

RECONHECIMENTO DE PADRÕES NO EEG COM PROGRAMAÇÃO GENÉTICA

Alceo Carta Neto [Bolsista PIBIC-CNPq], Heitor Silvério Lopes [orientador], Valfredo Pilla Júnior [orientador]

Laboratório de Bioinformática,
Centro de Pós-Graduação em Engenharia e Informática - CPGEI
Centro Federal de Educação Tecnológica do Paraná - CEFET-PR
Av. Sete de Setembro, 3165 - Curitiba/PR, Brasil - CEP 80230-901

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

Resumo - Este artigo apresenta uma aplicação de programação genética (GP) em reconhecimento de padrões. O objetivo deste trabalho é gerar uma regra capaz de classificar corretamente sinais relacionados a movimentos da mão direita. Os dados a serem classificados são 39 preditores no domínio do tempo extraídos de sinais de EEG, aplicados em paralelo à programação genética. Os resultados mostraram uma precisão preditiva similar a outros métodos previamente utilizados, porém os resultados não utilizaram os dados de todos os preditores, o que representa uma simplificação significativa para a construção de um sistema BCI.

Palavras-chave: Interface Cérebro-Computador, Reconhecimento de Padrões, Programação Genética.

Abstract - This paper shows an application of genetic programming (GP) in pattern recognition. The aim of this work is to generate a rule capable of correctly classifying EEG signals related to the right hand movements. The data used are 39 time-domain predictors extracted from EEG signals from three volunteers. Results show a predictive accuracy similar to other methods previously used. Nevertheless, final classifiers do not use data from all predictors, what represents a significant simplification and improvement for the construction of a BCI system.

Key-words: Brain-Computer Interface, Pattern Recognition, Genetic Programming.

INTRODUÇÃO

Os sistemas BCI (*Brain-Computer Interface*) [1] [2] têm como objetivo o auxílio de pessoas portadoras de deficiências através do reconhecimento de padrões no sinal de EEG. A Dessincronização Relacionada ao Movimento (MRD – *Movement Related Desynchronization*), é um padrão relacionado ao movimento de membros, observado na banda mu (8 – 12 Hz) do sinal de EEG.

Em 1998 foi iniciado o projeto BCI do Laboratório de Bioinformática no CPGEI/CEFET-PR. O projeto já conta com um banco de dados de sinais de EEG [3], um *software* desenvolvido com um *kit* DSP (*Digital Signal Processor*) TMS320C542 para cálculo da Transformada Rápida de Fourier em tempo real [4] e um estudo de preditores no domínio do tempo aplicados ao reconhecimento de padrões relacionados ao movimento, utilizando uma rede neural LVQ (*Learning Vector Quantization*) [5] [6].

Neste trabalho foi aplicado o paradigma da programação genética, com a finalidade de se obter um classificador capaz de distinguir a MRD presente no sinal de EEG da atividade de fundo. A atividade realizada pelos voluntários foi um movimento único da mão direita.

METODOLOGIA

Os sinais de EEG foram coletados através de um conjunto de eletrodos e aplicados a um amplificador diferencial de banda passante entre 8 e 100Hz e de ganho 30000, com uma taxa de amostragem de 1280 amostras/s e resolução de 0,01 μ V [7]. Os sinais foram dizimados para

taxa de 256 amostras/s e aplicados a um filtro passa-faixa com banda entre 7 e 13Hz, de ordem 64, limitando-os à banda mu.

A partir de 39 algoritmos preditores no domínio do tempo, desenvolvidos em etapa anterior do Projeto BCI [5] [6], foram extraídos parâmetros dos sinais, como a média absoluta ou o número de inclinações ascendentes. Os parâmetros extraídos por cada um dos 39 preditores formam um conjunto de sinais, dos quais são tomados pontos, em instantes de tempo iguais, dentro dos intervalos referentes às duas classes, formando a base de dados utilizada. Os padrões são vetores de dimensão 40, formados por informações de cada um dos preditores e pela identificação da classe a que pertencem. Foram utilizados os dados de 3 voluntários, individualmente e combinados, com 50% dos dados separados para teste e 50% para treinamento. Os melhores classificadores foram posteriormente testados para todos os voluntários.

Com a finalidade de encontrar um classificador capaz de distinguir as duas classes, foi utilizada a programação genética. Recentemente a programação genética vem sendo utilizada como uma ferramenta para o reconhecimento de padrões [8] [9], devido à sua capacidade de encontrar relações entre os dados sob a forma de expressões matemáticas. Ela possui a capacidade de encontrar um programa de computador dentro de um espaço possível de programas, que produz uma saída numérica desejada de acordo com as entradas.

O aplicativo de programação genética utilizado foi o Lil-gp versão 1.1 [10]. O conjunto de funções utilizado é o mesmo empregado por KOZA [11] para a regressão simbólica simples: adição, subtração, multiplicação, divisão protegida, seno, cosseno, logaritmo protegido e exponencial. Para encontrar o melhor conjunto de parâmetros de controle foram testadas variações do conjunto de funções, tamanho da população, número máximo de gerações, tamanho máximo da regra, operadores, subpopulações e esquemas de trocas de indivíduos entre as subpopulações.

