

**UNIVERSIDADE FEDERAL DE VIÇOSA
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

ÁLVARO FURLANI BASTOS

**ESTUDO DE MÉTODOS CLASSIFICATÓRIOS DA
ATENÇÃO SELETIVA AUDITIVA A PARTIR DE
POTENCIAIS CEREBRAIS EVOCADOS**

**VIÇOSA
2014**

ÁLVARO FURLANI BASTOS

**ESTUDO DE MÉTODOS CLASSIFICATÓRIOS DA
ATENÇÃO SELETIVA AUDITIVA A PARTIR DE
POTENCIAIS CEREBRAIS EVOCADOS**

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 490 - Monografia e Seminário e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Orientador: Prof. Dr. Leonardo Bonato Felix.

VIÇOSA
2014

ÁLVARO FURLANI BASTOS

**ESTUDO DE MÉTODOS CLASSIFICATÓRIOS DA ATENÇÃO
SELETIVA AUDITIVA A PARTIR DE POTENCIAIS CEREBRAIS
EVOCADOS**

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 490 - Monografia e Seminário e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Aprovada em 17 de Junho de 2014.

COMISSÃO EXAMINADORA

Prof. Dr. Leonardo Bonato Felix - Orientador
Universidade Federal de Viçosa

Prof. Dr. Alexandre Santos Brandão - Membro
Universidade Federal de Viçosa

Prof. B.Sc. Paulo Fábio Figueiredo Rocha - Membro
Universidade Federal de Viçosa

Aos meus pais Rita e Euzébio, e a todos que torcem pelo meu sucesso.

Agradecimentos

Inicialmente agradeço a Deus por me propiciar saúde e determinação para percorrer o caminho na busca de meus objetivos, superando cada obstáculo encontrado, ao mesmo tempo colocando pessoas ao meu redor que sempre me entusiasmaram (ou me embebedaram) nas diversas situações.

À minha família, em especial meus pais Rita e Euzébio, minha irmã Nathália e meus sobrinhos Hugo, Cauã e Júlia por estarem sempre presentes quando precisei e pela confiança depositada em mim, mesmo em meus momentos de ausência.

Aos professores dos diversos departamentos da UFV com os quais tive contato, por possibilitarem que eu adquirisse o conhecimento utilizado nesse trabalho e em minha futura vida profissional. Especial agradecimento para os professores do Departamento de Engenharia Elétrica que me influenciaram positivamente de muitas formas.

Ao CNPq (PIBIC) pelo apoio financeiro concedido.

Por último, mas não menos importante, a todos os meus amigos nos 5 anos em Viçosa (os quais não ousarei citar para não correr o risco de injustamente esquecer algum), responsáveis por amenizar as dificuldades e stress nas situações de incerteza, assim como proporcionar tantos momentos inesquecíveis.

“It is paradoxical, yet true, to say, that the more we know, the more ignorant we become in the absolute sense, for it is only through enlightenment that we become conscious of our limitations. Precisely one of the most gratifying results of intellectual evolution is the continuous opening up of new and greater prospects.”

Nikola Tesla

Resumo

A existência de pessoas com limitações motoras desperta o interesse no desenvolvimento de alternativas que possibilitem sua comunicação com outras pessoas e até mesmo o controle de dispositivos externos por meio de suas ondas cerebrais. Uma possibilidade é utilizar o foco de sua atenção seletiva auditiva detectada nos potenciais cerebrais coletados em seu escalpo através de eletroencefalografia. Para que esse método possa ser utilizado eficientemente em Interfaces Cérebro-Computador, é necessário a escolha apropriada das técnicas de classificação de sua atenção. Os dados para análise nesse trabalho correspondem aos sinais coletados de 14 indivíduos (sem presença de disordens auditivas) sob estimulação de tom AM binaural, sendo orientados a focar sua atenção nos estímulos das orelhas esquerda e direita, respectivamente. Esses dados foram processados para eliminação de ruídos e artefatos oculares, além da redução de dimensão e obtenção dos parâmetros de um modelo autorregressivo que o represente. Os classificadores analisados foram Análise de Discriminante, Máquina de Vetores de Suporte, Rede Neural Artificial e *k-Nearest Neighbors*; cada um deles foi testado individualmente para obter a melhor configuração, as quais foram posteriormente combinadas para atuar simultaneamente. Na análise individual, as maiores taxas de acerto concentraram-se entre 83,3 e 91,7% para cerca de 17 segundos de coleta de dados; por outro lado, ao utilizá-los em conjunto (onde a classificação final é dada de acordo com a maioria das classificações individuais), a taxa de acerto alcançou 100%. Outro fator analisado foi a Taxa de Transferência de Informação, cujo valor médio foi 2 bits por minuto. A elevada taxa de acerto obtida, assim como a Taxa de Transferência de Informação, indicam a aplicabilidade dessas técnicas de classificação para Interfaces Cérebro-Computador, pelo menos para a realização de tarefas simples.

Abstract

The existence of people with physical limitations creates interest in the development of alternatives that enable communication with other people and even control of external devices using their brain waves. One possibility is to use the focus of their auditory selective attention detected in brain potentials collected in their scalp through electroencephalography. For the efficient use of this method in Brain-Computer Interfaces, it is necessary to choose the appropriate techniques for classifying their attention. The data for analysis in this work corresponds to the signals collected from 14 individuals (without auditory disorders) under binaural stimulation tone AM, which were instructed to focus their attention on stimuli at left and right ears, respectively. These data were processed to eliminate noise and ocular artifacts, as well as dimension reduction and obtainment of parameters of an autoregressive model to represent it. The classifiers analyzed are Discriminant Analysis, Support Vector Machines, Artificial Neural Network and k -Nearest Neighbors; each one was tested individually for the best configuration, which were subsequently combined to act simultaneously. In the individual analyzes, the highest success rates were between 83,3 and 91,7% for about 17 seconds of data collection; on the other hand, using them together (where the final classification is given according to the majority of individual classifications), the success rate reached 100%. Another factor considered is the Information Transfer Rate whose values ranged around 2 bits per minute. The high success rate, as well as Information Transfer Rate, indicate the applicability of these classification techniques in Brain-Computer Interfaces, at least for simple tasks.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	11
1.1	Objetivos	12
2	Revisão Bibliográfica	13
2.1	Potencial Evocado	13
2.2	Atenção Seletiva Auditiva	14
2.3	Interface Cérebro-Computador	14
2.4	Taxa de Transferência de Informação	16
2.5	Processamento de Sinais	16
2.5.1	Rejeição de artefatos	16
2.5.2	Análise de Componentes Principais	17
2.5.3	Modelo Autorregressivo	18
2.6	Métodos de Classificação	19
2.6.1	Análise de Discriminante	19
2.6.2	Máquina de Vetores de Suporte	20
2.6.3	Redes Neurais Artificiais	22
2.6.4	<i>k-Nearest Neighbors</i>	23
2.6.5	Combinação de Classificadores	25

3	Materiais e Métodos	26
3.1	Protocolo Experimental	26
3.2	Estimulação	27
3.3	Aquisição de Sinais	28
3.4	Treinamento e Classificação	29
4	Resultados e Discussões	32
5	Conclusões	38
	Referências	39
	Anexo A – Número de Componentes da PCA	42

Lista de Figuras

1	Design e operação básica de uma ICC (adaptada de [1]).	15
2	Hiperplano de separação de classes da Análise de Discriminante Linear. . .	20
3	Classificação binária bidimensional através de Máquina de Vetores de Su- porte.	21
4	Mapeamento de pontos para dimensão superior.	22
5	Estrutura de uma Rede Neural Artificial.	23
6	Representação do método <i>k</i> -NN bidimensional (2 e 5 classes, respectiva- mente).	24
7	Fluxo de processos para o reconhecimento de padrões auditivos.	26
8	Linha temporal da estimulação em cada um dos ouvidos.	28
9	Sistema Internacional 10-20 para posicionamento de eletrodos.	29
10	Variação da taxa de acerto e taxa de transferência de informação em função do número de janelas consideradas.	37
11	<i>Scree plot</i> dos arquivos com estimulação no ouvido esquerdo.	43
12	<i>Scree plot</i> dos arquivos com estimulação no ouvido esquerdo.	44

Lista de Tabelas

1	Parâmetros de estimulação do tom AM.	28
2	Taxas máximas de acerto para Análise de Discriminante (dadas em %). . .	32
3	Taxas máximas de acerto para Máquina de Vetores de Suporte (dadas em %).	32
4	Taxas máximas de acerto para Rede Neural Artificial com 1 camada oculta (dadas em %).	33
5	Taxas máximas de acerto para Rede Neural Artificial com 2 camadas ocultas e treinamento por <i>Levenberg-Marquardt</i> (dadas em %).	33
6	Taxas máximas de acerto para Rede Neural Artificial com 2 camadas ocultas e treinamento por Gradiente Conjugado Escalonado (dadas em %). . .	33
7	Taxas máximas de acerto para <i>k-Nearest Neighbors</i> com pesos iguais (dadas em %).	34
8	Taxas máximas de acerto para <i>k-Nearest Neighbors</i> com pesos inversos (dadas em %).	34
9	Análise conjunta das melhores configurações dos classificadores.	36
10	Variância acumulada em função do número de Componentes Principais . .	42

1 *Introdução*

Diversos estudos interdisciplinares procuram por formas alternativas de comunicação para permitir que pessoas expressem seus desejos e controlem dispositivos sem a necessidade de movimentos musculares. Uma das alternativas em desenvolvimento são as Interfaces Cérebro-Computador (ICC), que utilizam as ondas cerebrais para permitir que pessoas com alguma limitação muscular se comuniquem com outras pessoas, controlem membros artificiais ou o ambiente ao seu redor (por exemplo, temperatura e luminosidade). O estudo de ICC envolve a análise de técnicas invasivas ou não-invasivas para detectar a atividade cerebral, avaliação dos sinais de controle (tipos de padrões da atividade cerebral que podem ser utilizadas com êxito para comunicação), desenvolvimento de algoritmos para transformação dos sinais cerebrais em comandos computacionais e desenvolvimento de novas aplicações para esses sistemas [2]. Tais interfaces tem produzido resultados satisfatórios para pessoas com paralisia total ou parcial (resultante de acidente vascular cerebral, paralisia cerebral, esclerose lateral amiotrófica), uma vez que suas funções cognitivas permanecem intactas e são capazes de aprender a se comunicar através de uma interface comandada por sinais elétricos do cérebro [3].

Entre as várias técnicas de mapeamento da atividade cerebral, tais como ressonância magnética funcional (fMRI), espectroscopia de infravermelho, a eletroencefalografia (EEG) tem sido a mais utilizada por ser não-invasiva, econômica, prontamente aplicável e com excelente resolução temporal [4]. EEG corresponde à medição neurofisiológica da atividade elétrica cerebral a partir do registro de eletrodos conectados no escalpo humano, refletindo a soma ponderada dos potenciais pós-sinápticos. O avanço da tecnologia disponível possibilitou melhor entendimento da atividade cerebral, demonstrando correlação entre sinais de EEG e movimentos reais ou imaginados e atividades mentais [5].

