

一种基于多层神经网络的盲信号分离算法^{*})

赵 峰¹ 李维勤² 张军英¹

(西安电子科技大学计算机学院 西安 710071)¹ (西安交通大学电子科学与技术系 西安 710049)²

摘 要 基于多层神经网络,提出一种盲信号分离算法。该算法不对信号的密度模型做任何假设,通过多层神经网络估计任意信号的概率密度函数,并由此估计信号的评价函数。同其他方法相比,该方法不仅具有更好的分离性能,而且收敛速度较快。该方法可直接应用于所有以非线性函数代替评价函数的盲信号分离算法。实验验证了方法的有效性。

关键词 盲信号分离,独立分量分析,概率密度估计

Blind Source Separation Algorithm Using Multilayer Networks

ZHAO Feng¹ LI Wei-Qin² ZHANG Jun-Ying¹

(School of Computer Science and Engineering, Xidian University, Xi'an 710071)¹

(Department of Electronic Science and Technology, Xi'an Jiaotong University, Xi'an 710049)²

Abstract An algorithm for blind source separation based on multilayer networks is proposed. Using a multilayer network density estimation technique, the algorithm may estimate the unknown probability density functions and its derivative of the score functions of the source signals and the algorithm is truly blind to the particular underlying distribution of the mixed signals. The new algorithm not only outperforms other methods, but also the approach has a better convergence property. The method can be applied to all the blind source separation algorithms where the score function is obtained by a nonlinear function. Simulation results show good performances of the proposed algorithm on both demixing and convergence to the desired solutions.

Keywords Blind source separation, Independent component analysis, Multilayer networks

近几年来,盲信号分离已成为信号处理学界和神经网络学界共同感兴趣的研究热点领域,并获得了迅速的发展^[1~6]。所谓盲信号分离,是指在不知源信号和传输通道参数的情况下,根据输入信源信号的统计特性,仅由观测信号恢复出源信号的过程。

普遍认为,盲信号分离算法的性能依赖于真实源信号的概率密度函数,但实际上这样的密度函数经常是未知的。提出了一些采用灵活模式估计源信号的密度函数^[5,7,8],这些方法虽然在某些情况下可以产生好的分离效果,但是算法的收敛性以及对于任意密度的信号估计都没有考虑^[9]。

常用的概率密度估计方法是核方法^[10]。然而该方法的问题是对 kernel 宽度选择很敏感,另一个缺点是在密度函数的低概率部分容易造成不平滑现象。

文[11]提出了一种基于多层神经网络的密度函数估计。和以前方法的区别是,该方法直接估计信号的概率分布函数,直接求导就可得到密度函数。Magdon-Ismail 指出该方法优于最优核参数的核方法^[11]。

本文利用多层神经网络估计信号的密度函数及导数,进而估计评价函数。这种方法是无参化的,不需要特别的概率模型,可用于所有以非线性函数代替评价函数的盲信号分离算法中,仿真实验获得了良好的分离效果。

1 盲信号分离原理

盲信号分离问题可以简单描述为:设有 n 个观测信号 x_i

($i=1,2,\dots,n$),每个观测信号是由 m 个独立源信号 s_j ($j=1,2,\dots,m$) 线性混合而来,则可记 $X=A \cdot S$ 。其中, $X=[x_1, x_2, \dots, x_n]^T$, $S=[s_1, s_2, \dots, s_m]^T$, 分别为观测信号和源信号, A 为混合矩阵。独立分量分析的目的就是找出分离矩阵 B , 使 $Y=B \cdot X$ 尽可能接近源信号 S 。

