

Laboratório de Modelagem, Análise e Controle de Sistemas Não-Lineares

Departamento de Engenharia Eletrônica

Universidade Federal de Minas Gerais

Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, MG Brasil

Fone: +55 3409-4866 - Fax: +55 3409-4850

aguirre@cpdee.ufmg.br



Técnicas de Reconhecimento de Padrões para a Classificação de Sinais Fisiológicos

Cristina Sady Coelho da Rocha

Tese submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de Doutor em Engenharia Elétrica.

Orientadores: Prof. Luis Antonio Aguirre, PhD
Ubiratan Freitas, Dr.
Prof. Christophe Letellier, Dr.

Belo Horizonte, Março de 2013

“É mais fácil amar o retrato. Eu já disse que o que se ama é uma ‘cena’. ‘Cena’ é um quadro belo e comovente que existe na alma antes de qualquer experiência amorosa. A busca amorosa é a busca da pessoa que, se achada, irá completar a cena. Antes de te conhecer eu já te amava... E então, inesperadamente, nos encontramos com o rosto que já conhecíamos antes de o conhecer. E somos então possuídos pela certeza absoluta de haver encontrado o que procurávamos. A cena está completa. Estamos apaixonados.”

Retratos de Amor, de Rubem Alves

“Compreendi que a vida não é uma sonata que, para realizar sua beleza, tem de ser tocada até o fim. Dei-me conta, ao contrário, de que a vida é um álbum de minissonatas. Cada momento de beleza vivido e amado, por efêmero que seja, é uma experiência completa que está destinada à eternidade. Um único momento de beleza e amor justifica a vida inteira.”

Concerto para Corpo e Alma, de Rubem Alves

Dedicatória

*Ao meu grande amor
Murilo Gomes.*

*“In the vastness of space and
the immensity of time, it is
my joy to share a planet and
an epoch with Murilo.”*

Adaptação de *Cosmos*, de Carl Sagan¹

¹Diante da vastidão do espaço e da imensidade do tempo, é uma alegria para mim partilhar um planeta e uma época com Murilo.

Agradecimentos

Ao meu amor Murilo, por fazer parte de minha história, pelo amor, pela amizade, pelo companheirismo, pelo incentivo desde o início desta longa caminhada, por dividir comigo o gosto pelas ciências biomédicas, por ter sido a grande fonte de inspiração e motivação para este trabalho e por ter preenchido os meus dias com música, poesia e alegria! Não há palavras que possam expressar o amor que sinto por você!

À minha família pelo amor, apoio e incentivo.

Ao meu orientador Luis Antonio Aguirre, pelos ensinamentos, pelas preciosas discussões e contribuições, pelas sugestões, pelo aprendizado, pelo incentivo, pelos momentos de alegria e pela grande amizade! Você é um grande exemplo de amor e dedicação à pesquisa e à profissão!

Ao meu co-orientador Ubiratan Freitas, pelas contribuições e pelo apoio durante o estágio sanduíche na França.

Ao Prof. Eduardo Mazoni, pela amizade, pela paciência, pelo apoio, pelo incentivo e pelo senso de humor brincalhão e divertido que preenche os nossos árduos dias de labuta com alegria!

Ao Prof. Carlos Júlio, por sempre ter aberto as portas para trabalharmos em conjunto, pela confiança, pelas sugestões, pelo incentivo, pelo amor à pesquisa e por ter me mostrado os primeiros caminhos para a entrada na Engenharia Biomédica.

Ao Prof. Antônio de Pádua Braga, pela simpatia, amizade e sugestões.

Ao Prof. Hani, pela amizade, apoio, incentivo e entusiasmo.

Ao Prof. Antônio L. P. Ribeiro, do HC-UFMG, por disponibilizar as séries de VFC estudadas neste trabalho.

Aos meus amigos do MACSIN, CPH, LITC, NEPEB, dentre outros alunos do PPGEE, em especial, Tales, Giovani, Gledson, Roger, Mara, Dair, Fernanda Wanderley, Victor Costa, André Paim, Fred, Edson, Éverthon, Cássio, Renilson, Dimas, Rodrigo, Anny, Wendy e Grazi, pelo apoio, amizade, discussões, cooperação, incentivo, pelas ajudas iniciais com o MATLAB, pela troca de conhecimentos, pelas risadas e pelos momentos de alegria.

Aos amigos do CORIA, em especial, Dounia e Elise, pela receptividade, pela ajuda no aprendizado da língua francesa, pelas sugestões e discussões iniciais, pelas risadas à longa distância.

À amiga da Física, Natália Landin, pelo apoio, pela imensa amizade de longos anos e pelas ajudas iniciais com o Linux, Fortran, Latex, dentre outros.

Aos amigos Ricardo Chácara e Ricardo Nunes, pelo apoio e longa amizade.

Ao CNPq, CNRS e FAPEMIG pelo apoio financeiro.

Epígrafe

“(.. .) Por que o sonho de ontem foi pleno de ternura e lágrimas, o de anteontem foi brincalhão e exuberante, um anterior foi aventureiro e de uma busca sombria e constante? Por que razão nesse desfrute belezas indescritíveis da música, por que paio no ar e vôo naquele outro, com enlevo de uma águia, em direção a picos distantes? Tais criações, que dão margem e desafogo aos nossos impulsos de ternura, de humor, de aventura, ou a nosso anseio de música de montanhas – cada qual terá à mão seus próprios exemplos mais notáveis –: são interpretações de nossos estímulos nervosos durante o sono, interpretações muito livres, muito arbitrárias, de movimentos do sangue e das vísceras, da pressão do braço e das cobertas, dos sons do sino da torre, dos cata-ventos, dos noctívagos e outras assim.

(.. .) A vida de vigília não tem essa liberdade de interpretação que tem a vida que sonha, é menos inventiva e desenfreada – mas devo acrescentar que nossos impulsos, nas horas despertas, igualmente não fazem senão interpretar, os estímulos nervosos e, conforme suas necessidades, estabelecer as “causas” deles? que não há diferença essencial entre sonhos e vida desperta? que, mesmo comparando estágios de cultura bem diversos, a liberdade de interpretação desperta, em um, não fica atrás da liberdade do outro em sonhos? que também nossos juízos e valorações morais são apenas imagens e fantasias sobre um processo fisiológico de nós desconhecido, uma espécie de linguagem adquirida para designar certos estímulos nervosos? que tudo isso que chamamos de consciência é um comentário, mais ou menos fantástico, sobre um texto não sabido, talvez não “sabível”, porém sentido?

(.. .) O que são, então, nossas vivências? São muito mais aquilo que nelas pomos do que o que nelas se acha! Ou deveríamos até dizer que nelas não se acha? Que viver é inventar? (.. .)”

Aurora, de Friedrich Nietzsche

Sumário

Resumo	xv
Abstract	xvii
Lista de Figuras	xxii
Lista de Tabelas	xxiv
Lista de Símbolos	xxv
Lista de Abreviações	xxvii
1 Introdução	1
1.1 Motivação e Relevância da Contribuição	1
1.2 Objetivos	5
1.3 Contribuições Esperadas	6
1.4 Organização do Texto	6
2 Fisiologia do Sono e da Variabilidade da Frequência Cardíaca	9
2.1 Registros Polissonográficos	10
2.1.1 As variáveis neurológicas	11
2.1.2 As variáveis cardiorrespiratórias	14
2.2 Arquitetura do Sono	18
2.2.1 Macroarquitetura do Sono	19
2.2.2 Microarquitetura do Sono	22

2.3	Organização do Sono	27
2.3.1	O Hipnograma	27
2.3.2	Os parâmetros e os índices de análise do sono	28
2.4	Ciclo Sono-Vigília	30
2.4.1	Processos de Regulação dos Estágios de Sono e de Vigília	32
2.5	Funções Fisiológicas do Sono	34
2.5.1	O Metabolismo Cerebral	35
2.5.2	A Função Cardiovascular	35
2.6	Distúrbios do Sono e Patologias Ligadas ao Sono	38
2.7	Fisiologia da Variabilidade da Frequência Cardíaca	38
2.7.1	O Sinal da Variabilidade da Frequência Cardíaca	38
2.7.2	Controle Neural da Frequência Cardíaca	42
3	Ferramentas de Dinâmica Não-Linear Aplicadas à Análise e Classificação de Sinais Fisiológicos	43
3.1	Introdução	43
3.2	Análise Topológica	49
3.2.1	Dinâmica Simbólica	49
3.2.2	Entropia	61
3.3	Máquinas de Vetores de Suporte	64
3.3.1	Técnicas de Otimização para a determinação do Hiperplano Ótimo	70
3.3.2	Hiperplano ótimo para padrões não-separáveis	72
3.3.3	Núcleo do produto interno	73
4	Classificação da Variabilidade da Frequência Cardíaca por meio de Sequências Simbólicas	77
4.1	Introdução	77
4.1.1	Disfunção Autonômica	79
4.1.2	Disautonomia Parassimpática e Disfunção Ventricular Esquerda	81
4.1.3	Disfunção Autonômica e Prognóstico por meio de Técnicas de Análise da Variabilidade da Frequência Cardíaca	83

4.1.4	Variabilidade da Frequência Cardíaca e Dinâmica Simbólica . . .	85
4.2	Objetivos	89
4.3	Material e Métodos	89
4.3.1	Protocolo Experimental	89
4.3.2	Aquisição das séries e Pré-processamento	92
4.3.3	Extração de Características	93
4.3.4	Classificação	101
4.4	Resultados	105
4.4.1	Análise da VFC por meio da Dinâmica Simbólica	105
4.4.2	Análise da VFC por meio de Testes Estatísticos - ANOVA simples	106
4.4.3	Resultados obtidos na Classificação	106
4.5	Discussão	109
4.6	Conclusões do Capítulo	111
5	Classificação Automática dos Estágios do Sono por meio de Sinais Fisiológicos	
	Disponíveis a um Ventilador	113
5.1	Introdução	113
5.2	Materiais e Métodos: Dados e Processamento dos Sinais	117
5.2.1	Pacientes	117
5.2.2	Protocolo Experimental	118
5.2.3	Registro de Dados	121
5.2.4	Pré-processamento do Fotopletismograma	125
5.2.5	Pré-processamento da Vazão	126
5.3	Métodos: Arquitetura do Classificador e Extração de Características . .	127
5.3.1	Extração de Características	128
5.3.2	Classificador	131
5.3.3	Classificação Dependente do Paciente	135
5.3.4	Classificação Independente do Paciente	136
5.3.5	Teste de Pertinência de Características	138
5.3.6	Efeito do Pós-Processamento no Desempenho do Classificador . .	138

5.3.7	Análise Prospectiva da Adaptação à Ventilação Não Invasiva . . .	139
5.3.8	Estimação de Hipnogramas de Noites sem Polissonografias . . .	140
5.4	Resultados	142
5.4.1	Classificação	142
5.4.2	Pós-Processamento do Classificador	142
5.4.3	Adaptação à VNI	143
5.5	Discussão	148
5.6	Conclusões do Capítulo	150
6	Conclusões e Propostas de Continuidade	151
6.1	Propostas de Continuidade	152
6.1.1	Limitações do Trabalho	153
	Bibliografia	169
A	Hipnogramas Estimados	171
A.1	Estimação de Hipnogramas para noites sem Polissonografia via Modelos	171
B	Testes de Hipóteses	177
B.1	Análise de Variância (ANOVA)	177
B.2	Teste de Wilcoxon	179
B.3	Coeficiente de Kappa Cohen	181

Resumo

A regulação da frequência cardíaca, da pressão arterial e de outras variáveis cardiovasculares é mediada, principalmente, pelo sistema nervoso autônomo, junto com os mecanismos homeostáticos intrínsecos e extrínsecos. Complexas alterações são observadas nessas variáveis e a análise de sua variabilidade tornou-se uma importante ferramenta não-invasiva para estudar as interações simpato-vagais em diversas condições fisiopatológicas, tais como o infarto do miocárdio, a neuropatia diabética e a doença de Chagas. A análise da variabilidade da frequência cardíaca (VFC) aliada às técnicas de análise de sistemas dinâmicos não-lineares têm se mostrado ferramentas promissoras para tais estudos. Métodos de modelagem têm sido utilizados, na literatura, para a síntese e análise de sinais fisiológicos, tais como, séries temporais de intervalos RR (derivados de sinais de ECG) e sinais respiratórios. A escolha da técnica de modelagem não-linear parece ser mais aceitável para representar, de forma confiável, os processos biológicos, uma vez que esses processos são susceptíveis a apresentar algum tipo de não-linearidade. Este trabalho apresenta a aplicação da dinâmica simbólica, de padrões estatísticos ordinais e de outras técnicas de dinâmica não-linear para distinguir entre o comportamento dinâmico da VFC em pacientes saudáveis e chagásicos. Para isso, foi utilizada a entropia de Shannon, calculada a partir de sequências infinitas de símbolos dos tacogramas. Os resultados mostraram que foi possível distinguir duas classes: indivíduos saudáveis e indivíduos chagásicos com 84% de acerto; e três classes: indivíduos saudáveis, indivíduos chagásicos com e sem cardiopatia com 70% de acerto, utilizando uma máquina de vetor de suporte (SVM) multiclases.

