

AVALIAÇÃO DE TÉCNICAS DE DETECÇÃO DE *SPIKES*

Renato A. Castellani *, Marcio Eisencraft* e Maria D. Miranda**

*Universidade Federal do ABC, Santo André, Brasil

** Escola Politécnica da Universidade de São Paulo, São Paulo, Brasil

renato.castellani@ufabc.edu.br

Resumo: A necessidade de observação de regiões específicas do cérebro para estimulação controlada, seja com fins medicinais ou de interação homem-máquina, vem despertando um crescente interesse em técnicas eficientes de detecção de *spikes*. Neste artigo, técnicas usuais de detecção baseadas em limiar são avaliadas com medidas de taxas de falso alarme e de falhas de detecção em função da relação sinal-ruído e do nível do limiar utilizado. O objetivo final é associar as vantagens e limites de bom desempenho das técnicas de detecção com as características das diferentes séries temporais consideradas.

Abstract: The observation of specific regions of brain for controlled stimulation, either for medicinal purposes or for human-computer interaction, has attracted a growing interest in efficient approaches for spike detection. In this paper, some usual techniques based on detection threshold are evaluated measuring false alarm and detection rates as a function of the signal to noise ratio and of the threshold level used. The ultimate objective is to associate the advantages and performance limits of the detection techniques with characteristics of the different time series considered.

Palavras-chave: sinais neuronais, *spikes*, detecção, *spike sorting*.

Introdução

Avanços tecnológicos recentes possibilitam a aquisição e o registro da atividade de um grande número de neurônios em animais acordados e com comportamento natural utilizando-se eletrodos intracranianos extracelulares [1]. Atuando-se de forma baseada nas atividades neurais podem-se realizar terapias para perdas de memória, cognição, paralisia e epilepsia, além de permitir o estudo da percepção, entendimento, movimento e memória e interação homem-computador e cérebro-máquina [2]. Apesar da aquisição, registro e processamento eficientes das atividades neurais serem pré-requisito para vários tipos de estudo de funções cerebrais, ainda, sob diversos aspectos, representam um desafio teórico e metodológico atual [2, 3].

Uma forma de estudar os sinais neuronais é por meio da análise de séries temporais adquiridas por redes de eletrodos alocados em determinadas posições do cérebro [3, 4]. Essa série é formada por trechos curtos de sinais (*spikes*), que refletem potenciais de ação dos neurônios próximos aos eletrodos, e ruído. Os *spikes* captados podem variar em duração, amplitude e forma. Entretanto, em estudos de decodificação neural, geralmente eles são interpretados apenas como disparos, não importando seu formato específico [3].

Os sensores podem captar a atividade de mais de um neurônio. Usualmente, em situações de estudo, observa-se na série temporal a atividade predominante de dois a dez neurônios [2]. Portanto, os *spikes* que chegam ao eletrodo são provenientes de neurônios distintos e muitas vezes, deseja-se analisá-los separadamente. Assim, nos últimos anos, muitos algoritmos foram propostos para detectar e classificar estes sinais, separando-os pelo neurônio que o emitiu. Nesse contexto, o problema de detecção e classificação é conhecido como *spike sorting* [1-5]. Apesar de muitas abordagens interessantes para realizar o *spike sorting* terem sido propostas, nenhuma solução foi aceita como universal e este assunto continua em aberto.

Uma das etapas fundamentais do *spike sorting* é a detecção dos *spikes*, ou seja, encontrar os instantes em que eles ocorrem. Uma das dificuldades é que usualmente existe uma quantidade significativa de ruído presente na série temporal, o que pode mascarar a presença de *spikes* ou causar falsas detecções [2]. Além disso, por não se ter controle sobre a geração dos sinais que se deseja detectar, o problema da detecção torna-se aqui muito mais complicado do que aquele que aparece em sistemas de comunicação usuais, em que se conhecem perfeitamente os sinais transmitidos [6]. Portanto, se a detecção não é feita de forma precisa, todo o processo de *spike sorting* pode ser comprometido.

