

METODOLOGIA ADAPTATIVA PARA ATUALIZAÇÃO DOS PARÂMETROS DO CSP E LDA UTILIZANDO UMA ABORDAGEM BASEADA NO FK.

Odislei E. Nascimento * Wendy Y. Eras-Herrera *
Fabricio J. Erazo-Costa *

* *Departamento de Engenharia Elétrica, Universidade Federal de Ouro Preto, MG, (e-mail: odislei.nascimento@aluno.ufop.edu.br)*

Abstract: This article proposes the use of CSP (Common Spatial Patterns) as a feature extraction method for the recognition of patterns present in brain signals from motor imagery. These signals are classified using the LDA (Linear Discriminant Analysis) classifier. In order to improve the classification, an update is proposed, both in the CSP and in the LDA, based on the Kalman filter algorithm (KF). The KF has a prediction and correction structure that allows an improvement in the estimation of states in relation to linear classifiers. The results demonstrate an improvement in classification accuracy when comparing the conventional time-invariant CSP and LDA methods with the adaptive methods.

Resumo: Este artigo propõe a utilização do CSP (Common Spatial Patterns), como método de extração de características para o reconhecimento dos padrões presentes em sinais cerebrais, provenientes de uma imagética motora. Esses sinais são classificados a partir do classificador LDA (Linear Discriminant Analysis). Afim de melhorar a classificação, é proposto uma atualização, tanto no CSP, quanto no LDA, baseando-se no algoritmo do filtro de Kalman (FK). O FK possui uma estrutura de predição e correção que permite uma melhora na estimação de estados em relação a classificadores lineares. Os resultados demonstram uma melhoria da precisão na classificação quando se é comparado os métodos CSP e LDA convencionais invariantes no tempo, com os métodos adaptativos.

Keywords: brain-computer interface; common spatial patterns; Kalman filter.

Palavras-chaves: Interface cérebro-computador; Padrões Espaciais Comuns; Filtro de Kalman.

1. INTRODUÇÃO

Uma interface cérebro-máquina (ICM) é um sistema que permite a comunicação entre um sujeito e dispositivos externos controlados por atividade cerebral (Tsui, 2009). Existem várias tecnologias que permitem o registro desta atividade, sendo a eletroencefalografia (EEG) uma das mais utilizadas por ser portátil, não invasiva e de baixo custo (Ramoser and Pfurtscheller, 2000). A imagética motora (IM) é definida como um processo cognitivo ativo durante o qual a representação de uma ação específica é reproduzida internamente via memória de trabalho, sem qualquer manifestação externa do ato motor (Jackson and Doyon, 2001). Especificamente, as ICMs baseadas por IM (ICM-IM) são ativadas com a detecção da IM (tarefa motora) e podem ser aplicadas no controle de robôs, próteses e órteses com o objetivo de melhorar a qualidade de vida do sujeito ou como um meio de neuroreabilitação (Khan and Puthusserypady, 2020).

As ICMs podem ser consideradas modulares, passando por estágios como aquisição de sinais, pré-processamento,

* Reconhecimento do suporte financeiro deve vir nesta nota de rodapé.

extração de características e aplicação da interface (Mobaien, 2016). Nas ICM-IM, as características são extraídas para obter informações relacionadas com o fenômeno fisiológico das tarefas motoras como a sincronização/dessincronização relacionada ao evento (ERS/ERD), em que existe uma diminuição de energia (nos sinais de EEG) antes (banda alfa) e durante (bandas alfa e beta) a tarefa (Jeon and Whang, 2011). Desta forma, diferentes técnicas como a densidade espectral de potência (Saa and Gutierrez, 2010) e a transformada de Wavelet (Chaurasiya and Ghosh, 2015) têm sido utilizadas de forma unidimensional, enquanto técnicas como análise de componentes principais (Chaurasiya and Ghosh, 2015), análise de componentes independentes (LeVan and Gotman, 2006), padrões espaciais comuns (CSP, do inglês *Common Spatial Patterns*) (Mobaien, 2016), têm sido utilizadas no ambiente multidimensional.