Diversos paradigmas tem sido testados em pacientes com diversos graus de limitação. Em muitos deles, longos períodos de imobilidade e degeneração das células piramidais do córtex motor dificultam a produção de movimentos imaginados; em outros casos, a visão pode estar comprometida (dificuldade em ajustar o foco ou orientar a retina em diferentes

posições), além da perda de concentração ao manter a visão fixa por longo tempo [3]. Outra desvantagem presente em estímulos dependentes da visão é a contaminação do sinal por potenciais visuais evocados pelo piscar dos olhos e movimento do globo ocular [4].

Para superar essas limitações, vários estudos foram desenvolvidos para analisar se a resposta auditiva em regime permanente é modulada pela atenção seletiva auditiva, e conseqüentemente poder ser utilizada como paradigma em ICC. Apesar dos primeiros resultados serem contraditórios, estudos posteriores comprovaram a modulação das respostas originadas no córtex, da mesma forma que ocorre com os estímulos visuais [6]. Sendo um paradigma que requer relativamente pouco treinamento e que, em geral, não está comprometido nos pacientes, o estímulo auditivo tornou-se uma opção amplamente utilizada em Interfaces Cérebro-Computador.

1.1 **Objetivos**

O objetivo do presente trabalho é analisar diversas técnicas de classificação da atenção seletiva auditiva a partir de potenciais cerebrais evocados obtidos no escalpo humano. A classificação será binária, representando o foco da atenção em cada um dos ouvidos (os quais serão estimulados por sinais sonoros com diferentes frequências de modulação). Além de analisar o desempenho individual de cada um dos classificadores, também será considerado o funcionamento conjunto e simultâneo de vários algoritmos, de forma a abordar um ponto importante ainda pouco presente na literatura atual [7].

Considerando trabalhos prévios realizados nessa mesma linha de pesquisa [8, 9], o presente estudo tratará de forma abrangente as principais vantagens dos métodos de classificação já testados, analisando novos métodos e propondo novas alternativas para sua utilização.

2 *Revisão Bibliográfica*

Esse capítulo fornece uma revisão bibliográfica dos tópicos a serem considerados posteriormente, apresentando os aspectos relevantes à execução do presente trabalho. Uma visão mais detalhada (porém não necessária nesse contexto) pode ser encontrada nas referências utilizadas.

2.1 **Potencial Evocado**

Potencial evocado é o potencial elétrico gerado pelo sistema nervoso de seres humanos e outros animais em resposta à apresentação de um estímulo externo [10]. Esses sinais podem ser registrados a partir do córtex cerebral, tronco cerebral, medula espinhal, apresentando, em geral, baixa amplitude (na ordem de microvolts) quando comparados a sinais de eletroencefalograma (EEG) e eletrocardiograma (ECG). A presença contínua desses sinais de fundo exige o emprego de técnicas de promediação e processamento de sinais para revelar a atividade cerebral evocada [11].

As respostas evocadas são classificadas em transitórias e de regime permanente; no primeiro caso, os estímulos possuem baixa taxa de repetição ($< 2Hz$), possibilitando que cada resposta termine antes do início do próximo estímulo. A resposta de regime permanente ocorre quando os estímulos são apresentados a uma taxa suficientemente elevada ($> 5Hz$), acarretando sobreposição das respostas de diferentes estímulos.

Há três formas de potencial evocado amplamente utilizadas em clínica médica e geralmente medidas no escalpo: auditivo, originado no tronco cerebral; visual; e somatossensorial, induzido pela estimulação elétrica do nervo periférico. O potencial evocado auditivo, que será o objeto de estudo desse trabalho, reflete a atividade neural no nervo e núcleo coclear, córtex auditivo e tronco cerebral, podendo ser detectado de forma não-invasiva mesmo em indivíduos sedados ou anestesiados.

A resposta auditiva de regime permanente (*ASSR - Auditory Steady-State Response*)

ocorre durante a apresentação de um som a uma taxa suficientemente elevada, tal como o tom senoidal modulado em amplitude (tom AM). A ASSR evocada por esse tipo de tom caracteriza-se pelo aumento de energia na frequência de modulação (e seus harmônicos) no espectro de potência do sinal registrado. Em humanos, os tons AM mais utilizados possuem frequência em torno de 40 Hz ou entre 80 e 110 Hz [12].

2.2 Atensão Seletiva Auditiva

A atenção seletiva auditiva corresponde à capacidade humana de focar sua atenção em determinado estímulo de interesse, em detrimento de estímulos distratores. Esse fato pode ser explicado pela limitação de processamento do mecanismo de atenção, segundo o qual apenas uma tarefa pode ser executada de cada vez e a capacidade de processamento em paralelo é limitada a partir de certo ponto do sistema nervoso [13].

Diversos estudos relatam a influência da atenção seletiva auditiva nas ASSR's, como exemplo [12]:

- aumento da amplitude das ASSR's do estímulo focado, além da redução da resposta ao estímulo ignorado;
- aumento da resposta nos eletrodos frontocentrais ao focar a atenção no estímulo de 40 Hz; nenhuma alteração significativa ao utilizar frequência de 20 Hz;
- lateralização das ASSR's (aumento das respostas no hemisfério direito) e especificidade em frequência para as respostas evocadas na condição de atenção.

Esses resultados indicam que a amplitude da ASSR é apropriada para a classificação da atenção seletiva auditiva do indivíduo, desde que os parâmetros dos estímulos sejam selecionados adequadamente.

2.3 Interface Cérebro-Computador

Interface cérebro-computador (*BCI - Brain-Computer Interface*) é uma forma de comunicação direta entre o cérebro e dispositivos externos, os quais recebem mensagens e comandos do usuário. Seu objetivo é permitir que pessoas com severas desordens neuromusculares possam expressar seus desejos para a pessoa que o acompanha ou até mesmo

utilizar recursos computacionais básicos e comandar neuropróteses de uma forma não-muscular. Esse sistema representa uma importante opção de comunicação e controle para pessoas com deficiências motoras, além de poder ser utilizado como uma forma adicional de controle para pessoas sem deficiências em circunstâncias especiais [14]. A Figura 1 representa a estrutura básica de uma ICC, onde os sinais cerebrais são obtidos através de eletrodos conectados no escalpo e processado para extrair as características do sinal que representam a intenção do usuário, as quais são transformadas em comandos para a operação do dispositivo externo.

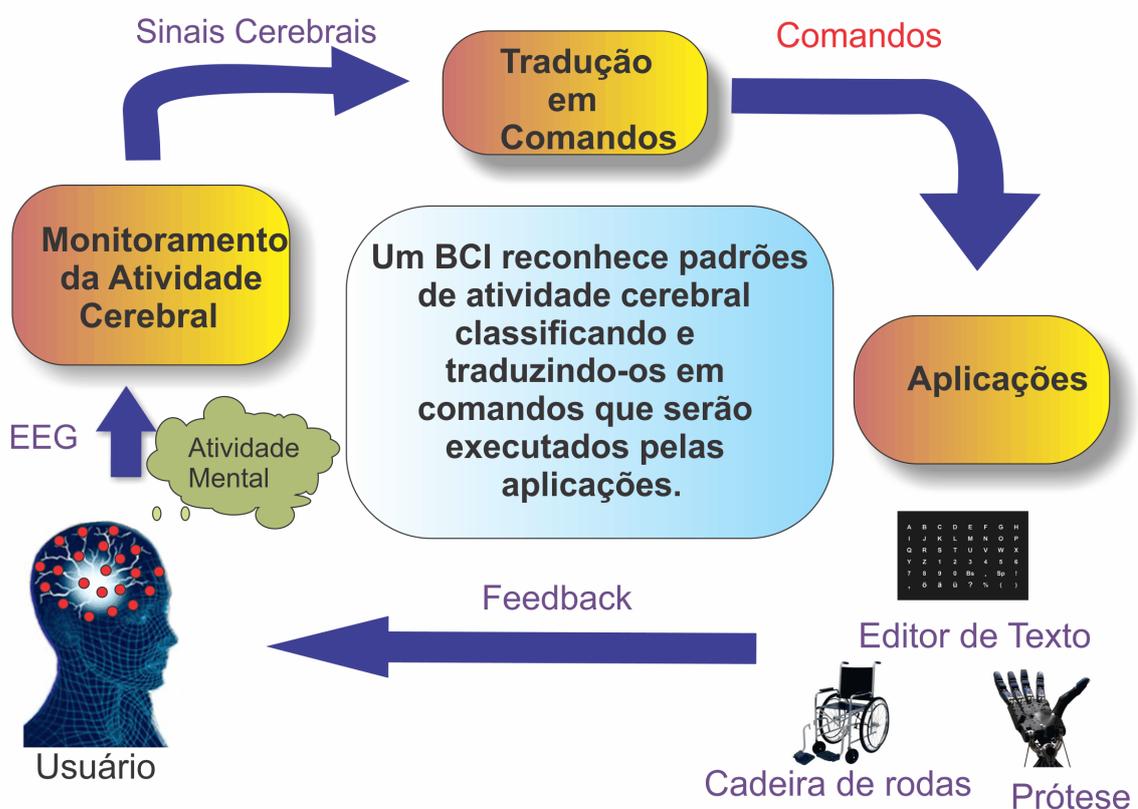


Figura 1: Design e operação básica de uma ICC (adaptada de [1]).

Uma vez que os pacientes com distúrbios neurológicos em geral não apresentam limitação em suas habilidades cognitivas, torna-se possível utilizar sua atividade cerebral como forma de comunicação através de ICC [4]. A obtenção dos sinais cerebrais pode ser realizada através de eletroencefalografia, magnetoencefalografia (MEG), tomografia por emissão de positrons (PET), ressonância magnética funcional (fMRI) e eletrocorticografia (ECoG). A maioria das ICCs desenvolvidas são baseadas nos sinais de EEG, pois este representa um método não-invasivo, prático e de menor custo, o que favorece a sua comercialização em grande escala [15].

O grande desafio das ICC's é obter corretamente o sinal originado a partir do estímulo externo, uma vez que sua amplitude pode ser dezenas de vezes menor do que a amplitude gerada pela atividade elétrica cerebral espontânea, além dos diversos ruídos adicionados durante a coleta dos potenciais. Outra limitação é a baixa taxa de transferência de informação; entretanto valores comumente observados de 10-25 comandos por minuto ainda são suficientes para simples tarefas de comunicação.

2.4 Taxa de Transferência de Informação

A Taxa de Transferência de Informação (TTI ou *ITR - Information Transfer Rate*) é uma forma de avaliar uma ICC, representando a quantidade de informação transmitida/recebida pela interface por unidade de tempo. Seu valor é uma estimativa matemática e é dado por

$$TTI = s \cdot \left[\log_2(N) + P \cdot \log_2(P) + (1 - P) \cdot \log_2\left(\frac{1 - P}{N - 1}\right) \right], \quad (2.1)$$

onde N corresponde ao número de comandos (nesse caso 2: foco da atenção na esquerda ou direita), s à quantidade de comandos executados por minuto e P à probabilidade de acerto [16].

