

基于独立分量分析的降噪技术

张智林¹, 皮亦鸣¹, 孙志坚²

(1. 电子科技大学电子工程学院 成都 610054; 2. 青岛理工大学理学院 山东 青岛 266033)

【摘要】介绍了新兴的独立分量分析技术的基本概念和原理,以及具有代表性的算法,即FastICA算法、EASI算法、非线性PCA算法和基于自然梯度的最大似然估计算法。通过降噪仿真实验,并采用均方误差作为降噪的性能指数,对这些算法与传统的自适应信号处理算法进行比较。所得实验结果表明,独立分量分析算法在降噪上的效果优于自适应信号处理算法。因此在降噪上具有较大的应用价值。

关键词 独立分量分析; 降噪; 自适应信号处理; 非线性PCA; 最大似然估计
中图分类号 TN911.4 文献标识码 A

Independent Component Analysis Based Denoising Technology

ZHANG Zhi-lin¹, PI Yi-ming¹, SUN Zhi-jian²

(1. School of Electronic Engineering, UEST of China Chengdu 610054;

2. College of Science, Qingdao Technological University Shandong Qingdao 266033)

Abstract The paper introduces a new technology of signal processing: independent component analysis, including its basic concept, principles, and some representative algorithms, such as FastICA, EASI, Nonlinear PCA, and natural gradient algorithm based maximum likelihood estimation. In a denoising simulation experiment with the mean square error criterion, these algorithms are compared to the classic algorithms of adaptive signal processing, such as LMS and RLS. Results show that in denoising application ICA algorithms are superior to the classic adaptive algorithms. Thus ICA algorithms have large value in denoising application, deserving further study and promoting.

Key words independent component analysis; denoising; adaptive signal processing; nonlinear PCA; maximum likelihood estimation

现实世界中所获得的信号往往是源信号与噪声的混合。为了从接收到的信号中滤除噪声而恢复源信号,多年来发展了很多技术,比如利用自适应信号处理技术进行噪声的对消^[1,2]。独立分量分析(Independent Component Analysis, ICA)是近年来兴起的信号处理技术^[3]。ICA假设各源信号之间是独立的,其目的是把接收到的混合信号分解为相互独立的成分,而分离出的各成分即为源信号。由于现实世界中有用信号与噪声往往是独立的,因此ICA在降噪上有着很大的潜力。

1 ICA的基本概念、原理和算法

1.1 ICA的基本概念

ICA基本模型可描述为:

收稿日期: 2003-12-30

基金项目: 国家高校博士点专项科研基金资助项目(20030614001)

作者简介: 张智林(1980-),男,硕士生,主要从事盲信号处理、通信信号处理、自适应信号处理方面的研究。

$$\mathbf{x}(t) = \mathbf{A}s(t) \quad (1)$$

式中 \mathbf{A} 为 $N \times N$ 的混合矩阵; $\mathbf{x}(t)$ 为 $V \times 1$ 维观测信号矢量; $s(t)$ 为 N 个统计独立的源信号组成的列向量。ICA 的目的就是寻求一线性变换 \mathbf{W} (即分离矩阵), 通过它能由观测信号 $\mathbf{x}(t)$ 恢复源信号 $s(t)$:

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) = \mathbf{W}\mathbf{A}s(t) \quad (2)$$

式中 $\mathbf{y}(t)$ 即为 $s(t)$ 的估计矢量(为了书写方便, 以下省略时间序号 t)。在 ICA 基本模型中作了两个基本的假设^[4]: 1) N 个源信号之间瞬时统计独立; 2) 源信号中至多只有一个信号为高斯分布。

