

文章编号: 100622467(2004)0520794205

低信噪比神经元锋电位信号的分类方法

张溥明^{1,2}, 吴金勇¹, 梁培基², 陈爱华³, 袁景淇¹

(1. 上海交通大学自动化系, 上海 200030; 2. 上海交通大学生物医学工程系, 上海 200030;
3. 中国科学院上海生命科学研究院, 上海 200031)

摘 要: 基于离散序列小波变换和主元分析, 对低信噪比的神经元锋电位信号提出了新的分类方法。通过对原始信号进行尖峰检测, 获得尖峰信号样本, 对每个样本进行离散序列小波变换之后, 再对小波变换系数进行主元分析, 选取主元进行聚类, 实现对原始信号的分类。将该方法应用于多电极细胞外记录的小鸡视网膜神经节细胞电活动信号分析, 并据此推断出某电极附近的神经节细胞的个数。仿真结果表明, 在低信噪比情况下, 该方法比单纯通过小波变换进行分类的方法更有效。

关键词: 神经元; 锋电位信号分类; 低信噪比; 小波变换; 主元分析; 视网膜

中图分类号: TP 273 **文献标识码:** A

Neural Spike Sorting under Low Signal-to-Noise Ratio

ZHANG Pu^{1,2}, WU Jin¹, LIANG Pei², CHEN Ai³, YUAN Jing¹

(1. Dept. of Automation, Shanghai Jiaotong Univ., Shanghai 200030, China;

2. Dept. of Biomedical Eng., Shanghai Jiaotong Univ., Shanghai 200030;

3. Shanghai Inst. for Biological Science, Chinese Academy of Sciences, Shanghai 200031)

Abstract: A new method was proposed to sort neural spikes under low signal-to-noise ratio by using discrete time wavelet transformation and principal component analysis. The events that represent spikes were extracted from the raw recording using a peak detecting algorithm, and a segment of fixed number of points was selected to straddle every event. Then the discrete time wavelet transformation was applied to analyze every spike event. After those coefficients of the discrete time wavelet transformation were analyzed by principal component analysis technique, a clustering analysis was applied to the principal components. Consequently the spikes were sorted according to the clustering result. The proposed method was also used to analyze the multi-channel extracellular recording of chicken's isolated retina, from which the number of ganglion cells around the electrode was deduced. The efficiency of the method was well shown by the results of simulation. When the signal-to-noise ratio is low, the proposed method outperforms the method based on wavelet transformation.

Key words: neurons; spike sorting; low signal-to-noise ratio; wavelet transformation; principal component analysis; retina

收稿日期: 2003204213

基金项目: 国家自然科学基金(60375039); 上海市科委重点基金(02JC14008)资助项目

作者简介: 张溥明(1974), 女, 湖南株洲人, 讲师, 2002年毕业于浙江大学信息学院, 获工学博士学位, 现主要从事生物神经信息研究
电话(Tel): 021264070495; E-mail: pmzhang@sjtu.edu.cn

神经元对行为信息的编码是神经科学研究的一个基本问题。目前大部分的电生理实验采用单细胞刺激的细胞内或细胞外记录技术,将每个神经元作为一个独立单位。然而,信息处理过程依赖于多个相关神经元的相互作用与整合,多电极细胞外记录能同时记录到多个神经元的活动,为研究神经元之间的群体活动特性提供了实验基础。但是记录到的信号是多个神经元信号的叠加,而且还包含噪声,这就引出了神经元锋电位信号分类的问题。

国外有不少学者致力于神经元锋电位信号的检测和分类研究,提出了主元分析法^[1,2]、特征集分析法^[3]、模糊聚类^[4]、匹配测试^[5,6]、人工神经网络识别法^[7,8]、阵列处理^[9]和基于小波变换的方法^[10-12]等多种分析方法。近年来兴起的小波变换方法,它的多分辨率分析的特性对于神经信号的分析非常有效。Leterlier等^[12]利用离散小波变换对锋电位进行分类,并且与主元分析方法和特征集分析方法进行了比较,说明小波变换的方法能更有效地分离波形和幅值非常类似的信号。但是,Leterlier在选择小波变换系数进行人工聚类分析时具有很大的人为随机性,实用性较差。进一步的仿真试验发现,当信噪比较低时,该方法失效,而来自实际过程的信号,特别是细胞外记录的信号,信噪比往往都比较低,因此希望提出一种新的分析方法来对低信噪比的神经元锋电位信号进行有效的分类。

本文基于离散序列小波变换和主元分析,对低信噪比的神经元锋电位信号提出了新的分类方法,并将该方法成功地应用于小鸡视网膜神经节细胞锋电位信号的分类。

1 数据预处理与尖峰检测

采用多电极记录技术在细胞外记录的神经元电活动信号信噪比通常都比较低,而且背景噪声的来源很多^[13],所以在进行信号分类之前必须进行数据预处理,以提高数据的信噪比,便于以后的信号检测和分类。数据预处理通常包括频谱分析、带通滤波^[14]等。

从信号中找出可以表达为锋电位的事件,最简单的方法为设定阈值抽取尖峰信号。虽然这种方法简单易行,但是当神经元电活动非常频繁,大量锋电位信号重叠时,通过这种方法可能抽取到因为信号之间重叠而严重变形的尖峰,使得尖峰样本有效性降低,为随后的分类带来了困难。在对小鸡视网膜神经节细胞锋电位信号进行分类的实际应用中,结合电生理的先验知识,对锋电位信号的峰峰值、幅宽等

都设定阈值范围,确保了检测到的尖峰信号样本的有效性。

2 神经元锋电位信号的离散序列小波变换分类方法

小波变换利用具有一定的时间和频率分辨率的基函数来分析信号,是变分辨率的时频联合分析方法,是一种非常理想的进行时变信号处理的数学工具^[15]。实际信号是经过采样获得的一系列离散信号,因此,在工程应用中离散序列小波变换应用最广。

对每个尖峰取同样长度的采样信号,即对每个尖峰取同样长度的采样点数 s ,构成一个离散序列,然后进行小波变换。对于某段信号,假设检测到 N 个尖峰,这样就有 N 个离散序列,如果最大分解次数取 M ,则经小波变换之后,将由低频概貌向量 r_i 和每一次分解得到的高频细节向量 d_i ($k=1, 2, \dots, M$)构成 N 个系数样本:

$$q^i = \{r_i^1, d_i^1, d_i^2, \dots, d_i^{M-1}\} \\ i = 1, 2, \dots, N$$

系数样本的长度 p 与小波基函数的种类有关。以Matlab小波工具箱中的小波基函数'db4'为例,如果对尖峰取 $s=64$ 个采样点,小波分解扩展模式选取'periodization',则每个小波变换系数样本的长度 $p=64$ 。

为实现信号分类,对小波变换系数进行聚类分析,这相当于将 N 个样本在64维空间进行聚类,这无疑是不可行的,必须进行降维处理。Leterlier的方法是把 N 个小波变换系数样本全部重叠绘制在一张图上,然后人为判断选择若干个均值和方差都较大的系数来达到降维的目的,再进行人工聚类分析^[12]。这种方法虽然简单易行,但随机性很大,不同的人可能选出不同的系数,所选系数的个数也可能不一致,这很可能导致最后聚类分析的结果产生差异,从而前功尽弃。而且,对于低信噪比的情况,这种方法也不适用,详见仿真和应用部分。因此本文在此基础上进一步对小波变换系数进行主元分析,通过选取前若干个主元来实现对小波变换系数的降维,再进行聚类分析实现对信号的分类。

3 引入主元分析的神经元锋电位信号分类新方法

主元分析方法是一种广泛应用并颇受推崇的降低变量维数的分析方法。它将变量进行组合变换,产生新的隐式变量,即主元,使得主元的个数较少,又

尽可能多地保留了原有变量的有用信息

传统的主元分析方法为去除数据量纲的影响而要求先对数据进行归一化和标称化处理,使得数据初始化为均值为 0 而且方差为 1,然后再进行奇异值分解.而这里的小波变换系数是无量纲的,并且对不同样本之间的差异必须非常关注.如果此时仍然先进行归一化和标称化预处理,将会损失有用的信息.因此,这里直接对小波变换系数进行奇异值分解,即

$$Q = USV^T \quad (1)$$

式中: $Q = [q^1 \ q^2 \ \dots \ q^N]^T$ 为由小波变换系数样本组成的 $N \times p$ 维矩阵; $U^T U = I$, U 为 $N \times p$ 维矩阵; $V^T V = I$, V 为 $p \times p$ 维矩阵; S 为 $p \times p$ 维对角阵, 对角线元素为 Q 的奇异值, 从大到小依次排列, 依次为每个主元对应的奇异值. 定义

$$Z = QV \quad (2)$$

式中, Z 的列向量为为主元(得分向量). 定义主元 z_i 的方差贡献率为 r_i , 则有

$$r_i = \frac{(S_{i,i})^2}{\sum_{k=1}^p (S_{k,k})^2} \quad (3)$$

由于没有对小波变换系数进行预处理, 故这里的方差贡献率在数学意义上与传统的主元分析方法中的定义是不一致的. 但为了衡量奇异值分解之后每个主元的影响, 这里仍称之为方差贡献率. 前 m 个主元的累计方差贡献率

$$R_m = \sum_{k=1}^m r_k \quad (4)$$

表示这 m 个主元从 p 个变量中提取的信息量

对 N 个样本的小波变换系数进行了主元分析之后, 根据预先确定 R_m 的阈值大小, 确定选用前多少个主元代替原来 p 个系数变量, 然后再对主元进行聚类分析. 这样既有效地对小波变换系数进行了降维, 又尽可能多地保留了原有数据的有用信息

4 聚类分析和仿真

在选取了前 m 个主元之后, 对这 m 维数据集进行聚类分析即可实现对锋电位信号的分类. 聚类分析的方法很多, 对于细胞外记录的神经元信号, 由于背景噪声复杂, 信噪比较低, 而且通过阈值检测很难排除所有异常的尖峰样本(如多个锋电位信号叠加产生的尖峰信号等), 故应采用鲁棒的聚类分析方法, 以保证分类结果的可靠性. 在对小鸡视网膜神经节细胞锋电位信号进行分类的实际应用中, 减法聚类法^[16]体现了良好的鲁棒性

首先模拟产生神经元信号. 假设信号由 3 个神经元发出的锋电位信号和噪声信号叠加而成, 并假设噪声服从正态分布, 均值为 0, 均方差为 0.033, 生成一个 64 000 点的噪声序列, 形成噪声背景. 假设 3 个神经元独立发出的锋电位信号模板如图 1(a) 所示. 信号模板宽度均取 64 个采样点宽. 然后在噪声背景上叠加 100 个信号 I 的模板, I 的峰值点叠加于噪声背景上随机选取的点. 以同样的方法在噪声背景上随机叠加 100 个信号 II 和 100 个信号 III, 最后形成的模拟信号如图 1(b) 所示

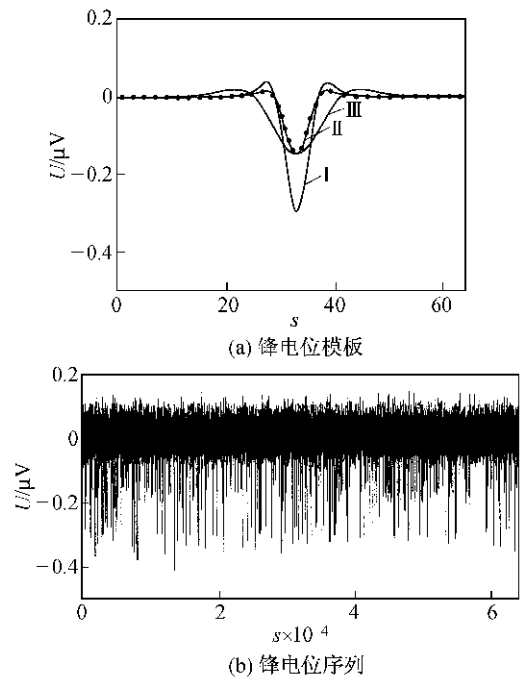


图 1 模拟锋电位信号模板和序列

Fig 1 Artificial spike templates and spike train

由正态分布统计特性, 噪声信号的最大峰值接近 0.1, 而真实锋电位信号的峰值分别为 0.30 和 0.15, 仅为噪声最大峰值的 3.0 倍和 1.5 倍, 可见该模拟信号的信噪比是比较低的. 结合模拟信号特征, 可得计算信噪比的公式为

$$SNR = 10 \lg \frac{\sum_{j=1}^3 y_j}{\sigma^2} \quad (5)$$

式中: $y_j = \frac{1}{s} \sum_{k=1}^s y_{j,k}^2$, $y_{j,k}$ 为信号 j ($j = I, II, III$) 在 k 点的值, s 为采样点数, 此处为 64; σ 为噪声的均方差. 这样, 计算得到模拟信号的信噪比为 4.07 dB.