O CSP realiza uma filtragem espacial para descorrelacionar os dados de duas classes, encontrando uma matriz de projeção com o objetivo de maximizar as variâncias entre duas classes (máxima variância da classe 1 e mínima da classe 2, ou vice-versa). O CSP tem sido bastante utilizado na classificação entre IM de diferentes classes como da mão direita e da mão esquerda (Müller-Gerking and Flyvbjerg,

1999; Ramoser and Pfurtscheller, 2000) oferecendo taxas de acerto que poderiam ser utilizadas em aplicações em tempo real para a ICM-IM. O CSP possui um modelo fixo robusto mas não consegue seguir a natureza dinâmica das ICMs.

O filtro de Kalman (FK) é um algoritmo que estima os estados de um sistema a partir de um modelo e um conjunto de medições ruidosas (Kalman, 1960). O FK tem um baixo esforço computacional e possui características recursivas que podem ser utilizadas para a estimação de parâmetros, tanto do classificador, quanto no método de extração de características. No trabalho desenvolvido por Tsui (2009), utiliza-se o FK para realizar a atualização dos parâmetros do classificador análise de discriminante linear (LDA, do inglês *Linear Discriminant Analysis*), utilizando como características as bandas de potência em um ambiente ICM-IM. Esta metodologia mostrou uma melhora na taxa de acerto utilizando o método proposto adaptativo do LDA, comparado com o classificador LDA com parâmetros invariantes no tempo. O estudo desenvolvido por Mobaien (2016) propõe um método inspirado no FK estendido para atualizar tanto a matriz de projeção CSP (utilizada para calcular as características), quanto o classificador baseado na distância euclidiana. O método CSP adaptativo apresenta uma melhora na taxa de acerto comparado com o método convencional.

O CSP convencionalmente é utilizado em conjunto com o LDA como métodos de extração de características e classificação respectivamente em problemas de discriminação de tarefas motoras (Tsui, 2009). Estes modelos são geralmente invariantes no tempo e dependem exclusivamente dos dados utilizados no treinamento. Devido à natureza dinâmica da ICM-IM, métodos adaptativos que atualizem os parâmetros variantes no tempo dos modelos são necessários para melhorar o desempenho da interface. Desta forma, o presente trabalho propõe a implementação de um algoritmo baseado na estrutura do FK, afim de atualizar os parâmetros do LDA e a matriz de características do CSP com a entrada de novos trechos de sinal enviados para a classificação. Na abordagem proposta são combinadas as duas metodologias propostas por Tsui (2009) e Mobaien (2016) com o intuito de melhorar a taxa de acerto do classificador LDA.

2. MATERIAIS E MÉTODOS

2.1 Protocolo Experimental

O conjunto de dados de EEG estudado é proveniente do trabalho desenvolvido por (Brunner and Pfurtscheller, 2008) e consta do registro de 22 sinais (Figura 1a) de EEG com uma frequência de amostragem $f_s = 250Hz$ de 9 indivíduos saudáveis das tarefas de IM: da mão direita (MD), mão esquerda (ME), língua e pés.

Os sujeitos foram posicionados em uma poltrona na frente de um computador. A sessão do trecho iniciava ($t=-3s$) com a apresentação de uma cruz na tela por 2s e por um breve estímulo sonoro. Posteriormente, uma seta era mostrada ($t=-1s$) por 1,25s para indicar o tipo de IM que seria realizada, em que: seta para esquerda (IM ME), seta para direita (IM MD), seta para cima (IM da língua), seta para baixo (IM dos pés). No instante de tempo

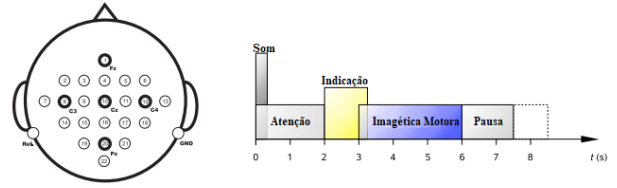


Figura 1. a) Disposição dos 22 eletrodos utilizados para coleta dos sinais, b) Metodologia de temporização da coleta de dados.

$t=0s$ o sujeito realizava a IM por 3s. Uma pausa de aproximadamente 1,5s a 2,5s era colocada para o início do próximo trecho.