大多数盲信号分离算法最基本的原理是最小互信息算法^[12],即:

$$W_{opt} = \arg \min_W I(y_1, \dots, y_N) \quad (1)$$

应用信息理论知识,此表达式可以写为

$$\min_W \left\{ \sum_{i=1}^N H(y_i) - \log |\det W| - H(x) \right\} \quad (2)$$

由于 $H(x)$ 与 W 无关,目标函数可写为

$$L(W) = - \sum_{i=1}^N E[\log p_i(y_i)] - \log |\det W| \quad (3)$$

这里 $p_i(\cdot)$ 表示概率密度函数。一般 y_i 的评价函数通过下式给出:

$$\varphi(y_i) = \frac{p'(y_i)}{p(y_i)} \quad (4)$$

依据自然梯度迭代算法^[13],可得到

$$W(t+1) = W(t) + \eta \{ I - \varphi(y(t)) y^T(t) \} W(t) \quad (5)$$

在上面的源信号分离算法中,对评价函数源信号估计的近似程度直接影响分离的效果。传统的分离算法中,非线性函数往往只能凭经验选取。Cardoso 指出,对源分布不正确的假设将导致分离性能的降低,有时候甚至完全不能获得源信号^[4]。

^{*})国家自然科学基金(60574039,60371044)、国防预研项目(413070501)资助课题。赵 峰 博士,主研方向:智能信息处理、雷达目标识别;李维勤 博士,主研方向:智能信息处理、信源盲分离;张军英 教授,博导,主研方向:神经网络、智能信息处理。

2 基于多层神经网络的盲分离算法

2.1 密度函数估计

为了估计源信号的密度函数,我们采用基于多层神经网络的密度估计技术^[14]。让 $x_m \in R, m=1, \dots, M$, 代表给定的数据点。设 x_m 的概率密度函数是 $p(x)$, 分布函数是 $F(x) = \int_{-\infty}^x p(t) dt$ 。神经网络的输出是 $H(x, w)$, w 代表神经网络的权值。一旦训练完成, 则有 $H(x, w) = F(x)$ 。设 $y = F(x)$ 服从 $[0, 1]$ 的均匀分布, y 的密度函数为 $p_Y(y)$, 是网络的目标。

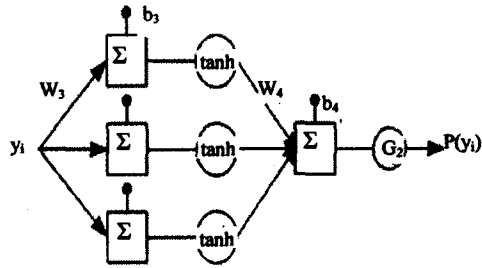
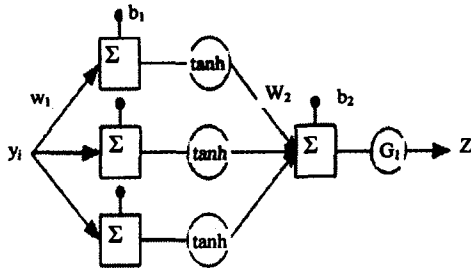


图1 神经网络模型

2.2 评价函数的估计

让 $y_m, m=1, 2, \dots, M$ 是神经网络的输入, 不失一般性, 我们假设 $y_{11} \leq y_{12} \leq \dots \leq y_{1M}$ 。让 $z_m, m=1, 2, \dots, M, z_1 \leq z_2 \leq \dots \leq z_M$ 服从 $[0, 1]$ 的均匀分布, 是网络的输出。

我们使用一个由三个神经元组成的单隐层神经网络, 隐层神经元的激活函数为 $\tanh(\cdot)$, 如图 1(a) 所示。一旦神经网络收敛, 则有

$$H_1(y_i, w) = G_1(b_2 + \sum_{i=1}^3 [w_2(i) \cdot \tanh(y_i \cdot w_1(i) + b_1(i))]) \quad (8)$$

故 y_i 的密度函数为

$$\begin{aligned} p(y_i) &= \partial H_1(y_i, w) / \partial y_i \\ &= G_1'(y_i, w) \cdot \sum_{i=1}^3 [\tanh'(w_1(i) \cdot y_i + b_1(i)) \cdot w_2(i) \cdot w_1(i)] \end{aligned} \quad (9)$$

为了估计 $p'(y_i)$, 直接对 $p(y_i)$ 求一阶导数, 或者数值微分方法往往不准确。由于神经网络具有很好的逼近非线性函数的能力, 这里我们也使用相同的神经网络来估计 $p'(y_i)$ 。