Estudos recentes vêm se preocupando também com a classificação automática do sono, que é de grande importância na clínica médica, uma vez que diversas patologias

e distúrbios do sono, tais como a apnéia obstrutiva do sono, a narcolepsia e a insônia, estão relacionadas com a qualidade do sono. Diversos autores utilizaram técnicas de dinâmica não-linear e modelagem autorregressiva para aplicação no estagiamento do sono. A ventilação não invasiva (VNI), um tratamento reconhecido para insuficiência respiratória crônica hipercápnica, é predominantemente aplicada à noite. Entretanto, a qualidade do sono é raramente avaliada devido à complexidade tecnológica necessária para tal avaliação. Uma nova técnica para a classificação automática do sono foi proposta, neste trabalho, para pacientes submetidos à VNI. Essa nova técnica requer somente sinais disponíveis a ventiladores domiciliares (fluxo de ar e saturação de oxigênio) e o photoplethmograma (PPG), um sinal também já disponível em alguns ventiladores. Conseqüentemente, não são necessários sinais de EEG, EOG, EMG ou ECG. Características cardiorrespiratórias são extraídas dos três sinais e utilizadas como entradas em uma SVM multiclass. Dois diferentes tipos de classificação do sono foram investigados: um distingue três estágios (vigília, sono REM e sono não-REM), e o outro retorna cinco estágios (vigília, sono REM e estágios N1, N2 e N3). Classificadores dependentes e independentes do paciente foram testados e os hipnogramas obtidos foram comparados com os obtidos pela classificação visual/manual de um especialista do sono. Taxas de acerto médias de 91% e de 84% foram obtidas com classificadores dependentes do paciente para três e cinco estágios, respectivamente. Para classificadores independentes do paciente, foram obtidas taxas de acerto médias de 78% e 62% para três e cinco estágios, respectivamente. Reduções de 4,5% e de 5% nas taxas de acerto foram observadas, respectivamente, para três e cinco estágios, quando características relacionadas com o PPG e com o fluxo foram retiradas. Os resultados sugerem que, a longo prazo, a avaliação do sono e a monitorização noturna em domicílio é viável em pacientes tratados com VNI. Essa técnica poderia, ainda, ser integrada aos ventiladores.

O presente trabalho visa, portanto, a aplicar técnicas de reconhecimento de padrões para a classificação da cardiopatia chagásica por meio da extração de características de sinais de VFC. Além disso, este trabalho também tem como objetivo, classificar os estágios do sono de pacientes ventilados por meio de sinais coletados sem a utilização de eletrodos.

Abstract

The regulation of heart rate, blood pressure and other cardiovascular variables is mediated primarily by the autonomic nervous system, along with the homeostatic mechanisms intrinsic and extrinsic. Complex changes are observed in these variables and analysis of its variability has become an important noninvasive tool to study the sympathetic-vagal interactions in many pathophysiological conditions such as myocardial infarction, diabetic neuropathy and Chagas disease. The analysis of heart rate variability (HRV) coupled with a technical analysis of nonlinear dynamical systems have shown to be promising tools for such studies. Modeling methods have been used in the literature for the synthesis and analysis of physiological signals such as time series of RR intervals (ECG signals derived) and respiratory signals. The choice of nonlinear modeling techniques seems to be more acceptable to depict reliably the biological processes, since these processes are likely to present some nonlinearity. This paper presents the application of symbolic dynamics, ordinal pattern statistics and other techniques of nonlinear dynamics to distinguish between the dynamic behavior of HRV in healthy and chagasic subjects. For this, we used the Shannon entropy, calculated from finite sequences of symbols of tachograms. The results demonstrate that it was possible to distinguish two classes: healthy individuals and chagasic individuals with 84% accuracy; and three classes: healthy individuals, chagasic individuals with and without cardiopathology with 70% accuracy, using a Support Vector Machine (SVM) multiclass classification.

Recent studies also have been concerned with the automatic sleep classification, which is of great importance in clinical medicine, since several diseases and sleep

disorders such as obstructive sleep apnea, narcolepsy and insomnia, are related to the quality sleep. Several authors use nonlinear dynamics techniques and autoregressive modeling for application in sleep staging. Non invasive ventilation (NIV), a recognized treatment for chronic hypercapnic respiratory failure, is predominantly applied at night. Nevertheless, the quality of sleep is rarely evaluated due to the required technological complexity. A new technique for automatic sleep staging is here proposed for patients treated by NIV. This new technique only requires signals (airflow and hemoglobin oxygen saturation) available in domiciliary ventilators plus a photo-plethysmogram, a signal already managed by some ventilators. Consequently, no EEG, EOG, EMG nor ECG are needed. Cardio-respiratory features are extracted from the three selected signals and used as input to a SVM multi-class classifier. Two different types of sleep scoring were investigated: one distinguish three stages (wake, REM sleep and nonREM sleep), and one returns five stages (wake, REM sleep, N1, N2 and N3 stages). Patient-dependent and patient-independent classifiers were tested comparing the resulting hypnograms with those obtained from visual/manual scoring by a sleep specialist. An average accuracy of 91% (84%) was obtained with three-stage (five-stage) patient-dependent classifiers. With patient-independent classifiers, an average accuracy of 78% (62%) was obtained when three (five) sleep stages were scored. Also if the PPG-based and flow features are left out, a reduction of 4.5% (resp. 5%) in accuracy is observed for the three-stage (resp. five-stage) cases. Our results suggest that long-term sleep evaluation and nocturnal monitoring at home is feasible in patients treated by NIV. Our technique could be even integrated into ventilators.

The present work aims to apply pattern recognition techniques for classification of Chagas heart disease by extracting features of HRV signals. Furthermore, this study also aims to classify the sleep stages in ventilated patients using signals collected without electrodes.

Lista de Figuras

1.1	Número de artigos que fazem referência ao sono e à respiração entre os anos de 1991 e 2010. Fonte: ISI WEB OF SCIENCE, novembro de 2012. Palavras-chave: <i>sleep and breathing</i>	2
1.2	Número de artigos que fazem referência à classificação automática dos estágios do sono entre os anos de 1991 e 2010. Fonte: ISI WEB OF SCIENCE, novembro de 2012. Palavras-chave: <i>automatic and sleep and classification</i>	3
2.1	Posição dos eletrodos durante a polissonografia. O lado esquerdo da Figura indica a posição dos eletrodos para o registro do EEG e o lado direito para o registro do EMG e do EOG. CEE = canto exterior do olho esquerdo; CED = canto exterior do olho direito.	14
2.2	Comparação entre os manuais de 1968 (R&K) e o de 2007 (AASM).	23
2.3	Hipnograma	28
2.4	(a) Série de intervalos RR de um indivíduo controle. (b) Segmento supostamente estacionário de 5mim. Fonte: (Gomes 2001).	39
3.1	Partição topológica do mapa de primeiro retorno da função logística para $\mu = 3,9999$ e $N = 10000$ pontos. O ponto crítico está localizado no ponto de máximo da parábola, isto é, $x_c = 0,5$ para qualquer valor de μ	53

3.2	Partição topológica do mapa de primeiro retorno da função logística para (a) $\mu = 3,2$, (b) $\mu = 3,5545$ e $N = 10000$ pontos. O ponto crítico está localizado no ponto de máximo da parábola, isto é, $x_C = 0,5$ (<i>kneading sequence</i>) para qualquer valor de μ	54
3.3	Partição estatística do mapa de primeiro retorno da função logística para (a) $\mu = 3,7$, (b) $\mu = 3,9$, (c) $\mu = 3,98$ e (d) $\mu = 3,999$; $N = 10000$ pontos. O ponto crítico está localizado no ponto de máximo da parábola, isto é, $x_C = 0,5$ para qualquer valor de μ . A partição topológica (C) coincide com a estatística (C_e) em $\mu = 3,999$	56
3.4	Histogramas das sequências simbólicas para uma série com dinâmica caótica obtida por meio do mapa logístico com $\mu = 3,9999$ (a) $q = 2$, $n = 6$; (c) $q = 3$, $n = 4$; (e) $q = 4$, $n = 3$) e para uma série com dinâmica estocástica (variável aleatória ou ruído branco) (b) $q = 2$, $n = 6$; (d) $q = 3$, $n = 4$; (f) $q = 4$, $n = 3$).	58
3.5	Histogramas das sequências simbólicas. Os cálculos foram feitos a partir de 10000 pontos, com $q = 2$ e $n = 6$: (a) $\mu = 3,70$; (b) $\mu = 3,90$; (c) $\mu = 3,98$; (d) $\mu = 3,9999$	59
3.6	Mapa não-linear $\varphi(\cdot)$ do espaço de entrada para o espaço de características.	65
3.7	Ilustração da idéia de um hiperplano ótimo para padrões linearmente separáveis. As setas indicam os vetores de suporte.	67
3.8	Interpretação geométrica das distâncias algébricas de pontos até o hiperplano ótimo para um caso bidimensional.	68
3.9	Arquitetura do SVM, em que $K(\cdot)$ é a função <i>kernel</i> , M é o número de vetores de suporte e P é a dimensão da entrada.	76
4.1	Mapa de primeiro retorno de um tacograma e as variabilidades dos pontos de curta e longa duração	95
4.2	Partição estatística do mapa de primeiro retorno definida ao longo da reta perpendicular à reta x_1	96

4.3	Partição estatística do mapa de primeiro retorno definida ao longo da reta paralela à reta x_1	97
4.4	Etapas do processo de extração de características via DS.	99
4.5	PEO's representados a partir da permutação das amostras do sinal, segundo sua amplitude.	100
4.6	PEO's para séries temporais. Padrões para $n = 3$. Frequência de ocorrência dos padrões π para a função seno (a) e para o ruído branco (b). . .	102
4.7	Histogramas normalizados de (a) um indivíduo do grupo controle; (b) um indivíduo do grupo 2; (c) um indivíduo do grupo 3.	105
4.8	Gráfico <i>boxplot</i> representando os valores de entropia médios encontrados para os três grupos. A reta central representa a mediana das entropias, as bordas da caixa são os percentis 25% e 75% e as retas nas extremidades são os valores máximos e mínimos de entropia.	106
5.1	Montagem da eletroencefalografia. A colocação dos eletrodos está circulada em laranja. Fonte: Figura modificada de (Wikipedia 2013a). . . .	121
5.2	Polissonografia sob ventilação não invasiva: exemplo de um paciente do estudo no laboratório do sono (Adaptado de (Naeck 2011)).	123
5.3	Desenho esquemático da posição dos sensores durante uma polissonografia completa. Fonte: Figura modificada de (Wikipedia 2013b).	124
5.4	As amplitudes de oscilações do PPG foram calculadas usando os máximos e mínimos do sinal de PPG original (gráfico superior). Os intervalos de batimentos cardíacos foram calculados usando cruzamentos por zero no sinal de PPG filtrado e interpolado (gráfico inferior).	127
5.5	(a) Mediana da amplitude de oscilações do pletismograma. (b) Hipnograma da noite N_2 do paciente P_1	128
5.6	Esboço do processo de classificação.	132
5.7	Etapas da classificação.	137

5.8	Hipnograma produzido por um classificador dependente do paciente (a) Segundo um neurologista para o cenário de uma noite. (b) Caso do paciente P_{12} . Neste caso, $\bar{A}_{cc} \approx 92\%$	144
5.9	Hipnograma produzido por um classificador dependente do paciente (a) Segundo um neurologista para o cenário de duas noites. (b) Caso do paciente P_{12} . Neste caso, $\bar{A}_{cc} \approx 83\%$	145
A.1	Hipnogramas estimados para o paciente (a) P_1 : noite N_3 (gráfico su- perior) e noite N_4 (gráfico inferior); (b) P_8 : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior).	172
A.2	Hipnogramas estimados para o paciente (a) P_9 : noite N_3 (gráfico supe- rior) e noite N_4 (gráfico inferior); (b) P_{10} : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior).	173
A.3	Hipnogramas estimados para o paciente (a) P_{11} : noite N_3 (gráfico su- perior) e noite N_4 (gráfico inferior); (b) P_{12} : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior).	174
A.4	Hipnogramas estimados para o paciente (a) P_{13} : noite N_3 (gráfico supe- rior) e noite N_4 (gráfico inferior); (b) P_2 : noite N_3 (gráfico superior), P_3 : noite N_3 (gráfico do meio) e P_{14} : noite N_4 (gráfico inferior).	175

Lista de Tabelas

2.1	Tabela resumida das características dos estágios do sono, segundo classificação de (Rechtschaffen e Kales 1968)	21
4.1	Lista de Características	103
4.2	Desempenho do SVM	107
4.3	<i>p</i> – valores (<i>p</i>) para a comparação das médias das características entre os grupos. Os índices <i>i</i> representam os índices listados na Tabela 4.1 em ordem crescente de <i>p</i> . Os números subscritos em <i>p</i> representam os grupos que estão sendo comparados. Os valores em negrito indicam que as características relacionadas à DS e aos PEO's são significativas.	108
4.4	Desempenho do SVM multiclass para 2 e 3 classes	108
4.5	Desempenho do SVM multiclass para 2 e 3 classes para grupos de características específicos	109
5.1	Características clínicas dos treze pacientes (6 mulheres e 7 homens) incluídas no estudo de (Naeck 2011).	118
5.2	Programação do procedimento de ventilação não invasiva	119
5.3	Sinais registrados durante cada polissonografia nas noites N_1 , N_2 e N_{15} . Com exceção da pneumotacografia e da Tosca 500, todos os sensores são da marca CIDELEC.	122
5.4	Lista de características calculadas para cada época de 30 s.	130