Nas próximas seções, técnicas de detecção baseadas em limiar são avaliadas com medidas de taxas de falso alarme e de falhas de detecção em função da relação sinal-ruído e do limiar de detecção para um neurônio disparando a uma taxa constante. Desta forma,

pretende-se verificar qual a técnica mais adequada para se utilizar num posterior estudo de *spike sorting*.

Materiais e Métodos

Em uma primeira etapa, que é o caso deste estudo, a fim de possibilitar uma comparação justa entre os métodos de detecção é necessário conhecer a série temporal que se deseja avaliar. Por esse motivo, considera-se que a série temporal é formada por disparos periódicos de um único neurônio e assume-se que a forma de onda do *spike* não varia ao longo do tempo. Além disso, considera-se que a sequência já está amostrada a uma taxa conveniente de forma que o tratamento é todo feito em tempo discreto.

Seja $s(n)$ as amostras da forma de onda de um *spike* associado a um certo neurônio com $0 \leq n < N_s - 1$. O sinal medido por um eletrodo extracelular pode ser modelado como

$$x(n) = s(n) * h(n) + r(n), \quad (1)$$

sendo “*” a operação de convolução discreta,

$$h(n) = \sum_{k=1}^{N_k} \delta(n - p_k), \quad (2)$$

$\delta(n)$ a função impulso unitário, p_k o instante em que ocorre o k -ésimo disparo, N_k o número de *spikes* na sequência e $r(n)$ uma função-amostra de ruído branco, gaussiano, de média nula e desvio-padrão σ_r . O objetivo do processo de detecção é determinar os instantes de disparo p_k .

Nas simulações computacionais apresentadas aqui, utilizam-se como sinais $s(n)$ os disponibilizados por Rutishauser et al. [1]. Esses sinais correspondem a *spikes* detectados, alinhados e amostrados a uma taxa de 100 kHz. Exemplos de sinais $s(n)$ são mostrados na

Figura 1. Cada *spike* é representado por $N_s = 256$ amostras.

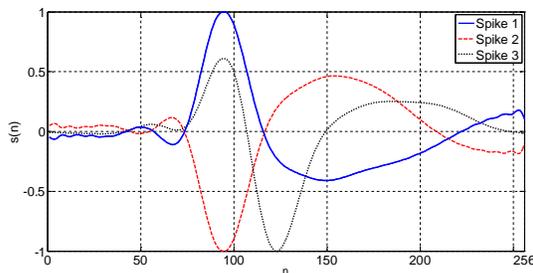


Figura 1: Spikes gravados por eletrodos extracelulares [1].

Considerou-se aqui que os disparos ocorrem de forma periódica, ou seja, $p_k = kT$, sendo T o número de amostras entre disparos. Para esse caso, a potência ou variância σ_r^2 do ruído adicionado é calculada como

$$\sigma_r^2 = \frac{P_s}{\text{SNR}} \quad (3)$$

sendo P_s a potência média do *spike* utilizado,

$$P_s = \frac{1}{N_s} \sum_{n=0}^{N_s-1} s^2(n), \quad (4)$$

e SNR a relação sinal-ruído que se deseja avaliar. É usual representar a SNR em decibéis,

$$\text{SNR}_{\text{dB}} = 10 \log(\text{SNR}). \quad (5)$$

Como considera-se que apenas um neurônio está disparando, uma vez detectado um *spike*, assume-se que não ocorrem outros *spikes* numa janela de $N_s = 256$ amostras, ou 2,56 ms, que é a duração do *spike*.

Na Figura 2(a) é mostrado um exemplo de trecho de sinal $x(n)$ com $N_x = 10000$ amostras. No exemplo, considerou-se como $s(n)$ o *Spike 1* da Figura 1 e uma $\text{SNR}_{\text{dB}} = 5$ dB. O neurônio dispara a cada $T = 1024$ amostras, o que equivale a um período de aproximadamente 10 ms. Na Figura 2(b) é mostrado o trem de impulsos $h(n)$ associado. Um processo de detecção deve idealmente ter como saída os instantes p_k em que estes impulsos ocorrem dado o sinal da Figura 2(a).

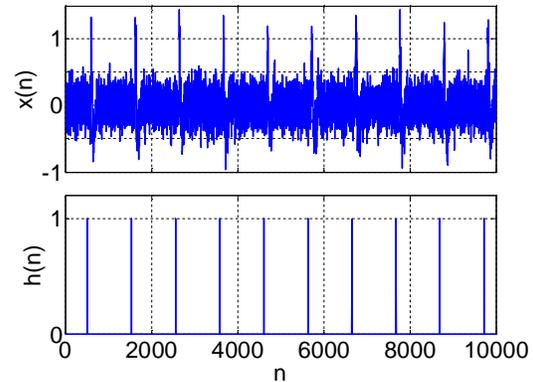


Figura 2: (a) Sinal gerado a partir do *Spike 1* com $\text{SNR}_{\text{dB}} = 5$ dB; (b) trem de impulsos $h(n)$.

Em geral, os métodos de detecção de *spikes* têm dois passos principais: a pré-ênfase do sinal e a aplicação de um limiar. A pré-ênfase tem por objetivo realçar os *spikes* frente ao ruído. Após essa fase, considera-se uma janela, ou seja, um pequeno trecho do sinal pré-

processado em um intervalo de tempo definido, e aplica-se o limiar. O sinal de *spike* é detectado se o sinal pré-processado apresentar amplitude maior que o limiar considerado [2, 4].

Observa-se na literatura que no pré-processamento algumas opções utilizadas são o módulo do sinal, a sua potência instantânea, o operador de energia não-linear (NEO, do inglês *Non-linear Energy Operator*), o desvio-padrão janelado ou até não pré-processar, isto é, aplicar o limiar ao próprio sinal medido [1, 2, 4, 7-10].

O NEO implica em transformar $x(n)$ em $y_{\text{NEO}}(n)$ dado por [8, 9]

$$y_{\text{NEO}}(n) = x^2(n) - x(n + \Delta)x(n - \Delta). \quad (6)$$

O atraso Δ depende da frequência de amostragem utilizada e deve ser escolhido de modo a assegurar a melhor capacidade de realçar os *spikes* em relação ao ruído. O valor de $y_{\text{NEO}}(n)$ é tanto maior, quanto maior é $x^2(n)$ em relação a $x(n - \Delta)$ e $x(n + \Delta)$. Em outras palavras, o valor de $y_{\text{NEO}}(n)$ é tanto maior quando maior é a potência instantânea e a frequência do sinal $x(n)$ [2]. Como um *spike* é caracterizado por altas frequências localizadas e um aumento na potência instantânea, este método tem uma vantagem sobre métodos que se preocupam apenas com um aumento na potência instantânea do sinal sem relacioná-lo com a frequência.

A pré-ênfase com desvio-padrão janelado [1, 7] consiste na seguinte operação sobre $x(n)$:

$$y_{\text{DP}}(n) = \left\{ \frac{1}{N} \sum_{i=1}^N (x(n-i) - x_m(n))^2 \right\}^{\frac{1}{2}}, \quad (7)$$

sendo

$$x_m(n) = \frac{1}{N} \sum_{i=1}^N x(n-i) \quad (8)$$

uma média-móvel de comprimento N .

Na Figura 3 estão mostrados exemplos de pré-processamento para uma mesma sequência. O mais simples é o caso (a) com o sinal bruto; em (b) é mostrado o seu valor absoluto $|x(n)|$; em (c) a sua potência instantânea $x^2(n)$; em (d) o operador de energia não linear $y_{\text{NEO}}(n)$ com $\Delta = 15$ e em (e) o desvio-padrão janelado $y_{\text{DP}}(n)$. Aqui se utiliza $N = 20$, como em [1].

Em geral, o limiar é definido como $L\sigma_r$, sendo $L > 1$.

