音信号服从超高斯分布,选用 \tanh 函数作为分离矩阵的激活函数.

通过矩估计,计算出信号的三阶矩和四阶矩,再根据上述方法,选择合适的激活函数实现语音信号的盲源分离.

2.3 算法的实现步骤

综上所述,基于分段激活函数的盲分离算法实现步骤如下:

- 1) 对观测信号 x 进行预处理,即去均值和白化;
- 2) 设置初始的分离矩阵 W 和初始学习步长 η ;
- 3) 利用公式 $Y(t) = WX(t)$ 计算分离信号 y_i ;
- 4) 用矩估计的方法计算信号的三阶矩和四阶矩,根据图 1 给出的分段激活函数的范围,选择合适的激活函数;
- 5) 将选择的激活函数代入式(2)中,求出分离矩阵 W_{k+1} ;
- 6) 判断分离矩阵 W_{k+1} 是否收敛,即 $|W_{k+1} - W_k| < \varepsilon$ 是否成立,其中, ε 是一个无限小的正数.若收敛,则输出分离信号;反之,则返回到第 3) 步,直到收敛;
- 7) 得到所有的估计值 y_i ,输出分离信号 $Y(t)$.

3 仿真实验及其性能分析

3.1 仿真实验

语音信号分离实验中,应用 Matlab 仿真软件,选取两路语音信号,其中一路为音乐信号,一路为谈话信号(图 2a)来进行分离.采样频率为 10 kHz,采样点数为 3×10^4 .先将两路语音信号按下列随机矩阵 A 进行混合,进而得到一组混合信号(图 2b),

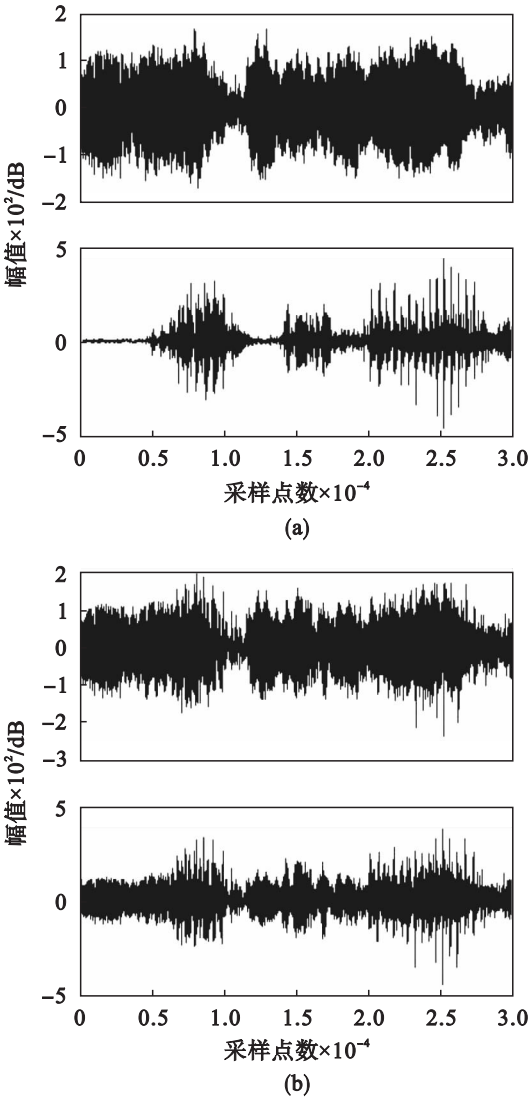


图 2 原始语音信号及其混合后语音信号
Fig. 2 The original signals and their mixed signals
(a)—原始两路语音信号;(b)—混合后语音信号.

$$A = \begin{bmatrix} 0.228\ 9 & 0.152\ 3 \\ 0.913\ 3 & 0.825\ 8 \end{bmatrix}. \tag{9}$$

分别采用传统的自然梯度算法以及基于皮尔逊系统的改进算法对混合后的两组语音信号进行分离,分离后的结果如图 3 所示.

直观上可以看到两种算法都可以实现语音信号的盲分离.

3.2 性能分析

本文选用分离后输出信号的均方误差作为评价指标^[8].定义均方误差如下:

$$MSE_i = \frac{1}{N} \sum_{k=1}^N [y_i(k) - s_i(k)]^2. \tag{10}$$

式中, N 表示采样点数.可以通过均方误差的变化来评价算法的分离性能,均方误差越小,说明分离的效果越好.两种算法的均方误差曲线如图 4 所示.