2.2 Common Spatial Patterns

O CSP é uma técnica multidimensional que permite encontrar a máxima variância entre duas classes (máxima variância da classe 1 e mínima da classe 2 ou vice-versa). O algoritmo CSP consiste em encontrar a matriz de projeção W_{CSP} que permita projetar os dados do sinal $E \in R^{r \times p}$ de um único trecho de EEG, em que r é o número de canais e p é o número de amostras. O sinal filtrado $S \in R^{2r \times p}$ é dado por:

$$S = W_{CSP}^T E \quad (1)$$

em que as primeiras componentes de S proporcionam máxima variância para a classe 1 e mínima variância para a classe 2, enquanto as últimas componentes de S fornecem a mínima variância para classe 1 e máxima variância para a classe 2. As características são encontradas a partir da seleção das primeiras e últimas M componentes do sinal filtrado S como:

$$X_i = \log \left(\frac{\text{var}(S_i)}{\sum_{i=1}^{2M} \text{var}(S_i)} \right). \quad (2)$$

2.3 Análise Discriminante Linear

O LDA é um classificador que visa encontrar um hiperplano linear afim de separar duas ou mais classes. Para o caso de classificação binária, com conjuntos de dados para as classes 1 e 2. Os parâmetros do LDA são definidos pelos coeficientes a_1 e a_0 dados por:

$$a_1 = \sum^{-1} (\mu_1 - \mu_2) \quad (3)$$

$$a_0 = -0.5(\mu_1^T \sum^{-1} \mu_1 - \mu_2^T \sum^{-1} \mu_2) \quad (4)$$

em que \sum^{-1} é a matriz de covariância comum às duas classes e μ_x é a média das características de cada classe, $x = 1, 2$. Esses parâmetros são utilizados para testar e classificar os sinais através de uma função de teste dada por:

$$X(a_1) + a_0. \quad (5)$$

em que X é dado pelas características do sinal a ser classificado obtido em (2). Caso essa função seja maior ou igual a zero, o sinal pertence a classe 1, caso contrário, ele pertence a classe 2.

2.4 Filtro de Kalman

O filtro de Kalman é um algoritmo recursivo que estima os estados a partir de um modelo e um conjunto de medições

ruidosas (Kalman, 1960). O modelo do sistema dinâmico é representado em forma de espaço de estados como:

$$x_k = A_{k-1}x_{k-1} + B_k u_{k-1} + w_{k-1}, \quad (6)$$

$$y_k = C_k x_k + v_k, \quad (7)$$

em que: $A_{k-1} \in R^{n \times n}$ é a matriz de transição, $B_{k-1} \in R^{n \times m}$ é a matriz de entrada do sistema, $C_k \in R^{m \times n}$ é a matriz de saída e todas elas são conhecidas. x_{k-1} é o vetor das medições passadas do sinal de EEG e $y_k \in R^m$ é o vetor do sinal de EEG medido.

O algoritmo do FK compreende duas etapas: a *etapa de predição* e a *etapa de assimilação de dados*. Para a inicialização, assuma que o vetor aleatório x_0 é Gaussiano com média $\hat{x}_{0|0}$ e covariância $P_{0|0}^{xx} = E[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T]$, ambas conhecidas.

O FK pode ser utilizado para estimar a matriz de projeção do CSP. Desta forma, a *etapa de predição* é responsável pela predição dos valores da matriz de projeção a partir das informações disponíveis *a priori*, dada por:

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1}, \quad (8)$$

$$P_{k|k-1}^{xx} = A_{k-1}P_{k-1|k-1}^{xx}A_{k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T, \quad (9)$$

$$\hat{y}_{k|k-1} = C_k\hat{x}_{k|k-1}, \quad (10)$$

$$P_{k|k-1}^{yy} = C_k P_{k|k-1}^{xx} C_k^T + R_k, \quad (11)$$

$$P_{k|k-1}^{xy} = P_{k|k-1}^{xx} C_k^T. \quad (12)$$

em que $P_{k|k-1}^{xx} = E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T]$, $P_{k|k-1}^{yy} = E[(y_k - \hat{y}_{k|k-1})(y_k - \hat{y}_{k|k-1})^T]$, $P_{k|k-1}^{xy} = E[(x_k - \hat{x}_{k|k-1})(y_k - \hat{y}_{k|k-1})^T]$ são as matrizes de covariância do erro de predição, covariância da inovação e covariância cruzada, respectivamente.

A *etapa de assimilação de dados* incorpora uma nova informação do sinal medido com o intuito de melhorar a extração de características dos dados desse sinal:

$$K_k = P_{k|k-1}^{xy} (P_{k|k-1}^{yy})^{-1}, \quad (13)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - \hat{y}_{k|k-1}), \quad (14)$$

$$P_{k|k}^{xx} = P_{k|k-1}^{xx} - K_k P_{k|k-1}^{yy} K_k^T. \quad (15)$$

2.5 Método Adaptativo para Atualizar os Parâmetros do CSP e LDA

A abordagem proposta atualiza os parâmetros do método de extração de características CSP a partir de um fator de esquecimento e os parâmetros do classificador LDA utilizando uma abordagem baseada no FK, de forma a combinar as metodologias utilizadas por Tsui (2009) e Mobaien (2016).

Neste trabalho, a abordagem proposta envolve duas etapas: *etapa de atualização dos parâmetros CSP* e *etapa de atualização dos parâmetros LDA*. Os parâmetros adaptados são: a matriz de projeção W_{CSP} do método CSP e os coeficientes a_1 e a_0 do classificador LDA.

Para o método de extração de características CSP, encontra-se um modelo com as observações de treinamento afim de encontrar uma matriz de projeção W_{CSP} . As matrizes de projeção desse passo são utilizadas posteri-

ormente para calcular uma matriz W_{CSP} atualizada. Para isso, é utilizado:

$$\bar{C}_{k|k}^r = (1 - \alpha)\bar{C}_{k-1|k-1}^r + \alpha C_x \quad (16)$$

em que r indica a classe (mão esquerda ou mão direita), C_x indica a matriz de covariância espacial normalizada do novo trecho x e α é o fator de esquecimento que designa a efetividade, maior ou menor, da nova informação em relação a matriz CSP antiga. Esse valor varia entre 0 e 1. Após atualizar a matriz de covariância, uma nova matriz W_{CSP} é calculada e atualizada usando as equações (1) e (2). Essa matriz depende das matrizes de covariância antigas e da matriz atualizada, para que seja feita a extração de características a partir da W_{CSP} atualizada.

A atualização dos parâmetros do LDA é feita a partir de equações baseadas no FK, que foram adaptadas para um contexto de classificação de sinais de EEG provenientes de IM. Desta forma, a *etapa de predição* é dada por

$$W_{k|k-1} = [a_1; a_0] \quad (17)$$

$$P_{k|k-1} = P_{k-1|k-1} + Q_{k-1} \quad (18)$$

em que

$$Q_{k-1} = \max\{0, (u_{k|k-1} - u_{k-1|k-1})\}I \quad (19)$$

$$u = y(1 - y) \quad (20)$$

em que $P_{k|k-1}$ é a variância de $W_{k|k-1}$, I é uma matriz identidade, u representa a incerteza da saída do classificador y . Nessa etapa, é possível verificar algumas diferenças entre o método do FK e o utilizado para atualização do LDA. A equação (8) representa o sinal de análise, enquanto a equação (17) representa os parâmetros do LDA a serem atualizados. Por outro lado, para a *etapa de assimilação de dados* é dada por

$$W_{k|k} = W_{k|k-1} + K_k (Z_k - y_{k|k-1}) \quad (21)$$

$$P_{k|k} = P_{k|k-1} - K_k u_{k|k-1} (P_{k|k-1} \phi_k)^T \quad (22)$$

em que

$$K_k = \frac{P_{k|k-1}}{C_{cons} + u_{k|k-1} s_{k|k-1}^2} \quad (23)$$

$$s_{k|k-1}^2 = \phi_k^T P_{k|k-1} \phi_k \quad (24)$$

$$\phi = [X; 1] \quad (25)$$

em que ϕ é o vetor de características do sinal, K_k é o ganho do filtro de Kalman, C_{cons} é um parâmetro de aprendizado que varia de indivíduo para indivíduo, e z é a identificação do sinal, sendo 1 para mão esquerda, e 0 para mão direita.