让 $y_m, m=1, 2, \dots, M$ 是网络的输入, $p(y_{11}), p(y_{12}), \dots, p(y_{1M})$ 是网络的输出, 如图 1(b) 所示。当神经网络收敛时, 有

$$H_2(y_i, w) = G_2(b_4 + \sum_{i=1}^3 [w_4(i) \cdot \tanh(y_i \cdot w_3(i) + b_3(i))]) \quad (10)$$

$$\begin{aligned} p'(y_i) &= \partial H_2(y_i, w) / \partial y_i \\ &= G_2'(y_i, w) \cdot \sum_{i=1}^3 [\tanh'(w_3(i) \cdot y_i + b_3(i)) \cdot w_4(i) \cdot w_3(i)] \end{aligned} \quad (11)$$

最终, 我们发现评价函数为

$$\begin{aligned} \varphi(y_i) &= \frac{G_2'(y_i, w) \cdot \sum_{i=1}^3 [\tanh'(w_3(i) \cdot y_i + b_3(i)) \cdot w_4(i) \cdot w_3(i)]}{G_1'(y_i, w) \cdot \sum_{i=1}^3 [\tanh'(w_1(i) \cdot y_i + b_1(i)) \cdot w_2(i) \cdot w_1(i)]} \end{aligned} \quad (12)$$

3 实验与结果

信号源如图 2 所示。信号的归一化峭度分别为 0.5149, -0.3754, 0.2450, 0, 每个信号源的样本个数都取 1000。实验中取本文的算法以及扩展最大熵化算法 (Entended-Informax)^[4], 两种算法的学习率都取 $\eta=0.05$ 。实验中, 我们使用

显然有下式成立:

$$p_Y(y) = \frac{p(x)}{\left| \frac{dy}{dx} \right|} = \frac{p(x)}{\left| \frac{dF(x)}{dx} \right|} = 1, 0 \leq y \leq 1 \quad (6)$$

因此, 当输入信号的密度函数为 $p(x)$ 时, 假如 $H(x, w)$ 近似为 $F(x)$, 则神经网络的输出服从 $[0, 1]$ 的均匀分布。我们训练神经网络, 使得网络的输出服从 $[0, 1]$ 的均匀分布。一旦神经网络训练完成, $H(x, w)$ 近似为 $F(x)$, 则有

$$p(x) = \partial H(x, w) / \partial x \quad (7)$$

只有一个隐层、三个神经元的网络结构。其中, 隐层的激励函数为 $\log \text{sig} = 1 / (1 + \exp(-x))$, 训练样本随机抽取 100 个。图 3 所示为随机混合后的信号, 图 4 所示为本文算法迭代 1000 次的分离结果, 图 5 是最大熵化算法相对应的分离结果。从图中可以看到, 我们的算法比较好地恢复了源信号。

为比较以上两种方法的收敛性能, 一般采用对广义矩阵 $Q = (q_{ij}) = W \cdot A$ 的误差作为算法性能的度量^[14], 即

$$\text{Error} = \sum_{i=1}^n \left(\sum_{j=1}^n \frac{|q_{ij}|}{\max_k |q_{ik}|} - 1 \right) + \sum_{j=1}^n \left(\sum_{i=1}^n \frac{|q_{ij}|}{\max_k |q_{kj}|} - 1 \right) \quad (13)$$

图 6 给出了两种算法的性能随迭代次数变化的情况。从所获得的误差曲线可以看出, 扩展最大熵法和本文提出的算法均收敛, 但本文的算法有更快的收敛速度, 表明了我们算法的优越性。

