

# Análise de Preditores no Domínio do Tempo para Classificadores LVQ de Padrões de EEG Relacionados ao Movimento

Ivana Naomi Mukai<sup>1</sup>, Valfredo Pilla Jr<sup>1,2</sup>, Heitor Silvério Lopes<sup>1,2</sup>

<sup>1</sup>Laboratório de Bioinformática

<sup>2</sup>Departamento de Eletrônica

Centro Federal de Educação Tecnológica do Paraná

Av. 7 de setembro, 3165 - Curitiba, Brasil

valfredo@daeln.cefetpr.br, hslopes@cpgei.cefetpr.br

## Abstract

In this work is presented an analysis of time domain predictors applied to the detection of EEG patterns related to hand movements. 39 predictors were applied to EEG signals collected from 2 volunteers during the movement tasks. The Pareto's multiobjective optimization was applied for both cases, individually. The optimized parameters were the speed of processing and the classification performance. The data processed by the predictors selected were used as input patterns to training classifiers based on Kohonen-LVQ (*Learning Vector Quantization*) neural networks.

## 1. Introdução

Entre as diversas abordagens aplicáveis no projeto de dispositivos de assistência a portadores de deficiências físicas severas apresentam-se os Sistemas de Interface Cérebro Computador (BCI – *Brain-Computer Interface*). As BCIs baseiam-se na detecção de artefatos induzidos no sinal eletroencefalográfico (EEG) através da realização de movimentos, tarefas mentais ou por estímulos externos [1,2,3]. Nestes sistemas, artefatos eletroencefalográficos são associados a um “alfabeto” de comandos que podem ser aplicados na operação de um microcomputador ou de uma cadeira de rodas, entre outras.

Neste contexto, este trabalho analisa preditores no domínio do tempo aplicados na detecção da MRD (*Movement Related Desynchronization*) da banda mu produzida pela realização do movimento de um membro. São empregados sinais previamente coletados de dois voluntários. Estes sinais foram anteriormente utilizados na elaboração de um sistema classificador no qual utilizou-se a FFT (*Fast Fourier Transform*) como preditor [4]. A estrutura do classificador é constituída por uma rede neural Kohonen-LVQ (*Learning Vector Quantization*) [5], devido a sua velocidade de treinamento e pequena dimensão.

A seguir, são revisadas algumas das características do sinal de EEG e apresentada a arquitetura original do experimento. Em seguida, são relacionados os diversos preditores no domínio do tempo experimentados e o critério de Pareto para a otimização multiobjetivos,

assim como o processo de treinamento e avaliação das redes neurais LVQ. Finalmente, os resultados são apresentados e discutidos para dois voluntários, mostrando o desempenho do classificador para esta abordagem.

## 2. O sinal eletroencefalográfico

O sinal eletroencefalográfico humano pode ser definido como o reflexo da atividade elétrica do cérebro. Este sinal é complexo e estocástico por natureza. Ainda, outras variáveis aumentam sua complexidade, como o local da coleta do sinal sobre o escalpo, o nível de consciência e o estado mental do indivíduo, entre outras [6].

Um sinal de EEG típico frequentemente possui uma componente aleatória sobreposta por algum ritmo conhecido, especialmente as ondas alfa (8-13Hz), beta (>13Hz) e teta (4-8Hz) [7]. Tomadas sobre o escalpo, estas ondas possuem amplitude de pico da ordem de algumas unidades até algumas poucas centenas de microvolts. Particularmente, quando a derivação eletroencefalográfica é tomada na região do córtex sensoriomotor encontra-se presente também o ritmo mu (8-12Hz), cuja banda sobrepõe-se ao da onda alfa. Porém, este ritmo possui forte correlação com os processos de preparação e realização de atividades motoras, enquanto o ritmo alfa possui correlação com o nível de consciência e estado mental.

O ritmo mu é sujeito à um efeito denominado de Dessincronização Relacionada ao Movimento (MRD – *Movement Related Desynchronization*) [8]. Este efeito se caracteriza pela diminuição da amplitude do ritmo mu quando da realização do movimento contralateral, se o sinal foi tomado sobre a região do córtex relacionado aos processos de atuação e/ou controle do movimento.