Mantendo o número de comandos fixo, o valor da TTI é diretamente proporcional à taxa de acerto P e ao número de comandos por minuto s . O valor de s pode ser aumentado ao se utilizar menos tempo para a classificação; entretanto, isso em geral diminui a taxa de acerto da interface. Dessa forma, é necessário analisar a relação entre essas duas variáveis para obter as condições ótimas de operação.

2.5 Processamento de Sinais

2.5.1 Rejeição de artefatos

Sinais coletados por eletroencefalografia geralmente estão contaminado por artefatos provenientes da movimentação dos olhos, piscamento palpebral, movimentação muscular e linha de base. A amplitude de tais artefatos, particularmente dos movimentos oculares, apresentam elevada amplitude (superior à do sinal de interesse). Mesmo que eles estejam distribuídos dispersamente pelo sinal, seus valores influenciam os potenciais evocados médios e afetam os resultados do experimento [17]. Dessa forma, rejeitar esses artefatos

é um ponto crucial na análise da atividade cerebral.

Existem diversos métodos baseados em inspeção visual ou técnicas estatísticas para realizar a rejeição de artefatos do sinal coletado. O método aqui utilizado é o do desvio padrão, que visa reduzir os artefatos de alta amplitude e pouco frequentes (induzidos por movimentos musculares). Esse método é adaptativo à característica de cada sinal e menos subjetivo em relação às técnicas de inspeção visual [18].

O método consiste em calcular o desvio padrão de um segmento de referência de aproximadamente 20s de duração que apresente toda a dinâmica do sinal e seja isento de artefatos. O limiar para rejeição é definido em função desse desvio padrão (geralmente 3 vezes seu valor, o que gera um intervalo de confiança de 99,73% para uma distribuição normal).

O sinal original é dividido em vários segmentos de curta duração para análise; o trecho em questão será rejeitado caso contenha mais de 5% de amostras contínuas ou mais de 10% de quaisquer amostras fora do intervalo de confiança estabelecido. Ao eliminar o trecho de um sinal proveniente de algum eletrodo, os dados referentes a esse intervalo de tempo também será rejeitado de todos os outros eletrodos, de forma a manter o mesmo número de trechos em todas as variáveis do banco de dados.

2.5.2 Análise de Componentes Principais

Em muitas situações envolvendo análise de sinais, é comum encontrar conjunto de dados de elevada dimensão em que determinado grupo de variáveis apresenta comportamento relacionado entre si. Isso permite substituir esse grupo por um conjunto menor, eliminando a redundância e diminuindo a dimensão dos dados analisados.

A Análise de Componentes Principais (*PCA - Principal Component Analysis*) é um método quantitativamente rigoroso que possibilita atingir essa simplificação, sendo um dos mais empregados para descrição de padrões de bancos de dados multivariáveis. Seus objetivos são extrair as informações mais importantes do conjunto de dados; reduzir a dimensão dos dados, eliminando a correlação entre variáveis; simplificar a descrição do banco de dados; analisar a estrutura das observações e variáveis.

Para atingir esses objetivos, o método obtém novas variáveis denominadas *componentes principais* que representam combinações lineares das variáveis originais. A primeira componente obrigatoriamente possui a maior variância, e por isso ela é a mais representativa das propriedades dos dados iniciais. As componentes seguintes devem ser ortogonais

a todas as componentes já computadas anteriormente, além de possuir a maior variância possível. O valor de cada uma das novas variáveis do banco de dados pode ser interpretado geometricamente como as projeções das observações sobre as componentes principais [19]. Esse método tem a propriedade de reter as informações importantes nas primeiras componentes, enquanto que as variações associadas a erro experimental, imprecisão de medição e aproximações estão contidas nas componentes finais.

O número de componentes principais é igual ou menor ao número de variáveis originais. Entretanto, decidir se uma componente representa ou não variação representativa dos dados é incerto em muitas situações. Dessa forma, a determinação do número de componentes principais é um aspecto central da análise, uma vez que reter um número inapropriado de componentes resulta em perda de informação ou inclusão de ruído, causando distorção no posterior reconhecimento de padrões [20].

Existem muitos métodos designados ao cálculo do número de componentes apropriado, sendo sua escolha muito subjetiva. Por motivos empíricos, alguns desses métodos são mais utilizados na literatura por serem intuitivamente razoáveis e apresentarem bons resultados na prática [21]. O critério mais óbvio é o “Percentual Acumulado da Variância Total”, em que escolhe-se o total de variância que as componentes principais selecionadas devem apresentar em relação aos dados originais; o número de componentes a serem utilizadas corresponde ao menor valor de variáveis que exceda o limiar adotado.

Outra técnica comumente utilizada é o gráfico *scree plot*; esse gráfico é constituído pela ordem da componente principal no eixo horizontal e o respectivo autovalor no eixo vertical (o autovalor é representativo da fração da variância explicada por cada componente). Como cada componente obtida pela PCA sempre apresenta variância menor em relação a todas às componentes anteriores, esse gráfico é monotonicamente decrescente. A última componente a ser retida corresponde ao ponto em que o gráfico apresenta uma mudança brusca de seu coeficiente angular. Analogamente, a retenção de novas componentes principais pode ser interrompida quando o gráfico se tornar aproximadamente linear (não necessariamente horizontal); caso haja mais de um trecho linear, a parada ocorre no ponto com maior autovalor.

2.5.3 Modelo Autorregressivo

Modelagem paramétrica é uma técnica utilizada para análise temporal em que um modelo matemático é obtido para representar o sinal original. O objetivo é obter um modelo que aproxime bem o comportamento observado no sinal, de forma que ele possa

ser utilizado em alguma aplicação, como nos problemas de classificação [22].

O modelo matemático mais utilizado é uma função de transferência racional com parâmetros livres ajustados de acordo com o sinal a ser representado. Caso todos esse parâmetros estejam no denominador da função de transferência, o modelo é denominado autorregressivo (AR). O AR corresponde a um modelo estocástico útil na representação de séries temporais, sendo um filtro com entrada ruído branco de média 0 e desvio padrão 1 [23]. O valor atual do processo é estimado a partir da soma ponderada dos p (ordem do modelo) valores passados, dado por

$$y(k) = a_1y(k-1) + a_2y(k-2) + \dots + a_p y(k-p) + a_t, \quad (2.2)$$

onde os coeficiente a_1, a_2, \dots, a_p representam os pesos dos valores anteriores no cálculo do valor atual.

Dados provenientes do eletroencefalograma representam uma série temporal e podem ser aproximados por um modelo AR de ordem finita [24]. A escolha apropriada da ordem é de central importância na obtenção do modelo AR e, em geral, esse valor não é conhecido *a priori*. É recomendado adotar a menor ordem que represente o sinal adequadamente, de forma a reduzir o custo computacional na fase de classificação. Dessa forma, a ordem do modelo AR é um dos parâmetros a ser ajustado durante o treinamento e validação dos métodos classificatórios.

2.6 Métodos de Classificação

Essa seção apresenta uma breve descrição dos métodos de classificação utilizados no presente trabalho, priorizando as características de interesse nesse contexto. Ao adotar alguma interpretação geométrica do método em questão, será considerado o caso bidimensional para simplificar a representação gráfica.

2.6.1 Análise de Discriminante

A Análise de Discriminante é um classificador baseado na criação de hiperplanos para separar os dados representantes de cada classe [25]. Para o problema de classificação binária, a classe do ponto de interesse é determinada pelo lado do hiperplano (definido por (2.3), em que x é o vetor de entrada e w e w_0 são determinados por treinamento) em que esse ponto se encontra; mais especificamente, deve-se considerar o sinal de $f(x)$,

conforme representado na Figura 2 para um discriminante linear.

$$x : f(x) = w_0 + w^T x = 0 \quad (2.3)$$

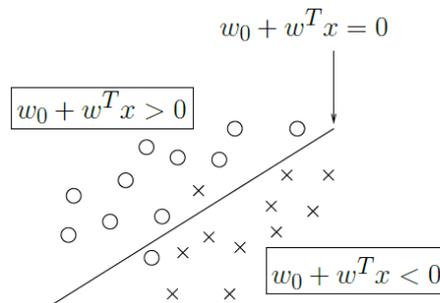


Figura 2: Hiperplano de separação de classes da Análise de Discriminante Linear.

A Análise de Discriminante Linear, também conhecida como Fisher's LDA (*Linear Discriminant Analysis*) assume que os dados possuem distribuição normal, com a mesma matriz de covariância para as duas classes. O hiperplano de separação é obtido pela projeção que maximiza a distância entre as médias das duas classes e reduz a variância entre as classes [7]. A classe atribuída aos novos dados corresponde àquela que apresentar o menor custo de classificação incorreta. Outra possibilidade é o discriminante quadrático, onde a matriz de covariância varia para cada classe.

Esse método tem sido amplamente utilizado em Interfaces Cérebro-Computador devido ao baixo custo computacional. Além disso, é um classificador simples e que geralmente fornece bons resultados.

2.6.2 Máquina de Vetores de Suporte

Máquina de Vetores de Suporte, geralmente referida como SVM (*Support Vector Machines*), é o classificador linear binário mais popular em aprendizado supervisionado devido à sua alta precisão e capacidade de analisar dados de elevada dimensão. O método consiste em obter um hiperplano como a fronteira de decisão, tal que a margem de separação seja máxima, isto é, ocorra a maior distância possível entre o hiperplano e os pontos de treinamento. A obtenção da margem de separação máxima garante que os erros de generalização do classificador serão reduzidos, assumindo que os novos pontos a serem classificados apresentem a mesma distribuição dos pontos de treinamento [26].

A Figura 3 ilustra um problema de classificação binário bidimensional. O diagrama da esquerda apresenta três possíveis fronteiras de decisão; apesar de classificar todos os pontos

corretamente, alguns exemplos estão muito próximos da linha de separação, aumentando a chance de novos pontos serem classificados incorretamente. Por outro lado, a figura da direita representa a linha de separação com margem máxima (distância entre as linhas pontilhadas), situação na qual o erro de generalização do classificador é minimizado. Os pontos mais próximos do hiperplano de separação são denominados *vetores de suporte* (pontos representados por círculos maiores na figura), sendo os pontos mais informativos para o processo de classificação [25].

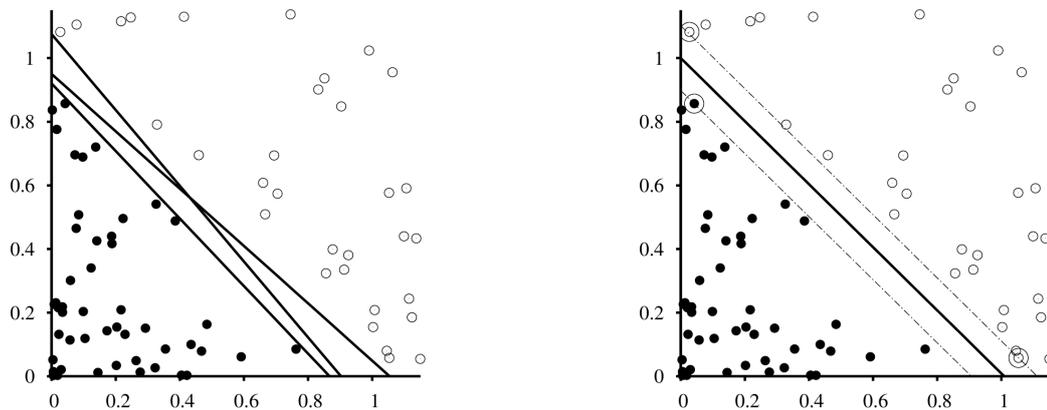


Figura 3: Classificação binária bidimensional através de Máquina de Vetores de Suporte.