5.5	Resultados para 5 estgios (classes) do sono. \bar{A}_{cc} representa o valor mdio de A_{cc} definido em (5.6). \bar{k} representa a mdia do coeficiente Kappa Cohen. DP: dependente do paciente; IP: independente do paciente e o nmero entre parentesis representa o nmero de noites registradas.	143
5.6	Resultados para 3 estgios (classes) do sono. Veja as definies na legenda da Tabela 5.5	143
5.7	Resultados para 5 estgios (classes) do sono retirando caractersticas especficas. Veja legenda da Tabela 5.5 para definies.	144
5.8	Resultados para 3 estgios (classes) do sono retirando caractersticas especficas. Veja legenda da Tabela 5.5 para definies.	145
5.9	Teste pareado de <i>Wilcoxon</i> para significncia estatstica dos diferentes desempenhos quando caractersticas especficas <i>no</i> so utilizadas. O asterisco indica os valores que <i>no</i> so estatisticamente significativos.	146
5.10	Resultados para 5 e 3 estgios (classes) do sono para a classificao com filtro. Veja legenda da Tabela 5.5 para definies.	146
5.11	Teste de <i>Wilcoxon</i> pareado para significncia estatstica entre os Δ 's das noites N_2 e N_{15} de cada paciente para classificaes de 5 e 3 estgios. Os nmeros em negrito indica os valores que <i>so</i> estatisticamente significativos.	147
B.1	Evidncia de rejeio de H_0 em funo do <i>p-valor</i>	179

Lista de Símbolos

- q = Número de partições;
 n = Tamanho das sequências de símbolos;
 s_i = Sequências simbólicas;
 H = Entropia de Shannon;
 π = Padrão ordinal;
 \mathbf{w} = Vetor de pesos;
 b = viés;
 $\Phi(\mathbf{w})$ = Função de custo;
 $J(\mathbf{w}, b, \alpha)$ = Função lagrangiana;
 α_i = Multiplicadores de Lagrange;
 ϵ_i = Variáveis de folga;
 $K(\mathbf{x}, \mathbf{x}_i)$ = Núcleo do produto interno (*kernel*);
 C = Parâmetro de regularização da SVM;
 σ = Parâmetro do núcleo RBF da SVM;
 A_{cc} = Taxa de acerto;
 k = coeficiente de Kappa Cohen

Lista de Abreviações

AASM	<i>American Academy of Sleep Medicine;</i>
ASDA	<i>American Sleep Disorders Association, ASDA;</i>
CAP	Padrões Cíclicos Alternantes (<i>Cyclic Altternating Pattern</i>);
EEG	Eletroencefalograma;
ECG	Eletrocardiograma;
EMG	Eletromiograma;
EOG	Eletro-oculograma;
DAC	Disfunção Autonômica Cardíaca;
DNL	Dinâmica Não-Linear;
DS	Dinâmica Simbólica;
FC	Frequência Cardíaca;
HF	Faixa de altas frequências (<i>High Frequency Band</i>) da densidade de potência espectral do sinal de VFC: 0,15 a 1,2Hz para ratos e 0,15 a 0,4Hz para humanos;
LF	Faixa de baixas frequências (<i>Low Frequency Band</i>) da densidade de potência espectral do sinal de VFC: 0,04 a 0,15Hz para humanos;

NN	Intervalos RR normais (<i>Normal-to-Normal RR intervals</i>):
MD	Micro-despertar;
REM	<i>Rapid Eye Movement</i> ;
não-REM	<i>non-Rapid Eye Movement</i> ;
PEO's	Padrões Estatísticos Ordinais;
PPG	Fotopletismograma;
PSD	Densidade de potência espectral (<i>Power Spectral Density</i>)
QRS	Complexo das ondas Q, R e S do eletrocardiograma; caracteriza a contração do ventrículo cardíaco
R&K	Rechtschaffen e Kales;
RMSE	<i>Rate Mean Square Error</i> ;
RNA	Rede Neural Artificial;
RR	Intervalo de tempo entre duas ondas R sucessivas do ECG; é um estimador da duração do batimento cardíaco;
SNA	Sistema Nervoso Autônomo;
SNC	Sistema Nervoso Central;
SP	Sono Paradoxal;
S_pO_2	Saturação de oxigênio da hemoglobina;
SVM	Máquina de Vetor de Suporte (<i>Support Vector Machine</i>);
VFC	Variabilidade da Frequência Cardíaca;
VNI	Ventilação Não-Invasiva

Introdução

1.1 Motivação e Relevância da Contribuição

O sono é um estado de consciência e capacidade de resposta reduzidas, tanto aos estímulos externos quanto internos sendo, entretanto, um processo fisiológico ativo (Shneerson 2005). Nos humanos o sono representa tipicamente um terço da vida adulta e, aproximadamente, a metade da vida dos lactentes (Rabarimanantsoa-Jamous 2008). Assim, a eficiência e a qualidade do sono são primordiais para uma boa recuperação física e mental. Além disso, muitas patologias, como a síndrome da apnéia obstrutiva, a narcolepsia e a insônia estão intimamente ligadas à fisiologia e aos estágios do sono.

A evolução dos estágios do sono foi amplamente divulgada por (Rechtschaffen e Kales 1968) com a análise de parâmetros, técnicas e reconhecimento de formas de ondas de três sinais fisiológicos: eletroencefalograma (EEG), eletromiograma (EMG) mentoniano e submentoniano e eletro-oculograma (EOG). Esses estágios foram classificados em: vigília, estágio de movimento rápido do olhos (*Rapid Eye Movement, REM*) e estágio em que não há movimento rápido dos olhos, que se subdivide em quatro outros estágios, que serão descritos no Capítulo 2.

Apesar de a ciência do sono ser uma disciplina estabelecida, principalmente a partir da década de 1960 com a publicação do manual de (Rechtschaffen e Kales 1968), muitos estudos clínicos ainda não consideram a variação do sono adicionada a fatores cronobiológicos na interpretação dos dados. O aumento do conhecimento sobre o sono e sobre as técnicas de análise de sinais fisiológicos tende a aumentar o impacto do sono na ciência médica. Desde então, houve um crescimento no número de artigos publicados que tratam do sono. A Figura 1.1 ilustra o número de artigos publicados

nos últimos 20 anos que fazem referência ao sono e a patologias respiratórias ligadas ao sono.

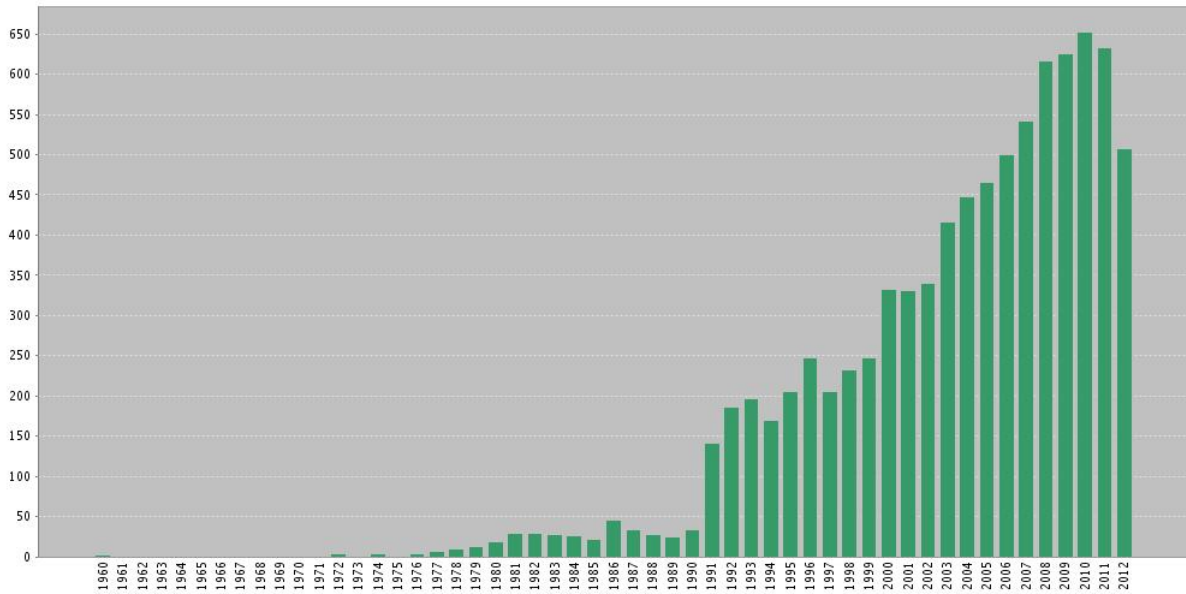


Figura 1.1: Número de artigos que fazem referência ao sono e à respiração entre os anos de 1991 e 2010. Fonte: ISI WEB OF SCIENCE, novembro de 2012. Palavras-chave: *sleep and breathing*.

Por outro lado, ainda são poucos os estudos que fazem alusão a alguma técnica de classificação automática dos estágios do sono, como mostra a Figura 1.2. Nesta Figura ainda pode-se verificar que este tipo de análise é algo muito recente, uma vez que houve um aumento significativo do número de artigos publicados por ano apenas no último ano, indicando que ainda há questões em aberto no que concerne à classificação dos estágios do sono. Além disso, os poucos estudos feitos nesta área ainda são controversos, principalmente no que concerne à própria conceituação dos estágios pelos neurologistas, uma vez que ainda não há uma padronização dos manuais existentes na área (Kales et al. 2001; Himanen e Hasan 2000; Novelli, Ferri e Bruni 2010).

Um procedimento bastante disseminado na prática clínica é a ventilação mecânica não invasiva (VNI). Tal procedimento pode ser aplicado tanto durante o dia quanto durante a noite. À noite, a assistência respiratória aplicada mediante máscaras apropriadas é indicada para pacientes com síndrome de apneia obstrutiva do sono e também

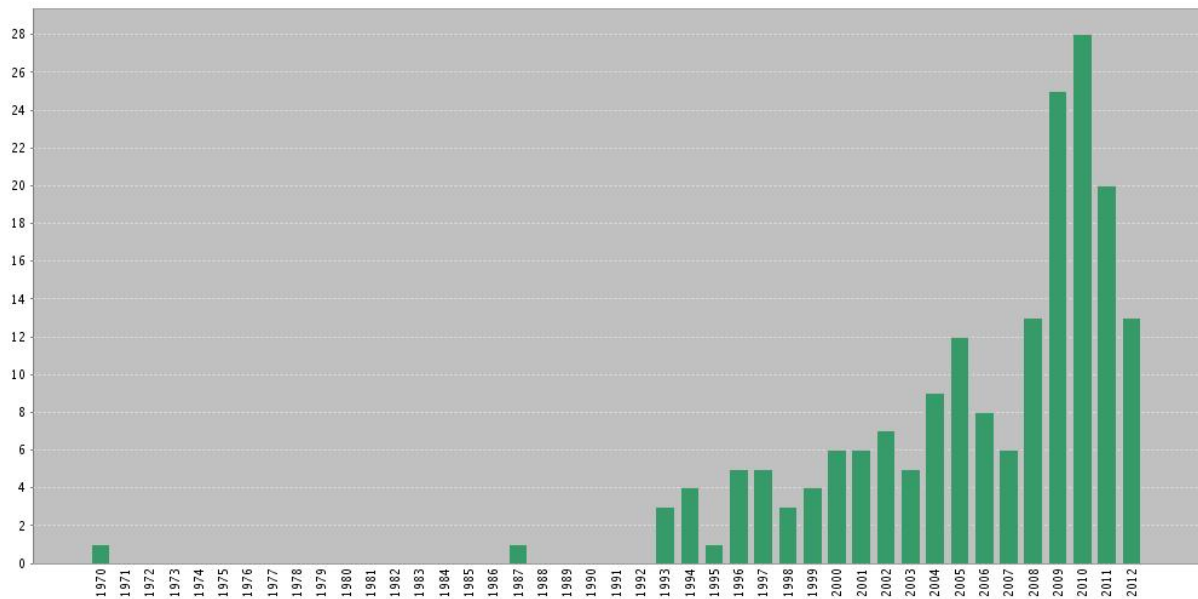


Figura 1.2: Número de artigos que fazem referência à classificação automática dos estágios do sono entre os anos de 1991 e 2010. Fonte: ISI WEB OF SCIENCE, novembro de 2012. Palavras-chave: *automatic and sleep and classification*.

para pacientes com disfunções respiratórias crônicas associadas à doença pulmonar obstrutiva crônica e à síndrome de obesidade mórbida (Rodrigues et al. 2010). As interações paciente-ventilador tornam-se especialmente complexas durante o sono, pois características específicas da dinâmica ventilatória podem variar de acordo com os estágios de sono. Nesse contexto, torna-se importante garantir que a dinâmica do sono não dificulte ou mesmo inviabilize a instituição de um suporte ventilatório apropriado. A interação inversa também é crítica; isto é, a ventilação deve ser aplicada sem deteriorar a qualidade do sono do indivíduo de modo significativo. Um procedimento clássico para a monitoração de pacientes em exames noturnos é a polissonografia. A polissonografia possibilita registrar simultaneamente diversos sinais fisiológicos. O processamento de sinais biomédicos permite, então, acessar informações sobre variáveis características de um ou mais subsistemas fisiológicos a partir das medições disponíveis.