1.2 ICA 的基本原理和典型算法

1.2.1 负熵最大化和 FastICA 算法

文献[4, 5]提出了快速固定点 ICA 算法。它是一种快速收敛、鲁棒的离线(off-line)批处理算法。该算法所基于的判据是负熵最大化, 其定义如下:

$$J(\mathbf{y}) = H_G(\mathbf{y}) - H(\mathbf{y}) \quad (3)$$

式中 $\mathbf{y} = \mathbf{W}^T \mathbf{z}$, \mathbf{W} 是分离矩阵, \mathbf{z} 为白化后的观测列向量, $H(\mathbf{y})$ 是随机矢量 \mathbf{y} 的联合微分熵, $H_G(\mathbf{y})$ 是与 \mathbf{y} 具有相同协方差阵的高斯分布的微分熵。由于直接计算该公式很复杂, 因此提出了逼近负熵的判据:

$$J(\mathbf{y}) = [E\{G(\mathbf{y})\} - E\{G(\mathbf{v})\}]^2 \quad (4)$$

式中 \mathbf{v} 为一零均值、单位方差的高斯变量, G 为非线性函数, 文中取如下形式:

$$G(\mathbf{y}) = \frac{1}{a} \lg \cosh(a\mathbf{y}) \quad (5)$$

式中 $1 < a < 2$, 常常取 $a=1$ 。在 \mathbf{W} 正交的约束条件下, 通过对式(4)求极大值, 可解得:

$$\mathbf{w}_i \leftarrow E\{z\mathbf{g}(\mathbf{w}_i^T \mathbf{z})\} - E\{g'(\mathbf{w}_i^T \mathbf{z})\} \mathbf{w}_i \quad (6)$$

$$\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^T)^{-1/2} \mathbf{W} \quad (7)$$

这里 $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N)$, g 为 G 的导数。

1.2.2 EASI 算法

文献[6]在 1996 年提出了 EASI 算法。其归一化算法如下:

$$\mathbf{W}(t+1) = \mathbf{W}(t) - \lambda(t) \left[\frac{\mathbf{y}(t)\mathbf{y}(t)^T - \mathbf{I}}{1 + \lambda(t)\mathbf{y}(t)^T \mathbf{y}(t)} + \frac{\mathbf{g}(\mathbf{y}(t))\mathbf{y}(t)^T - \mathbf{y}(t)\mathbf{g}(\mathbf{y}(t))^T}{1 + \lambda(t)|\mathbf{y}(t)^T \mathbf{g}(\mathbf{y}(t))|} \right] \mathbf{W}(t) \quad (8)$$

式中 \mathbf{y} 、 \mathbf{W} 含义同上。 $\lambda(t)$ 为步长因子, 可为固定的标量。非线性函数可取为: $g(\cdot) = \tanh(\cdot)$ 。

1.2.3 采用自然梯度的极大似然估计算法^[3,5]

由 ICA 基本模型 $\mathbf{x}(t) = \mathbf{A}s(t)$, 令矩阵 \mathbf{B} 为 \mathbf{A} 的逆矩阵, 得到 \mathbf{x} 的概率密度函数, 其对数似然函数为:

$$\lg L(\mathbf{B}) = \sum_{t=1}^T \sum_{i=1}^n \lg p_i(\mathbf{b}_i^T \mathbf{x}(t)) + T \lg |\det \mathbf{B}| \quad (9)$$

式中 $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)^T$ 。最大化对数似然函数, 可得到关于参数 \mathbf{B} 的最佳估计。似然函数是独立分量的概率密度的函数。这使得该问题非常的复杂, 于是采用间接方法, 即由有限的参数确定一类概率密度函数, 用这类函数来逼近独立分量的密度函数, 这样就得到了极大似然估计算法: 首先计算 $\gamma_i = (1 - \mu_r)\gamma_i + \mu_r E\{-\tanh(y_i)y_i + (1 - \tanh(y_i))^2\}$, 其中 $\mathbf{y} = \mathbf{B}\mathbf{x}$, $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$ 。当 $\gamma_i > 0$ 时, 则定义 $g_i(y_i) = -2 \tanh(y_i)$, 否则定义 $g_i(y_i) = \tanh(y_i) - y_i$ 。再计算得:

$$\mathbf{B} = \mathbf{B} + \mu[\mathbf{I} + \mathbf{g}(\mathbf{y})\mathbf{y}^T] \mathbf{B} \quad (10)$$

若不收敛, 则再次循环迭代直至收敛。这里 $\mathbf{g}(\mathbf{y}) = (g_1(y_1), g_2(y_2), \dots, g_n(y_n))^T$ 。

1.2.4 非线性 PCA 的 RLS 算法^[4]