O hiperplano de separação é definido por

$$f(x) = w \cdot x + b = 0, \quad (2.4)$$

onde x é o vetor de entrada, w o vetor peso e b o *bias*; a classificação de cada novo ponto é feita a partir do sinal da função $f(x)$. A obtenção do vetor peso é realizada através de uma expressão convexa e seu mínimo global pode ser encontrado eficientemente. Além disso, os pesos associados a cada ponto de treinamento é zero, com exceção dos vetores de suporte; o número de vetores de suporte usualmente é muito menor que o número total de pontos de treinamento, representando uma vantagem do SVM em relação a outros métodos paramétricos [27].

Outra vantagem do classificador SVM é a possibilidade de analisar facilmente dados não-linearmente separáveis. A escolha apropriada de um mapeamento não-linear para uma dimensão suficientemente elevada sempre permite separar duas classes com um hiperplano (de forma geral, N pontos sempre serão separáveis em espaços de dimensão $N - 1$ ou maior [26]). Essa transformação é realizada através de uma função *kernel*, sendo mais utilizadas as polinomiais e a RBF (*Radial Basis Function* [28]). O processo de mapeamento está

ilustrado na Figura 4. A primeira figura apresenta uma fronteira de decisão não-linear definida por $x_1^2 + x_2^2 \leq 1$; ao utilizar o mapeamento definido por $f_1 = x_1^2$, $f_2 = x_2^2$ e $f_3 = \sqrt{2}x_1x_2$, os dados originais são transformados para o espaço tridimensional, onde são claramente separáveis por um hiperplano.

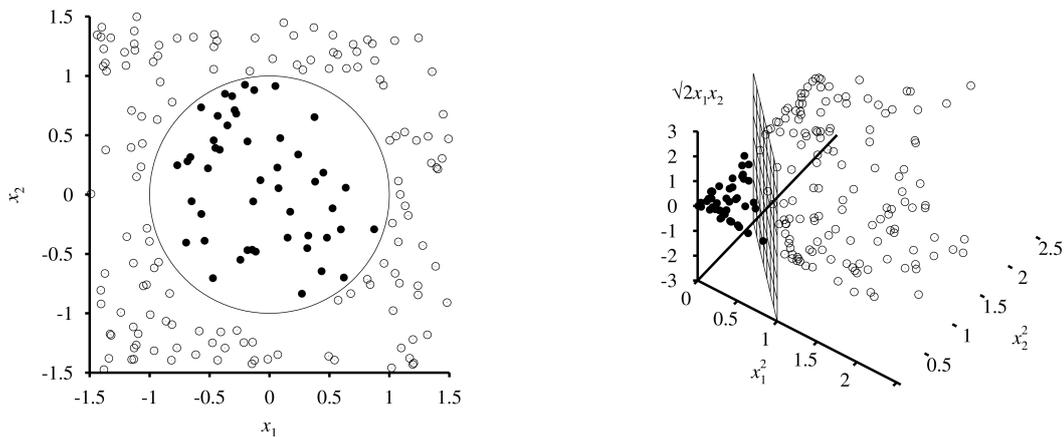


Figura 4: Mapeamento de pontos para dimensão superior.

Por fim, a complexidade do classificador SVM é dependente do número de vetores de suporte em vez da dimensão do espaço transformado. Assim, ele tende a ser menos propenso a problemas de sobre-ajuste (*overfitting*) em relação a outros métodos.

2.6.3 Redes Neurais Artificiais

Rede Neural Artificial (RNA) é um classificador cujo modelo computacional é inspirado no sistema nervoso central humano e animal, porém de complexidade significativamente menor. Esses modelos são capazes de aprendizado e reconhecimento de padrões, sendo melhor ajustáveis a funções não-lineares. Sua estrutura é composta por entradas, saídas e camadas ocultas, onde essas últimas são formadas por neurônios interconectados com todos os outros neurônios das camadas adjacentes para a troca de informações [7]. Os neurônios representam a unidade processadora da rede neural, sendo capazes de realizar funções coletivamente e em paralelo, ao invés de existir uma delimitação de subtarefas para cada unidade [29]. A Figura 5 representa uma estrutura simples de rede neural com apenas 1 camada oculta contendo 3 neurônios.

Cada conexão entrada-neurônio e neurônio-neurônio é associada a um peso, cujo valor é adaptativo e representa os parâmetros numéricos que são ajustados através do algoritmo de aprendizado (há uma grande variedade de algoritmos para realizar essa tarefa). Em geral, os pesos da rede neural são determinados de forma incremental à medida que novos

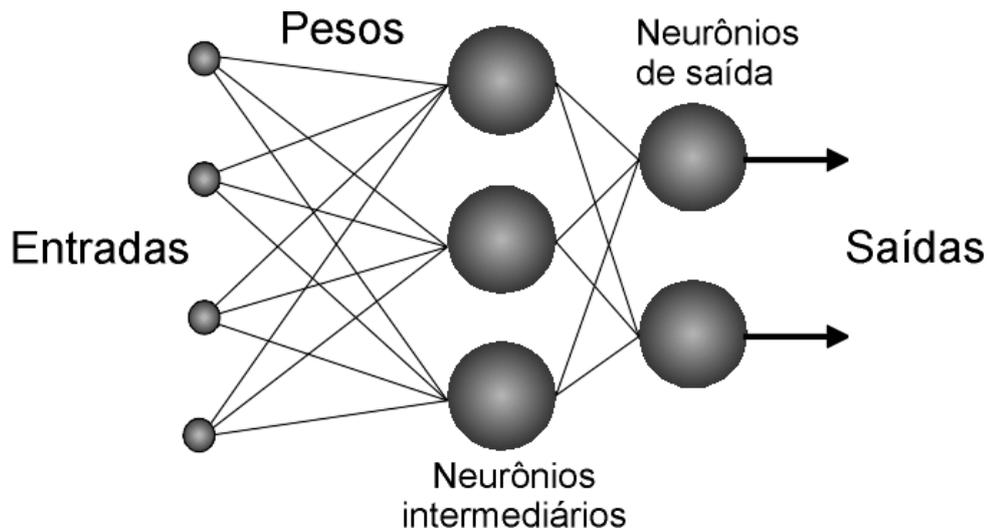


Figura 5: Estrutura de uma Rede Neural Artificial.

valores entrada-saída estão disponíveis; isso diminui o esforço computacional, uma vez que não é necessário a presença de todo o conjunto de treinamento ao mesmo tempo [30].

Após a ponderação de todas as entradas pelos respectivos pesos, um combinador linear agrega todos esses valores para produzir um potencial de ativação. Outro parâmetro da rede é o limiar de ativação, que representa um patamar de comparação para gerar o valor de saída da rede; o potencial produzido será excitatório ou inibitório, em função do valor da diferença entre a saída do combinador linear e o limiar de ativação ser positivo ou negativo, respectivamente.

Para garantir que a saída de cada neurônio seja limitada, o potencial de ativação é aplicado a uma função de ativação, sendo que o uso de funções não-lineares possibilita o uso desse classificador para dados complexos. A mais utilizada nesse processo é a sigmóide, por apresentar um equilíbrio entre os comportamentos lineares e não-lineares [31].

2.6.4 *k-Nearest Neighbors*

O método de classificação *k-Nearest Neighbors* (*k-NN*) é relativamente simples, consistindo em classificar o ponto de interesse de acordo com a classe dominante entre os k pontos mais próximos contidos no conjunto de treinamento. O processo de treinamento consiste em armazenar os vetores multidimensionais de entrada e a respectiva classe a qual pertence. A Figura 6 [32] ilustra o algoritmo para o caso bidimensional; a figura da esquerda contém um conjunto formado por duas classes e um novo ponto a ser classificado. Para $k = 1$, o algoritmo o classifica como sendo da classe positiva, enquanto que para $k = 5$, a classificação é negativa. A figura da direita, contendo 5, classes distintas é

denominada Diagrama de Voronoi e ilustra a superfície de decisão criada ao adotar $k = 1$, onde cada polígono convexo representa a região mais próxima do ponto de treinamento.

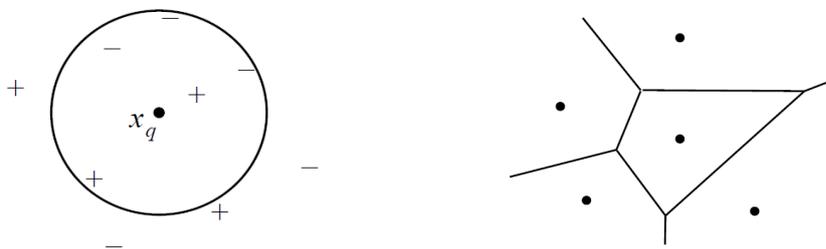


Figura 6: Representação do método k -NN bidimensional (2 e 5 classes, respectivamente).

A melhor escolha do valor de k é realizada de forma heurística, sendo que valores grandes reduzem o efeito de ruídos na classificação, apesar de tornar as fronteiras entre as classes menos distintas. Além disso, em classificações binárias, é comum adotar um valor ímpar para evitar a ocorrência de empate na contagem dos vizinhos mais próximos.

A determinação dos pontos mais próximos é dependente da métrica utilizada. Tipicamente, as distâncias são medidas através da métrica de Minkowski, definida por

$$L_p(x_j, x_q) = \sqrt[p]{\sum_i |x_{j,i} - x_{q,i}|^p}, \quad (2.5)$$

onde x_q e x_j são os pontos considerados. Com $p = 2$, L_p é equivalente à distância euclidiana (geralmente utilizada se as propriedades medidas são similares); para $p = 1$, L_p é a métrica de Manhattan [26].

A acurácia do método pode ser melhorada atribuindo pesos aos pontos de treinamento, priorizando aqueles mais próximos ao ponto a ser classificado (por exemplo, o peso pode ser definido sendo o inverso da distância). Nesse caso, pode-se permitir que todos os pontos do conjunto de treinamento sejam considerados na classificação, uma vez que os mais distantes terão pouca influência; entretanto, isso reduz a velocidade do algoritmo.

Em geral, esse método de classificação gera bons resultados em espaços de baixa dimensão e muitos dados e seu treinamento é robusto contra a influência de ruídos. Além disso, a atribuição de pesos para cada ponto ameniza a influência de dados ruidosos isolados.

2.6.5 Combinação de Classificadores

Atualmente há uma tendência de utilizar combinação de diversos classificadores em Interfaces Cérebro-Computador para melhorar sua performance. Uma das propostas é considerar como classe final a que for escolhida majoritariamente pelos classificadores individuais (*Sistema de Votação*). Esse tipo de combinação é uma alternativa para diminuir a variância do classificador, que influencia a taxa de erro de classificação e reflete a sua sensibilidade ao conjunto de treinamento utilizado [7].