A análise dos estágios do sono é um exemplo de grande relevância na pesquisa clínica. Por meio da aplicação de técnicas de reconhecimento de padrões em sinais fisiológicos coletados sem a utilização de eletrodos seria possível a construção de uma

ferramenta de classificação automática desses estágios, que seria de grande auxílio na clínica médica, uma vez que o estagiamento do sono geralmente é feito por meio da análise e interpretação visual dos registros polissonográficos. A análise polissonográfica requer a coleta dos sinais em hospitais especializados e profissionais especializados para a análise, interpretação e codificação dos dados. Assim, uma ferramenta de classificação automática que utilize *apenas* sinais disponíveis a um ventilador proporcionaria a coleta dos dados em domicílio, além de ser uma ferramenta de baixo custo. Além disso, a análise do sono é muito importante no auxílio ao diagnóstico e prognóstico de diversas patologias ligadas ao sono, como algumas doenças respiratórias (apneia, hipoventilação, asma e bronquites crônicas, dentre outras) e cardiovasculares (hipertensão, infarto do miocárdio, dentre outras) (Shneerson 2005). Todo o estudo sobre o sono realizado, neste trabalho, se refere a pacientes adultos.

A análise simbólica é uma ferramenta natural para o processamento dos sinais coletados na polissonografia pois a evolução dos estágios de sono é usualmente codificada em uma série simbólica denominada hipnograma. Ademais, as séries das variáveis respiratórias podem ser simbolizadas de acordo com o perfil de interação paciente ventilador (Rodrigues et al. 2010). Por outro lado, as variáveis simbólicas/categóricas obtidas a partir da polissonografia são de naturezas distintas. Algumas delas são binárias, indicando apenas a ocorrência ou não de um dado evento. Outras, em contrapartida, assumem valores sobre alfabetos com maior número de símbolos. Além disso, determinadas escalas de tempo são significativamente distintas. Essas características, e outras não mencionadas, usualmente dificultam o processamento de registros experimentais das referidas variáveis. Assim, um estudo interessante iniciado por (Rodrigues et al. 2010) vem sendo feito com o intuito de avaliar as propriedades dos gráficos de recorrência para as análises conjuntas de séries de assincronismos e do hipnograma. Uma análise de hipnogramas gerados a partir de classificadores será apresentada no Capítulo 5.

Séries simbólicas advindas de sinais de VFC serão utilizadas, neste trabalho, para a classificação de pacientes chagásicos. A doença de Chagas é um dos principais problemas médico-sociais do Brasil, acometendo três a quatro milhões de brasileiros, com

elevado impacto econômico e na morbi-mortalidade dos pacientes. O curso clínico da doença é extremamente variável e, apesar de muitos indivíduos permanecerem assintomáticos por longos períodos, aproximadamente um terço dos pacientes infectados desenvolvem doença cardíaca potencialmente letal, incluindo arritmias ventriculares malignas e insuficiência cardíaca (Biolo, L. E N. 2010). Entretanto, esta patologia não está restrita à América Latina. Indivíduos com doença de Chagas foram identificados em países não endêmicos da Europa, na Austrália, no Japão, Canadá e nos Estados Unidos, em que há um grande crescimento na imigração vinda de países endêmicos (Ribeiro et al. 2012). O diagnóstico e prognóstico da doença de Chagas, geralmente, são feitos mediante técnicas invasivas, como a dosagem sorológica e a análise por imagens, e de alto custo. Assim a classificação de pacientes saudáveis ou chagásicos seria uma ferramenta não invasiva, de baixo custo de grande auxílio no diagnóstico da doença de Chagas.

Diante desse panorama, o presente trabalho pretende verificar a real utilidade de algumas ferramentas de análise em problemas de classificação. Espera-se, portanto, que este trabalho possa contribuir com estudos futuros para o desenvolvimento de protocolos médicos que possam ser padronizados e aplicados na clínica médica para o diagnóstico e prognóstico de doenças cardiovasculares que estão intimamente relacionadas com a variação do ritmo cardíaco e para o desenvolvimento de ferramentas de auxílio na classificação automática dos estágios do sono.

1.2 Objetivos

O presente trabalho apresentará uma metodologia baseada em técnicas de reconhecimento de padrões para a classificação da disfunção autônoma da VFC e dos estágios do sono. Assim, os principais objetivos são:

- Classificar os estágios do sono em pacientes com patologias respiratórias por meio de sinais fisiológicos coletados sem a utilização de eletrodos.
- Classificar pacientes saudáveis, chagásicos não cardiopatas e chagásicos cardiopatas

por meio de séries simbólicas e índices tempo-frequenciais obtidos a partir de sinais de variabilidade da frequência cardíaca (VFC).

1.3 Contribuições Esperadas

Assim as principais contribuições desta tese consistem:

1. No desenvolvimento de uma ferramenta de classificação automática dos estágios do sono baseada em técnicas de reconhecimento de padrões. Essa questão será detalhada no Capítulo 5.
2. Na aplicação de séries simbólicas para a classificação da disfunção autonômica via VFC em banco de dados de chagásicos. Essa questão será detalhada no Capítulo 4.

1.4 Organização do Texto

Este texto está organizado da seguinte forma: o Capítulo 2 apresenta uma revisão bibliográfica sobre fisiologia do sono e da variabilidade da frequência cardíaca, com ênfase na arquitetura do sono, na coleta de sinais polissonográficos e em patologias relacionadas ao sono. O Capítulo 3 apresenta uma revisão sobre técnicas de análise e classificação de sinais fisiológicos, com ênfase em sinais de VFC e polissonográficos. Exemplos das aplicações dessas técnicas são apresentados nos Capítulos 4 e 5, os quais englobam estudos de casos de classificação de sinais fisiológicos. Nesses Capítulos são apresentadas as metodologias de classificação e os resultados desta tese com citações das principais publicações geradas por esses resultados e com as respectivas discussões sobre as classificações dos estágios do sono e sobre a classificação da disfunção autonômica na doença de Chagas. Optou-se por fazer a divisão dos resultados e da discussão em dois capítulos com o intuito de construir um texto mais didático para o leitor, uma vez que há duas bases de dados diferentes (uma constituída de sinais de VFC e outra de sinais polissonográficos), dois protocolos experimentais e parte da

metodologia utilizam métodos específicos para cada caso. Assim, o objetivo geral deste trabalho consiste em resolver dois problemas de classificação. E, por fim, no Capítulo 6, são apresentadas as conclusões e algumas propostas para continuidade deste trabalho.

Fisiologia do Sono e da Variabilidade da Frequência Cardíaca

“(..) um dia no inverno, quando eu voltava para casa, minha mãe, vendo que eu estava com frio, ofereceu-me chá, algo que eu normalmente não bebia. No começo eu resisti e então, sem nenhuma razão particular, mudei de idéia. Ela serviu um desses bolinhos rechonchudos chamados ‘madalenas’, que parecem ter sido moldados numa concha de crustáceo. Então, mecanicamente, desanimado após um dia melancólico e com a perspectiva de um amanhã monótono, encostei meus lábios numa colherada do chá no qual eu tinha mergulhado um pedaço de bolinho. Tão logo aquele líquido quente misturado com pedacinhos do bolinho tocou meu palato, senti um calafrio e parei para entender o que estava acontecendo comigo. Um prazer indescritível invadiu meus sentidos. Um prazer isolado, destacado, sem alusão da sua origem. De uma só vez as vicissitudes da vida se tornaram indiferentes para mim, seus desastres inócuos, sua brevidade ilusória – esta nova sensação teve em mim um efeito de preenchimento equivalentes ao do amor e de sua preciosa essência; ou melhor, esta essência não estava em mim, ela agora era eu.”

Caminhos de Swann, de Marcel Proust

Este capítulo tem por objetivo abordar a fisiologia do sono e da variabilidade da frequência cardíaca e os mecanismos que determinam as interrelações do sono com

os diversos sistemas do organismo, com ênfase no controle neural e nos sistemas respiratório e cardiovascular.

2.1 Registros Polissonográficos

Em 1924, Hans Berger, deu origem aos primeiros registros de eletroencefalograma (EEG) humano. Em 1937, durante o estudo do sono de gato, M. B. Rheinberger, recorreu à eletroencefalografia para completar os dados de observações visuais (Rheinberger e Jasper 1937); paralelamente, A. L. Loomis utilizou este novo método na investigação do sono de um homem (Loomis, Harvey e Hobart 1937). À esta época, a maioria dos autores (Blake, Gerard e Kleitman 1939; Davis et al. 1938) distinguiam um estágio de *vigília* e quatro estágios de sono diferentes. Estes estudos revelavam que o sono fisiológico era acompanhado de *fusos* de ondas lentas corticais e subcorticais. A existência de períodos de atividade cortical intensa, com ondas rápidas, durante o sono foi reconhecida vinte anos mais tarde (Rimbaud, Passouant e Cadilhac 1955), enquanto Aserinsky e Kleitman (Aserinsky e Kleitman 1955) classificavam o sono em quatro estágios de *sono não-REM* (*non Rapid Eye Movement sleep*), para sonos sem movimentos oculares rápidos e um estágio de *sono REM* (*Rapid Eye Movement sleep*) para sonos com movimentos oculares rápidos, associados a uma atonia muscular total. Dement (Dement e Tononi 1957) mostram que a atividade cortical rápida é acompanhada de movimentos dos olhos, podendo se traduzir numa atividade onírica. O trabalho de Michel Jouvet (Jouvet 1961) com gatos, tratava-se, mais particularmente, do sono com movimentos oculares rápidos, para quantificar o estágio de *sono paradoxal* (sono REM). Em 1968, uma nova classificação, desenvolvida e sistematizada por (Rechtschaffen e Kales 1968), se apoiou nas três vias da *polissonografia* (EEG, EMG e EOG). Essa classificação é ainda usada, atualmente, como base para a análise do sono, mas é importante notar que uma nova classificação, que se aproxima da desenvolvida por (Rechtschaffen e Kales 1968) foi recentemente publicada (Iber et al. 2007b; Silber 2007) pela Academia Americana de Medicina do Sono (*American Academy of Sleep Medicine, AASM*). Os detalhes dessas duas últimas classificações e as comparações entre elas serão desenvolvidas ao longo desta

Seção. De toda maneira, essas classificações se baseiam em duas escalas de observações diferentes: uma *macro-arquitetura* global e uma *micro-arquitetura* definida por eventos mais pontuais.

O registro polissonográfico é recolhido ao longo de uma noite de sono, com 8 a 12 h de duração, para caracterizar variáveis fisiológicas, como o sono e parâmetros cardiorrespiratórios e musculares (medida dos movimentos dos membros inferiores). Existem vários sistemas de aquisição de sinais polissonográficos, sendo que, nesta tese, foi usado o sistema CID108L, fornecido pela sociedade CIDELEC (*CIDELEC SA, Saint-Gemme sur Loire*). O sistema CID108L consiste num sistema de controle que recolhe sinais analógicos e os transforma em sinais digitais a uma frequência de amostragem fixa. Esse sistema é ligado a um computador para a armazenagem dos dados. Para uma polissonografia completa, são distinguidos os sensores de sinais cardiorrespiratórios, permitindo a análise da respiração, e sensores de sinais neurológicos, permitindo a análise do sono. Detalhes sobre o protocolo experimental serão descritos no Capítulo 5.

2.1.1 As variáveis neurológicas

As variáveis neurológicas coletadas durante uma polissonografia serão descritas a seguir.

O eletroencefalograma

O *eletroencefalograma* (EEG) é um registro da atividade elétrica cerebral por meio de eletrodos colocados sobre o escalpo. Quando os eletrodos são colocados diretamente na superfície do cérebro, este registro passa a denominar-se *eletrocortigrama*. Em estudos humanos, o EEG é registrado em pontos previamente padronizados, deste modo, os EEG's podem ser registrados aproximadamente a partir dos mesmos pontos de um indivíduo em diferentes tempos ou em vários pontos análogos em diferentes indivíduos. Os registros da superfície do cérebro ou mesmo da superfície externa da cabeça demonstram que existe atividade elétrica contínua no cérebro. Tanto a intensidade quanto os padrões dessa atividade elétrica são determinados pelos níveis de excitação

de diferentes partes do sistema nervoso central (SNC) resultantes de *sono*, *vigília* ou doenças cerebrais como a *epilepsia* e as *psicoses*. Logo, o EEG é um importante instrumento de diagnóstico na neurologia clínica, sendo particularmente útil em pacientes com epilepsia.

O EEG é formado por ondas de várias frequências. As frequências dominantes dependem de diversos fatores, inclusive do estágio de vigília, da idade do indivíduo, da localização dos eletrodos de registro e da ausência ou presença de medicamentos ou doença. Quando um adulto normal, em vigília, fica relaxado e com os olhos fechados, as frequências dominantes do EEG registradas sobre os lobos parietal e occipital são de cerca de 8 a 13 Hz, os *ritmos alfa*. Se for pedido ao indivíduo que abra os olhos, o EEG se tornará menos sincronizado, e a frequência dominante aumentará para 13 a 30 Hz, o que é chamado *ritmo beta*. Os *ritmos delta* (0,5 a 4 Hz) e *teta* (4 a 7 Hz) são observados durante o sono (Berne et al. 2004; Shneerson 2005).