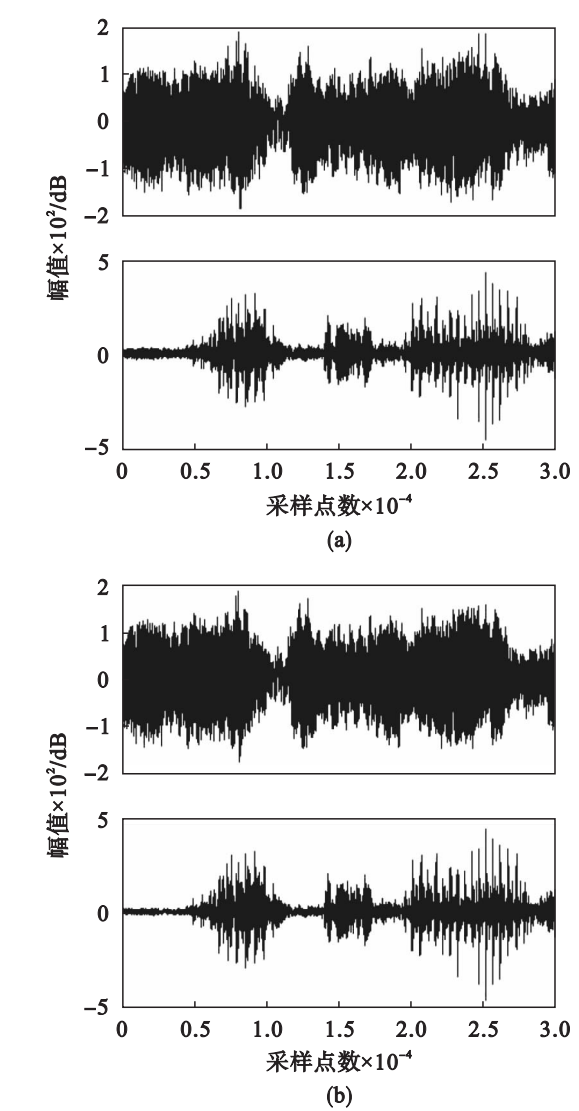


图 3 不同算法对语音信号的分离
Fig. 3 The signals separated with different algorithms
(a)—传统算法；(b)—改进算法。

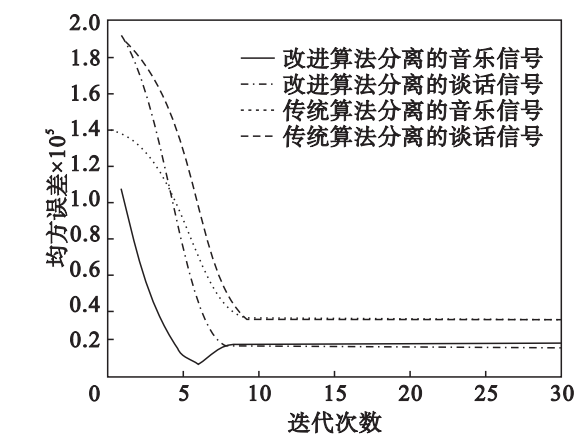


图 4 两种分离算法的均方误差曲线
Fig. 4 The mean square error curves of two algorithms

为了更加清楚直观地看到算法的优越性,给出两种算法收敛时的均方误差,原始算法和改进算法的均方误差分别为 1.25×10^{-6} 和 0.49×10^{-6} 。

从两种算法的均方误差值可以看出,基于皮尔逊系统的语音信号盲分离算法与传统的自然梯度算法相比,收敛速度基本保持不变;但是收敛后的均方误差值却明显要小很多,这说明改进后的算法在分离性能方面优于传统的自然梯度算法。

4 结 论

本文通过引入皮尔逊系统,将其作为一种激活函数与其他的激活函数相结合,提出一种新的分段激活函数. 利用矩估计的方法,不同语音信号根据自身特点选择合适的激活函数,代入分离矩阵进行盲分离. 与原始的自然梯度盲分离算法相比,改进的算法在基本保证收敛速度的同时,明显降低了输出信号的均方误差,即提高了算法的分离性能。

参考文献：

[1] Choi C H, Chang W, Lee S Y. Blind source separation of speech and music signals using harmonic frequency dependent independent vector analysis [J]. *Electronics Letters*, 2012, 48 (2) : 124 – 125.

[2] Liang Y, Chen G, Naqvi S M R, et al. Independent vector analysis with multivariate student's t-distribution source prior for speech separation [J]. *Electronics Letters*, 2013, 49 (16) : 1035 – 1036.

[3] Hesse C W, James C J. On semi-blind source separation using spatial constraints with applications in EEG analysis [J]. *IEEE Transactions on Biomedical Engineering*, 2006, 53 (12) : 2525 – 2534.

[4] Karvanen J, Koivunen V. Blind separation methods based on Pearson system and its extensions [C] // *Proceedings of the Second International Workshop on Independent Component Analysis and Blind Signal Separation*. Espoo, 2002 : 663 – 673.

[5] Amari S. Natural gradient works efficiently in learning [J]. *Neural Computation*, 1998, 10 (2) : 251 – 276.

[6] Karvanen J, Eriksson J, Koivunen V. Pearson system based method for blind separation [C] // *Proceedings of the Second International Workshop on Independent Component Analysis and Blind Signal Separation*. Espoo, 2000 : 585 – 590.

[7] 季策, 胡祥楠, 朱丽春, 等. 改进的高阶收敛 FastICA 算法 [J]. *东北大学学报: 自然科学版*, 2011, 32 (10) : 1390 – 1393. (Ji Ce, Hu Xiang-nan, Zhu Li-chun, et al. Improved higher order convergent FastICA algorithm [J]. *Journal of Northeastern University: Natural Science*, 2011, 32 (10) : 1390 – 1393.)

[8] 程舒慧. 动态盲源分离及其在生物医学信号处理中的应用研究 [D]. 合肥: 安徽大学, 2011. (Cheng Shu-hui. Research of dynamic blind source separation and its application to biomedical signal processing [D]. Hefei: Anhui University, 2011.)