Mesmo em situações em que os classificadores não são independentes, uma pequena diferença entre suas hipóteses é suficiente para reduzir a correlação entre seus erros e, conseqüentemente, tornar a análise conjunta muito útil [26]. De forma geral, essa técnica fornece melhores resultados do que os classificadores atuando individualmente, uma vez que suas regras de decisão são diferentes e fornecem informações complementares.

3 *Materiais e Métodos*

Esse capítulo aborda as etapas executadas ao longo do desenvolvimento desse trabalho, conforme a visão geral apresentada na Figura 7. Os dados coletados pelo EEG são processados para atenuação da interferência de ruídos e artefatos e redução da dimensão. A próxima fase é a obtenção dos parâmetros ótimos de cada algoritmo, sendo posteriormente utilizados na classificação de novos dados.



Figura 7: Fluxo de processos para o reconhecimento de padrões auditivos.

As seções seguintes descrevem detalhadamente cada uma dessas etapas.

3.1 **Protocolo Experimental**

Os dados utilizados correspondem aos sinais de EEG coletados de 14 indivíduos masculinos (totalizando 28 estimulações) com idade compreendida entre 19 e 28 anos e sem patologias neurológicas aparentes. Para verificar a capacidade auditiva de cada voluntário, e por conseguinte a validade de sua participação no estudo, realizou-se os testes de meatoscopia e audiometria convencional (250 a 8000 HZ) na Divisão de Saúde da UFV pela fonoaudióloga Lenilda Viana, sendo todos eles classificados como *limiares auditivos dentro dos padrões de normalidade bilateralmente* [12].

As coletas dos sinais de EEG foram realizadas no laboratório do Núcleo Interdisciplinar de Análise de Sinais (NIAS) do Departamento de Engenharia Elétrica da UFV. O local é equipado com sistema de refrigeração, cabine audiométrica e cadeira reclinável.

O protocolo experimental foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos da UFV com o projeto intitulado “*Investigação da atenção seletiva auditiva para interface cérebro-computador*”.

3.2 Estimulação

Os estímulos auditivos consistem de um tom senoidal puro modulado em amplitude (tom AM) criado pela multiplicação de duas ondas senoidais (a de maior frequência corresponde à portadora e a menor é a envoltória modulante), dado por

$$s(t) = A \cdot \frac{1}{1 + \lambda} \cdot \text{sen}(2\pi f_c t) \cdot (1 + \lambda \cdot \text{sen}(2\pi f_m t)), \quad (3.1)$$

onde A é a amplitude do sinal, λ é a profundidade de modulação (entre 0 e 1), f_c é a frequência da portadora e f_m é a frequência da moduladora [33]. A amplitude do sinal foi ajustada de forma que a intensidade sonora de cada tom AM seja 75 dB SPL e utilizou-se profundidade de modulação de 100% (por gerar ASSR de maior amplitude).

As frequências de modulação foram modificadas para reduzir o espalhamento espectral, sendo seus valores ajustados de forma que ocorra um número primo de oscilações em cada trecho do sinal de comprimento fixo [34]. Dado o sinal a ser detectado, com período T_0 em cada ciclo, o número de oscilações (L) deve ser inteiro e o comprimento da janela ($L \cdot T_0$) deve ser múltiplo inteiro do período de amostragem (T_s), isto é

$$\frac{f_s}{f_m} = \frac{N}{L}, \quad (3.2)$$

onde N é o número de pontos de cada janela e f_s e f_m são as frequências de amostragem e da moduladora, respectivamente.

Utilizou-se 1024 pontos por janela e frequência de amostragem de 601,5 Hz. As frequências da moduladora foram escolhidas arbitrariamente na faixa 30-50 Hz. As portadoras para os estímulos da esquerda e direita foram 500 e 2000 Hz, respectivamente. A Tabela 1 representa as frequências de estimulação, além de conter as frequências corrigidas.

A estimulação em cada ouvido teve duração de 2 minutos, com intervalo de 30 segundos entre cada uma. De forma a instruir o voluntário a focar sua atenção e facilitar a tarefa cognitiva, o estímulo de interesse foi executado por 10 segundos em apenas uma orelha. Posteriormente, os dois sinais são apresentados simultaneamente. A evolução

Tabela 1: Parâmetros de estimulação do tom AM.

	Moduladora original (Hz)	Moduladora corrigida (HZ)	Portadora (Hz)
Ouvido esquerdo	32	31,132	500
Ouvido direito	38	39,356	2000

temporal desse processo está ilustrada na Figura 8.

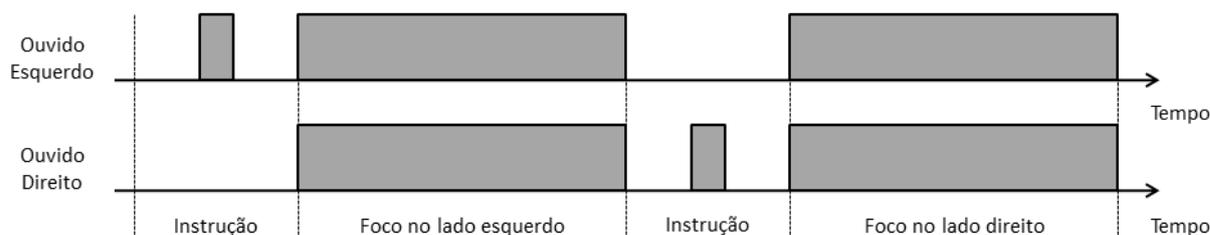


Figura 8: Linha temporal da estimulação em cada um dos ouvidos.

Os estímulos sonoros foram apresentados ao voluntário através do fone de inserção E-A-RTONE 5A da *Aearo Technologies*, que é introduzido no canal auditivo por meio de um tubo plástico flexível, com plugue de terminação macia (capaz de atenuar mais de 30 dB de ruídos externos). Outras vantagens são a maior atenuação transcraniana (evita a estimulação contralateral) e o melhor acoplamento auricular (minimização das perdas de energia acústica) [35].

Todo o experimento foi conduzido no interior de uma cabine audiométrica, com isolamento acústico para evitar a interferência de ruídos ambientes no procedimento. Além disso, os voluntários foram orientados a permanecer de olhos fechados durante a estimulação para facilitar a concentração e reduzir a presença de artefatos oculares.

3.3 Aquisição de Sinais

A coleta dos dados foi realizada através do amplificador de sinais biológicos *BrainNet BNT-36*, com filtro passa-altas e passa-baixas de 0,1 e 100 Hz, respectivamente, um filtro notch de 60 Hz e frequência de amostragem de 601,5 Hz.

Os sinais do EEG foram registrados a partir de eletrodos de prata não-invasivos, com diâmetro de 10 mm e impedância inferior a 5 k Ω . Os eletrodos foram posicionados no escalpo do voluntário de acordo com o Sistema Internacional 10-20, com referência biauricular, conforme representado na Figura 9 [36].

Total e gráfico *scree plot*. Para o primeiro método, adotou-se um limiar acumulado de 95%, sendo necessário pelo menos 5 componentes principais para que os dados de todos os indivíduos exceda esse valor. Em relação ao *scree plot*, também foram necessárias 5 componentes para que todos os gráficos se tornassem aproximadamente lineares. Dessa forma, esse trabalho considera no máximo 5 componentes dos dados obtidos com cada estimulação. Os valores detalhados para cada arquivo do banco de dados referentes a esses dois critérios de parada encontram-se no Anexo A.

A última etapa do processamento dos sinais correspondeu à obtenção dos parâmetros do modelo autorregressivo, adotando-se que a ordem máxima desses modelos seria 10. Apesar de ser uma escolha arbitrária, percebeu-se que o aumento da ordem a partir desse valor não melhora a performance dos métodos de reconhecimento e classificação de padrões.

Após a preparação dos dados, realizou-se os processos de treinamento do métodos classificatórios, considerando 10 janelas dos sinais coletados do EEG. Os bancos de dados foram divididos aleatoriamente, sendo 16 utilizados para treinamento e os 12 restantes na validação. Para cada um dos métodos considerados, utilizou-se diversas combinações para o número de componentes principais e ordem do modelo AR, sendo que esses valores variaram de 1 a 5 e 1 a 10, respectivamente. Além disso, cada método possui diversos parâmetros específicos que podem ser modificados para otimizar as taxas de acerto; a lista a seguir contém as configurações que foram analisadas nesse trabalho.

1. Análise de Discriminante

Adotou-se dois tipos de discriminantes: linear e quadrático.

2. Máquina de Vetores de Suporte

O parâmetro variado nesse algoritmo foi a função de mapeamento dos dados, sendo utilizadas as seguintes funções *kernel*: linear, polinomiais de ordem 2, 3, 4 e 5, e *Radial Basis Function*.

3. Rede Neural Artificial

As redes analisadas são formadas por 1 ou 2 camadas ocultas, 1, 2 ou 3 neurônios por camada e função de treinamento *Levenberg-Marquardt* ou Gradiente Conjugado Escalonado.

4. *k-Nearest Neighbors*

Foram utilizadas as métricas euclidiana e de Manhattan, atribuindo pesos a cada ponto (pesos iguais ou inversamente proporcionais à distância). O número de pontos vizinhos considerados foi 1, 3 e 5 para pesos iguais, e de 1 a 6 para as situações com peso inverso. O motivo do uso de valores pares apenas nos casos com pesos do tipo inverso é para evitar possíveis empates no processo de classificação.

Para cada uma das diversas combinações analisadas, calculou-se a taxa de acerto a partir dos dados de validação. Dessa forma, foi possível determinar os melhores parâmetros para cada classificador, sendo essas configurações utilizadas posteriormente na análise conjunta dos diversos métodos classificatórios. Além disso, analisou-se a Taxa de Transferência de Informação obtida por cada método.

4 *Resultados e Discussões*

Esse capítulo trata dos resultados obtidos a partir de cada método de classificação, além de comparar suas performances e analisar seu funcionamento conjunto.

A variação do número de componentes principais e da ordem do modelo AR resulta em 50 treinamentos para cada configuração de parâmetros. As tabelas a seguir apresentam as taxas percentuais máximas de acerto obtidas em cada configuração, além do par de valores formado pelo número de componentes e ordem do modelo AR em que tal acerto foi alcançado. Para o caso da Análise de Discriminante, a utilização de um discriminante quadrático resulta em taxa de acerto maior, conforme Tabela 2.

Tabela 2: Taxas máximas de acerto para Análise de Discriminante (dadas em %).

Tipo de discriminante	
Linear {PCA AR}	Quadrático {PCA AR}
75,0 {5 7}	83,3 {1 5, 1 6, 1 7, 1 8, 3 10}

Como pode ser verificado na Tabela 3, a ordem da função polinomial *kernel* da Máquina de Vetores de Suporte não afeta as taxas de acerto obtidas nesse estudo, sendo por isso que valores superiores a 5 não foram utilizados. Além disso, a função de mapeamento *Radial Basis Function* apresentou o melhor resultado, conforme ocorre frequentemente com esse algoritmo de classificação.

Tabela 3: Taxas máximas de acerto para Máquina de Vetores de Suporte (dadas em %).

Função <i>Kernel</i>	Porcentagem {PCA AR}
Linear	83,3 {1 10, 2 8, 5 8}
Polinomial de ordem 2	83,3 {1 7, 1 8, 2 8, 4 8}
Polinomial de ordem 3	83,3 {1 7, 1 8, 2 8, 3 8, 5 8}
Polinomial de ordem 4	83,3 {1 5, 1 7, 1 8, 1 9, 1 10, 2 7, 5 8}
Polinomial de ordem 5	83,3 {1 7, 1 8, 1 10, 3 9}
<i>Radial Basis Function</i>	91,7 {1 5}

Os resultados da classificação por Rede Neural Artificial encontram-se nas Tabelas 4, 5 e 6. Percebe-se que o treinamento através do Gradiente Conjugado Escalonado (*trainscg*) em geral apresenta maiores taxas de acerto em relação ao algoritmo *Levenberg-Marquardt* (*trainlm*). Outra observação importante é que redes com menos camadas ocultas e poucos neurônios por camada, e conseqüentemente menos complexas, corresponderam às maiores taxas de acerto.

Tabela 4: Taxas máximas de acerto para Rede Neural Artificial com 1 camada oculta (dadas em %).

Neurônios por camada	Função de treinamento	
	<i>trainscg</i> {PCA AR}	<i>trainlm</i> {PCA AR}
1	91,7 {5 8}	83,3 {1 7, 1 8}
2	83,3 {1 7, 1 8}	83,3 {1 8, 2 8, 4 8}
3	83,3 {1 6}	83,3 {1 6, 1 7, 1 8}

Tabela 5: Taxas máximas de acerto para Rede Neural Artificial com 2 camadas ocultas e treinamento por *Levenberg-Marquardt* (dadas em %).

Camada 1	Neurônios da camada 2		
	1 {PCA AR}	2 {PCA AR}	3 {PCA AR}
1	83,3 {1 8, 3 6, 4 8}	83,3 {4 8}	83,3 {3 7}
2	83,3 {1 8, 2 8, 4 8, 5 8}	75,0 {1 3, 1 4, 1 5, 4 5}	75,0 {1 6, 1 8, 1 9, 2 8}
3	83,3 {1 6}	83,3 {1 10}	83,3 {1 5, 1 8, 5 7}

Tabela 6: Taxas máximas de acerto para Rede Neural Artificial com 2 camadas ocultas e treinamento por Gradiente Conjugado Escalonado (dadas em %).

Camada 1	Neurônios da camada 2		
	1 {PCA AR}	2 {PCA AR}	3 {PCA AR}
1	83,3 {4 8}	91,7 {1 6}	83,3 {1 9}
2	83,3 {1 7, 5 6}	83,3 {1 7, 3 9, 5 6, 5 8}	75,0 {1 9, 1 10}
3	83,3 {1 6, 1 8, 3 8}	83,3 {5 8}	83,3 {4 6}

Em relação ao método *k-Nearest Neighbors*, a métrica adotada não influenciou nos resultados, assim como a adoção de pesos iguais ou inversos também geraram taxas de acerto semelhantes, de acordo com as Tabelas 7 e 8. Além disso, a utilização de 3 pontos como vizinhos corresponde à melhor combinação.

Tabela 7: Taxas máximas de acerto para k -Nearest Neighbors com pesos iguais (dadas em %).

Número de vizinhos (k)	Métrica	
	Euclidean ($p = 2$) {PCA AR}	Manhattan ($p = 1$) {PCA AR}
1	83,3 {1 6, 1 7}	83,3 {1 6, 1 7, 2 7}
3	91,7 {1 5}	91,7 {1 5}
5	91,7 {1 6}	91,7 {1 6}

Tabela 8: Taxas máximas de acerto para k -Nearest Neighbors com pesos inversos (dadas em %).

Número de vizinhos (k)	Métrica	
	Euclidean ($p = 2$) {PCA AR}	Manhattan ($p = 1$) {PCA AR}
1	83,3 {1 6, 1 7}	83,3 {1 6, 1 7, 2 7}
2	83,3 {1 6, 1 7}	83,3 {1 6, 1 7, 2 7}
3	91,7 {5 8}	91,7 {4 8}
4	83,3 {1 6, 1 7}	83,3 {1 6, 1 7, 2 7}
5	83,3 {1 6, 1 7, 3 5, 4 5}	83,3 {1 6, 1 7, 2 6, 3 5, 4 5}
6	83,3 {1 6, 1 7, 4 5}	83,3 {1 6, 1 7, 5 6}

Os valores apresentados anteriormente demonstram que os métodos de classificação considerados nesse estudo podem ser ajustados de forma a obter elevadas taxas de acerto e produzir resultados similares entre si. Outra análise que pode ser feita na tentativa de melhorar a performance de classificação é utilizar os vários métodos anteriores em conjunto através do Sistema de Votação. Para realizar tal combinação, selecionou-se a(s) melhor(es) combinação(ões) de cada método individual e obteve-se a respectiva classe da atenção seletiva auditiva. A seguir estão listados os parâmetros para cada classificador que resultaram na taxa de acerto mais elevada.

1. Análise de Discriminante

- Discriminante quadrático, PCA = 3 e AR = 10

2. Máquina de Vetores de Suporte

- Função de mapeamento do tipo *Radial Basis Function*, PCA = 1 e AR = 5

3. Rede Neural Artificial

- Rede composta de 1 camada oculta com 1 neurônio, função de treinamento Gradiente Conjugado Escalonado, PCA = 5 e AR = 8
- Rede composta de 2 camadas ocultas, com 1 neurônio na primeira camada e 2 na segunda, função de treinamento Gradiente Conjugado Escalonado, PCA = 1 e AR = 6

4. *k*-Nearest Neighbors

- Métrica euclidiana, pesos inversamente proporcionais à distância, 3 pontos vizinhos, PCA = 5 e AR = 8
- Métrica de Manhattan, pesos inversamente proporcionais à distância, 3 pontos vizinhos, PCA = 4 e AR = 8

A Tabela 9 representa a classificação obtida por cada método para todos os arquivos do conjunto de validação, sendo que os valores 0 e 1 corresponde à atenção seletiva auditiva na direita e esquerda, respectivamente. A classificação final foi determinada a partir da classe dominante fornecida individualmente por cada um dos métodos considerados. Note que enquanto a taxa máxima de acerto desses métodos foi 91,7%, a utilização em conjunto dos mesmos resultou em 100% de acerto. Uma situação que merece destaque é o empate no número de cada classe, como ocorreu com um dos arquivos analisados; nesse caso, a classificação final foi baseada na soma das taxas de acerto gerais de cada um dos métodos. Mais especificamente, esse problema ocorreu no décimo arquivo de entrada da Tabela 9, sendo que os classificadores que retornaram 1 apresentam taxas de acerto de 83,3, 91,7 e 91,7%, enquanto que essas taxas para os que retornaram 0 são 91,7, 91,7 e 91,7%. Dessa forma, esse arquivo foi classificado como 0. Além disso, percebe-se que a utilização de apenas uma configuração de RNA e *k*-NN gera os mesmos resultados finais. Outro ponto a ser destacado é a possibilidade de obter 100% de acerto utilizando um número menor de métodos; por exemplo, isso pode ser obtido ao se combinar a Máquina de Vetores de Suporte, uma Rede Neural Artificial¹ e um *k*-Nearest Neighbors³, sendo que essa combinação impede a ocorrência de empate no Sistema de Votação por ser um número ímpar de classificadores.

Para melhor entendimento da escolha de 10 janelas para análise, a Figura 10 ilustra as taxas de acerto obtidas por cada uma das configurações de classificadores utilizadas anteriormente ao variar o número de janelas consideradas. Fica evidente que a performance de cada algoritmo é dependente das janelas adotadas, e a situação com 10 janelas

Tabela 9: Análise conjunta das melhores configurações dos classificadores.

Entrada	Algoritmo de classificação						Final
	AD	SVM	RNA ¹	RNA ²	k-NN ³	k-NN ⁴	
1: 1	1	1	1	1	1	1	1
2: 1	1	1	1	1	1	1	1
3: 1	1	1	0	1	1	0	1
4: 1	1	1	1	1	1	1	1
5: 1	1	1	1	1	1	1	1
6: 0	0	0	0	0	0	0	0
7: 0	1	0	0	0	0	0	0
8: 0	0	0	0	0	0	0	0
9: 0	0	0	0	0	0	0	0
10: 0	1	1	0	1	0	0	0
11: 0	0	0	0	0	1	0	0
12: 0	0	0	0	0	0	0	0
Acerto (%)	83,3	91,7	91,7	91,7	91,7	91,7	100

corresponde aos resultados ótimos. Nota-se que uma das redes neurais apresenta taxa de acerto de 100% ao se utilizar apenas 9 janelas, entretanto a performance dos outros classificadores piora nessa situação.

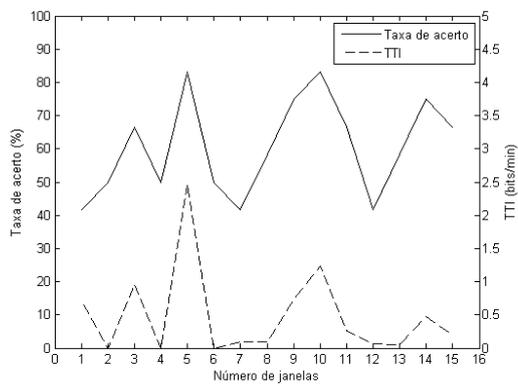
Outro fator importante na análise da aplicabilidade de determinado método em uma Interface Cérebro Computador é a Taxa de Transferência de Informação (TTI). Esses valores também estão representados na Figura 10 para as configurações de classificadores consideradas. Tomando-se como referência as melhores taxas de acerto (10 janelas), a TTI é da ordem de 2 bits por minuto. Observa-se estreita ligação entre esse valor e as taxas de acerto, uma vez que ele é dependente da precisão do classificador; além disso, há uma tendência de redução da TTI com o tempo ocasionada pelo número de comandos executados por minuto. Para uma das configurações de rede neural, a descontinuidade observada no gráfico da Figura 10(d) corresponde ao acerto de 100%, situação na qual a TTI não está definida.

¹Rede composta de 1 camada oculta com 1 neurônio, função de treinamento Gradiente Conjugado Escalonado, PCA = 5 e AR = 8

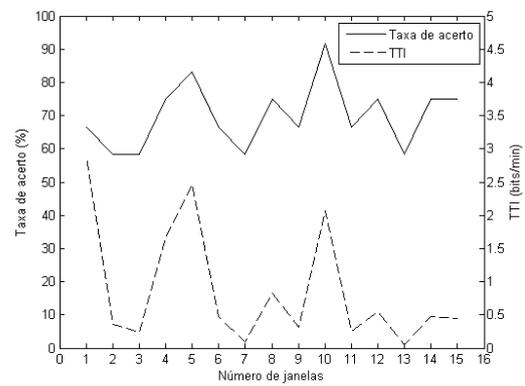
²Rede composta de 2 camadas ocultas, com 1 neurônio na primeira camada e 2 na segunda, função de treinamento Gradiente Conjugado Escalonado, PCA = 1 e AR = 6

³Métrica euclidiana, pesos inversamente proporcionais à distância entre os pontos, 3 pontos vizinhos, PCA = 5 e AR = 8

⁴Métrica de Manhattan, pesos inversamente proporcionais à distância entre os pontos, 3 pontos vizinhos, PCA = 4 e AR = 8



(a) Análise de Discriminante



(b) Máquina de Vetores de Suporte

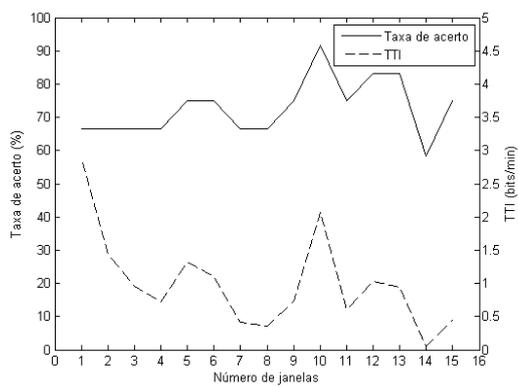
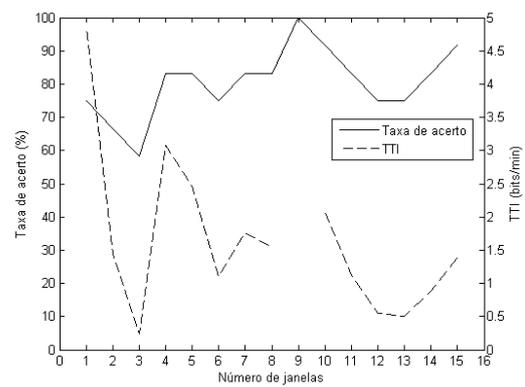
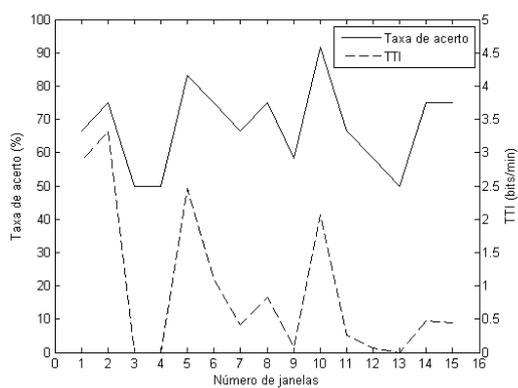
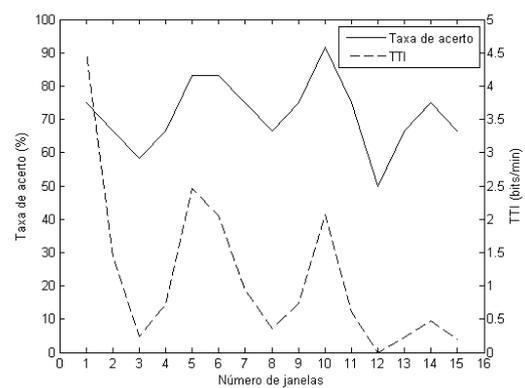
(c) Rede Neural Artificial ¹(d) Rede Neural Artificial ²(e) k -Nearest Neighbors ³(f) k -Nearest Neighbors ⁴

Figura 10: Variação da taxa de acerto e taxa de transferência de informação em função do número de janelas consideradas.

5 *Conclusões*

Os resultados obtidos nesse trabalho permitem destacar alguns pontos referentes à utilização de ASSR's durante atenção seletiva auditiva em Interfaces Cérebro-Computador. Inicialmente, a escolha do classificador a ser utilizado e a sintonização de seus parâmetros é o ponto fundamental para o bom funcionamento da ICC, uma vez que pequenas variações em suas configurações acarretam diferenças significativas nas taxas de acerto obtidas. Apesar de cada um dos classificadores analisados ter apresentado boa performance (acertos acima de 80%), ficou evidente que a utilização conjunta e simultânea dos mesmos é uma forma de aprimorar o resultado final. Além disso, atribuir pesos a cada classificador durante o Sistema de Votação baseado em sua performance geral mostrou-se eficiente na determinação da classe final, especialmente em casos de empate. A capacidade de detecção correta do foco da atenção seletiva auditiva é um indicativo de que as ASSR's são moduladas sob as condições de atenção adotadas nesse estudo.

As Taxas de Transferência de Informação alcançadas da ordem de 2 bits por minuto, apesar de serem relativamente baixas, já são suficientes para que o indivíduo possa realizar tarefas simples, como controle de luminosidade ou utilização de um processador de textos. Entretanto, para que as ICC baseadas na atenção auditiva sejam amplamente utilizadas, é necessário a investigação de formas para aumentar a TTI.

A complexidade e custo computacional dos classificadores corresponde ao processo de treinamento para a obtenção de sua configuração ótima, sendo que a posterior classificação de novos dados é realizada de forma simples e rápida. Isso sugere que as técnicas de classificação podem ser utilizadas em modo *online*, isto é, a classificação é realizada simultaneamente com a coleta de dados do escalpo humano.

Referências

- [1] MOLINA, G. G. *Electroencephalogram processing in Motor Imagery Based BCI: A tutorial*. Amsterdam, November 2008.
- [2] HOFFMANN, U. et al. An efficient p300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, v. 167, p. 115–125, 2008.
- [3] HILL, N. J. et al. An auditory paradigm for brain-computer interfaces. *Advances in Neural Information Processing Systems*, v. 17, 2004.
- [4] KIM, D.-W. et al. Classification of selective attention to auditory stimuli: toward vision-free brain-computer interfacing. *Journal of Neuroscience Methods*, v. 197, p. 180–185, 2011.
- [5] KALLENBERG, M. *Auditory Selective Attention as a method for a Brain Computer Interface*. Music, Mind and Machine Group - Radboud University Nijmegen, 2006.
- [6] LOPEZ, M. A. et al. Evidences of cognitive effects over auditory steady-state responses by means of artificial neural networks and its use in brain-computer interfaces. *Neurocomputing*, v. 72, p. 3617–3623, 2009.
- [7] LOTTE, F. et al. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, v. 4, 2007.
- [8] FERREIRA, A. C. F. *Investigação da atenção seletiva auditiva para interface cérebro-computador usando máquinas de vetor de suporte*. Junho 2011. Departamento de Engenharia Elétrica, Universidade Federal de Viçosa.
- [9] NICÁCIO, L. R. *Investigação da atenção seletiva auditiva para interface cérebro-computador usando magnitude quadrática de coerência*. Janeiro 2011. Departamento de Engenharia Elétrica, Universidade Federal de Viçosa.
- [10] CHIAPPA, K. H. *Evoked Potentials in Clinical Medicine*. Philadelphia: Lippincott Williams & Wilkins, 1997.
- [11] FELIX, L. B. *Detecção Objetiva de Respostas Multivariável aplicada ao Eletroencefalograma e a Potenciais Evocados*. Dissertação (Mestrado) — Engenharia Elétrica, Universidade Federal de Minas Gerais, 2006.
- [12] RANAUDO, F. de S. *Classificação em tempo real da Atenção Seletiva Auditiva usando Potenciais Evocados em Regime Permanente e Coerência Espacial*. Dissertação (Mestrado) — Engenharia Biomédica, Universidade Federal do Rio de Janeiro, 2012.

- [13] GIARD, M.-H. et al. Neurophysiological mechanisms of auditory selective attention in humans. *Frontiers in Bioscience*, v. 5, p. 84–94, 2000.
- [14] WOLPAW, J. R. et al. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, v. 113, p. 767–791, 2002.
- [15] MILLAN, J. R. et al. Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges. *Frontiers in Neuroscience*, v. 4, p. 1–15, 2010.
- [16] OBERMAIER, B. et al. Information transfer rate in a five-classes brain-computer interface. *Neural Systems and Rehabilitation Engineering*, v. 9, p. 283–288, 2001.
- [17] DELORME, A.; MAKEIG, S.; SEJNOWSKI, T. Automatic artifact rejection for EEG data using high-order statistics and Independent Component Analysis. *International Workshop on ICA*, 2001.
- [18] AMABILE, R. A. N. et al. Remoção automática de artefatos em sinais de EEG usando Local SSA. *Congresso Brasileiro de Automática*, 2008.
- [19] ADBI, H.; WILLIAMS, L. J. Principal Component Analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, v. 2, 2010.
- [20] NETO, P. R. P.; JACKSON, D. A.; SOMERS, K. M. How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, v. 49, p. 974–997, 2005.
- [21] JOLLIFE, I. T. *Principal Component Analysis*. 2. ed. New York: Springer, 2002.
- [22] PARDEY, J.; ROBERTS, S.; TARASSENK, L. A review of parametric modelling techniques for EEG analysis. *Medical Engineering & Physics*, v. 18, p. 2–11, 1996.
- [23] AGUIRRE, L. A. *Introdução à Identificação de Sistemas - Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. 3. ed. Belo Horizonte: Editora UFMG, 2007.
- [24] HEFFTNER, G.; ZUCCHINI, W.; JAROS, G. G. The electromyogram (EMG) as a control signal for functional neuromuscular stimulation - autoregressive modeling as a means of EMG signature discrimination. *IEEE Transactions on Biomedical Engineering*, v. 35, p. 230–237, 1988.
- [25] DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. 2. ed. New York: John Wiley & Sons, 2001.
- [26] RUSSELL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3. ed. Upper Saddle River: Prentice Hall, 2010.
- [27] BEN-HUR, A.; WESTON, J. A user’s guide to support vector machines. *Methods in Molecular Biology*, v. 609, p. 223–239, 2010.
- [28] HSU, C.-W.; CHANG, C.-C.; LIN, C.-J. *A Practical Guide to Support Vector Classification*. Department of Computer Science, National Taiwan University, 2010.
- [29] MARQUES, J. P. *Pattern Recognition: Concepts, Methods and Applications*. 1. ed. New York: Springer, 2001.

-
- [30] YEGNANARAYANA, B. *Artificial Neural Networks*. 1. ed. New York: PHI Learning, 2009.
- [31] HAYKIN, S. S. *Neural Networks: A Comprehensive Foundation*. 2. ed. New Jersey: Prentice Hall, 1999.
- [32] MITCHELL, T. M. *Machine Learning*. New York: McGraw-Hill, 1997.
- [33] JOHN, M. S. et al. Multiple auditory steady-state responses: Stimulus and recording parameters. *Audiology*, v. 37, p. 59–82, 1998.
- [34] FELIX, L. B. et al. Avoiding spectral leakage in objective detection of auditory steady-state evoked responses in the inferior colliculus of rat using coherence. *Journal of Neuroscience Methods*, v. 144, p. 249–255, 2005.
- [35] AEARO. *Instructions for the Use of E-A-RTONE*. Aearo Company - Auditory Systems Production, Indianapolis, 1999.
- [36] BCI2000. *User Tutorial: EEG Measurement Setup*. Disponível em http://www.bci2000.org/wiki/index.php/User_Tutorial:EEG_Measurement_Setup. Acessado em 21/05/2014.