As ondas do EEG são derivadas de potenciais pós-sinápticos excitatórios e inibitórios alternantes que ocorrem nos neurônios corticais em decorrência de aferência talamocortical e outras. Os potenciais são produzidos principalmente por correntes extracelulares que se dirigem verticalmente através do córtex quando da geração de potenciais sinápticos nas células piramidais. As correntes extracelulares, geradas pelos potenciais de ação, são de baixa amplitude, rápidas e assíncronas demais para serem registradas com eletrodos para EEG.

Embora uma onda de curta duração no EEG algumas vezes seja denominada *espícula*, este termo não se refere aos potenciais de ação. Os potenciais registrados no EEG têm amplitudes entre 0 e 200 μV (Berne et al. 2004; Guyton e Hall 2006). Estes potenciais refletem a organização de muitas células piramidais, que se dispõem com seus dendritos apicais alinhados em paralelo, formando uma camada de dipolos. Um pólo desta camada é orientado para a superfície cortical e o outro para a substância branca subcortical¹.

¹Observe que o sinal de uma onda de EEG não indica, em si, se as células piramidais estão sendo excitadas ou inibidas. Por exemplo, um potencial negativo no EEG pode ser gerado no córtex por excitação de dendritos apicais ou por inibição perto dos somas. Inversamente, uma onda positiva no EEG pode ser produzida por inibição de dendritos apicais ou por excitação perto dos somas.

A coleta do EEG utiliza o sistema internacional 10–20 de posicionamento dos eletrodos desenvolvido por (Jasper 1958). O manual de (Rechtschaffen e Kales 1968) recomenda a utilização de um único eletrodo em C3 ou C4 e de um eletrodo de referência colocado sob o lobo auricular (A1 ou A2). As derivações utilizadas são, então, C3-A2 ou C4-A1 (contralateral). A Figura 2.1 indica essas derivações. Em alguns casos, talvez possa ser necessária acrescentar uma derivação O1-A2 ou O2-A1 para melhor visualização do ritmo α .

O eletromiograma

O eletromiograma (EMG) permite medir a atividade elétrica dos músculos mentonianos e dos membros inferiores, graças a eletrodos colocados na posição submentoniana ou sobre uma perna. O EMG mentoniano é utilizado na classificação do sono, enquanto o EMG das pernas pode ser utilizado para identificar um movimento brusco ou diagnosticar os movimentos periódicos dos membros inferiores (sintoma das *pernas inquietas*). A presença ou ausência de atividade muscular (tônus) é importante para determinar se o paciente está acordado, ou em sonos REM ou não-REM, para isto, os eletrodos são, geralmente, posicionados na região mentoniana ou em região submentoniana, como mostra a Figura 2.1.

O eletro-oculograma

O eletro-oculograma (EOG) registra, indiretamente, os movimentos dos olhos durante o sono. Movimentos dos olhos são importantes no diagnóstico da sonolência e, em particular, se o sono REM ou a vigília estiver presente. Dois eletrodos são colocados nos cantos externos dos olhos, permitindo assim medir uma diferença de potencial entre a córnea e a retina e, um outro eletrodo fixo é posicionado ao nível do lobo auricular; esta diferença de potencial varia conforme os movimentos dos globos oculares, que estão no mesmo sentido para os dois olhos e que podem ser horizontais ou verticais (Shneerson 2005). A Figura 2.1 indica a posição dos eletrodos para o registro do EOG do sono.

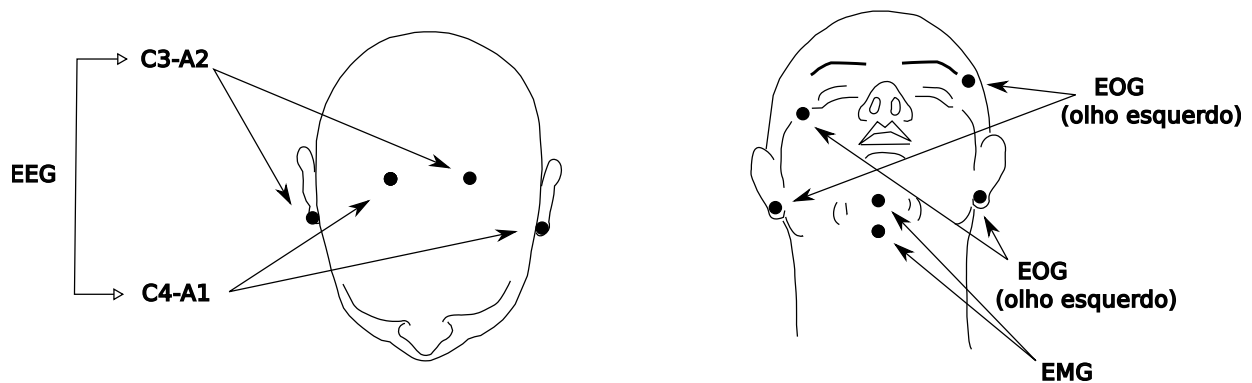


Figura 2.1: Posição dos eletrodos durante a polissonografia. O lado esquerdo da Figura indica a posição dos eletrodos para o registro do EEG e o lado direito para o registro do EMG e do EOG. CEE = canto exterior do olho esquerdo; CED = canto exterior do olho direito.

O eletrocardiograma

Há ainda a medida da atividade elétrica do coração, que também pode ser coletada durante uma polissonografia, mas que não foi aferida durante o protocolo para a análise do sono desta tese. Essa medida é representada pelo eletrocardiograma (ECG), e foi utilizada no protocolo estabelecido para a análise da frequência cardíaca em pacientes chagásicos.

A variabilidade da frequência cardíaca é obtida por meio da análise sucessiva de intervalos RR do ECG. Este sinal e o seu controle, que é mediado pelo sistema nervoso autônomo (SNA), será descrito na seção 2.6.

2.1.2 As variáveis cardiorrespiratórias

Análise dos esforços inspiratórios do paciente

Os esforços inspiratórios do paciente podem ser analisados com o auxílio de métodos invasivos, tais como a pressão esofágica, ou ainda, por métodos não-invasivos, tais como a medida da mobilidade torácico-abdominal:

- *A mobilidade torácico-abdominal*² reflete cada esforço inspiratório do paciente com base na variação da circunferência do tórax e do abdômen. Essa variável é medida

²Conjunto de movimentos de um organismo. Faculdade de se mover.

por meio de cintas (faixas emborrachadas) torácicas e abdominais, semelhantes a dois cintos colocados um ao redor do tórax e do abdômen. Essas faixas são, geralmente, de material têxtil (emborrachadas) e equipadas com um extensômetro. Seu princípio é baseado na mudança da resistência elétrica causada por seu alongamento. As curvas obtidas pela transformação de variações de resistência em variações de tensão refletem as mudanças de dimensão da caixa torácica e do abdômen a cada ciclo respiratório. A desvantagem desse método pode provir da diversidade de corpulência dos pacientes. De fato, em alguns pacientes, os movimentos tóraco-abdominais são pouco visíveis e uma cinta solta ou relaxada durante a noite pode tornar imperceptíveis os esforços inspiratórios. Embora não seja considerada como um método padrão para a identificação dos esforços inspiratórios, a mobilidade torácico-abdominal continua a ser usada em laboratórios de sono. Na verdade, é um método não-invasivo e fácil de implementar em comparação com a pressão esofágica.

- *A pressão esofágica* é considerada como a medida de referência dos esforços inspiratórios. O método consiste em inserir uma sonda esofágica pelo nariz ou pela boca. A sonda pode ser em forma de balão; neste caso mede-se a diferença de pressão entre o balão e a atmosfera. Essa medida também pode ser baseada em um extensômetro situado em uma das extremidades da sonda (a que se encontra no esôfago do paciente); neste caso, o esforço inspiratório é traduzido por uma variação de pressão no extensômetro principal produzindo, assim uma variação de tensão. Uma calibração do extensômetro permite transformar diretamente essa variação de tensão em variação de pressão. A desvantagem desta técnica é o tempo de estabelecimento da sonda e, sobretudo, os eventuais reflexos do paciente que podem modificar a posição da sonda durante o registro, e por conseguinte, afetar a medida. A pressão esofágica não é usada rotineiramente para a avaliação dos esforços respiratórios durante o sono noturno.

A medida da pressão esofágica ou da mobilidade torácico-abdominal depende, em grande parte, da estabilidade de seus respectivos sensores. Essa estabilidade pode ser

muito afetada pelo comportamento do paciente, que pode se mover, tirar os sensores ou, ainda, mudar de posição sobre sua cama. A análise do comportamento do paciente é, então, primordial para a interpretação dos outros sinais polissonográficos durante a aquisição dos dados.

Análise da posição do paciente durante o sono

Ao longo de uma noite, as mudanças de posição ou os movimentos do paciente podem ajudar a decidir se o sinal cardiorrespiratório ou neurológico possui artefato ou não. Além disso, alguns eventos respiratórios, tais como a apneia, são mais frequentes em posição dorsal do que em posição ventral ou lateral.

- *A posição do paciente* é registrada por meio de um sensor de posição e indica a posição do paciente segundo a qual está deitado (dormindo): dorsal, ventral, do lado direito ou do lado esquerdo. Essa variável é medida graças a um sensor de posição colocado sobre o esterno. A posição do paciente é avaliada a cada 30s para monitorar as variações de posição do paciente para explicar o surgimento de artefatos sob a curva de esforços inspiratórios do paciente. Cada mudança de posição pode induzir um microdespertar (MD), assim o registro da posição do paciente permite diferenciar um MD de origem comportamental de um MD de origem respiratória devido a um evento de apnéia, por exemplo.
- *Os movimentos do paciente* são registrados graças a um *actímetro*. Esse pequeno gravador, do tamanho de um relógio, é usado no pulso e armazena na memória os movimentos graças a uma célula piezo-elétrica sensível à aceleração. Esses movimentos podem ser originados de um despertar breve ou de um MD do paciente.

Análise da ventilação

A análise das variações do fluxo e da pressão no sistema ventilatório permite a detecção de eventos, tais como a apnéia e os assincronismos paciente-ventilador. A

consequência desses eventos sobre o plano ventilatório é, essencialmente, medida pela saturação de O_2 no sangue arterial.

- A *oximetria* transcutânea permite quantificar a saturação de O_2 . Mede-se a saturação de O_2 por meio de um oxímetro de pulso, constituído de uma pinça digital encaixada no dedo do paciente. O princípio do oxímetro baseia-se na observação dos espectros de absorção da oxiemoglobina e da hemoglobina, que são diferentes. A oximetria é feita através de feixes luminosos de diferentes comprimentos de onda, na faixa do vermelho e do infravermelho. A absorção específica da hemoglobina e da oxiemoglobina permite determinar a porcentagem de hemoglobina e de oxiemoglobina livre e, conseqüentemente, estimar a saturação de O_2 . Na verdade, 98% do oxigênio presente no sangue está na forma associada à hemoglobina, formando a oxiemoglobina. O oxímetro de pulso também permite registrar a frequência de pulso, em ciclo por minuto, por meio de um *pletismograma* de pulso³.
- A *pressão nasal* é medida por meio de uma sonda nasal inserida no interior das narinas e ligada a um sensor de pressão situado no módulo de aquisição. Dentro do sensor de pressão, há uma membrana que limita a pressão dentro do circuito de ventilação e a pressão atmosférica. Uma variação da pressão do ar provoca um deslocamento da membrana que gera uma variação de tensão. Essa última reflete, então, a pressão do ar medida de forma contínua.

Outras variáveis

Quando o paciente não é ventilado, o diagnóstico de síndrome de apnéia do sono, pode ajudar na análise do ronco⁴. Na realidade, em caso de obstrução das vias aéreas, o escoamento torna-se turbulento, o que gera sons tanto na inspiração quanto na expiração. Os *sons traqueais* são registrados graças a um microfone colocado na base do

³A pletismografia estuda as variações do volume sanguíneo. A cada batimento cardíaco, esse volume muda. Um sensor fotoelétrico, tal como um oxímetro de pulso, pode deduzir o volume sanguíneo e a frequência cardíaca do pulso a partir da absorção do feixe de luz que ele envia.

⁴Os roncos são geralmente definidos por sons de intensidade superior a 76 dB, numa banda passante de 20 a 200 Hz e com duração superior a 0,25 s.

pescoço em posição susternal. A câmara acústica do microfone deve aderir perfeitamente à epiderme. O módulo de aquisição registra igualmente as energias inspiratórias e expiratórias do paciente por integração do sinal sonoro.

No caso de uma polissonografia sob ventilação, todas as variáveis citadas acima são utilizadas, com exceção da pressão nasal que é substituída por duas outras variáveis:

- *O fluxo* no sistema de ventilação que é medido graças a um fluxômetro, em geral, um *pneumotacógrafo de Fleish nº 1*. O princípio de medida consiste em impor uma resistência ao fluxo de ar por interposição de uma grade que induz um escoamento laminar; a equação de Bernoulli pode, então, ser aplicada e o fluxo é obtido por meio da medida de diferença de pressão entre os dois pontos do escoamento laminar.
- *A pressão do ar*, que substitui a pressão nasal, é medida na máscara com a ajuda de um *sensor de pressão diferencial*. Dentro do sensor de pressão, há uma membrana que limita a pressão dentro do circuito de ventilação e a pressão atmosférica. Uma variação da pressão do ar provoca um deslocamento da membrana que gera uma variação de tensão. Essa última reflete, então, a pressão do ar medida de forma contínua.

