

ICA 去除 EEG 中眼动伪差和工频干扰方法研究

万柏坤,朱 欣,杨春梅,高 扬

(天津大学生物医学工程与科学仪器系,天津 300072)

摘 要: 眼动伪差和工频干扰是临床脑电图(EEG)中常见噪声,严重影响其有用信息提取.本文尝试采用独立分量分析(Independent Component Analysis,ICA)方法分离 EEG 中此类噪声.通过对早老性痴呆症(Alzheimer disease,AD)患者临床 EEG 信号(含眼动伪差和混入工频干扰,信噪比仅 0dB)作 ICA 分析,比较了最大熵(Infomax)和扩展最大熵(Extended Infomax)ICA 算法的分离效果,证实虽然最大熵算法可以分离出眼动慢波,但难以消除工频干扰,为此需采用扩展的最大熵算法;并知 ICA 方法在极低信噪比时也有较好的抗干扰性,且在处理非平稳信号时有好的鲁棒性;文中还结合近似熵(approximate entropy,ApEn)分析说明利用 ICA 去除干扰后有助于恢复和保持原始 EEG 信号的非线性特征.研究表明 ICA 方法在生物医学信号处理中具有潜在的重要应用价值,值得深入研究和推广.

关键词: 脑电(EEG);眼动伪差;工频干扰;独立分量分析(ICA);最大熵(Infomax)

中图分类号: TH772+.2 **文献标识码:** A **文章编号:** 0372-2112(2003)10-1571-04

Study on Applying Independent Component Analysis to Remove Blink Artifacts and Power Noise in EEG

WAN Bai-kun, ZHU Xin, YANG Chun-mei, GAO Yang

(Department of Biomedical Engineering, Tianjin University, Tianjin 300072, China)

Abstract: Blink artifacts and power noise are constantly found to strongly influence the acquisition and analysis of EEG signals. In this paper, by comparing the efficiencies of two ICA algorithms—Infomax ICA and Extended Infomax ICA methods in extracting blink artifacts and power noise in the EEG signals, it was shown that ICA algorithms were insensitive to disturbance in the conditions of low signal-noise-ratio, and ICA algorithms demonstrated a strong robustness in processing non-stationary signals. Though blink slow waves could be extracted by infomax algorithm, but power noise was unlikely to be removed by it. Therefore, Extended Infomax ICA algorithm should be used. By applying Extended Infomax algorithms, blink artifacts and power noise contained in the 16-channel EEG signals of Alzheimer-disease patients were removed successfully (the lowest signal-noise-ratio for power noise can be -40dB). Meanwhile, it proved by calculating approximation entropy (ApEn) that ICA algorithms could preserve the nonlinear characteristics of EEG after removing the interference.

Key words: EEG; blink artifacts; power noise; independent component analysis; infomax

1 引言

临床脑电图(electroencephalograms, EEG)含有反映大脑皮层神经电气活动随时间-空间变化的信息,已成为医学临床诊断脑部疾病和观察大脑功能的常规手段.但 EEG 信号具有非平稳性,且十分微弱(幅值仅 μV 量级),极易受到肌电和工频干扰,严重影响其有用信息的提取和分析.因此,寻找鲁棒性能好,分离效率高的消除噪声方法成为人们关注的研究目标.

独立分量分析(Independent Component Analysis, ICA)是近年研究开发的信源分解技术,其基本思路是将观测信号按照统计独立的原则通过优化算法分解为若干独立来源成分.将 ICA 用于消除 EEG 中干扰噪声和模式提取已成为近年的研究热点^[1-5].

本文尝试采用最大熵(Infomax)^[6]和扩展最大熵(Extended Infomax)^[7]ICA 算法消除 EEG 中眼动伪差和工频干扰,并对于算法在极低信噪比和 EEG 非平稳特性影响下的去噪效果进行分析 and 评价,同时还结合近似熵(approximate entropy, ApEn)分析讨论经 ICA 去除干扰噪声后对 EEG 信号非线性特征的影响,以期优选出高效适用的 ICA 消噪方法.