Os melhores resultados foram obtidos em um máximo de 100 gerações, com 3 subpopulações de 5000 indivíduos cada e com operadores e probabilidades diferentes para cada uma, conforme mostrado na tabela 1. As subpopulações 1 e 2, com parâmetros mais agressivos, geram uma diversidade maior de indivíduos. A cada 5 gerações, os 50 melhores indivíduos destas duas subpopulações são trocados pelos 50 piores indivíduos da subpopulação 3 que, por ter parâmetros de controle menos agressivos é responsável por um ajuste mais fino das regras.

Tabela 1 – Parâmetros de controle

| | subpopulação 1 | subpopulação 2 | subpopulação 3 |
|-------------------------------------|----------------|----------------|----------------------|
| Probabilidade de reprodução | 0,10 | 0,01 | 0,10 |
| Método de seleção para reprodução | melhores | Melhores | Melhores |
| Probabilidade de recombinação | 0,90 | 0,69 | 0,90 |
| Método de seleção para recombinação | <i>fitness</i> | <i>fitness</i> | torneio (200 indiv.) |
| Probabilidade de mutação | 0 | 0,30 | 0 |
| Método de seleção para mutação | - | Melhores | - |

Os padrões são classificados de acordo com o resultado numérico obtido pela regra. Caso este resultado seja menor do que zero, o padrão é classificado como padrão sem movimento; caso contrário é classificado como padrão com movimento.

A função de *fitness* da programação genética, apresentada na equação 1, foi definida como a média geométrica entre a sensibilidade (*se*) e especificidade (*es*) [12]. A sensibilidade e a especificidade são as taxas de acerto considerando somente os padrões pertencentes à classe com movimento e sem movimento, respectivamente.

O desempenho do classificador foi medido da mesma forma, com os dados de teste.

$$fitness(i) = \sqrt{se(i).es(i)} \quad (1)$$

RESULTADOS

Os primeiros treinamentos foram realizados para a base de dados de cada voluntário individualmente, gerando uma regra específica para cada voluntário. Posteriormente, com a finalidade de verificar a generalidade das melhores regras individuais, as mesmas foram aplicadas às bases de dados dos outros voluntários. Ainda buscando uma generalidade maior, foram realizados treinamentos com uma combinação dos dados dos 3 voluntários, e o melhor classificador encontrado foi testado individualmente para cada voluntário e com os dados de teste combinados.

O desempenho calculado a partir da sensibilidade e da especificidade (à esquerda) e a taxa percentual de acerto (à direita) de cada classificador encontrado pela programação genética estão listados na tabela 2. As colunas referenciam aos dados de treinamento, utilizados na busca das regras (individuais de cada voluntário e combinados), e as linhas referenciam aos dados utilizados no teste dos classificadores.

Tabela 2 – Resultados obtidos com dados de teste

| Treinamento \ Teste | voluntário 1 | | voluntário 2 | | voluntário 3 | | voluntários 1, 2 e 3 | |
|----------------------|--------------|-------|--------------|-------|--------------|-------|----------------------|-------|
| Voluntário 1 | 0,7021 | 70,7% | 0,5179 | 52,7% | 0,5443 | 54,4% | 0,5536 | 56,0% |
| Voluntário 2 | 0,4606 | 48,4% | 0,5851 | 60,7% | 0,5474 | 55,8% | 0,5718 | 58,0% |
| Voluntário 3 | 0,5029 | 50,4% | 0,5107 | 54,1% | 0,6382 | 64,1% | 0,5771 | 57,9% |
| Voluntários 1, 2 e 3 | 0,5356 | 53,7% | 0,5421 | 56,4% | 0,5887 | 58,9% | 0,5708 | 57,5% |
| Média | 0,5503 | 55,8% | 0,5390 | 45,7% | 0,5797 | 58,3% | 0,5683 | 57,4% |

O classificador obtido com os dados do voluntário nº 3 foi o que obteve melhor desempenho médio. Quanto à generalidade, a taxa de acerto média dos classificadores obtidos com os dados dos voluntários 1, 2, 3 e combinados foram respectivamente 50,8%, 54,4%, 56,4% e 57,3%. Como exemplo de regra gerada pelo sistema, a menor regra obtida foi aquela para o voluntário nº 2, mostrada nas equações 2 e 3.