3. RESULTADOS E DISCUSSÃO

Os sinais de EEG da IM MD e IM ME dos 9 indivíduos foram segmentados em trechos de 3s ($t = 0s$ até $t = 3s$) totalizando 53 trechos por cada classe e sujeito. Estes sinais foram filtrados digitalmente utilizando um filtro digital *Butterworth* (fase zero) de ordem 4, de 8 a 30Hz. Destes 53 trechos de cada classe, 35 e 18 trechos foram selecionados para treinamento e teste, respectivamente. Além disso, foram segmentados trechos de 3s de EEG espontâneo (ESP) até completar os 53 trechos.

Para extração de características foram utilizadas as $M = 3$ primeiras e últimas componentes do W_{CSP} (utilizando os

35 trechos de treinamento por classe) para encontrar as características de cada trecho conforme dado em (2). Da mesma forma, os coeficientes a_1 e a_0 dados por (3) e (4) foram encontrados para o classificador LDA utilizando as 35 amostras de treinamento por classe.

A abordagem proposta atualiza os parâmetros do CSP e LDA para as classificações binárias de IM MD vs. ESP, IM ME vs. ESP e IM MD vs. IM ME. Para comparar a metodologia adaptativa proposta, foram considerados os casos com CSP com modelo de parâmetros fixos obtidos a partir das observações do treinamento (CF), CSP com a matriz de projeção W_{CSP} atualizada (CA) a cada entrada de um novo trecho e com inicialização do modelo CF, LDA com modelo com parâmetros fixos obtidos a partir das observações do treinamento (LF) e LDA com modelo com coeficientes atualizados (LA) a cada entrada de um novo trecho e com inicialização do modelo LF. Desta forma, foram realizadas quatro metodologias combinadas: CF LF, CA LF, CF LA, CA LA.

Na Figura 2 mostra-se as taxas de acerto dos métodos CF LA (traçado em cor azul) e CA LA (traçado em cor vermelha) obtidas para o indivíduo S_2 com diferentes valores de C_{cons} . Para encontrar estes valores as 35 amostras de treinamento foram divididas em 20 amostras para o treinamento e 15 amostras para a validação. Observa-se que os valores de $C_{cons} = 0$ mostram os maiores valores de taxa de acerto. Desta forma, a abordagem proposta considera $C_{cons} = 0$ (equação 23) para os modelos com LA e $\alpha = 0.05$ (equação 16) para os modelos com CA.

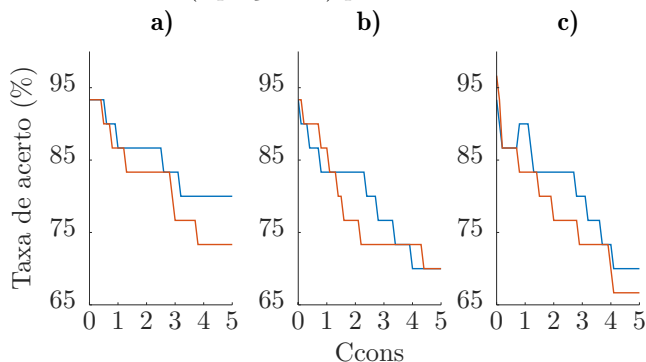


Figura 2. Taxas de acerto (%) dos métodos CF LA e CA LA obtidas para o indivíduo S_2 com diferentes valores de C_{cons} para classificar: a) IM MD vs. ESP, b) IM ME vs. ESP e c) IM MD vs. IM ME. Os traçados mostram o método CF LA (linha contínua azul) e CA LA (linha contínua vermelha).

As Tabelas 1-3 mostram as taxas de classificação para IM MD vs. ESP, IM ME vs. ESP e IM MD vs. IM ME, respectivamente, utilizando os métodos CF LF, CA LF, CF LA e CA LA. Na Tabela 1 observa-se que a média das taxas de acertos dos 9 indivíduos é maior para o caso de CF LA com aproximadamente 7% superior ao caso clássico CF LF e em torno de 10% e 1% superior para o método CA LF e CA LA, respectivamente, quando comparados com o método CF LF. Observa-se que nos indivíduos S_2 , S_5 e S_9 houve um aumento na taxa de acerto entre 13% a 15% de CF LF e CA LF comparado com CF LA e CA LA. Entretanto, somente houve uma redução da taxa de acerto para o indivíduo S_7 quando foram utilizados os métodos CF LA ou CA LA.