A figura 1 apresenta a resposta temporal média do espectro da banda mu calculada através da FFT a partir dos sinais coletados de um dos voluntários, quando este realizava o movimento (em t=6s) de um membro [4]. A média da primeira sessão é constituída por 23 janelas e a média da segunda sessão por 29 janelas. A MRD é presente na forma de uma vale localizado em instante próximo ao do início do movimento do membro. O movimento é constituído por uma única elevação seguida do retorno da mão direita à posição de repouso.

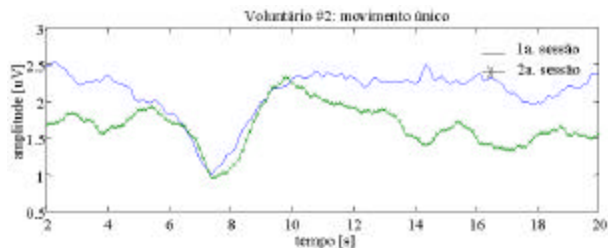


Figura 1: Resposta temporal média para a banda mu do voluntário 2 (médias de 2 sessões de coleta de sinais independentes).

### 3. Metodologia

#### 3.1 Aquisição de sinais

Os sinais de EEG utilizados neste trabalho foram coletados de dois voluntários, utilizando um amplificador eletroencefalográfico diretamente conectado a um computador de mesa [9]. O EEG foi tomado sobre a área do córtex sensoriomotor esquerdo. A taxa de amostragem utilizada foi de 1280 amostras / segundo, com uma resolução padrão de 16 bits e posteriormente dizimados para uma taxa de amostragem de 256 amostras / segundo, com o auxílio de um filtro digital FIR (*Finite Impulse Response*) de ordem 128 e frequência de corte 128Hz. Estes sinais foram então filtrados por um passa-faixa digital FIR de 7-13Hz, de ordem 64, para atenuar ruídos e sinais fora da faixa de frequência relacionada ao movimento.

Os voluntários participaram de duas sessões cada um, nas quais os sinais eram coletados por uma derivação diferencial tomada sobre o córtex sensoriomotor do hemisfério esquerdo (figura 2). Os voluntários, segundo uma sinalização, realizavam um único movimento da mão (em  $t=6s$ ) que consistia no seu giro em torno do punho para cima e no retorno à posição de repouso.

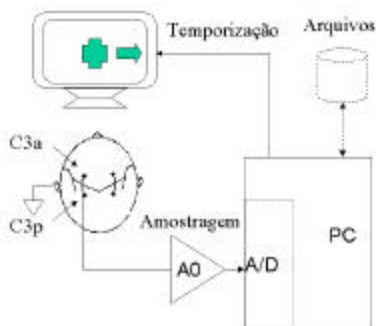


Figura 2: Aquisição dos sinais eletroencefalográficos. C3a: eletrodo anterior a C3 (Sistema Internacional de Eletroencefalografia); C3b: eletrodo posterior a C3.

#### 3.2 Descrição dos preditores

Os 39 preditores experimentados neste trabalho foram tomados de relatos da literatura [10,11], em que os mesmos foram empregados na detecção de artefatos

do tipo pico e onda, que são padrões eletroencefalográficos associados à epilepsia. Estes preditores operam os sinais de EEG no domínio do tempo. No presente trabalho, os padrões a serem detectados não se caracterizam pelo padrão pico e onda e sim pela presença da MRD da banda mu.

A tabela 1 relaciona os preditores apresentados em [10] e a tabela 2a e 2b os apresentados em [11].

#### 3.2 Treinamento da rede LVQ

A rede neural utilizada como classificador é do tipo Kohonen-LVQ (*Learning Vector Quantization*)[5], treinada com o algoritmo da rede LVQ2. O treinamento desta rede trata-se do ajuste dos vetores de referência. A inicialização dos vetores de referência foi realizada através do algoritmo *fuzzy c-means*[12].

O objetivo final do treinamento da rede neural é o reconhecimento de padrões de EEG definidos pelas classes com movimento (isto é, com a presença da MRD) e sem movimento da mão (isto é, atividade de fundo)[4]. Os padrões de entrada da rede neural foram constituídos por conjuntos de vetores contendo 9 amostras consecutivas calculadas através do(s) preditor(es) em análise. Estas amostras correspondem a uma janela de tempo de 1,125s. Assim, os padrões de entrada foram tomados dos sinais originais no intervalo em que ocorre a presença da MRD ( $t=6,5s$  a  $7,5s$ ) para a classe com movimento e a partir de  $t=10s$  para a classe sem movimento. Estes sinais foram então processados pelos preditores e utilizados no treinamento da rede. Os dados foram divididos em 2/3 para treinamento e 1/3 para teste da rede. Para cada voluntário, foram utilizadas o total de 144 janelas do sinal da classe com movimento e 48 janelas da classe sem movimento para treinamento e teste do classificador.

A avaliação da rede LVQ foi feita através do cálculo das ocorrências positivas e negativas verdadeiras, ou seja, dos acertos do classificador; e das ocorrências positivas e negativas falsas, ou seja, dos erros do classificador calculados para o conjunto de teste. Através dessas ocorrências verdadeiras e falsas é possível definir a sensibilidade e especificidade [13], dadas respectivamente, pelas equações a seguir:

$$se(i) = \frac{vp(i)}{vp(i) + fn(i)} \quad (1)$$

$$es(i) = \frac{vn(i)}{vn(i) + fp(i)} \quad (2)$$

onde:

$i$ :  $i$ -ésima época de treinamento;

$vp$ : (verdadeiros positivos) número de classificações corretas da classe com movimento;

$fn$ : (falsos negativos) número de classificações incorretas classe com movimento;

$vn$ : (verdadeiros negativos) número de classificações corretas classe sem movimento;

*fp*: (falsos positivos) número de classificações incorretas classe sem movimento.

O índice de desempenho da rede neural é dado por valores entre 0 e 1, ou seja, 0% e 100% de acerto, e é obtido pela equação 3:

$$de(i) = \sqrt{se(i).es(i)} \quad (3)$$

### 3.3 Critério de análise

A análise dos preditores pretende identificar os que possuem o melhor desempenho individual, de forma que sua escolha represente a solução ótima no processo de separação da MRD da atividade de fundo do sinal eletroencefalográfico.

Tabela 1: Lista de preditores [10].

Descrição	Sigla
Valor médio absoluto	media_abs
Número de cruzamentos por zero	crz_zero
Número de inversões de inclinação	inv_inclinacao
Comprimento da forma de onda	comp_onda
Valor médio	valor_medio

Tabela 2a: Lista de preditores [11].

Descrição	Sigla
Desvio padrão das amplitudes	amsd
Inclinação das amplitudes	amsk
Média das amplitudes vértice-a-vértice <sup>a</sup>	amvvn
Desvio padrão das amplitudes vértice-a-vértice	amvvsd
Coefficiente de variação das amplitudes vértice-a-vértice	amvvcv
Média das inclinações absolutas das amplitudes	slmn
Desvio padrão das inclinações absolutas das amplitudes	sbsd
Coefficiente de variação das inclinações absolutas das amplitudes	slev
Média das inclinações absolutas vértice-a-vértice <sup>b</sup>	slvvn
Desvio padrão das inclinações absolutas vértice-a-vértice	slvvsd
Coefficiente de variação das inclinações absolutas vértice-a-vértice	slvvcv
Média das curvaturas ( $d^2x/dt^2$ ) nos vértices	cuvvn
Desvio padrão das curvaturas nos vértices	cuvvsd
Coefficiente de variação das curvaturas nos vértices	cuvvcv
Média dos períodos vértice-a-vértice	tivvn

<sup>a</sup> Vértice - qualquer pico ou vale. Vértice-a-vértice - o intervalo (de amplitude ou tempo) entre dois vértices consecutivos.

<sup>b</sup> Inclinação vértice-a-vértice é a diferença de amplitude entre dois

vértices consecutivos dividida pela diferença de tempo entre os mesmos dois vértices.