$$\begin{cases} a = \cos\left(e^{\sin(inv_incl)}\right) + \log\left\{\cos\left[\log\left(\left|tivvdsd\right|\right) + timxmn + tivvcv\right]\right\} \\ b = \sin(crz_zero + slvvmn) + \sin(comp_onda * rtpn) + \sin(cuvvmn) + slvvsd \\ c = tivvmn * \sin\left[\log\left(\left|tivvmn * rcpn\right|\right)\right] \\ d = \sin(2 * media_abs + crz_zero) + \cos(crz_zero) + e^{\sin(comp_onda * rtpn)} \\ f = e^{\sin(tivvmn - amsd - tivvcv)} * \sin\left\{\log\left[\cos\left(\left|tivvmn - ctmxmn\right|\right)\right]\right\} \end{cases} \quad (2)$$

$$resultado = a + e^{\frac{b}{c}} + e^{\frac{d}{f}} \quad (3)$$

DISCUSSÃO E CONCLUSÕES

Uma das vantagens que a programação genética apresenta em relação a outros métodos tradicionalmente empregados em reconhecimento de padrões, como as redes neurais, é que a

regra encontrada pela programação genética em geral não utiliza todos os dados de entrada. Assim a programação genética elimina informação desnecessária, o que é muito interessante, pois para construir um sistema BCI seria necessário extrair os preditores do sinal de EEG em tempo real.

Os resultados obtidos neste trabalho podem ser considerados relativamente satisfatórios levando-se em consideração a qualidade dos sinais utilizados. Isto é considerado com os resultados obtidos com outras técnicas [6] [7].

Conforme o esperado, a regra encontrada para todos os voluntários obteve resultados satisfatórios tanto para os testes individuais para cada voluntário quanto para o teste com os dados combinados. Isso indica um maior grau de generalidade do classificador. Já as regras geradas para cada voluntário estão mais ajustadas para o próprio voluntário, apresentaram um desempenho mais elevado para o voluntário para o qual a regra foi encontrada em relação aos outros.

O classificador gerado para o voluntário nº 1 apresentou um desempenho muito superior para o próprio voluntário quando comparado ao desempenho obtido para os outros voluntários. Esse fato indica que a regra está sobreajustada para o voluntário. As regras encontradas para os voluntários 2 e 3 também apresentaram melhor desempenho para os respectivos voluntários, porém os resultados para os outros voluntários foram melhores (maior generalidade).

REFERÊNCIAS

- [1] PFURTSCHELLER, G., KALCHER, J., NEUPER, C., FLOTZINGER, D., PREGENZER, M., On-line EEG classification during externally-paced hand movements using a neural network-based classifier. **Electroencephalography and Clinical Neurophysiology**, v. 99, n. 5, p. 416-425, 1996.
- [2] WOLPAW, J. R., BIRBAUMER, N., HEETDERKS, W. J., McFARLAND, D. J., PECKHAM, P. H., SCHALK, G., DONCHIN, E., QUATRANO, L. A., ROBINSON, C. J., VAUGHAN, T. M., Brain-computer interface technology: a review of the first international meeting. **IEEE Transactions on Rehabilitation Engineering**, v. 8, n. 2, p. 164-173, 2000.
- [3] PILLA JÚNIOR, V., **Protocolo de aquisição de sinais para o sistema BCI**. CPGEI, Relatório Técnico, 1998.
- [4] MUKAI, I. N., PILLA JÚNIOR., V., LOPES, H. S., Processamento digital de sinais de uma interface cérebro-computador. **Seminário do Programa institucional de Bolsas de Iniciação Científica - PIBIC'99, 4p.; julho, 1999.**
- [5] MUKAI, I. N., PILLA JÚNIOR., V., LOPES, H. S., Reconhecimento de padrões de EEG relacionados ao movimento. **Seminário do Programa institucional de Bolsas de Iniciação Científica – PIBIC 2000, 4p.; julho, 2000.**
- [6] CARTA NETO, A., PILLA JÚNIOR., V., LOPES, H. S., Análise de preditores no domínio do tempo aplicados ao reconhecimento de padrões de EEG relacionados ao movimento. **Seminário do Programa institucional de Bolsas de Iniciação Científica – PIBIC 2001.**
- [7] PILLA JÚNIOR, V., **Detecção de padrões eletroencefalográficos relacionados ao movimento para uma interface cérebro-computador**. Curitiba, 1999; Dissertação (Curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Engenharia Biomédica) - CPGEI, CEFET-PR.
- [8] MARCHESI, B., **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.
- [9] KISHORE, J. K., PATNAIK, L. M., MANI, V., AGRAWAL, V. K., Application of Genetic Programming for Multicategory Pattern Classification. **IEEE Transactions on Evolutionary Computation**, v. 4, n. 3, p.242-258; setembro, 2000.
- [10] ZONGKER, D., PUNCH, B., RAND, B., **Lil-gp 1.1 – User's Manual**. Genetic Algorithms Research and Applications Group (GARAGE), Department of Computer Science, Michigan State University, 1996.
- [11] KOZA, J. R., **Genetic Programming: on the programming of computers by means of natural selection**. Cambridge: MIT Press, 1992.
- [12] LOPES, H. S., **Analogia e Aprendizado Evolucionário: Aplicação em Diagnóstico Clínico**. Florianópolis, 1996. Tese (Doutor em Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina.