

Luiz Henrique Giovanini Marrega

*Avaliação da Relevância de Atributos
Posturográficos Globais e Estruturais para a
Identificação do Acidente Vascular Encefálico
utilizando Modelos de Classificação*

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Informática.

Curitiba

2015

Luiz Henrique Giovanini Marrega

*Avaliação da Relevância de Atributos
Posturográficos Globais e Estruturais para a
Identificação do Acidente Vascular Encefálico
utilizando Modelos de Classificação*

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Informática.

Área de concentração: *Descoberta do Conhecimento e Aprendizagem de Máquina.*

Orientador: Prof. Dr. Julio Cesar Nievola

Coorientador: Profa. Dra. Elisangela Ferretti Manffra

Curitiba

2015

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central

M358a
2015 Marrega, Luiz Henrique Giovanini
Avaliação de relevância de atributos posturográficos globais e estruturais para a identificação do acidente vascular encefálico utilizando modelos de classificação / Luiz Henrique Giovanini Marrega ; orientador, Julio Cesar Nievola ; coorientador, Elisangela Ferretti Manffra. – 2015.
xvii, 131 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2015
Bibliografia: f. 107-126

1. Mineração de dados. 2. Aprendizado do computador. 3. Algoritmo genético. 4. Informática. I. Nievola, Julio Cesar. II. Manffra, Elisangela Ferretti, 1973-. III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática. IV. Título.

CDD 20. ed. – 004

ATA DE DEFESA DE DISSERTAÇÃO DE MESTRADO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

DEFESA DE DISSERTAÇÃO DE MESTRADO Nº 03/2015

Aos 15 dias do mês de Abril de 2015 realizou-se a sessão pública de Defesa da Dissertação “ **Avaliação da Relevância de Atributos Posturográficos Globais e Estruturais para a Identificação do Acidente Vascular Encefálico utilizando Modelos de Classificação**” apresentado pelo aluno **Luiz Henrique Giovanini Marrega**, como requisito parcial para a obtenção do título de Mestre em Informática, perante uma Banca Examinadora composta pelos seguintes membros:

Prof. Dr. Júlio César Nievola
PUCPR (Orientador)

Júlio Cesar Nievola APROVADO
(assinatura) (Aprov/Reprov)

Prof.^a Dr.^a Elisangela Ferretti Manffra
PUCPR/PPGTS (co-orientadora)

Elisangela Manffra APROV.
(assinatura) (Aprov/Reprov)

Prof. Dr. Emerson Cabrera Paraiso
PUCPR

Emerson Cabrera Paraiso APROV.
(assinatura) (Aprov/Reprov)

Prof. Dr. Marcelo Eduardo Pellenz
PUCPR

Marcelo Eduardo Pellenz APROV.
(assinatura) (Aprov/Reprov)

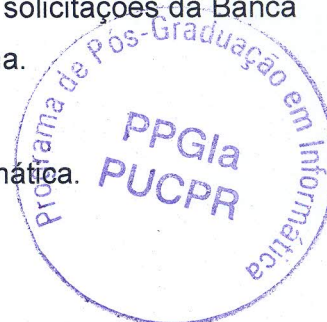
Prof. Dr. Paulo José Abatti
UTFPR

Paulo José Abatti APROV.
(assinatura) (Aprov/Reprov)

Conforme as normas regimentais do PPGIa e da PUCPR, o trabalho apresentado foi considerado APROVADO (aprovado/reprovado), segundo avaliação da maioria dos membros desta Banca Examinadora. Este resultado está condicionado ao cumprimento integral das solicitações da Banca Examinadora registradas no Livro de Defesas do programa.

Andreia Malucelli
Prof.^a Dr.^a Andreia Malucelli.

Coordenadora do Programa de Pós-Graduação em Informática.



*Dedico este trabalho a todos aqueles que,
de forma direta ou indireta, contribuíram
com a sua concretização.*

*Em especial, à minha mãe Margaret e à
minha noiva Adriana.*

Agradecimentos

Agradeço primeiramente a Deus pelas condições favoráveis durante a realização das atividades do mestrado (boas condições de saúde, vivência familiar amorosa, situação econômica confortável, etc.) e também por ter me conduzido por caminhos que levaram a um grande amadurecimento, pessoal e profissional, durante todo este período.

Um agradecimento especial ao meu orientador, Prof. Dr. Julio Cesar Nievola, por ter me guiado com tamanha excelência durante toda esta jornada, pela flexibilidade em relação ao tema de pesquisa, pelas ideias claras e objetivas, pelo senso crítico apurado e, principalmente, pelo apoio incondicional durante os momentos adversos.

Um agradecimento especial também à minha coorientadora, Profa. Dra. Elisangela Ferretti Manffra, com quem já tive o prazer de trabalhar em outros projetos e que vem contribuindo significativamente com o meu amadurecimento desde a época da graduação. Obrigado por aceitar participar deste estudo, por fornecer a base de dados, pelo olhar aguçado e experiente principalmente durante a manipulação dos dados e, sobretudo, por ter me iniciado neste tema de pesquisa no passado.

À Simone Massaneiro Silva, que fez a coleta da base de dados utilizada por este estudo durante suas atividades no PPGTS da PUCPR. Simone, esta pesquisa não seria possível sem o seu árduo trabalho de coleta, obrigado!

À Cheila, secretária do PPGIa que, dotada de imenso carisma e disposição, sempre me auxiliou com as questões burocráticas de maneira ágil e eficiente.

À toda a minha família, por quem tenho imenso orgulho e carinho. Em especial à minha mãe Margaret que, fazendo papel de mãe e de pai, sempre proveu-me um lar para o qual era bom retornar. Obrigado pelo seu carinho incondicional.

À minha noiva e futura esposa, Adriana Lívia, pelo carinho e apoio incondicionais. Obrigado por compreender os meus momentos de ausência, por vibrar junto comigo a cada etapa vencida, por me aconselhar durante os momentos difíceis, enfim, por desempenhar o papel de companheira em todas as suas atribuições.

À CAPES, pelo apoio financeiro através da bolsa de estudos.

A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê.

Arthur Schopenhauer

Lista de Figuras

- Figura 2.1** - Ilustração das estratégias posturais (a) do tornozelo, (b) do quadril e (c) do passo. 9
- Figura 2.2** - (a) Representação ilustrativa de uma plataforma de força e seus eixos x (direção ântero-posterior), y (direção médio-lateral) e z (direção vertical) de medida das componentes de força e torque das trajetórias do CP. (b) Posicionamento de um indivíduo sobre uma plataforma de força real. 11
- Figura 2.3** - (a) Ilustração da posição espacial do CP de um indivíduo posicionado sobre uma plataforma de força. As oscilações do CP nas direções AP e ML podem ser visualizadas (b) espacial e conjuntamente, através do estatocinesiógrama, ou (c,d) temporal e individualmente, através dos estabilogramas. 12
- Figura 2.4** - Comparação entre um sinal (a) determinístico e outro (b) aleatório, ambos possuindo a mesma variabilidade (capturada pelo desvio padrão) porém estruturas temporais (mensuradas pela entropia amostral) completamente distintas..... 14
- Figura 2.5** - Elipse (linha preta tracejada) cuja área engloba 95% das trajetórias do CP (linha cinza contínua) dentro do estatocinesiógrama. 20
- Figura 2.6** - Espectro de potências de um estabilograma com a indicação das frequências de pico (F_{pico}), média ($F_{\text{média}}$), em 50% (F_{50}) e em 80% (F_{80}) da concentração espectral total..... 21
- Figura 2.7** - Atrator do sistema de Lorenz..... 23
- Figura 2.8** - (a) Curva Dc versus m para o modelo de Lorenz em que Dc atinge um valor de saturação a partir de $m = 3$. Ilustração do atrator do modelo em um espaço de fases (b) tridimensional e (c) bidimensional, em que o ponto cinza denota um falso cruzamento de duas trajetórias. Nos dois casos, foi utilizado $\tau = 6$ [CAM05]..... 25
- Figura 2.9** - Soma de correlação em escala logarítmica para o atrator do modelo de Lorenz reconstruído com $m = 1, 2, \dots, 10$. As linhas tracejadas delimitam a região da curva utilizada para a estimativa da dimensão de correlação, a qual considera desde 0.5% até 75% de pares de pontos do atrator reconstruído. 27

- Figura 2.10** - Ilustração das três fases (delimitadas pelas linhas pontilhadas) que compõem a curva EP versus $L0$ (linha sólida) de um atrator caótico. A inclinação de uma linearização do segmento da fase II (linha tracejada) representa uma estimativa do máximo expoente de Lyapunov do sistema..... 29
- Figura 2.11** - Ilustração do processo de escalamento de uma série temporal original para $\eta = 2$ (painel superior) e $\eta = 3$ (painel inferior). 31
- Figura 2.12** - Curvas médias da entropia multiescala para séries temporais do CP de 14 adultos e 14 idosos nas direções AP (esquerda) e ML (direita). Nos cálculos, foi utilizado $M = 2$ e $\Gamma = 0.2$ 31
- Figura 2.13** - Ilustração da auto-similaridade (ou fractalidade) em diferentes escalas de tempo presente em uma série temporal do CP na direção ântero-posterior. 32
- Figura 2.14** - Ilustração da faixa de valores assumidos pelo expoente de escala (α) para processos aleatórios fracionais do tipo Gaussiano e Browniano. São indicados também os valores de α para uma série de ruído Gaussiano puro ($\alpha = 0.5$), ruído $1/f$ puro ($\alpha = 1$) e ruído Browniano puro ($\alpha = 1.5$). 33
- Figura 2.15** - À esquerda, são mostradas as linearizações (linhas cinzas sólidas) das curvas logarítmicas Fw versus w (linhas pretas sólidas) obtidas para uma série temporal de 60 segundos das trajetórias do CP (painel superior à direita) e sua versão *surrogate* de amostras (painel inferior à direita). 35
- Figura 2.16** - Gráfico de recorrência com $r = 20$ para a série temporal $x(n)$ do modelo de Lorenz em regime caótico. Para a reconstrução do atrator, utilizou-se dimensão de imersão $m = 5$ e atraso temporal $\tau = 5$ 37
- Figura 2.17** - Representação do indivíduo $x_2 + y$ em forma de árvore de sintaxe..... 46
- Figura 2.18** - Ilustração da operação de cruzamento entre dois programas quaisquer (pais) para a criação de dois novos indivíduos (filhos). 47
- Figura 3.1** - Diagrama de blocos que expressa a metodologia empregada no presente estudo através de cinco etapas gerais. 57
- Figura 3.2** - À esquerda, é exibida a derivada de 1ª ordem do erro de previsão calculado para uma série do CP na direção ML de um indivíduo saudável na condição OA. O pico deste gráfico (linha tracejada) foi utilizado para identificar o início da fase II, enquanto

que a queda abaixo do limiar de 50% do valor de pico (linha pontilhada) demarca o fim desta fase. À direita, é exibida a curva do erro de previsão com as demarcações da fase II. 65

Figura 3.3 - Representação da composição das bases *APG*, *APE* e *APGE* em função da inclusão de determinados atributos posturográficos dentre todos os 47 disponíveis para cada uma das 114 instâncias. 71

Figura 3.4 - Ilustração do método empregado neste trabalho para a avaliação dos efeitos da seleção de atributos nas bases de dados em estudo, tomando como exemplo as bases *APG* e *APG-selec*. 73

Figura 3.5 - Ilustração do método empregado neste trabalho para a avaliação da relevância de atributos posturográficos globais e estruturais durante a classificação de sujeitos saudáveis e de pacientes pós-AVE. As setas representam a realização de uma comparação estatística entre as duas médias apontadas. 75

Figura 4.1 - Curva *DC versus m*, $m = 1, 2, \dots, 15$ (linha —o— preta) para uma série do CP na direção AP de um indivíduo pós-AVE na condição OA. A linha cinza pontilhada representa o critério de Takens utilizado para a escolha das dimensões m e *DC* do sistema. 78

Figura 4.2 - Comparação entre os valores médios dos parâmetros *DC* (canto superior esquerdo), λ_{max} (canto superior direito), *EntAm* (canto inferior esquerdo) e α (canto inferior direito) calculados para as séries originais do CP (barras em branco) e para suas versões *surrogate* de fases (barras em preto) e *surrogate* de amostras (barras em cinza). As *cruzes* indicam diferença estatisticamente significativa ($p < 0.05$) entre um determinado parâmetro calculado para um determinado tipo de série e o mesmo parâmetro calculado para os outros tipos de séries em uma dada coordenada do CP. 79

Figura 4.3 - Comparação entre os valores médios dos APG de domínio temporal obtidos para os indivíduos saudáveis (barras em branco) e pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* e as *cruzes* posicionadas em uma determinada coordenada indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente. 80

Figura 4.4 - Comparação entre os valores médios dos APG de domínio espaço-temporal obtidos para os indivíduos saudáveis (barras em branco) e para os pacientes

pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* indicam diferença estatística com nível p de significância < 0.001 81

Figura 4.5 - Comparação entre os valores médios dos APG de domínio espectral calculados para os indivíduos saudáveis (barras em branco) e para os pacientes pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* e as *cruzes* posicionadas em uma determinada coordenada indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente. ... 82

Figura 4.6 - Comparação entre os valores médios dos APE calculados para os indivíduos saudáveis (barras em branco) e para os pacientes pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* e as *cruzes* posicionadas em uma determinada coordenada indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente. 83

Figura 4.7 - Comparação entre as acurácias médias conjuntas dos classificadores para as versões integral e reduzida das bases APG (painel da esquerda), APE (painel central) e APGE (painel da direita). As barras de erro representam o intervalo de confiança de 95% sobre os valores médios. As *estrelas* e as *cruzes* indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente. 85

Figura 4.8 - Comparação entre as acurácias médias conjuntas dos classificadores para as bases de dados integrais (painel da esquerda) e reduzidas (painel da direita). As barras de erro representam o intervalo de confiança de 95% sobre os valores médios. As *estrelas* e as *cruzes* indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente. 85

Lista de Tabelas

Tabela 2.1 - Relação de métricas posturográficas globais comuns na literatura com seus respectivos domínios. A terceira coluna apresenta alguns estudos da área da posturografia que utilizaram as métricas em questão.....	15
Tabela 2.2 - Relação das métricas de avaliação estrutural do equilíbrio mais comuns na literatura com a indicação das respectivas propriedades do SCP descritas pelas mesmas. A terceira coluna apresenta estudos da área da posturografia que consideraram as métricas em questão.....	21
Tabela 2.3 - Relação de técnicas utilizadas na literatura para a avaliação de correlações de longa duração em séries temporais experimentais. A terceira coluna exhibe exemplos de estudos que utilizaram as técnicas em questão no contexto da avaliação do equilíbrio.	32
Tabela 3.1 - Relação dos algoritmos de classificação utilizados nesta pesquisa com os respectivos parâmetros de configuração adotados. A última coluna atribui uma sigla para as diferentes configurações de execução de cada algoritmo.	73
Tabela 4.1 - Relação do tempo médio gasto durante a extração dos atributos posturográficos considerados nesta pesquisa.	77
Tabela 4.2 - Relação dos atributos selecionados a partir das bases de dados <i>APG</i> , <i>APE</i> e <i>APGE</i>	84
Tabela 4.3 - Relação das acurácias individuais dos classificadores utilizados neste estudo para os diferentes conjuntos de atributos posturográficos.	86

Lista de Símbolos

$x(t)$: Série temporal contínua
$x(n)$: Série temporal discreta
x_0	: Condição inicial
$u(n)$: Espaço de fases
$u_R(n)$: Espaço de fases reconstruído
τ	: Atraso temporal ou <i>delay</i>
m	: Dimensão de imersão
D_C	: Dimensão de correlação
$C_m(r)$: Soma de correlação
r	: Raio de busca da Soma de Correlação
$F_{x,y,z}$: Componente de força na direção x, y ou z
$M_{x,y,z}$: Componente de torque na direção x, y ou z
h	: Altura da base de apoio do indivíduo
D_{med}	: Distância média
D_{rms}	: Distância RMS
A_{des}	: Amplitude de deslocamento
V_{med}	: Velocidade média
DP_V_{med}	: Desvio padrão da velocidade média
V_p	: Velocidade de pico
V_{rms}	: Velocidade RMS
V_{medT}	: Velocidade média total
F_{med}	: Frequência espectral média
F_p	: Concentração de $p\%$ de potência espectral
t_A	: Tempo de amostragem
f_A	: Frequência de amostragem
λ_{max}	: Máximo expoente de Lyapunov
$EntAp$: Entropia aproximada
$EntAm$: Entropia amostral
$EntME$: Entropia Multiescala
TR_{RQA}	: Taxa de recorrência
DET_{RQA}	: Determinismo
ENT_{RQA}	: Entropia RQA
η	: Fator de escala utilizado na entropia multiescala
H	: Expoente de Hurst
\hat{H}	: Expoente de Hurst aproximado
α	: Expoente de escala
w	: Comprimento das janelas não sobrepostas
$M(i, j)$: Matriz de recorrência
f_c	: Frequência de corte

Lista de Abreviaturas

AD	: Árvore de Decisão
AM	: Aprendizagem de Máquina
AP	: Ântero-Posterior
APE	: Atributos Posturográficos Estruturais
APG	: Atributos Posturográficos Globais
AVE	: Acidente Vascular Encefálico
CFS	: <i>Correlation-based Feature Selection</i>
CG	: Centro de Gravidade
CP	: Centro de Pressão
DAS	: <i>Stabilogram Diffusion Analysis</i>
DFA	: <i>Detrended Fluctuation Analysis</i>
E	: Condição Olhos Abertos sobre Superfície de Espuma
EP	: Erro de Previsão
EMD	: <i>Empirical Mode Decomposition</i>
FIM	: Função de Informação Mútua
FSA	: <i>Fast Simulated Annealing</i>
GB	: Giga Byte
IA	: Inteligência Artificial
IMF	: <i>Intrinsic Mode Function</i>
k-NN	: <i>k-Nearest Neighbours</i>
LDA	: <i>Linear Discriminant Analysis</i>
LVQ	: <i>Learning Vector Quantization</i>
ML	: Médio-Lateral
NB	: Naive Bayes
OA	: condição Olhos Abertos
OF	: condição Olhos Fechados
PC	: Paralisia Cerebral
PG	: Programação Genética
PUCPR	: Pontifícia Universidade Católica do Paraná
R/S	: <i>Rescaled Range Analysis</i>
RBF	: Ruído Browniano Fracional
RGF	: Ruído Gaussiano Fracional
RMS	: <i>Root Mean Square</i>
RNA	: Rede Neural Artificial
RQA	: <i>Recurrence Quantification Analysis</i>
SCP	: Sistema de Controle Postural
SVM	: <i>Support Vector Machines</i>
SWV	: <i>Scaled Windowed Variance</i>
WEKA	: Waikato Environment for Knowledge Analysis

Resumo

A manutenção do equilíbrio é uma tarefa motora extremamente importante para a realização das atividades de vida diária de forma independente e segura. Esta tarefa tem sido estudada com o auxílio da posturografia, uma técnica que permite quantificar os perfis de oscilação corporal de um indivíduo em uma determinada condição de equilíbrio na forma de uma série temporal dos deslocamentos do Centro de Pressão (CP). Através da parametrização do sinal do CP, é possível avaliar certas características sobre o Sistema de Controle Postural (SCP) do indivíduo (ou grupo de indivíduos) em estudo, tal como a susceptibilidade a sofrer quedas e também os efeitos de programas de reabilitação. Na literatura, duas perspectivas de parametrização do balanço postural foram estabelecidas: a *global* e a *estrutural*. No primeiro caso, atributos posturográficos globais (APG) são utilizados para descrever uma única propriedade do SCP, a estabilidade; já no segundo caso, atributos posturográficos estruturais (APE) são calculados para descrever diversas outras propriedades do SCP, tal como o número de graus de liberdade ativos, a regularidade, etc. Diante do exposto, o objetivo deste trabalho é avaliar a relevância do uso simultâneo entre atributos globais e estruturais em comparação com o uso de atributos de apenas uma ou outra categoria em particular para, com o auxílio de algoritmos clássicos de Aprendizagem de Máquina (AM) e também da Programação Genética (PG), diferenciar os perfis de balanço postural produzidos por indivíduos saudáveis e pós-Acidente Vascular Encefálico (AVE). Para isso, foi utilizada uma base de dados contendo 114 perfis de balanço postural (ou instâncias), 57 oriundos de pacientes pós-AVE e os outros 57 associados a indivíduos saudáveis pareados. Então, por meio de rotinas computacionais desenvolvidas em ambiente MATLAB, foram extraídos 28 APG e 18 APE de cada instância. Na sequência, seis modelos de classificação foram treinados e avaliados em três situações distintas: utilizando (i) somente APG, (ii) somente APE e (iii) APG e APE em conjunto. Como métrica de avaliação de desempenho, foi utilizada a acurácia média conjunta calculada entre os seis classificadores. Finalmente, avaliou-se os efeitos da redução de dimensionalidade dos conjuntos integrais de atributos na acurácia média conjunta dos modelos preditivos. Dentre os resultados, o uso simultâneo entre APG e APE rendeu, antes da seleção de atributos, uma acurácia média conjunta estatisticamente equivalente ($p = 0.218$) àquela resultante do uso de apenas APG (66.8%) e, após a seleção de atributos, uma acurácia média conjunta estatisticamente equivalente ($p = 0.705$) àquela oriunda do uso de apenas APE (68.8%). Diante destes achados pode-se concluir que, para a identificação das estratégias de equilíbrio adotadas por indivíduos saudáveis e pós-AVE com o auxílio de classificadores tradicionais e da PG, o uso de parâmetros posturográficos apenas globais (quando não realizada uma seleção de atributos) ou apenas estruturais (quando realizada uma seleção de atributos) possui a mesma relevância do que o uso simultâneo entre parâmetros das duas categorias, sendo estas primeiras as opções mais vantajosas neste contexto por exigir menores esforços computacional e intelectual.

Palavras-Chave: Equilíbrio Postural, AVE, Mineração de Dados, Estudos de Séries Temporais.

Abstract

Maintaining balance is an essential motor task for humans to perform their daily activities in a safe and independent way. This task has been studied with the help of posturography, a technique that allows to quantify the postural sway profiles of a person in a given balance task by recording a time series of its center-of-pressure (CP) displacements. The parameterization of COP signal allows one to investigate some properties of the Postural Control System (PCS), such as the risk of falling and the effects of rehabilitation programs. Two distinct views of CP parameterization are available in the literature: the global and the structural. The metrics derived from the former view are called global posturographic features (GPF) and used to describe only one property of PCS, the stability, whereas the metrics derived from the latter are known as structural posturographic features (SPF) and computed do index several other properties of PCS, such as the number of active degrees-of-freedom, the regularity, etc. The main purpose of this work is to evaluate the relevance of combining global and structural posturographic features in comparison with using features from only one category to distinguish healthy subjects from stroke patients with the help of classical Machine Learning algorithms as well as Genetic Programming. For this purpose, a dataset of 114 postural sway profiles (or instances) was used, where 57 of them were derived from post-stroke patients and the 57 remaining were produced by healthy matched subjects. Then, a MATLAB routine was implemented to extract 28 GPF and 18 SPF from each instance. Next, six popular classification models were trained and evaluated under three experiments: using (i) only GPF, (ii) only SPF and (iii) combining GPF and SPF. The average joint accuracy among the six models were adopted as performance metric in each experiment. Finally, the effects of feature selection on the classifiers' performance were investigated. Among the results, before the features selection, the combination of GPF and SPF yielded a statistically equivalent ($p = 0.218$) performance than the use of only GPF (66.8%) and, after the features selection, a statistically equivalent ($p = 0.705$) performance than the use of only SPF (68.8%). Based on this findings, it is possible to infer that the combination of global and structural posturographic features has the same relevance for distinguish healthy subjects from stroke patients with popular classification methods than the use of only global features (when a features selection is not performed) or only structural features (when a features selection is performed). These latter options are most advantageous because they involve the computation of features from only one category, requiring less efforts than compute features from both categories.

Keywords: Postural Sway, Stroke, Data Mining, Time Series Studies.

Sumário

Lista de Figuras	vii
Lista de Tabelas	xi
Lista de Símbolos	xii
Lista de Abreviaturas	xiii
Resumo	xiv
Abstract	xv
Capítulo 1 - Introdução	1
1.1. Motivação	3
1.2. Objetivo	5
1.3. Hipótese de Pesquisa	5
1.4. Escopo	5
1.5. Contribuição Científica	6
1.6. Organização do Documento	7
Capítulo 2 - Descrição do Problema	8
2.1. Avaliação do Equilíbrio Postural.....	9
2.1.1. Sistema de Controle Postural em Situação Saudável e no AVE	9
2.1.2. Posturografia.....	10
2.1.3. Perspectivas de Análise e Parametrização do Sinal do CP.....	13
2.1.3.1. Atributos Posturográficos Globais.....	15
2.1.3.2. Atributos Posturográficos Estruturais.....	21
2.1.4. Estudos Correlatos	39
2.2. Classificação do Equilíbrio Postural.....	42
2.2.1. Principais Métodos de Classificação	42
2.2.2. Programação Genética	45
2.2.3. Treinamento e Avaliação dos Modelos de Classificação	49
2.2.4. Seleção de Atributos	50
2.2.5. Estudos Correlatos	51
2.3. Conclusão do Capítulo	55
Capítulo 3 - Metodologia	57
3.1. Seleção da Base de Dados Posturográficos	58
3.2. Extração de Atributos Posturográficos	59

3.2.1. Rotina de Processamento Global.....	59
3.2.2. Rotina de Processamento Estrutural.....	60
3.2.3. Avaliação do Tempo de Execução das Rotinas.....	67
3.3. Avaliação da Natureza dos Dados Posturográficos.....	68
3.4. Análise Estatística dos Atributos Posturográficos.....	69
3.5. Avaliação dos Métodos de Classificação.....	70
3.5.1. Bases de Dados Secundárias.....	70
3.5.2. Seleção de Atributos.....	72
3.5.3. Relevância dos Atributos Posturográficos Globais e Estruturais.....	74
Capítulo 4 - Avaliação dos Resultados Experimentais	76
4.1. Resultados.....	77
4.1.1. Tempo Gasto na Extração de Atributos.....	77
4.1.2. Dados Originais <i>versus</i> Dados Surrogate.....	78
4.1.3. Análise Estatística dos Atributos Posturográficos.....	80
4.1.4. Desempenho da Seleção de Atributos.....	84
4.1.5. Desempenho dos Modelos de Classificação.....	85
4.2. Discussão.....	87
4.2.1. Esforço Requerido Durante a Extração dos Atributos.....	87
4.2.2. Natureza dos Dados Posturográficos.....	88
4.2.3. Avaliação Tradicional da Relevância dos Atributos Posturográficos.....	90
4.2.4. Avaliação da Relevância Conjunta dos Atributos Posturográficos.....	96
4.2.5. Algumas Observações Sobre as Avaliações Individual e Conjunta dos Parâmetros Posturográficos.....	102
Conclusão e Trabalhos Futuros	104
Referências	107
ANEXOS	127

Capítulo 1

Introdução

A manutenção do equilíbrio é uma tarefa motora de extrema importância para a realização das atividades de vida diária de forma independente e segura. Do ponto de vista dos seres humanos, o controle do equilíbrio é encarado como sendo uma tarefa simples e usual, pois é um pré-requisito básico para a maioria de suas atividades de vida diária; entretanto, do ponto de vista fisiológico, esta é uma tarefa bastante complexa. Isso porque, para o controle do equilíbrio, o sistema nervoso central recruta diferentes tipos de informações sensoriais do indivíduo (e.g., visual, vestibular e proprioceptiva) em um trabalho integrado com as propriedades passivas de seu sistema músculo esquelético, sendo este mecanismo biológico conhecido na literatura como Sistema de Controle Postural (SCP) (e.g., [DON07; GHO10; ROE06; NEG13]). Sendo assim, a escolha de uma estratégia postural adequada para a manutenção do equilíbrio em um determinado ambiente dependerá das condições sensoriais e motoras do indivíduo [MER10].

O controle do equilíbrio tem sido amplamente avaliado pela literatura com o auxílio da posturografia, uma técnica utilizada para mensurar alguma variável relacionada ao balanço postural do indivíduo durante a postura ereta quieta ou durante a execução de uma determinada tarefa na postura em pé [DUA10; SIL12]. No primeiro caso, que será objeto de interesse no presente estudo, captura-se um padrão específico

de balanço postural conhecido como *oscilação postural espontânea*, a qual representa uma oscilação produzida naturalmente pelo indivíduo durante o controle do equilíbrio na postura ereta quieta [LEB00]. Neste cenário, uma variável posturográfica amplamente mensurada para descrever a oscilação postural espontânea do sujeito é o centro de pressão (CP) [DUA00b; DUA10; FRE06], a qual representa um resultado combinado entre os mecanismos biológicos recrutados para o controle do equilíbrio e a força da gravidade [DUA00b; DUA01], e que detêm informações sobre as estratégias posturais adotadas durante a manutenção do equilíbrio [VET10]. Mais especificamente, utilizando em geral um dispositivo eletrônico conhecido como plataforma de força [DUA01], os deslocamentos do CP são capturados na forma de séries temporais (e.g., [DOY05; HAA04; TAL13]).

Entre outros cenários, a posturografia tem sido utilizada para avaliar o controle do equilíbrio em pacientes acometidos pelo acidente vascular encefálico (AVE) (e.g., [GEN08; HAA04; ROE06; SIL12]), pela paralisia cerebral (PC) (e.g., [DEF09; DON08]) e pela esclerose múltipla (e.g., [NEG13]). Isso porque as inúmeras sequelas decorrentes destas afecções acarretam uma série de conflitos sensoriais e, como resultado, reduzem a capacidade do indivíduo de se manter em pé de forma estável [SIL12]. O estudo do equilíbrio nesta situação tem grande relevância pois pode revelar a susceptibilidade destes indivíduos a sofrer quedas (e.g., [FOR95]) e também os efeitos de programas de reabilitação (e.g., [HAA04; ROE06]). Por exemplo, o estudo de Forster e Young [FOR95] relata uma incidência de quedas de até 73% em pacientes pós-AVE com até seis meses de lesão.

Para avaliar certas características do controle do equilíbrio em uma determinada população, não basta apenas mensurar a oscilação corporal de voluntários de pesquisa com o auxílio da posturografia; é necessário também aplicar alguma técnica de análise sobre os dados adquiridos. Na literatura, existem basicamente duas perspectivas de análise e parametrização do sinal do CP: a *global* e a *estrutural* [DUA10; KIR13; STE11]. No primeiro caso, *parâmetros posturográficos globais* são utilizados para extrair informações relevantes a partir do tamanho (ou magnitude) dos padrões de oscilação do CP, tanto no domínio do tempo como também no domínio da frequência [DUA10; KIR12]. Já na perspectiva de análise estrutural, procura-se extrair informações a partir dos padrões temporais de oscilação do CP com o auxílio de *parâmetros posturográficos estruturais* [KIR12].

A partir das duas perspectivas supracitadas, diversas métricas de avaliação postural têm sido propostas e utilizadas com sucesso na literatura para descrever certas propriedades dos mecanismos biológicos de controle motor [KIR13]. Mais especificamente, os atributos de avaliação global têm sido associados unicamente à *estabilidade* postural [KIR12], ao passo que os atributos de avaliação estrutural vêm sendo utilizados para descrever diversas outras propriedades do SCP, tais como a *regularidade* (ou *previsibilidade*) (e.g., [MAD11; TAL13]), os *graus de liberdade ativos* (e.g., [DON07; ROE06; GUR13]), a *sensibilidade* (e.g., [DON07; ROE06; PAS05]), a *complexidade* (e.g., [DUA08; KIR12]) e *auto similaridade* (ou *fractalidade*) (e.g., [DUA01; AMO07]), entre outras. Nesse contexto, não existe um consenso na literatura sobre qual métrica posturográfica é a mais relevante para a avaliação do equilíbrio em uma ou outra população [DUA10]. Ao invés disso, estudos da área têm chamado a atenção para a importância de se considerar diferentes métricas, globais e estruturais, para uma descrição mais completa sobre o SCP [GHO10; HAN05; TAL13].

1.1. Motivação

Neste ponto da discussão, cabe apresentar duas lacunas identificadas pelo autor na literatura da avaliação postural que serviram de motivação para a realização do presente estudo. Em primeiro lugar, questionou-se a utilização de atributos posturográficos pertencentes exclusivamente a uma única categoria (global ou estrutural) por estudos da área (e.g., [BAR02; BLA08; CAR01; CAS13; COR01; DEL03; DEM08; DUA10; LAF04; ROE06; SIL12]) pois, como mencionado anteriormente, ambas as perspectivas de análise do equilíbrio têm revelado com sucesso diferentes propriedades dos mecanismos biológicos de controle postural. A segunda lacuna reside no fato de que, mesmo quando são consideradas métricas posturográficas de avaliação global e estrutural, os estudos seguem, em geral, um procedimento padrão: tentar diferenciar sistemas fisiológicos em condições distintas (e.g., diferentes estados de saúde ou tarefas de equilíbrio distintas) comparando estatisticamente um mesmo parâmetro calculado a partir das séries temporais do CP produzidas pelos sistemas em questão (e.g., [DON08; GHO10; HAN05; KIR12]). Tal abordagem conduz o pesquisador a uma entre duas conclusões possíveis acerca de cada atributo posturográfico considerado no estudo: *o atributo é* ou *o atributo não é* capaz de

diferenciar os sistemas fisiológicos avaliados, ou seja, sucesso ou falha, sem meio termo. Sendo assim, não se pode afirmar aqui que os diferentes atributos posturográficos, os quais descrevem propriedades particulares do SCP, estão sendo utilizados de forma conjunta. Na verdade, o que se faz é identificar o(s) atributo(s) que melhor diferencia(m) os sistemas avaliados dentre todos aqueles considerados no estudo de uma forma individual, comparando-se estatisticamente o mesmo atributo entre sistemas (ou populações).

Por outro lado, alguns estudos no campo da inteligência artificial têm utilizado com sucesso métodos supervisionados de Aprendizagem de Máquina para identificar sistemas fisiológicos distintos com base em atributos posturográficos. A maioria destes estudos está relacionada ao desenvolvimento de dispositivos para a detecção de quedas em tempo real voltados ao público idoso (e.g., [ALH14; CHO11; KER12; LIU10]). Além disso, tais métodos foram utilizados também para identificar as estratégias posturais adotadas por indivíduos jovens saudáveis (e.g., [RAS09; SAR14]), idosos saudáveis (e.g., [RAS09]) e até mesmo por uma criança saudável em processo de crescimento [BOU10], todos eles em tarefas específicas de controle do equilíbrio. No presente trabalho, acredita-se também que o emprego de métodos de classificação seja mais apropriado para o problema em questão visto que, por meio desta abordagem, todos os atributos posturográficos considerados no estudo (ou ao menos um subconjunto contendo os mais discriminantes deles) serão utilizados de maneira conjunta para descrever as populações (ou classes) avaliadas na forma de um modelo matemático de previsão. Entretanto, os estudos que utilizam esta abordagem têm se limitado apenas ao uso de atributos posturográficos globais. Por isso, nesta pesquisa, pretende-se avaliar a relevância de atributos posturográficos globais e estruturais para a descrição do equilíbrio postural em diferentes populações utilizando técnicas supervisionadas de classificação.

Finalmente, é importante mencionar que o AVE é uma das principais causas de morte e de incapacidade funcional em adultos no Brasil e no mundo [POR14], figurando assim como um grave problema mundial de saúde pública. Neste cenário, o estudo do equilíbrio na presença do AVE pode trazer diversas contribuições, tal como a indicação do risco destes indivíduos a sofrer acidentes envolvendo quedas (e.g., [FOR95]) e também de bons descritores quantitativos dos efeitos de programas de reabilitação motora (e.g., [HAA04; ROE06]). É possível também contribuir com o

aprimoramento de sistemas de detecção de quedas, os quais têm sido utilizados para minimizar os riscos de acidentes domésticos envolvendo pessoas com comprometimentos de controle motor (e.g., [ALH14]). Todas estas contribuições podem beneficiar a população acometida pelo AVE, proporcionando-lhe melhorias na qualidade de vida.

1.2. Objetivo

O objetivo geral deste trabalho é avaliar a relevância do uso conjunto de atributos globais e estruturais comuns na área da posturografia em comparação com o uso de atributos de apenas uma ou outra categoria em particular para, com o auxílio de algoritmos clássicos de Aprendizagem de Máquina (AM) e também da Programação Genética (PG), diferenciar os perfis de oscilação postural produzidos por sistemas fisiológicos saudáveis e pós-AVE.

1.3. Hipótese de Pesquisa

A hipótese da presente pesquisa é a de que a combinação entre parâmetros posturográficos das categorias global e estrutural é mais eficiente para diferenciar indivíduos saudáveis e pós-AVE com o auxílio dos métodos de classificação considerados em comparação com o uso de atributos de apenas uma ou outra categoria em particular.

1.4. Escopo

O escopo deste trabalho limita-se à avaliação do desempenho conjunto entre seis determinados modelos de classificação formulados para diferenciar os perfis de oscilação postural produzidos por indivíduos saudáveis e por pacientes pós-AVE a partir de três conjuntos distintos de atributos posturográficos, sendo eles: conjunto de atributos de categoria global (28 de cada instância), conjunto de atributos de categoria estrutural (18 de cada instância) e, finalmente, conjunto de atributos globais e estruturais, composto pela união dos dois conjuntos anteriores. De maneira intuitiva, a extração dos referidos atributos posturográficos também faz parte do escopo deste

trabalho, a qual é realizada para uma base de dados coletada por um outro estudo [SIL12] e composta por 114 perfis de balanço postural (i.e., 114 instâncias), 57 de cada população (ou classe). Além disso, inseriu-se no escopo desta pesquisa a avaliação dos efeitos da seleção de atributos no desempenho conjunto dos classificadores. Finalmente, além da análise utilizando os modelos de classificação —que é o foco deste trabalho —, é realizada também uma análise estatística tradicional sobre os atributos posturográficos apenas com o intuito de investigar os aspectos em comum entre estas duas abordagens de análise mencionadas, bem como os aspectos particulares de cada uma delas. Neste contexto, é importante salientar que uma comparação entre o desempenho das duas abordagens supracitadas para distinguir indivíduos saudáveis e pós-AVE não faz parte do escopo deste trabalho.

1.5. Contribuição Científica

As principais contribuições científicas desta pesquisa são (i) a indicação de qual situação de uso de parâmetros posturográficos (apenas globais, apenas estruturais ou globais e estruturais em conjunto) é mais relevante para a identificação de indivíduos saudáveis e de pacientes pós-AVE com o auxílio de modelos preditivos clássicos, (ii) a indicação dos parâmetros em si e, por fim, (iii) a relação custo/benefício envolvida na utilização de tais parâmetros, ou seja, a indicação dos esforços intelectual e computacional necessários para a extração de parâmetros de uma ou outra categoria em detrimento da taxa de acerto conjunta dos classificadores alcançada com o seu uso. Todas estas indicações podem ser utilizadas por estudos futuros para um delineamento mais eficiente do método, possibilitando por exemplo priorizar o cálculo de atributos mais eficazes para diferenciar as populações em questão e, de forma complementar, evitar o desperdício de tempo e de esforço com o cálculo de parâmetros que exibiram um baixo poder de discriminação nesta pesquisa. Além disso, por meio da indicação da relação custo/benefício supracitada, o pesquisador/profissional da saúde pode avaliar com maior clareza a viabilidade do uso de parâmetros posturográficos de uma ou outra categoria em função de determinados aspectos relevantes, tal como o tempo hábil para se familiarizar com os procedimentos matemáticos envolvidos nos cálculos dos parâmetros, a necessidade de um *feedback* em tempo real do paciente — em caso de

avaliações em hospitais ou clínicas de reabilitação — e, conseqüentemente, o tempo disponível para a extração dos parâmetros, entre outros.

Outra contribuição relevante a ser mencionada é que este estudo gera demanda por novas pesquisas envolvendo a aplicação de ferramentas de mineração de dados no contexto da avaliação do equilíbrio postural, um cenário pouco explorado na literatura atual. Mais especificamente, novos estudos podem ser delineados para, por exemplo, validar os resultados reportados neste trabalho, os quais podem estar condicionados às características da amostra utilizada, ou também para verificar se tais resultados se repetem na presença de outros fatores além do AVE que também exercem influência no controle do equilíbrio do indivíduo, tal como a idade avançada, a condição física e até mesmo outros tipos de afecções neurológicas (e.g., paralisia cerebral, esclerose múltipla, etc.).

1.6. Organização do Documento

Com relação à estruturação do presente documento, será inicialmente realizada por meio do capítulo 02 uma discussão sobre o problema investigado nesta pesquisa. Mais especificamente, será discorrido sobre o fenômeno das oscilações posturais do corpo humano existentes durante a manutenção do equilíbrio na postura ereta quieta, tanto em situação saudável como também na presença do AVE. Além disso, serão também exploradas as duas principais perspectivas de análise — juntamente com suas métricas e demais procedimentos matemáticos — utilizadas na literatura para a avaliação do controle do equilíbrio em situações adversas (e.g., idade avançada, condição física comprometida, presença de acometimento neurológico, etc.). Por fim, serão apresentados os principais métodos supervisionados de classificação da literatura de AM, os quais serão utilizados neste trabalho com o intuito de diferenciar os perfis de oscilação postural produzidos por indivíduos saudáveis e pós-AVE. Na sequência, no capítulo 03, será apresentado o método delineado para esta pesquisa, seguida por uma exposição e discussão acerca dos resultados experimentais alcançados no estudo por meio do capítulo 04. Então, serão feitas as considerações finais do trabalho, a saber, a conclusão e algumas sugestões de pesquisas futuras, seguidas finalmente pela indicação das referências consultadas durante a realização do presente estudo.

Capítulo 2

Descrição do Problema

No presente capítulo, será inicialmente discorrido por meio da seção 2.1 sobre o controle do equilíbrio motor em situação saudável e também na presença do AVE, seguido pela explicação dos principais parâmetros de avaliação posturográfica oriundos das perspectivas de análise global e estrutural da oscilação corporal humana. Então, na seção 2.2, serão apresentados alguns dos principais métodos supervisionados de classificação da área de Aprendizagem de Máquina, os quais serão utilizados neste estudo com o intuito de diferenciar sistemas fisiológicos saudáveis e pós-AVE com base nos parâmetros posturográficos abordados na seção 2.1. Por fim, a seção 2.3 traz as conclusões do presente capítulo.

2.1. Avaliação do Equilíbrio Postural

2.1.1. Sistema de Controle Postural em Situação Saudável e no AVE

Composto pelas componentes sensorial e motora do sistema nervoso central trabalhando de forma integrada [SIL12], o sistema de controle postural (SCP) é o mecanismo fisiológico dos seres humanos responsável pela manutenção do equilíbrio [DUA00a], uma tarefa motora indispensável para a realização da maioria das nossas atividades de vida diária e que por isso está intimamente relacionada com a independência funcional do sujeito. Em outras palavras, é atribuída ao SCP a tarefa de determinar constantemente a estratégia postural mais adequada para o controle do equilíbrio em virtude das atuais condições sensoriais do indivíduo bem como das atuais condições do ambiente onde o mesmo se situa (e.g., superfície de contato regular/irregular, rígida/macia, ambiente com baixa/alta iluminação). Neste contexto, Duarte [DUA00a] *apud* Nashner e McCollum [NAS85] aponta as três estratégias principais utilizadas pelo sujeito para a manutenção da postura ereta, sendo elas as estratégias (i) do tornozelo, (ii) do quadril e (iii) do passo (Figura 2.1). Em situação saudável, observam-se padrões de movimento articular que envolvem uma combinação simultânea entre as três estratégias supracitadas [DUA00a].

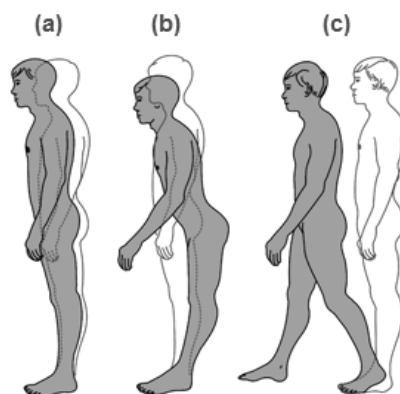


Figura 2.1 - Ilustração das estratégias posturais (a) do tornozelo, (b) do quadril e (c) do passo.

Fonte: Shumway-Cook e Woollacott [SHU03].

Como o SCP é formado pelas componentes sensorial e motora do sistema nervoso central, a escolha das estratégias posturais durante a manutenção do equilíbrio depende, entre outros fatores, das condições sensoriais e motoras do indivíduo [MER10]. Sendo assim, as diversas sequelas fisiológicas decorrentes de afecções

nerológicas em geral, tal como o acidente vascular encefálico (AVE) e a paralisia cerebral (PC), conferem ao indivíduo acometido maiores dificuldades na tarefa de controle do equilíbrio do que em situação saudável. Daqui em diante, o AVE será utilizado como afecção neurológica de referência pois o estudo do equilíbrio nesta situação é objeto de interesse por esta pesquisa.

O AVE é caracterizado como um conjunto de alterações neurológicas causadas por falhas no suprimento sanguíneo do encéfalo [WOL00]. Atualmente, além de figurar como uma das principais causas de morte no Brasil e no mundo, o AVE é uma das patologias que mais incapacitam o indivíduo acometido para a realização de suas atividades de rotina diária de forma independente e segura [AMI14; POR14], tornando-o assim dependente de terceiros. As inúmeras sequelas decorrentes do AVE (consultar [SIL12] para uma discussão detalhada) expõem o sujeito a uma série de conflitos sensoriais, aumentando significativamente o risco de acidentes envolvendo quedas. Isso porque estas pessoas utilizam estratégias posturais alternativas e que nem sempre são efetivas para mantê-las em equilíbrio [PYO04]. No estudo de Forster e Young [FOR95], por exemplo, foi relatada uma incidência de quedas de até 73% em pacientes pós-AVE com até seis meses de lesão.

2.1.2. Posturografia

Conforme discutido previamente, o SCP é o mecanismo biológico dos seres humanos responsável pelo controle do equilíbrio. Tratando-se especificamente da postura ereta quieta, o equilíbrio é alcançado através do constante "gerenciamento" (ou compensação) da chamada *oscilação postural espontânea* produzida pelo corpo [HOR06]. Em termos mais específicos, a oscilação postural espontânea está associada ao chamado *centro de gravidade* (CG) do corpo, que é o ponto de aplicação da força da gravidade sobre o mesmo [WIN90; ROD07]. Dessa forma, o principal objetivo do SCP é a conservação do CG em um ponto fixo no espaço [SIL12]. Porém, em virtude das perturbações que agem no indivíduo de forma interna (e.g., movimento involuntário dos órgãos, contrações musculares) e externa (e.g., força da gravidade), o CG oscila continuamente ao redor de um ponto fixo [FRE06; SIL12], originando assim a oscilação postural espontânea [LEB00]. Nesse contexto, o *centro de pressão* (CP) é o ponto de aplicação das forças de reação do solo em resposta aos movimentos do CG e,

consequentemente, é um descritor da oscilação postural espontânea. Em outras palavras, as oscilações do CP descrevem as respostas neuromusculares em função dos deslocamentos do CG [SIL12], representando assim um resultado combinado entre os mecanismos fisiológicos de controle postural do indivíduo e a força da gravidade [DUA00a]. Dessa forma, o CP carrega informações tanto sobre as oscilações posturais produzidas pelo indivíduo como também sobre as estratégias motoras utilizadas pelo mesmo durante o controle do equilíbrio [VET10].

De acordo com Duarte [DUA10], a oscilação postural espontânea pode ser avaliada de maneira qualitativa, por meio da observação, e também de forma quantitativa, com o auxílio de equipamentos de medição. Focando a presente discussão na avaliação quantitativa, o equilíbrio pode ser estudado através da posturografia, uma técnica utilizada para mensurar o balanço postural espontâneo ou uma variável associada [DUA10]. Neste cenário, os estudos da área têm utilizado com frequência os deslocamentos temporais do CP para descrever o balanço postural do sujeito [DUA00b; DUA10; FRE06]. Para isso, um equipamento conhecido como *plataforma de força* é normalmente utilizado com o intuito de capturar as trajetórias do CP na forma de séries temporais (e.g., [DOY05; GHO10; HAA04; TAL13]). Este dispositivo consiste em uma placa de metal na qual estão distribuídos sensores de força — normalmente do tipo célula de carga ou piezoelétrico — com o objetivo de mensurar três componentes de força, F_x , F_y e F_z , e também as componentes de torque associadas a estas forças, sendo elas M_x , M_y e M_z , em que x , y e z representam respectivamente as direções ântero-posterior (AP), médio-lateral (ML) e vertical [DUA10] (Figura 2.2).

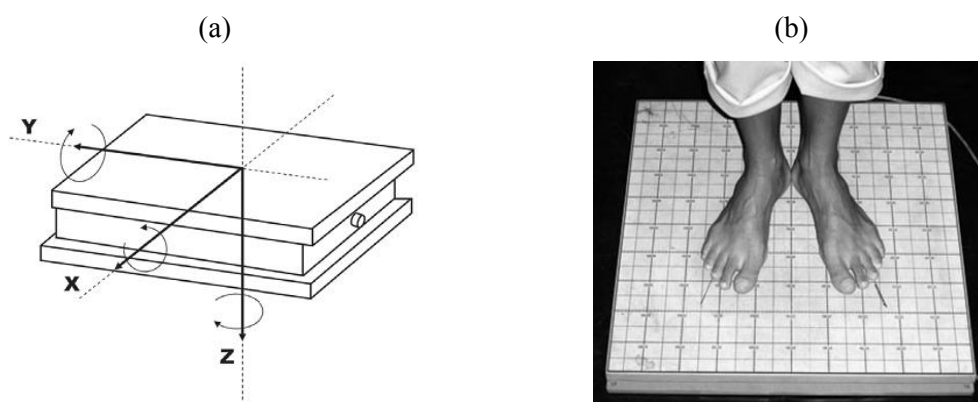


Figura 2.2 - (a) Representação ilustrativa de uma plataforma de força e seus eixos x (direção ântero-posterior), y (direção médio-lateral) e z (direção vertical) de medida das componentes de força e torque das trajetórias do CP. (b) Posicionamento de um indivíduo sobre uma plataforma de força real.

Fonte: Duarte e Freitas [DUA10].

O CP corresponde a uma medida de posição definida por duas coordenadas na superfície da plataforma de força em função da orientação do indivíduo avaliado [DUA10]. Convencionando o posicionamento do indivíduo sobre a plataforma conforme a Figura 2.3-a, os sinais de força e de momento adquiridos pela plataforma podem ser utilizados para calcular as trajetórias do CP ao longo das direções AP (CP_{AP}) e ML (CP_{ML}) com o auxílio das seguintes expressões [DUA00a; DUA10; SIL12]:

$$CP_{AP} = (-h \cdot F_x - M_y) / F_z \quad (2.1)$$

$$CP_{ML} = (-h \cdot F_y + M_x) / F_z, \quad (2.2)$$

sendo h a altura da base de apoio acima da plataforma de força (por exemplo, um tapete ou uma espuma posicionados sobre a plataforma). A representação temporal das trajetórias do CP em uma determinada direção em particular (AP ou ML) constitui um *estabilograma* (Figura 2.3-c,d), ao passo que a visualização espaço-temporal destas trajetórias nas direções AP e ML simultaneamente é realizada por meio do chamado *estatocinesiograma* (Figura 2.3-b) [MAI05].

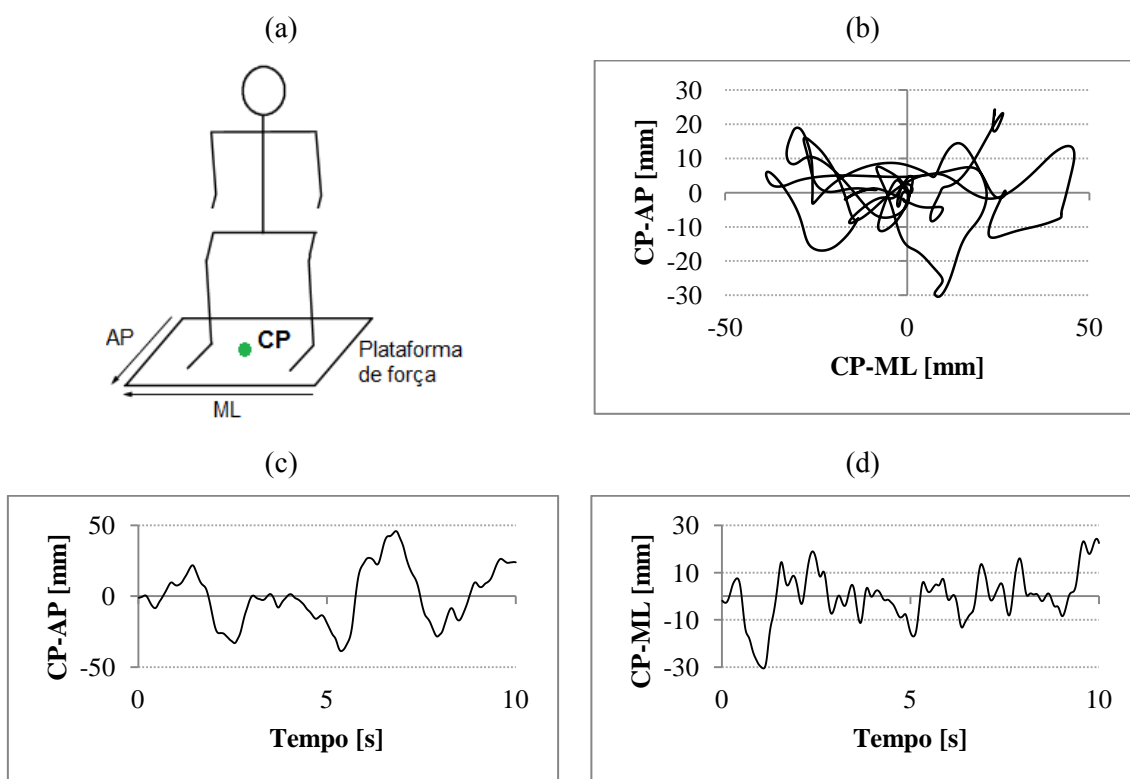


Figura 2.3 - (a) Ilustração da posição espacial do CP de um indivíduo posicionado sobre uma plataforma de força. As oscilações do CP nas direções AP e ML podem ser visualizadas (b) espacial e conjuntamente, através do estatocinesiograma, ou (c,d) temporal e individualmente, através dos estabilogramas.

Fonte: adaptado de [DUA01].

2.1.3. Perspectivas de Análise e Parametrização do Sinal do CP

Ao longo dos anos, duas perspectivas principais de análise e parametrização do sinal do CP foram estabelecidas na literatura da área da posturografia: a *global* e a *estrutural* [DUA10]. Conforme será discutido na subseção 2.1.4, estas análises têm sido empregadas no sinal do CP adquirido nas mais diversas situações (e.g., diferentes condições de saúde, características antropométricas, condição física, idade, ambiente) com o intuito de investigar certas propriedades dos mecanismos de controle postural, as quais podem ser utilizadas para avaliar o risco de acidentes envolvendo quedas (e.g., [FOR95]) e também os efeitos de programação de reabilitação (e.g., [DON07; HAA04; ROE06]) sobre uma determinada população, entre outros fatores. Nesta pesquisa, os atributos (ou métricas) provenientes de técnicas de avaliação global do equilíbrio serão denominados *atributos posturográficos globais* (APG), ao passo que àqueles oriundos de técnicas de avaliação estrutural serão denominados *atributos posturográficos estruturais* (APE).

Na perspectiva de análise global do equilíbrio, procura-se extrair informações relevantes a partir do tamanho (ou amplitude) dos padrões de deslocamento do CP, tanto no domínio do tempo como também no domínio da frequência [DUA10]. Além disso, considera-se que as oscilações do CP representam algum tipo de "erro motor" cometido pelo indivíduo durante a realização da tarefa de equilíbrio de forma que, quanto maiores os níveis de variabilidade postural, menor a estabilidade do sujeito [KIR13]. Dessa forma, são esperados níveis mais baixos de oscilação para sistemas fisiológicos saudáveis em comparação com sistemas com algum tipo de comprometimento (e.g., estado de saúde e/ou condição física debilitados, idade avançada, etc.). Diante do exposto, os mais diversos APG propostos na literatura (e.g., [BAR02; DEM08]) têm sido utilizados para descrever uma única propriedade dos mecanismos biológicos de controle motor: a *estabilidade* [KIR12; KIR13].

Já na perspectiva de análise estrutural do equilíbrio, são extraídas informações a partir das estruturas (ou padrões) temporais das flutuações do sinal do CP [KIR12]. Ainda, de acordo com Duarte e Freitas [DUA10], a análise estrutural identifica subunidades nos dados posturográficos e as relaciona aos processos de controle motor. Nesse contexto, ao contrário dos atributos de avaliação global que têm sido utilizados para descrever apenas a *estabilidade postural*, os atributos de avaliação estrutural do

equilíbrio têm sido capazes de descrever com sucesso diversas outras propriedades do SCP, tal como o *número de graus de liberdade ativos* (e.g., [DON07; ROE06; GUR13]), a *sensibilidade a pequenas perturbações locais* (e.g., [DON07; ROE06; PAS05]), o *grau de regularidade* (ou *previsibilidade*) (e.g., [MAD11; TAL13]) e de *complexidade* (e.g., [DUA08; COS05; KIR12]), a *auto-similaridade* (ou *fractalidade*) (e.g., [AMO07; DON07; DUA01]), entre outras.

A principal diferença entre parâmetros posturográficos globais e estruturais reside no fato de que, no primeiro caso, são capturadas informações a partir do tamanho dos padrões de oscilação — tanto no domínio do tempo como também no domínio da frequência — enquanto que, no segundo caso, utiliza-se a ocorrência (ou organização) temporal daqueles padrões de oscilação. Tal situação é ilustrada na Figura 2.4 em que são comparados dois sinais distintos, ambos com a mesma quantidade de variabilidade — a qual pode ser capturada por meio do desvio padrão (vide subseção 2.1.3.1) — porém com ocorrências temporais — que podem ser descritas pela entropia amostral (vide subseção 2.1.3.2) — completamente distintas.

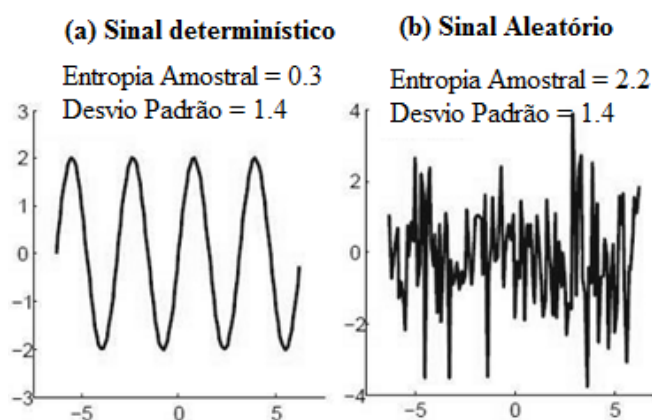


Figura 2.4 - Comparação entre um sinal (a) determinístico e outro (b) aleatório, ambos possuindo a mesma variabilidade (capturada pelo desvio padrão) porém estruturas temporais (mensuradas pela entropia amostral) completamente distintas.

Fonte: adaptado de Kirchner [KIR13].

Atualmente, diversos parâmetros globais e estruturais estão disponíveis na literatura para a parametrização das trajetórias do CP. Por exemplo, somente no estudo de Baratto et al. [BAR02], foram investigadas 38 variáveis derivadas do sinal do CP a partir das duas perspectivas de análise já explanadas. No entanto, não existe um consenso na literatura sobre qual métrica posturográfica é a mais relevante para descrever o controle do equilíbrio em uma determinada população ou tarefa motora

[DUA10]. Ao invés disso, como tais variáveis têm sido capazes de descrever com sucesso diferentes propriedades dos mecanismos biológicos de controle postural, estudos da área têm chamado a atenção para a importância de se considerar diferentes métricas, globais e estruturais, para uma descrição mais completa sobre tais mecanismos [GHO10; HAN05; TAL13]. Nas duas subseções seguintes, serão abordados os atributos globais e estruturais mais utilizados pelos estudos da área da posturografia, juntamente com as propriedades do SCP descritas pelos mesmos.

2.1.3.1. Atributos Posturográficos Globais

Na literatura, uma série de atributos posturográficos globais têm sido utilizados para descrever a *estabilidade* dos mecanismos de controle postural por meio da análise da variabilidade das trajetórias do CP [KIR12; KIR13]. Na Tabela 2.1, encontram-se relacionados alguns dos atributos globais mais comuns na área da posturografia pertencentes a três diferentes domínios, sendo eles: (i) temporal, (ii) espaço-temporal e (iii) espectral. Tanto os atributos temporais como também os espectrais são parâmetros unidimensionais, ou seja, são calculados individualmente nas direções AP e ML dos deslocamentos do CP (i.e., extraídos dos estabilogramas). Já os atributos espaço-temporais são parâmetros bidimensionais, ou seja, são calculados considerando as direções AP e ML simultaneamente (i.e., baseiam-se no estatocinesiograma).

Tabela 2.1 - Relação de métricas posturográficas globais comuns na literatura com seus respectivos domínios. A terceira coluna apresenta alguns estudos da área da posturografia que utilizaram as métricas em questão.

Domínio	Métrica	Estudos relacionados
Temporal	Distância média (<i>Dmed</i>)	Carpenter et al. [CAR01], Demura et al. [DEM08], Ganesan et al. [GAN13], Genthon et al. [GEN08], Rugelk e Sevsek [RUG11], Vette et al. [VET10], Zok et al. [ZOK08].
	Distância <i>root-mean-square</i> (<i>Drms</i>)	Cabeza-Ruiz et al. [CAB11a], Carpenter et al. [CAR01], Deffeyes et al. [DEF09], Demura et al. [DEM08], Duarte e Freitas [DUA10], Ganesan et al. [GAN13].
	Desvio padrão (<i>DP</i>)	Castro [CAS13], Demura et al. [DEM08], Donker et al. [DON07], Duarte e Freitas [DUA10], Ghomashchi et al. [GHO10], Gurses e Celik [GUR13], Kirchner et al. [KIR12], Kirchner [KIR13], Madeleine et al. [MAD11], Tallon et al. [TAL13].

Domínio	Métrica	Estudos relacionados
Temporal	Amplitude de deslocamento (<i>Ades</i>)	Blaszczyk [BLA08], Castro [CAS13], Corriveau et al. [COR01], Deffeyes et al. [DEF09], Doyle et al. [DOY05], Duarte e Freitas [DUA10], Ganesan et al. [GAN13], Kirchner et al. [KIR12], Kirchner [KIR13], Lafond et al. [LAF04], Zok et al. [ZOK08].
	Velocidade média (<i>Vmed</i>)	Blaszczyk [BLA08], Cabeza-Ruiz et al. [CAB11a], Castro [CAS13], Corriveau et al. [COR01], Duarte e Freitas [DUA10], Demura et al. [DEM08], Ganesan et al. [GAN13], Kirchner et al. [KIR12], Kirchner [KIR13], Lafond et al. [LAF04], Rugelk e Sevsek [RUG11], Vette et al. [VET10], Zok et al. [ZOK08].
	Desvio padrão da velocidade média (<i>DP_Vmed</i>)	Demura et al. [DEM08].
	Velocidade de pico (<i>Vp</i>)	Doyle et al. [DOY05].
	Velocidade <i>root-mean-square</i> (<i>Vrms</i>)	Demura et al. [DEM08], de Haart et al. [HAA04], Silva [SIL12].
Espaço-Temporal	Velocidade média total (<i>VmedT</i>)	Duarte e Freitas [DUA10].
	Sway path (<i>SP</i>)	Baratto et al. [BAR02], Deffeyes et al. [DEF09], Donker et al. [DON07; DON08], Han et al. [HAN05], Lebedowska e Syczewska [LEB00].
	Deslocamento da Oscilação Total (<i>DOT</i>)	Blaszczyk [BLA08], Castro [CAS13], Duarte e Freitas [DUA10], Kirchner et al. [KIR12], Kirchner [KIR13], Rugelk e Sevsek [RUG11], Silva [SIL12].
	Área do CP (<i>Area</i>)	Baratto et al. [BAR02], Blaszczyk [BLA08], Cabeza-Ruiz et al. [CAB11a], Castro [CAS13], Deffeyes et al. [DEF09], Demura et al. [DEM08], Doyle et al. [DOY05], Duarte e Freitas [DUA10], Kirchner et al. [KIR12], Kirchner [KIR13], Lafond et al. [LAF04], Oliveira et al. [OLI09], Rugelk e Sevsek [RUG11], Tallon et al. [TAL13].
Espectral	Frequência espectral média (<i>Fmed</i>)	Cabeza-Ruiz et al. [CAB11a], Carpenter et al. [CAR01], Duarte e Freitas [DUA10], Ganesan et al. [GAN13], Genthon et al. [GEN08], Lafond et al. [LAF04], Silva [SIL12], Zok et al. [ZOK08].
	Concentração de <i>p%</i> da potência espectral (<i>Fp</i>)	Baratto et al. [BAR02], Demura et al. [DEM08], Duarte e Freitas [DUA10], Gurses e Celik [GUR13], Kirchner et al. [KIR12], Kirchner [KIR13].

Distância Média

A distância média (D_{med}) representa a distância entre o valor médio do sinal do CP na direção AP (CP_{ap}) ou ML (CP_{ml}) e a origem ($CP_{ap} = 0$ ou $CP_{ml} = 0$) [GAN13]:

$$D_{med_{ap|ml}} = \frac{1}{N} \sum_{i=1}^N |CP_{ap|ml}(i)|, \quad (2.3)$$

em que o parâmetro N é o comprimento (i.e., o número total de amostras) do sinal.

Distância RMS

A distância RMS (*Root Mean Square*) (D_{rms}) representa uma forma alternativa de cálculo de distância no sinal do CP [DUA10; GAN13]:

$$D_{rms_{ap|ml}} = \sqrt{\frac{1}{N} \sum_{i=1}^N CP_{ap|ml}^2(i)} \quad (2.4)$$

Desvio Padrão

O desvio padrão (DP) mede a dispersão das trajetórias do CP ao longo do tempo em relação a um valor médio constante [DUA10; GAN13]:

$$DP_{ap|ml} = \sqrt{\frac{\sum_{i=1}^N [CP_{ap|ml}(i) - D_{med_{ap|ml}}]^2}{N - 1}} \quad (2.5)$$

Amplitude de Deslocamento

A amplitude de deslocamento (A_{des}) representa a distância entre os deslocamentos máximo e mínimo do sinal do CP [DUA10; GAN13]:

$$A_{des_{ap|ml}} = |\max[CP_{ap|ml}(n)] - \min[CP_{ap|ml}(n)]| \quad (2.6)$$

Velocidade Média

Em ambas as direções AP e ML, uma série temporal da velocidade dos deslocamentos do CP pode ser calculada da seguinte forma [GAN13]:

$$V_{ap|ml}(n) = \frac{CP_{ap|ml}(n+1) - CP_{ap|ml}(n)}{t_a}, \quad n = 1, 2, \dots, N-1 \quad (2.7)$$

em que t_a representa o tempo de amostragem do sinal de posição do CP. Então, a velocidade média (V_{med}) em cada direção pode ser calculada como [DUA10]:

$$V_{med_{ap|ml}} = \frac{1}{N} \sum_{i=1}^N V_{ap|ml}(i) \quad (2.8)$$

Desvio Padrão da Velocidade Média

O desvio padrão da velocidade média ($DP_{V_{med}}$) está associado com a dispersão do sinal de velocidade das trajetórias do CP em relação ao seu nível médio (V_{med}) [DEM08]:

$$DP_{V_{med_{ap|ml}}} = \sqrt{\frac{\sum_{i=1}^N [V_{ap|ml}(i) - V_{med_{ap|ml}}]^2}{N-1}} \quad (2.9)$$

Velocidade de Pico

A velocidade de pico (V_p) representa o maior valor observado no sinal de velocidade das trajetórias do CP [DOY05]:

$$V_{p_{ap|ml}} = \max(|V_{ap|ml}(n)|) \quad (2.10)$$

Velocidade RMS

A velocidade RMS (V_{rms}) representa uma forma alternativa de cálculo da velocidade dos deslocamentos do CP ao longo do tempo [DEM08]:

$$V_{rms_{ap|ml}} = \sqrt{\frac{1}{N} \sum_{i=1}^N V_{ap|ml}^2(i)} \quad (2.11)$$

Velocidade Média Total

Baseada no estatocinesiógrama, a velocidade média total (V_{medT}) considera a velocidade dos deslocamentos do CP nas direções AP e ML simultaneamente [DUA10]:

$$V_{medT} = \frac{f_a}{N} \sum_{i=1}^N \sqrt{V_{ap}^2(i) + V_{ml}^2(i)}, \quad (2.12)$$

sendo f_a a frequência de amostragem do sinal de posição do CP.

Sway Path

O *sway path* (SP) é um parâmetro clínico utilizado para capturar a quantidade de atividade postural exibida no estatocinesiógrama [DON08; HAN05]:

$$SP = \sum_{i=1}^{N-1} \sqrt{[CP_{ap}(i+1) - CP_{ap}(i)]^2 + [CP_{ml}(i+1) - CP_{ml}(i)]^2} \quad (2.13)$$

Deslocamento da Oscilação Total

Baseado no estatocinesiógrama, o deslocamento da oscilação total (DOT) representa o comprimento das trajetórias do CP sobre a base de suporte do indivíduo [DUA10]:

$$DOT = \sum_{i=1}^N \sqrt{CP_{ap}^2(i) + CP_{ml}^2(i)} \quad (2.14)$$

Área do CP

Este parâmetro é utilizado para estimar a dispersão das trajetórias do CP dentro do estatocinesiograma através da área de uma elipse que engloba 95% destas trajetórias [DUA10], conforme ilustra a Figura 2.5.

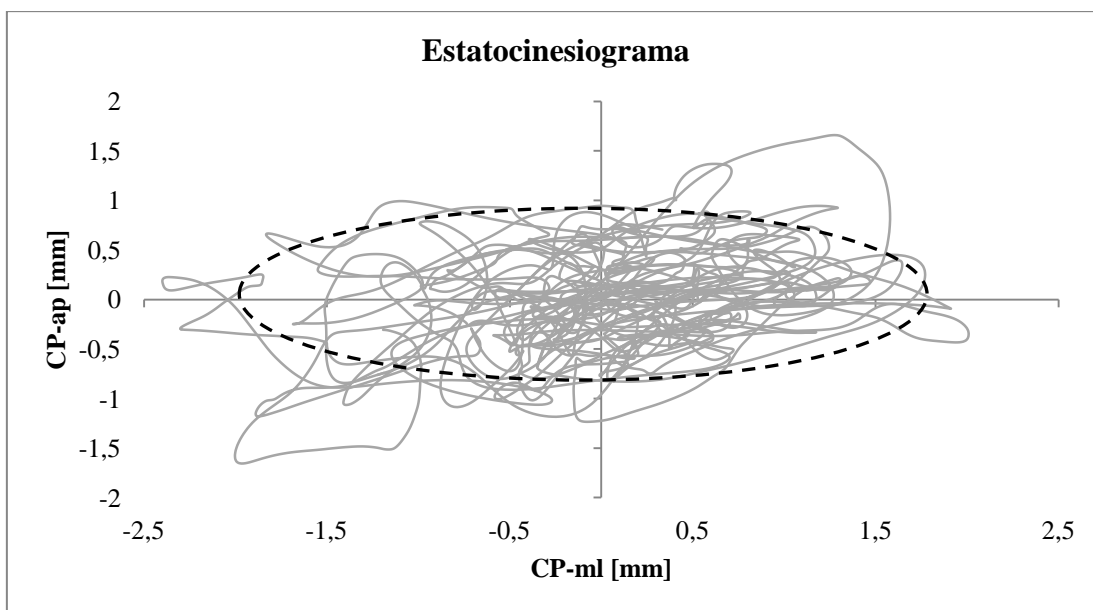


Figura 2.5 - Elipse (linha preta tracejada) cuja área engloba 95% das trajetórias do CP (linha cinza contínua) dentro do estatocinesiograma.

Fonte: o autor.

Frequência Espectral Média

A frequência espectral média (F_{med}) representa o valor médio do espectro de potências das trajetórias do CP em cada uma das direções de deslocamento [DUA10; GAN13] (Figura 2.6).

Concentração de Potência Espectral

Este parâmetro estima o valor de frequência até onde, partindo da origem, encontra-se concentrada $p\%$ da potência total do espectro do sinal do CP em cada uma das direções de deslocamento [DUA10; KIR12] (Figura 2.6).

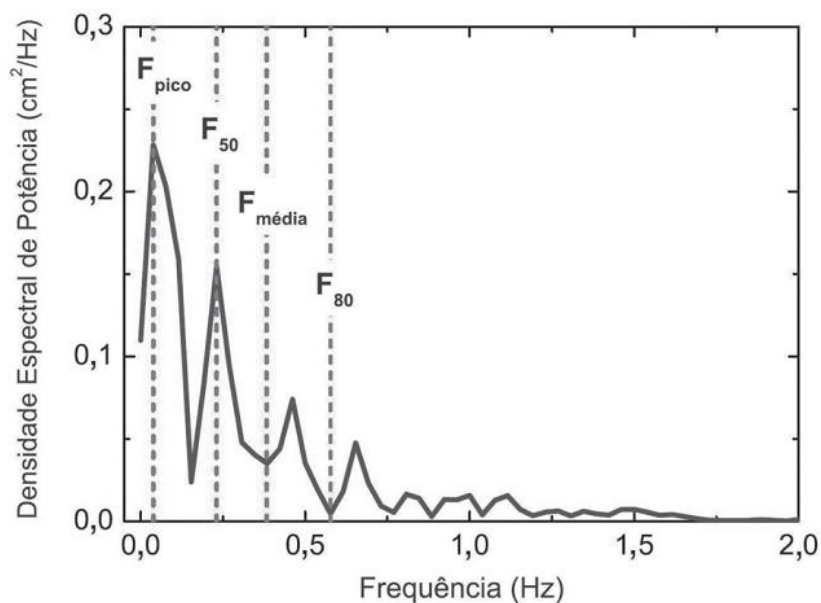


Figura 2.6 - Espectro de potências de um estabilograma com a indicação das frequências de pico (F_{pico}), média ($F_{\text{média}}$), em 50% (F_{50}) e em 80% (F_{80}) da concentração espectral total.

Fonte: Duarte e Freitas [DUA10].

2.1.3.2. Atributos Posturográficos Estruturais

Conforme elucidado previamente, enquanto os atributos globais têm sido utilizados para descrever apenas a *estabilidade* postural, os atributos de avaliação estrutural do equilíbrio têm sido capazes de descrever diversas outras propriedades do SCP por meio da análise da organização temporal das trajetórias do CP. A Tabela 2.2 traz uma relação dos atributos estruturais mais utilizados pela literatura da área da posturografia, juntamente com as respectivas propriedades dos mecanismos de controle motor descritas pelos mesmos. É importante salientar que todos os parâmetros relacionados na Tabela 2.2 são unidimensionais, ou seja, são extraídos individualmente nas direções AP e ML das trajetórias do CP.

Tabela 2.2 - Relação das métricas de avaliação estrutural do equilíbrio mais comuns na literatura com a indicação das respectivas propriedades do SCP descritas pelas mesmas. A terceira coluna apresenta estudos da área da posturografia que consideraram as métricas em questão.

Métrica	Propriedade	Estudos relacionados
Dimensão de Correlação	Graus de liberdade ativos envolvidos no SCP	Gurses e Celik [GUR13], Donker et al. [DON07], Doyle et al. [DOY05], Ghomashchi et al. [GHO10], Han et al. [HAN05], Madeleine et al. [MAD11], Pascolo et al. [PAS05], Roerdink et al. [ROE06]

Métrica	Propriedade	Estudos relacionados
Máximo Expoente de Lyapunov	Estabilidade local do SCP	Donker et al. [DON07], Pascolo et al. [PAS05], Roerdink et al. [ROE06]
Entropia Amostral	Grau de regularidade das trajetórias do CP	Duarte e Sternad [DUA08], Donker et al. [DON07; DON08], Kirchner et al. [KIR12], Kirchner [KIR13], Madeleine et al. [MAD11], Roerdink et al. [ROE06], Tallon et al. [TAL13]
Entropia Multiescala	Grau de complexidade das trajetórias do CP	Duarte e Sternad [DUA08], Costa et al. [COS07], Kirchner [KIR13]
Expoentes de escala e de Hurst	Auto-similaridade das trajetórias do CP	Amoud et al. [AMO07], Delignières et al. [DEL03], Donker et al. [DON07], Duarte e Sternad [DUA08], Duarte e Zatsiorsky [DUA00b], Duarte e Zatsiorsky [DUA01], Kirchner et al. [KIR12], Kirchner [KIR13], Roerdink et al. [ROE06]
Taxa de recorrência, Determinismo e Entropia de recorrência	Regularidade e complexidade das estruturas determinísticas do sinal do CP	Negahban et al. [NEG13], Samaei et al. [SAM13], Seigle et al. [SEI09], Tallon et al. [TAL13]

Dimensão de Correlação

No contexto da avaliação do equilíbrio postural, a dimensão de correlação (D_C) é um parâmetro capaz de capturar a dimensionalidade das estruturas temporais das trajetórias do CP, refletindo em uma estimativa do número de graus de liberdade ativos envolvidos nos mecanismos de controle motor [GUR13; ROE06].

Dada uma série temporal experimental, o primeiro passo para a estimativa da dimensão de correlação consiste na reconstrução do *atrator* do sistema descrito por aquela série dentro do chamado *espaço de fases*. Mais especificamente, o espaço de fases consiste em um sistema de coordenadas em que os eixos estão associados às variáveis de estado (ou variáveis independentes) do sistema em estudo, sendo o atrator a representação da dinâmica daquele sistema dentro do espaço de fases [GAN01]. Como exemplo, a Figura 2.7 exibe o atrator do modelo de Lorenz [LOR63], um sistema dinâmico de comportamento caótico determinístico regido por três variáveis de estado: x , y e z .

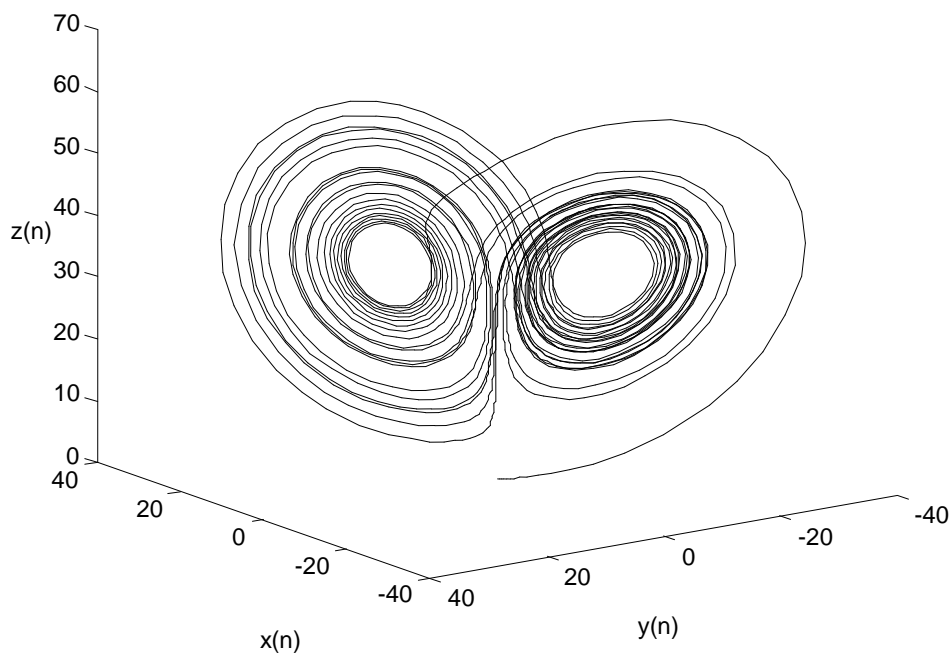


Figura 2.7 - Atrator do sistema de Lorenz.

Fonte: o autor.

Para sistemas teóricos como o de Lorenz em que são conhecidas as equações que regem o comportamento de todas as variáveis de estado, é possível gerar séries temporais que descrevam estas variáveis e, posteriormente, visualizar o atrator dentro do espaço de fases (vide Figura 2.7 para o sistema de Lorenz). Entretanto, quando são utilizadas séries temporais experimentais, nem todas as variáveis de estado do sistema são conhecidas pois, na prática, é impossível registrar simultaneamente todo o conjunto de variáveis de estado do fenômeno em observação [CAM05; CAM07; GAN01]. Felizmente, segundo o teorema da imersão introduzido por Takens [TAK81], é possível estimar o comportamento de todas as m variáveis de estado de um determinado sistema a partir das observações de uma única variável, permitindo assim a *reconstrução* do atrator daquele sistema em um espaço de fases m -dimensional. Para este propósito, um dos métodos mais utilizados na literatura é o *método das coordenadas defasadas* ou *das coordenadas de atraso temporal* proposto por Ruelle e Takens [RUE79; TAK81], em que o espaço de fases reconstruído $u_R(n)$ é expresso como uma composição de vetores de estado m -dimensionais defasados entre si por uma componente de atraso τ , assim como:

$$u_R(n) = (x(n), x(n + \tau), x(n + 2\tau), \dots, x(n + [m - 1]\tau)), \quad (2.15)$$

sendo $x(n)$ a série temporal que descreve a única variável de estado observada do sistema, τ o *atraso temporal* (ou *delay*) e m a *dimensão de imersão* do espaço de fases, ou seja, o número de eixos utilizados para a exibição do atrator, representando assim uma estimativa do (até então desconhecido) número total de variáveis de estado do sistema em estudo [CAM05; FIG00; GAN01; NAS05].

Neste cenário, alguns métodos foram propostos na literatura para a estimativa do valor de τ a partir de séries temporais experimentais; dentre eles, destaca-se o uso do primeiro mínimo local da Função de Informação Mútua (FIM) [FRA86], definida como:

$$FIM(\tau) = \sum_{\substack{a \in A \\ b \in B}} P_{A,B}(a, b) \times \log_2 \left[\frac{P_{A,B}(a, b)}{P_A(a) \cdot P_B(b)} \right], \quad \tau = 1, \dots, \tau_{max}, \quad (2.16)$$

em que A e B representam respectivamente os conjuntos dos N primeiros elementos das séries temporais $x(n)$ e $x(n + \tau)$, $\tau = 1, \dots, \tau_{max}$, sendo a um elemento do conjunto A e b um elemento do conjunto B . $P_A(a)$ e $P_B(b)$ representam as probabilidades de se encontrar, respectivamente, o elemento a em uma escolha do conjunto A e o elemento b em uma escolha do conjunto B e, finalmente, $P_{A,B}(a, b)$ é a probabilidade de que um dos vetores atrasados $x(n + \tau)$ tenha a como primeiro elemento e b como segundo elemento [CAM05; GAN01]. Neste contexto, quando o valor do atraso é muito pequeno, os conjuntos A e B serão praticamente idênticos, ao passo que um valor muito elevado de τ produzirá conjuntos completamente descorrelacionados. No entanto, o valor apropriado de τ — que corresponde ao primeiro mínimo local da FIM [FRA86] — garante um grau de correlação adequado entre os conjuntos. Em outras palavras, a FIM mede a dependência geral entre a série original e a série defasada, fornecendo uma estimativa adequada e robusta do atraso temporal para a reconstrução do atrator [CAM05; GAN01; VAS07].

Com relação à dimensão de imersão (m), este parâmetro corresponde ao número total de variáveis de estado do sistema dinâmico em estudo e, conseqüentemente, ao número de eixos a ser utilizado no espaço de fases para uma visualização adequada do seu atrator. Por exemplo, para o sistema teórico de Lorenz [LOR63] que é descrito por três variáveis de estado, tem-se $m = 3$ (ver Figura 2.7). No entanto, para séries temporais experimentais, o número de variáveis de estado do sistema gerador é inicialmente desconhecido, ou seja, não se conhece o valor de m *a priori*. Nesse contexto, uma estimativa do valor de m pode ser obtida com o auxílio da relação D_C

versus m da série; isso porque, para sistemas caóticos determinísticos, a D_C atinge um valor de saturação a partir do momento em que a dimensão m do espaço de fases permite o desdobramento completo do seu atrator [CAM05; CAM07; GAN01]. Tomando como exemplo uma série $x(n)$ que descreve a dinâmica da variável de estado x do modelo de Lorenz, a Figura 2.8-a mostra que existe uma convergência na relação D_C *versus* m daquela série a partir de $m = 3$, indicando assim que o sistema gerador possui três variáveis de estado e que, portanto, seu atrator pode ser reconstruído adequadamente em um espaço de fases tridimensional (Figura 2.8-b). Ainda, o valor de D_C correspondente a $m = 3$ pode ser assumido como a dimensão de correlação do atrator do sistema. A Figura 2.8-c mostra que a imersão do atrator de Lorenz em apenas duas dimensões também é possível, sendo porém inadequada pois a "falta de espaço" para o desdobramento das órbitas resulta em falsos cruzamentos de algumas das trajetórias.

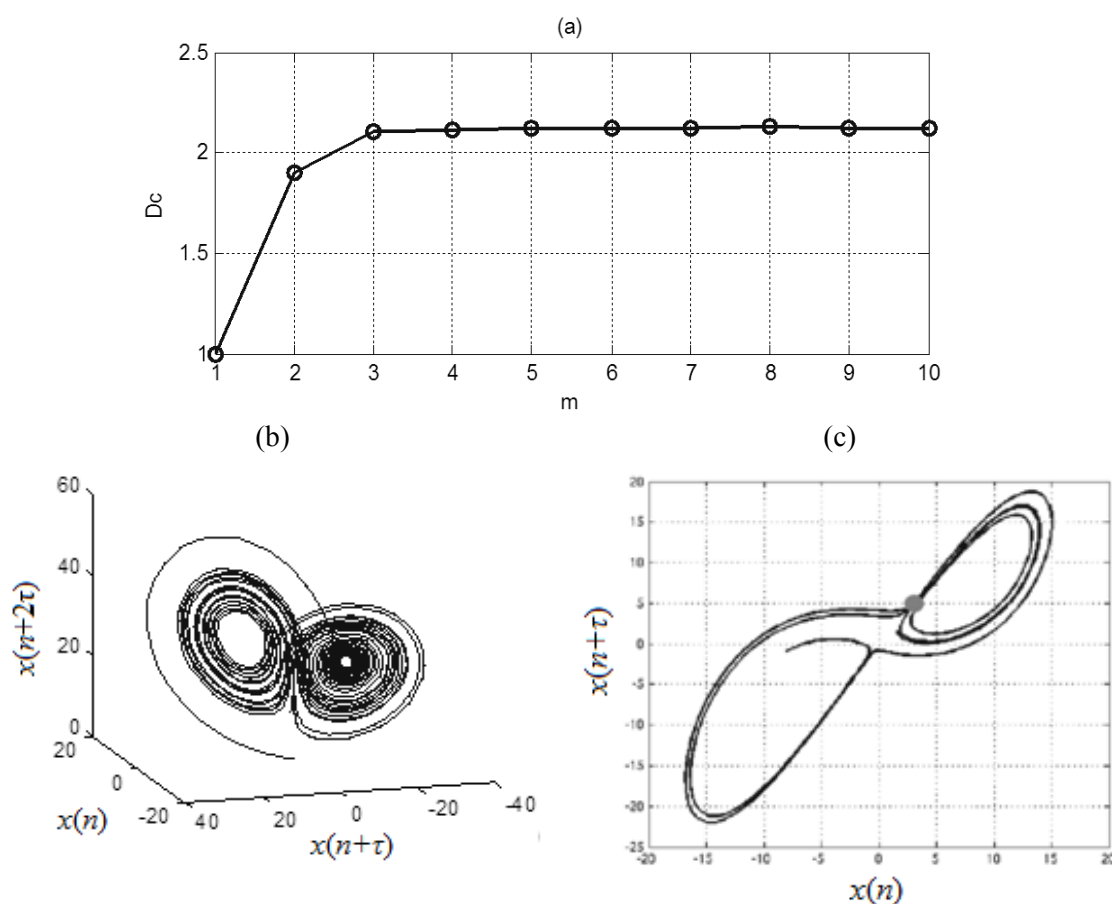


Figura 2.8 - (a) Curva D_C *versus* m para o modelo de Lorenz em que D_C atinge um valor de saturação a partir de $m = 3$. Ilustração do atrator do modelo em um espaço de fases (b) tridimensional e (c) bidimensional, em que o ponto cinza denota um falso cruzamento de duas trajetórias. Nos dois casos, foi utilizado $\tau = 6$ [CAM05].

Fonte: adaptado de Campanharo [CAM07].

Assim, conforme explicado para o sistema de Lorenz, é possível estimar os parâmetros m e D_C de um sistema dinâmico qualquer descrito por uma série temporal a partir da relação D_C versus m daquela série. No entanto, para a reconstrução do atrator de sistemas descritos por séries temporais experimentais, Takens [TAK81] sugere que a dimensão m do espaço de fases obedeça o seguinte critério:

$$m \geq 2D_C + 1 \quad (2.17)$$

A dimensão de correlação (D_C) introduzida por Grassberger e Procaccia [GRA83] pode ser representada pela seguinte expressão:

$$D_C = \lim_{r \rightarrow 0} \frac{\log_{10} [C_m(r)]}{\log_{10} [r]}, \quad (2.18)$$

onde $C_m(r)$ é a soma de correlação dada por:

$$C_m(r) = \frac{1}{N^2} \sum_{i,j=1, i \neq j}^N \theta(r - |X_i - X_j|), \quad (2.19)$$

sendo o índice m de $C_m(r)$ a dimensão de imersão do atrator reconstruído, N o número total de vetores de estado (ou pontos) do atrator e $\theta(z)$ a função degrau unitário, na qual $\theta(z) = 1$ para $z \geq 0$ e $\theta(z) = 0$ caso contrário. X_i e X_j são vetores de estado reconstruídos, enquanto que $|X_i - X_j|$ representa a distância Euclidiana entre tais vetores. O parâmetro r corresponde ao raio de uma esfera de busca por pares de vetores X_i, X_j dentro do espaço de fases.

É importante perceber na equação (2.18) que, para pequenos valores de r , a soma de correlação comporta-se de acordo com a seguinte lei de potência: $C_m(r) \propto r^{D_C}$ e, dessa forma, o cálculo da D_C pode ser realizado por meio de um gráfico $C_m(r)$ versus r em escala logarítmica [CAM07]. Entretanto, de acordo com Roerdink et al. [ROE06], o número de pontos do atrator considerados no cálculo da soma de correlação é muito baixo quando r tende a zero na equação (2.18), podendo interferir na precisão da estimativa da D_C . Diante disso, aqueles autores propõem que a dimensão de correlação seja estimada dentro de uma região de $C_m(r)$ delimitada entre r_{min} e r_{max} , raios em que

0.5% e 75% de pares de pontos são respectivamente englobados pela esfera de busca da soma de correlação [ROE06]. Neste contexto, a D_C pode ser obtida da seguinte forma:

$$D_C = \frac{\log_{10} C_m(r_{max}) - \log_{10} C_m(r_{min})}{\log_{10}(r_{max}) - \log_{10}(r_{min})} \quad (2.20)$$

A Figura 2.9 exibe a curva $C_m(r)$ versus r em escala logarítmica para o atrator do modelo de Lorenz reconstruído em um espaço de fases m -dimensional com $m = 1, 2, \dots, 10$. É possível perceber que, a partir de $m = 3$ até $m = 10$, a soma de correlação apresenta inclinação praticamente constante dentro da região delimitada pelas linhas tracejadas, as quais representam os limites inferior (0.5% de pares de pontos) e superior (75% de pares de pontos) sugeridos no estudo de Roerdink et al. [ROE06] para a estimativa da D_C . Em outras palavras, a dimensão de correlação atinge um valor de saturação quando $m = 3$ pois, como discutido previamente, esta é a dimensionalidade do espaço de fases que garante a representação adequada do atrator de Lorenz.

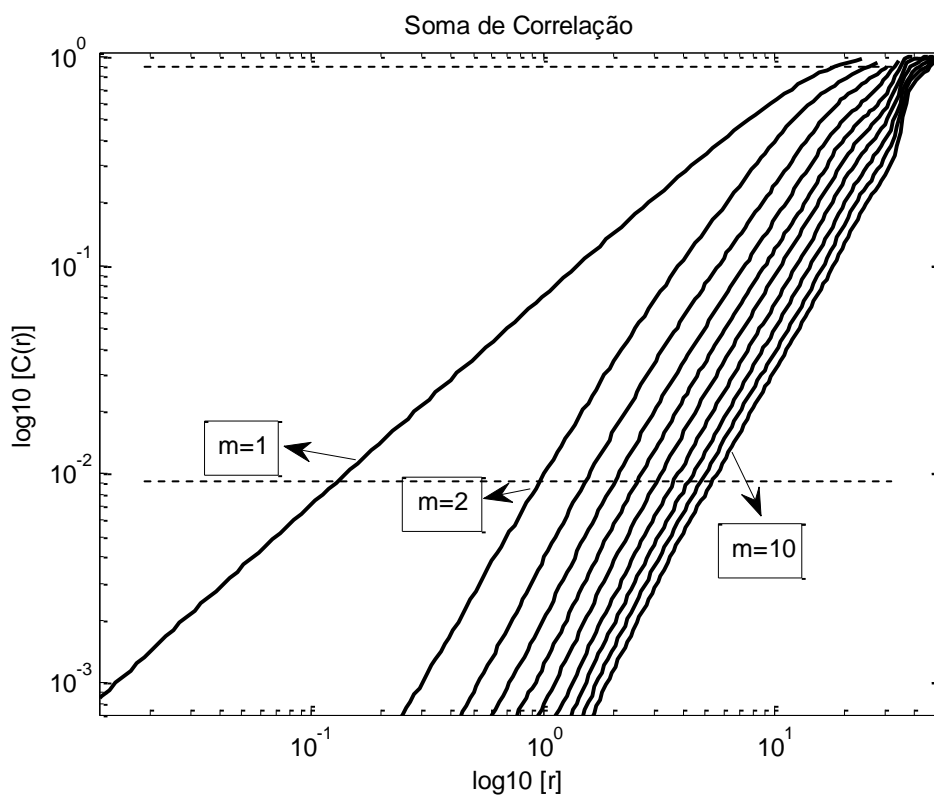


Figura 2.9 - Soma de correlação em escala logarítmica para o atrator do modelo de Lorenz reconstruído com $m = 1, 2, \dots, 10$. As linhas tracejadas delimitam a região da curva utilizada para a estimativa da dimensão de correlação, a qual considera desde 0.5% até 75% de pares de pontos do atrator reconstruído.

Fonte: o autor.

Máximo Expoente de Lyapunov

De acordo com Wolf et al. [WOL85], os chamados *expoentes de Lyapunov* são parâmetros capazes de descrever a taxa média em que trajetórias inicialmente vizinhas de um atrator se separam (expoente positivo) ou se aproximam (expoente negativo), em escala exponencial, dentro do espaço de fases ao longo da evolução temporal do sistema em estudo. Em sistemas caóticos, a alta sensibilidade às condições iniciais implica na existência de pelo menos um expoente de Lyapunov positivo [CAM05; CAM07]. No cenário do controle do equilíbrio, o maior (ou máximo) expoente de Lyapunov (λ_{max}) é utilizado para avaliar a estabilidade local dos mecanismos de controle postural, ou seja, para avaliar a sensibilidade destes mecanismos a pequenas perturbações locais [ROE06].

Segundo Campanharo et al. [CAM05], diversos métodos foram propostos na literatura para o cálculo dos expoentes de Lyapunov. Dentre eles, destaca-se o algoritmo proposto por Wolf et al. [WOL85] que permite estimar — caso existam — os expoentes positivos a partir de uma série temporal experimental. Alternativamente, é possível utilizar também o método introduzido por Sato et al. [SAT87] e Kurths [KUR87] que, similarmente ao algoritmo de Wolf, prevê a estimativa de λ_{max} a partir da análise do chamado *erro de previsão (EP)*, o qual está relacionado com a taxa média de divergência de trajetórias vizinhas do atrator em função de uma distância L_0 de separação inicial entre aquelas trajetórias. Neste cenário, para uma série temporal caótica, a curva *EP versus L_0* pode ser segmentada em três fases distintas [PAR98] (Figura 2.10), sendo elas:

Fase I: representa um período inicial de transiente em que as órbitas do atrator estão convergindo em direção ao máximo expoente de Lyapunov. Esta fase é caracterizada por uma inclinação elevada na curva do erro de previsão;

Fase II: é caracterizada por uma taxa de divergência aproximadamente constante entre órbitas vizinhas do atrator. A inclinação de uma linearização deste segmento na curva *EP versus L_0* representa uma estimativa de λ_{max} ;

Fase III: representa um período em que a taxa de divergência entre órbitas vizinhas deixa de ser exponencial. Esta fase é caracterizada pela convergência da curva *EP versus L_0* para um valor de saturação;

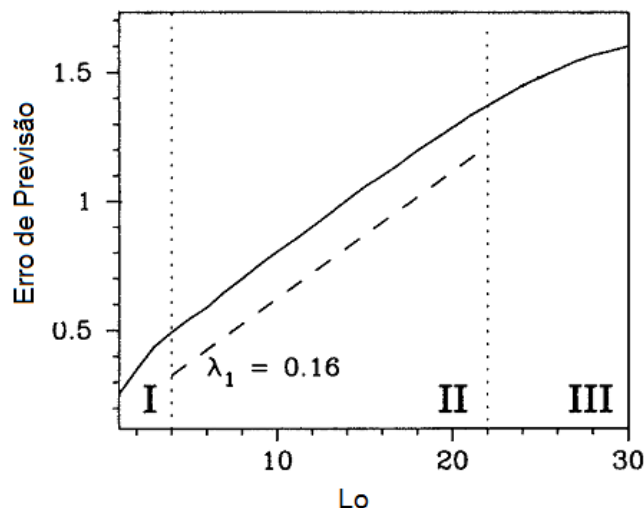


Figura 2.10 - Ilustração das três fases (delimitadas pelas linhas pontilhadas) que compõem a curva EP versus L_0 (linha sólida) de um atrator caótico. A inclinação de uma linearização do segmento da fase II (linha tracejada) representa uma estimativa do máximo expoente de Lyapunov do sistema.

Fonte: adaptado de Parlitz [PAR98].

Entropia Amostral

De maneira geral, a entropia é uma medida que avalia o *grau de regularidade* de uma série temporal, ou seja, ela mede o grau de surpresa (ou incerteza) de um sinal ao longo de sua evolução temporal [MOU05; ROE06; SIL10]. Mais especificamente, pequenos valores de entropia correspondem a um alto grau de regularidade e podem sugerir comportamento determinístico para os dados em estudo, ao passo que valores elevados de entropia representam alto grau de aleatoriedade, sugerindo assim natureza estocástica para os dados em questão.

No contexto da avaliação postural, as métricas mais utilizadas na análise da regularidade do sinal do CP são a *entropia aproximada* proposta por Pincus [PIN91] e a *entropia amostral* ($EntAm$) introduzida por Richman e Moorman [RIC00]. Vale ressaltar que a entropia amostral foi proposta com o intuito de suprir algumas limitações da entropia aproximada [SIL10] e, como consequência, ela tem apresentado resultados mais consistentes e também menor sensibilidade à determinadas características do sinal em estudo (e.g., comprimento) [CHE05]. De acordo com Donker et al. [DON08], sistemas fisiológicos doentes ou com idade avançada costumam produzir padrões mais regulares de variabilidade corporal (i.e., com menor entropia) quando comparados a sistemas saudáveis.

Dada uma série temporal de comprimento k , a entropia amostral calcula a probabilidade condicional de que uma sequência similar de comprimento M permanecerá similar após a inclusão de uma amostra vizinha (ou seja, para $M + 1$ amostras) dentro de uma tolerância Γ . Em outras palavras, o parâmetro Γ representa um limiar de distância utilizado para definir se os segmentos de comprimento M e $M + 1$ representam padrões similares [DON07; ROE06; SIL10]. De acordo com Borg e Laxaback [BOR10], para garantir uma boa estimativa da entropia amostral, é preciso que o sinal em estudo contenha entre 10^M e 20^M amostras. Diante disso, estudos na literatura sugerem o cálculo da $EntAm$ para $M = 2$ (e.g., [CHE06; KIR12]). Em relação ao limiar de distância, estudos sugerem que se considere uma tolerância de 15% do desvio padrão dos dados, ou seja, $\Gamma = 0.15$ tendo em vista uma série normalizada (i.e., com média nula e desvio padrão unitário) (e.g., [DUA08; COS05]).

Entropia Multiescala

Motivados pelo fato de que o cálculo da entropia amostral não considera as múltiplas escalas de tempo existentes em uma série temporal, Costa et al. [COS02; COS05] introduziram a chamada *entropia multiescala* ($EntME$) com o objetivo de medir o *grau de complexidade* de um sinal. Isso porque, segundo os autores, nem sempre o aumento no valor da entropia amostral reflete a elevação no grau de complexidade de uma série de dados, podendo ocorrer apenas pela única escala de tempo da série que está sendo considerada nos cálculos.

A entropia multiescala consiste em calcular a entropia amostral para um determinado sinal sobre diferentes escalas de tempo (Figura 2.12). Neste contexto, as chamadas *séries escaladas* são obtidas calculando-se a média do sinal original dentro de janelas consecutivas e não sobrepostas de comprimento η (Figura 2.11). Finalmente, para a comparação da entropia multiescala calculada em duas situações distintas (e.g., diferentes populações, como ilustrado na Figura 2.12), Costa et al. [COS07] sugerem que seja utilizada a área sob a curva $EntAm$ versus η , a qual pode ser estimada através da seguinte expressão:

$$EntME = \sum_{i=1}^k EntAm(i), \quad (2.21)$$

sendo k o valor do máximo fator de escala considerado.

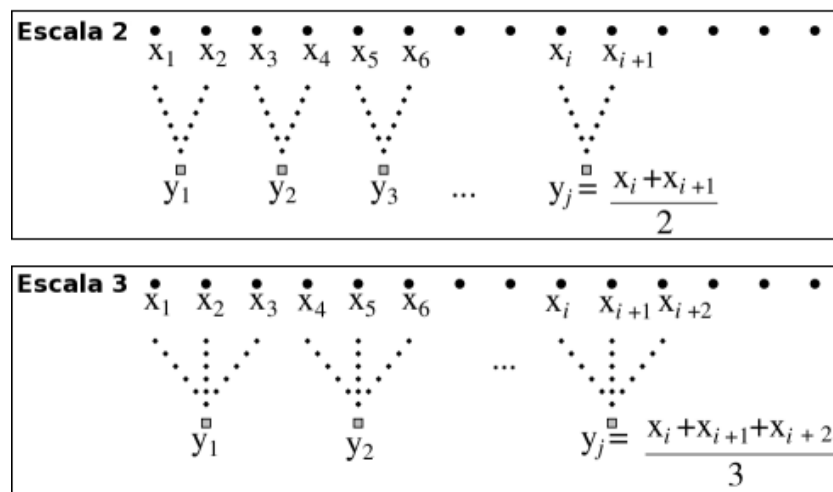


Figura 2.11 - Ilustração do processo de escalamento de uma série temporal original para $\eta = 2$ (painel superior) e $\eta = 3$ (painel inferior).

Fonte: adaptado de Silva [SIL10].

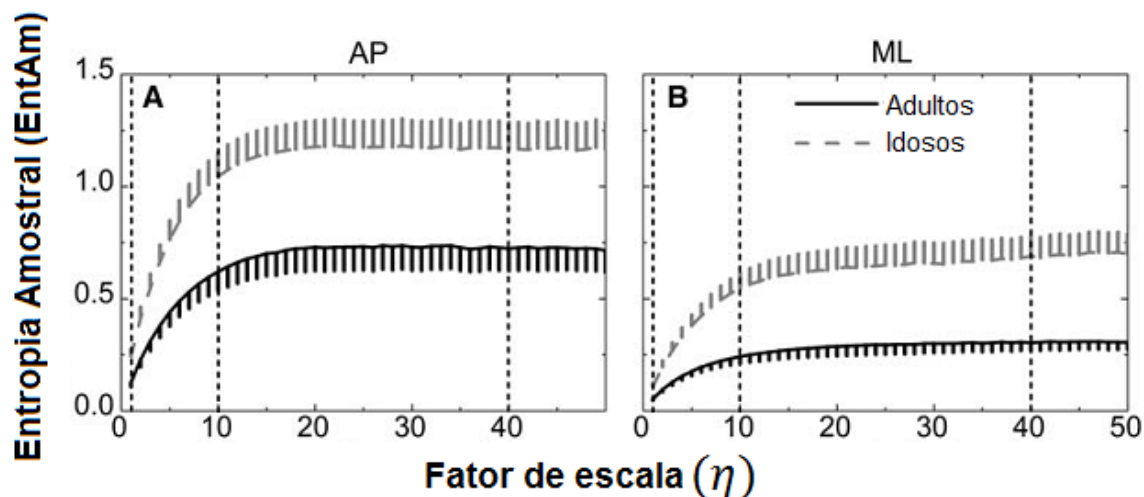


Figura 2.12 - Curvas médias da entropia multiescala para séries temporais do CP de 14 adultos e 14 idosos nas direções AP (esquerda) e ML (direita). Nos cálculos, foi utilizado $M = 2$ e $\Gamma = 0.2$.

Fonte: Duarte e Sternad [DUA08].

Expoentes de Escala

Da maneira geral, processos que exibem correlações de longa duração apresentam características similares sobre diferentes escalas de tempo (Figura 2.13), sendo assim denominados *processos auto-similares* ou *fractais* [DUA08]. Em relação aos mecanismos biológicos de controle postural, alguns estudos na literatura encontraram correlações de longa duração no sinal do CP para diferentes escalas de tempo (e.g., [COL93; DUA00b; DUA01; DUA08]), sugerindo assim que o SCP possui características de auto-similaridade (ou fractalidade).

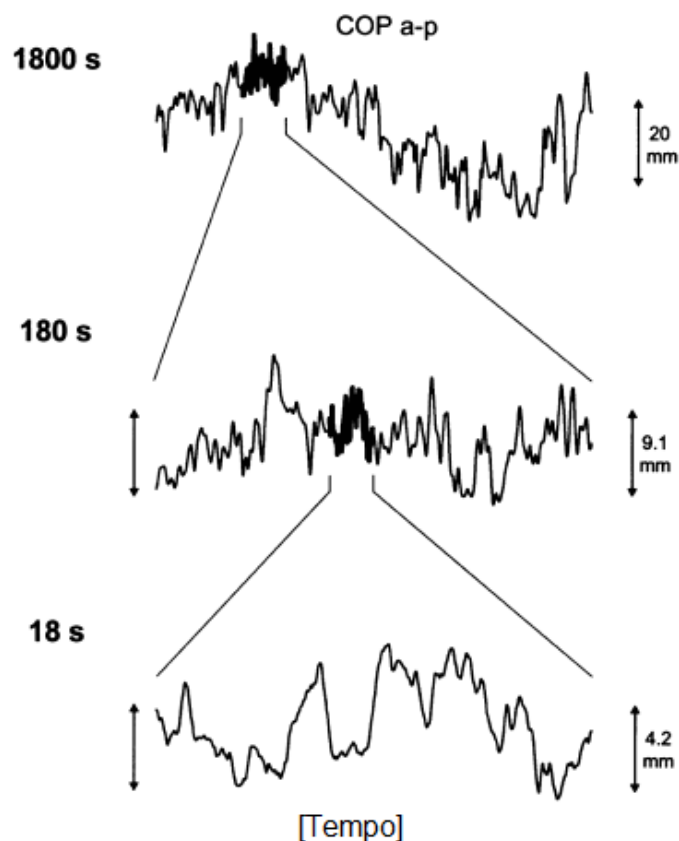


Figura 2.13 - Ilustração da auto-similaridade (ou fractalidade) em diferentes escalas de tempo presente em uma série temporal do CP na direção ântero-posterior.

Fonte: traduzido de Duarte e Zatsiorsky [DUA00b].

Diversas técnicas têm sido propostas na literatura para a investigação de correlações de longa duração a partir de séries temporais experimentais, assim como mostra a Tabela 2.3.

Tabela 2.3 - Relação de técnicas utilizadas na literatura para a avaliação de correlações de longa duração em séries temporais experimentais. A terceira coluna exhibe exemplos de estudos que utilizaram as técnicas em questão no contexto da avaliação do equilíbrio.

Técnica	Parâmetro resultante	Estudos relacionados
<i>Detrended Fluctuation Analysis</i> (DFA)	Expoente de escala (α)	Amoud et al. [AMO07], Delignières et al. [DEL03], Donker et al. [DON07], Duarte e Sternad [DUA08], Duarte e Zatsiorsky [DUA01], Kirchner et al. [KIR12], Roerdink et al. [ROE06]
<i>Scaled Windowed Variance</i> (SWV)	Expoente de Hurst estimado (\hat{H})	Kirchner et al. [KIR12], Kirchner [KIR13], Delignières et al. [DEL06]
<i>Stabilogram Diffusion Analysis</i> (SDA)	Expoente de Hurst estimado (\hat{H})	Amoud et al. [AMO07], Schiffman et al. [SCH05]
<i>Rescaled Range Analysis</i> (R/S)	Expoente de Hurst estimado (\hat{H})	Delignières et al. [DEL03], Duarte e Zatsiorsky [DUA00b]

Dentre as técnicas indicadas na Tabela 2.3, destaca-se o uso da DFA introduzida por Peng et al. [PEN95] em estudos da área da posturografia. Esta técnica resulta em uma métrica denominada *expoente de escala* (α). Com exceção da DFA, todas as outras técnicas indicadas na Tabela 2.3 fornecem uma estimativa \hat{H} do chamado *expoente H de Hurst*. Ambos os expoentes α e H descrevem a correlação e a persistência de uma série temporal [RIB10; ROE06], estando relacionados pela seguinte expressão [DEL06; KIR13]:

$$H = \begin{cases} \alpha, & \text{se o processo é do tipo Gaussiano fracional} \\ \alpha - 1, & \text{se o processo é do tipo Browniano fracional,} \end{cases} \quad (2.22)$$

lembrando que tanto o ruído Gaussiano fracional (RGF) como também a sua função integrada, o ruído Browniano fracional (RBF), são processos fractais amplamente utilizados pela literatura para modelar sistemas reais que apresentam correlações de longa duração (consultar [DUA00a] para mais detalhes). A Figura 2.14 ilustra a faixa de valores que o expoente de escala pode assumir para processos de diferentes naturezas.

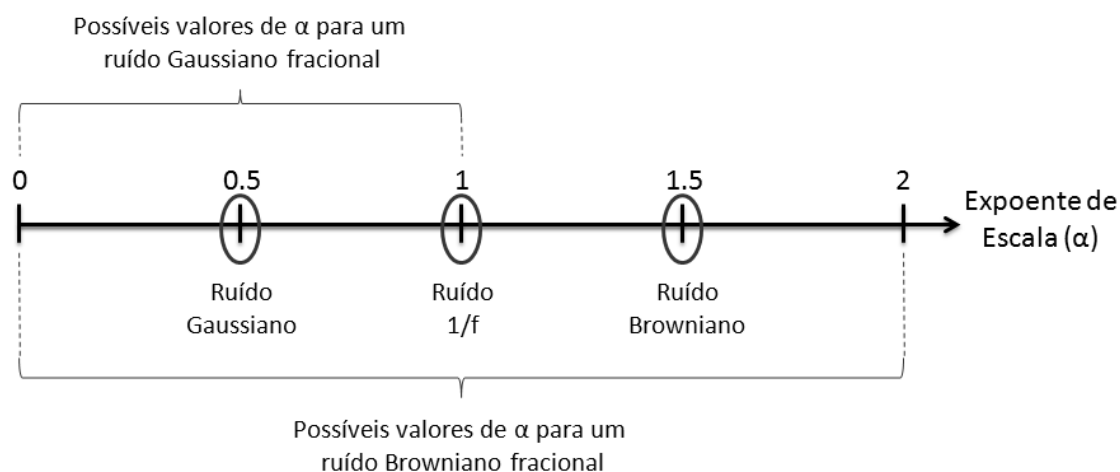


Figura 2.14 - Ilustração da faixa de valores assumidos pelo expoente de escala (α) para processos aleatórios fracionais do tipo Gaussiano e Browniano. São indicados também os valores de α para uma série de ruído Gaussiano puro ($\alpha = 0.5$), ruído $1/f$ puro ($\alpha = 1$) e ruído Browniano puro ($\alpha = 1.5$).

Fonte: o autor.

Em relação à interpretação dos expoentes, tomando como referência o expoente de Hurst, $H = 0.5$ sugere a presença de comportamento estocástico, ao passo que $H \neq 0.5$ sugere a existência de algum tipo de correlação nos dados em estudo. Quando

$0 \leq H < 0.5$, o processo em questão encontra-se em regime antipersistente, ou seja, o futuro e o passado estão negativamente correlacionados pois existe uma probabilidade maior que 50% de que um valor positivo seja precedido por um valor negativo e vice-versa. Em contrapartida, quando $0.5 < H \leq 1$, diz-se que o processo está em regime persistente pois apresenta uma tendência, ou seja, a probabilidade de um valor positivo ser precedido por outro valor positivo (e vice-versa) é maior que 50% [COL93; RIB10; ROE06].

Dada uma série temporal $x(n)$ de comprimento k (i.e., $x(n)$ com $n = 1, 2, \dots, k$), o primeiro passo da técnica DFA para a estimação de α consiste em remover a delimitação da série por meio do processo de integração [DUA08; KIR12; ROE06], tal como:

$$y(n) = \sum_{i=1}^n x(i), \quad n = 1, 2, \dots, k \quad (2.23)$$

Isso porque as trajetórias do CP estão sempre delimitadas dentro de uma área de suporte [ROE06]. Na sequência, a série integrada $y(n)$ é dividida em janelas não sobrepostas de comprimento w em que $w < k$. Em cada janela, a tendência local $y_w(n)$ é estimada por um ajuste linear através do método dos mínimos quadrados, e então é subtraída da série integrada. A flutuação da série integrada sem a componente de tendência pode ser estimada da seguinte forma [AMO07; DUA08; KIR13]:

$$F(w) = \sqrt{\frac{1}{k} \sum_{j=1}^k [y(j) - y_w(j)]^2} \quad (2.24)$$

Segundo Duarte e Sternad [DUA08], para uma série temporal que descreve um sistema biológico complexo, a curva $F(w)$ versus w obedece uma relação de potência, ou seja, $F(w) \propto w^\alpha$, sendo α o expoente de escala. Assim, o valor de α pode ser estimado por meio da inclinação de uma linearização de $F(w)$ versus w em escala logarítmica [AMO07; DUA08; KIR13], assim como ilustra a Figura 2.15 para uma série original do CP e também para a sua versão *surrogate* de amostras.

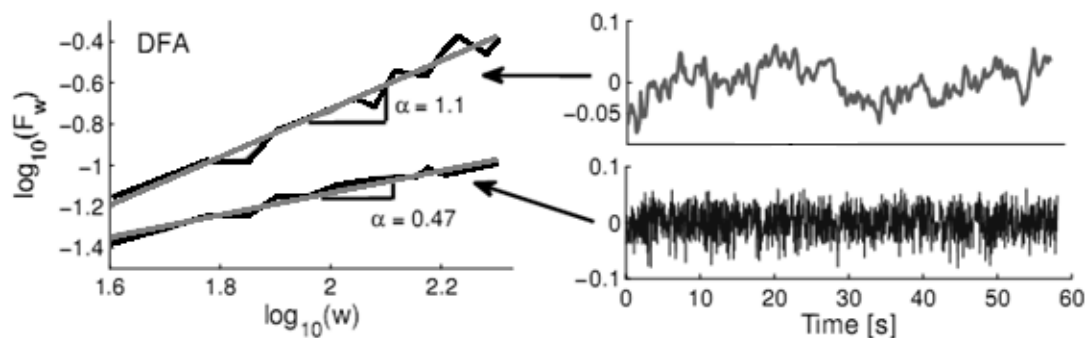


Figura 2.15 - À esquerda, são mostradas as linearizações (linhas cinzas sólidas) das curvas logarítmicas $F(w)$ versus w (linhas pretas sólidas) obtidas para uma série temporal de 60 segundos das trajetórias do CP (painel superior à direita) e sua versão *surrogate* de amostras (painel inferior à direita).

Fonte: Kirchner [KIR13].

No exemplo ilustrado pela Figura 2.15, observa-se $\alpha = 0.47$ para a série *surrogate* de amostras, valor muito próximo ao expoente resultante de uma série de ruído Gaussiano ($\alpha = 0.5$); isso porque, conforme será discutido na subseção 3.3, as observações de uma série *surrogate* de amostras não possuem correlação temporal alguma. Já para a série original do CP observa-se $\alpha = 1.1$, sugerindo que o comportamento do sinal pode ser representado por um RBF (ver Figura 2.14). Neste caso, seguindo a expressão (2.22), o expoente de Hurst é dado por $H = \alpha - 1 = 0.1$, sugerindo assim um regime anti-persistente para o conjunto de dados.

Conforme explicado anteriormente, para se estimar o expoente de Hurst a partir do expoente de escala, é preciso avaliar primeiramente se os dados em estudo são descritos mais adequadamente por um RGF ou por um RBF. Neste cenário, a literatura aponta uma limitação da técnica DFA. Isso porque, quando os dados são modelados por um RGF, a DFA fornece boas estimativas para o expoente de escala, mas tais estimativas perdem precisão quando os dados seguem um RBF, situação na qual Delignieres et al. [DEL06] consideraram mais apropriado estimar diretamente o expoente H de Hurst com o auxílio da técnica SWV (ver Tabela 2.3) introduzida por Cannon et al. [CAN97]. Além disso, o estudo de Bryce e Sprague [BRY12] apontou diversas críticas com relação ao uso da DFA, sugerindo que seja sempre utilizada alguma outra técnica auxiliar para a validação dos resultados.

Dada novamente uma série temporal $x(n)$ de comprimento k (i.e., $x(n)$ com $n = 1, 2, \dots, k$), o primeiro passo da técnica SWV para a estimação do expoente de Hurst consiste em dividir a série em janelas consecutivas e não sobrepostas de comprimento w . Na sequência, a tendência de cada janela é estimada e removida do

signal. Em seguida, calcula-se o desvio padrão para cada uma das w_j ($j = k/w$) janelas do sinal, assim como:

$$DP_{w_j} = \sqrt{\frac{1}{w-1} \sum_{i=l}^{w+(l-1)} [x(i) - \bar{x}]^2} \quad \text{para } l = (w(j-1)) + 1, \quad (2.25)$$

em que \bar{x} representa a média da série temporal em uma determinada janela w_j . Então, calcula-se a média do desvio padrão entre todas as janelas do sinal, ou seja:

$$S_w = \frac{1}{k/w} \sum_{j=1}^{k/w} DP_{w_j} \quad (2.26)$$

Por fim, o expoente de Hurst pode ser estimado através da inclinação de uma linearização da curva S_w versus w em escala logarítmica [CAN97].

Recurrence Quantification Analysis

Segundo Kantz e Schreiber [KAN04], a recorrência de um determinado sinal pode ser calculada com o auxílio de uma matriz bidimensional $M(i, j)$, tal como:

$$M(i, j) = \theta(r - |X_i - X_j|), \quad (2.27)$$

em que $\theta(\cdot)$ é a função degrau unitário de Heaviside, r é a distância de uma esfera de busca centrada em X_i , e X_i, X_j são vetores de estado reconstruídos em um espaço de fases m -dimensional e defasados entre si por um atraso τ . Dessa forma, se a órbita do atrator em um dado instante j retorna até a vizinhança onde estava no instante i (i.e., é englobada pela esfera de busca), então $M(i, j) = 1$; caso contrário, $M(i, j) = 0$ [GOM12; KAN04]. Os gráficos de recorrência (ou *recurrence plots*) introduzidos por Eckmann et al. [ECK87] consistem em plotar a matriz $M(i, j)$ com diferentes cores para os valores binários, como por exemplo, cor preta para $M(i, j) = 1$ e cor branca para $M(i, j) = 0$ (Figura 2.16).

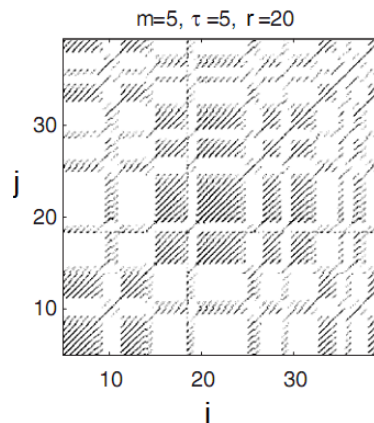


Figura 2.16 - Gráfico de recorrência com $r = 20$ para a série temporal $x(n)$ do modelo de Lorenz em regime caótico. Para a reconstrução do atrator, utilizou-se dimensão de imersão $m = 5$ e atraso temporal $\tau = 5$.

Fonte: Marwan et al. [MAR07].

De acordo com Vieira et al. [VIE12], a análise dos gráficos de recorrência é subjetiva, podendo conduzir diferentes pesquisadores à conclusões divergentes. Diante disso, a quantificação destes gráficos foi inicialmente proposta na literatura por Zbilut e Webber Jr [ZBI92; WEB94] e posteriormente aprimorada por Marwan et al. [MAR02a; MAR03]. Neste contexto, a extração de métricas de quantificação a partir dos gráficos de recorrência ficou conhecida como Análise de Quantificação de Recorrência ou *Recurrente Quantification Analysis* (RQA).

Entre outros cenários, a técnica RQA tem sido empregada na área da posturografia com o objetivo de investigar determinadas propriedades dos mecanismos de controle postural. Mais especificamente, dentre todas as métricas oriundas desta técnica (consultar [MAR06] para mais informações), os estudos de avaliação do equilíbrio costumam considerar a *Taxa de Recorrência* (TR_{RQA}) e o *Determinismo* (DET_{RQA}) para investigar a regularidade das estruturas determinísticas derivadas do controle postural, bem como a *Entropia de Recorrência* (ENT_{RQA}) para examinar a complexidade daquelas estruturas (e.g., [NEG13; TAL13; SAM13]). Neste contexto, é importante esclarecer que a Taxa de Recorrência e que o Determinismo não medem diretamente a regularidade (ou previsibilidade) do sinal do CP como no caso da entropia amostral apresentada anteriormente, e sim das estruturas do gráfico de recorrência produzido a partir daquele sinal. O mesmo vale para a Entropia de Recorrência e para a entropia multiescala em relação à propriedade de complexidade.

Em relação às definições matemáticas, a Taxa de Recorrência (TR_{RQA}) é uma medida de densidade de pontos no gráfico de recorrência expressa por:

$$TR_{RQA} = \frac{1}{k^2} \sum_{i,j=1}^k M_{i,j}, \quad (2.28)$$

sendo k o número total de pontos (ou vetores de estado) do atrator reconstruído no espaço de fases. Por sua vez, o Determinismo (DET_{RQA}) é uma medida associada à taxa de linhas diagonais no gráfico de recorrência, contendo assim informações sobre o determinismo (ou a previsibilidade) do sistema em estudo. Isso porque, para sistemas determinísticos, o gráfico de recorrência exibe longas diagonais e poucos pontos isolados (ver Figura 2.16 para o sistema de Lorenz), ao passo que processos estocásticos produzem gráficos de recorrência com nenhuma ou então pequenas diagonais. O Determinismo pode ser estimado da seguinte forma:

$$DET_{RQA} = \frac{\sum_{l=l_{min}}^k l \cdot P(l)}{\sum_{i,j}^k M_{i,j}}, \quad (2.29)$$

ou seja, é a relação entre as estruturas diagonais de diferentes comprimentos l do gráfico de recorrência acumuladas no parâmetro $P(l)$ e o número total de pontos do gráfico. Para o cálculo do Determinismo, é necessário estipular um comprimento mínimo l_{min} para a contagem das diagonais. Quando $l_{min} = 1$, o Determinismo torna-se equivalente à Taxa de Recorrência.

Finalmente, o cálculo da Entropia de Recorrência (ENT_{RQA}) leva em consideração a distribuição de frequência dos comprimentos das linhas diagonais do gráfico de recorrência, assim como:

$$ENT_{RQA} = - \sum_{l=l_{min}}^k \xi(l) \ln [\xi(l)], \quad (2.30)$$

em que:

$$\xi(l) = \frac{P(l)}{\sum_{l=l_{min}}^k P(l)} \quad (2.31)$$

Assim como no caso do Determinismo, o cálculo da Entropia de Recorrência exige que se defina uma distância mínima (l_{min}) para a contagem das estruturas diagonais.

2.1.4. Estudos Correlatos

A posturografia tem sido utilizada para avaliar o controle do equilíbrio nas mais diversas situações, tal como em diferentes estados de saúde, características antropométricas, condição física, idade, tarefa motora, etc. Destaca-se porém a avaliação do equilíbrio na presença de acometimentos neurológicos cujas sequelas trazem comprometimentos motores e sensoriais ao indivíduo, tais como o acidente vascular encefálico (e.g., [GEN08; HAA04; GHO10; ROE06]), a paralisia cerebral (e.g., [DON08; DEF09]) e a esclerose múltipla (e.g., [NEG13]). Nestes estudos, os pesquisadores costumam comparar a habilidade individual de diferentes métricas posturográficas para descrever (ou diferenciar) os perfis de oscilação corporal produzidos por duas populações distintas (e.g., pacientes acometidos e indivíduos saudáveis pareados).

No estudo de Silva [SIL12], por exemplo, foram avaliados os efeitos de perturbações sensoriais durante o controle do equilíbrio utilizando somente atributos posturográficos globais. Aquela pesquisa contou com 19 adultos pós-AVE e 19 adultos saudáveis pareados, sendo que cada um dos voluntários foi submetido a três tarefas distintas de equilíbrio na postura ereta quieta. Então, para cada indivíduo, foram registradas séries temporais das trajetórias do CP durante 60 segundos em cada tarefa. Na sequência, para a avaliação do controle do equilíbrio, a autora utilizou cinco atributos globais, sendo eles: amplitude de deslocamento RMS, velocidade RMS, deslocamento da oscilação total, área do estatocinesiograma e frequência espectral média. Dentre os resultados, Silva [SIL12] constatou que algumas das variáveis consideradas exibiram valores estatisticamente superiores para o grupo de indivíduos saudáveis, os quais conhecidamente possuem maior habilidade em manter o equilíbrio de forma mais segura em comparação com pacientes pós-AVE.

Já no estudo de Roerdink et al. [ROE06], os mecanismos biológicos de controle postural foram avaliados somente através de atributos posturográficos estruturais. Naquele trabalho, um grupo de adultos saudáveis e outro grupo de pacientes adultos pós-AVE foram submetidos a três tarefas de equilíbrio motor na postura ereta quieta. Então, as trajetórias do CP foram capturadas durante 30 segundos para cada voluntário em cada uma das tarefas. Na sequência, a avaliação do controle do equilíbrio foi realizada por meio de quatro atributos globais, a saber, a dimensão de correlação, o

máximo expoente de Lyapunov, a entropia amostral e o expoente de Hurst. Dentre os resultados, os autores observaram maior dimensionalidade (i.e., mais graus de liberdade ativos) e menor entropia (i.e., maior grau de regularidade) nas dinâmicas posturais dos pacientes, sendo possível diferenciar os grupos com significância estatística em apenas algumas das tarefas motoras. Diante do exposto, os autores concluíram que, para lidar com as limitações motoras e sensoriais decorrentes do AVE, os pacientes recrutaram mecanismos de controle adicionais, o que justificaria o número elevado de graus de liberdade em relação aos indivíduos saudáveis. Ademais, conclui-se que o maior grau de regularidade presenciado nas trajetórias do CP dos pacientes deve-se ao maior envolvimento cognitivo daqueles indivíduos com o SCP durante a manutenção do equilíbrio.

No trabalho de Han et al. [HAN05], foi recrutado um grupo de adultos jovens saudáveis e outro grupo de indivíduos idosos com histórico de desequilíbrio e quedas frequentes. Aquelas populações foram submetidas a quatro tarefas distintas de controle do equilíbrio na postura ereta quieta, situações em que tiveram seus deslocamentos do CP registrados durante 20 segundos. Então, os autores utilizaram um único atributo global denominado *sway path* em conjunto com um único atributo estrutural, a dimensão de correlação, com o intuito de diferenciar as populações em estudo. Dentre os resultados, conclui-se que o *sway path* é um bom indicador do nível da dificuldade da tarefa motora, ao passo que a dimensão de correlação descreveu mais adequadamente a eficiência das estratégias posturais adotadas pelos indivíduos durante as diferentes condições de equilíbrio.

O estudo de Ghomashchi et al. [GHO10] reuniu um grupo de adultos jovens saudáveis e outro grupo de pacientes adultos pós-AVE com tempo de lesão inferior a um ano, ambos submetidos a uma única tarefa de equilíbrio na postura ereta quieta em que capturou-se as oscilações do CP durante 20 segundos para cada voluntário. Primeiramente, aqueles autores motivaram-se a investigar a natureza dos dados posturográficos adquiridos quanto à determinística ou estocástica, utilizando para isso a técnica das séries *surrogate* (vide subseção 3.3). Como resultado, obtiveram evidências de que os dados originais não provêm de processos estocásticos. Na sequência, foi comparada a utilização da dimensão de correlação frente ao desvio padrão para diferenciar as populações em estudo. Após analisar os resultados, os autores concluíram que o desvio padrão foi capaz de diferenciar os grupos, mas incapaz de fornecer

informações sobre a estrutura das oscilações posturais, o que só foi possível com o auxílio da dimensão de correlação. Neste cenário, ao contrário dos achados supracitados do estudo similar de Roerdink et al. [ROE06], Ghomashchi et al. [GHO10] encontraram maior dimensionalidade para as dinâmicas posturais dos indivíduos saudáveis.

Na pesquisa de Deffeyes et al. [DEF09], foram avaliadas as estratégias de equilíbrio de crianças na postura sentada. Para tal, foi recrutado um grupo de crianças com desenvolvimento normal e outro grupo de crianças com atrasos no desenvolvimento, todas diagnosticadas como detentoras ou propensas ao desenvolvimento de paralisia cerebral. Então, foram registradas as oscilações do CP durante 10 segundos para cada um dos indivíduos de pesquisa. Como métricas de avaliação postural, foram utilizados quatro atributos globais, a saber, a amplitude, a amplitude RMS, o *sway path* e a área do estatocinesiograma, em conjunto com outros três atributos estruturais, sendo eles a entropia aproximada, o máximo expoente de Lyapunov e a dimensão de correlação. Dentre os resultados, o máximo expoente de Lyapunov foi o único parâmetro capaz de diferenciar estatisticamente as populações em estudo, estando as crianças saudáveis associadas a expoentes mais elevados, ou seja, apresentando mecanismos de controle postural mais sensíveis à pequenas perturbações locais. Diante disso, os autores concluíram que as crianças saudáveis repetiram com pouca frequência padrões similares de equilíbrio durante as tarefas posturais em comparação com as crianças com atrasos no desenvolvimento.

No trabalho de Madeleine et al. [MAD11], os autores estudaram o controle postural de indivíduos com traumas na coluna cervical. Para este propósito, foi recrutado um grupo de adultos saudáveis e outro grupo de pacientes adultos traumáticos, ambos submetidos a três tarefas distintas de equilíbrio na postura ereta quieta, situações em que foram coletadas séries de 45 segundos das trajetórias do CP. Na sequência, a partir das séries, foram calculados dois atributos posturográficos globais, a saber, o desvio padrão e o coeficiente de variação, e também outros dois atributos estruturais, sendo eles a entropia amostral e a dimensão de correlação. Após as devidas análises, os autores verificaram menor dimensionalidade nas dinâmicas posturais dos pacientes. Além disso, os dois grupos exibiram aumento na entropia amostral (i.e., diminuição na regularidade das trajetórias do CP) para as tarefas motoras cognitiva e sensorialmente mais complexas.

No estudo de Negahban et al. [NEG13], os autores avaliaram o controle do equilíbrio na presença da esclerose múltipla, uma afecção neurológica degenerativa. Para isso, foi recrutado um grupo de 23 pacientes acometidos e um outro grupo de 23 pacientes saudáveis, todos submetidos a três tarefas distintas de equilíbrio dentro das quais foram coletadas séries temporais das trajetórias do CP. Durante as análises, foi empregada a técnica RQA nos dados posturográficos para a extração dos seguintes atributos estruturais: TR_{RQA} , DET_{RQA} e ENT_{RQA} . Após as análises, os autores concluíram que o balanço postural dos dois grupos tornou-se menos regular e complexo em estruturas determinísticas a medida em que a tarefa de equilíbrio tornava-se mais difícil.

2.2. Classificação do Equilíbrio Postural

Conforme mencionado previamente no Capítulo 01, o principal objetivo desta pesquisa consiste em avaliar a relevância de atributos posturográficos globais e estruturais para, com o auxílio de algoritmos de Aprendizagem de Máquina, diferenciar os perfis de oscilação postural produzidos por indivíduos saudáveis e pós-AVE. Para isso, serão considerados neste estudo alguns dos métodos (ou modelos) de classificação mais populares na área de Mineração de Dados, bem como um método com menor difusão em problemas da área da saúde, a saber, a Programação Genética (PG).

2.2.1. Principais Métodos de Classificação

Na tarefa de classificação, diz-se que a aprendizagem é realizada de maneira *supervisionada*. Isso porque o modelo matemático de previsão de novas instâncias é construído sob a supervisão de instâncias já "rotuladas", ou seja, instâncias cuja classe é previamente conhecida [ARA11; HAN11]. A seguir, serão sucintamente apresentados alguns dos métodos de classificação mais populares na área de Mineração de Dados (consultar [GOR11; HAN11] para informações mais detalhadas). De maneira geral, cada um dos modelos a seguir se difere dos demais em relação ao procedimento de construção do modelo matemático de previsão.

(i) *Árvore de Decisão (AD)*: com este método, o modelo de previsão é construído recursivamente por meio da estratégia dividir-para-conquistar, na qual um problema complexo é segmentado em problemas menores [FUC08]. Como o próprio nome sugere, a representação do modelo de previsão dá-se na forma de uma árvore de decisão constituída por nós e folhas. Os nós da árvore representam testes sobre determinados atributos das instâncias, ao passo que as folhas indicam as classes nas quais as instâncias podem ser categorizadas [HAN11; GOR11]. Neste contexto, a inferência de uma classe para uma nova instância é um processo indutivo, ou seja, a árvore induz o usuário até uma determinada classe (ou nó-folha) através da avaliação de certos atributos (ou nós-teste), motivo pelo qual a AD propicia uma aprendizagem indutiva [FUC08; GOR11]. Devido à facilidade de interpretação, os modelos de classificação em forma de AD são amplamente empregados em diversos segmentos, tal como para o auxílio ao diagnóstico médico (e.g., [MOR12]) e também para previsões financeiras no mercado de ações (e.g., [CHA11]), entre outros.

(ii) *Naive Bayes (NB)*: baseado no teorema da probabilidade condicional de Bayes [FIN74], o NB é um algoritmo estatístico de aprendizagem Bayesiana que parte da seguinte suposição simplificadora: todos os atributos das instâncias são condicionalmente independentes, ou seja, eles não possuem nenhuma correlação entre si [HAN11]. O modelo matemático de previsão construído pelo NB consiste em um conjunto de probabilidades *a priori*, estimadas a partir da frequência de ocorrência dos M atributos e das N classes presentes no conjunto de instâncias de treinamento. Então, dada uma nova instância, o modelo utiliza as probabilidades *a priori* para calcular a probabilidade condicional daquela nova instância pertencer a cada uma das N classes do problema em estudo. Por fim, a instância será rotulada com a classe de maior probabilidade condicional.

(iii) *Máquinas de Vetor de Suporte (SVM)*: proposta na década de 1990 por Vladimir Vapnik [VAP95], a técnica SVM é baseada na teoria de aprendizagem estatística [GOR11] e se destaca pela sua sólida fundamentação matemática, assim como pelo elevado desempenho alcançado em aplicações práticas [CAM09; SEM02]. A partir de um conjunto de exemplos de treinamento, o SVM busca separar as classes através de um hiperplano linear ótimo. No entanto, quando as classes não são linearmente separáveis, o SVM utiliza uma estratégia de separação conhecida como *soft*

margin ou *margem de separação flexível*, em que são tolerados ruídos (ou *outliers*) e também erros de classificação [SAN02]. Alternativamente à estratégia *soft margin*, o SVM adota também outra estratégia que é baseada nas chamadas *funções kernel*, as quais são utilizadas para mapear as instâncias de treinamento em um espaço de mais alta dimensão, no qual elas possam ser separadas por um hiperplano linear [GOR11; HAN11]. Finalmente, após formular o modelo de previsão, o processo de classificação com o SVM consiste em verificar de que lado da fronteira de decisão encontra-se uma nova instância e, então, classificá-la com a classe predominante daquele lado do hiperplano.

(iv) *Redes Neurais Artificiais* (RNA): uma RNA utiliza unidades de processamento cuja estrutura e funcionamento se baseiam nos neurônios do sistema nervoso humano [CAM09; NET06]. Para a tarefa de classificação, um dos algoritmos de aprendizagem supervisionada baseados em RNA mais utilizados é o de retropropagação (*backpropagation*). Primeiramente, este algoritmo cria uma rede de neurônios — conectados entre si por meio de *sinapses* — que apresenta, no mínimo, uma camada de entrada e outra de saída. Normalmente, uma RNA apresenta também ao menos uma camada intermediária (ou oculta) e, neste caso, ela é conhecida como *perceptron de múltiplas camadas* (ou *multilayer perceptron*) [OLI05]. A cada sinapse da rede é atribuído inicialmente um valor aleatório de peso. O objetivo de cada neurônio é, a partir das entradas ponderadas pelos pesos sinápticos, calcular o valor de uma determinada função matemática conhecida como *função de ativação*. Então, na fase de treinamento, os exemplos são apresentados à rede na camada de entrada e, após computados os valores de saída, o erro entre o valor obtido e o valor esperado é retropropagado através da rede e utilizado para recalculer os pesos sinápticos, ajustando assim o modelo aos exemplos de treinamento [HAN11]. A principal desvantagem das redes neurais reside na forma em que o conhecimento é representado. Isso porque este método cria um modelo "caixa-preta", ou seja, a relação entrada-saída não é facilmente interpretável por pessoas, ao contrário dos modelos em forma de árvore de decisão [HAN11].

(v) *k-Vizinhos Mais Próximos* (*k*-NN): este método supervisionado de classificação é baseado em um tipo de aprendizagem conhecido como *aprendizado tardio/preguiçoso* (ou *lazy learning*) em que nenhum modelo de previsão é treinado. Ao

invés disso, os exemplos de treinamento são simplesmente armazenados, e o aprendizado somente ocorre quando solicita-se a categorização de uma nova instância [CAM09; HAN11]. Dentre os métodos de aprendizado tardio, destaca-se o k -NN (*k-Nearest Neighbours* ou *k-Vizinhos Mais Próximos*) no qual o processo de classificação de uma nova instância é baseado na análise de seus k vizinhos mais similares, sendo k um valor positivo e tipicamente pequeno [COS13; GOR11]. A métrica mais comum para avaliar a *similaridade* ou *proximidade* de duas instâncias é a distância Euclidiana, mas podem ser utilizadas também outras medidas, tais como as distâncias de Manhattan e de Minkowski [GOR11; HAN11]. O primeiro passo para a classificação com o algoritmo k -NN consiste em calcular as N distâncias entre o novo exemplo e os N exemplos de treinamento previamente armazenados. Então, a nova instância será classificada pelo voto majoritário de seus k vizinhos mais próximos, ou seja, a classe que possuir mais exemplos dentre os k vizinhos selecionados será atribuída à nova instância [COS13]. Se $k = 1$, por exemplo, então a nova instância será rotulada com a mesma classe que a do seu vizinho mais similar.

2.2.2. Programação Genética

Desenvolvida em 1989 por John Koza [KOZ89], a Programação Genética (PG) é um método computacional baseado na teoria da Evolução de Darwin que considera que os indivíduos mais bem adaptados às condições do ambiente em que se situam estão mais aptos a sobreviver. Além disso, a carga genética daqueles indivíduos é transmitida para seus descendentes, sofrendo progressivas mutações de geração em geração para melhor se ajustar às condições do atual ambiente. Dessa forma, com o passar de algumas gerações, observa-se uma evolução natural da população, ou seja, os indivíduos sobreviventes detêm características naturalmente selecionadas [CAB11b; SOU06]. No paradigma da PG, os indivíduos de uma população consistem em programas de computador, sendo compostos basicamente pelos seguintes conjuntos de símbolos [CAB11b; MAI08]:

Conjunto de funções (F): é formado por operadores aritméticos (+, -, *, /), funções matemáticas (e.g., seno, cosseno, logaritmo), operadores lógicos (e.g., E, OU, XOR), expressões lógicas (e.g., se-então, faça-enquanto, para), entre outros;

Conjunto de terminais (T): comporta todas as variáveis (ou atributos) do problema e também constantes numéricas.

Mais especificamente, os indivíduos na PG são representados por meio de uma árvore de sintaxe [SOU06], que consiste em uma combinação entre determinados elementos dos conjuntos F e T . Por exemplo, um indivíduo com a forma $x^2 + y$ consiste em uma dentre as diversas combinações possíveis entre os conjuntos $F=\{+, *\}$ e $T=\{x, y\}$, estando sua árvore de sintaxe ilustrada na Figura 2.17.

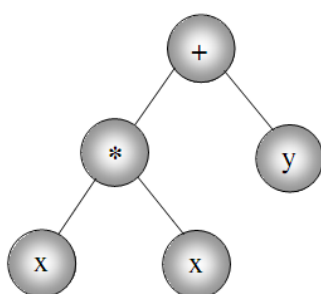


Figura 2.17 - Representação do indivíduo $x^2 + y$ em forma de árvore de sintaxe.
Fonte: Souza [SOU06].

O processo computacional de evolução natural realizado no algoritmo da PG pode ser brevemente descrito nas seguintes etapas [ROD02]:

-
- (i) Criar aleatoriamente uma população inicial de programas, os quais variam de forma e de tamanho;
 - (ii) Executar os seguintes passos até que um determinado *critério de parada* seja satisfeito:
 - (ii-a) Avaliar todos os programas da população atual através de uma função heurística de aptidão (ou *fitness*);
 - (ii-b) Selecionar os melhores programas de acordo com a função de *fitness* e, sobre eles;
 - (ii-c) Aplicar os operadores genéticos de *reprodução*, *mutação* e *cruzamento*;
 - (ii-d) Reinsere na população os novos programas modificados, originando assim uma *nova geração* de programas.
 - (iii) Retornar o programa que apresenta o melhor *fitness*, i.e., aquele que melhor se adaptou ao problema em estudo.
-

Conforme indicado na etapa (ii-c) acima, o processo de evolução natural da população envolve a aplicação de três operadores genéticos principais sobre determinados indivíduos, sendo eles [MAI08; ROD02; SOU06]:

(i) *Reprodução*: seleciona um indivíduo da população atual e o copia para a próxima geração sem nenhuma modificação. Dessa forma, vão existir duas versões do mesmo indivíduo na próxima geração;

(ii) *Mutação*: após selecionar um indivíduo, escolhe aleatoriamente uma sub-árvore de sua árvore de sintaxe (ver Figura 2.17) e a substitui por uma nova sub-árvore criada também de forma aleatória. Sendo assim, este operador não cria novos indivíduos, e sim apenas modifica determinados indivíduos da geração atual e os reinsere na nova geração;

(iii) *Cruzamento*: seleciona primeiramente dois indivíduos da população atual — denominados *pais* — em função de suas funções de aptidão. Em seguida, escolhe de maneira aleatória uma sub-árvore das árvores de sintaxe dos pais e, realizando permutações nestas sub-árvores, cria dois novos indivíduos conhecidos como *filhos* (Figura 2.18). Então, os indivíduos filhos são inseridos na nova geração.

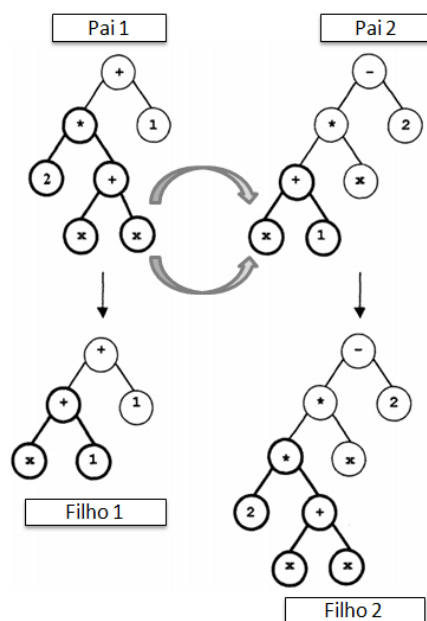


Figura 2.18 - Ilustração da operação de cruzamento entre dois programas quaisquer (pais) para a criação de dois novos indivíduos (filhos).

Fonte: Rodrigues [ROD02].

Conforme mencionado, o processo de evolução natural da PG ocorre até que um determinado critério de parada seja satisfeito. Neste cenário, um dos critérios mais empregados consiste em atingir um certo número máximo de gerações, ou seja, um número máximo de iterações do algoritmo. Outro critério bastante utilizado consiste em estipular um limiar a ser atingido pela função de *fitness*, ou seja, evoluir a população até que uma solução considerada satisfatória seja alcançada [KOZ92]. Em todos os casos, a função de *fitness* é utilizada para expressar o quão próximo cada um dos programas de uma determinada geração encontra-se da solução ideal do problema [ROD02], ou seja, para identificar os melhores e os piores indivíduos.

Diante do exposto, fica evidente que a execução do algoritmo da PG requer que sejam definidos *a priori* os seguintes parâmetros [SOU06; ROD02]:

(i) *Elementos dos conjuntos de funções e de terminais;*

(ii) *Tamanho da população inicial:* é o número de programas criados aleatoriamente no início do algoritmo e sua escolha deve ser criteriosa, visto que ela definirá o tamanho do espaço de busca por novos indivíduos. Neste cenário, uma pequena população restringirá o espaço de busca a poucas combinações, ao passo que uma população inicial de tamanho excessivo poderá provocar um esforço computacional exacerbado sem grandes vantagens nos resultados;

(iii) *Taxa de cruzamento:* representa a taxa de utilização do operador de cruzamento em cada geração. Se essa taxa for muito baixa, o algoritmo poderá gastar muito tempo para produzir um indivíduo bem adaptado ao problema. Porém, se a taxa de cruzamento for excessivamente alta, poderá ocorrer uma situação indesejada conhecida como *convergência prematura*, em que as características de alguns indivíduos bem adaptados mas não ótimos dominam rapidamente a população. Neste caso, ocorrerá a chamada *convergência para um máximo local*, situação em que elimina-se a possibilidade de buscar melhores soluções para o problema;

(iv) *Taxa de mutação:* de maneira similar ao parâmetro anterior, a taxa de mutação define o percentual de mutações a serem realizadas em cada geração. Altas taxas de mutação podem tornar aleatória a busca por melhores indivíduos;

(v) *Critério de parada.*

Entre outros cenários, a PG tem sido utilizada em problemas de classificação sendo que, nesta situação, cada programa (ou indivíduo) da população consiste em um modelo matemático de previsão distinto. Em cada geração, o desempenho dos classificadores é avaliado sobre um determinado conjunto de instâncias de treinamento com o auxílio da função de *fitness*. Então, aqueles mais bem adaptados ao problema têm maiores chances de serem reinseridos (copiados ou modificados) nas gerações seguintes [CAB11b]. Finalmente, o modelo de classificação resultante da PG será o programa com a melhor aptidão dentre todos aqueles da última geração.

A representação de um modelo de classificação construído com a PG pode ser realizada por meio da chamada *função discriminante* [ESP10], a qual consiste em uma equação matemática que representa a árvore de sintaxe do classificador (ver Figura 2.17). Dessa forma, a função discriminante envolverá uma combinação entre diferentes tipos de operadores e de terminais, sendo estes últimos os atributos das instâncias em estudo. Então, o processo de classificação de uma nova instância consiste em utilizar os valores dos atributos desta instância diretamente na função discriminante e avaliar o seu resultado. Por exemplo, para um problema binário (e.g., classes "sim" e "não"), um resultado positivo representaria a atribuição da classe "sim", enquanto que um resultado negativo sugeriria a atribuição da classe "não".

2.2.3. Treinamento e Avaliação dos Modelos de Classificação

Uma das técnicas utilizadas para o treinamento e avaliação de modelos de classificação é a *validação cruzada* (ou *cross-validation*). O primeiro passo desta técnica consiste em dividir aleatoriamente o conjunto de instâncias de treinamento em N partições (*folds*) mutuamente exclusivas e de tamanhos aproximadamente iguais. Na sequência, $N-1$ partições são utilizadas para treinar o modelo, enquanto a partição restante será utilizada para testá-lo, ou seja, para avaliar o seu desempenho. Este processo é repetido N vezes, de forma que os conjuntos de treino e de teste nunca se repetem. Então, o desempenho final do modelo é calculado através da média dos desempenhos alcançados durante as N etapas do processo supra elucidado [HAN11; OLI05].

Diversas métricas foram propostas na literatura de AM para a avaliação do desempenho de modelos de classificação. Dentre elas, destaca-se a *taxa de acerto* ou

acurácia, uma métrica de avaliação geral que considera o percentual de instâncias do conjunto de teste que são corretamente rotuladas pelo modelo. Entretanto, existem situações em que a taxa de acerto pode "mascarar" determinadas limitações do modelo de classificação, como por exemplo quando as classes das instâncias de treinamento estão desbalanceadas (i.e., mais exemplos de uma classe do que da outra). Nestes casos, é mais apropriado utilizar métricas de avaliação mais específicas, as quais consideram o desempenho do modelo para uma ou outra classe em particular, tal como a *sensibilidade*, a *especificidade* e a *medida-F*. Considerando um problema de duas classes, como por exemplo "sim" e "não", a sensibilidade e a especificidade correspondem ao percentual de instâncias de teste das classes "sim" e "não" corretamente rotuladas pelo modelo, respectivamente. Por sua vez, a medida-F leva em conta tanto a sensibilidade como também a especificidade do modelo, refletindo assim a sua capacidade de prever corretamente instâncias pertencentes às duas classes em questão, ou seja, sem ser tendencioso para uma ou outra classe em particular (consultar [GOR11; HAN11] para informações mais detalhadas).

2.2.4. Seleção de Atributos

A seleção de atributos, também conhecida como redução de dimensionalidade, consiste em extrair um subconjunto de atributos mais discriminantes a partir de uma lista de atributos candidatos, os quais podem ser utilizados com maior eficácia por algoritmos de aprendizagem de máquina para diferenciar as classes do problema em questão [JAI97]. Diversos trabalhos na literatura têm obtido resultados promissores com a seleção de atributos tendo em vista os mais variados modelos de classificação e os mais diversos contextos de pesquisa, tais como em estudos envolvendo microarranjos de DNA [BOR06; GAR11] e em problemas de identificação do equilíbrio postural [SAR13], entre outros.

Os diversos métodos de seleção de atributos propostos na literatura podem ser divididos em duas categorias principais: *wrapper* e *filter* [HAN11]. No primeiro caso, a estratégia consiste em buscar um subconjunto de atributos ótimo para um determinado modelo de classificação em particular, o qual irá variar de classificador para classificador [HSU08]. Mais especificamente, através da abordagem *wrapper*, um dado classificador é treinado para diferentes subconjuntos de atributos até que se atinja um

determinado critério de parada, tal como um limiar de desempenho considerado adequado [HAN11]. Já na abordagem *filter*, os métodos de redução de dimensionalidade visam avaliar a relevância/redundância de cada um dos atributos considerados na pesquisa de forma independente a qualquer modelo de classificação. Para isso, ao invés do desempenho de um dado classificador, os atributos são ranqueados por meio de outras métricas de "peso", tais como a informação mútua e a correlação e, por fim, seleciona-se um subconjunto contendo os atributos de mais alto posto a partir de um determinado limiar [DIN14].

Dentre os diversos métodos de redução de dimensionalidade baseados na abordagem *filter*, pode-se destacar o uso do CFS (*Correlation-based Feature Selection* ou Seleção de Características baseada em Correlação) proposto por Hall [HAL00] na literatura de AM. Este método utiliza a correlação para avaliar a relevância dos atributos considerados no problema, partindo da hipótese de que *"um bom subconjunto contém atributos altamente relacionados com a classe, e não correlacionados uns com os outros"*. O funcionamento do algoritmo CFS consiste em calcular uma matriz de correlação atributo-atributo e atributo-classe e, então, utilizar o *best first search* — um algoritmo de exploração de grafos — para encontrar o subconjunto de atributos mais promissor de acordo com a regra especificada. O critério de parada é atingido quando a expansão consecutiva de um número preestabelecido de atributos, normalmente cinco [HAL99], não produz nenhuma melhoria no mérito heurístico do subconjunto.

2.2.5. Estudos Correlatos

Conforme discutido previamente na subseção 2.1.4, os estudos da área da posturografia normalmente objetivam-se a diferenciar o controle do equilíbrio em situações distintas (e.g., diferentes populações ou diferentes tarefas motoras) através de comparações estatísticas entre determinados atributos posturográficos. Todavia, alguns outros estudos da área de Inteligência Artificial (IA) têm realizado esta tarefa com o auxílio de métodos de classificação. Tais estudos estão geralmente relacionados com o desenvolvimento de dispositivos de detecção de quedas em tempo real visando, entre outras finalidades, minimizar o risco de acidentes envolvendo pessoas idosas (e.g., [ALH14; CHO11; KER12; LIU10]) e até mesmo aprimorar o controle do equilíbrio de robôs bípedes (e.g., [KIM11]). Além disso, alguns outros estudos têm utilizado modelos

de classificação para identificar as estratégias posturais adotadas por indivíduos jovens saudáveis (e.g., [RAS09; SAR14]), idosos saudáveis (e.g., [RAS09]) e até mesmo por uma criança saudável em processo de crescimento [BOU10], cada um deles em tarefas específicas de equilíbrio motor.

No estudo de Saripalle et al. [SAR13], os autores avaliaram a acurácia de alguns classificadores durante a identificação das estratégias posturais adotadas por 14 jovens saudáveis em 11 tarefas distintas de equilíbrio. Para este propósito, foram primeiramente registrados os deslocamentos do CP destes voluntários por 40 segundos em cada condição e, a partir destes dados, foram calculados 23 atributos posturográficos globais. Visando aprimorar a acurácia dos modelos e também diminuir o custo computacional da classificação, foram selecionados apenas os atributos mais discriminantes de cada tarefa motora (ou classe) com o auxílio da técnica FSA (*Fast Simulated Annealing*). Na sequência, utilizando a técnica de validação cruzada com quatro partições, os autores treinaram e avaliaram os classificadores SVM com função *kernel* polinomial, SVM com função *kernel* Gaussiana, *k*-NN e LDA (*Linear Discriminant Analysis*) para a identificação das classes. Mais especificamente, cada classificador foi utilizado sobre cada um dos 11 conjuntos de atributos mais discriminantes. Dentre os resultados, do ponto de vista geral, o pior desempenho médio foi de 78.8% com o LDA, enquanto que o *k*-NN alcançou a melhor taxa média de acerto, 92.4%. Entretanto, para determinadas classes em particular, aquele estudo relatou taxa de acerto de 100% com os quatro classificadores. Dessa forma, para a identificação de estratégias posturais distintas com maior precisão, aqueles autores sugerem que sejam utilizados diferentes classificadores de forma conjunta.

No trabalho de Rasku [RAS09], o autor utilizou o *k*-NN para a classificação do equilíbrio motor em dois experimentos distintos, sendo eles: (i) para uma mesma população, utilizando as tarefas posturais como classes do problema e (ii) para uma mesma tarefa postural, utilizando diferentes populações como classes. Nos dois experimentos, mensurou-se por 60 segundos as trajetórias do CP de cada sujeito em cada condição, e extraiu-se um único atributo posturográfico global destes sinais: o deslocamento da oscilação total do estatocinesiograma. Nos dois casos, foi utilizado o *k*-NN variando o parâmetro *k* desde 1 até 19 e em passos de 2, ou seja, apenas os valores ímpares compreendidos naquela faixa. Isso porque, em ambos os experimentos, apenas duas classes estavam envolvidas no problema, de forma que utilização de vizinhos

ímpares evita empates no esquema de voto majoritário. O desempenho do classificador foi avaliado através da sensibilidade e da especificidade, ou seja, o percentual de instâncias de cada classe corretamente rotuladas. No experimento (i), um grupo de 41 adultos jovens saudáveis foi submetido a duas tarefas experimentais distintas na posição ereta quieta: (a) permanecer com os olhos abertos e (b) com os olhos fechados sobre uma plataforma de força, enquanto a mesma sofria perturbações rotacionais de até 6°. Dentre os resultados, o classificador apresentou maior eficiência para as tarefas (a) e (b) com $k = 19$ (94%) e $k = 15$ (89%), respectivamente. Já no experimento (ii), um grupo de 33 jovens e outro grupo de 33 idosos, ambos saudáveis, foram submetidos a outras duas tarefas de equilíbrio motor. Na primeira tarefa, o k -NN atingiu eficiência máxima de 74% para a classe de sujeitos saudáveis (com $k = 19$) e de 73% para a classe de idosos (com $k = 15$). Na segunda tarefa, para indivíduos saudáveis e idosos, a eficiência máxima foi de 92% ($k = 19$) e 74% ($k = 1$), respectivamente.

No estudo realizado por Moustakidis et al. [MOU08], os autores utilizaram métodos de classificação para identificar as condições de caminhada em função da velocidade e da carga carregada pelo sujeito. Para isso, 40 indivíduos foram recrutados e orientados a caminhar sobre uma plataforma de força em 11 condições distintas envolvendo a alteração da velocidade da caminhada, da carga carregada e também da forma de distribuição da carga carregada sobre o corpo dos voluntários (e.g., simétrica ou assimetricamente distribuída). Então, o sinal de força de reação do pé com o solo de cada voluntário foi mensurado de 1 à 3 segundos em cada condição experimental, sendo capturados no total 2800 padrões de balanço postural (instâncias) uniformemente distribuídas entre as 11 condições experimentais (classes). De cada instância, foram extraídos até 100 atributos posturográficos utilizando a transformada *wavelet*. Dentre as investigações do estudo, os autores avaliaram a acurácia de sete classificadores em função do número de atributos utilizados. Os classificadores utilizados foram o SVM com *kernel* Gaussiano e *kernel* ajustado via Algoritmo Genético, o k -NN ($k = 1$ e 2), o LVQ (*Learning Vector Quantization*), o Fuzzy k -NN, a Árvore de Decisão (algoritmo C4.5), o LDA (*Linear Discriminant Analysis*) e o MLH (*Maximum Likelihood*). Em relação aos resultados, o LDA apresentou a melhor acurácia (97.83%) apenas para 75 atributos utilizados. Já em todos os outros casos, a melhor taxa de acerto foi obtida com o SVM com *kernel* ajustado via Algoritmo Genético, com destaque para a acurácia de 98.21% quando utilizadas todas as 100 características.

No trabalho de Boughorbel et al. [BOU10], os autores avaliaram a taxa de acerto de quatro classificadores durante a identificação das estratégias posturais adotadas por uma criança menor de dois anos de idade. Para isso, uma única criança saudável foi recrutada para realizar o mesmo experimento em três etapas distintas: com (i) 14 meses, (ii) 16 meses e (iii) 19 meses de idade. O experimento consistiu em deixar a criança livre sobre um tapete plano sensível à pressão, o qual foi utilizado para mensurar por aproximadamente 10 minutos os padrões motores daquela criança. Os autores dividiram a superfície do tapete em diversos anéis concêntricos e, a cada quadro de 25 segundos, calculou-se a pressão exercida sobre a área de cada anel. Então, foram extraídos quatro atributos globais para descrever os padrões posturais, sendo eles a média, o desvio padrão, a curvatura e a obliquidade. Cada instância foi rotulada com uma entre seis classes associadas ao padrão do movimento executado pela criança, sendo elas: "sentado", "em pé", "deitado de costas", "deitado de bruços", "deitado de lado" e "outro" para posições que não se enquadravam em nenhuma das anteriores. O experimento contou com 3848 instâncias na etapa (i), 9065 instâncias na etapa (ii) e 8675 instâncias na etapa (iii). Na sequência, foram utilizados os classificadores linear, quadrático, k -NN e SVM com uma estratégia de combinação de voto para a classificação das instâncias. Dentre os resultados, os autores observaram erros médios de classificação de 0.64%, 1.28% e 1.53% para o experimento nas etapas (i), (ii) e (iii) respectivamente.

Na pesquisa realizada por Kerdegari et al. [KER12], visando auxiliar a população idosa na tarefa de manutenção do equilíbrio durante as atividades de vida diária, os autores propuseram um sistema para a detecção de quedas em tempo real. Entretanto, naquele estudo, não foram utilizadas séries temporais do CP e sim sinais de aceleração provenientes de acelerômetros posicionados em 50 voluntários jovens e saudáveis, todos instruídos a realizar 20 tarefas cotidianas sendo que onze delas envolviam a queda do sujeito. No total, aquele trabalho contou com 6962 instâncias e 29 atributos posturográficos relacionados basicamente à aceleração, à velocidade e à posição dos sujeitos durante as tarefas, ou seja, atributos de avaliação global. Então, para classificar os padrões de movimento em atividades que "oferecem" e que "não oferecem" risco de quedas, os autores utilizaram as implementações do perceptron multicamadas, Naive Bayes, árvore de decisão, SVM, ZeroR e OneR disponíveis no *software* WEKA. Nesse contexto, 75% e 25% das instâncias foram utilizadas respectivamente para treinar e avaliar os modelos de previsão. Dentre os resultados, a

rede neural apresentou a melhor taxa de acerto (90.15%), seguida pelo Naive Bayes (88.57%). A pior acurácia foi obtida com o ZeroR (66.49%).

Em [KIM11], os autores empregaram um método de classificação para detectar quedas em robôs bípedes em tempo real, pois consideraram este um aspecto indispensável para a inserção desta máquina em ambientes ocupados por seres humanos. Neste cenário, os autores submeteram um robô à duas condições experimentais: (i) equilíbrio estável (sem perturbação) e (ii) aplicação de força externa, capaz de causar a queda do robô. Então, para cada situação, foram extraídas cinco características: a aceleração do corpo do robô nos eixos x , y e z e as trajetórias do centro de pressão nas direções ântero-posterior (AP) e médio-lateral (ML). Cada instância foi associada a um rótulo (classe) referente à condição experimental, sendo ele "estável" ou "queda". Infelizmente, os autores daquele estudo reportaram apenas o número de instâncias da classe "estável", a saber, 200. Em seguida, a validação cruzada com 10 partições foi utilizada para treinar/avaliar o classificador SVM, sendo posteriormente observada uma acurácia de 98.5%. Sendo assim, aquele estudo utilizou o SVM para a detecção do estado do robô em tempo real ou, mais precisamente, a cada 0.01 segundos. Caso o modelo sugerisse uma queda, estratégias de *hardware* eram utilizadas para aumentar a base de suporte do robô com o intuito de mantê-lo em equilíbrio.

2.3. Conclusão do Capítulo

Duas perspectivas principais de análise do controle do equilíbrio foram estabelecidas na literatura, a global e a estrutural. Enquanto a parametrização global do sinal do CP visa capturar o tamanho (ou amplitude) dos padrões de oscilação postural, a parametrização estrutural procura descrever as estruturas temporais do balanço corporal. Visto que ambas as perspectivas supracitadas têm sido empregadas com sucesso em diferentes tipos de avaliação posturográfica, e também que não existe um consenso na literatura sobre qual métrica é a mais relevante para a descrição do equilíbrio em uma ou outra população, evidencia-se a importância de se considerar diferentes métricas (ou atributos) tanto de domínio global como também estrutural para uma melhor compreensão sobre os mecanismos de controle motor em estudo.

Os métodos de classificação são ferramentas computacionais de grande relevância para a descoberta de conhecimento em diversas áreas de pesquisa, inclusive

na área da posturografia. Porém, alguns cuidados devem ser tomados durante a construção e também durante a avaliação dos modelos matemáticos de previsão, principalmente quando existe um desbalanço entre as classes do problema em estudo, uma situação bastante comum em pesquisas envolvendo fenômenos biológicos [HAN11]. Além disso, em diversos problemas da área da saúde, um modelo de classificação pode servir também como um sistema de amparo ao diagnóstico, cenário no qual torna-se fundamental a facilidade de interpretação do modelo por profissionais da área da saúde. Em outras palavras, modelos do tipo "caixa branca" (e.g., funções discriminantes ou árvore de decisão) são preferíveis em relação aos chamados "caixa preta" (e.g., redes neurais ou SVM).

De maneira geral, os estudos da área da posturografia costumam avaliar a relevância individual de cada um dos parâmetros considerados na pesquisa para diferenciar as estratégias de equilíbrio adotadas em diferentes situações (e.g., diferentes estados de saúde, tarefas motoras, etc.). Para este propósito, são normalmente empregadas análises estatísticas tradicionais. Alternativamente, com o auxílio de métodos de classificação, é possível investigar a relevância conjunta dos parâmetros posturográficos durante a diferenciação supracitada. Isso porque os modelos matemáticos de previsão são formulados através de uma interação entre todas as variáveis consideradas na pesquisa, ou então por um subconjunto contendo as mais discriminantes dentre elas.

Capítulo 3

Metodologia

Neste capítulo, será explanado em detalhes o método delineado para a presente pesquisa cujo panorama geral é ilustrado na Figura 3.1 em forma de diagrama de blocos.

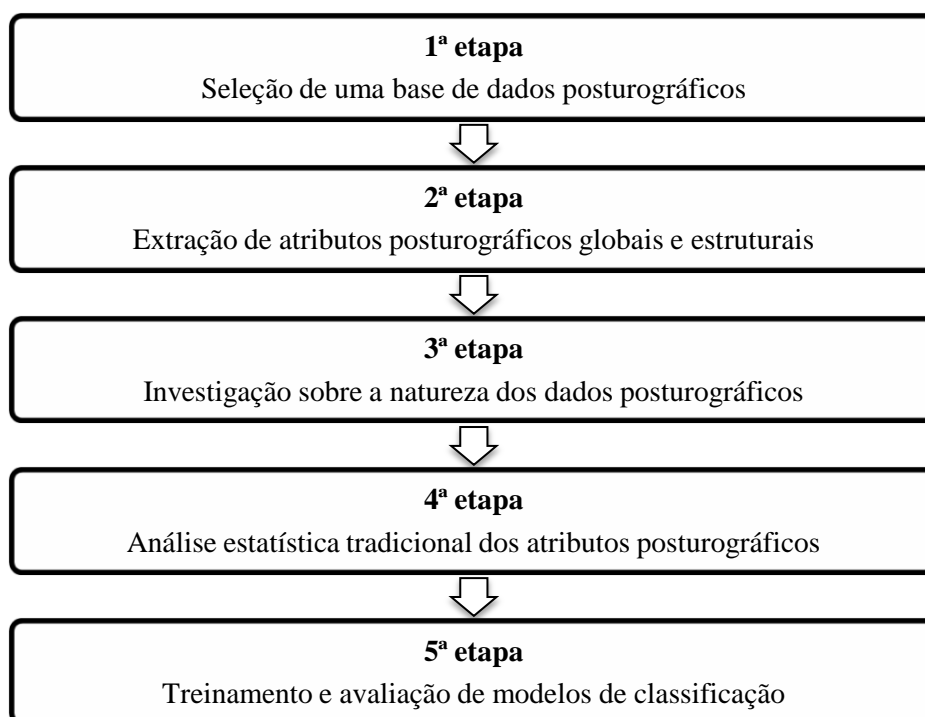


Figura 3.1 - Diagrama de blocos que expressa a metodologia empregada no presente estudo através de cinco etapas gerais.

Fonte: o autor.

3.1. Seleção da Base de Dados Posturográficos

Neste trabalho, foi utilizada a base de dados posturográficos coletada por Silva [SIL12] durante sua pesquisa de mestrado apresentada ao Programa de Pós-Graduação em Tecnologia em Saúde da PUCPR. A manipulação de tais dados pelo presente estudo foi aprovada pelo Comitê de Ética em Pesquisa da PUCPR sob o parecer de número 991.103 (ANEXO A) e encontra-se formalizada por meio de um Termo de Compromisso de Utilização de Dados (ANEXO B).

No estudo de Silva [SIL12], o controle do equilíbrio foi avaliado na presença do acidente vascular encefálico e também em diferentes tarefas posturais. Para este propósito, a autora primeiramente recrutou 19 pacientes adultos pós-AVE pareados (55.1 ± 6.7 anos, 1.65 ± 0.09 metros de estatura e 72.9 ± 12.3 kg de massa corporal) e outros 19 adultos saudáveis (53.6 ± 5.9 anos, 1.62 ± 0.07 metros de estatura e 66.1 ± 10.2 kg de massa corporal). É importante ressaltar que, entre os 19 indivíduos pós-AVE recrutados, constatou-se a ausência de outros distúrbios neurológicos além das sequelas decorrentes do AVE que pudessem influenciar no controle do equilíbrio postural, tal como a presença de Alzheimer/Parkinson, demência, disfunção cerebelar ou doença vestibular. Então, com o auxílio de uma plataforma de força, foram coletadas séries temporais dos deslocamentos do CP nas direções AP e ML de cada voluntário durante 60 segundos em cada uma das seguintes tarefas de manutenção do equilíbrio na postura ereta quieta: (i) mantendo os olhos abertos (condição OA), (ii) mantendo os olhos fechados (condição OF) e (iii) mantendo os olhos abertos e permanecendo sobre um tapete de espuma posicionado entre os pés e a plataforma (condição E). Para as coletas, foi adotada uma frequência de amostragem de 100Hz, resultando assim em séries do CP com comprimento de 6000 amostras cada, as quais foram registradas em formato de planilha Excel. Após a coleta, os dados foram submetidos a uma filtragem passa-baixas do tipo Butterworth de 2ª ordem com frequência de corte igual a 10Hz. Dessa forma, a base de dados coleta por Silva [SIL12] consiste em um total de 114 instâncias (38 voluntários \times 3 tarefas posturais), sendo 57 delas oriundas do grupo (ou classe) de pacientes pós-AVE e outras 57 produzidas pelo grupo de indivíduos saudáveis. Em outras palavras, cada instância corresponde às séries do CP nas direções AP e ML de um determinado indivíduo em uma determinada tarefa.

3.2. Extração de Atributos Posturográficos

Após a seleção de uma base de dados posturográficos, foram desenvolvidas duas rotinas computacionais distintas em ambiente MATLAB R2013b para a análise daqueles dados: a *rotina de processamento global* e a *rotina de processo estrutural*, responsáveis pela extração de 28 atributos posturográficos globais (ver Tabela 2.1) e de 18 atributos estruturais (ver Tabela 2.2), respectivamente, a partir das séries do CP. Tais rotinas serão descritas em detalhes nas duas subseções seguintes.

3.2.1. Rotina de Processamento Global

Neste estudo, foram consideradas todas as métricas de avaliação global do equilíbrio listadas na Tabela 2.1 e discutidas previamente na subseção 2.1.3.1, as quais são amplamente difundidas na literatura da posturografia. É importante lembrar que, dentre aquelas métricas, 08 são de domínio temporal, 04 de domínio espaço-temporal e outras 04 de domínio espectral. Como já foi discutido, as métricas espaço-temporais são bidimensionais, ou seja, levam em consideração as trajetórias do CP nas direções AP e ML simultaneamente, ao passo que as métricas temporais e espectrais são unidimensionais, isto é, são calculadas individualmente para as direções AP e ML. Dessa forma, foram extraídos no total 28 APG de cada uma das instâncias.

O funcionamento da rotina de processamento global desenvolvida nesta pesquisa pode ser resumido nas seguintes etapas:

Para cada uma das 114 instâncias:

- (i) Lê as séries temporais do CP nas direções AP e ML registradas em planilha Excel.
 - (ii) Remove o nível médio das séries.
 - (iii) Calcula todos os APG listados na Tabela 2.1.
 - (iv) Registra os APG em uma planilha Excel de saída.
-

O nível médio do sinal do CP depende tão somente da posição absoluta do sujeito sobre a plataforma de força [DUA10], ou seja, não carrega nenhuma informação relevante sobre o controle do equilíbrio em si; dessa forma, sua remoção é um procedimento comumente adotado pelos estudos da área da posturografia (e.g., [DON07; ROE06; SIL12]).

O cálculo dos atributos posturográficos globais considerados nesta pesquisa foi realizado através da implementação computacional de suas respectivas equações matemáticas, todas previamente apresentadas na subseção 2.1.3.1. Neste contexto, para o cálculo do atributo *área do CP*, foi utilizada a técnica de análise de componentes principais introduzida por Oliveira et al. [OLI94] para a determinação dos eixos da elipse que englobava 95% das trajetórias do estatocinesiograma [DUA00a; DUA10; SIL12]. Os atributos espectrais foram calculados utilizando uma implementação disponibilizada na literatura [DUA10], sendo que a *concentração de p% da potência espectral* foi calculada para $p = 50, 80$ e 95 [BAR02; KIR12; KIR13].

3.2.2. Rotina de Processamento Estrutural

Em relação à análise estrutural do balanço corporal, foram considerados neste estudo todos os 09 atributos posturográficos estruturais listados na Tabela 2.2 e previamente elucidados na subseção 2.1.3.2, os quais são amplamente empregados em estudos de avaliação do equilíbrio. Como todos aqueles atributos são calculados individualmente para as direções AP e ML de deslocamento do CP, foram extraídos no total 18 APE de cada uma das instâncias.

O funcionamento da rotina de processamento estrutural dos dados posturográficos implementada neste estudo pode ser resumido nas seguintes etapas:

Para cada uma das 114 instâncias:

Para cada uma das direções de deslocamento do CP (AP e ML) em particular:

- (i) Lê a série temporal do CP naquela direção.
 - (ii) Realiza o *detrending* da série.
 - (iii) Gera uma série *surrogate* de fases e uma série *surrogate* de amostras.
 - (iv) Para cada uma das três séries (original, *surrogate* de fases e *surrogate* de amostras):
 - (iv-a) Calcula todos os APE apresentados previamente na subseção 2.1.3.2.
 - (v) Registra os APE em uma planilha Excel de saída.
-

É importante ressaltar que todo o processamento estrutural dos dados foi realizado com o auxílio do TSTOOL [TST09], um pacote de funções implementadas em MATLAB e destinadas à análises não-lineares de séries temporais.

O método de *detrending* mencionado na etapa (ii) da rotina principal varia de acordo com o parâmetro estrutural a ser calculado. Isto porque alguns dos APE considerados nesta pesquisa são sensíveis à estacionariedade do sinal e, conforme constatado em um estudo anterior [MAR12], os dados do CP em uso exibem características de processos não-estacionários, sendo necessário assim o uso de alguma técnica para a obtenção da estacionariedade dos dados. Mais especificamente, os APE deste estudo sensíveis à não-estacionariedade são a dimensão de correlação, o máximo expoente de Lyapunov — pois ambos dependem da reconstrução do atrator do sistema no espaço de fases [THI00] — e a entropia multiescala [COS07]. Dessa forma, para o cálculo destes parâmetros e também da entropia amostral — derivada da entropia multiescala para a escala de tempo original (de aquisição) dos sinais —, o *detrending* dos dados do CP foi realizado com o auxílio da técnica *Empirical Mode Decomposition* (EMD) (e.g., [COS05]). Já para o cálculo dos demais atributos estruturais, o *detrending* consistiu apenas em remover o nível médio do sinal (e.g., [DON07; GHO10; ROE06]).

De maneira sucinta, a técnica EMD permite decompor um dado sinal em um conjunto de *Intrinsic Mode Functions* (IMFs), cada uma delas com componentes espectrais predominantes distribuídas sobre diferentes faixas de frequência. Dessa forma, diferentes IMFs descrevem o sinal original sobre diferentes escalas de tempo. Neste estudo, assim como realizado em [COS07], o método de *detrending* dos dados por meio da técnica EMD consistiu em subtrair de cada série original suas cinco últimas IMFs, as quais exibiram 99% da potência espectral total limitada abaixo de 2Hz para todos os sinais do CP. Conforme será detalhado mais adiante, a entropia multiescala será avaliada neste estudo para escalas de tempo entre 0.01 s (100Hz) e 0.5 s (2Hz); portanto, a remoção de componentes espectrais limitadas abaixo de 2Hz garante que nenhuma informação relevante para a análise multiescala será eliminada.

Conforme será discutido em detalhes na subseção 3.3, com a intenção de investigar a natureza dos dados posturográficos quanto à caótica determinística ou estocástica, cada um dos APE considerados nesta pesquisa foi calculado para as séries originais do CP e também para as suas versões *surrogate* de fases e *surrogate* de amostras (ver subseção 3.3), as quais foram geradas com o auxílio das respectivas funções do TSTOOL:

$$serie_surrogate_fases = \mathbf{surrogate1}(serie_original) \quad (3.1)$$

$$serie_surrogate_amostras = \mathbf{surrogate3}(serie_original), \quad (3.2)$$

sendo *serie_original* uma variável contendo o sinal original do CP. Na sequência, será detalhado o método implementado na rotina para o cálculo de cada um dos 09 atributos posturográficos estruturais supramencionados.

Dimensão de Correlação

A dimensão de correlação (D_C) foi estimada através do levantamento da relação D_C versus m em conjunto com a aplicação do teorema de Takens, ambos explanados previamente na subseção 2.1.3.2. Tal procedimento foi realizado com o auxílio de uma sub-rotina cujo funcionamento encontra-se resumido abaixo:

Para cada série do CP lida pela rotina principal:

- (i) Estima o valor do atraso temporal (τ).
 - (ii) Para uma dimensão de imersão $m = 1, 2, \dots, 15$:
 - (ii-a) Reconstrói o atrator do sistema em um espaço de fases m -dimensional.
 - (ii-b) Estima o valor da dimensão de correlação (D_C) do atrator reconstruído.
 - (ii-c) Registra o valor de m e o respectivo valor de D_C em uma matriz de saída.
 - (iii) Aplica o critério de Takens na matriz de saída.
 - (iv) Retorna para a rotina principal os valores adequados de m e de D_C da série.
-

O atraso temporal citado na etapa (i) foi estimado através do primeiro mínimo local da função de informação mútua (FIM) (vide subseção 2.1.3.2, equação (2.16)), a qual foi computada com o auxílio da seguinte função do TSTOOL:

$$FIM = \mathbf{amutual}(serie, \tau_{max}), \quad (3.3)$$

em que *serie* é um vetor contendo o sinal do CP em análise e τ_{max} contém o valor máximo de atraso para o qual a informação mútua será calculada. Esta função tem como retorno um vetor *FIM* de $\tau_{max} + 1$ posições, contendo a informação mútua da série para $\tau = 0, 1, \dots, \tau_{max}$. No estudo similar de Roerdink et al. [ROE06], o primeiro mínimo da FIM para um grupo de pacientes pós-AVE foi observado (na média) em $\tau = 11$ para a coordenada ML e em $\tau = 10$ para a coordenada AP enquanto que, para um grupo de indivíduos saudáveis pareados, estes mínimos foram encontrados (na média) em $\tau = 11$ para ambas as direções AP e ML de deslocamento do CP. Sendo

assim, para se obter uma ampla visualização dos desdobramentos da FIM, adotou-se $\tau_{max} = 100$ no presente estudo.

Na seqüência, a reconstrução do atrator do sistema mencionada na etapa (ii-a) da sub-rotina foi realizada com o método das coordenadas defasadas (vide equação (2.15)), o qual encontra-se implementado na função *embed* do TSTOOL:

$$atrator = \mathbf{embed}(serie, \tau, m), \quad (3.4)$$

sendo τ o valor do atraso temporal obtido na etapa (i) e m o valor da dimensão escolhida para o espaço de fases. Esta função tem como retorno a matriz *atrator* com $c = 1, \dots, m$ colunas, sendo que cada c -ésima coluna contém a série defasada de $(c - 1)\tau$. Então, a dimensão de correlação do atrator reconstruído foi estimada através da soma de correlação (vide subseção 2.1.3.2, equação (2.19)), calculada com o TSTOOL da seguinte forma:

$$[C_m, r] = \mathbf{corrsum}(atrator, r_{min}, r_{max}, nr), \quad (3.5)$$

sendo r_{min} e r_{max} os raios relativos — em relação ao diâmetro do atrator — mínimo e máximo da esfera de busca utilizada na busca por pares de pontos, enquanto que nr representa a quantidade de raios r que serão atribuídos à esfera durante a busca tal que $r_{min} \leq r \leq r_{max}$. A função tem como retorno o vetor r contendo os valores dos raios atribuídos à esfera de busca, e também o vetor C_m contendo a soma de correlação para os raios mencionados, ambos os vetores com nr posições. Neste estudo, a soma de correlação foi calculada para $nr = 200$ raios de busca compreendidos entre $r_{min} = 0.005$ e $r_{max} = 0.75$, ou seja, considerando desde 0.5% até 75% dos pares de pontos do atrator reconstruído [DON07; ROE06]. Então, o valor da D_C foi estimado como sendo a inclinação de uma linearização da curva C_m versus r em escala logarítmica [DON07; ROE06] (vide subseção 2.1.3.2, equação (2.20)).

Conforme indicado na etapa (ii) da sub-rotina, o procedimento supracitado de reconstrução do atrator e posterior estimação da dimensão de correlação foi realizado para $m = 1, 2, \dots, 15$ [ROE06] visando o levantamento da relação D_C versus m da série. Após tal levantamento, o critério de Takens (vide subseção 2.1.3.2, equação (2.17)) foi implementado para auxiliar na escolha da dimensão m e sua respectiva D_C mais apropriadas para descrever a série analisada [DON07; ROE06].

Máximo Expoente de Lyapunov

A estimativa do maior expoente de Lyapunov (λ_{max}) foi realizada com o auxílio do erro de previsão (vide subseção 2.1.3.2), calculado com o TSTOOL da seguinte forma:

$$ErroPrev = \mathbf{largelyap}(atrator, L_0^{max}), \quad (3.6)$$

em que L_0^{max} é a maior distância inicial de separação entre órbitas inicialmente vizinhas do atrator para a qual o erro de previsão será calculado. A função tem como retorno o vetor $ErroPrev$ de $L_0^{max} + 1$ posições, o qual contém o erro de previsão para $L_0 = 0, 1, \dots, L_0^{max}$. No trabalho de Parlitz [PAR98], para a estimação do máximo expoente de Lyapunov a partir de uma série de dados caóticos experimentais, o erro de previsão foi calculado tomando-se $L_0^{max} = 30$. Neste trabalho, para garantir o desdobramento de todas as três fases do erro de previsão (vide subseção 2.1.3.2 para mais detalhes), adotou-se $L_0^{max} = 100$.

Conforme discutido previamente, a estimação de λ_{max} pelo erro de previsão exige a identificação da fase II da curva $ErroPrev$ versus L_0 , cujo início e fim estão associados a mudanças no coeficiente angular desta curva. Sendo assim, tal determinação foi realizada com o auxílio da taxa de variação (ou derivada de 1ª ordem) do erro de previsão (Figura 3.2). Neste cenário, a distância que marca o início da fase II ($L_0^{início}$) foi estimada através do pico da derivada do erro de previsão, ao passo que a distância que demarca o final desta fase (L_0^{fim}) foi tomada como a distância para a qual o erro de previsão decaiu abaixo de 50% do seu valor de pico (Figura 3.2, painel à esquerda). Inclinações inferiores à este limiar foram consideradas pertencentes à fase III, a qual representa o período de convergência de $ErroPrev$ versus L_0 para um valor de saturação. Tais limiares foram arbitrados empiricamente através de uma análise dos dados em estudo uma vez que nenhuma orientação sobre a identificação da fase II na curva do erro de previsão para dados posturográficos foi encontrada na literatura. Finalmente, o máximo expoente de Lyapunov foi estimado como sendo a inclinação de uma linearização da fase II na curva do erro de previsão (Figura 3.2, painel à direita) [PAR98].

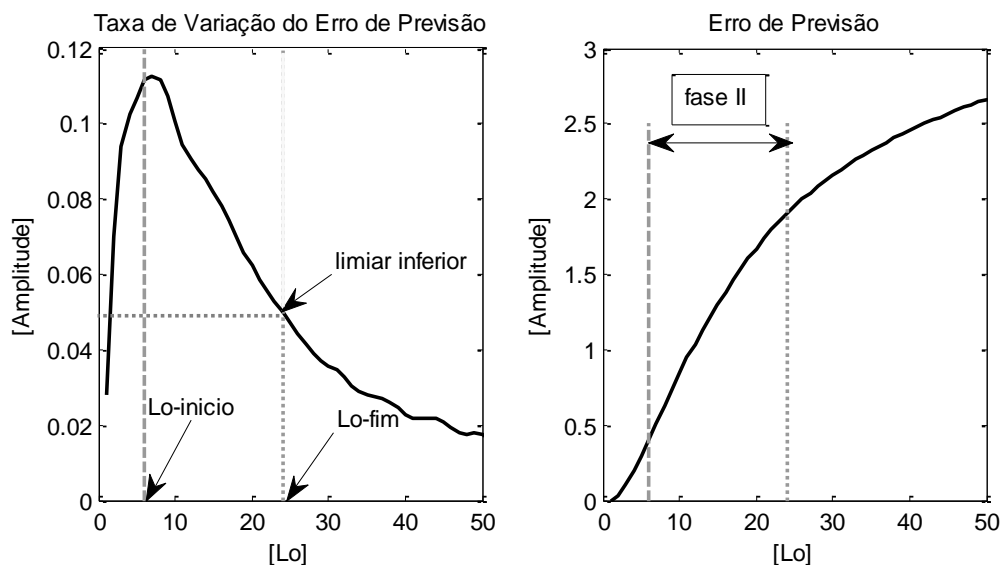


Figura 3.2 - À esquerda, é exibida a derivada de 1ª ordem do erro de previsão calculado para uma série do CP na direção ML de um indivíduo saudável na condição OA. O pico deste gráfico (linha tracejada) foi utilizado para identificar o início da fase II, enquanto que a queda abaixo do limiar de 50% do valor de pico (linha pontilhada) demarca o fim desta fase. À direita, é exibida a curva do erro de previsão com as demarcações da fase II.

Fonte: o autor.

Entropia Amostral

No estudo de Rhea et al. [RHE11], os autores verificaram que a entropia amostral (*EntAm*) é um parâmetro sensível às colinearidades do sinal em estudo, as quais se fazem presentes quando utiliza-se uma frequência de amostragem muito superior do que o exigido pelo teorema de Nyquist para capturar fenômenos de baixa frequência, assim como observado nos estudos da área da posturografia em geral para a quantificação da oscilação postural na forma de séries temporais do CP. Diante disso, para o cálculo da *EntAm*, o sinal do CP em análise foi primeiramente reamostrado a uma taxa de 20Hz [KIR12; KIR13] visando a redução de suas colinearidades. Com isso, o comprimento da série foi redimensionado de 6000 para 1200 amostras, as quais continuavam a representar 60 segundos de observação das trajetórias do CP. Na sequência, o sinal foi normalizado de forma a apresentar média nula e desvio padrão unitário (e.g., [KIR12; ROE06]). Então, a partir do sinal reamostrado e normalizado, a entropia amostral foi calculada com o auxílio da toolbox *PhysioNet* [PHY11] para MATLAB tomando os parâmetros $M = 2$ e $\Gamma = 0.15$ [BOR10; CHE06; COS05; KIR13].

Entropia Multiescala

Neste estudo, assim como no trabalho de Duarte e Sternad [DUA08], a entropia multiescala do sinal do CP foi calculada para $\eta = 1, 2, \dots, \eta_{max}$ com $\eta_{max} = 50$ amostras (ou 0.5 s). Neste contexto é importante lembrar que, conforme discutido anteriormente na subseção 2.1.3.2, a literatura sugere a existência de pelo menos 10^M amostras no sinal em estudo para garantir uma estimativa adequada da entropia amostral [BOR10]. Sendo assim, como foi adotado $M = 2$ neste trabalho, são necessárias pelo menos $10^2 = 100$ amostras. Visto que o comprimento do sinal do CP é de $N = 6000$ amostras cada, o valor $\eta_{max} = 50$ adotado neste estudo satisfaz o critério mencionado pois, utilizado tal fator de escala, o sinal resultante apresentará $N/\eta_{max} = 120$ amostras. Por fim, a área sob a curva da entropia multiescala ($EntME$) foi calculada através da implementação da equação (2.21) sugerida por Costa et al. [COS07].

Expoente de Escala e Expoente de Hurst

Neste estudo, as propriedades fractais do sinal do CP foram avaliadas com o auxílio das técnicas DFA e SWV previamente discutidas na subseção 2.1.3.2, as quais resultam respectivamente em um expoente de escala α e em um expoente de Hurst estimado \hat{H} . Para isso, foram implementadas todas as equações matemáticas envolvidas nestas técnicas, as quais encontram-se descritas na subseção 2.1.3.2. É importante ressaltar que, antes da extração dos expoentes em questão, o sinal do CP em análise foi reamostrado a 20Hz [KIR12; KIR13] visando reduzir suas colinearidades [RHE11].

No presente trabalho, os cálculos das técnicas DFA e SWV foram realizados dividindo o sinal do CP em janelas não sobrepostas de comprimento w com $w = w_{min}, w_{min} + 2, w_{min} + 4, \dots, w_{max}$, sendo $w_{min} = 20$ amostras (ou 1 segundo) e $w_{max} = 200$ amostras (ou 10 segundos) [KIR12; KIR13]. Em outras palavras, o sinal foi dividido desde 60 até 06 janelas não sobrepostas de mesmo comprimento. No caso da técnica DFA, a tendência linear dentro de cada uma das janelas não sobrepostas do sinal foi estimada através de uma regressão linear [KIR13; ROE06].

Após os cálculos supracitados, foram estimadas as relações $F(w)$ versus w com a DFA e S_w versus w com a SWV. Finalmente, a inclinação de uma linearização destas curvas em escala logarítmica foi utilizada para estimar os valores dos expoentes de escala e de Hurst, respectivamente [CAN97; KIR12; ROE06].

Recurrence Quantification Analysis

A estimação das medidas Taxa de Recorrência (TR_{RQA}), Determinismo (DET_{RQA}) e Entropia (ENT_{RQA}) previamente discutidas na subseção 2.1.3.2 foi realizada com o auxílio da toolbox CRP v5.17 desenvolvida e aprimorada por Marwan et al. [MAR02a; MAR02b; MAR02c; MAR07]. Esta toolbox contém funções desenvolvidas em MATLAB destinadas ao cálculo de gráficos de recorrência e também de diversas métricas oriundas da técnica RQA a partir de séries temporais. Para o cálculo dos atributos supracitados, adotou-se um comprimento mínimo $l_{min} = 4$ para a contagem das diagonais do gráfico de recorrência [TAL13], sendo utilizada a norma Euclidiana para as medições de distância [NEG13].

3.2.3. Avaliação do Tempo de Execução das Rotinas

Assim como mencionado no capítulo 01, o principal objetivo deste trabalho consiste em avaliar a relevância de certos atributos posturográficos globais e estruturais para diferenciar indivíduos saudáveis de pacientes pós-AVE com o auxílio de modelos de classificação. Neste contexto, além da eficiência propriamente dita dos classificadores perante um ou outro conjunto de atributos, levou-se em conta também o tempo gasto com o cálculo daqueles atributos por meio das rotinas computacionais desenvolvidas nesta pesquisa. Isso porque o estudo do equilíbrio não se limita apenas à análises realizadas após uma coleta de dados, circunstância em que o pesquisador normalmente dispõe de tempo hábil para desenvolver análises mais complexas e elaboradas sobre os dados posturográficos. Na verdade, conforme discutido na subseção 2.2.5, alguns estudos na área de AM têm utilizado métodos de classificação para o desenvolvimento de sistemas de detecção de quedas em tempo real, cenário no qual o tempo gasto com a extração das métricas posturográficas é de relevância crucial.

Sendo assim, durante a execução das rotinas de processamento global (ver subseção 3.2.1) e de processamento estrutural (ver subseção 3.2.2), mensurou-se o tempo médio de extração de cada uma das métricas posturográficas consideradas nesta pesquisa. Além disso, mensurou-se também o tempo total de execução de cada uma das rotinas supracitadas, ou seja, o tempo gasto com a extração de todos os atributos globais e estruturais para todas as 114 instâncias da base de dados em estudo.

3.3. Avaliação da Natureza dos Dados Posturográficos

A literatura da área da posturografia não é unânime quanto à natureza da oscilação postural humana; isso porque, enquanto alguns estudos têm associado esta oscilação à processos aleatórios (e.g., [DUA00b; COL93]), os achados de alguns outros trabalhos permitiram não somente distingui-la de processos estocásticos como também associá-la a processos de comportamento caótico determinístico (e.g., [LAD09; NEW93]) (consultar [GHO10] para uma discussão mais detalhada acerca deste impasse). Por este motivo, os trabalhos da área normalmente realizam investigações sobre o comportamento dos dados posturográficos em estudo quanto à caótico determinístico ou estocástico.

Para o propósito supracitado, um procedimento amplamente utilizado na literatura consiste em comparar estatisticamente um determinado APE calculado a partir das séries originais do CP e de suas chamadas *versões surrogate* [THE92] (e.g., [DON07; GHO10; ROE06]), as quais representam séries de comportamento predominantemente estocástico que foram geradas a partir da manipulação de algumas propriedades dos dados originais [THE92]. Isto porque, ao contrário dos APG, os APE capturam a organização temporal do sinal em estudo, estando aptos assim para caracterizar o comportamento evolutivo daquele sinal. Portanto, a existência de diferença estatística entre um determinado APE calculado para um conjunto de dados originais e para um conjunto de dados *surrogate* sugere que os dois conjuntos não provêm da mesma população, ou seja, que os dados originais não são oriundos de um processo estocástico tal como os dados *surrogate* [GHO10].

Na literatura da área da posturografia, duas técnicas em particular têm sido utilizadas com maior frequência para a geração de séries *surrogate*, sendo elas: (i) embaralhar temporalmente as amostras da série original [SCH89] (e.g., [DON07; ROE06]), gerando assim uma *série surrogate de amostras* e (ii) embaralhar as fases do espectro da série original após a transformada de Fourier [THE92] (e.g., [DON07; GHO10; ROE06]), gerando assim uma *série surrogate de fases*. No primeiro caso, são preservadas as propriedades estatísticas da série original (e.g., média, variância), mas a correlação temporal do sinal é completamente aniquilada após o embaralhamento das amostras [ROE06]. Já no segundo caso, o embaralhamento das fases do espectro não altera a distribuição de potências espectrais do sinal original de forma que

autocorrelação entre as amostras é preservada [GHO10; ROE06], porém, todas as outras informações codificadas nas fases do espectro são perdidas [GHO10].

Conforme mencionado na subseção 3.2.2, todos os APE considerados nesta pesquisa foram calculados tanto para as séries originais do CP como também para suas versões *surrogate* de fases e *surrogate* de amostras. Entretanto, para a avaliação da natureza do conjunto de dados original, foram consideradas apenas a dimensão de correlação, o máximo expoente de Lyapunov, a entropia amostral e o expoente de escala [DON07; GHO10; ROE06]. Em outras palavras, com o intuito de evitar comparações redundantes, foram desconsideradas nesta etapa a entropia multiescala — associada com a entropia amostral já considerada — bem como o expoente de Hurst, uma medida associada ao já incluso expoente de escala. Finalmente, por estarem relacionadas com o gráfico de recorrência do sinal do CP ao invés do sinal original em si, as métricas de recorrência também foram desconsideradas nesta etapa do presente estudo.

Assim, após avaliar a normalidade da distribuição das amostras, o teste não-paramétrico U de Mann-Whitney foi utilizado para comparar estatisticamente os valores médios da dimensão de correlação, do máximo expoente de Lyapunov, da entropia amostral e do expoente de escala obtidos a partir das séries originais do CP e de suas versões *surrogate* de fases e *surrogate* de amostras. Estas comparações foram realizadas separadamente nas direções AP e ML para cada um dos grupos de sujeitos em estudo (pós-AVE e saudável). Em todos os casos, os valores médios dos parâmetros supracitados foram calculados sobre as três tarefas de equilíbrio postural (OA, OF, E) [DON07; ROE06].

3.4. Análise Estatística dos Atributos Posturográficos

Antes da utilização dos modelos de classificação, foi realizada primeiramente uma análise estatística sobre todos os atributos posturográficos calculados neste estudo. Conforme já discutido na subseção 2.1.4, esta prática é amplamente empregada nos estudos de avaliação do equilíbrio com o propósito de identificar a relevância individual de um determinado parâmetro para diferenciar os sistemas fisiológicos sob investigação. Assim, após avaliar a normalidade da distribuição das amostras, o teste não-paramétrico U de Mann-Whitney foi utilizado para comparar entre os grupos de sujeitos (saudável e pós-AVE) o valor médio sobre as três tarefas de equilíbrio de cada um dos atributos posturográficos utilizados nesta pesquisa, globais e estruturais. Estas comparações

foram realizadas separadamente nas direções AP e ML para os atributos unidimensionais (e.g., velocidade média, dimensão de correlação) e simultaneamente nas duas direções para os atributos bidimensionais ou espaço-temporais (e.g., velocidade média total, área do CP).

3.5. Avaliação dos Métodos de Classificação

3.5.1. Bases de Dados Secundárias

Primeiramente, é importante lembrar que a base de dados posturográficos utilizada neste trabalho é composta por 114 instâncias no total (38 voluntários \times 3 tarefas posturais), sendo 57 delas provenientes da classe de indivíduos saudáveis e as outras 57 oriundas da classe de pacientes pós-AVE. Mais especificamente, cada uma destas instâncias consistem em duas séries temporais dos deslocamentos do CP — uma na direção AP e outra na direção ML — de um determinado indivíduo em uma determinada tarefa de equilíbrio (OA, OF ou E). Então, conforme explanado na subseção 3.2, foram extraídos 46 atributos posturográficos numéricos a partir de cada uma das instâncias supracitadas, sendo 28 deles de avaliação global (APG) e os outros 18 de avaliação estrutural (APE) do balanço postural. Além disso, a tarefa de equilíbrio da instância também é uma característica útil para a sua descrição, sendo este um atributo do tipo nominal. Portanto, cada uma das 114 instâncias em estudo poderia ser representada com até 47 atributos.

Conforme discutido no capítulo 01, o principal objetivo desta pesquisa consiste em avaliar a relevância de APG e de APE comuns na área da posturografia para, com o auxílio de métodos de classificação, diferenciar indivíduos saudáveis de pacientes pós-AVE. Para este propósito, foram confeccionadas três bases de dados secundárias distintas para posterior análise com os modelos de classificação (Figura 3.3), sendo elas:

(i) Base *APG*: composta pelas 114 instâncias representadas por 29 atributos cada, sendo eles a tarefa de equilíbrio (atributo nominal) em conjunto com 28 APG (atributos numéricos);

(ii) Base *APE*: composta pelas 114 instâncias representadas por 19 atributos cada, sendo eles a tarefa de equilíbrio (atributo nominal) em conjunto com 18 APE (atributos numéricos);

(ii) Base *APGE*: composta pelas 114 instâncias representadas por 47 atributos cada, sendo eles a tarefa de equilíbrio (atributo nominal) juntamente com 28 atributos globais e com 18 atributos estruturais (atributos numéricos). Em outras palavras, esta base representa uma composição entre as duas bases anteriores.

É importante mencionar que, tendo em vista a utilização de determinados classificadores cujo paradigma leva em conta a distância entre os atributos das instâncias, todos os atributos numéricos das três bases supracitadas foram primeiramente normalizados para valores entre zero e um.

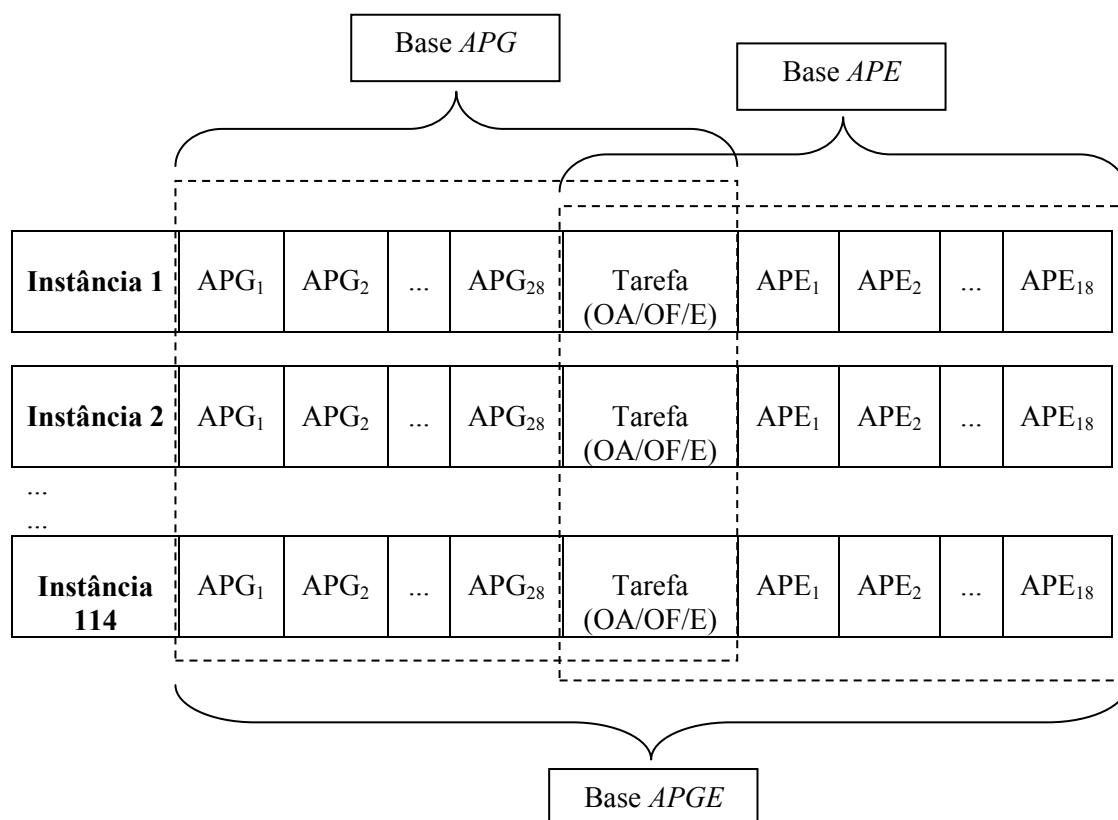


Figura 3.3 - Representação da composição das bases *APG*, *APE* e *APGE* em função da inclusão de determinados atributos posturográficos dentre todos os 47 disponíveis para cada uma das 114 instâncias.

Fonte: o autor.

3.5.2. Seleção de Atributos

Na literatura de AM, diversos estudos têm utilizado com sucesso a seleção de atributos para aprimorar o processo de aprendizagem tendo em vista os mais variados modelos de classificação nas mais diversas áreas de aplicação (e.g., microarranjos de DNA [BOR06; GAR11]). Já foi demonstrada também a relevância da redução de dimensionalidade em problemas de classificação do equilíbrio postural [SAR13]. No presente trabalho, visto que o objetivo principal diz respeito à avaliação média conjunta de diferentes modelos de classificação, foi realizada uma seleção de atributos baseada na abordagem *filter* (vide subseção 2.2.4) sobre as bases *APG*, *APE*, e *APGE* (ver Figura 3.3), dando origem assim a três novas bases: *APG-selec*, *APE-selec* e *APGE-selec*, respectivamente. Para isso, por meio do *software* WEKA [HAL09], foi utilizado o algoritmo CFS com o auxílio da técnica *best first search forward* (vide subseção 2.2.4) com critério de parada de cinco atributos que não elevam o mérito heurístico do subconjunto selecionado [HAL99; GHA12].

Na sequência, para avaliar os efeitos da redução de dimensionalidade das bases em estudo, 06 modelos populares de classificação foram treinados e avaliados sobre as bases integrais (*APG*, *APE* e *APGE*) e também sobre as bases selecionadas (*APG-selec*, *APE-selec* e *APGE-selec*). Como métrica de avaliação de desempenho dos modelos, foi utilizada a acurácia ou taxa de acerto (ver subseção 2.2.3) uma vez que tanto as bases integrais como também as selecionadas são balanceadas (i.e., possuem o mesmo número de instâncias para cada classe). Mais especificamente, para cada uma das bases supracitadas, calculou-se a acurácia média conjunta entre os seis modelos de classificação considerados. Por fim, após avaliar a normalidade da distribuição das amostras, o teste não-paramétrico U de Mann-Whitney foi utilizado para comparar a acurácia média conjunta dos modelos entre as bases $APG \leftrightarrow APG-selec$, $APE \leftrightarrow APE-selec$ e $APGE \leftrightarrow APGE-selec$. A Figura 3.4 ilustra todo o processo descrito neste parágrafo, tomando como exemplo as bases *APG* e *APG-selec*.

Mais especificamente, para a realização do procedimento supracitado, foram utilizados os 06 modelos de classificação relacionados na Tabela 3.1, visando uma avaliação mais abrangente dos classificadores, alguns modelos foram executados mais de uma vez com determinadas modificações em seus parâmetros de configuração. No caso da PG, por exemplo, o algoritmo *GeneticProgramming* [LEV08] foi executado utilizando primeiramente um conjunto de funções mais simples, envolvendo apenas as

quatro operações aritméticas fundamentais. Na sequência, foram adicionadas funções mais complexas ao conjunto, a saber, a raiz quadrada, a potência, o logaritmo na base 10 e as funções trigonométricas seno e cosseno.

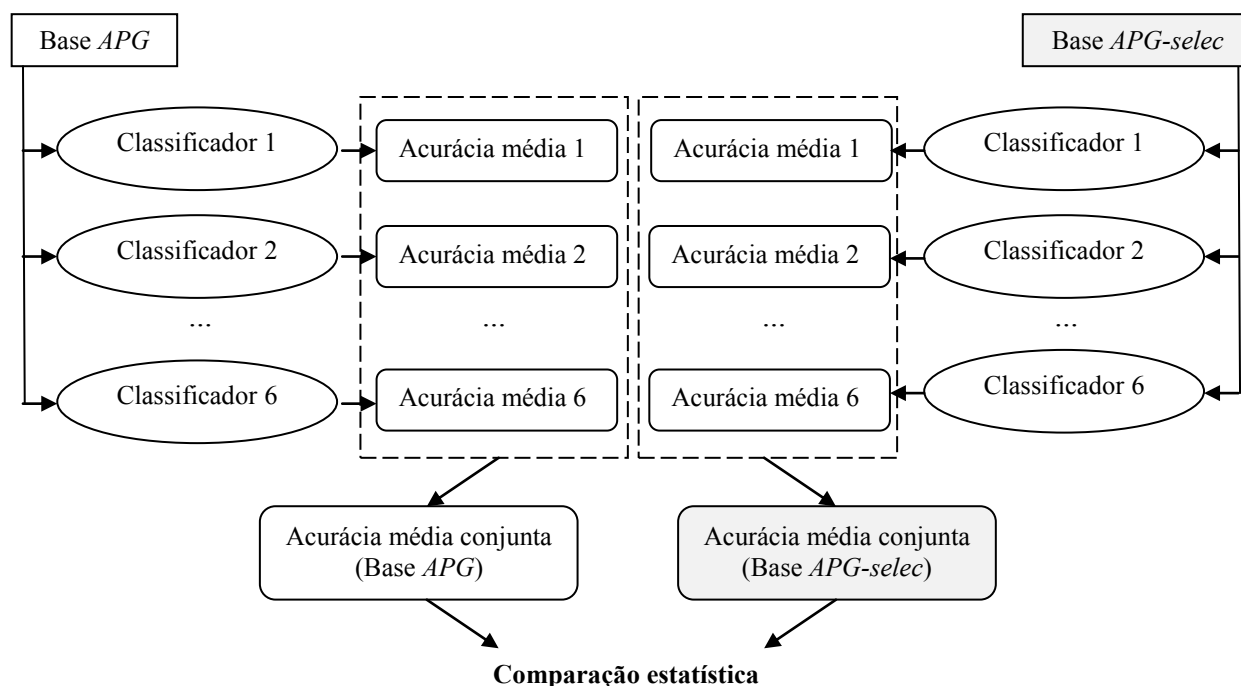


Figura 3.4 - Ilustração do método empregado neste trabalho para a avaliação dos efeitos da seleção de atributos nas bases de dados em estudo, tomando como exemplo as bases *APG* e *APG-selec*.

Fonte: o autor.

Tabela 3.1 - Relação dos algoritmos de classificação utilizados nesta pesquisa com os respectivos parâmetros de configuração adotados. A última coluna atribui uma sigla para as diferentes configurações de execução de cada algoritmo.

Modelo de classificação	Algoritmo no WEKA	Configurações de execução do algoritmo	Sigla
Programação Genética	<i>GeneticProgramming</i>	1000 indivíduos, 100 gerações, T={atributos, constantes aleatórias}, F={+,-,*,/}, demais parâmetros padrão	PG ₁
		1000 indivíduos, 100 gerações, T={atributos, constantes aleatórias}, F={+,-,*,/, raiz, potência, log10, seno, cosseno}, demais parâmetros padrão	PG ₂
Árvore de Decisão	<i>J48</i>	Todos os parâmetros padrão	AD
Naive Bayes	<i>NaiveBayes</i>	Todos os parâmetros padrão	NB
SVM	<i>LibSVM</i>	<i>Kernel radial</i> , custo = 1.0, demais parâmetros padrão	SVM ₁
		<i>Kernel radial</i> , custo = 10.0, demais parâmetros padrão	SVM ₂
		<i>Kernel sigmoidal</i> , custo = 1.0, demais parâmetros padrão	SVM ₃
		<i>Kernel sigmoidal</i> , custo = 10.0, demais parâmetros padrão	SVM ₄

Modelo de classificação	Algoritmo no WEKA	Configurações de execução do algoritmo	Sigla
<i>Multilayer Perpectron</i>	<i>MultilayerPerceptron</i>	Tempo de treinamento igual a 10 mil épocas, tamanho do conjunto de validação igual a 5%, demais parâmetros padrão	MLP ₁
		Tempo de treinamento igual a 10 mil épocas, tamanho do conjunto de validação igual a 10%, demais parâmetros padrão	MLP ₂
<i>k</i> vizinhos mais próximos	<i>IBk</i>	Parâmetro <i>k</i> = 1, 3 e 5, demais parâmetros padrão.	<i>k</i> -NN

Tendo em vista a quantidade de instâncias disponível nas bases em análise (114 no total, 57 de cada classe), todos os modelos de classificação relacionados na Tabela 3.1 foram treinados e testados com o auxílio da técnica *leave-one-out* dentro de 15 repetições para cada base. Esta técnica é equivalente a validação cruzada (ver subseção 2.2.3) quando se assume a quantidade de partições (*folds*) como sendo o número total de instâncias da base.

3.5.3. Relevância dos Atributos Posturográficos Globais e Estruturais

Após a obtenção da acurácia média conjunta entre os classificadores relacionados na Tabela 3.1 para as bases integrais e selecionadas, foi avaliada a relevância dos atributos posturográficos globais e estruturais considerados nesta pesquisa durante o processo de identificação das populações em estudo através dos métodos de classificação mencionados. Em outras palavras, dentre o uso isolado de atributos globais, o uso isolado de atributos estruturais e o uso conjunto de atributos globais e estruturais, avaliou-se qual destas três situações permitiu distinguir com maior precisão os indivíduos saudáveis dos pacientes pós-AVE (e vice-versa) com o auxílio dos modelos populares de classificação considerados neste estudo. Para este propósito, após avaliar a normalidade da distribuição das amostras, o teste não-paramétrico U de Mann-Whitney foi utilizado para comparar a acurácia média conjunta dos classificadores entre as três bases integrais e, posteriormente, entre as três bases selecionadas, assim como ilustra a Figura 3.5.

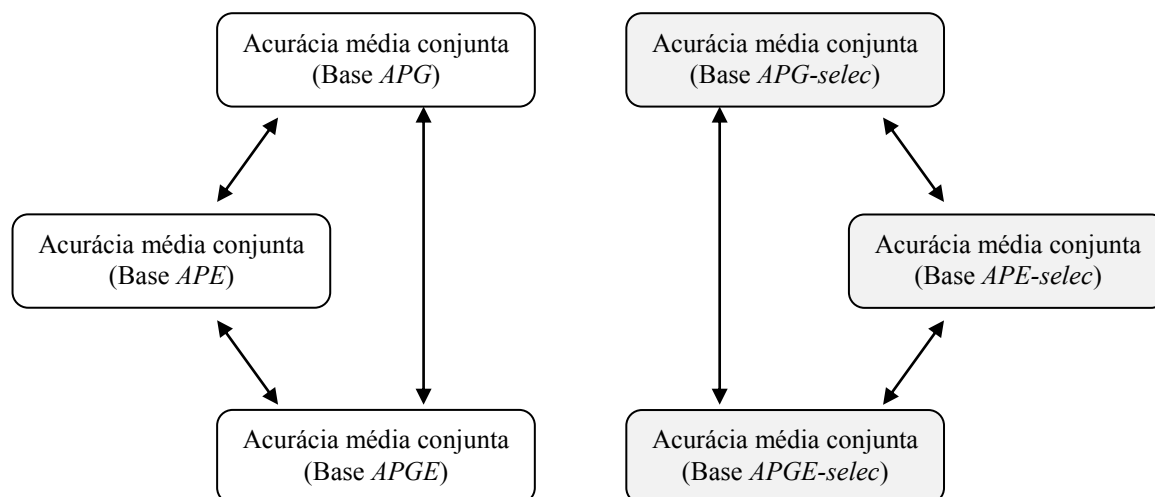


Figura 3.5 - Ilustração do método empregado neste trabalho para a avaliação da relevância de atributos posturográficos globais e estruturais durante a classificação de sujeitos saudáveis e de pacientes pós-AVE. As setas representam a realização de uma comparação estatística entre as duas médias apontadas.

Fonte: o autor.

Capítulo 4

Avaliação dos Resultados Experimentais

No presente capítulo, serão primeiramente apresentados os resultados provenientes de cada um dos experimentos realizados nesta pesquisa, os quais encontram-se detalhadamente descritos no capítulo anterior. Na sequência, será realizada uma discussão acerca de tais resultados tendo por base a hipótese de pesquisa inicialmente levantada pelo autor e também os achados de outros estudos correlatos.

4.1. Resultados

4.1.1. Tempo Gasto na Extração de Atributos

Durante a etapa de extração dos atributos posturográficos considerados neste estudo, constatou-se que as rotinas de processamento global e estrutural levaram aproximadamente 1.33 segundos e 6235.80 segundos (ou 1.73 horas), respectivamente, para realizar todos os cálculos implementados tendo em vista todas as 114 instâncias disponíveis na base utilizada. A Tabela 4.1 mostra uma relação do tempo médio gasto com a extração de cada atributo a partir das instâncias mencionadas. É importante ressaltar que, durante o processamento dos dados, cada uma das rotinas supracitadas foi executada sem outros processos de usuário em paralelo em um *notebook* com processador i5 de 1.8GHz, 06 GB de memória RAM e sistema operacional *Windows* 8.1 de 64 bits. De forma mais genérica (i.e., independentemente das configurações do computador utilizado), os APG foram extraídos 4688.6 vezes mais rápido do que os APE para os dados do CP. É importante salientar também que as 1.73 horas gastas pela rotina de processamento estrutural referem-se apenas à análise dos dados originais, ou seja, não levam em conta a extração de APE a partir dos dados *surrogate*.

Tabela 4.1 - Relação do tempo médio gasto durante a extração dos atributos posturográficos considerados nesta pesquisa.

Categoria	Tipo	Atributo posturográfico	Tempo médio de extração por instância	Tempo médio total por instância
Global	Unidimensional	D_{med}	0.24 ± 0.89 ms	11.68 ± 5.29 ms
		D_{rms}	0.18 ± 0.18 ms	
		DP	0.43 ± 0.16 ms	
		A_{des}	0.16 ± 0.66 ms	
		V_{med}	0.11 ± 0.03 ms	
		$DP_{V_{med}}$	0.36 ± 0.10 ms	
		V_p	0.08 ± 0.31 ms	
		V_{rms}	0.15 ± 0.05 ms	
		F_{med}, F_{50}, F_{80} e F_{95} *	8.83 ± 2.13 ms	
	Bidimensional	V_{medT}	0.23 ± 0.43 ms	
		SP	0.43 ± 0.10 ms	
		DOT	0.14 ± 0.16 ms	
		$Area$	0.34 ± 0.08 ms	
Estrutural	Unidimensional	D_c	44.82 ± 1.27 s	54.70 ± 1.55 s
		λ_{max}	1.52 ± 0.05 s	
		$EntAm$	0.07 ± 0.01 s	
		$EntME$	2.98 ± 0.03 s	
		α	1.62 ± 0.04 s	
		H	1.64 ± 0.03 s	
		TR_{ROA}, DET_{ROA} e ENT_{ROA} *	2.04 ± 0.13 s	

* Parâmetros extraídos de forma conjunta através de uma mesma função computacional.