Há, no entanto, um segundo parâmetro a ser otimizado: a velocidade de processamento. A implementação em tempo real de um sistema BCI requer preditores (assim como um algoritmo classificador) que demandem o menor esforço computacional possível.

Assim, dois critérios devem ser otimizados simultaneamente:

- mínimo tempo de processamento
- máximo desempenho na classificação

Tabela 2b: Lista de preditores (cont.) [11].

Descrição	Sigla
Desvio padrão dos períodos vértice-a-vértice	tivvsd
Coefficiente de variação dos períodos vértice-a-vértice	tivvcv
Número de cruzamentos médios por segundo (histerese = 25%AMSD) <sup>c</sup>	ctmxmn
Média do período de cruzamento médio	timxmn
Desvio padrão do período de cruzamento médio	timxsd
Coefficiente de variação do período de cruzamento médio	timxsd
Amplitude de 50-100% da banda de frequência / desvio padrão <sup>d</sup>	fb5099
Amplitude de 25-50% da banda de frequência / desvio padrão <sup>d</sup>	fb2550
Amplitude de 12-25% da banda de frequência / desvio padrão <sup>d</sup>	fb1225
Amplitude de 6-12% da banda de frequência / desvio padrão <sup>d</sup>	fb0612
Amplitude de 3-6% da banda de frequência / desvio padrão <sup>d</sup>	fb0306
Amplitude de 0-3% da banda de frequência / desvio padrão <sup>d</sup>	fb0003
Máxima amplitude pico-a-pico / desvio padrão	racr
Média das amplitudes positivas / média das amplitudes negativas <sup>e</sup>	rapn
Média do tempo de subida / média do tempo de descida <sup>f</sup>	rtrf
Média do período de cruzamento das amplitudes médias positivas / média do período de cruzamento das amplitudes médias negativas <sup>e</sup>	rtpn
Curvatura média dos vértices positivos / curvatura média dos vértices negativos <sup>e</sup>	rcpn

<sup>c</sup> Para que um cruzamento médio seja detectado, o sinal deve cruzar o valor médio do sinal e ser maior que o fator de histerese acima ou abaixo do último cruzamento médio. Neste caso, o fator de histerese é 25% do amsd.

<sup>d</sup> Calcula a relação da soma das amplitudes de 50-100% da banda de frequência pelo amsd.

<sup>e</sup> Amplitudes positiva e negativa são aquelas acima e abaixo da amplitude média, respectivamente.

<sup>f</sup> Tempos de subida e descida são os tempos em que a amplitude está crescendo ou decrescendo, respectivamente.

Um problema de otimização multiobjetivos pode ser solucionado de maneira similar a um problema de otimização para um único critério [14]. Pode-se determinar os valores das variáveis de projeto que otimizem os parâmetros de desempenho. A solução é então obtida através do conjunto ótimo de Pareto definido pelos pontos dominantes do sistema.

A condição para determinar os pontos dominantes pela otimização de Pareto é dada pela equação 4:

$$(x < p y) \Leftrightarrow (\forall i)(x_i \leq y_i) \wedge (\exists i)(x_i < y_i) \quad (4)$$

onde  $x < p y$  : vetor  $x$  é parcialmente menor que  $y$ ,  $x$  e  $y$  são vetores compostos pelas variáveis de projeto.

Obedecendo tal condição, o ponto  $x$  domina o ponto  $y$ . Se um ponto não é dominado por outro qualquer ele é dito um ponto dominante, não-dominado ou não-inferior. O total de pontos dominantes define então o conjunto ótimo da solução.

## 4. Resultados

Os sinais processados pelos preditores correspondem aos arquivos de dados referentes ao EEG com um movimento adquiridos de dois voluntários.

Devido ao grande número de preditores, foram feitas algumas análises dos mesmos com a finalidade de determinar os melhores parâmetros para a entrada do sistema classificador.

### 4.1 Análise pelo critério da otimização de Pareto

Os objetivos adotados para tal critério foram os seguintes:

- O cálculo dos algoritmos dos preditores deve gastar o menor tempo possível de processamento;
- Os preditores devem obter o melhor desempenho possível durante o treinamento da rede neural, quando treinados individualmente.

Os preditores que obedecerem os objetivos propostos serão os preditores que melhor servirão como entrada do classificador para reconhecer os padrões relacionados ao movimento da mão.

Para o primeiro objetivo adotado, foi determinado o tempo de processamento (normalizado entre 0 e 1) necessário para o processamento de cada algoritmo utilizando apenas um dos arquivos de dados como referência, através de um computador Pentium 233MHz.

Para o segundo objetivo, os preditores foram treinados individualmente através de uma rede neural do tipo LVQ e avaliados de acordo com o seu desempenho dado pela equação 3. Os parâmetros da rede como:

constante de aprendizagem  $\alpha$  de 0,01; número de neurônios intermediários da camada competitiva igual à combinação de 1 a 1 até 6 a 6 neurônios para cada classe; o número de épocas de 250 vezes o total de vetores de referência, incluídas ambas as classes, foram todos baseados nos experimentos realizados por Pilla Júnior [9].

Este critério foi adotado para o treinamento de uma rede neural desbalanceada (com 75% dos dados de entrada da classe de atividade de fundo e 25% da classe com movimento) específica para o voluntário 1 e outra para o voluntário 2 para o movimento único da mão.

Assim, foi traçado o gráfico de tempo de processamento normalizado *versus* 1-desempenho da rede neural de cada preditor para os dados de todos os voluntários e para os dados do voluntário 1, representados pelas figuras 3 e 4, cujos valores são apresentados nas tabelas 3 e 4. O conjunto dos pontos ótimos corresponde aos pontos dominantes determinados pelo critério da otimização de Pareto (equação 4). Para o critério de Pareto, os melhores pontos correspondem àqueles que estão mais abaixo e mais à esquerda dos gráficos das figuras 3 e 4. Para a rede treinada com os dados somente do voluntário 1, os pontos dominantes, circulados na figura 3, foram: *amvvsd*, *amsd*, *media\_abs* e *valor\_medio*. Para a rede treinada com os dados somente do voluntário 2, os pontos dominantes, circulados na figura 4, foram: *timxsd*, *sbsd*, *comp\_onda*, *incl\_ascend*, *incl\_descend*, *media\_abs* e *valor\_medio*.

Tabela 3: Tempo de processamento e (1-desempenho) da rede LVQ dos preditores treinados com os dados do voluntário 1.

Preditor	Tempo de processamento normalizado	1-desempenho (rede LVQ)
amvvsd	0,252052463	0,262253
amsd	0,105577017	0,271131
Media_abs	0,054800571	0,28193
Valor_medio	0,051318515	0,350481

Tabela 4: Tempo de processamento e (1-desempenho) da rede LVQ dos preditores treinados com os dados do voluntário 2.

Preditor	Tempo de processamento normalizado	1-desempenho (rede LVQ)
timxsd	0,306961902	0,219375
sbsd	0,117564908	0,275611
Comp_onda	0,080210908	0,3125
Incl_ascend	0,076772512	0,378133
Incl_descend	0,078242222	0,378133
Media_abs	0,054800571	0,413698
Valor_medio	0,051318515	0,413698

## 5 Discussão e conclusões

Determinados os preditores ótimos obtidos pelo critério de Pareto, duas redes neurais foram treinadas (uma para cada voluntário) utilizando uma concatenação do conjunto de preditores ótimos, específicos para cada voluntário. Antes de serem concatenados, os preditores (agrupados por preditor e por voluntário) foram normalizados para corresponderem a uma distribuição normal com média zero e variância igual a 1.

O treinamento foi individualizado para cada voluntário. Nesta configuração, o melhor desempenho obtido para o voluntário 1 foi de 56,3%. Para o voluntário 2 o desempenho foi de 53,4%.

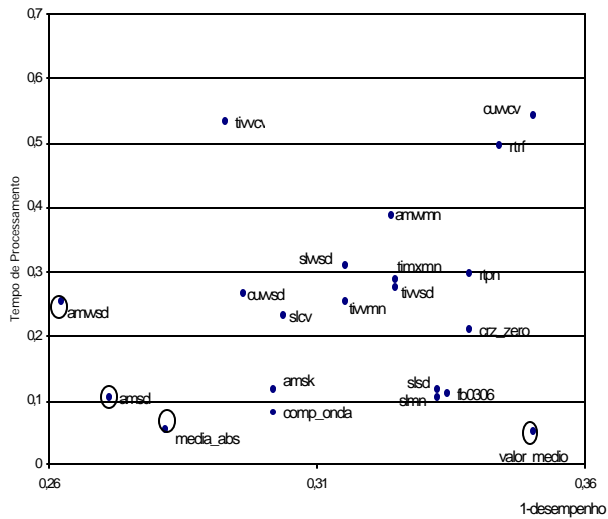


Figura 3: Plano de Pareto para a determinação do conjunto de preditores ótimos para o voluntário 1. Os círculos marcam os pontos dominantes.

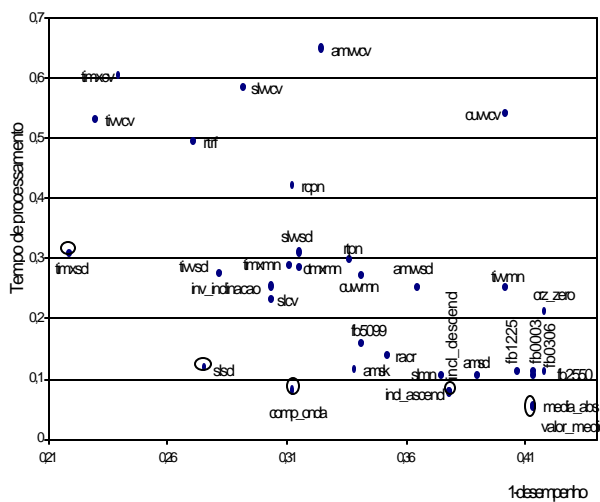


Figura 4: Plano de Pareto para a determinação dos preditores ótimos para o voluntário 2. Os círculos marcam os pontos dominantes.

Para o voluntário 1, o conjunto de preditores ótimos apresentou melhor resultado quando treinados individualmente em comparação à rede com os preditores treinados juntos, considerando o mesmo tempo de treinamento. Os valores do desempenho dos preditores *amvvsd*, *amsd*, *media\_abs* e *valor\_medio* quando treinados individualmente para uma rede desbalanceada foram 73,8%, 72,8%, 71,8% e 64,9%, respectivamente, o que indica um percentual superior à da rede desbalanceada com os preditores ótimos treinados juntos, cujo valor foi de 56,3%.

O mesmo ocorreu para o voluntário 2, em que os preditores *timxsd*, *sbsd*, *comp\_onda*, *incl\_ascend*, *incl\_descend*, *media\_abs* e *valor\_medio* apresentaram como desempenho individual 78,1%, 72,4%, 68,8%, 6,2%, 62,2%, 58,6% e 58,6%, respectivamente, que são valores maiores comparados ao desempenho da rede com todos os preditores ótimos como entrada que foi de 53,3%.

Em trabalhos anteriores [4], o desempenho de classificação obtido foi de 64% e de 88% para os voluntários 1 e 2, respectivamente. Nestes casos, foi utilizado como preditor a resposta temporal da banda mu calculada através da FFT. Assim, os preditores individuais aqui apresentados obtiveram melhor desempenho para o voluntário 1 e desempenho inferior para o voluntário 2 que os previamente obtidos através de um preditor no domínio da frequência. Estes resultados serão posteriormente estendidos quando a metodologia for reaplicada em um grande número de casos, para sinais coletados de vários outros voluntários.

De acordo com o critério de análise adotado, muitos dos preditores calculados foram descartados por não apresentarem resultados satisfatórios. Uma explicação é que as referências bibliográficas em que foram baseadas a escolha de tais parâmetros [10,11] utilizam os mesmos para detecção de padrões epileptiformes e não para a detecção de padrões relacionados ao movimento.