由于产生模拟信号时只加入了随机噪声, 故不需要进行去除工频噪声和直流分量的预处理. 首先采用阈值检测法对原始信号进行尖峰检测, 阈值取

峰峰值 > 0.15 μV , 得到 350 个尖峰样本 由于信噪比较低, 噪声信号影响较大, 故检测到的尖峰信号明显多于真实锋电位信号

对每个尖峰信号取 64 个采样点构成离散序列, 然后利用 Matlab 小波工具箱来进行小波变换, 选用小波基函数为 'db4', 最大变换次数取 5 对 64 个点构成的离散序列, 经过变换之后得到由 64 个小波变换系数组成的样本, 由 350 个这样的样本组成了 350×64 维的矩阵 Q. 将 350 个样本的小波变换系数重叠绘制, 结果如图 2 所示

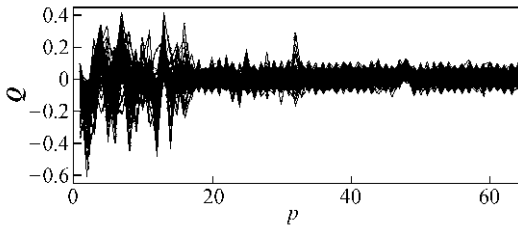


图 2 350 个样本的小波变换系数

Fig 2 350 samples' wavelet transformation coefficients

按照 Letelier 的方法^[12], 目测选取 $r_{-5}(2)$ 、 $d_{-4}(3)$ 、 $d_{-4}(4)$ 和 $d_{-3}(4)$ 4 个系数进行聚类分析, 效果很不理想, 部分结果如图 3 所示 类似地选取其他系数进行聚类分析, 效果也同样不理想 可见, Letelier 的方法对于信噪比较低的模拟信号无法进行分类, 而且系数选择过程无确定准则可以遵循, 随机性较大

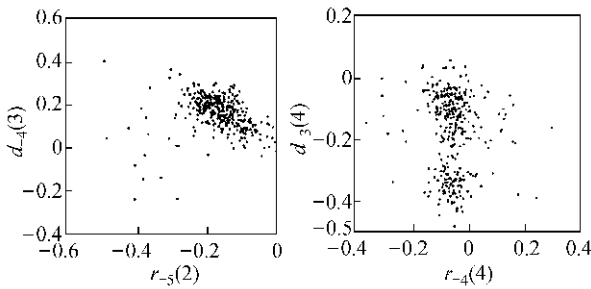


图 3 小波变换系数聚类结果

Fig 3 Clustering results by using wavelet transformation coefficients

按照先前提出的基于离散序列小波变换和主元分析的分类方法, 对矩阵 Q 进行奇异值分解, 得到前 2 个主元的方差贡献率为 59.86%, 大于预先确定 R_m 的阈值 50%, 选取这 2 个主元应用减法聚类法进行分析, 得到 3 个聚类中心, 以各聚类中心为中心取边长为 0.2 的正方形得到聚类结果如图 4 所示. 显然, 聚类结果比较理想, 从模拟信号中完全分离出 3 种信号 (· 为异常样本点).

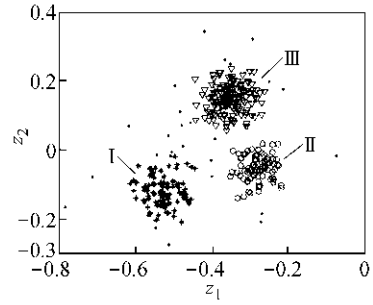


图 4 用文中提出的方法进行聚类的结果

Fig 4 Clustering result by using the proposed method

5 视网膜神经节细胞锋电位信号分类

视网膜主要进行视觉信息的初级加工和处理 为研究视网膜神经节细胞群体活动的特性, 应用多电极系统 (Multi Channel Systems, MCS) 对一小块保持功能活性的新生小鸡视网膜的神经节细胞进行细胞外记录 神经节细胞的电活动通过 60 道电极进行同步记录, 然后由商业软件进行采样 图 5 所示为在无光刺激下某一道电极记录的信号 (已经过预处理去除了工频噪声和直流分量), 为时 30 s, 采样频率为 20 kHz