Tabela 1. Taxa de acerto (%) para a classificação de IM MD vs. ESP utilizando os métodos CF LF, CA LF, CF LA e CA LA.

	CF LF	CA LF	CF LA	CA LA
S1	94,44	94,44	97,22	97,22
S2	75,00	72,22	94,44	94,44
S3	97,22	97,22	97,22	97,22
S4	97,22	91,67	97,22	97,22
S5	72,22	72,22	94,44	94,44
S6	91,67	91,67	97,22	97,22
S7	94,44	91,67	88,89	83,33
S8	91,67	86,11	97,22	91,67
S9	77,78	72,22	91,67	94,44
Média	87,96	85,49	95,06	94,14

Tabela 2. Taxa de acerto (%) para a classificação de IM ME vs. ESP utilizando os métodos CF LF, CA LF, CF LA e CA LA.

	CF LF	CA LF	CF LA	CA LA
S1	88,89	86,11	94,44	88,89
S2	80,56	77,78	94,44	94,44
S3	91,67	94,44	94,44	97,22
S4	91,67	91,67	97,22	97,22
S5	75,00	80,56	94,44	94,44
S6	94,44	94,44	94,44	86,11
S7	100,00	100,00	100,00	100,00
S8	97,22	94,44	97,22	97,22
S9	91,67	91,67	97,22	97,22
Média	90,12	90,12	95,99	94,75

Tabela 3. Taxa de acerto (%) para a classificação de IM MD vs. IM ME utilizando os métodos CF LF, CA LF, CF LA e CA LA.

	CF LF	CA LF	CF LA	CA LA
S1	83,33	86,11	94,44	91,67
S2	72,22	72,22	91,67	94,44
S3	100,00	97,22	100,00	97,22
S4	91,67	94,44	94,44	94,44
S5	75,00	75,00	97,22	97,22
S6	86,11	83,33	97,22	97,22
S7	86,11	88,89	97,22	97,22
S8	97,22	94,44	91,67	97,22
S9	83,33	77,78	97,22	97,22
Média	86,11	85,49	95,68	95,99

Na Tabela 2 mostra-se um padrão análogo ao comparado com a Tabela 1. Observa-se que houve um aumento de quase 5% na média quando foram utilizados os métodos CF LA e CA LA. Além disso, observa-se que o indivíduo S_5 obteve uma melhora na taxa de acerto de aproximadamente 10%. Na Tabela 3 mostra-se uma média da taxa de acerto maior para o método CA LA. Observa-se que os indivíduos S_2 e S_5 apresentam uma taxa de acerto superior a 20%.

Na Figura 3 mostra-se a evolução da taxa de acerto na classificação da IM MD vs. ESP, IM ME vs. ESP e IM MD vs. IM ME para o indivíduo S_2 com 36 amostras de teste utilizando os métodos CF LF (traçado em azul), CA LF (traçado em vermelho), CF LA (traçado em amarelo) e CA LA (traçado em roxo). Nestas metodologias foi considerado que quando um acerto é atingido, é somado 1, caso contrário, 0. Observa-se que para as metodologias CF LA e CA LA seguem quase uma linha reta enquanto que para os métodos CF LF e CA LF este desempenho baixa principalmente na metade das amostras de teste.

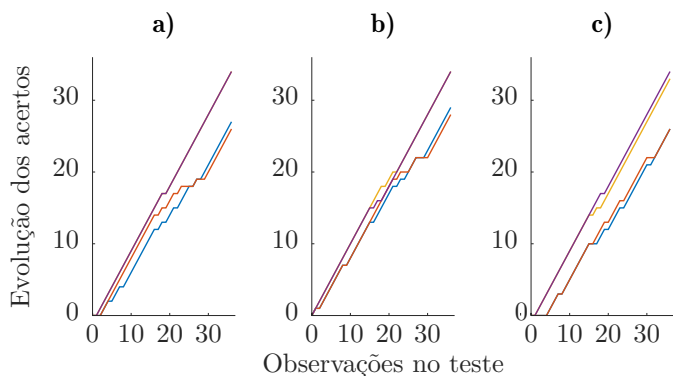


Figura 3. Evolução da taxa de acerto (%) para classificar: a) IM MD vs. ESP, b) IM ME vs. ESP e c) IM MD vs. IM ME, obtidas para o indivíduo S_2 utilizando os métodos CF LF, CA LF, CF LA e CA LA. Os traçados mostram o método CF LF (linha azul), CA LF (linha vermelha), CF LA (linha amarelo) e CA LA (linha roxa).