4.1.2. Dados Originais versus Dados Surrogate

A Figura 4.2 mostra uma comparação entre os valores médios da dimensão de correlação (D_C), do máximo expoente de Lyapunov (λ_{max}), da entropia amostral ($EntAm$) e do expoente de escala (α) sobre as três tarefas de equilíbrio (OA, OF, E) calculados separadamente para os dois grupos de sujeitos (pós-AVE e saudável) nas direções AP e ML a partir das séries originais do CP e também de suas versões *surrogate* de fases e *surrogate* de amostras. Na realidade, para as séries *surrogate* de amostras, não foi possível estimar o valor médio dos parâmetros D_C e λ_{max} ; isso porque a relação D_C versus m , $m = 1, 2, \dots, 15$ de nenhuma daquelas séries foi capaz de satisfazer o critério de Takens utilizado neste estudo, ou seja, não foi possível determinar um valor de m para a reconstrução do atrator de nenhuma daquelas séries no espaço de fases (Figura 4.1). Como pode ser visto na Figura 4.2, com exceção do parâmetro α para as séries originais e *surrogate* de fases (Figura 4.2, canto inferior direito), todas as outras métricas posturográficas calculadas para os dados originais, *surrogate* de fases e *surrogate* de amostras exibiram diferença estatisticamente significativa ($p < 0.05$) entre si.

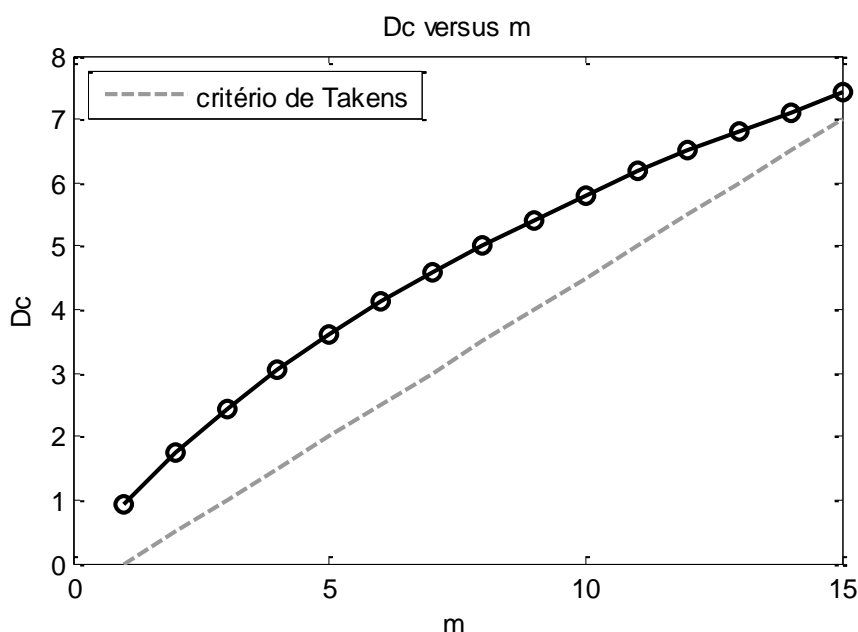


Figura 4.1 - Curva D_C versus m , $m = 1, 2, \dots, 15$ (linha —o— preta) para uma série do CP na direção AP de um indivíduo pós-AVE na condição OA. A linha cinza pontilhada representa o critério de Takens utilizado para a escolha das dimensões m e D_C do sistema.

Fonte: o autor.

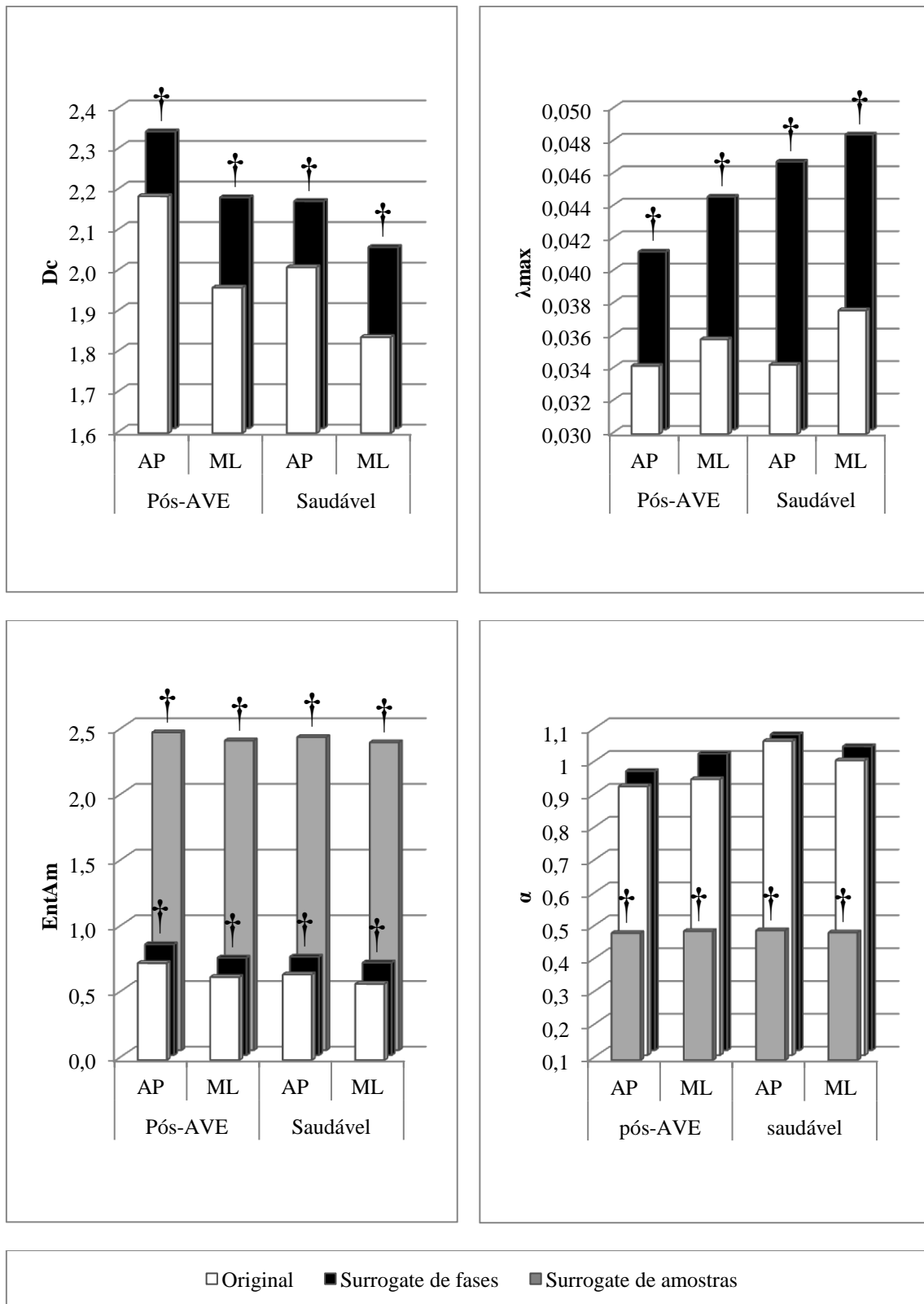


Figura 4.2 - Comparação entre os valores médios dos parâmetros D_c (canto superior esquerdo), λ_{max} (canto superior direito), $EntAm$ (canto inferior esquerdo) e α (canto inferior direito) calculados para as séries originais do CP (barras em branco) e para suas versões *surrogate* de fases (barras em preto) e *surrogate* de amostras (barras em cinza). As *cruzes* indicam diferença estatisticamente significativa ($p < 0.05$) entre um determinado parâmetro calculado para um determinado tipo de série e o mesmo parâmetro calculado para os outros tipos de séries em uma dada coordenada do CP.

4.1.3. Análise Estatística dos Atributos Posturográficos

Na comparação dos APG de domínio temporal entre indivíduos saudáveis e pacientes pós-AVE (Figura 4.3), foi observada diferença estatisticamente significativa apenas para os parâmetros referentes à velocidade dos deslocamentos do CP (Figura 4.3, coluna da direita) tanto na direção AP ($p < 0.001$) como também na direção ML ($p < 0.05$), sendo eles a velocidade média ($Vmed$), o desvio padrão da velocidade média (DP_Vmed), a velocidade de pico (Vp) e a velocidade RMS ($Vrms$).

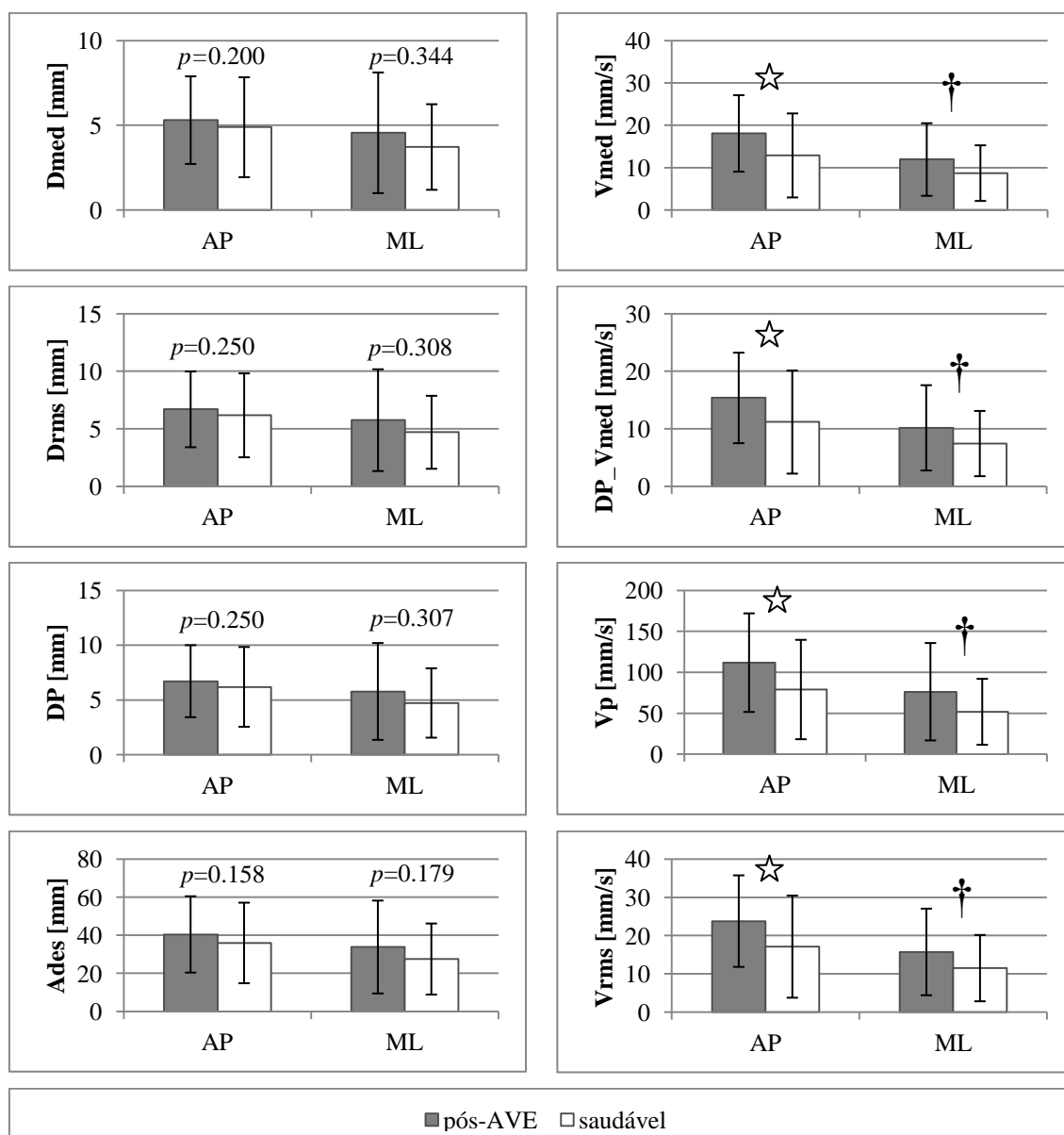


Figura 4.3 - Comparação entre os valores médios dos APG de domínio temporal obtidos para os indivíduos saudáveis (barras em branco) e pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* e as *cruzes* posicionadas em uma determinada coordenada indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente

Em relação aos APG de domínio espaço-temporal, foi observada diferença estatisticamente significativa ($p < 0.001$) entre os grupos de sujeitos apenas na comparação entre a velocidade média total (V_{medT}) (Figura 4.4, canto superior esquerdo) e entre o *Sway Path* (SP) (Figura 4.4, canto superior direito).

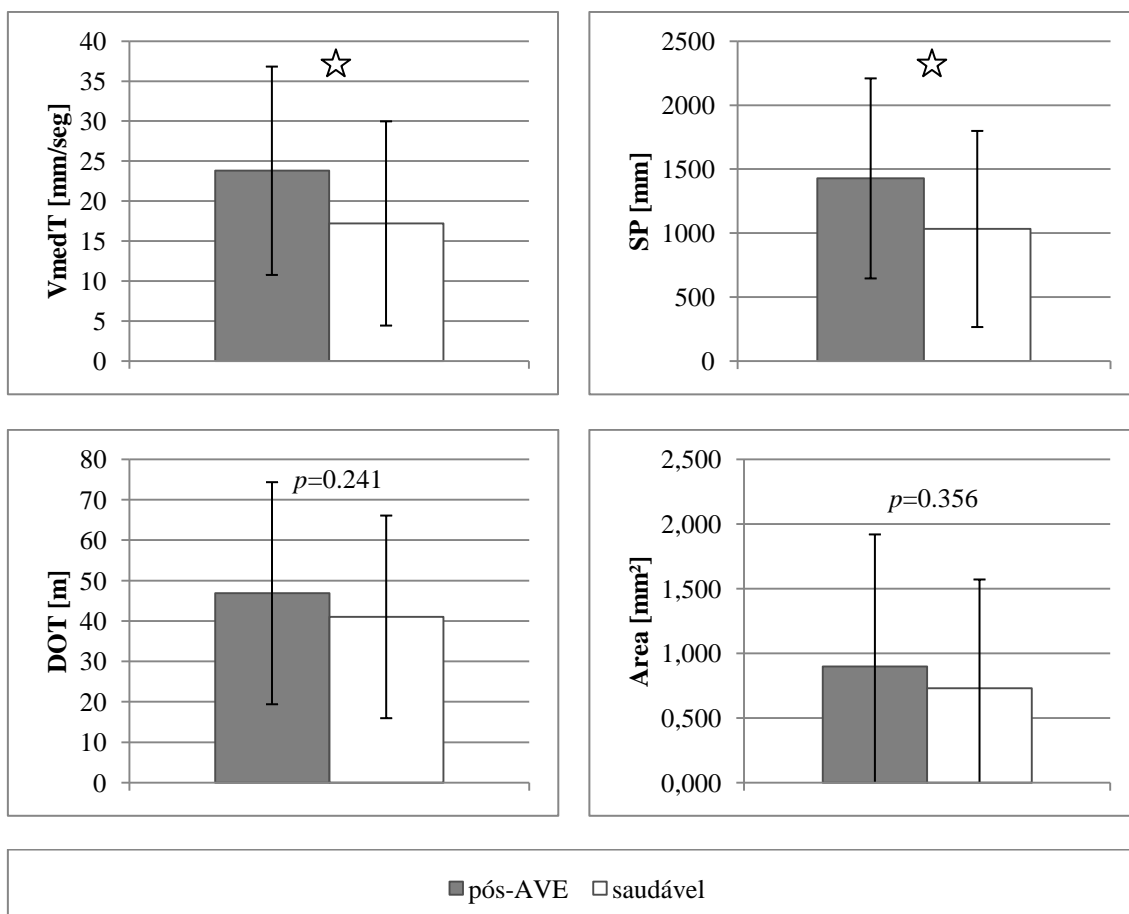


Figura 4.4 - Comparação entre os valores médios dos APG de domínio espaço-temporal obtidos para os indivíduos saudáveis (barras em branco) e para os pacientes pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* indicam diferença estatística com nível p de significância < 0.001 .

A Figura 4.5 mostra uma comparação entre os valores médios sobre as três tarefas de equilíbrio dos APG de domínio espectral calculados para indivíduos saudáveis e pós-AVE nas coordenadas AP e ML de deslocamento do CP. Neste cenário, foi constatada diferença estatisticamente significativa tanto na direção AP ($p < 0.001$) como também na direção ML ($p < 0.05$) para todos os atributos em questão, sendo eles a frequência espectral média (F_{med}) e as frequências de concentração de 50% (F_{50}), de 80% (F_{80}) e de 95% (F_{95}) da potência espectral total.

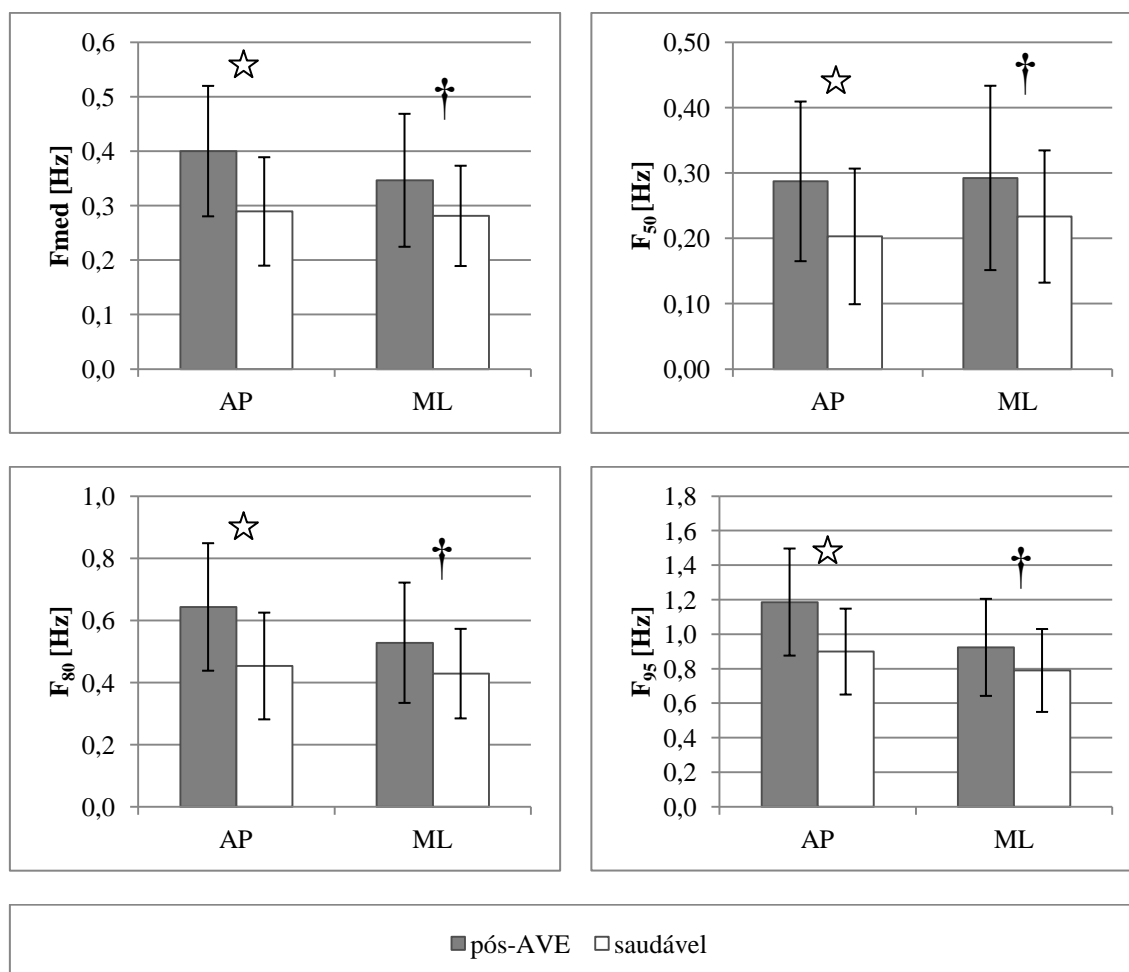


Figura 4.5 - Comparação entre os valores médios dos APG de domínio espectral calculados para os indivíduos saudáveis (barras em branco) e para os pacientes pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* e as *cruzes* posicionadas em uma determinada coordenada indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente.

Em relação aos atributos posturográficos estruturais, a Figura 4.6 exibe uma comparação entre os valores médios sobre as três tarefas de equilíbrio da dimensão de correlação (D_C), do máximo expoente de Lyapunov (λ_{max}), da entropia amostral ($EntAm$), da entropia multiescala ($EntME$), dos expoentes de escala (α) e de Hurst (H), da taxa de recorrência (TR_{RQA}), do determinismo (DET_{RQA}) e, finalmente, da entropia de recorrência (ENT_{RQA}) calculados para os grupos de indivíduos saudáveis e pós-AVE nas direções AP e ML. É possível perceber que apenas o parâmetro λ_{max} não apresentou diferença estatística em nenhuma das coordenadas do CP, enquanto que o parâmetro α diferiu significativamente ($p < 0.001$) apenas na coordenada AP. Todos os demais parâmetros exibiram diferença estatística tanto na coordenada AP ($p < 0.001$) como também na ML ($p < 0.05$).

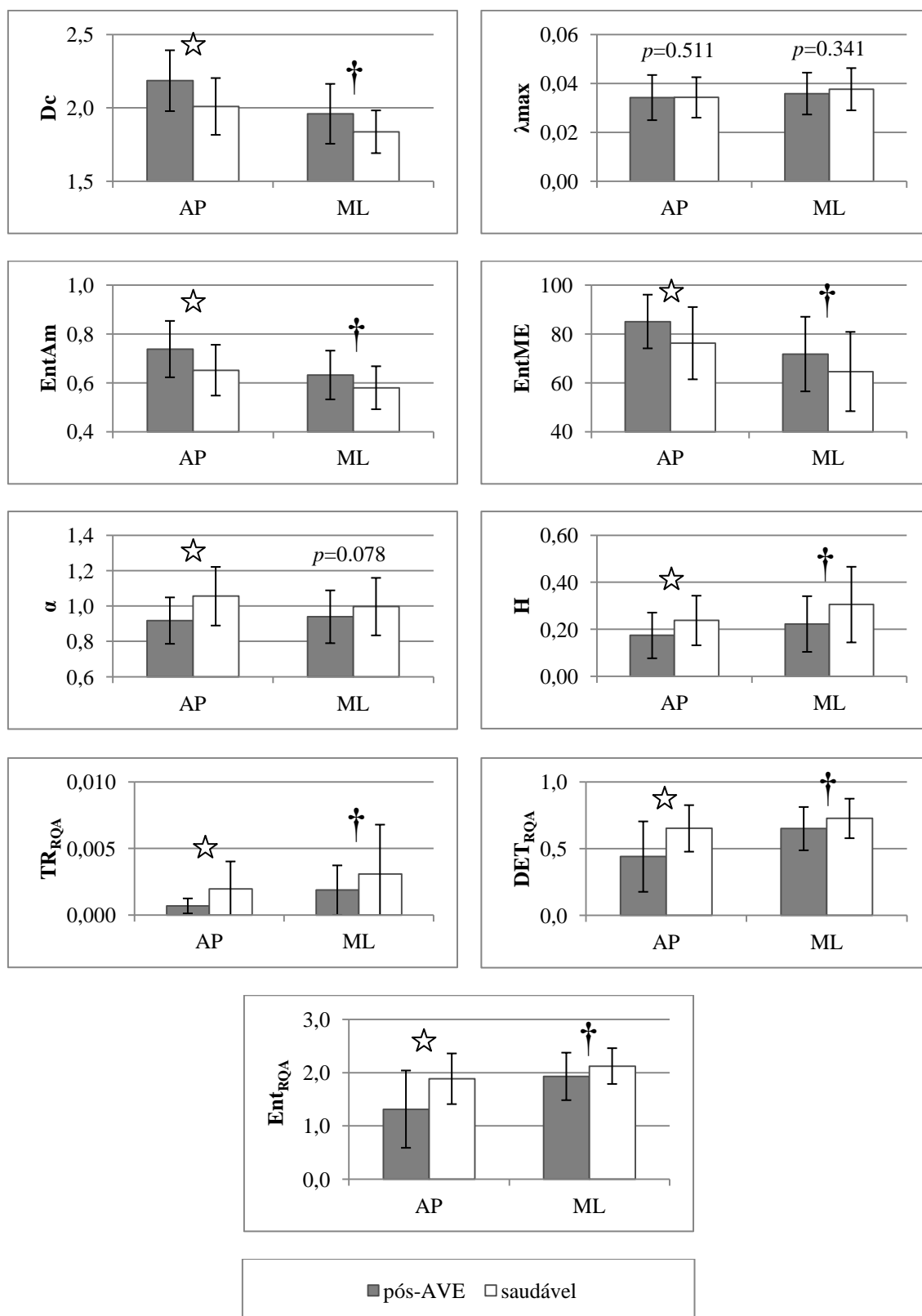


Figura 4.6 - Comparação entre os valores médios dos APE calculados para os indivíduos saudáveis (barras em branco) e para os pacientes pós-AVE (barras em cinza). As barras de erro representam o desvio padrão sobre os valores médios. As *estrelas* e as *cruzes* posicionadas em uma determinada coordenada indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente.

4.1.4. Desempenho da Seleção de Atributos

Durante o procedimento de redução de dimensionalidade das bases preparadas para as análises com os modelos de classificação, foram selecionados 07 atributos da base *APG*, 07 atributos da base *APE* e 11 atributos da base *APGE* (Tabela 4.2), os quais deram origem às novas bases *APG-selec*, *APE-selec* e *APGE-selec*, respectivamente.

Tabela 4.2 - Relação dos atributos selecionados a partir das bases de dados *APG*, *APE* e *APGE*.

Base original	Atributos selecionados	Coordenada
<i>APG</i>	<i>Vmed</i>	AP
	<i>DP Vmed</i>	AP
	<i>Fmed</i>	AP
	<i>F₅₀</i>	AP
	<i>F₈₀</i>	AP
	<i>F₉₅</i>	AP
	<i>F₉₅</i>	ML
<i>APE</i>	<i>D_C</i>	AP
	<i>D_C</i>	ML
	<i>EntAm</i>	AP
	α	AP
	<i>H</i>	AP
	<i>TR_{ROA}</i>	AP
	<i>ENT_{ROA}</i>	AP
<i>APGE</i>	<i>DP Vmed</i>	AP
	<i>Vp</i>	AP
	<i>F₈₀</i>	AP
	<i>F₉₅</i>	ML
	<i>D_C</i>	AP
	<i>D_C</i>	ML
	<i>EntAm</i>	AP
	α	AP
	<i>H</i>	AP
	<i>TR_{ROA}</i>	AP
	<i>ENT_{ROA}</i>	AP

Através da comparação estatística entre o desempenho dos classificadores para as bases de dados integrais e suas respectivas versões reduzidas (Figura 4.7), constatou-se que a redução de dimensionalidade proporcionou um aumento significativo ($p < 0.001$) no desempenho conjunto dos modelos preditivos para as bases *APE* e *APGE*, elevando a taxa média de acerto de 65.7% para 68.8% no primeiro caso (Figura 4.7, painel central) e de 67.4% para 69.0% no segundo caso (Figura 4.7, painel da direita). Entretanto, o efeito contrário foi observado para a base *APG* onde, após a seleção de atributos, a acurácia média conjunta dos modelos foi significativamente ($p < 0.05$) reduzida de 66.8% para 65.5% (Figura 4.7, painel da esquerda).

Acurácia Média Conjunta dos Classificadores

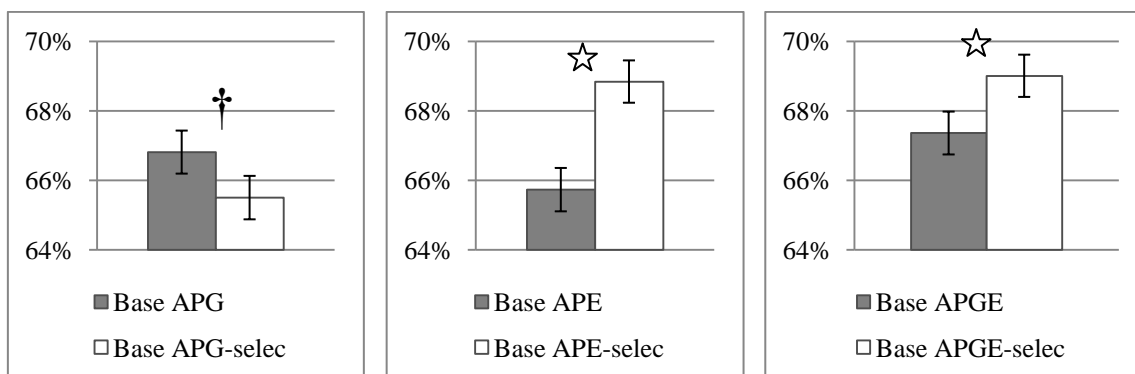


Figura 4.7 - Comparação entre as acurácias médias conjuntas dos classificadores para as versões integral e reduzida das bases *APG* (painel da esquerda), *APE* (painel central) e *APGE* (painel da direita). As barras de erro representam o intervalo de confiança de 95% sobre os valores médios. As *estrelas* e as *cruzes* indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente.

4.1.5. Desempenho dos Modelos de Classificação

Tendo em vista as bases integrais preparadas neste estudo para as análises com os modelos de classificação (vide subsecção 3.5.1), a base *APE* rendeu a menor ($p < 0.05$) taxa de acerto média conjunta entre os classificadores, 65.7%, ao passo que as bases *APG* e *APGE* resultaram acurácias médias conjuntas estatisticamente equivalentes ($p = 0.218$) iguais a 66.8% e 67.4%, respectivamente (Figura 4.8, painel da esquerda). Já para as bases de dados reduzidas (Figura 4.8, painel da direita), a menor ($p < 0.001$) taxa de acerto média conjunta foi obtida para a base *APG-selec*, 65.5%, enquanto que as bases *APE-selec* e *APGE-selec* renderam acurácias médias conjuntas estatisticamente equivalentes ($p = 0.705$) iguais a 68.8% e 69.0%, respectivamente.

Acurácia Média Conjunta dos Classificadores

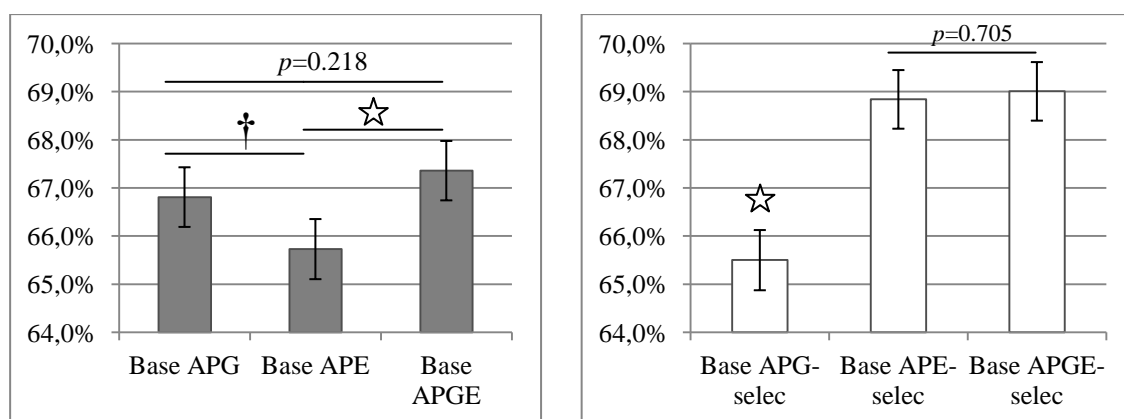


Figura 4.8 - Comparação entre as acurácias médias conjuntas dos classificadores para as bases de dados integrais (painel da esquerda) e reduzidas (painel da direita). As barras de erro representam o intervalo de confiança de 95% sobre os valores médios. As *estrelas* e as *cruzes* indicam diferença estatística com nível p de significância < 0.001 e < 0.05 respectivamente.

Na Tabela 4.3, são exibidas as acurácias individuais de cada um dos modelos de classificação considerados nesta pesquisa (vide Tabela 3.1) resultantes dos conjuntos de atributos posturográficos integrais (*APG*, *APE* e *APGE*) e selecionadas (*APG-selec*, *APE-selec* e *APGE-selec*).

Tabela 4.3 - Relação das acurácias individuais dos classificadores utilizados neste estudo para os diferentes conjuntos de atributos posturográficos.

Base	Acurácia Individual dos Classificadores (%)												
	PG ₁	PG ₂	AD	NB	SVM ₁	SVM ₂	SVM ₃	SVM ₄	MLP ₁	MLP ₂	1-NN	3-NN	5-NN
<i>APG</i>	64,2	63,3	56,1	67,5	67,5	73,7	64,0	70,2	67,9	67,8	68,4	67,5	70,2
<i>APE</i>	64,0	66,3	70,2	71,1	70,2	64,0	70,2	65,8	62,5	65,2	64,0	61,4	59,6
<i>APGE</i>	64,9	65,0	66,7	72,8	66,7	70,2	67,5	75,4	65,3	67,3	64,9	63,2	65,8
<i>APG-selec</i>	65,6	66,8	65,8	69,3	67,5	67,5	70,2	66,7	61,6	64,4	53,5	64,9	67,5
<i>APE-selec</i>	67,0	67,2	65,8	69,3	71,1	66,7	71,9	68,4	66,0	68,4	70,2	70,2	72,8
<i>APGE-selec</i>	67,0	68,1	64,9	73,7	73,7	65,8	72,8	69,3	64,4	67,7	68,4	70,2	71,1

4.2. Discussão

4.2.1. Esforço Requerido Durante a Extração dos Atributos

Ao invés de investigar melhorias em determinados métodos de classificação ou qualquer outra tarefa focada no aprimoramento de algoritmos de aprendizagem supervisionada, o principal viés desta pesquisa residiu no pré-processamento da base de dados posturográficos coletada originalmente por Silva [SIL12], mais especificamente, na extração de 25 variáveis posturográficas (16 de avaliação global e 09 de avaliação estrutural) comuns no campo da avaliação do equilíbrio a partir de cada uma das 228 séries temporais do CP disponíveis na base supracitada. Neste contexto, estimou-se o "esforço computacional" para a extração de tais atributos através do tempo de execução das rotinas implementadas neste estudo.

Como pode ser vislumbrado na Tabela 4.1, enquanto eram necessários em média 54.7 s para a extração dos APE a partir de uma determinada instância, os APG foram calculados (na média) em apenas 11.7 ms para a mesma instância. Em outras palavras, os atributos globais considerados nesta pesquisa foram extraídos em média 4688.6 vezes mais rapidamente do que os atributos estruturais. Esta situação já era esperada uma vez que os métodos implementados na rotina de processamento global são consideravelmente mais simples, envolvendo apenas equações matemáticas determinísticas (ver subseção 2.1.3.1). Em contrapartida, na rotina de processamento estrutural, foram implementados diversos métodos não-lineares e estatísticos, os quais necessitavam ser executados diversas vezes para estimação de um valor médio. No caso da dimensão de correlação, por exemplo, aquele parâmetro foi estimado dentro de 15 repetições para cada reconstrução do atrator em um espaço de fases m -dimensional fazendo $m = 1, 2, \dots, 15$. Além disso, para cada repetição, a soma de correlação — que é uma função computacionalmente custosa — foi calculada para 200 raios de busca distintos. É importante ressaltar que os tempos indicados na Tabela 4.1 para a extração do λ_{max} e das métricas de recorrência são relativamente baixos pois, para aqueles cálculos, aproveitou-se os valores do atraso temporal (τ) e da dimensão de imersão (m) obtidos durante a estimação da D_C .

Além de um maior esforço computacional, o cálculo dos atributos de avaliação estrutural do equilíbrio exigiu também um maior esforço intelectual pois, para este

propósito, fez-se necessário um estudo aprofundado sobre a teoria dos sistemas dinâmicos caóticos determinísticos. Além disso, enquanto os APG estão relacionados a uma única propriedade dos mecanismos de controle motor, a estabilidade, os APE descrevem uma série de outras propriedades daqueles mecanismos cujo entendimento é razoavelmente mais complexo, existindo inclusive divergências na literatura sobre a interpretação de alguns parâmetros desta categoria no contexto do equilíbrio (por exemplo, consultar [KIR13] para divergências sobre a interpretação da entropia amostral). Outro fator de complexidade em relação aos APE é a avaliação da natureza dos dados posturográficos, a qual será discutida na próxima subseção. Esta avaliação é necessária pois as métricas estruturais dependentes da reconstrução de um atrator — D_C , λ_{max} e métricas de recorrência — podem fornecer informações incoerentes caso os dados em estudo apresentem natureza predominantemente estocástica. Já no caso dos atributos globais, apenas o "tamanho" (ou amplitude) das trajetórias do CP são levadas em conta, sendo o comportamento da sua ocorrência temporal, portanto, irrelevante.

4.2.2. Natureza dos Dados Posturográficos

Durante a extração dos APE a partir dos dados originais do CP e suas versões *surrogate*, não foi possível determinar uma dimensão de imersão m apropriada para a reconstrução do atrator de nenhuma das séries *surrogate* de amostras. Em outras palavras, a relação D_C versus m daquelas séries não satisfaz o critério de Takens ($m \geq 2D_C + 1$) para nenhum $m = 1, 2, \dots, 15$ (ver Figura 4.1), o que tornou impraticável a estimação de qualquer métrica dependente da reconstrução do atrator, tal como a própria dimensão de correlação e também o máximo expoente de Lyapunov. Esta mesma situação já foi relatada por estudos similares (e.g., [DON07; ROE06]) pois, como esperado, a dimensão do atrator (D_C) de um sistema aleatório — que é o caso de uma série *surrogate* — nunca se estabiliza no espaço de fases, independentemente do número de eixos (m) utilizado para representá-lo.

Por outro lado, a Figura 4.2 mostra que foi possível estimar tanto D_C como λ_{max} para as séries *surrogate* de fases, ou seja, foi possível determinar o valor de m para a reconstrução do atrator de todas aquelas séries. Esta situação pode parecer contraditória em um primeiro momento pois, assim como as séries *surrogate* de amostras, os dados *surrogate* de fases também possuem uma natureza predominantemente estocástica

[DOY05; GHO10]. Por isso, é importante lembrar que a reconstrução do atrator pelo método das coordenadas defasadas depende exclusivamente da correlação temporal do sinal em estudo (ver subseção 2.1.3.2), a qual é inexistente em uma série *surrogate* de amostras. Entretanto, a série *surrogate* de fases preserva a autocorrelação dos dados originais [ROE06] de forma que, mesmo não convergindo para um valor de saturação, sua relação D_C versus m é capaz de satisfazer o critério de Takens [TAK81] adotado neste estudo. Ainda assim, em concordância com outros estudos correlatos (e.g., [CAM07; DON07; DOY05; ROE06]), os valores médios dos parâmetros D_C e λ_{max} para as séries *surrogate* de fases foram estatisticamente ($p < 0.05$) superiores àqueles oriundos dos dados originais (Figura 4.2, painéis superiores), tanto para os indivíduos saudáveis como também para os pacientes pós-AVE em ambas as direções AP e ML, sugerindo assim a presença de componentes predominantemente determinísticas nos sinais originais do CP.

Em relação à entropia amostral e ao expoente de escala para os dados originais e *surrogate* (Figura 4.2, painéis inferiores), os resultados deste trabalho também vão de encontro com a literatura correlata (e.g., [DON07; DOY05; ROE06]) e apenas reforçam a conclusão de que os sinais do CP em estudo possuem natureza predominantemente determinística. No caso da entropia amostral (Figura 4.2, canto inferior esquerdo), foram observados valores médios estatisticamente ($p < 0.05$) superiores para as séries *surrogate* de fases e *surrogate* de amostras em todos os casos, ou seja, aqueles dados exibiram um grau de regularidade significativamente menor do que os dados originais para as duas populações nas coordenadas AP e ML de deslocamento do CP.

Com relação ao expoente de escala (Figura 4.2, canto inferior direito) percebe-se que, para indivíduos saudáveis e pós-AVE nas direções AP e ML, o valor médio de α decaiu significativamente ($p < 0.05$) para 0.5 — valor associado à sistemas estocásticos — após o embaralhamento temporal dos dados originais (i.e., para as séries *surrogate* de amostras), mas permaneceu estatisticamente ($p > 0.05$) inalterado quando embaralhou-se apenas as fases do espectro de Fourier daqueles dados e preservou-se a sua correlação temporal (i.e., para as séries *surrogate* de fases). Esta situação evidencia a presença de fortes correlações temporais nos dados originais do CP e, ao mesmo tempo, sugere a existência de estruturas dinâmicas determinísticas naqueles dados [ROE06].

4.2.3. Avaliação Tradicional da Relevância dos Atributos Posturográficos

Neste trabalho, antes da utilização dos métodos de classificação, foi realizada primeiramente uma avaliação estatística tradicional com o intuito de investigar a relevância individual de cada um dos atributos posturográficos considerados nesta pesquisa para diferenciar os grupos de sujeitos em estudo, procedimento este amplamente utilizado pelos trabalhos da área da posturografia.

Atributos Posturográficos Globais

Como pode ser vislumbrado na Figura 4.3, dentre todos os 08 APG de domínio temporal, nenhum dos 04 parâmetros relacionados com os deslocamentos do CP propriamente ditos foi capaz de diferenciar estatisticamente ($p < 0.05$) os grupos de sujeitos saudáveis e pós-AVE. No entanto, todas as outras 04 métricas associadas à velocidade dos deslocamentos do CP permitiram tal diferenciação (Figura 4.3, coluna da direita). De maneira similar, destacaram-se também todos os 04 atributos globais de domínio espectral escolhidos neste trabalho, os quais permitiram diferenciar com significância estatística ($p < 0.05$) as populações em estudo (Figura 4.5). Já em relação aos APG de domínio espaço-temporal, somente 02 dentre os 04 parâmetros considerados foram úteis para a tarefa em discussão (Figura 4.4, painéis superiores).

De maneira geral, sendo ou não sendo possível a distinção com significância estatística ($p < 0.05$) entre as populações, todos os APG deste estudo exibiram valores médios inferiores para o grupo dos sujeitos saudáveis em relação ao grupo dos pacientes pós-AVE em ambas as direções AP e ML (ver Figura 4.3, Figura 4.4 e Figura 4.5), coincidindo assim com os achados de outros estudos similares (e.g., [GHO10; HAN05; SIL12]). Tais resultados já eram esperados visto que, conforme mencionado na subseção 2.1.3.1, a avaliação global das trajetórias do CP remete à variabilidade postural do sujeito durante a manutenção do equilíbrio e, neste contexto, sabe-se que sistemas fisiológicos saudáveis produzem menor variabilidade do que aqueles com algum tipo de comprometimento (e.g., estado de saúde debilitado, idade avançada).

Por fim, é importante destacar também que, em todos os casos nos quais foi possível diferenciar estatisticamente as populações através dos APG unidimensionais (i.e., os temporais e os espectrais), esta distinção aconteceu com maior significância

para a coordenada AP ($p < 0.001$) do que para a coordenada ML ($p < 0.05$) de deslocamento do CP (ver Figura 4.3 e Figura 4.5). Em outras palavras, os perfis de variabilidade postural na direção AP foram mais relevantes para distinguir as estratégias de equilíbrio adotadas pelas populações em estudo. Em um trabalho recente, foram relatadas evidências de que a variabilidade do CP na direção AP é mais discriminante quando são manipuladas as condições visuais durante o controle do equilíbrio, ao passo que a variabilidade na coordenada ML torna-se mais relevante quando são alteradas as características da superfície de contato dos pés com a plataforma de força [GAN13]. Dessa forma, visto que a base de dados em estudo é composta por duas tarefas de equilíbrio com manipulação visual (condições OA e OF) e por apenas uma única tarefa de manipulação da superfície de equilíbrio (condição E), os resultados deste trabalho coincidem com os achados do estudo supracitado.

Atributos Posturográficos Estruturais

Em relação aos 09 APE considerados neste trabalho, o máximo expoente de Lyapunov (λ_{max}) (Figura 4.6, canto superior direito) foi o único parâmetro incapaz de diferenciar os grupos de sujeitos em pelo menos uma direção de deslocamento do CP adotando-se um nível de significância mínimo de 5% durante as comparações estatísticas. Entretanto, no estudo similar de Roerdink et al. [ROE06], os autores conseguiram diferenciar indivíduos saudáveis de pacientes pós-AVE com o auxílio desta métrica em ambas as direções AP e ML adotando também um nível de significância de 5% nas avaliações estatísticas. A disparidade entre os resultados deste trabalho e os achados de Roerdink et al. [ROE06] pode estar relacionada com certas características particulares das amostras avaliadas em cada estudo, ou também com o algoritmo utilizado para o cálculo do máximo expoente de Lyapunov. Neste trabalho, tal parâmetro foi estimado através da análise em escala logarítmica do erro de previsão (consultar subseção 2.1.3.2 para mais detalhes) enquanto que, no estudo de Roerdink et al. [ROE06], a estimação foi realizada com o auxílio do método proposto por Rosentein et al. [ROS93]. Diante do exposto, trabalhos futuros são necessários para investigar com maior precisão os fatores que acarretaram a discrepância entre os achados deste estudo e os resultados da literatura.

Em relação ao expoente de escala (α), tal parâmetro foi capaz de diferenciar estatisticamente ($p < 0.05$) as populações apenas na coordenada AP do CP (Figura 4.6, terceiro painel da primeira coluna). Mais especificamente, o valor médio de α para os pacientes pós-AVE na coordenada AP situou-se entre 0.5 e 1.0, diferindo significativamente daquele valor médio entre 1.0 e 1.5 observado para os indivíduos saudáveis. Outro estudo na literatura obteve estes mesmos resultados durante a avaliação do equilíbrio em indivíduos jovens e idosos [DUA08]. Neste cenário, é importante lembrar que o expoente de escala descreve a "aspereza" de uma série temporal, sendo que quanto maior o valor de α , mais "suave" é a aparência da evolução temporal da série (i.e., exibe transições menos abruptas) [DON07; PEN95]. Assim, como pode ser visto na Figura 2.14, os achados supracitados sugerem que o balanço postural produzido tanto por sistemas fisiológicos saudáveis como também por sistemas comprometidos pelo AVE ou pela idade avançada se distinguem de ruído do tipo $1/f$, assemelhando-se a outros tipos de processos. A oscilação corporal dos sujeitos saudáveis assemelhou-se mais a um ruído Browniano ($\alpha = 1.5$), que é um processo com transições mais suaves ao longo do tempo e que possui apenas correlações locais [DUA08; KIR13]. Já os pacientes pós-AVE deste estudo, assim como os indivíduos idosos do estudo de Duarte e Sternad [DUA08], produziram oscilações posturais mais semelhantes ao ruído branco ($\alpha = 0.5$), cuja evolução temporal exibe transições menos "suaves" ao longo do tempo devido à sua natureza completamente imprevisível.

Com exceção do máximo expoente de Lyapunov e do expoente de escala, todos os outros 07 APE foram capazes de diferenciar estatisticamente os dois grupos de sujeitos em ambas as direções AP ($p < 0.001$) e ML ($p < 0.05$) de deslocamento do CP (Figura 4.6). No caso da dimensão de correlação (Figura 4.6, canto superior esquerdo), foram observados valores médios de D_C significativamente maiores para os pacientes pós-AVE. Estes resultados coincidem com os achados de Roerdink et al. [ROE06], sendo que aqueles autores também avaliaram o controle postural em indivíduos saudáveis e pós-AVE mediante 03 tarefas distintas de equilíbrio, sendo elas: sobre uma plataforma de força com os olhos abertos, com os olhos fechados, e com os olhos abertos realizando uma segunda tarefa cognitiva. Em contrapartida, após avaliar o controle do equilíbrio nas mesmas populações em questão somente através da condição experimental "sobre uma plataforma de força com os olhos abertos", Ghomashchi et al. [GHO10] encontraram resultados divergentes, ou seja, obtiveram uma D_C média

estatisticamente ($p = 0.000$) superior para os indivíduos saudáveis, tanto na direção AP como também na ML. Uma possível interpretação para estes achados seria que, para lidar com tarefas motoras mais simples e que fornecem menores riscos de queda (e.g., permanecer sobre superfícies regulares sem sofrer privações sensoriais), o maior número de graus de liberdade ativos — descritos pela D_C — observado por Ghomashchi et al. [GHO10] para os indivíduos saudáveis reflete a maior "eficiência" do SCP daquela população frente aos pacientes doentes. No entanto, quando a tarefa motora torna-se mais complexa, os pacientes pós-AVE passam a recrutar processos de controle adicionais (i.e., mais graus de liberdade ativos) visando compensar as deficiências de seus mecanismos de controle postural. Por fim, é importante salientar que, tanto nesta pesquisa como também nos estudos de Roerdink et al. [ROE06] e de Ghomashchi et al. [GHO10] que envolveram as mesmas populações, o parâmetro D_C mostrou-se um bom descritor do equilíbrio em ambas as direções AP e ML.

Em relação à entropia amostral (Figura 4.6, segundo painel da primeira coluna), as trajetórias do CP dos pacientes pós-AVE produziram uma *EntAm* média estatisticamente superior em comparação com o grupo dos sujeitos saudáveis nas direções AP ($p < 0.001$) e ML ($p < 0.05$). Dessa forma, dentre as duas populações avaliadas neste trabalho, observou-se um menor grau de regularidade (ou previsibilidade) para os perfis de balanço postural produzidos pelos pacientes pós-AVE. No estudo similar de Roerdink et al. [ROE06], aqueles autores também observaram menor regularidade nas trajetórias do CP de um grupo de pacientes pós-AVE em comparação com indivíduos saudáveis na direção AP, mas o oposto foi relatado para a coordenada ML. Além disso, no trabalho de Duarte e Sternad [DUA08] envolvendo sujeitos jovens e idosos, os maiores valores de entropia nas direções AP e ML também foram associados aos indivíduos com mecanismos de controle postural mais debilitados: os idosos. Por outro lado, o estudo de Donker et al. [DON08] constatou menor regularidade (i.e., maior entropia amostral) no balanço postural de crianças saudáveis em comparação com crianças acometidas por paralisia cerebral. Por fim, Madaleine et al. [MAD11] não conseguiram diferenciar sujeitos saudáveis de indivíduos com traumas na coluna cervical através da *EntAm*. Todos estes relatos sobre a entropia amostral sugerem que pode não haver um padrão de regularidade nas trajetórias do CP produzidas por sistemas fisiológicos com diferentes tipos de comprometimento em comparação com sistemas saudáveis, como por exemplo: "*indivíduos com algum tipo de*

acometimento produzem um balanço postural mais/menos regular do que sujeitos saudáveis".

No caso da entropia multiescala, observou-se um valor médio estatisticamente superior daquele parâmetro para os pacientes pós-AVE em comparação com os indivíduos saudáveis em ambas as direções AP ($p < 0.001$) e ML ($p < 0.05$) de deslocamento do CP (Figura 4.6, segundo painel da segunda coluna). Em um estudo similar de avaliação do equilíbrio, Duarte e Sternad [DUA08] também obtiveram valores estatisticamente ($p < 0.05$) superiores de *EntME* para indivíduos idosos quando comparados à sujeitos jovens saudáveis. Tais resultados sugerem que os mecanismos biológicos de controle postural de pessoas comprometidas tanto pela idade avançada como também pelo AVE possuem maior grau de complexidade do que os mecanismos de pessoas jovens e/ou saudáveis.

Através do expoente H de Hurst (Figura 4.6, terceiro painel da segunda coluna), constatou-se um comportamento de antipersistência ($0 < H < 0.5$) no balanço postural das duas populações em estudo, ou seja, os deslocamentos do CP dos indivíduos saudáveis e pós-AVE exibiram correlações negativas em ambas as direções AP e ML. Além disso, em concordância com os resultados do expoente de escala, os perfis de oscilação corporal produzidos pelos indivíduos saudáveis se assemelharam mais à um ruído Browniano ($H = 0.5$), o qual apresenta correlações locais e transições mais "suaves" ao longo do tempo [DUA08], ao passo que a oscilação postural dos pacientes pós-AVE assemelhou-se mais à um ruído branco ($H = 0$), ou seja, à um processo de natureza completamente aleatória e com transições temporais mais abruptas. Tais achados são condizentes também com os resultados da D_C e da *EntAm* alcançados neste trabalho. Isso porque, em comparação com os indivíduos saudáveis, os pacientes pós-AVE recrutaram processos de controle adicionais (i.e., maior D_C) durante as tarefas de manutenção do equilíbrio, produzindo assim perfis de balanço postural com mais alto grau de irregularidade (i.e., maior *EntAm*) e que se assemelham mais à um processo com dinâmica aleatória (H próximo de zero, α próximo de 0.5), apresentando assim transições mais abruptas ao longo da evolução temporal (i.e., maior "aspereza"). Um outro estudo de avaliação do equilíbrio também constatou valores estatisticamente ($p < 0.05$) inferiores de H para um grupo de pacientes pós-AVE em comparação com outro grupo de sujeitos saudáveis, entretanto, apenas na coordenada AP [ROE06]. Complementarmente, no estudo de Kirchner et al. [KIR12], os autores observaram

valores estatisticamente ($p < 0.05$) inferiores de H para o balanço postural produzido em uma tarefa mais complexa de equilíbrio em comparação com outra tarefa mais simples, ambas realizadas por um grupo de adultos saudáveis. Portanto, tais achados sugerem que, quanto piores as condições de manutenção do equilíbrio — seja em relação ao estado de saúde ou às condições sensoriais da tarefa —, maior a "aspereza" e irregularidade das trajetórias do CP ao longo do tempo.

Em relação às medidas de recorrência consideradas neste trabalho (Figura 4.6, três últimos painéis), observou-se valores médios estatisticamente inferiores dos parâmetros TR_{RQA} , DET_{RQA} e ENT_{RQA} para os pacientes pós-AVE em ambas as direções AP ($p < 0.001$) e ML ($p < 0.05$). Portanto, em termos de estruturas determinísticas dos gráficos de recorrência produzidos a partir dos dados posturográficos, os indivíduos doentes exibiram menor regularidade e também menor complexidade em relação aos sujeitos saudáveis. No estudo de Negahban et al. [NEG13], os autores observaram resultados similares quando compararam situações distintas de controle postural que, naquele caso, não estavam associadas à condição de saúde dos sujeitos, mas sim às condições de equilíbrio em que as populações foram submetidas. Mais especificamente, aqueles autores observaram que os perfis de balanço postural de indivíduos saudáveis e de pacientes com esclerose múltipla tornaram-se menos regulares (i.e., diminuição da TR_{RQA} e da DET_{RQA}) e também menos complexos (i.e., diminuição da ENT_{RQA}) em estruturas determinísticas a medida em que intervenções sensoriais eram realizadas a fim de dificultar a tarefa de equilíbrio. Devido à inexistência de outros estudos de avaliação postural de indivíduos saudáveis e pós-AVE baseada na técnica RQA, não é possível tecer uma discussão mais aprofundada sobre o comportamento das métricas de recorrência consideradas neste trabalho para as populações supracitadas.

Finalmente, assim como observado para os APG, todos os APE capazes de diferenciar estatisticamente os dois grupos de sujeitos o fizeram com maior significância na coordenada AP ($p < 0.001$) do que na coordenada ML ($p < 0.05$) (Figura 4.6). Estes resultados podem sugerir que, quando são manipuladas as condições visuais durante o controle do equilíbrio, não somente a variabilidade relatada em [GAN13] mas também a organização temporal das trajetórias do CP na direção AP são mais relevantes para a descrição do equilíbrio em indivíduos saudáveis e pós-AVE em comparação com a coordenada ML.

4.2.4. Avaliação da Relevância Conjunta dos Atributos Posturográficos

Ao invés do procedimento tradicionalmente empregado nos estudos de avaliação do equilíbrio, o qual consiste em analisar individualmente o poder de descrição de cada um dos parâmetros posturográficos calculados, o foco deste trabalho residiu na investigação da relevância de tais parâmetros para, em conjunto, distinguirem os perfis de oscilação corporal produzidos por indivíduos saudáveis e pós-AVE. Esta tarefa somente foi possível com o auxílio de métodos de classificação, os quais consistem em ferramentas computacionais capazes de diferenciar — com uma determinada precisão — os grupos (ou classes) a partir de uma análise conjunta das variáveis disponíveis; no caso, os atributos posturográficos.

Redução de Dimensionalidade

Primeiramente em relação ao conjunto de atributos posturográficos globais considerados, a redução de dimensionalidade realizada na base *APG* provocou uma diminuição estatisticamente significativa ($p < 0.05$) na acurácia média conjunta dos modelos de previsão avaliados neste trabalho (Figura 4.7, painel da esquerda). Em outras palavras, a exclusão de determinados atributos globais resultou em uma perda significativa de informação relevante sobre as populações do ponto de vista dos classificadores mencionados. Nesse contexto, é importante lembrar que a literatura tem associado os atributos globais em geral à uma única propriedade dos mecanismos de controle postural, a estabilidade [KIR13; STE11]. Portanto, os resultados mencionados sugerem que todas as 16 métricas globais consideradas neste estudo (vide Tabela 2.1) são relevantes para descrever a estabilidade postural de indivíduos saudáveis e pós-AVE através de métodos tradicionais de classificação. Entretanto, em um outro estudo similar, a redução de dimensionalidade sobre uma base de atributos posturográficos globais (154 instâncias, 23 atributos cada) foi capaz de aprimorar a acurácia de classificadores populares para diferenciar determinadas tarefas motoras realizadas por um único grupo de indivíduos jovens saudáveis [SAR13]. Diante do exposto, é possível que a redução de dimensionalidade sobre atributos posturográficos globais seja vantajosa apenas para diferenciar perfis de oscilação corporal produzidos pela mesma população sob diferentes condições de manutenção do equilíbrio, tal como realizado em

Saripalle et al. [SAR13], mas perde a eficácia quando o objetivo é distinguir os perfis posturais produzidos por populações distintas, assim como neste trabalho. Em contrapartida, é possível que o conflito entre os resultados reportados neste estudo e os achados de Saripalle et al. [SAR13] deva-se aos diferentes conjuntos de atributos globais utilizados em cada trabalho. Dessa forma, estudos futuros são necessários para avaliar tais hipóteses.

Já com relação aos atributos posturográficos estruturais, os quais estão associados à diversas propriedades do SCP, a redução de dimensionalidade na base *APE* proporcionou um aumento estatisticamente significativo ($p < 0.001$) na acurácia média conjunta dos classificadores (Figura 4.7, painel central). Dessa forma, evidencia-se que algumas das propriedades descritas pelos parâmetros estruturais considerados neste trabalho são mais relevantes do que outras para diferenciar indivíduos saudáveis de pacientes pós-AVE (e vice-versa). Mais especificamente, dentre estas propriedades mais relevantes (Tabela 4.2), encontram-se o número de graus de liberdade ativos do SCP descrito pela D_C , o grau de regularidade da oscilação corporal descrito pela *EntAm*, o tipo de correlação temporal existente no balanço postural descrito pelos parâmetros α e H e, por fim, a regularidade e a complexidade das estruturas determinísticas dos gráficos de recorrência descritas respectivamente pelos parâmetros TR_{RQA} e ENT_{RQA} . Neste mesmo contexto, foram desconsideradas a sensibilidade dos mecanismos de controle postural à pequenas perturbações locais (λ_{max}) e a complexidade da ocorrência temporal do balanço postural (*EntME*). Finalmente, foi desconsiderado também o parâmetro DET_{RQA} que, no entanto, exerce função similar à métrica TR_{RQA} já considerada pela seleção de atributos. Infelizmente, a validação destes resultados é impraticável devido à inexistência de outros estudos com metodologia similar na literatura atual.

No caso da utilização conjunta dos atributos globais e estruturais, a redução de dimensionalidade realizada na base *APGE* resultou em um aumento estatisticamente significativo ($p < 0.001$) da acurácia média conjunta dos classificadores (Figura 4.7, painel da direita), tal como observado para a base *APE*. No entanto, é importante lembrar que a seleção de atributos na base *APG* produziu o efeito contrário, ou seja, redução no poder de previsão dos modelos. Uma possível interpretação para estes achados é que, quando utilizou-se apenas atributos globais (base *APG*), a redução de dimensionalidade causou uma perda significativa de informação sobre a única

propriedade descrita por aqueles parâmetros: a estabilidade postural das populações. Porém, após a adição de atributos estruturais ao conjunto de atributos globais (base *APGE*), diversas outras propriedades passaram também a descrever os mecanismos de controle postural das populações em estudo, sendo que a redução de dimensionalidade permitiu selecionar apenas as mais discriminantes entre elas.

É importante salientar também que, dentre os 11 atributos posturográficos selecionados a partir da base *APGE*, 04 são globais e 07 são estruturais, sendo que estes últimos são exatamente os mesmos selecionados a partir da base *APE* (ver Tabela 4.2). Primeiramente, estes resultados sugerem que tanto a variabilidade como também a organização temporal das trajetórias do CP são relevantes para distinguir indivíduos saudáveis e pós-AVE com o auxílio de métodos de classificação. Além disso, pode-se concluir também que as propriedades do SCP descritas pelos parâmetros estruturais D_C , $EntAm$, α , H , TR_{RQA} e ENT_{RQA} (ver Tabela 4.2) são relevantes para a distinção das populações mencionadas, tanto na presença como também na ausência de atributos de avaliação global do equilíbrio.

Desempenho dos Modelos de Classificação

Conforme pode ser vislumbrado na Figura 4.8, as bases de dados reduzidas utilizadas neste estudo (painel da direita) proporcionaram melhores performances pelos modelos de classificação em comparação com o uso das bases integrais (painel da esquerda). Dessa forma, direcionando a discussão para as bases reduzidas, a utilização exclusiva de atributos estruturais selecionados (base *APE-selec*) rendeu uma taxa de acerto média conjunta estatisticamente ($p < 0.001$) superior em relação ao uso exclusivo de atributos globais selecionados (base *APG-selec*) (Figura 4.8, painel da direita). Esta situação já era esperada uma vez que os atributos estruturais descrevem diversas propriedades dos mecanismos biológicos de controle postural (e.g., graus de liberdade ativos, regularidade, complexidade, etc.), ao passo que os atributos globais descrevem apenas uma única propriedade, a estabilidade.

Além disso, esperava-se também que a combinação entre todas as propriedades supracitadas do SCP alcançada através da utilização conjunta de atributos globais e estruturais selecionados (base *APGE-selec*) resultasse um desempenho superior em relação ao uso de apenas atributos globais selecionados (base *APG-selec*) ou de apenas

atributos estruturais selecionados (base *APE-selec*). Entretanto, em desacordo com esta premissa, a base *APGE-selec* rendeu uma acurácia média conjunta estatisticamente equivalente ($p = 0.705$) em comparação com a base *APE-selec* (Figura 4.8, painel da direita). Tal achado sugere que, do ponto de vista dos classificadores considerados nesta pesquisa, não existem vantagens em se calcular e adicionar atributos posturográficos globais a um determinado conjunto de atributos estruturais (ver Tabela 4.2) para a identificação de sujeitos saudáveis e pós-AVE. Mais especificamente, diante das propriedades do SCP descritas pelos parâmetros estruturais D_C , $EntAm$, α , H , TR_{RQA} e ENT_{RQA} (base *APE-selec*), a adição de informações sobre a estabilidade postural dos sujeitos (i.e., a adição de atributos globais) não apresentou nenhuma relevância para distinguir as populações em questão. Por fim é importante ressaltar que, devido à inexistência de outros trabalhos com metodologia similar na literatura para a validação dos resultados, estudos futuros são necessários para reforçar todas as conclusões supracitadas.

Relação Custo/Benefício do Uso de Atributos Globais e Estruturais

A Figura 4.8 mostra que, para os conjuntos integrais de atributos utilizados neste estudo (painel da esquerda), o uso exclusivo de atributos globais (base *APG*) e a combinação entre atributos globais e estruturais (base *APGE*) renderam as melhores taxas de acerto médias conjuntas entre os classificadores, as quais são estatisticamente equivalentes ($p = 0.218$). Neste caso, portanto, a utilização da base *APG* representa a opção mais vantajosa para diferenciar as populações em estudo pois envolve apenas o cálculo de parâmetros posturográficos da categoria global, ao passo que a base *APGE* envolve a extração de parâmetros das duas categorias, global e estrutural, requerendo assim um maior esforço intelectual e computacional. Aplicando este mesmo raciocínio para os conjuntos de atributos reduzidos desta pesquisa (Figura 4.8, painel da direita) é possível concluir que, naquele caso, a utilização exclusiva de atributos estruturais selecionados (base *APE-selec*) é a opção mais vantajosa para diferenciar as populações através dos modelos de classificação considerados. Sendo assim, para fins de comparação entre as opções mais vantajosas supracitadas, tem-se uma acurácia média conjunta de 66.8% resultante da base *APG* no caso das bases integrais (Figura 4.8,

painel da esquerda) *versus* uma acurácia média conjunta de 68.8% oriunda da base *APE-selec* no caso das bases reduzidas (Figura 4.8, painel da direita).

Neste ponto da discussão, é importante lembrar que o cálculo de parâmetros posturográficos estruturais possui diversos fatores de complexidade em relação ao cálculo de parâmetros globais. Em primeiro lugar, pode-se mencionar a investigação do comportamento do sinal do CP quanto a determinístico ou estocástico, procedimento necessário para o cálculo das métricas estruturais dependentes da reconstrução de um atrator — D_C , λ_{max} e métricas de recorrência — mas totalmente dispensável para o cálculo de atributos globais, pois consideram apenas a variabilidade em amplitude daquele sinal. Neste cenário, a avaliação da natureza do balanço postural — realizada neste trabalho através da técnica das séries *surrogate* (ver subseção 3.3) — insere complexidade e eleva o tempo dedicado à análise dos dados. Além disso, a aplicabilidade dos parâmetros estruturais sensíveis à natureza do sinal se restringe ao resultado daquela avaliação de forma que, se forem constatadas componentes predominantemente estocásticas no sinal do CP, a extração de tais parâmetros não é recomendada. Esta restrição pode ser prejudicial em estudos de longo prazo (e.g., artigo científico, dissertação, tese), exigindo uma eventual redução na amostra do estudo através da exclusão de determinados sinais, mas pode ser ainda mais problemática em análises de tempo real onde o profissional da saúde (no caso de avaliações hospitalares/clínicas) ou um determinado dispositivo eletrônico (no caso de sistemas de detecção de queda) necessita de um *feedback* imediato sobre o desempenho/situação de um paciente. Ademais, como elucidado previamente na subseção 3.2.2, alguns dos parâmetros estruturais disponíveis na literatura são sensíveis também à não-estacionariedade do sinal em estudo (e.g., dimensão de correlação, máximo expoente de Lyapunov, entropia multiescala), sendo necessária neste caso a utilização de técnicas para remover esta não-estacionariedade — tal como a técnica EMD utilizada nesta pesquisa (ver subseção 3.2.2) —, enquanto que a precisão dos parâmetros globais não é influenciada por este fator.

Outro ponto a ser lembrado ainda é a teoria de fundamentação dos parâmetros posturográficos globais e estruturais. No primeiro caso, por extraírem informações sobre a variabilidade da oscilação postural, os parâmetros globais envolvem análises matemáticas determinísticas e pouco sofisticadas sobre o sinal do CP, tal como o cálculo de médias, desvio padrão, amplitudes de pico, etc. (consultar a subseção 2.1.3.1

para mais detalhes). Por outro lado, alguns dos parâmetros estruturais são fundamentados na teoria de sistemas dinâmicos caóticos determinísticos, envolvendo análises matemáticas estatísticas e potencialmente mais complexas sobre o sinal do CP. Sendo assim, a familiarização com parâmetros posturográficos da categoria estrutural exige do pesquisador, em geral, um esforço intelectual mais significativo em relação à familiarização com parâmetros de categoria global. Ademais, a maior complexidade matemática existente na análise estrutural do balanço postural é um outro fator que atribui maior dificuldade e que eleva o tempo gasto com o processamento dos dados frente à análise global.

Neste cenário, retomando a comparação entre as opções mais vantajosas de diferenciação entre indivíduos saudáveis e pós-AVE através dos métodos de classificação utilizados nesta pesquisa, foram necessário apenas 1.33 segundos para a extração de todos os atributos globais que compõem a base *APG* (vide Tabela 2.1), a qual rendeu uma acurácia média conjunta de 66.8% (Figura 4.8, painel da esquerda). Em contrapartida, foram gastos aproximadamente 5415,57 segundos (ou 1.5 horas)¹ para o cálculo dos atributos estruturais que compõem a base *APE-selec* (vide Tabela 4.2), a qual resultou uma taxa de acerto média conjunta de 68.8% (Figura 4.8, painel da direita). Em outras palavras, a base *APE-selec* foi extraída aproximadamente 4071.9 vezes mais lentamente do que a base *APG* e produziu uma taxa de acerto média conjunta superior em apenas 2.0%.

Diante de todos os fatores discutidos nesta subseção, a viabilidade do uso do subconjunto selecionado de atributos estruturais (base *APE-selec*) frente ao conjunto integral de atributos globais (base *APG*) considerados nesta pesquisa é, certamente, passível de questionamentos, principalmente pelo fato desta primeira base ter sido obtida às custas de um processamento consideravelmente mais complexo e computacionalmente mais custoso (i.e., mais lento) do que a segunda base, e por ter rendido uma acurácia média conjunta superior em apenas 2.0% conforme mencionado no parágrafo anterior, ou seja, por não ter apresentado grandes vantagens na diferenciação de sujeitos saudáveis e pós-AVE através de modelos populares de

¹ Para a estimação deste tempo de execução, somou-se o tempo médio de extração por instância relatado na Tabela 4.1 para os parâmetros estruturais D_C , $EntAm$, α , H , TR_{RQA} e ENT_{RQA} , os quais compõem a base *APE-selec*. Com exceção da D_C , os outros parâmetros mencionados foram selecionados apenas em uma coordenada do CP (vide Tabela 4.2) e, portanto, foi considerado a metade do tempo indicado na Tabela 4.1 para estes parâmetros durante a soma supracitada. Por fim, o valor resultante da soma foi multiplicado pelo número total de instâncias: 114.

classificação. De maneira mais geral, ou seja, para outros contextos de avaliação do equilíbrio motor em estudos futuros, a viabilidade do cálculo de parâmetros posturográficos de uma ou outra categoria pode depender de diversos fatores, tal como a familiaridade do pesquisador (ou o tempo hábil para se familiarizar) com os procedimentos matemáticos envolvidos no cálculo dos parâmetros, da necessidade de se obter um *feedback* do paciente em tempo real e, conseqüentemente, do tempo disponível para a extração dos parâmetros, do objetivo da pesquisa, entre outros.

4.2.5. Algumas Observações Sobre as Avaliações Individual e Conjunta dos Parâmetros Posturográficos

Neste trabalho, duas abordagens distintas foram empregadas com o intuito de tentar diferenciar os perfis de oscilação postural produzidos por indivíduos saudáveis e pós-AVE através da análise de uma determinada relação de atributos posturográficos globais e estruturais. Na primeira abordagem, a chamada avaliação tradicional, investigou-se a *relevância individual* de cada um dos parâmetros considerados nesta pesquisa com o auxílio de comparações estatísticas, conforme descrito em detalhes na seção 3.4 e discutido na subseção 4.2.3. Já na segunda abordagem — que é o foco deste trabalho —, avaliou-se a *relevância conjunta* de tais parâmetros para a identificação das populações em estudo através de determinados métodos de classificação, assim como descrito detalhadamente na seção 3.5 e discutido na subseção 4.2.4. Neste último caso, em outras palavras, investigou-se quais parâmetros posturográficos eram mais relevantes para, em grupo, discriminarem o balanço postural produzido por sistemas fisiológicos saudáveis e acometidos pelo AVE. Neste ponto da discussão, é importante salientar que o presente trabalho não teve como objetivo comparar a eficiência das duas abordagens de avaliação supracitadas e, dessa forma, nenhuma discussão nesse sentido será realizada nesta subseção. Ao invés disso, será discorrido sobre alguns aspectos em comum e alguns outros aspectos particulares observados durante a análise dos resultados em relação às abordagens de avaliação individual e conjunta dos parâmetros posturográficos.

Em primeiro lugar, após a etapa de seleção de atributos realizada durante a avaliação da relevância conjunta dos parâmetros posturográficos (ver subseção 3.5.2 para mais detalhes), constatou-se que todos os atributos selecionados a partir das bases

APG (composta somente por atributos globais), *APE* (composta somente por atributos estruturais) e *APGE* (composta pela união das duas bases anteriores) (ver Tabela 4.2) foram capazes também de diferenciar os grupos de sujeitos a partir da análise tradicional dos dados (ver Figura 4.3, Figura 4.4, Figura 4.5 e Figura 4.6). Em outras palavras, constatou-se que os melhores atributos posturográficos para diferenciar as populações através de uma interação conjunta — por meio dos modelos preditivos — foram também capazes de fazê-lo de forma individual, através de uma simples comparação estatística entre grupos. Entretanto, é importante ressaltar que a situação contrária não foi observada, ou seja, nem todos os atributos capazes de distinguir as populações na avaliação tradicional foram escolhidos durante a seleção de atributos para a avaliação conjunta. Diante deste achado, é possível inferir que certos parâmetros posturográficos podem não se relacionar muito bem em grupo para diferenciar as populações em questão, mesmo sendo capazes de diferenciá-las quando analisados individualmente. De forma complementar, é possível inferir também que certos parâmetros podem apresentar um bom desempenho em grupo para diferenciar as populações mesmo não sendo capazes de diferenciá-las individualmente. Esta última inferência advém do fato do conjunto integral de atributos globais (base *APG*) — composto por parâmetros tanto capazes como também incapazes de diferenciar individualmente os grupos de sujeitos — ter rendido uma acurácia média conjunta dos classificadores estatisticamente ($p < 0.001$) superior do que o subconjunto selecionado de atributos globais (base *APG-selec*) (Figura 4.7, painel da esquerda), composto apenas por atributos com maior poder individual de discriminação.

Ainda no contexto da seleção de atributos, outro aspecto a ser discutido é que, com exceção do atributo global F_{95} na coordenada ML selecionado a partir das bases *APG* e *APGE* e do atributo estrutural D_C na coordenada ML selecionado a partir das bases *APE* e *APGE*, todos os outros atributos posturográficos selecionados para a avaliação conjunta a partir dos três conjuntos de atributos considerados neste estudo (*APG*, *APE* e *APGE*) pertencem à coordenada AP (ver Tabela 4.2). Esta situação vai de encontro com os resultados da avaliação tradicional dos dados pois, durante tal análise, todos os parâmetros que foram capazes de diferenciar estatisticamente os grupos de sujeitos o fizeram com maior significância para a coordenada AP ($p < 0.001$) do que para a coordenada ML ($p < 0.05$) (ver Figura 4.3, Figura 4.4, Figura 4.5 e Figura 4.6), conforme discutido anteriormente na subseção 4.2.3. Neste cenário é interessante

relembrar que, do ponto de vista da análise estatística tradicional, já se conhecia a relevância da coordenada AP frente a ML para diferenciar duas populações submetidas à intervenções visuais de equilíbrio através de parâmetros posturográficos globais [GAN13] e, neste trabalho, emergiram evidências desta mesma relevância para tal diferenciação baseada em atributos posturográficos estruturais. Além disso, os resultados discutidos neste parágrafo sugerem também a relevância da coordenada AP para, através da abordagem de avaliação conjunta baseada em modelos de classificação, diferenciar sujeitos saudáveis de pacientes pós-AVE através de um determinado subconjunto de atributos estruturais (base *APE-selec*), bem como deste mesmo subconjunto de atributos estruturais acrescido de certos atributos globais (base *APGE-selec*) (ver Tabela 4.2), situações estas em que o procedimento de seleção de atributos elevou a taxa de acerto média conjunta dos modelos preditivos com significância estatística ($p < 0.001$) (ver Figura 4.7). No entanto, estudos futuros são necessários para reforçar as inferências mencionadas, as quais podem estar condicionadas tanto à certas características como também ao tamanho da amostra utilizada neste trabalho.

Conclusão e Trabalhos Futuros

Este trabalho teve por objetivo avaliar a relevância do uso conjunto entre atributos das categorias global e estrutural comuns na área da posturografia em comparação com o uso de atributos de apenas uma ou outra categoria em particular para diferenciar as estratégias de equilíbrio adotadas por indivíduos saudáveis e por pacientes pós-AVE, utilizando para isso tanto métodos populares de classificação como também a Programação Genética, um método menos usual e com baixa difusão em problemas da área da saúde, mas que vem sendo utilizado com sucesso em outras áreas do conhecimento e cujo modelo preditivo resultante pode ser facilmente interpretado por profissionais da área da saúde.

Primeiramente, com relação aos conjuntos integrais de atributos (i.e., sem a realização de uma redução de dimensionalidade), comprova-se nesta pesquisa que o uso simultâneo dos atributos globais e estruturais considerados rende o mesmo desempenho médio conjunto entre os classificadores avaliados em comparação com o uso de apenas atributos globais. Além disso, após a realização de uma redução de dimensionalidade, o uso simultâneo dos atributos globais e estruturais passa a resultar o mesmo desempenho

médio conjunto entre os classificadores do que o uso de apenas atributos estruturais, constatações estas que conflitam com a hipótese de pesquisa inicialmente estipulada pelo autor. Tais achados devem ser vislumbrados com otimismo pela comunidade científica visto que, nas duas circunstâncias avaliadas neste trabalho (com e sem seleção de atributos), a utilização de parâmetros posturográficos de uma única categoria exibiu a mesma capacidade de diferenciar indivíduos saudáveis e pós-AVE em comparação com o uso simultâneo de atributos das duas categorias, sendo que este segundo caso intuitivamente exige maiores esforços do pesquisador/profissional da saúde em relação ao primeiro.

Em projetos futuros, novos estudos podem ser conduzidos para, por exemplo, validar os resultados reportados neste trabalho. É importante ressaltar, porém, que certas características da amostra além do estado de saúde — tais como estatura, massa corporal e idade — também estão relacionadas com o funcionamento dos mecanismos de controle postural do sujeito e, portanto, podem conduzir diferentes pesquisadores a diferentes conclusões, mesmo que tenha sido estudada a mesma condição de saúde nos dois casos (e.g., saudável, pós-AVE). Dessa forma, estudos futuros devem levar em conta as características da amostra durante a comparação de seus resultados com a literatura. Além disso, trabalhos futuros podem verificar se os resultados desta pesquisa se repetem na presença de outros fatores além do AVE e que também exercem influência no sistema de controle postural, tal como a idade avançada, a condição física e até mesmo outros tipos de acometimento neurológico (e.g., paralisia cerebral, esclerose múltipla, etc.).

Em ambos os cenários supracitados, como os atributos posturográficos estruturais são, de maneira geral, mais complexos de se compreender, mais lentos de se calcular e com aplicação mais restrita — pois dependem do comportamento do perfil de balanço postural em análise — em comparação com os atributos globais, a decisão pelo uso de parâmetros de uma ou de outra categoria — ou, ainda, de ambas — vai depender de diversos aspectos, tal como o tempo hábil do pesquisador/profissional da saúde para se familiarizar com os procedimentos matemáticos envolvidos nos cálculos dos parâmetros, da necessidade de se obter um *feedback* em tempo real do paciente e, conseqüentemente, do tempo disponível para a extração dos parâmetros, do objetivo do estudo, etc.

Neste trabalho, os classificadores considerados renderam acurácias médias conjuntas entre 65 e 70% em todos os experimentos realizados, os quais envolveram o reconhecimento de perfis posturais de duas populações: indivíduos saudáveis e pós-AVE. Sendo assim, a utilização de tais modelos é mais apropriada do que a simples tentativa de acerto ao acaso, a qual detêm uma probabilidade de acerto de 50% para um problema envolvendo duas populações. Entretanto, novos estudos podem ser conduzidos na tentativa de elevar a acurácia dos modelos de classificação em problemas similares. Para isso, seria possível considerar outras métricas posturográficas — tanto globais como estruturais — além daquelas tradicionais utilizadas nesta pesquisa, bem como métricas de avaliações clínicas realizadas com os sujeitos de pesquisa, tal como a escala de equilíbrio de Berg. Além disso, poderia ser utilizada também uma combinação de múltiplos classificadores com esquema de votação.

Como sugestão adicional de trabalho futuro, seria possível avaliar o desempenho individual — ao invés do desempenho conjunto, como feito neste trabalho — de diferentes modelos de classificação para distinguir o equilíbrio postural em duas ou mais populações de interesse. Neste caso, para cada classificador considerado, seria mais apropriado realizar uma redução de dimensionalidade no conjunto de atributos posturográficos utilizando a estratégia *wrapper*, a qual permite identificar as variáveis com maior poder de discriminação para um dado classificador em particular.

Referências

- [ALH14] ALHIMALE, L.; ZEDAN, H.; AL-BAYATTI, A. *The implementation of an intelligent and video-based fall detection system using a neural network*. Applied Soft Computing, v. 18, p. 59-69, mar. 2014.
- [AMI14] AMINOFF, M; JOSEPHSON, S. *Aminoff's Neurology and General Medicine*. Ed Academic Press, 5ª ed. 2014.
- [AMO07] AMOUD, A; ABADI, M.; HEWSON, D.J.; MICHEL-PELLEGRINO, V.; DOUSSON, M.; DUCHÊNE, J. *Fractal time series analysis of postural stability in elderly and control subjects*. Journal of NeuroEngineering and Rehabilitation, v. 4, n. 12, mai. 2007.
- [ARA11] ARAÚJO, P.V. *Classificação automática de processos em sistemas operacionais*. Dissertação de mestrado, 74 p. Pós-Graduação em Informática, Pontifícia Universidade Católica do Paraná. Curitiba, 2011.
- [BAN06] BANDEIRA, H.T. *Modelo caótico e a memória da cinética dos canais iônicos*. Dissertação de Mestrado, 61 p. Programa de Pós-Graduação em Biometria, Universidade Federal Rural de Pernambuco. Recife, 2006.
- [BAR02] BARATTO, L.; MORASSO, P.G.; RE, C.; SPADA, G. *A new look at posturographic analysis in the clinical context: sway-density vs. other parameterization techniques*. Motor Control, v. 6, n. 3, p. 246-270, jul. 2002.
- [BLA08] BKASZCZYK, J.W. *Sway ratio - a new measure for quantifying postural stability*. Acta Neurobiologiae Experimentalis, vol. 68, n. 1, p. 51-57, 2008.
- [BOR06] BORGES, H.B. *Redução de dimensionalidade em bases de dados de expressão gênica*. Dissertação de mestrado, 123 p. Pós Graduação em Informática, Pontifícia Universidade Católica do Paraná. Curitiba, 2006.
- [BOR10] BORG, F.G.; LAXABACK, G. *Entropy of balance - some recent*

results. Journal of NeuroEngineering and Rehabilitation, v. 7, n. 38, jul. 2010.

- [BOU10] BOUGHORBEL, S.; BRUEKERS, F.; BREEBAART, J. *Baby-posture classification from pressure-sensor data*. International Conference on Patter Recognition, p. 556-559, ago. 2010.
- [BRY12] BRYCE, R.M.; SPRAGUE, K.B. *Revisiting detrended fluctuation analysis*. Scientific Reports, v. 2, n. 315, mar. 2012.
- [BUE06] BUENO, L.P.P. *Dinâmicas emergentes na família de memórias associativas bidirecionais caóticas e sua habilidade para saltar passos*. Tese de doutorado, 163 p. Escola de Engenharia de São Carlos, Universidade de São Paulo. São Carlos, 2006.
- [CAB11a] CABEZA-RUIZ, R.; GARCÍA-MASSÓ, X.; CENTENO-PRADA, R.A.; BEAS-JIMÉNEZ, J.D.; COLADO, J.C.; GONZÁLEZ, L.M. *Time and frequency analysis of the static balance in young adults with Down syndrome*. Gait and Posture, v. 33, n. 1, p. 23-28, jan. 2011.
- [CAB11b] CABRAL, R.V. *Programação genética baseada em árvores para classificação com uma classe com ênfase na geração de anomalias*. Dissertação de mestrado, 94 p. Programa de Pós-Graduação em Informática, Universidade Federal do Paraná (UFPR). Curitiba, 2011.
- [CAM05] CAMPANHARO, A.S.L.O.; MACAU, E.E.N.; RAMOS, F.M. *Detectando a presença de caos em uma série temporal*. Congresso Nacional de Matemática Aplicada e Computacional. São Paulo, 2005.
- [CAM07] CAMPANHARO, A.S.L.O. *Análise de sinais turbulentos na copa da floresta Amazônica: em busca de comportamento caótico e estruturas coerentes*. Dissertação de mestrado, 122 p. Pós-Graduação em Computação Aplicada, Instituto Nacional de Pesquisas Espaciais (INPE). São José dos Campos, 2007.
- [CAM09] CAMILO, C.O.; SILVA, J.C. *Mineração de dados: conceitos, tarefas,*

métodos e ferramentas. Relatório Técnico, 28 p. Instituto de Informática, Universidade Federal de Goiás. 2009.

- [CAR01] CARPENTER, M.G.; FRANK, J.S.; WINTER, D.A.; PEYSAR, G.W. *Sampling duration effects on center of pressure summary measures*. *Gait and Posture*, v. 13, n. 1, p. 35-40, fev. 2001.
- [CAS13] CASTRO, O.G. *Deslocamento do centro de pressão nas posições bipodal e unipodal em crianças*. Dissertação de mestrado, 52 p. Faculdade de Educação Física, Universidade de Brasília. Brasília, 2013.
- [CHA11] CHANG, T.S. *A comparative study of artificial neural networks, and decision trees for digital game content stock price prediction*. *Exper Systems with Applications*, v. 38, n. 12, p. 1486-1451, nov. 2011.
- [CHE05] CHEN, X.; SOLOMON, I.C.; CHON, K.H. *Comparison of the Use of Approximate Entropy and Sample Entropy: Applications to Neural Respiratory Signal*. *Conference Proceedings IEEE Engineering in Medicine and Biology Society*, v. 4, p. 4212-4215, 2005.
- [CHE06] WI-TING, C.; ZHI-ZHONG, W.; XIAO-MEI, R. *Characterization of surface EMG signals using improved approximate entropy*. *Journal of Zhejiang University Science B*, v. 7, n. 10, p. 844-848, 2006.
- [CHO11] CHOI, Y.; RALHAN, A.S.; KO, S. *A study on machine learning algorithms for fall detection and movement classification*. *International Conference on Information Science and Applications*, p. 1-8, abr. 2011.
- [COL93] COLLINS, J.J.; DE LUCA, C.J. *Open-loop and closed-loop control of posture: a random walk analysis of center-of-pressure trajectories*. *Experimental Brain Research*, v. 95, n. 2, p. 308-318, 1993.
- [COR01] CORRIVEAU, H.; HÉRBERT, R.; PRICE, F.; RAÎCHE, M. *Postural Control in the Elderly: An Analysis of Test-Retest and Interrater Reliability of the COP-COM Variable*. *Archives of Physical Medicine and Rehabilitation*, v. 82, n. 1, p. 80-85, jan. 2001.

- [COS02] COSTA, M.; GOLDBERGER, A.L.; PENG, C.K. *Multiscale Entropy Analysis of Complex Physiologic Time Series*. Physical Review Letters, v. 89, n. 6, aug. 2002.
- [COS05] COSTA, M.; GOLDBERGER, A.L.; PENG, C.K. *Multiscale Entropy Analysis of Biological Signals*. Physical Review E., v. 71, fev. 2005.
- [COS07] COSTA, M.; PRIPLATA, A.A.; LIPSITZ, L.A.; WU, Z.; HUANG, N.E.; GOLDBERG, A.L.; PENG, C.K. *Noise and poise: Enhancement of postural complexity in the elderly with a stochastic-resonance-based therapy*. Europhysics Letters, v.77, mar. 2007.
- [COS13] COSTA, A.B. *Aprendizado de máquina em representações tridimensionais*. Dissertação de mestrado, 92 p. Faculdade de Computação, Universidade Federal do Mato Grosso do Sul. Campo Grande, 2013.
- [DAV03] DAVIDS, K.; GLAZIER, P.; ARAÚJO, D.; BARLETT, R. *Movement systems as dynamical systems - the functional role of variability and its implications of sports medicine*. Sports Medicine, v. 33, n. 4, p. 245-260, 2003.
- [DEF09] DEFFEYES, J.E.; HARBOURNE, R.T.; KYVELIDOU, A.; STUBERG, W.A.; STERGIOUS, N. *Nonlinear analysis of sitting postural sway indicates developmental delay in infants*. Clinical Biomechanics, v. 24, n. 7, p. 564-570, ago. 2009.
- [DEL03] DELIGNIÉRES, D.; DESCHAMPS, T.; LEGROS, A.; CAILLOU, N. *A methodological note on non-linear time series analysis: Is Collins and De Luca (1993)'s open- and closed-loop model a statistical artifact?* Journal of Motor Behaviour, v. 35, n. 1, p. 86-97, mar. 2003.
- [DEL06] DELIGNIERES, D.; RAMDANI, S.; LEMOINE, L.; TORRE, K.; FORTES, M.; NINOT, G. *Fractal analyses for 'short' time series: A re-assessment of classical methods*. Journal of Mathematical Psychology, v.

50, n. 6, p. 525-544, dez. 2006.

- [DEL09] DELIGNIERES, D.; TORRE, K. *Fractal dynamics of human gait: a reassessment of the 1996 data of Hausdorff et al.* Journal of Applied Physiology, v. 106, n. 4, p. 1272-1279, abr. 2009.
- [DEM08] DEMURA, S.; KITABAYASHI, T.; AOKI, H. *Body-sway characteristics during a static upright posture in the elderly.* Geriatrics and Gerontology International, v. 8, n. 3, . 188-197, set. 2008.
- [DIN00] DINGWELL, J.B.; CUSUMANO, J.P. *Nonlinear time series analysis of normal and pathological human walking.* Chaos, v. 10, n. 4, p. 848-863, dez. 2000.
- [DIN07] DINGWELL, J.B.; KANG, H.G. *Differences Between Local and Orbital Dynamic Stability During Human Walking.* Journal of Biomechanical Engineering, v. 129, n. 4, p. 586-593, ago. 2007.
- [DIN14] DINAKARAM, S.; THANGAIAH, P.R.J. *Comparative Analysis of Filter-Wrapper Approach for Random Forest Performance on Multivariate Data.* Int. Conf. Intell. Comput. Appl., p. 174-178, 2014.
- [DON07] DONKER, S.F.; ROERDINK, M.; GREVEN, A.J.; BEEK, P.J. *Regularity of center-of-pressure trajectories depends on the amount of attention invested in postural control.* Experimental Brain Research, v. 181, n. 1, p. 1-11, jul. 2007.
- [DON08] DONKER, S.F.; LEDEBT, A.; ROERDINK, M.; SAVELSBERGH, G.J.P.; BEEK, P.J. *Children with cerebral palsy exhibit greater and more regular postural sway than typically developing children.* Experimental Brain Research, v. 184, n. 3, p. 363-370, jan. 2008.
- [DOY05] DOYLE, T.L.; NEWTON, R.U.; BURNETT, A.F. *Reliability of traditional and fractal dimension measures of quiet stance center of pressure in young, healthy people.* Archives of Physical Medicine and Rehabilitation, v. 86, n. 10, out. 2005.

- [DUA00a] DUARTE, M. *Análise estabilográfica da postura ereta humana quase-estática*. Tese de doutorado, 86 p. Escola de Educação Física e Esporte, Universidade de São Paulo. São Paulo, 2000.
- [DUA00b] DUARTE, M.; ZATSIORSKY, V.M. *On the fractal properties of natural human standing*. *Neuroscience Letters*, v. 283, n. 3, p. 173-176, abr. 2000.
- [DUA01] DUARTE, M.; ZATSIORSKY, V.M. *Long-range correlations in human standing*. *Physics Letters A*, v. 283, n. 1-2, p. 124-128, mai. 2001.
- [DUA08] DUARTE, M.; STERNAD, D. *Complexity of human postural control in young and older adults during prolonged standing*. *Experimental Brain Research*, v. 191, n. 3, p. 265-276, nov. 2008.
- [DUA10] DUARTE, M.; FREITAS, S.M.S.F. *Revisão sobre posturografia baseada em plataforma de força para avaliação do equilíbrio*. *Revista Brasileira de Fisioterapia*, v. 14, n. 3, p. 183-192, mai. 2010.
- [DUF14] DUFFELL, L.D.; SOUTHGATE, D.F.L.; GULATI, V.; MCGREGOR, A.H. *Balance and Gait Adaptations in Patients with Early Knee Osteoarthritis*. *Gait and Posture*, n. 14, jan. 2014.
- [ECK87] ECKMANN, J.P.; KAMPHORST, S.O.; RUELLE, D. *Recurrence plots of dynamical systems*. *Europhysics Letters*, v. 4, n. 973, 1987.
- [EMM00] EMMERIK, R.E.; WEGEN, E.E. *On variability and stability in human movement*. *Journal of Applied Biomechanics*, v. 16, n. 4, p. 394-406, 2000.
- [ESP10] ESPEJO, P.G.; VENTURA, S.; HERRERA, F. *A survey on the application of genetic programming to classification*. *Trans. Sys. Man Cyber Part C*, v. 40, n. 2, p. 121-144, 2010.
- [FIG00] FIGUEIREDO, J.C.B. *Equações diferenciais não lineares com três*

retardos: estudo detalhado das soluções. Tese de doutorado, 146 p. Instituto de Física, Universidade de São Paulo. São Paulo, 2000.

- [FIN74] DE FINNETI, B. *Theory of Probability*. Volumes 1 and 2, John Wiley & Sons, New York, 1974.
- [FOR95] FORSTER, A.; YOUNG, J. *Incidence and consequences of falls due to stroke: a systematic inquiry*. British Medical Journal, v. 331, p. 83-86, jul. 1995.
- [FRA86] FRASER, A.M.; SWINNEY, H.L. *Independent coordinates for strange attractors from mutual information*. Physical Review A, v. 33, n. 2, p. 1134-1140, fev. 1986.
- [FRE06] FREITAS, S.M.S.F.; DUARTE, M. *Métodos de análise do controle postural*. Universidade de São Paulo, 2006.
- [FUC08] FUCKNER, M. *Combinação de classificadores usando IAD*. Dissertação de mestrado, 79 p. Pós-Graduação em Informática, Pontifícia Universidade Católica do Paraná. Curitiba, 2008.
- [GAN01] GANDUR, M.C. *Comportamento dinâmico complexo em despelamento de fitas adesivas*. Tese de doutorado, 101 p. Instituto de Química, Universidade de Campinas (UNICAMP). Campinas, 2001.
- [GAN13] GANESAN, M.; LEE, Y.; ARUIN, A.S. *The effect of lateral or medial wedges on control of postural sway in standing*. Gait and Posture, n. 13, dez. 2013.
- [GAR11] GARCIA, R.B.B. *Seleção de atributos usando critérios multiobjetivo baseado em enxame para seleção de genes em microarranjos (SAMO-ESMA)*. Dissertação de mestrado, 119 p. Programa de Pós-Graduação em Informática, Pontifícia Universidade Católica do Paraná. Curitiba, 2011.
- [GEN08] GENTHON, N.; ROUGIER, P.; GISSOT, A.S.; FROGER, J.; PÉLISSIER, J. PÉRENNOU, D. *Contribution of each lower limb to*

uprighth standing in stroke patients. Journal of the American Heart Association, v. 39, n. 6, p. 1793-1799, jun. 2008.

- [GHA12] GHADERI, H.; KABIRI, P. *Fourier transform and correlation-based feature selection for fault detection of automobile engines*. AISP 16th CSI Int. Symp. Artif. Intell. Signal Process., p. 514–519, 2012.
- [GHO10] GHOMASHCHI, H.; ESTEKI, A.; SPROTT, J.C.; NASRABADI, A.M. *Identification of dynamic patterns of body sway during quiet standing: is it a nonlinear process?* International Journal of Bifurcation and Chaos, v. 20, n. 4, p. 1269-1278, 2010.
- [GOM12] GOMES, R.S.R. *Arquitetura de processamento paralelo com unidade de processamento gráfico de propósito geral para aplicação em séries temporais de dados ambientais*. Dissertação de mestrado, 56 p. Programa de Pós-Graduação em Física Ambiental, Universidade Federal do Mato Grosso. Cuiabá, 2012.
- [GOR11] GORUNESCU, F. *Data mining: concepts, models and techniques*. Intelligent Systems Reference Library, v. 12. Ed Springer, 2011.
- [GRA07] GRANATA, K.P. *Reply to the letter to the editor*. Gait and Posture, v. 26, p. 239-330, 2007.
- [GRA83] GRASSBERGER, P.; PROCACCIA, I. *Characterization of Strange Attractors*. Physical Review Letters, v. 50, n. 5, p. 346-349, 1983.
- [GUR13] GURSES, S.; CELIK, H. *Correlation dimension estimates of human postural sway*. Human Movement Science, v. 32, n. 1, p. 48-64, 2013.
- [HAA04] HAART, M.; GEURTS, A.C.; HUIDEKOPER, S.C.; FASOTTI, L.; LIMBEEK, J.V. *Recovery of Standing Balance in Postacute Stroke Patients: A Rehabilitation Cohort Study*. Archives of Physical Medicine and Rehabilitation, v. 85, n. 6, p. 886-895, jun. 2004.
- [HAL99] HALL, M.A. *Correlation-based feature selection for machine learning*. The University of Waikato, 1999.

- [HAL00] HALL, M.A. Correlation-based feature selection for discrete and numeric class machine learning. Hamilton, p. 1-10, 2000.
- [HAL09] HALL, M. et al. *The WEKA data mining software: an update*. SIGKDD Explorations, vol. 11, p. 10-18, jun. 2009.
- [HAN05] HAN, J.; MOUSSAVI, Z.; SZTURM, T.; GOODMAN, V. *Application of nonlinear dynamics to human postural control system*. Proceedings of the IEEE Engineering in Medicine and Biology Society, v. 7, p. 6885-6888, 2005.
- [HAN11] HAN, J.; KAMBER, M.; PEI, J. *Data mining: concepts and techniques*. Ed Morgan Kaufmann, 3^a ed. 2011.
- [HAR09] HARBOURNE, R.T.; STERGIOU, N. *Movement variability and the use of nonlinear tools: principles to guide physical therapist practice*. Physical Therapy, v. 89, n. 3, p. 267-282, mar. 2009.
- [HOR06] HORAK, F.B. *Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls?* Oxford University Press on behalf of the British Geriatrics Society, v. 35, n. 2, p. 7-11, 2006.
- [HOR87] HORAK, F.B. *Clinical measurement of postural control in adults*. Physical Therapy, v. 67, n. 12, p. 1881-1885, dez. 1987
- [HSU08] HSU, H.-H.; HSIEH, C.-W.; LU, M.-D. *A hybrid feature selection mechanism*. 8a. Int. Conf. Intell. Syst. Des. Appl., v. 2, 2008.
- [JAI97] JAIN, A.; ZONGKER, D. *Feature selection: evaluation, application, and small sample performance*. IEEE Trans. Pattern Anal. March. Intell., v. 19, n. 2, p. 153-158, 1997.
- [KAN04] KANTS, H.; SCHREIBER, T. *Nonlinear time series analysis*. Editora Cambridge, 2^a edição, 388 p., jan. 2004.
- [KER12] KERDEGARI, H.; SAMSUDIN, K.; RAMLI, A.R.; MOKARAM, S.

- Evaluation of fall detection classification approaches*. International Conference on Intelligent and Advanced Systems, p. 131-136, jun. 2012.
- [KIM11] KIM, J.J.; KIM, Y.J.; LEE, J.J. *A machine learning approach to falling detection and avoidance for biped robots*. Proceedings of SICE Annual Conference, p. 562-567, set. 2011.
- [KIR12] KIRCHNER, M.; SCHUBERT, P.; SCHMIDTBLEICHER, D.; HAAS, C.T. *Evaluation of the temporal structure of postural sway fluctuations based on a comprehensive set of analysis tools*. Pshysica A, v. 391, n. 20, p. 4692-4703, out. 2012.
- [KIR13] KIRCHNER, M. *Characterising postural sway fluctuations in humans using linear and nonlinear methods*. Tese de doutorado, 173 p. Goethe Universitat in Frankfurt am Main, mai. 2013.
- [KOZ89] KOZA, J.R. *Hierarchical genetic algorithms operating on populations of computer programs*. Proceedings of the 11th International Joint Conference on Artificial Intelligent, v. 1, p. 768-774, 1989.
- [KOZ92] KOZA, J.R. *Genetic Programming: On the Programming of Computers by Mean of Natural Selection*. MIT Press, Massachusetts, 1989.
- [KUR87] KURTHS, J.; HERZEL, H. *An attractor in solar time series*. Physica D, v. 25, p. 165-172, 1987.
- [LAC08] LACOUR, M.; BERNARD-DEMANZE, L.; DUMITRESCU, M. *Posture control, aging, and attention resources: models and posture-analysis methods*. Neurophysiologie Clinique, v. 38, n. 6, p. 411-421, dez. 2008.
- [LAD07] LADISLAO, L.; FIORETTI, S. *Nonlinear analysis of posturographic data*. Medical and Biological Engineering and Computing, v. 45, n. 7, p. 697-688, jul. 2007.
- [LAF04] LAFOND, D.; CORRIVEAU, H.; HÉRBERT, R.; PRINCE, F.

- Intrasession reliability of center of pressure measures of postural steadiness in healthy elderly people.* Archives of Physical Medicine and Rehabilitation, v. 85, n. 6, p. 896-901, jun. 2004.
- [LAT02] LATASH, M.L.; SCHOLZ, J.P.; SCHONER, G. *Motor control strategies revealed in the structure of motor variability.* Exercise and Sport Sciences Reviews, v. 30, n. 1, p. 26-31, jan. 2002.
- [LEB00] LEBIEDOWSKA, M. K.; SYCZEWSKA, M. *Invariant Sway properties in children.* Gait and Posture, v. 12, n. 3, p. 200-204, dez. 2000.
- [LEV08] LEVASSEUR, Y. *Techniques de l'intelligence artificielle pour la reconnaissance d'objets biologiques dans une image.* Dissertação de mestrado. Departamento de Tecnologia, Universidade de Quebec. Montreal, 2008.
- [LIU10] LIU, C.L.; LEE, C.H.; LIN, P.M. *A fall detection system using k-nearest neighbor classifier.* Expert Systems with Applications, v. 37, n. 10, p. 7174-7181, out. 2010.
- [LOC13] LOCKHART, T.; STERGIOU, N. *New perspectives in human movement variability.* Annals of Biomedical Engineering, v. 41, n. 8, p. 1593-1594, ago. 2013.
- [LOR63] LORENZ, E.N. *Deterministic nonperiodic flow.* Journal of Atmospheric Sciences, v. 20, n. 2, p.130-148, mar. 1963.
- [LUC11] LUCENA, R.N.O. *Propriedades ergódicas do modelo geométrico do atrator de Lorenz.* Dissertação de mestrado, 82 p. Pós-Graduação em Matemática, Universidade Federal de Alagoas. Maceió, 2011.
- [MAD11] MADELEINE, P.; NIELSEN, M.; ARENDT-NIELSEN, L. *Characterization of postural control deficit in whiplash patients by means of linear and nonlinear analyses - a pilot study.* Journal of Electromyography and Kinesiology, v. 21, n. 2, p. 291-297, abr. 2011.

- [MAI05] MAINENTI, M.R.M. *Análise de sinais estabilométricos em testes com estímulos sonoros*. Dissertação de mestrado, 75 p. Pós-Graduação em Engenharia Biomédica, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2005.
- [MAI08] MAIA, R.D.; MENDES, E.M.A.M. *Identificação de Estrutura e Parâmetros de Modelos para Sistemas Dinâmicos Não-Lineares Utilizando Programação Genética*. XVII Congresso Brasileiro de Automática, Juiz de Fora, 2008.
- [MAR02a] MARWAN, N.; WESSEL, N.; MEYERFELDT, U.; SCHIRDEWAN, A.; KURTHS, J. *Recurrence-plot-based measures of complexity and their application to heart-rate-variability data*. Physical Review E, v. 66, N. 2, ago. 2002.
- [MAR02b] MARWAN, N.; THIEL, M.; NOWACZYK, N.R. *Cross Recurrence Plot Based Synchronization of Time Series*. Nonlinear Processes in Geophysics, v. 9, p. 325-331, 2002.
- [MAR02c] MARWAN, N.; KURTHS, J. *Nonlinear analysis of bivariate data with cross recurrence plots*, Phys. Lett. A, v. 302, n. 5-6, p. 299-307, 2002.
- [MAR03] MARWAN, N. *Encounters with neighbours*. Tese de doutorado, 159 p. Instituto de Física, Universitat Potsdam. 2003.
- [MAR06] MARGARIDO, M.A.; MEDEIROS JUNIOR, H. *Teste para mais de uma raiz unitária: uso do software SAS na elaboração de uma rotina para o teste dickey-pantula*. Revista do Programa de Estudos Pós-Graduação em Economia, v. 17, n. 1, 2006.
- [MAR07] MARWAN, N.; ROMANO, M.C.; THIEL, M.; KURTHS, J. *Recurrence Plots for the Analysis of Complex Systems*. Physics Reports, v. 438, n. 5-6, p. 237-329, jan. 2007.
- [MAR12] MARREGA, L.H.G.; SILVA, S.M.; MANFFRA, E.F. *Influence of Time Series Length on the Estimates of Correlation Dimension from Stroke*

- Patients Posturographic Data*. International Conference of the IEEE Engineering in Medicine and Biology Society, p. 545-545, 2012.
- [MER10] MERGNER, T. *A neurological view on reactive human stance control*. Annual Reviews in Control, v. 34, n. 2, p. 177-198, 2010.
- [MOR06] MORETTIN, P.A.; TOLOI, C.M.C. *Análise de séries temporais*. Ed. Edgard Blucher LTDA, São Paulo, 2006.
- [MOR12] MORAIS, D.C.S.; MORAIS, B.C.S.; MENEZES JÚNIOR, J.V.; GUSMÃO, C.M.G. *Sistema móvel de apoio a decisão médica aplicado ao diagnóstico de asma - intelimed*. Simpósio Brasileiro de Sistemas de Informação, v. 01, p. 528-539, 2012.
- [MOU05] MOULIN, G.D. *Técnicas de entropia e de geoestatística no dimensionamento de redes pluviométricas*. Dissertação de mestrado, 90 p. Pós-graduação em Engenharia Ambiental, Universidade Federal do Espírito Santo. Vitória, 2005.
- [MOU08] MOUSTAKIDIS, S.P.; THEOCHARIS, J.B.; GIAKAS, G. *Subject recognition based on ground reaction force measurements of gait signals*. IEEE Transactions on Systems, Man, and Cybernetics, v. 38, n. 6, p. 1476-1485, dez. 2008.
- [NAS05] NASCIMENTO, A.W.M. *Análise de caos em oscilações de baixa frequência em dispositivos não-lineares*. Dissertação de mestrado, 77 p. Pós-Graduação em Engenharia de Materiais, Universidade Federal de Itajubá. Itajubá, 2005.
- [NAS85] NASHNER, L.M.; MCCOLLUM, G. *The organization of human postural movements: a formal basis and experimental synthesis*. Behav. Br. Sc., v. 8, p. 135-172, 1985.
- [NEG13] NEGAHBAN, H.; SANJARI, M.A.; MOFATEH, R.; PARNIANPOUR, M. *Nonlinear dynamical structure of sway path during standing in patients with multiple sclerosis and in healthy controls is affected by*

changes in sensory input and cognitive load. Neuroscience Letters, v. 553, p. 126-131, out. 2013.

- [NET06] PELLI NETO, A. *Redes neurais artificiais aplicadas às avaliações em massa - estudo de caso para a cidade de Belo Horizonte.* Dissertação de mestrado, 96 p. Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais. Belo Horizonte, 2006.
- [NEW93] NEWELL, K.M.; EMMERIK, R.E.A.; LEE, D.; SPRAGUE, R.L. *On postural stability and variability.* Gait and Posture, v. 4, p. 225-230, 1993.
- [NEW97] NEWELL, K.M.; SLOBOUNOV, S.M.; SLOBOUNOVA, B.S.; MOLENAAR, P.C.M. *Short-term non-stationarity and the development of postural control.* Gait and Posture, v. 6, p. 56-62, 1997.
- [OLI05] OLIVEIRA, A.R. *Redes neurais artificiais aplicadas na detecção, classificação e localização de defeitos em linhas de transmissão.* Dissertação de mestrado, 120 p. Pós-Graduação em Engenharia Elétrica, Universidade Federal de Juiz de Fora. Juiz de Fora, 2005.
- [OLI09] OLIVEIRA, L.F.; VIEIRA, T.M.M.; MACEDO, A.R.; SIMPSON, D.M.; NADAL, J. *Postural sway changes during pregnancy: a descriptive study using stabilometry.* European Journal of Obstetrics and Gynecology and Reproductive Biology, v. 147, n. 1, p. 25-28, nov. 2009.
- [OLI94] OLIVEIRA, L.F.; SIMPSON, D.M.; NADAL, J. *Calculation of area of stabilometric signals using principal component analysis.* Physiological Measurement, v. 17, p. 305-312, 1994.
- [PAR98] PARLITZ, U. *Nonlinear time-series analysis.* Capítulo do livro *Nonlinear modeling - advanced black-box techniques*, p. 209-239, 1998.
- [PAS05] PASCOLO, P.B; MARINI, A.; CARNIEL, R.; BARAZZA, F. *Posture as a chaotic system and an application to the Parkinson's disease.*

Chaos, Solitons and Fractals, v. 24, n. 5, p. 1343-1346, 2005.

- [PEN09] PENG, C.K.; COSTA, M.; GOLDBERGER, A.L. *Adaptive data analysis of complex fluctuations in physiologic time series*. Adv. Adapt. Data Anal., v. 1, n. 1, p. 61-70, jan. 2009.
- [PEN95] PENG, C.K.; HAVLIN, S.; STANLEY, H.E.; GOLDBERGER, A.L. *Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series*. Chaos, v. 5, n. 1, p. 82-87, 1995.
- [PHY11] PHYSIONET, 2011. Disponível em <<http://physionet.org/physiotools/sampen/matlab/>>.
- [PIN91] PINCUS, S.M. *Approximate entropy as a measure of system complexity*. Proc. Natl. Acad. Sci., v. 88, n. 6, p. 2297-2301, mar. 1991.
- [POR14] PORTAL BRASIL. *Acidente vascular cerebral*. Disponível em: <<http://www.brasil.gov.br/saude/2012/04/acidente-vascular-cerebral-avc>>, último acesso em 04 fev. 2015.
- [PYO04] PYORIA, O.; ERA, P.; TALVITIE, U. *Relationships between standing balance and symmetry measurements in patients following recent strokes or older strokes*. Physical Therapy, v. 84, n. 2, p. 128-136, fev. 2004.
- [RAS09] RASKU, J. *A method for the classification of corrective activity in context dependent postural controlling tasks*. Computers in Biology and Medicine, v. 39, n. 10, p. 940-945, out. 2009.
- [RHE11] RHEA, C.K.; SILVER, T.A.; HONG, S.L.; RYU, J.H.; STUDENKA, B.E.; HUGHES, C.M.L.; HADDAD, J.M. *Noise and complexity in human postural control: interpreting the different estimations of entropy*. Plos one, v. 6, n. 3, mar. 2011.
- [RIB10] RIBEIRO, R.A. *Correlações nos DFA de diversos perfis geológicos, estudo de caso: Bacia de Campos - RJ*. Dissertação de mestrado, 79 p. Pós-Graduação em Ciências e Engenharia do Petróleo, Universidade

Federal do Rio Grande do Norte. Natal, 2010.

- [RIC00] RICHMAN, J.S.; MOORMAN, J.R. *Physiological time-series analysis using approximate entropy and sample entropy*. American Journal of Physiology, Heart and Circulatory Physiology, v. 278, n. 6, p. 2039-2049, jun. 2000.
- [ROD02] RODRIGUES, E.L.M. *Evolução de funções em programação genética orientada a gramáticas*. Dissertação de mestrado, 104 p. Programa de Pós-Graduação em Informática, Universidade Federal do Paraná (UFPR). Curitiba, 2002.
- [ROD07] RODRIGUES, S. *Influência da carga e forma de transporte do material escolar sobre a distribuição da força plantar e trajetória do centro de massa corporal*. Dissertação de mestrado, 80 p. Pós-Graduação em Fisioterapia, Universidade Metodista de Piracicaba. Piracicaba, 2007.
- [ROE06] ROERDINK, M.; DE HAART, M.; DAFFERTSHOFER, A.; DONKER, S.F.; GEURTS, A.C.H.; BEEK, P.J. *Dynamical structure of center-of-pressure trajectories in patients recovering from stroke*. Experimental Brain Research, v. 174, n. 2, p. 256-269, set. 2006.
- [ROS93] ROSENTEIN, M.T.; COLLINS, J.J.; DE LUCA, C.J. *A practical method for calculating largest Lyapunov exponents from small data sets*. Physica D, v. 65, p. 117-134, 1993.
- [RUE79] RUELLE, D. *Ergodic theory of differentiable dynamical systems*. IHES Publications Mathématiques, v. 50, p. 275-320, 1979.
- [RUG11] RUGELJ, D.; SEVSEK, F. *The effect of load mass and its placement on postural sway*. Applied Ergonomics, v. 42, n. 6, p. 860-866, nov. 2011.
- [SAM13] SAMAEI, F.; DANESHFAR, M.; BEYDOKHTI, S.S. *Using the nonlinear analysis of COP variations during the quiet standing*. International Journal of Engineering and Advanced Technology, v. 3, n. 2, p. 57-60, dez. 2013.

- [SAN02] SANTOS, E.M. *Teoria e aplicação de support vector machines à aprendizagem e reconhecimento de objetos baseado na aparência*. Dissertação de mestrado, 111 p. Pós-Graduação em Informática, Universidade Federal da Paraíba. Campina Grande, 2002.
- [SAR14] SARIPALLE, S.K.; PAIVA, G.C.; CLIETT, T.C.; DERAKHSHANI, R.R.; KING, G.W.; LOVELACE, C.T. *Classification of body movements based on posturographic data*. Human Movement Science, v. 33, p. 238-250, fev. 2014.
- [SAT87] SATO, S.; SANO, M.; SAWADA, Y. *Practical method of measuring the generalized dimension and largest Lyapunov exponent in high dimensional chaotic system*. Progress of Theoretical Physics, v. 77, n. 1, p. 1-5, ago. 1987.
- [SCH89] SCHEINKMAN, J.A.; LEBARON, B. *Nonlinear Dynamics and Stock Returns*. The Journal of Business, v. 62, n. 3, jul. 1989.
- [SCH05] SCHIFFMAN, J.M.; BENSEL, C.K.; HASSELQUIST, L.; GREGORCZYK, K.N.; PISCITELLE, L. *Effects of carried weight on random motion and traditional measures of postural sway*. Applied Ergonomics, v. 37, n. 5, p. 607-614, set. 2006.
- [SEI09] SEIGLE, B.; RAMDANI, S.; BERNARD, P.L. *Dynamical structure of center of pressure fluctuations in elderly people*. Gait and Posture, v. 30, n. 2, p. 223-226, ago. 2009.
- [SEM02] SEMOLINI, R. *Support vector machines, inferência transdutiva e o problema de classificação*. Dissertação de mestrado, 128 p. Faculdade de Engenharia Elétrica e Computação, Universidade Estadual de Campinas (UNICAMP). Campinas, 2002.
- [SHU03] SHUMWAY-COOK, A; WOOLLACOTT, M.H. *Controle motor: teoria e aplicações*. Ed Manoele. São Paulo, 2003.
- [SIL10] SILVA, L.E.V. *Ferramentas computacionais na análise da*

variabilidade da frequência cardíaca através do paradigma não extensivo no estudo de cardiopatias. Dissertação de mestrado, 95 p. Pós-Graduação em Física Aplicada à Medicina e Biologia, Universidade de São Paulo. Ribeirão Preto, 2010.

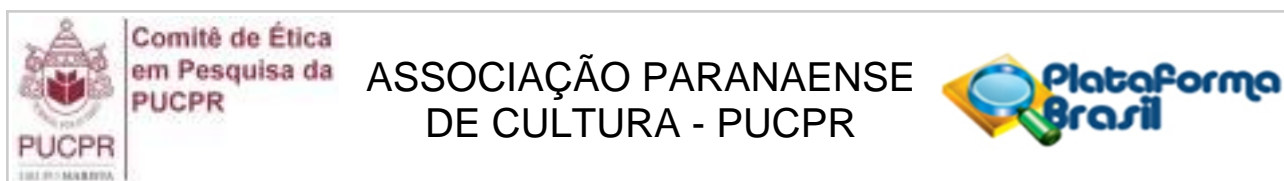
- [SIL12] SILVA, S.M. *Análise do controle postural de indivíduos pós-acidente vascular encefálico frente a perturbações dos sistemas visual e somatossensorial*. Dissertação de Mestrado, 114 p. Pós-Graduação em Tecnologia em Saúde, Pontifícia Universidade Católica do Paraná. Curitiba, 2012.
- [SOU06] SOUZA, L.V. *Programação genética e combinação de preditores para previsão de séries temporais*. Tese de doutorado, 138 p. Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná. Curitiba, 2006.
- [STE11] STERGIOU, N.; DECKER, L.M. *Human movement variability, nonlinear dynamics, and pathology: is there a connection?* Human Movement Science, v. 30, n. 5, p. 869-888, out. 2011.
- [TAK81] TAKENS, F. *Detecting strange attractors in turbulence*. Dynamical Systems and Turbulence, v. 898, p. 366-381, 1981.
- [TAL13] TALLON, G.; BLAIN, H.; SEIGLE, B.; BERNARD, P.L.; RAMDANI, S. *Dynamical and stabilometric measures are complementary for the characterization of postural fluctuations in older women*. Gait and Posture, v. 38, n. 1, p. 92-96, mai. 2013.
- [THE92] THEILER, J. et al. *Testing for nonlinearity in time series: the method of surrogate data*. Physica D, v. 58, p. 77-97, mar. 1992.
- [THI00] THIELO, M.R. *Análise e classificação de séries temporais não estacionárias utilizando métodos não-lineares*. Dissertação de mestrado, 103 p. Pós-Graduação em Computação, Universidade Federal do Rio Grande do Sul. Porto Alegre, 2000.

- [TST09] TSTOOL Home Page, 2009. Disponível em <<http://www.physik3.gwdg.de/tstool/>>.
- [VAP95] VAPNIK, V.N. *The nature of statistical learning theory*. Springer-Verlag, New York, 1995.
- [VAS07] VASCONCELLOS, R.M.G. *Reconstrução de espaços de estados aeroelásticos por decomposição em valores singulares*. Dissertação de mestrado, 127 p. Escola de Engenharia de São Carlos, Universidade de São Paulo. São Carlos, 2007.
- [VET10] VETTE, A.H.; MASANI, K.; SIN, V.; POPOVIC, M.R. *Posturographic measures in healthy young adults during quiet sitting in comparison with quiet standing*. Medical Engineering and Physics, v. 32, n. 1, p. 32-38, jan. 2010.
- [VIA10] VIANA JÚNIOR, E.R. *Espaços de parâmetros para um circuito caótico experimental*. Dissertação de mestrado, 102 p. Departamento de Física, Universidade Federal de Minas Gerais. 2010.
- [VIE12] VIEIRA, V.J.D.; COSTA, S.C.; COSTA, W.C.A. *Análise de quantificação de recorrência e análise discriminante aplicadas à classificação de sinais de vozes saudáveis e sinais de vozes patológicas*. Congresso Norte Nordeste de Pesquisa e Inovação, 2012.
- [WEB94] WEBBER, C.L.; ZBILUT, J.P. *Dynamical assessment of physiological systems and states using recurrence plot strategies*. Journal of Applied Physiology, v. 76. n. 2, p. 965-973, fev. 1994.
- [WIN90] WINTER, D.A.; PATLA, A.E.; FRANK, J.S. *Assessment of balance control in humans*. Medical Progress Through Technology, v. 16, p. 31-51, 1990.
- [WOL00] WOLFE, C.D. *The Impact of Stroke*. British Medical Bulletin, v. 56, n. 2, p. 275-286, 2000.

- [WOL85] WOLF, A.; SWIFTH, J.B.; SWINNEY, H.L.; VASTANO, J.A. *Determining Lyapunov exponents from a time series*. Physica D, v. 16, p. 285-317, 1985.
- [YAM95] YAMADA, N. *Chaotic swaying of the upright posture*. Human Movement Science, v. 14, n. 6, p. 711-726, dez. 1995.
- [ZBI92] ZBILUT, J.P.; WEBBER, C.L. *Embedding and delays as derived from quantification of recurrence plots*. Physics Letters A, v. 171, n. 3-4, p. 199-203, dez. 1992
- [ZOK08] ZOK, M.; MAZZÀ, C.; CAPOZZO, A. *Should the instructions issued to the subject in traditional static posturography be standardised?* Medical Engineering Physics, v. 30, n. 7, p. 913-916, set. 2008.

ANEXOS

ANEXO A - TERMO DE APROVAÇÃO DO COMITÊ DE ÉTICA EM PESQUISA



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: ANÁLISE E MODELAGEM DO CONTROLE DO EQUILÍBRIO EM INDIVÍDUOS PÓS-AVC

Pesquisador: Elisângela Ferretti Manffra

Área Temática:

Versão: 1

CAAE: 42106815.7.0000.0020

Instituição Proponente: Pontifícia Universidade Católica do Parana - PUCPR

Patrocinador Principal: MINISTERIO DA CIENCIA, TECNOLOGIA E INOVACAO

DADOS DO PARECER

Número do Parecer: 991.103

Data da Relatoria: 18/03/2015

Apresentação do Projeto:

Estudo exploratório baseado em estatísticas e técnicas de Mineração de Dados sobre uma base de dados secundária referente a indivíduos hígidos e pós-AVC. Os dados foram coletados entre julho de 2011 e setembro de 2011 a partir de 38 voluntários (19 adultos pós-AVC e 19 adultos hígidos pareados) no Centro Hospitalar de Reabilitação Ana Carolina Moura Xavier. O procedimento de coleta dos dados primários foi aprovado no CEP sob o parecer nº 319/2011. As etapas deste projeto consistem (i) no pré-processamento dos dados, mais especificamente, no cálculo de determinadas variáveis a partir das séries temporais do movimento do centro de pressão; (ii) na análise descritiva e comparativa das variáveis calculadas; (iii) na análise dos resultados por meio de técnicas de Mineração de Dados, e; (iv) na proposta de modelos teóricos do controle motor e sua comparação os com dados experimentais.

Objetivo da Pesquisa:

Objetivo Primário:

O objetivo primário desta pesquisa é avaliar a relevância do uso conjunto entre parâmetros posturográficos tradicionais das categorias global e estrutural em comparação com o uso de parâmetros de apenas uma categoria em particular para, com o auxílio de técnicas de Mineração

Endereço: Rua Imaculada Conceição 1155

Bairro: Prado Velho

CEP: 80.215-901

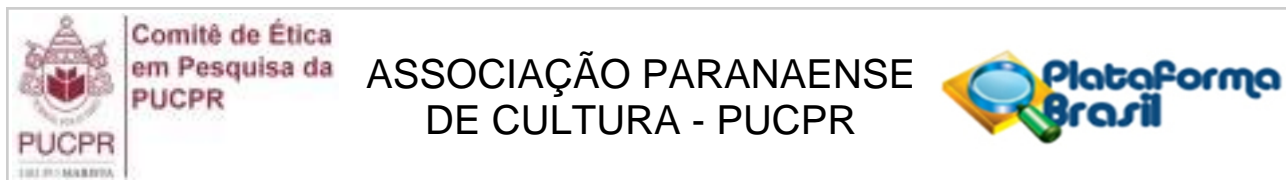
UF: PR

Município: CURITIBA

Telefone: (41)3271-2103

Fax: (41)3271-2103

E-mail: nep@pucpr.br



Continuação do Parecer: 991.103

de Dados, diferenciar os perfis de oscilação corporal produzidos por pacientes pós-AVC e indivíduos saudáveis pareados.

Objetivo Secundário:

O objetivo secundário desta pesquisa é avaliar o poder individual de descrição de diferentes parâmetros posturográficos tradicionais das categorias global e estrutural para, através de análises estatísticas, diferenciar os perfis de oscilação corporal produzidos (i) por diferentes populações na mesma condição de controle do equilíbrio e (ii) pela mesma população em diferentes condições. As populações avaliadas serão pacientes pós-AVC e indivíduos saudáveis pareados.

Avaliação dos Riscos e Benefícios:

Riscos:

Uma vez que este projeto envolve a utilização de uma base de dados secundária sem a identificação dos indivíduos, não existem riscos para os voluntários que participaram da coleta da base de dados primária. Também não existe risco de que os resultados sejam utilizados com fins de prejuízo ao acesso a tratamento e/ou planos de saúde, tendo em vista o compromisso assumido pelos pesquisadores de apenas utilizar os dados para fins descritivos e não para tomadas de decisão.

Benefícios:

Como benefício científico destaca-se a oportunidade de avaliar o quanto as técnicas de Mineração de Dados podem contribuir para a descrição dos mecanismos fisiológicos que determinam o controle postural de indivíduos pós-AVC. Como benefício científico complementar, tem-se a oportunidade de identificar variáveis com potencial para serem indicadores clínicos de pacientes pós-AVC, os quais podem vir a ser utilizados para avaliar resultados de tratamentos. A partir da identificação de algumas limitações da Mineração de Dados no atendimento dos objetivos será gerada demanda por novas pesquisas na área.

Comentários e Considerações sobre a Pesquisa:

PESQUISA COMPLEXA E IMPORTANTE.

Considerações sobre os Termos de apresentação obrigatória:

Foi apresentado o projeto de pesquisa, a folha de rosto, o TCUD, e termo de aprovação do comitê de ética em relação ao projeto cujos dados secundários fornecerão subsídios de análise para este projeto.

Recomendações:

Nenhuma.

Endereço: Rua Imaculada Conceição 1155

Bairro: Prado Velho

CEP: 80.215-901

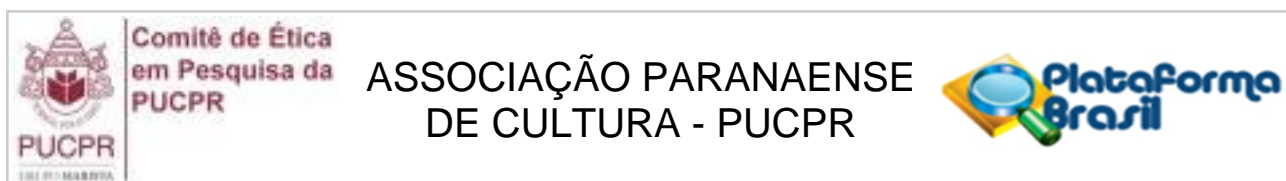
UF: PR

Município: CURITIBA

Telefone: (41)3271-2103

Fax: (41)3271-2103

E-mail: nep@pucpr.br



Continuação do Parecer: 991.103

Conclusões ou Pendências e Lista de Inadequações:

Aprovado.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

CURITIBA, 19 de Março de 2015

Assinado por:
NAIM AKEL FILHO
(Coordenador)

Endereço: Rua Imaculada Conceição 1155

Bairro: Prado Velho

CEP: 80.215-901

UF: PR

Município: CURITIBA

Telefone: (41)3271-2103

Fax: (41)3271-2103

E-mail: nep@pucpr.br


ANEXO B - TERMO DE COMPROMISSO DE UTILIZAÇÃO DE DADOS - TCUD

Nós, Luiz Henrique Giovanini Marrega, Simone Massaneiro Silva, Julio Cesar Nievola e Elisangela Ferretti Manffra, abaixo assinados, pesquisadores envolvidos no projeto de título **ANÁLISE E MODELAGEM DO CONTROLE DO EQUILÍBRIO EM INDIVÍDUOS PÓS-AVC**, nos comprometemos a manter a confidencialidade sobre os dados coletados por Simone Massaneiro Silva (abaixo assinada) para a sua dissertação de mestrado intitulada "*ANÁLISE DO CONTROLE POSTURAL DE INDIVÍDUOS PÓS-ACIDENTE VASCULAR ENCEFÁLICO FRENTE A PERTURBAÇÕES DOS SISTEMAS VISUAL E SOMATOSSENSORIAL*" e apresentada ao Programa de Pós-Graduação em Tecnologia em Saúde (PPGTS) da PUCPR no ano de 2012 sob a orientação da Prof.^a Dr.^a Elisangela Ferretti Manffra (abaixo assinada), bem como a privacidade de seus conteúdos, como preconizam os Documentos Internacionais e a Res. 466/12 do Conselho Nacional de Saúde.

Informamos que os dados a serem utilizados dizem respeito à manutenção do equilíbrio postural em indivíduos saudáveis e pós-AVC coletados entre julho de 2011 e setembro de 2011.

Curitiba, 20 de fevereiro de 2015

Envolvidos na manipulação e coleta dos dados

Nome completo	CPF	Assinatura
Luiz Henrique Giovanini Marrega	063.041.129-80	
Simone Massaneiro Silva	010.336.149-98	Simone M. Silva.
Julio Cesar Nievola	463.009.889-53	Julio Cesar Nievola
Elisangela Ferretti Manffra	874.269.009-91	Elisangela.