该算法是在 Yang Bin 提出的 PAST 算法中引入非线性函数而得的, 算法中采用了自适应滤波技术中的递归最小二乘算法:

$$\mathbf{q}(t) = \mathbf{g}(\mathbf{W}(t-1)\mathbf{z}(t)) = \mathbf{g}(\mathbf{y}(t)) \quad (11)$$

$$\mathbf{h}(t) = \mathbf{P}(t-1)\mathbf{q}(t) \quad (12)$$

$$\mathbf{m}(t) = \mathbf{h}(t) / [\beta + \mathbf{q}^T(t)\mathbf{h}(t)] \quad (13)$$

$$\mathbf{P}(t) = \frac{1}{\beta} \text{Tri}[\mathbf{P}(t-1) - \mathbf{m}(t)\mathbf{h}^T(t)] \quad (14)$$

$$\mathbf{r}(t) = \mathbf{z}(t) - \mathbf{W}^T(t-1)\mathbf{q}(t) \quad (15)$$

$$\mathbf{W}(t) = \mathbf{W}(t-1) + \mathbf{m}(t)\mathbf{r}^T(t) \quad (16)$$

式中 $\mathbf{z}(t)$ 、 $\mathbf{W}(t)$ 、 $g(\cdot)$ 含义均同上, 在实验中选取 $g(y) = y - \tanh(y)$, 记号 $\text{Tri}(\cdot)$ 表示把矩阵的上三角的元素复制到对应的下三角部分, β 是遗忘因子, 小于或等于1。 \mathbf{W} 和 \mathbf{P} 的初始化值可取单位矩阵。

2 ICA算法与自适应算法在降噪上的对比实验

文献[2]等人曾做过用自适应算法消除语音干扰的实验, 来展现自适应算法降噪的优势。本文参照其思路设计了类似的降噪实验, 但分别采用经典自适应算法和ICA典型算法进行降噪, 并以均方误差为降噪的指标, 以此来比较自适应算法与ICA算法的降噪性能。本文中采用的自适应算法是LMS和RLS算法; ICA算法是FastICA, EASI, 非线性PCA的RLS算法(NonPCA), 以及采用自然梯度的极大似然估计算法(NGML)。有用信号为周期曲线信号, 噪声为高斯白噪声。以降噪后的有用信号与真实有用信号之差的均方作为指标。取50次实验结果的平均值作为最终结果。

对于自适应算法, 主通道中SNR为10 dB, 参考通道SNR为-10 dB, 自适应滤波器有16个加权系数, 经过自适应处理后的信号是对消了噪声的有用信号。LMS算法中, 步长因子为0.01。RLS算法中, 遗忘因子为0.95, 逆矩阵初始值为 $50\mathbf{I}$, \mathbf{I} 为16阶的单位矩阵^[1]。对于ICA算法, 建模如下:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{A}\mathbf{s} = \begin{bmatrix} 1 & 0.3162 \\ 0.3162 & 1 \end{bmatrix} \begin{bmatrix} s \\ n \end{bmatrix} \quad (17)$$

式中 s 为有用信号(零均值单位方差), n 为零均值单位方差的高斯白噪声。混合矩阵的取值使得混合后的观测信号 x_1 的SNR为10 dB, x_2 的SNR为-10 dB。这样就和自适应算法具有可比性。分离出的两个信号之一必定是有用信号。ICA的各种算法的初始参数值及所选择的非线性函数, 均为1.2小节中的取值。

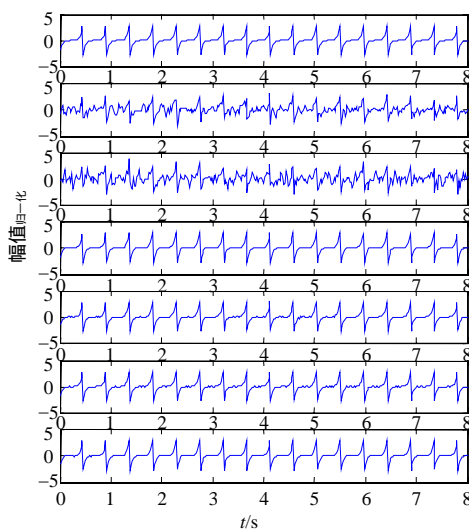


图1 恢复信号波形

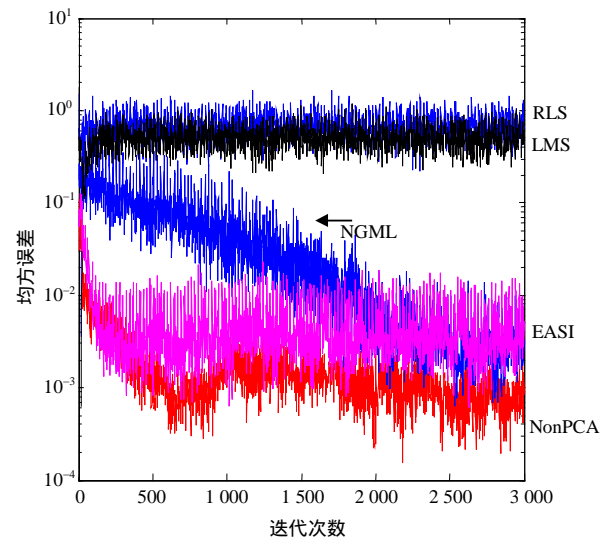


图2 各种算法的学习曲线

实验结果的波形如图1所示, 最上一行是原始波形, 往下依次是采用LMS, RLS, FastICA, EASI, NonPCA, NGML算法降噪后的波形。图2所示绘出了各种算法的迭代次数与均方误差的关系(曲线起伏较大是因为有用信号为周期信号)。每种算法的均方误差比较如表1所示。表2列出了各种算法在线收敛时大约所需要的运算量(FastICA是离线算法, 故没有比较)。

由实验结果不难发现: ICA算法在降噪效果上远远优于经典的自适应算法, 在中低信噪比(10dB)情况下

甚至高出2~3个数量级。在降噪效果上,最好的是FastICA,然后依次是基于PIS的非线性PCA的RLS算法,采用自然梯度的最大似然估计算法,以及EASI算法。相比之下经典的自适应算法效果最差。从在线收敛的计算量来看,自适应算法所需的计算量相对要少些。但EASI算法所需的计算量少于RLS算法。

此外还发现,在低信噪比(低于10dB)的情况下,采用自适应技术效果往往不佳,但采用ICA方法可以获得非常好的效果。

表1 自适应算法与ICA算法的比较(50次实验平均)

	LMS	RLS	FastICA	EASI	NonPCA	NGML
均方误差	0.518 3	0.678 1	0.000 5	0.003 7	0.001 2	0.002 0

表2 在线收敛时的大致计算量(50次实验平均)

	LMS	RLS	EASI	NonPCA	NGML
加减运算	2 000	17 000	3 000	9 000	30 000
乘除运算	2 000	26 000	7 000	24 000	53 000

3 结 束 语

通过本文的实验比较,可知ICA在降噪上有着很大的潜力。目前在工频干扰消除,图像噪声消除,生物医学噪声消除等方面都展开了ICA应用的研究,但还有待更深入的工作。比如提高算法的收敛速度,减少计算量,针对某些具体场合对ICA基本模型进行适当修改以及如何能消除乘性噪声等,都是今后研究的方向。

参 考 文 献

- [1] 龚耀寰. 自适应滤波——时域自适应滤波和智能天线, 第二版[M]. 北京: 电子工业出版社, 2003. 92-95, 420-421
- [2] Widrow B. Adaptive noise canceling: principles and applications[J]. Proc. IEEE, 1975,63(12): 1 692-1 716
- [3] Hyvärinen A, Oja E. Independent component analysis: algorithms and applications[J]. Neural Networks, 2000, 13(4-5): 411-430
- [4] Hyvärinen A, Karhunen J, Oja E. Independent component analysis[M]. New York: John Wiley & Sons, Inc, 2001:147-289
- [5] Hyvärinen A. Fast and robust fixed-point algorithm for independent component analysis[J]. IEEE Trans. On Neural Networks, 1999,10(3): 626-634
- [6] Cardoso J F, Hvam L B. Equivariant adaptive source separation[J]. IEEE Trans. On Signal Processing, 1996, 44 (12): 3 017-3 030
- [7] Cardoso J F. Informax and maximum likelihood for source separation[J]. IEEE Signal Processing Letters, 1997,4:112-114

编 辑 孙晓丹