2.2 Arquitetura do Sono

O sono é definido como um estado de inconsciência do qual uma pessoa pode ser despertada por um estímulo sensorial ou um outro estímulo. Deve ser distinguido do coma, que é um estado de inconsciência do qual a pessoa não pode ser despertada (Guyton e Hall 2006). Contudo, existem diversos estágios do sono, desde o muito leve até o muito profundo, embora a maioria dos pesquisadores do sono dividida-os em dois tipos distintos, cada um com qualidades específicas que serão descritas a seguir.

Durante cada noite, uma pessoa percorre estágios de dois tipos de sono que se alternam um com o outro. Esses estágios são chamados de (1) *sono de ondas lentas (sono não-REM)*, porque, neste tipo de sono, as ondas cerebrais são de grande amplitude e de

baixas frequências e (2) *sono com movimentos rápidos dos olhos (sono REM)*, porque, neste tipo de sono, os olhos realizam movimentos rápidos, apesar de a pessoa ainda estar dormindo.

2.2.1 Macroarquitetura do Sono

A macroarquitetura do sono revela uma descrição macroscópica do desenvolvimento e da organização dos diferentes aspectos do sono, descritos por seus estágios, bem como pelos estágios de alerta (vigília) eventuais. Os estágios do sono são classificados segundo regras precisas, estabelecidas por meio da leitura (interpretação) dos dados polissonográficos (EEG, EMG e EOG simultâneos), que são selecionados em períodos de 30 s⁵. Cada época será, então, codificada de maneira única; uma época corresponderá a um, e somente um, estágio do sono, mesmo que características de diferentes estágios estejam presentes na mesma época. É necessário que mais da metade da época, ou seja mais de 15 s de gravação correspondente a um único estágio, para que toda a época seja identificada neste estágio.

Os Estágios do Sono descritos por Rechtschaffen e Kales

A evolução dos estágios do sono foi amplamente divulgada por (Rechtschaffen e Kales 1968), em 1968, com a análise de parâmetros, técnicas e reconhecimento de formas de ondas de três sinais fisiológicos: EEG, EMG mentoniano e EOG. Esses estágios foram classificados em: *estágio de vigília*, *estágio REM* e *estágio não-REM*, que se subdivide em quatro outros estágios (estágios 1, 2, 3 e 4). Assim, de acordo com (Rechtschaffen e Kales 1968), seis ou sete estágios de sono e de vigília podem ser identificados ao longo de uma noite:

- *Estágio de Vigília*: O estágio de vigília pode ser ativo (*vigília com alerta*) ou passivo (*vigília relaxada*). As características típicas da vigília são a presença de uma ati-

⁵A janela ou época de 30 s não era escolhida nem arbitrariamente, nem segundo critérios fisiológicos, mas devido a restrições materiais impostas pelo dispositivo de registro. Durante as primeiras polissonografias, uma época era representada por uma página de registro de 30 cm passando através do aparelho que traçava a eletroencefalografia a uma velocidade de 10 mm/s. Atualmente, os registros são armazenados em computadores, mas a duração da época permaneceu inalterada.

vidade motriz traduzida por um EMG ativo e uma atividade desincronizada do EEG, com ondas α e θ relativamente rápidas, mas de baixa amplitude (tensão), traduzindo-se em uma atividade tônica dos neurônios: a frequência do EEG varia de acordo com o tipo de atenção. A *sonolência* – considerada como uma fase transitória entre a vigília relaxada (olhos fechados) e os estágio 1 do sono – se traduz por uma atividade α do EEG compreendida entre 8 e 12Hz.

- *Sono Lento Leve*: O sono lento leve compreende uma subdivisão do estágio não-REM e divide-se em dois estágios. O primeiro, chamado *Estágio 1*, é descrito por uma atividade de EEG de fundo com frequências variando de 3 a 7 Hz, essencialmente ondas α e θ , aproximando-se da atividade de vigília mais branda. Essa atividade de fundo é pontuada de *pontos em vértice* ⁶. O *Estágio 2* surge a partir do momento em que se observam os *complexos K* sob a atividade de fundo que fica globalmente inalterada, embora ainda um pouco mais lenta, e os *fusos do sono* ou *spindles*. A atividade muscular e os movimentos oculares diminuem consideravelmente.
- *Sono Lento Profundo*: O sono lento profundo também compreende uma subdivisão do estágio não-REM e é constituído pelos *Estágios 3 e 4*, caracterizado por uma transformação do ritmo do EEG de fundo que se traduz pelo aparecimento de ondas lentas, com frequências de grande amplitude ($> 75 \mu\text{V}$) compreendidas entre 0,5 e 2 Hz, chamadas ondas δ , características dos estágios 3 e 4. O sono lento profundo é traduzido por uma sincronização da atividade elétrica, registrada nas numerosas estruturas corticais e subcorticais, e devido a mecanismos neuronais do tipo *pacemaker*, o núcleo reticular isolado de qualquer aferência contínua a oscilar ritmicamente.
- *Estágio 5 ou de Sono Paradoxal ou Sono REM*: este estágio é definido por uma componente tônica durável, que representa a ausência total de tônus muscular e atividade do EEG próxima do estágio de vigília, novamente desincronizada e uma

⁶Os diferentes grafoelementos aparecem no EEG como eventos pontuais que serão detalhados na seção 2.2.2.

componente fásica, ou seja, de *movimentos oculares rápidos*; com presença de breves contrações musculares e de ondas do EEG em *dentes de serra*. Os padrões do EEG observados durante o sono REM são semelhantes aos de um indivíduo acordado, (em estágio de vigília ativa); as ondas são irregulares e de alta frequência, o que é normalmente sugestivo de atividade nervosa desincronizada, como a encontrada no estágio de vigília. Portanto, o sono REM é frequentemente chamado de *sono desincronizado*, porque não há sincronia no disparo dos neurônios, apesar de atividade cerebral significativa. Por este mesmo motivo, também é sugerido o termo *sono paradoxal (SP)* para este estágio do sono.

Assim, a classificação descrita por (Rechtschaffen e Kales 1968) pode ser resumida na Tabela 2.1.

Tabela 2.1: Tabela resumida das características dos estágios do sono, segundo classificação de (Rechtschaffen e Kales 1968)

Estágio	Ondas EEG	Frequência EEG (Hz)	EOG	EMG
Vigília ou 0	α	8-12	muito ativo	muito ativo
1	ritmos α, θ + pontos em vértice	3-7	movimentos oscilatórios, lentos e amplos	um pouco menos ativo
2	ritmo θ + complexos K + fusos do sono	2-8 id. 12-16	id.	id.
3	δ (20 à 50%) ondas muito lentas	1-2 0,5-1	movimentos lentos	tônus mais importante
4	δ (> 50%)	id.	id.	id.
5 ou SP	ritmo α, θ + ondas em <i>dentes de serra</i>	3-8	movimentos bruscos rápidos e amplos	muito fraco (atonia muscular)

A denominação dos estágios foi realizada de maneira lógica, em ordem cronológica de seu aparecimento; assim a progressão (normal) do sono se faz da vigília relaxada para os estágios do sono 1, 2, 3, 4 e REM, antes de retomar a um novo ciclo, a partir do estágio 1.

A classificação (Rechtschaffen e Kales 1968) foi autorizada antes de sua publicação, sendo admitida e utilizada como referência pelo conjunto de autores da pesquisa em matéria de sono até o fim dos anos 2007, data da publicação de um novo manual de codificação do sono emitido pela AASM (Iber et al. 2007b). As diferenças entre esses dois manuais são, de fato, mínimas; sendo que a segunda é considerada como uma atualização da primeira, segundo os neurologistas atuais e os pesquisadores do sono. Nota-se, entretanto, que a nova classificação não discrimina mais os estágios 3 e 4, unindo-os em um mesmo estágio denominado *N3*. Além disso, tal classificação faz uma mudança de nomenclatura: a vigília é denominada *estágio E*, os estágios 1 e 2 tornam-se, respectivamente, *N1* e *N2* e o estágio SP é chamado de *R*, aproximando-se, assim, da nomenclatura anglo-saxã⁷ ainda utilizada na literatura. São recomendadas também três derivações do EEG (F4-M1, C4-M1 e O2-M1), cujos eletrodos são posicionados nas regiões frontal, central e occipital (Moser 2009). A Figura 2.2 mostra as principais diferenças entre o manual de (Rechtschaffen e Kales 1968) e o da (Iber et al. 2007b), no que tange à arquitetura do sono.

Por fim, certos ritmos cerebrais que compõem o EEG não são mais utilizados como critérios de classificação por (Rechtschaffen e Kales 1968), entretanto, são produzidos em diferentes fases do sono e são identificados pela eletroencefalografia. É o caso dos ritmos β_1 e β_2 , que são ondas muito rápidas com banda de frequência compreendida entre 14 e 18 Hz para β_1 e entre 18 e 30 Hz para β_2 , ocorrendo durante o sono lento (Merica e Fortune 2005) e na ativação do SP (Brovelli et al. 2004). As ondas σ , que são fusos de 12 a 16 Hz, são observadas nos estágios 2 e 3. No entanto, a localização destes ritmos é difícil devido à sua origem difusa, o que impede seu uso na codificação dos sinais da polissonografia.

2.2.2 Microarquitetura do Sono

Os grafoelementos permitem aplicar as regras de codificação dos estágios do sono, segundo uma ou outra dentre as duas classificações; são eventos grafológicos pontuais,

⁷Esta nomenclatura substitui *N1*, *N2* e *N3* por *Non-REM Sleep* (estágios 1, 2 e 3), e *R* por *REM Sleep*.

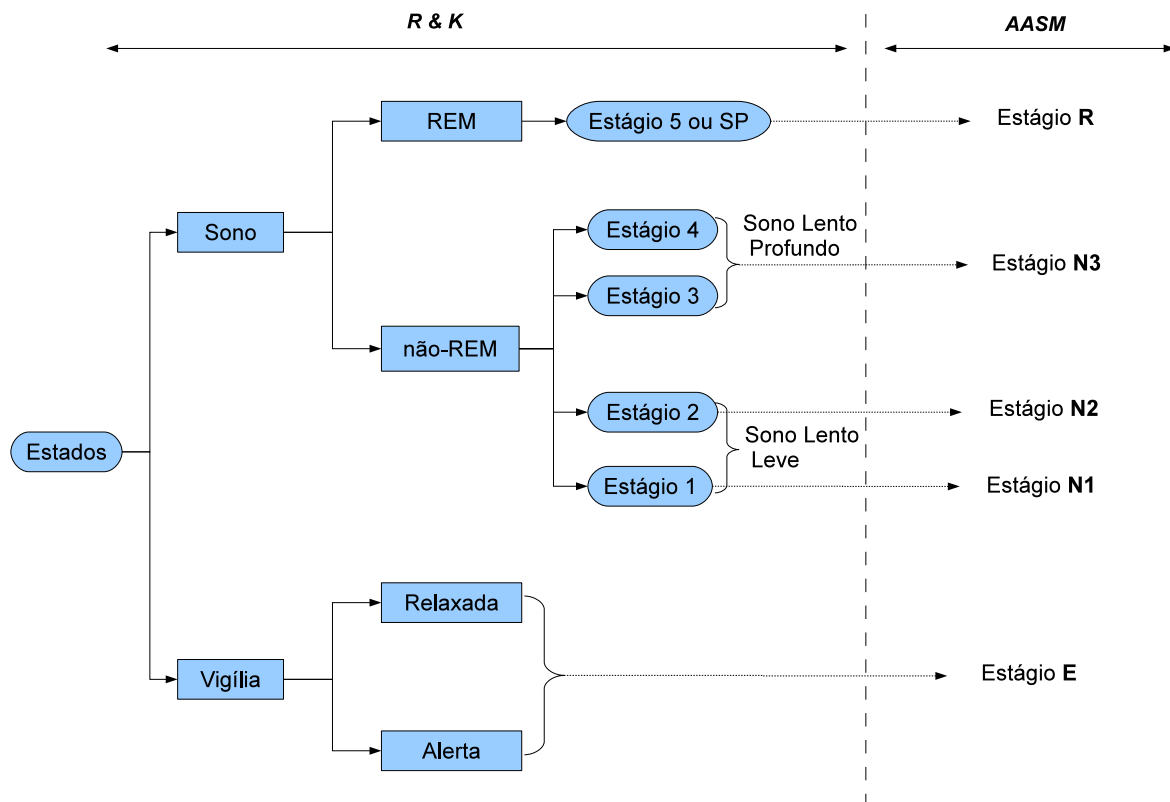


Figura 2.2: Comparação entre os manuais de 1968 (R&K) e o de 2007 (AASM).

que ocorrem ao longo do ritmo ou do EEG de fundo e fazem parte integrante da microarquitetura do sono, bem como os *microdespertares* (MD) ou os *padrões cíclicos alternantes* (*cyclic alternating pattern, CAP*) (Dauvilliers e Billiard 2004; Terzano e Parrino 2000).