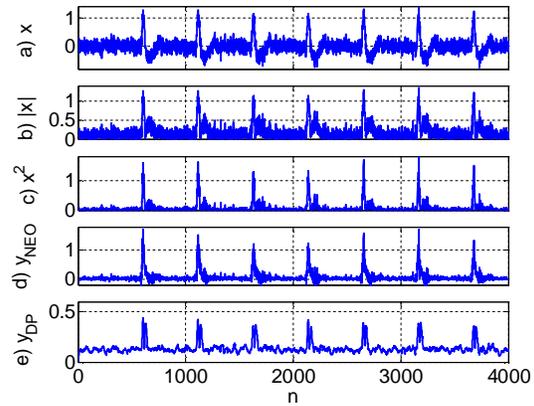


Figura 3: Pré-processamentos: (a) sinal bruto $x(n)$; (b) $|x(n)|$; (c) $x^2(n)$; (d) $y_{\text{NEO}}(n)$ com $\Delta = 15$ e (e) $y_{\text{DP}}(n)$ com $N = 20$.

O valor de σ_r é estimado a partir da sequência disponível e esta estimativa $\hat{\sigma}_r$ deve ser boa mesmo que muitos *spikes* estejam presentes na sequência observada. Essa estimativa pode ser feita de formas distintas. Em especial cabem destacar a estimativa que considera a mediana da sequência [5], a que considera o desvio padrão da própria sequência que contém os *spikes* e a que faz a média de $y_{\text{DP}}(n)$ [1].

Por exemplo, Quiroga et al. [5] detectam *spikes* utilizando $|x(n)|$, $\hat{\sigma}_r = \text{mediana} \left(\frac{|x|}{0,6745} \right)$ e $L = 4$.

Já Rutishauser et al. [1] aplicam o limiar a $y_{\text{DP}}(n)$, sendo o limiar calculado como L vezes o desvio-padrão dessa sequência com $L = 5$.

A fim de avaliar a eficiência dos métodos, considera-se a taxa de falsos alarmes (P_{FA}) e a taxa de falhas de detecção (P_{FD}). Definem-se estas taxas como [11]

$$P_{\text{FA}} = \left(\frac{\text{número de falsos alarmes}}{\text{número de verdadeiros negativos}} \right) \quad (9)$$

$$P_{\text{FD}} = \left(\frac{\text{número de falhas de detecção}}{\text{número de verdadeiros positivos}} \right).$$

Definimos verdadeiros positivos como as amostras que fazem partes dos *spikes* conhecidos e verdadeiros negativos como todas as outras amostras de $x(n)$. Falsos alarmes são todas as amostras em um *spike* detectado que não são parte de um *spike* verdadeiro e falhas de detecção são todas as amostras de um *spike* verdadeiro que não faz parte dos *spikes* detectados [11].

Como ilustração, na Figura 4 mostram-se exemplos de falsos alarmes e falhas de detecção em um sinal $x(n)$ para dois limiares de detecção diferentes. Na Figura 4(a), além da SNR_{dB} ser baixa, o limiar está colocado acima da posição ideal para uma boa detecção, o que causa muitas falhas de detecção. Na Figura 4(b), o limiar está abaixo de sua posição ideal, o que causa muitos falsos alarmes.

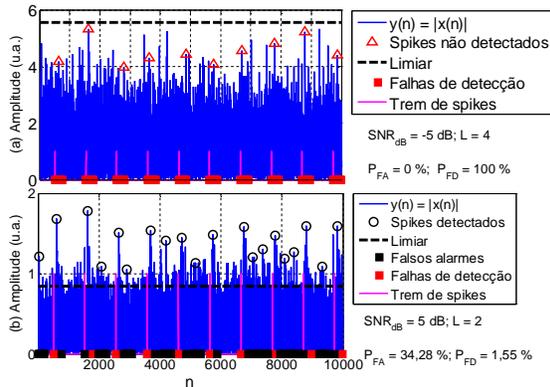


Figura 4: Exemplos de falsos alarmes e falhas de detecção.