ANEXO A – Número de Componentes da PCA

Tabela 10: Variância acumulada em função do número de Componentes Principais

Arquivo	Número de componentes				
	1	2	3	4	5
Esquerda 1	94,34%	98,59%	99,30%	99,55%	99,67%
Esquerda 2	91,98%	99,06%	99,39%	99,55%	99,69%
Esquerda 3	49,18%	90,14%	97,92%	98,73%	99,13%
Esquerda 4	92,55%	94,77%	96,40%	97,71%	98,32%
Esquerda 5	82,09%	91,06%	94,50%	97,30%	98,24%
Esquerda 6	85,50%	94,36%	97,21%	98,80%	99,22%
Esquerda 7	60,12%	80,38%	94,52%	96,09%	97,25%
Esquerda 8	80,35%	99,23%	99,69%	99,79%	99,87%
Esquerda 9	88,76%	98,23%	99,38%	99,69%	99,80%
Esquerda 10	97,98%	99,90%	99,94%	99,96%	99,97%
Esquerda 11	90,13%	97,86%	99,19%	99,53%	99,66%
Esquerda 12	83,74%	98,09%	99,26%	99,57%	99,72%
Esquerda 13	83,14%	98,09%	99,26%	99,57%	99,72%
Esquerda 14	93,08%	99,50%	99,70%	99,86%	99,94%
Direita 1	94,42%	98,16%	98,98%	99,31%	99,50%
Direita 2	92,21%	98,95%	99,50%	99,66%	99,80%
Direita 3	83,45%	96,50%	98,84%	99,19%	99,41%
Direita 4	66,87%	85,02%	90,56%	94,82%	96,40%
Direita 5	85,39%	91,73%	95,08%	97,30%	98,36%
Direita 6	91,12%	97,75%	99,55%	99,74%	99,81%
Direita 7	64,55%	82,50%	95,48%	96,74%	97,77%
Direita 8	79,91%	98,83%	99,61%	99,77%	99,86%
Direita 9	89,39%	98,33%	99,37%	99,76%	99,84%
Direita 10	97,95%	99,85%	99,92%	99,94%	99,96%
Direita 11	92,45%	98,51%	99,08%	99,38%	99,53%
Direita 12	83,42%	99,06%	99,52%	99,71%	99,81%
Direita 13	75,24%	97,96%	99,04%	99,33%	99,57%
Direita 14	93,59%	99,66%	99,83%	99,91%	99,94%

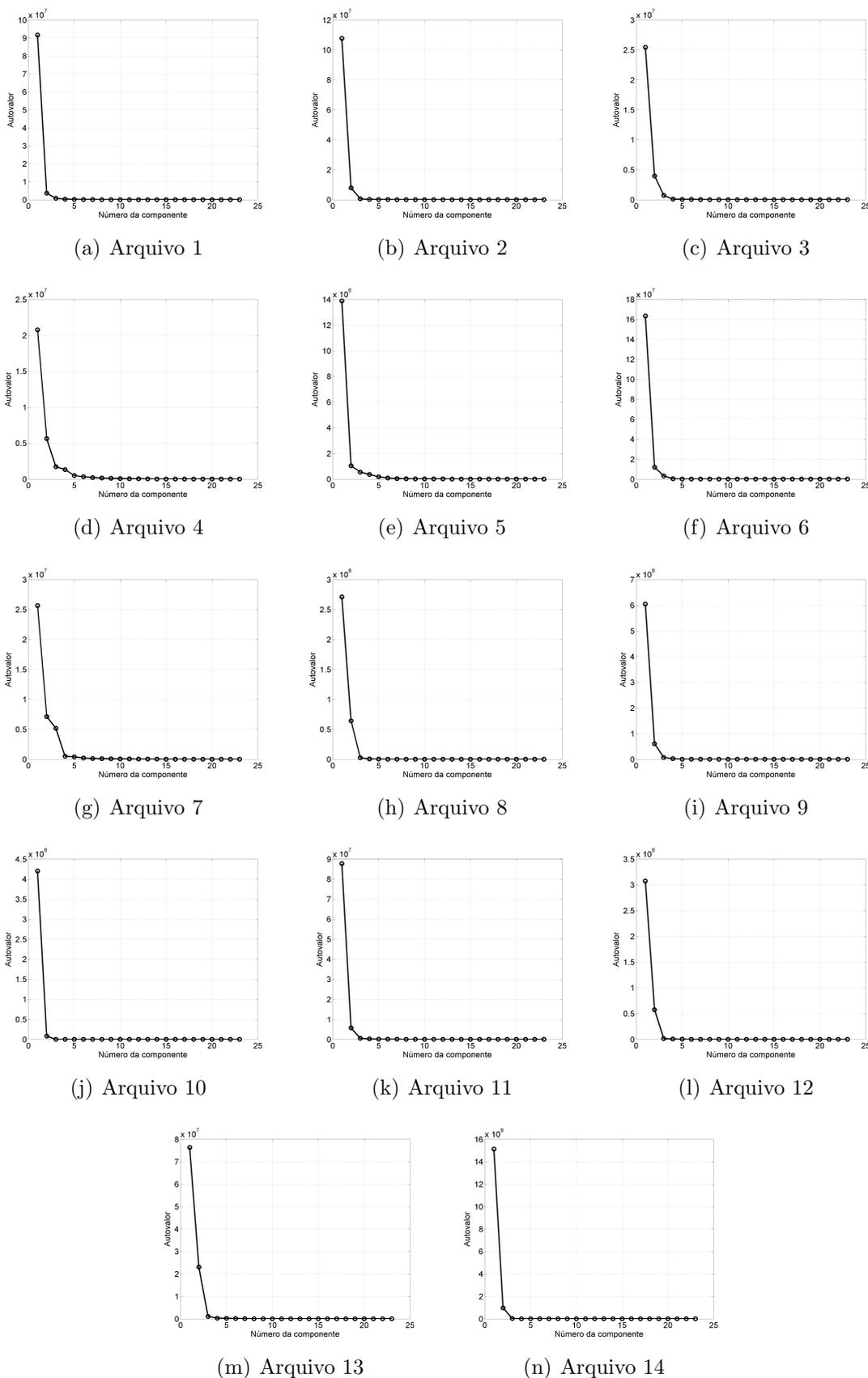
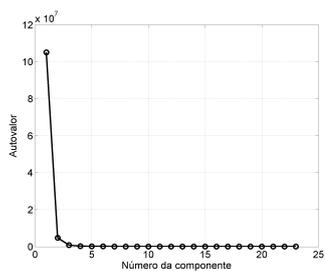
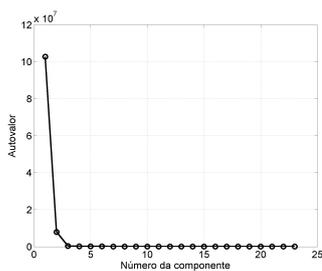


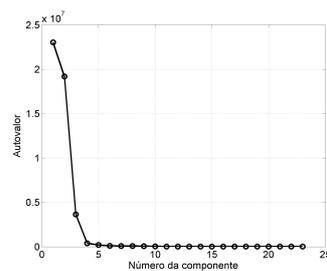
Figura 11: *Scree plot* dos arquivos com estimulação no ouvido esquerdo.



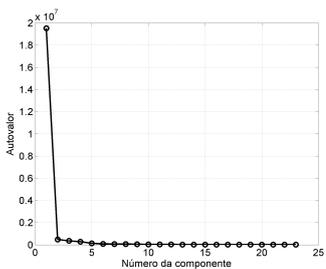
(a) Arquivo 1



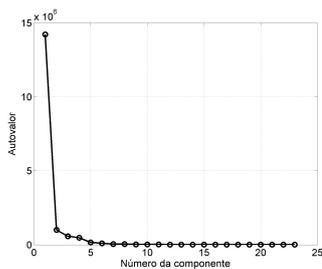
(b) Arquivo 2



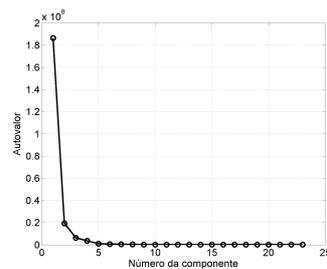
(c) Arquivo 3



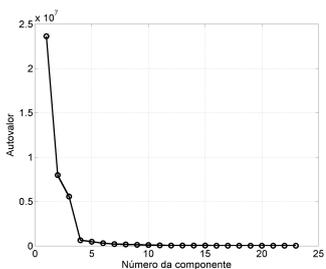
(d) Arquivo 4



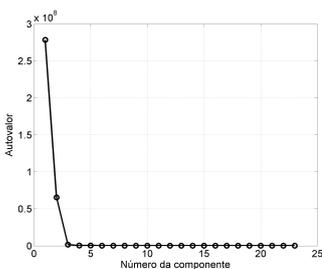
(e) Arquivo 5



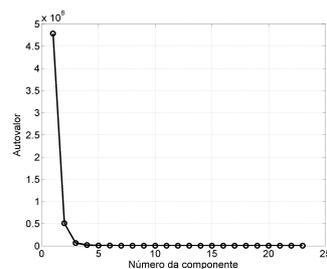
(f) Arquivo 6



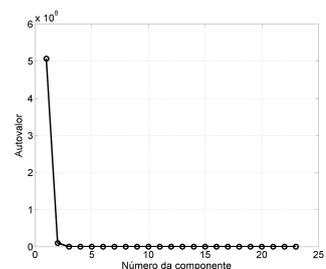
(g) Arquivo 7



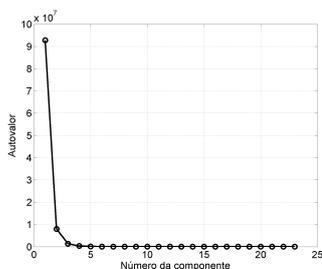
(h) Arquivo 8



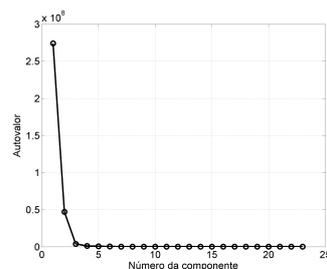
(i) Arquivo 9



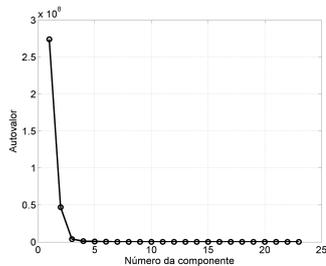
(j) Arquivo 10



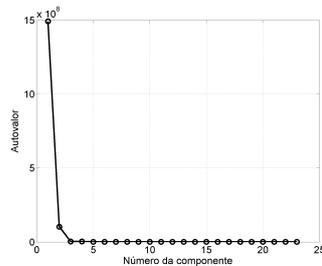
(k) Arquivo 11



(l) Arquivo 12



(m) Arquivo 13



(n) Arquivo 14

Figura 12: *Scree plot* dos arquivos com estimulação no ouvido esquerdo.