Eventos Fásicos

Os primeiros grafoelementos (GE) típicos do sono encontrados durante a sonolência são os pontos em vértice, nome dado às ondas bifásicas de grande amplitude ($>200 \mu\text{V}$), pontuais e simétricas que marcam a entrada no estágio 1 do sono; em seguida, a entrada no estágio 2 se traduz pelo surgimento dos fusos do sono, caracterizados por manifestações repentinas de ondas rápidas de baixas amplitudes, com faixa de frequência de 12 a 16 Hz, ocorrendo em intervalos de tempo de 1 a 2 s. Durante o estágio 2, os fusos do

sono são mais abundantes e podem aparecer em 3 a 6 s (Nicolat 2001). Ainda neste estágio, há o surgimento dos complexos K, que são ondas de altas frequências θ , com duração de 2 a 3 s, apresentando uma componente negativa de grande amplitude, imediatamente seguida de uma componente positiva. Os *pontos occipitais* do sono são pequenas ondas bifásicas provenientes de regiões posteriores e, frequentemente, observadas durante o estágio SP.

Os GE's são indicadores indispensáveis à classificação do sono por meio dos estágios, mas descrevem uma estrutura subjacente que torna-se interessante à observação sob uma escala mais fina.

Microdespertares

Em 1992, a Associação Americana de Distúrbios do Sono (*American Sleep Disorders Association, ASDA*) propôs uma definição de *microdespertares* (MD) (Association 1992), independente de qualquer regra de codificação. Os MD's ocorrem em qualquer estágio do sono e são caracterizados por uma mudança brusca de frequência, com duração de 3 a 15 s, com os ritmos do EEG de fundo, estando, então, representados. O tempo máximo de 15s não é explicitamente mencionado no manual da ASDA, mas é necessário, por si só, pois uma codificação do tipo microdespertar superior a 15 s tomaria mais da metade da época, que seria, então, classificada como vigília. Os MD's não caracterizam propriamente o indivíduo em estágio de sono, mas representam uma atividade do EEG modificada por influências externas como, por exemplo, um ruído ou a luminosidade. Para que o indivíduo tenha consciência de ter sido despertado durante a noite, a duração da alerta intra-sono deve ser superior a 3 min, ou seja, a seis épocas codificadas sucessivamente em estágio de vigília. Certas patologias cardio-respiratórias que induzem apnéia, podem dar origem aos MD's. A regra de codificação para os MD's permanece inalterada no manual de 2007, contudo um MD especificado durante o estágio SP deve ser obrigatoriamente acompanhado de um aumento de amplitude do EMG mentoniano durante, pelo menos, um segundo (Paquereau 2007). Finalmente, dois MD's sucessivos devem ser, pelo menos, separados por um período de 10 s de

recuperação dos ritmos do EEG de fundo. Esses eventos são, contudo, confusos nas transições de um estágio de sono para outro.

Os MD's não são eventos raros e um indivíduo sadio em estágio de sono pode apresentar até 10 MD's por hora de sono. O número de MD's por hora é um indicador do grau de *fragmentação do sono* que, quando é elevado, impede a progressão do sono para os estágios mais profundos e, portanto, altera consideravelmente a qualidade do sono. Esse número aumenta com a idade e representa um sinal fisiológico de maior fragmentação e de maior fragilidade do sono em pessoas idosas (Boselli et al. 1998; Carskadon, Brown e Dement 1982). Entretanto, é razoável saber que mesmo em SP, fase durante a qual ocorre o *sonho*, o cérebro pode, por vezes, estar em um estado fisiológico que parece completamente dissociado da realidade e do presente e, assim, manter sua vigilância para reagir a estímulos externos ou a eventos de apnéia.

Traçados Cíclicos Alternantes

Há vários anos, as equipes de pesquisa vêm se interessando pela estrutura do sono, levantando o problema, exclusivamente por meio de uma análise em estágios (*estagiamento do sono*), ignorando o aspecto dinâmico da progressão do sono e subestimando a importância dos eventos de curta duração, utilizando-os unicamente como marcadores de estágios. Assim, vários estudos se baseiam na análise dos *Traçados (ou Padrões) Cíclicos Alternantes (TCA) – Cyclic Alternating Pattern (CAP)* – para estudar a estrutura do sono: o CAP é identificado por uma *sequência de padrões periódica* do EEG, repetindo-se em períodos de menos de 60 s, geralmente em períodos de 40 s.

O CAP constitui-se de duas fases: a *fase A* ou *fase de ativação do CAP*, acompanhada de um aumento do tônus muscular e de uma ativação do sistema nervoso vegetativo (ou SNA), com duração de 10 a 12 s; e uma *fase B* ou *fase de desativação do CAP*, acompanhada de uma redução do tônus muscular e de atividades neurovegetativas, com duração de 20 a 30 s. Entre os CAP's encontra-se uma região chamada *não-CAP* caracterizada por estabilidade tônica e das funções neurovegetativas.

Um CAP começa, necessariamente, por uma *fase A* e termina em uma *fase B*. A defi-

nição de fase A está intimamente ligada à classificação em estágios, pois corresponde:

- aos ritmos α intermitentes e às sequências de pontos em vértice ao longo do estágio 1;
- às sequências de dois ou mais complexos K, seguidos ou não por ritmos β e/ou α ao longo do estágio 2;
- às manifestações repentinas de ondas δ , cujas amplitudes são, pelo menos, três vezes maior que a amplitude do EEG de fundo, ao longo dos estágios 3 e 4;
- e às fases de ativação transitória do EEG ou dos MD's, qualquer que seja o estágio considerado.

O CAP faz parte integrante do sono fisiológico (Terzano 1982), entretanto, inicialmente, foi definido como uma assinatura da insônia. O CAP foi associado à sonolência e ao sono lento, em correlação com as variações cíclicas das frequências cardíacas e respiratórias ou da pressão sanguínea e às ocorrências periódicas dos complexos K ou a outros eventos fásicos ligados à vigília (Lugaresi 1972). Estudos recentes mostraram, entretanto, que o CAP também está ligado às fases de sono profundo, tanto em indivíduos saudáveis, quanto em indivíduos que apresentam alguma patologia, mesmo que o critério do número de MD's seja instável (Terzano 1985). Por fim, parece que a ocorrência do CAP é robusta, mesmo em experiências em pacientes em coma prolongado ou menos severo: as variações cíclicas do EEG (ondas lentas de alta amplitude durante a fase A alternantes com atividades θ e δ de baixa amplitude durante a fase B), como funções fisiológicas, são identificadas (Fishgold e Mathis 1959); neste caso, podem ser espontâneas ou provocadas por estímulos sensoriais externos de diferentes modalidades (táctis, térmicos, acústicos, de dor, dentre outros). Independentemente de outros elementos do EEG, o CAP representa uma oscilação entre dois níveis de vigília, um nível alto, durante a fase A⁸ e um nível baixo, durante a fase B. A instabilidade da vigília ligada ao CAP está associada à uma ativação de outros parâmetros da polissonografia

⁸De fato, durante a fase A, por definição, entre 2 e 60 s, poderia haver um ou mais MD's, ou seja, uma época completa, classificada como vigília.

(atividade muscular) ou a funções fisiológicas (cardiorrespiratórias) durante a fase A e, sua atenuação durante a fase B. A ausência de CAP, definida como período não-CAP, gera uma instabilidade do nível de alerta, e se traduz por uma estabilidade do tônus muscular e das funções neurovegetativas. O aumento da frequência de CAP se traduz por um aumento da quantidade de sono lento e instável. As causas são múltiplas, frequentemente secundárias a sintomas ansioso-depressivos, a distúrbios respiratórios durante o sono, a movimentos periódicos dos membros inferiores, dentre outros. A proporção de CAP aumenta com a idade e diminui após a privação de sono. Assim, os CAP's correspondem à expressão polissonográfica de flutuações do SNA em resposta a diferentes estímulos endógenos ou exógenos (Dauvilliers e Billiard 2004).

2.3 Organização do Sono

2.3.1 O Hipnograma

O desenvolvimento temporal de um sono normal pode ser resumido em um *hipnograma*. Em um indivíduo sadio, o hipnograma mostra uma organização cíclica, em que cada ciclo é, geralmente, constituído por uma fase de sono lento, a partir do sono lento leve (estágios 1 e 2) ao sono lento profundo (estágios 3 e 4) e uma fase de SP (Figura 5.9). Assim, o hipnograma representa uma codificação da macro-arquitetura do sono. A vantagem desse método reside na possibilidade de se comparar os resultados de um estudo com outro, uma vez que esta codificação é padronizada.

Assim, ocorrem ao longo de uma noite de sono normal, 4 a 5 ciclos cuja duração decresce progressivamente durante a noite (tipicamente em torno de 90 min, para o primeiro e 60 min, para o segundo). A duração não é a única característica que varia durante a noite. Assim, no hipnograma apresentado na Figura 2.3, a dinâmica dos ciclos varia igualmente: os dois primeiros ciclos uma descendência progressiva da vigília até o SP; cada estágio foi realizado, neste caso. Os dois ciclos seguintes revelam transições (abruptas) entre os estágios 1 ou 2 e o SP: o sono lento profundo não é mais uma rota necessária para o SP. Finalmente, o último ciclo apresenta uma nova descendência

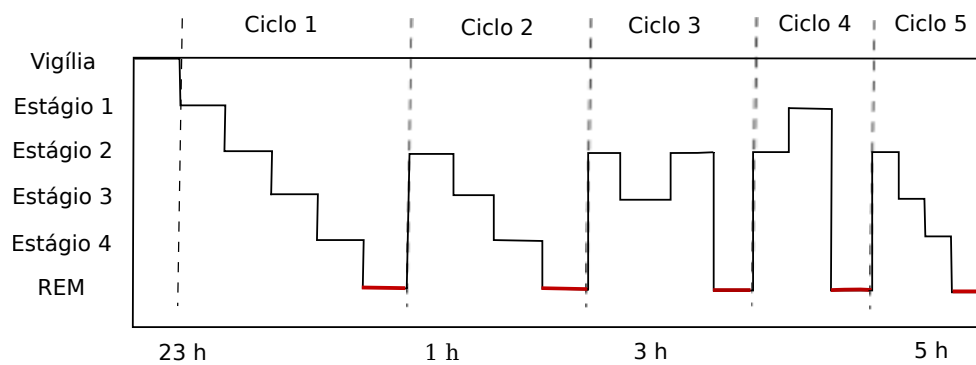


Figura 2.3: Hipnograma

progressiva via sono lento profundo. A supressão dos estágios 3 e 4 durante o ciclo, conduz à fases de SP mais longas (40-45 min para os ciclos 3 e 4 contra 25 min para os dois primeiros ciclos), apesar de uma redução da duração total do ciclo (70 min para o ciclo 4, por exemplo). Essas características são bastante gerais em indivíduos saudáveis. Tais características podem ser resumidas da seguinte maneira: um sono lento profundo relevante no início da noite e um SP prolongado no fim da noite.

2.3.2 Os parâmetros e os índices de análise do sono

Para avaliar a qualidade e a eficiência do sono, vários parâmetros de análise servem como referência. A maioria dos indicadores da qualidade do sono, utilizados na pesquisa clínica, são marcadores temporais, calculados a partir da interpretação das polissonografias. Os principais parâmetros são (Rabarimanantsoa-Jamous 2008):

- *A latência de sonolência (adormecimento)*: corresponde ao tempo decorrido entre a hora de se deitar e a ocorrência do primeiro estágio do sono. Num indivíduo saudável, uma latência de sonolência de mais de 30 min começa a ser considerada como patológica;
- *O período de sono total (PST)*: é a duração que separa o início do primeiro estágio do sono e a fim do último; esta duração varia de 5 a 10 h em um indivíduo saudável.
- *O número e a duração de despertares intrassono*: correspondem aos despertares que fazem parte da macroarquitetura, ou seja, ao menos uma época de 30 s classificada

como vigília. Os despertares intrassono são, também comuns, em indivíduos saudáveis, mas é difícil avaliar (estimar) um valor crítico a partir do qual o número de despertares começa a ser considerado como patológico;

- *O tempo de sono total (TST)*: corresponde ao período de sono total marcado pelos microdespertares intra-sono.

A partir dos parâmetros definidos, anteriormente, diferentes índices são calculados para que cada indivíduo seja comparado um com o outro, a fim de avaliar o sono de um indivíduo em relação a um outro, mesmo em comparação com valores normalizados. Os principais índices utilizados são (Rabarimanantsoa-Jamous 2008):

- *O índice de eficiência do sono*:

$$\frac{D_{SLP+SP}}{TST} = \frac{D_3 + D_4 + D_5}{TST}$$

em que D_{SLP+SP} é a duração do Sono Lento Profundo adicionada à duração do Sono Paradoxal, $D_{3,4,5}$ representam a duração dos estágios 3, 4 e 5 (SP) e TST é o Tempo de Sono Total, ou seja, a duração dos cinco estágios.

- *O índice de qualidade do sono*:

$$\frac{TST}{PST}$$

em que PST é o período de sono total. O índice de qualidade do sono representa, então, o tempo *realmente* ocupado pelo sono, e está compreendido entre 0 e 1. De fato:

$$\frac{TST}{PST} = \frac{PST - D_{despertares}}{PST} = 1 - \frac{D_{despertares}}{PST}$$

em que $D_{despertares}$ é a duração dos despertares intrassono. Se esta duração for nula, não houve despertares e o índice de qualidade do sono é igual a 1; caso contrário, o índice diminui quando o $D_{despertares}$ aumenta.

- *O número de MD's por hora:* é um indicador da fragmentação do sono. Quanto maior o número de MD's/h, mais fragmentado será o sono. O número de MD's/h de um indivíduo sadio está em torno de 10.
- *O número de mudanças de estágios:* às vezes relevante; este índice não é utilizado com frequência e deve ser discutido quase que *caso a caso*.
- *O número de ciclos:* sabendo-se que, às vezes, é difícil descrever um ciclo, principalmente quando a sequência dos estágios não é convencional, o que acontece frequentemente no caso de um distúrbio do sono. Então, o número de ciclos ao longo de uma noite de sono normal varia de 4 a 6, conforme o indivíduo.