Nas simulações, considerou-se como parâmetro $\max(P_{FA}, P_{FD})$ em função da SNR e do fator L que multiplica σ_r no cálculo do limiar. Note-se que $0 \leq \max(P_{FA}, P_{FD}) \leq 1$, sendo $\max(P_{FA}, P_{FD}) = 0$ se todos os *spikes* verdadeiros são detectados perfeitamente e não há falsos alarmes.

Avaliar $\max(P_{FA}, P_{FD})$ para diversos valores de L é relevante, pois como a estimativa do desvio-padrão do ruído pode não ser boa, é importante que os resultados sejam razoavelmente insensíveis ao limiar considerado.

Resultados e Discussão

Na Figura 5 são mostrados os valores de $\max(P_{FA}, P_{FD})$ em função da SNR_{dB} e do fator multiplicativo L do limiar $L\sigma_r$ para a detecção utilizando $|x(n)|$ e um limiar múltiplo da mediana da sequência [8]. Para cada par (SNR_{dB}, L) foram gerados 1000 sinais com a mesma estrutura da Figura 2(a). Quando $P_{FD} > P_{FA}$, utilizaram-se tons de azul e quando $P_{FA} > P_{FD}$ utilizaram-se tons de vermelho. Em particular, em [5] escolheu-se $L = 4$, que, de fato parece ser uma boa escolha para uma larga faixa de SNR_{dB} .

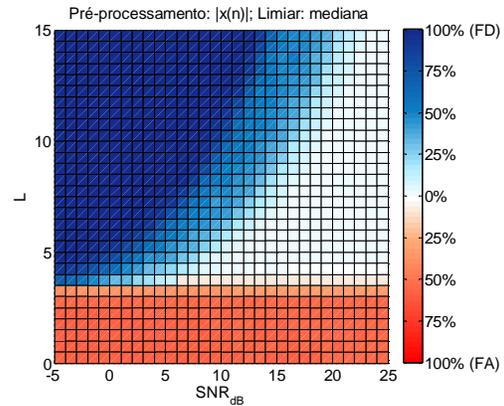


Figura 5: Valores de $\max(P_{FA}, P_{FD})$ para pré-processamento usando $|x(n)|$ e limiar com mediana.

Na Figura 6, repete-se o procedimento usado na Figura 5, empregando-se agora $y_{DP}(n)$ e o desvio-padrão [1]. A escolha feita na referência [1], $L = 5$, parece de fato uma boa escolha, pois é que apresenta $\max(P_{FA}, P_{FD}) \approx 0$ na maior faixa de SNR_{dB} .

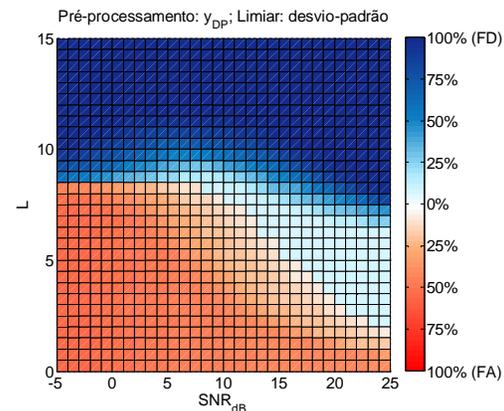


Figura 6: Valores de $\max(P_{FA}, P_{FD})$ para pré-processamento usando $y_{DP}(n)$ e limiar com desvio-padrão.

Finalmente, na Figura 7, mostram-se os resultados utilizando-se $y_{NEO}(n)$ com $\Delta = 15$ e a média. De todas as configurações estudadas, esta foi uma das que apresentou a maior faixa de $\max(P_{FA}, P_{FD}) \approx 0$, tendo um desempenho superior aos outros dois casos.

Os resultados apresentados mostram que a particular escolha da metodologia de detecção de *spikes* pode levar a maiores ou menores valores de P_{FA} e P_{FD} , o que pode certamente comprometer o desempenho de métodos de *spike sorting* que venham a utilizá-los.