2 ICA 算法原理

假设存在 N 道相互独立的未知源信号 $X = [x_1, x_2, \dots, x_N]$, 通过某个未知的待检测分析系统后线性叠加为在 M 个传感器上接收到的观测信号 $U = [u_1, u_2, \dots, u_M]$. 已经证明^[8]在 $M \geq N$ 条件下,如果源信号 X 不含一个以上的高斯

过程则存在混合矩阵 A , 使得 $U = AX$. ICA 算法的基本思路在于求解一矩阵 W , 使其作用于观测信号 U 所得估计信号 $X' = WU$, 在统计独立的意义下最逼近于未知源信号 X , 即要找到矩阵 W 使得信号 X 的各个分量尽可能地相互独立. 为此需建立一个合适的代价函数, 再采用某种优化算法分离源信号. 对于 EEG 信号, 代价函数可取各导联信号间的高阶统计量、互信息熵或最大似然估计, 优化算法则可采用牛顿迭代法、基于神经网络的自适应算法、随机梯度法、自然梯度法等^[5].

本文所采用的最大熵 (Infomax)^[6] 算法是根据常规的随机梯度法求解矩阵 W , 算法的目标判据为独立分量通过非线性环节后信息熵极大, 其调节公式为

$$\Delta W = \mu [(W^T)^{-1} + (1 + 2y) x^T] \quad (1)$$

由于算法中涉及矩阵求逆, 故又提出了用自然梯度代替常规梯度的改进算法, 以避免求逆过程, 减少计算量. 如此, 矩阵调节改为下式

$$\Delta W = \mu [(W^T)^{-1} + (1 + 2y) x^T] W^T W = \mu [I + (1 - 2y) u^T] W \quad (2)$$

其计算量和收敛速度均得到较大改进.

但上述算法仅适用于超高斯信号, 而眼动肌电和工频干扰之类噪声多属亚高斯信号. 为了使算法能同时适用于亚高斯信号, 又推导出所谓扩展的最大熵 (Extended Infomax)^[7] 算法, 其调节公式为

$$\Delta W = \mu [I - \tanh(u) u^T - uu^T] W, \quad k_{ii} = 1 (\text{超高斯信号}), \quad k_{ii} = -1 (\text{亚高斯信号}) \quad (3)$$

式中 K 为对角阵. 算法通过其主对角元素 K_{ii} 取不同值来区分超高斯和亚高斯信号, 以便将脑电数据中 EEG 有用成分和肌电、工频等干扰信号分解开来. 而简单的最大熵算法则难以做到这点.

3 实验数据和分析结果

本文所用脑电数据由天津医科大学总医院老年病室提供 (共 7 例早发性痴呆 (Alzheimer disease, AD) 病人), 采自天津医科大学总医院脑电图室的丹麦丹迪公司产 32 导联脑电检测系统. 实验中按照国际标准导联 10 - 20 系统在受试者头皮上安放 F_{p1} 、 F_{p2} 至 T_5 、 T_6 等 8 或 16 个导联电极 (参考电极置于左右耳垂), 同时采集脑电数据. A/D 采样频率为 512Hz, 原始数

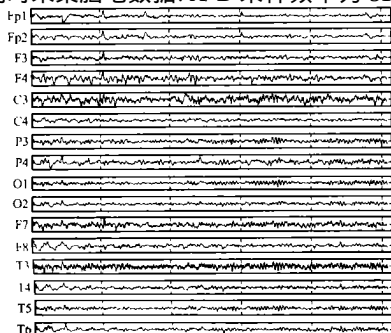


图 1 典型 AD 患者的原始脑电图, 其中 F_{p1} 和 F_{p2} 导联含有眼动干扰

据采集时间约 5 分钟, 选择脑电中无显著关机干扰的若干数据段 (每例约 4 - 5 段, 每段长度为 2560 点, 5 秒, 共 32 段) 供分析用. 图 1 所示为典型 AD 患者的脑电图 (仅给出 16 导联波形). 有经验的脑电图师可以看出 F_{p1} 和 F_{p2} 导联均受到眼动慢波干扰 (眼部水平运动), 而且 F_{p1} 导联还受到左眼眨眼所引起的肌电干扰; 图 2 是在图 1 数据中随机地叠加了 50Hz 工频干扰 (信噪比为 0dB).

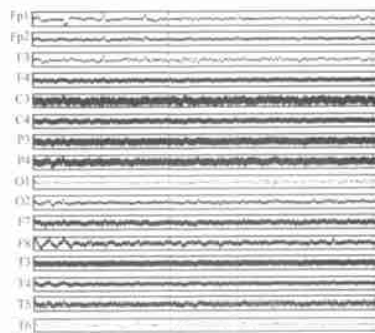


图 2 在图 1 脑电波形中随机加入工频干扰后的脑电图 (信噪比为 0dB)

分别采用上述最大熵算法和扩展最大熵算法对图 2 中脑电数据进行 ICA 分析 (计算中取 $M = N = 16$), 结果分别如图 3 和图 4 所示.

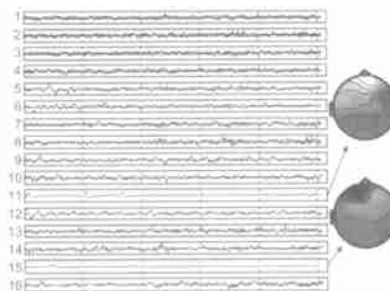


图 3 对图 2 采用最大熵算法分离所得 16 个独立分量, 箭头所示为相应分量的脑地形图.

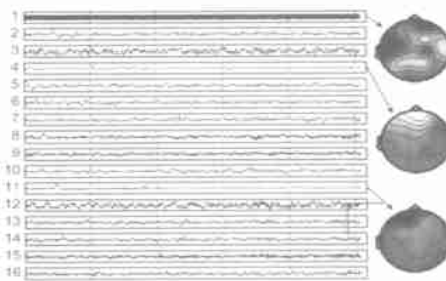


图 4 对图 2 采用扩展最大熵算法分离所得 16 个独立分量, 箭头所示为相应分量的脑地形图.

由图 3 可以看出, 最大熵算法分离出的第 11 分量为水平眼动慢波, 相应地形图 (箭头所指) 表明该分量集中在眼部 (白电平所示); 而第 15 分量则为左眼上下运动干扰, 相应地形图表明该分量集中在左眼侧 (图中黑电平所示); 但该算法未见分离出能对应 50Hz 工频干扰的分量.

图 4 是扩展最大熵算法的分离结果. 由图中第 4 分量和箭头所指地形图可知该分量对应于水平眼动慢波 (相当于图 3 中第 11 分量); 图中第 11 分量和相应的地形图表明该分量为左眼上下运动干扰 (相当于图 3 中第 15 分量); 而图 4 中第 1 分量为脑电数据中最高频率成分, 当为 50Hz 工频干扰 (箭头所示为其对应的脑地形图). 由于干扰是随机独立地加入各个导联信号中, 不满足空间连续性, 因而分离结果略有波动.

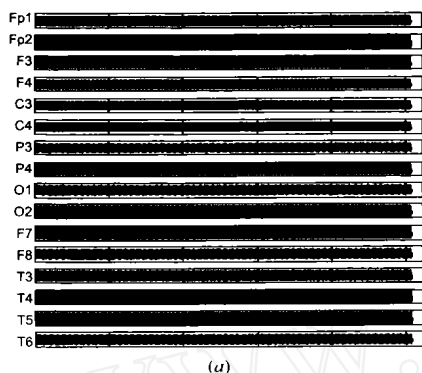
对照图 3 和图 4 可知, 虽然最大熵算法可有效地分离出属超高斯信号的眼动干扰, 但对属亚高斯信号的工频干扰则难以奏效, 而只有采用扩展最大熵算法才能同时将这两类噪

声同时分离出来.

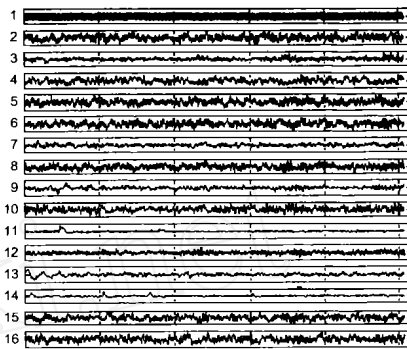
4 ICA 算法性能和讨论

4.1 算法的抗干扰性

为进一步考查 ICA 算法的抗干扰性, 本文采用了强工频干扰背景下的仿真 EEG 数据 (如图 5(a) 所示, 最低信噪比达 -40dB, 原有 EEG 波形基本被工频噪声所淹没), 观察 ICA 算法的处理效果. 图 5(b) 为运用扩展最大熵算法分离出的独立分量, 显然工频噪声 (第 1 分量) 和眼动伪差 (第 11 和 14 分量) 都较成功地被分离出来了.



(a)



(b)

图 5 (a) 加入强工频干扰的 EEG 信号, 最低信噪比达 -40dB, EEG 波形基本被噪声所淹没;

(b) 对图 5(a) 中信号采用扩展最大熵算法分离出的独立分量

4.2 信号非平稳性对算法的影响

EEG 信号是一类典型的非平稳信号. 为考察信号非平稳性对 ICA 算法分离噪声效果的影响, 采用扩展最大熵算法分别对图 2 的 EEG 信号不同时段进行了 ICA 分析. 图 6(a) 是对 EEG 信号前半段的分离结果, 图 6(b) 为取信号后半段的分离结果. 从图 1 和图 2 的脑电波形来看, 信号的前后两个半时段有较明显的差异: 例如, 虽然 F_{p1} 和 F_{p2} 导联信号在前后两个半时段均受到眼部水平运动的慢波干扰, 但 F_{p1} 导联信号在前半时段还受到左眼上下眨动所引起的肌电干扰. 同时, 各个导联信号在前后两个半时段也都受到工频干扰. 以上情况皆在 ICA 分析结果中反映出来. 如图 6 所示, (a) 中的第 8 分量和 (b) 中的第 5 分量皆对应眼部水平慢动干扰, 而 (a) 中的第

3 分量则是眼睛眨动干扰. 因仅在前半时段 (a) 中有左眼眨动, 故在后半时段 (b) 中未见眨眼干扰分量. 同时, 在 (a)、(b) 两个半时段中均成功地分离出了工频噪声 (皆对应 (a) (b) 的第 2 分量). 经对前述 32 个时段中存在上述情况的 11 段数据采用以上对比方法考察, 均得到较明显的类似分离结果, 表明 EEG 信号的非平稳性对 ICA 算法的分析结果影响较小, ICA 有较强的鲁棒性.

4.3 消除干扰后对信号非线性特征的影响

由于 EEG 信号具有明显的非线性特征, 对其使用 ICA 算法消除干扰噪声之后是否会改变其非线性特征呢? 为此, 本文选用度量时间序列复杂性和统计量化的非线性动力学参数: 近似熵 (ApEn) [6] 作为 EEG 信号的非线性特征指标来考察 ICA 算法对 EEG 信号非线性特征的影响. ICA 分析采用扩展最大熵算法. 考察方法为: 分别取仅含眼动的原始 EEG 信号 (图 1)、含眼动和工频干扰的 EEG 信号 (图 2)、仅剔除眼动干扰后的 EEG 信号 (采用最大熵算法滤波后重建的信号) 和剔除眼动及工频干扰后的 EEG 信号 (采用扩展最大熵算法滤波后重建的信号) 计算其近似熵的值, 结果如表 1 所示 (为减少计算时间仅取信号序列的前 500 个数据点).

一般来说, 我们希望对观测所得 EEG 数据作 ICA 分析后不改变原始 EEG 信号的非线性特征. 因此, 使用 ICA 算法分离眼动与工频干扰后 EEG 信号的近似熵应与原始 EEG 信号 (虽个别导联受眼动干扰) 的近似熵相近. 从表 1 的数据来看, $C4$ 、 $P3$ 、 $P4$ 、 $O1$ 、 $O2$ 、 $F8$ 、 $T3$ 、 $T4$ 、 $T5$ 、 $T6$ 等 10 个导联基本符合这种情况, 在剔除干扰后的近似熵变化相对较小 (10% 以内); 而 $Fp1$ 、 $Fp2$ 、 $F4$ 、 $F3$ 、 $C3$ 和 $F7$ 等 6 个导联在剔除干扰后的近似熵

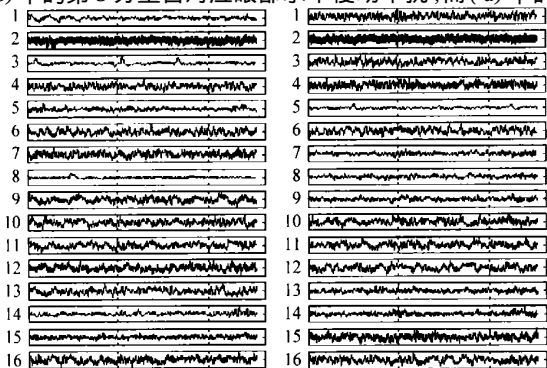


图 6 (a) 对图 2 中 EEG 信号的前半段采用扩展最大熵算法分离结果; (b) 对图 2 中 EEG 信号的后半段采用扩展最大熵算法分离结果

变化相对较大(均增加 10 %以上);其中以 Fp1、Fp2 和 F4 等 3 个导联的近似熵相对变化最大(超过 30 %).其原因可理解为:原始 EEG 波形中这 3 个导联明显受到眼动干扰,而规则的眼动可视为信噪比较高的确定信号(因而复杂度较低,近似熵较小),剔除眼动与工频干扰后只剩下纯正的 EEG 信号,自然复杂性相对会有所增加,则近似熵增大.

表 1 不同 EEG 信号的近似熵比较

导联	含眼动的 原始 EEG	含眼动和 工频干扰 的 EEG	仅剔除眼 动干扰后 的 EEG	剔除眼动 和工频干 扰后的 EEG
Fp1	0.71	1.23	1.18	0.94(+32.4%)
Fp2	0.58	0.99	0.93	0.80(+37.9%)
F3	0.91	1.34	1.36	1.02(+12.1%)
F4	0.53	0.70	0.69	0.69(+30.2%)
C3	0.75	0.68	0.69	0.84(+12.0%)
C4	0.66	0.64	0.66	0.66(0.00%)
P3	0.65	0.67	0.63	0.67(+3.0%)
P4	0.52	0.63	0.64	0.54(+3.8%)
O1	0.63	1.00	0.99	0.65(+3.2%)
O2	0.59	1.11	1.17	0.61(+3.4%)
F7	0.74	0.90	0.88	0.82(+10.8%)
F8	0.46	1.10	1.10	0.47(+2.2%)
T3	1.04	0.58	0.62	0.95(-8.7%)
T4	0.65	0.90	0.91	0.70(+7.7%)
T5	0.74	1.07	1.06	0.72(-2.7%)
T6	0.56	0.64	0.65	0.60(+7.1%)

注:近似熵的计算均取前 500 个数据点;第 4 列括号中的百分数为相对第 1 列变化值.

从表 1 还可以看到,除 C3、T3 以外所有导联在未剔除工频干扰时信号的近似熵多数都高于含眼动但未加工频干扰的原始波形,也高于剔除所有噪声干扰后的波形,或者基本相近(如 F4、C4、P3 和 T6).这是因为加入随机性较强的工频干扰后,由于所叠加信号间的非线性作用,一般使之近似熵增加;而对于 C3、T3 导联,由图 1 可明显看到这 2 个导联受到较强的工频干扰,原始波形的近似熵已经较高,再叠加进工频干扰时,因确定性较强的谐波信号占主导地位,其复杂性反倒降低,近似熵自然减小.上述情况与有关 EEG 近似熵性质研究结果类似^[9,10].总之,采用 ICA 方法消除 EEG 观测数据中的眼动和工频干扰,基本能有效地恢复和保持原始 EEG 信号的非线性特征.