Através dos resultados obtidos até o momento para os dados referentes ao movimento único da mão, pode-se concluir que os melhores valores de desempenho da rede neural do tipo LVQ são os obtidos pelo treinamento da rede desbalanceada por voluntário utilizando os preditores individualmente.

Para trabalhos futuros, existe a possibilidade de treinar a rede neural individual para mais voluntários para se fazer uma análise e comparação do desempenho. Também poderão ser utilizados os mesmos procedimentos adotados para treinamento da rede neural neste trabalho mas para os dados com movimentos repetidos da mão. Além disto, podem ser adotados outros critérios de análise como a análise da variância utilizada em Pilla Júnior [9] e também o cálculo de outros preditores que sejam capazes de identificar as classes com e sem movimento da mão.

Um estudo mais profundo de técnicas de combinação de preditores pode resultar ainda em melhoria de desempenho, associando-se inclusive

preditores do domínio da frequência com preditores temporais. A programação genética [4] ou outras técnicas de algoritmos evolucionários poderiam ser experimentadas no desempenho desta tarefa.

[14]D. E. Goldberg, Multiobjective Optimization. In: \_\_\_\_\_ **Genetic Algorithms, in Search Optimization and Machine Learning**, Addison-Wesley, p. 197-201, 1989.

## Referências

- [1]T. M. Vaughan J. R. Wolpaw, E. Donchin, EEG-based communication: prospects and problems. **IEEE Transactions on Rehabilitation Engineering**, v.4, n. 4, p. 425-430, 1996.
- [2]J. R. Wolpaw, D. J. McFarland, Multichannel EEG-based brain-computer communication. **Electroencephalography and Clinical Neurophysiology**, v. 90, n. 6, p. 444-449, 1994.
- [3]L. A. Farwell, E. Donchin, Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. **Electroencephalography and Clinical Neurophysiology**, v. 70, n. 6, p. 510-523, 1988.
- [4]V. Pilla Júnior, H. S. Lopes, Detection of movement-related desynchronization of the EEG using neural networks. **Proceedings of the World Congress on Medical Physics and Biomedical Engineering**, Chicago, USA, [CD-ROM], 2000.
- [5]T. Kohonen, **Self-Organizing Maps**. 2<sup>nd</sup> edition. Heidelberg : Springer-Verlag, 1997.
- [6]C. Janeczko, H. S. Lopes, A genetic approach to ARMA filter synthesis for EEG signal simulation. **Proceedings of Congress on Evolutionary Computation**, La Jolla, USA, v. 1, p. 805-810, 2000.
- [7]G. E. Chatrian, L. Bergamini, M. Dondey, D. W. Klass, M. Lennox-Buchthall, I. Petersén, A glossary of terms most commonly used by clinical electroencephalographers. **Electroencephalography and Clinical Neurophysiology**, v. 37, p. 538-548, 1974.
- [8]A. Stancak Junior, G. Pfurtscheller, The effects of handedness and type of movement on the contralateral preponderance of mu-rhythm desynchronization. **Electroencephalography and Clinical Neurophysiology**, v. 99, n. 2, p. 174-182, 1996.
- [9]V. Pilla Júnior, **Detecção de padrões eletroencefalográficos relacionados ao movimento para uma interface cérebro-computador**. Curitiba, 1999, 98p.; Dissertação (Curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Engenharia Biomédica) - CPGEI, CEFET-PR.
- [10]B. Marchesi, **Detecção de padrões epileptiformes em sinais de eletroencefalografia através do paradigma da programação genética**. Curitiba, 1998. 133p.; Dissertação (Curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Engenharia Biomédica) - CPGEI, CEFET-PR.
- [11]W. R. S. Webber, R. P. Lesser, R. Richardson, K. Wilson, An approach to seizure detection using an artificial neural network. **Electroencephalography and Clinical Neurophysiology**, v. 98, p. 250-272, 1996.
- [12]M. Friedman, A. Kandel, **Introduction to Pattern Recognition: statistical, structural, neural and fuzzy logic approaches**. 2<sup>nd</sup> edition. Singapore: World Scientific, 1999.
- [13]D. J. Hand, **Construction and Assessment of Classification Rules**. Chichesters: John Willey & Sons, 1997.