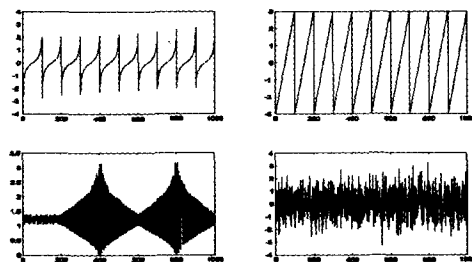


图2 源信号

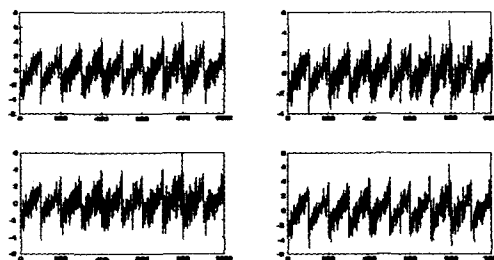


图3 混合信号

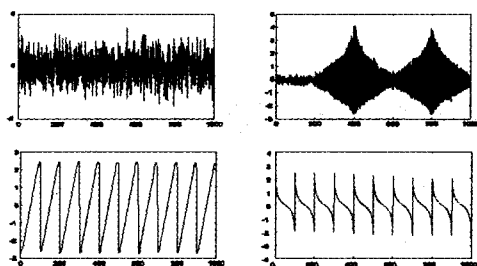


图4 本文算法的分离结果

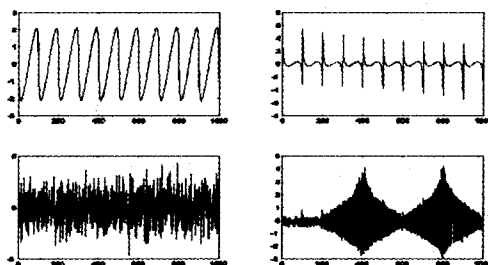


图5 最大熵算法的分离结果

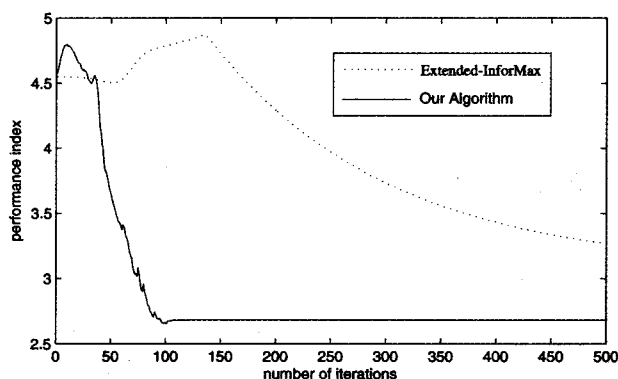


图6 本文的方法和扩展最大熵化算法的收敛曲线

结论 本文提出一种有效的估计评价函数的方法,并把该方法用于盲信号分离算法。该方法不需要源信号的概率密度模型,通过多层神经网络估计任意信号的概率密度函数,并由此估计信号的评价函数。该方法有能力估计任意信号的密度,具有很好的收敛特性,特别适合真实信号的混合。实验验证了其在分离性能和收敛速度上的有效性。

参考文献

- 1 Comon P. Independent component analysis, a new concept? *Signal Processing*, 1994, 36(3): 287~314
- 2 Bell A J, Sejnowski T. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 1995, 7(6):1129~1159
- 3 Nadal J P, Parga N. Non-linear neurons in the low-noise limit: A factorial code maximizes information transfer. *Network*, 1994, 4: 295~312
- 4 Cardoso J F. Blind signal separation; Statistical principles. *Proc IEEE Special Issue on Blind Identification and Estimation*, 1998, 9: 2009~2025
- 5 Lee T W, Girolami M, Sejnowski T J. Independent component analysis using an extended InfoMax algorithm for mixed sub-Gaussian and super-Gaussian sources. *Neural Computat*, 1999, 11(2):417~441
- 6 Amari S I, Chen T P, Cichocki A. Stability analysis of learning algorithms for blind source separation. *Neural Networks*, 1997, 10(8): 1345~1351
- 7 Karvanen J, Eriksson J, Koivunen V. Pearson system based method for blind separation. In: *Proc. Second Int Workshop on Independent Component Analysis and Blind Signal Separation*, 2000. 585~590
- 8 Vlassis N, Motomura Y. Efficient source adaptivity in independent component analysis. *IEEE Trans Neural Networks*, 2001, 12: 559~566
- 9 Boscolo R, Pan H, Vwani P. Roychowdhury. Independent Component Analysis Based on Nonparametric Density Estimation. *IEEE Trans Neural Networks*, 2004, 15(1): 154~161
- 10 Silverman B W. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1985
- 11 Magdon-Ismael M, Atiya A. Density Estimation and Random Variate Generation Using Multilayer Networks. *IEEE Trans Neural Networks*, 2002, 13(3)
- 12 Barlow H B. Unsupervised learning. *Neural Computation*, 1989, 1: 295~311
- 13 Cichocki A, Unbehauen R. Robust Neural Networks with On-line Learning for Blind Identification and Blind Separation of Sources. *IEEE Trans Circuits and Systems-I: Fundamental Theory and Applications*, 1996, 43: 894~906
- 14 Amari S, Cichocki A, Yang H H. A new learning algorithm for blind signal separation. In: *Touretzky D S, Mozer M C, Hasselmo M E, eds. Advances in Neural Information Processing Systems*, Cambridge, MA: MIT Press, 1996, 8(1): 757~7636