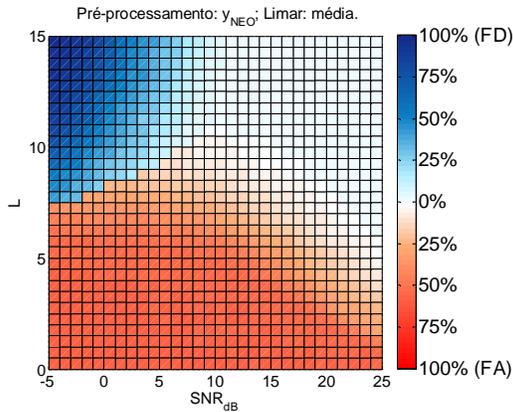


Figura 7: Valores de $\max(P_{FA}, P_{FD})$ para pré-processamento usando $y_{NEO}(n)$ e limiar com média.

Conclusões

Observam-se nitidamente três conjuntos de escolhas nos métodos de detecção de *spikes*: (a) estabelecer qual o pré-processamento de sinal a ser utilizado, (b) definir um limiar de detecção a partir de uma estimativa do desvio-padrão do ruído e (c) definir um intervalo mínimo em que se considera que só possa ter ocorrido um disparo. Estas escolhas levam a métodos de detecção com desempenhos distintos.

Neste estudo preliminar, analisaram-se as técnicas usualmente aplicadas na detecção de *spikes* como parte do processo de *spike sorting*. É importante que a detecção seja feita de forma adequada, mesmo em condições adversas de SNR. Note-se que é necessário obter σ_r a partir da sequência com os *spikes*, o que pode dificultar a escolha do fator multiplicativo do limiar.

Entre as possibilidades estudadas, a que usa o NEO e um limiar baseado em um múltiplo da média é a que fornece gama maior de valores de SNR e limiar sem o surgimento de falsos alarmes e falhas de detecção.

Um ponto importante a ser estudado é como as taxas de disparo afetam os limiares utilizados e, consequentemente, os valores de P_{FA} e P_{FD} . Utilizar outras técnicas para estimar σ_r , limiares adaptativos e modelos mais realistas para o ruído são trabalhos em andamento.

Agradecimentos

ME agradece ao CNPq e RAC agradece à CAPES pelo financiamento parcial desse trabalho.

Referências

- [1] U. Rutishauser, E. M. Schuman, and A. N. Mamelak, "Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo," *Journal of Neuroscience Methods*, vol. 154, pp. 204–224, 2006.
- [2] S. Gibson, J. W. Judy, and D. Markovi, "Spike sorting the first step in decoding the brain," *IEEE Signal Processing Magazine*, vol. 29, pp. 124–143, Jan 2012.
- [3] P. Dayan and L. F. Abbott, *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT Press, 2005.
- [4] I. Obeid, *A wireless multichannel neural recording platform for real-time brain machine interfaces*. Ph.D. Dissertation, Duke University, 2004.
- [5] R. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural Computation*, vol. 16, pp. 1661–1687, Aug 2004.
- [6] J. G. Proakis and M. Salehi, *Digital Communications*. McGraw Hill, 2008.
- [7] I. N. Bankman, K. O. Johnson, Schneider W. *Optimal detection, classification, and superposition resolution in neural waveform recordings*. *IEEE Trans Biomed Eng* 1993;40:836–41.
- [8] S. Mukhopadhyay and G. Ray, "A new interpretation of nonlinear energy operator and its efficacy in spike detection," *Biomedical Engineering, IEEE Transactions on*, vol. 45, pp. 180–187, feb. 1998.
- [9] J. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," in *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, pp. 381–384 vol.1, apr 1990.
- [10] Choi, J. H.; Jung, H. K. & Kim, T., A new action potential detector using the MTEO and its effects on spike sorting systems at low signal-to-noise ratios, *Biomedical Engineering, IEEE Transactions on*, 2006, 53, 738–746.
- [11] Gibson, S.; Judy, J. & Markovic, D., Technology-Aware Algorithm Design for Neural Spike Detection, Feature Extraction, and Dimensionality Reduction, *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 2010, 18, 469–478.