4. CONCLUSÕES

As metodologias adaptativas que atualizam os parâmetros do classificador do LDA permitiram um aumento considerável na taxa de acerto, enquanto que a atualização dos parâmetros do método de extração de características CSP não gerou impacto na taxa de acerto. Vale ressaltar que utilizando as duas metodologias de LDA adaptativas, CF LA e CA LA, os piores indivíduos obtiveram um aumento na taxa de acerto. Desta forma, os métodos LDA adaptativos mostraram-se como métodos promissores para aplicações de ICM-IM em tempo real.

Como proposta para trabalhos futuros, atualizar a matriz W_{CSP} de forma análoga ao classificador LDA, utilizando uma lógica recursiva e adaptativa baseada no algoritmo do FK, afim de verificar seu desempenho no método CSP adaptativo. Além disso, implementar novos métodos de extração de características, como as redes neurais ou o SVM (Support Vector Machine), afim de analisar o desempenho na taxa de acerto e o tempo de processamento desse novo classificador, em comparação com o método CSP.

AGRADECIMENTOS

Os autores agradecem a instituição de ensino Universidade Federal de Ouro Preto por viabilizar suas pesquisas.

REFERÊNCIAS

- Brunner, C., L.R.M.P.G.S.A. and Pfurtscheller, G. (2008). Bci competition 2008-graz data set a. *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, 16, 1–6.
- Chaurasiya, R. K., L.N.D. and Ghosh, S. (2015). Statistical wavelet features, pca, and svm based approach for eeg signals classification. *Latin American and Caribbean Conference for Engineering and Technology*, 9, 182–186.
- Jackson, P. L., L.M.F.M.F.R.C. and Doyon, J. (2001). Potential role of mental practice using motor imagery in neurologic rehabilitation. *American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation*, 82, 1133–1141.

- Jeon, Y., N.C.S.K.Y.J. and Whang, M.C. (2011). Event-related (de) synchronization (erd/ers) during motor imagery tasks: Implications for brain-computer interfaces. *International Journal of Industrial Ergonomics*, 41, 428–436.
- Kalman, R.E. (1960). A new approach to linear filtering and prediction problems.
- Khan, M. A., D.R.I.H.K. and Puthusserypadu, S. (2020). Review on motor imagery based bci systems for upper limb post-stroke neurorehabilitation: From designing to application. *Computers in Biology and Medicine*, 123, 1–17.
- LeVan, P., U.E. and Gotman, J. (2006). A system for automatic artifact removal in ictal scalp eeg based on independent component analysis and bayesian classification. *Clinical neurophysiology*, 117, 912–927.
- Mobaeni, A., B.R. (2016). Acsp: Adaptive csp filter for bci application. *Iranian Conference on Electrical Engineering (ICEE)*, 24th, 466–471.
- Müller-Gerking, J., P.G. and Flyvbjerg, H. (1999). Designing optimal spatial filters for single-trial eeg classification in a movement task. *Clin. Neurophysiol*, 110, 787–798.
- Ramoser, H., M.G.J. and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE Trans. Rehabil. Eng*, 8, 441–446.
- Saa, J.F.D. and Gutierrez, M.S. (2010). Eeg signal classification using power spectral features and linear discriminant analysis: A brain computer interface application. *Latin American and Caribbean Conference for Engineering and Technology*, 8, 1–7.
- Tsui, C. S. L., G.J.Q.R.S.J. (2009). A self-paced brain-computer interface for controlling a robot simulator: an online event labelling paradigm and an extended kalman filter based algorithm for online training. *Medical & biological engineering & computing*, 47.3, 257–265.