(上接第180页)

可见,运算结果输出可以正确地传递给下一级门电路信号Y的输入。那么,运算结果又是如何传递给下一级门电路信号X的输入是我们解决的问题。这里只需要增加一次“或”运算 $X=0 \vee F$ 即可,逻辑值“0”是非常容易实现的,就是初始化的状态。不难发现,如果下一级运算是“或”运算,则传递给输入信号X和Y的溶液可以直接混合,无需单独进行输入信号X的传递过程。所以,用多发夹结构的DNA分子表示的多值逻辑电路,采用“与-或”逻辑结构,是最简单的,不仅“与”门运算结果的输出无需求补转化,而且“或”门的所有输入信号都可以同时传递。

结束语 本文首次给出了一个基于DNA分子计算的三值逻辑电路模型,阐述了“与”、“或”运算的生物实现过程。因为这种发夹型DNA计算模型是基于表面方式的,所以随着生物芯片技术的飞速发展,构建这种基于DNA分子的多值逻辑计算芯片将成为可能。由于DNA分子计算本身的优点,该模型用于制造大规模低能耗的逻辑阵列电路是非常适合的,但是模型也有不足之处,(1)“与”运算过程中存在杂交竞争的问题;(2)运算过程中利用加热来解链的方法有待改进,例如文[4]的诱导型“发夹”结构,在表面上加入或清洗吡

啉二聚物,“发夹”闭合与打开非常容易控制。另外,由于DNA分子计算本身的复杂性和不确定性,模型的具体应用有待生物技术的进一步提高。

参考文献

- 1 Post E L. Introduction to a general theory of elementary propositions [J]. *Amer. J. Math.*, 1921, 43:163~185
- 2 Adleman L, et al. Molecular computations to combinatorial problems [J]. *Science*, 1994, 266:1021~1024
- 3 刘文斌,朱翔鸥,王向红,陈丽春. DNA 计算的研究进展[J]. *电子学报*, 2006, 34(11):2053~2057
- 4 Liu Wenbin, et al. A new DNA computing model for the NAND gate based on induced hairpin formation [J]. *BioSystems*, 2004, 77(1-3):87~92
- 5 Cukras A R, Faulhammer D, Lipton R J, Landweber L F. Chess games: A model for RNA-based computation [J]. *Biosystems*, 1999, 52:35~45
- 6 Yin Zhixiang, Zhang Fenyue, Xu Jin. DNA computing based on molecular beacons [J]. *Journal of Biomathematics*, 2003, 4(18): 497~501(in Chinese)
- 7 Nakatani K, Sando S, Saito I. Improved Selectivity for the Binding of Naphthylridine Dimer to Guanine-Guanine Mismatch. *Bioorg [J]. Med. Chem. Soc.*, 2001, 9:2381~2385
- 8 Nakatani K, Sando S, Saito I. Scanning of guanine-guanine mismatches in DNA by synthetic ligands using surface plasmon resonance assay [J]. *Nat. biotechnol.*, 2001, 19:51~55