5 结论

本文比较了最大熵和扩展最大熵 ICA 算法分离 EEG 临床数据中眼动伪差和工频噪声的效果,证实扩展最大熵算法能最有效地去除 EEG 中各种超高斯(眼动伪差)和亚高斯(工频干扰)噪声成分,并且有好的鲁棒性,在极低信噪比和 EEG 信号不平稳条件下都能有效地提取干扰并实现与原始 EEG 信号分离.这是一般基于二阶统计量分析方法所难以做到的.结合近似熵(ApEn)参数的分析还得知,利用 ICA 去除干扰后基本不影响且有助于恢复和保持原始 EEG 信号的非线性特征.以上结论说明 ICA 算法十分适合于 EEG 信号分析,有重

要的潜在应用价值,宜将之积极推广于临床工程之中.

参考文献:

- [1] Tzy-Ping Jung, et al. Extended ICA removes artifacts from electroencephalographic recordings[J]. Neural Info. Processing System, 1998, 10 (2): 894 - 900.
- [2] Ricardo Nuno Vigário. Extraction of ocular artifacts From EEG using independent component analysis[J]. Electroencephalography and Clinical Neurophysiology, 1997, 13(3): 395 - 404.
- [3] Karvanen J. et al. Pearson System Based Method for Blind Separation [A]. Proc. of Second Intl. Workshop on Independent Component Analysis and Blind Signal Separation[C], Helsinki 2000. 585 - 590.
- [4] Eriksson, J. et al. Source Distribution Adaptive Maximum Likelihood Estimation of ICA Model[A]. Proc. of Second Intl. Workshop on Independent Component Analysis and Blind Signal Separation[C], Helsinki 2000. 227 - 232.
- [5] 洪波等. ICA 在视觉诱发电位的少次提取与波形分析中的应用[J]. 中国生物医学工程学报, 2000, 9(3): 334 - 341.
- [6] Bell A. J. et al. An information maximization approach to blind separation and blind deconvolution[J]. Neural Computation, 1995, 7(6): 1129 - 1159.
- [7] Lee T W. et al. Independent component analysis using an extended infomax algorithm for mixed Subgaussian and Supergaussian sources[J]. Neural Computation, 1999, 11(2): 409 - 433.
- [8] Comon P. Independent Component Analysis, A new Concept[J]. 1994, 36: 287 - 314.
- [9] 洪波等. 近似熵、互近似熵的性质、快速算法及其在脑电与认知研究中的初步应用[J]. 信号处理, 1999, 15(2): 100 - 107.
- [10] 杨春梅等. 老年性痴呆患者的 EEG 近似熵特征初探[J]. 天津大学学报, 2002, 35(4): 521 - 524.

作者简介:



万柏坤 男, 1945 年生于江西南昌, 1968 年毕业于中国科技大学核电子学专业, 1982 年获中科院理学硕士学位, 先后在日本东京工业大学和东京医科齿科大学作访问学者, 现为天津大学生物医学工程系教授、博士生导师、系主任, 主要从事生物医学信息检测与处理, 任中国电子学会医学图像处理与分析专业委员会委员、中国生物医学工程学会生物信息与控制分会委员、教育部生物医学工程教学指导委员会委员、美国纽约科学院会员(ID 号 469859-7)。



朱欣 男, 1977 年 3 月生于湖南长沙, 1995 年免试进入天津大学攻读 5 年制生物医学工程专业, 2000 年 6 月获工学与文学双学士学位, 同年 9 月免试入天津大学研究生院, 2002 年 6 月获该专业工学硕士学位; 次年 4 月入日本会津大学大学院攻读博士学位, 主要研究方向为生物医学信号和图像处理, 心脏与神经电生理系统建模与仿真, 生物医学仪器与远程医疗, 已发表期刊论文 15 篇, 国内外会议论文 9 篇。