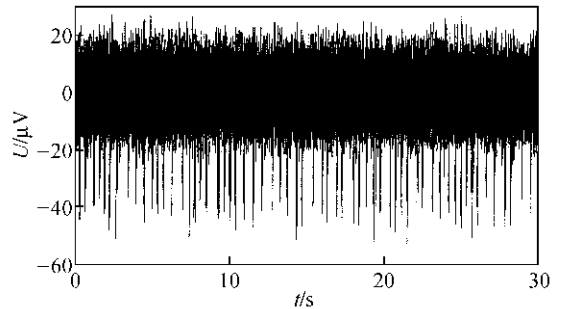


图 5 无光刺激下某道电极记录信号

Fig 5 One channel's recording without light stimulus

首先用阈值检测 (峰峰值 > 20 μV , 幅度 (0.5, 1.5) ms) 获取尖峰信号样本 421 个, 每个样本取 64 个采样点, 然后对这些样本进行离散序列小波变换, 再对小波变换系数进行主元分析, 得到前 2 个主元的方差贡献率为 64.28%, 选取这 2 个主元用减法聚类法进行聚类分析, 得到 2 个聚类中心, 以各聚类中心为中心取边长为 80 的正方形得到聚类结果如图 6 所示 显然, 聚类结果比较理想, 从信号中完全分离出 a 和 b 两种信号, 由于在进行尖峰检测时阈值条件比较苛刻, 所以选取的 421 个样本在聚类分析之后发现只有一个样本是异常的 虽然用苛刻的阈值条件选取尖峰样本时可能会遗漏一些样本, 但是该聚类分析结果还是能够说明该电极附近

至少有 2 个神经节细胞

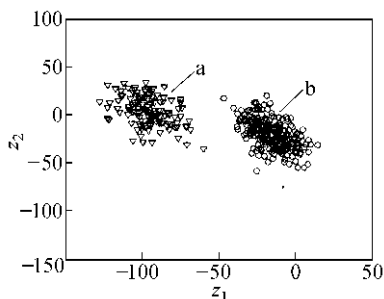


图 6 锋电位分类结果

Fig 6 Spike sorting result

参考文献:

- [1] Lewicki M S. A review of methods for spike sorting: the detection and classification of neural action potentials [J]. **Network: Computation Neural Systems**, 1998, 9(4): R53- R78
- [2] Zhang P M, Wu J Y, Zhou Y, *et al*. Spike sorting based on automatic template reconstruction with a partial solution to the overlapping problem [J]. **Journal of Neuroscience Methods**, 2004, 135 (1/2): 55- 65
- [3] Salganicoff M, Sarna M, Sax L, *et al*. Unsupervised waveform classification for multi-neural recordings: a real-time, software based system. I Algorithm and implementation [J]. **Journal of Neuroscience Methods**, 1988, 25(2): 181- 187.
- [4] Zouridakis G, Tam D C. Identification of reliable spike templates in multi-unit extracellular recordings using fuzzy clustering [J]. **Computer Methods and Programs in Biomedicine**, 2000, 61(1): 91- 98
- [5] Gadick R, Albus K. Real time separation of multi-neuron recordings with a DSP32C signal processor [J]. **Journal of Neuroscience Methods**, 1995, 57(2): 187- 193
- [6] Gadick R, Albus K. Performance of real time separation of multi-neuron recordings with a DSP32C microprocessor [J]. **Journal of Neuroscience Methods**, 1997, 75(2): 187- 192
- [7] Chandra R, Optican L M. Detection, classification, and superposition resolution of action potentials in multiunit single-channel recordings by an online real-time neural network [J]. **IEEE Transactions on Biomedical Engineering**, 1997, 44(5): 403- 412
- [8] Kim K H, Kim S J. Neural spike sorting under near 0dB signal-to-noise ratio using nonlinear energy operator and artificial neural network classifier [J]. **IEEE Transactions on Biomedical Engineering**, 2000, 47(10), 1406- 1411.
- [9] Bierer S M, Anderson D J. Multi-channel spike detection and sorting using an array processing technique [J]. **Neurocomputing**, 1999, 26/27: 947- 956
- [10] Zouridakis G, Tam D C. Multi-unit spike discrimination using wavelet transforms [J]. **Computers in Biology and Medicine**, 1997, 27(1): 9- 18
- [11] Krechel D, Kraft M, Sauer A. An algorithm for the analysis of single neuron activity by means of the wavelet packet transformation [A]. **Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences** [C]. Las Vegas: CSREA Press, 2000 601 - 607.
- [12] Letelier J C, Weber P P. Spike sorting based on discrete wavelet transform coefficients [J]. **Journal of Neuroscience Methods**, 2000, 101(2): 93- 106
- [13] Musial P G, Baker S N, Gerstein G L. Signal-to-noise ratio improvement in multiple electrode recording [J]. **Journal of Neuroscience Methods**, 2002, 115 (1): 29- 43
- [14] Chen A H, Zhou Y, Gong H Q, *et al*. Retinal ganglion cells response characteristics: multi-channel electrode recording study [J]. **Science in China (Ser C)**, 2003, 46(4): 414- 421.
- [15] 崔锦泰. 小波分析导论 [M]. 西安: 西安交通大学出版社, 1995
- [16] Rajesh N D, Raghu K. Robust clustering methods: a unified view [J]. **IEEE Transactions on Fuzzy Systems**, 1997, 5(2): 270- 293