2.4 Ciclo Sono-Vigília

A alternância atividade-reposo é um fenômeno que se encontra em conjunto no mundo dos seres vivos, animais ou vegetais, bem como em numerosos sistemas físicos: o estado passivo seguido do ativo, o sistema se tensiona, depois se relaxa etc. O repouso parece ser, então, a uma componente dinâmica natural.

A vigília é indispensável à sobrevivência; corresponde ao momento no qual nos satisfazemos de nossas necessidades imediatas, em que nos alimentamos, aprendemos. Essa atividade é traduzida por um comportamento motor e uma atividade cerebral intensa, acompanhada de um alto gasto de energia. A qualidade da vigília é tão importante para a sobrevivência das espécies que é muito provável que numerosos sistemas de regulação, tenham sido adquiridos durante a evolução: principalmente, três vias corticais situadas no tronco cerebral, que representariam um papel primordial na regulação da vigília por intermédio de um neurotransmissor – a acetilcolina – considerada como *mediador da vigília* (Guyton e Hall 2006). Esse neurotransmissor é secretado quando informações sensoriais transmitidas às células reticulares indicando a percepção do dia no mundo exterior. Esse neurotransmissor tem uma ação estimulante sobre a atividade cortical, facilitando a atenção, a execução de um gesto, da aprendizagem. Entretanto, existem outros neurotransmissores, secretados em outras estruturas, que

regulam a vigília, como a noradrenalina, que intervém em reações de vigília ligada ao estresse, a dopamina, que facilita a inibição dos movimentos, a serotonina, dentre outros (Dauvilliers e Billiard 2004; Guyton e Hall 2006; Shneerson 2005). De fato, é possível aumentar a qualidade da vigília por meio de drogas, do café, do chá, que têm uma ação estimulante no sistema de vigilância, seja uma ação psicológica, seja uma estimulação das funções mentais, como é o caso da cafeína. Por outro lado, a duração da vigília não pode ser aumentada por meio de drogas sem resultar num fenômeno de adaptação do organismo a uma substância tóxica, de vício (dependência de uma droga) ou de toxicidade. Enfim, uma vigília prolongada, resultará em um débito de sono subsequente, pois são produzidas, durante a vigília, certas substâncias responsáveis pelo sono.

O sono – ou diminuição reversível natural e periódica da percepção do meio externo com conservação de uma reatividade e das funções vegetativas – se apresenta *a priori* como um repouso do corpo. O corpo fica quase imóvel, as atividades cardíaca e respiratória ficam mais lentas, a temperatura corporal central diminui, a percepção do mundo exterior se atenua. Uma explicação intuitiva fundamentada na noção subjetiva de fadiga explicaria o mecanismo do sono: as atividades diurnas solicitam a ativação das funções cerebrais e musculares que demandam relaxamento. Uma segunda explicação estaria ligada à economia de energia realizada por uma diminuição da temperatura corporal que seria, então, uma variante da hibernação. Além disso, os mecanismos do sono não são considerados como fenômenos passivos e exclusivamente de repouso, mas um modo ativo, complementar da vigília. Assim, as estruturas cerebrais envolvidas no sono não são aqueles da vigília, mas correspondem ao início de um processo ativo diferente.

O sono lento e o SP são o resultado do funcionamento de redes neuronais diferentes, mas que os associa a uma rede executiva responsável pelo sono e uma rede permissiva, que controla o surto de sono. Por outro lado, a vigília é o resultado da atividade de numerosas estruturas redundantes (Guyton e Hall 2006). Entretanto, ainda não há uma explicação para a operação cíclica que ocorre entre o sono e a vigília. Sabe-se, apenas, que o surgimento do sono, ou da sonolência, depende de várias condições que podem

ocorrer em circunstâncias diversas. Sem considerar as doenças como a narcolepsia, a sonolência ou a sonolência diurna podem ocorrer em um indivíduo saudável, indicando a existência de um limite de vigília voluntário. Entretanto, em casos extremos ou patológicos, muitos acidentes rodoviários são causados por distúrbios de vigília diurna dos condutores. Por outro lado, o contrário também pode acontecer, por exemplo, um indivíduo sadio pode ter dificuldade de adormecer ou de dormir, mesmo que todas as condições exteriores sejam favoráveis ao sono. A insônia atinge mais de 30% da população entre 18 a 79 anos, se agravando com a idade. Pode ser transitória, mas torna-se crônica, portanto, patológica em mais de 10% da população. Assim, a transição do estágio de vigília ao estágio de sono e, vice-versa, não é completamente voluntária, nem sistemática, mas resulta da superposição de vários processos de regulação, destacando-se, dentre os quais, três processos: homeostático, circadiano e ultradiano.

2.4.1 Processos de Regulação dos Estágios de Sono e de Vigília

Processo Homeostático

O *processo homeostático* ou *processo S* se traduz por um comportamento acumulativo: a necessidade de dormir aumenta durante a vigília e diminui durante o sono. Isso resulta, em particular, pela oscilação da concentração de adenosina, produto de degradação do metabolismo das células cerebrais, consumindo ATP⁹ durante a vigília. Assim, o acúmulo de adenosina nos receptores específicos induzem uma necessidade crescente de sono. Essa teoria foi validada pela injeção de adenosina no tronco cerebral de animais adormecidos, provocando uma sonolência quase imediata, enquanto que o consumo de cafeína, substância antagonista dos receptores de adenosina, induz uma vigília prolongada.

Um outro aspecto desse processo é a atividade cerebral – em ondas θ – de vigília, correlacionada com a atividade em ondas lentas – tipicamente δ – durante o sono (Tabela 2.1). De fato, foi mostrado que uma vigília prolongada resultando em uma tendência crescente da atividade θ da vigília teria, como consequência, um aumento da atividade

⁹Tri-Fosfato de Adenosina.

em ondas lentas durante o sono seguinte. Por outro lado, uma redução de duração da vigília é acompanhada por uma diminuição, em fase, da atividade noturna seguinte, em ondas lentas. Assim, uma hipótese possível seria aquela, segundo a qual o sono lento profundo (aquele no qual se observa ondas lentas) teria uma função motora reparadora. Além disso, essa função reparadora responderia às necessidades locais: fisiologicamente, essa atividade de ondas lentas é mais importante em regiões frontais durante o primeiro ciclo do sono. O lobo frontal, que está envolvido em processos de atenção, logo particularmente ativo e solicitado em vigília, súbito uma atividade em ondas lentas notavelmente relevantes durante o primeiro ciclo do sono seguinte.

As experiências de estimulação preferencial de certas regiões cerebrais durante a vigília resultam em um aumento da atividade de ondas lentas *ad hoc* na primeira hora do sono, permitindo confirmar a hipótese de regulação regional da atividade das ondas lentas durante o sono (Dauvilliers e Billiard 2004).

Processo Circadiano

O *processo circadiano*¹⁰ regula o que é comumente chamado de *relógio biológico*. A maioria dos numerosos organismos vivos, dos mais simples aos mais complexos (organismos unicelulares, plantas, animais etc), apresentam ritmos biológicos intrínsecos, próximos de 24 h, que se traduzem em uma alternância relativamente periódica de todas as atividades fisiológicas. Assim, a alternância sono-vigília, a temperatura corporal central, a absorção da alimentação, as secreções hormonais (melatonina, cortisol etc) ou a expressão de certos genes são alguns dos marcadores dessa ritmicidade. Nos mamíferos, a quase totalidade desses ritmos é gerada nos núcleos supraquiasmáticos, que contêm todos os elementos necessários à geração de um relógio biológico, com uma periodicidade endógena. Entretanto, sinais externos permitem a sincronização das atividades de indivíduos de uma mesma espécie. Esses sincronizadores externos são, principalmente, a luz visível e a atividade física. Tais sincronizadores permitem aos núcleos supra-quiasmáticos, cuja periodicidade endógena é ligeiramente superior

¹⁰Do latim, *circa* = em torno de e *dies* = dia, ou seja, período entre 20 h e 28 h, *ultradiano* representa o período inferior a 28 h e *infradiano*, representa um período superior a 28 h.

a 24 h, de se (re)sincronizar sob condições ambientais. Nos mamíferos, o principal sincronizador é a luz percebida pelas células da retina, que informa ao organismo sobre a duração do dia e da noite, variando durante os anos segundo as estações, por intermédio da taxa de secreção de melatonina, cujo máximo, no homem, se situa entre 2 h e 5 h da manhã. Estudos realizados em voluntários sadios, isolados em cavernas (grutas) ou abrigos por vários meses, mostraram que existe uma sincronização interna progressiva entre as variações da temperatura corporal, que fica estável, e a ritmicidade da alternância sono-vigília, que se modifica nas primeiras horas do dia (após um período de 24,5 h), que pode adotar um período muito maior, a partir de 36 a 60 h, com proporção de 2/3 de vigília para 1/3 de sono. Esses resultados são favoráveis à hipótese da existência de um relógio biológico dual, comportando um *oscilador forte*, pouco dependente do ambiente, que dá origem às variações dos ritmos da temperatura central, da secreção de melatonina e de cortisol, do SP, e um *oscilador fraco*, mais sensível às condições exteriores, que origina a sincronização dos ritmos de sono-vigília (Dauvilliers e Billiard 2004).

Processo Ultradiano

O *processo ultradiano* controla a alternância (o balanço) sono lento-SP. Essa regulação se faz por intermédio de duas populações de neurônios distintos, *sommeil paradoxal on* e *sommeil paradoxal off*, cuja alternância de atividade provoca a alternância desses dois estágios do sono (Dauvilliers e Billiard 2004).

2.5 Funções Fisiológicas do Sono

As funções do sono são ainda pouco conhecidas, mas certamente, os estágios REM, não-REM e o estágio de vigília refletem diferentes funções. O sono é, em muitos aspectos, um estado vulnerável, devido à diminuição da consciência e da capacidade de resposta ao meio ambiente. O sono exerce papel importante em diversos sistemas do organismo humano que está diretamente ligado aos estágios do sono e à qualidade do sono. Cada um desses estágios implica um tipo de funcionamento cerebral particular,

que terá influência sobre os diferentes controles exercidos pelo sistema nervoso central (SNC), principalmente sobre as grandes funções vitais, dentre as quais: a função cardiovascular, ventilatória, endócrina, gastrointestinal, gênito-urinária, imunológica, termorreguladora, motora, autonômica e mental. As funções do sono se dividem em diversas categorias: bioquímicas, fisiológicas, neurológicas, psicológicas e sociais. A seguir, serão descritos alguns mecanismos de controle das funções vegetativas ligadas ao sono, com ênfase na função cardiovascular, que é a de interesse neste estudo. Para maiores detalhes sobre as funções fisiológicas do sono, o leitor interessado poderá consultar as principais referências: (Dauvilliers e Billiard 2004; Shneerson 2005). Uma visão geral sobre a fisiologia do sono é dada em (Guyton e Hall 2006) e detalhes sobre a anátomo-fisiologia dos componentes do sistema nervoso central (SNC) e do sistema nervoso autônomo (SNA) são apresentados em (Kandel, Schwarts e Jessel 2000; Guyton e Hall 2006; Aires 1999).

2.5.1 O Metabolismo Cerebral

A Função Autonômica

Durante o sono há um relativo aumento da atividade parassimpática em relação à simpática. O aumento na atividade simpática é também devido a fatores circadianos durante a noite. Por outro lado, o sistema simpático sofre mais influência dos estágios de sono-vigília do que dos ritmos circadianos. A atividade parassimpática aumenta do estágio 1 para o estágio 4, ao contrário do que acontece no estágio REM. Dentro dos limites de cada estágio, o balanço simpato-vagal permanece estável.

2.5.2 A Função Cardiovascular

A diminuição progressiva da atividade do sistema nervoso autônomo (SNA) durante o sono gera múltiplas variações de parâmetros cardiovasculares ao longo de uma noite. A existência de uma ritmicidade periódica, durante o sono, da frequência cardíaca, da pressão arterial e da respiração, bem como da pressão intracraniana, é bem

estabelecida, com ciclos próximos de 15 a 20 s e isto está relacionado com a provável presença de um gerador central comum (Dauvilliers e Billiard 2004).

Variações da Frequência Cardíaca

Mesmo antes do início do sono não-REM, a diminuição da frequência cardíaca, em paralelo com a redução da temperatura corporal, está relacionada à queda na taxa metabólica. A vasodilatação é devida à redução da atividade simpática e, possivelmente, ao acúmulo de adenosina (Shneerson 2005).

O sistema cardiovascular é mais estável durante o sono não-REM do que durante o sono REM. Durante o sono REM, a frequência cardíaca, embora particularmente instável (por causa das atividades fásicas), é mais elevada do que em sono lento, e isto está relacionado a um tônus simpático próximo do estágio de vigília relaxada. Essa forte atividade simpática durante o sono REM pode explicar a ocorrência de distúrbios do ritmo (principalmente as extra-sístoles ventriculares) e da condução cardíaca (como os bloqueios átrio-ventriculares e as pausas sinusais) em indivíduos livres de qualquer patologia cardíaca (Dauvilliers e Billiard 2004; Shneerson 2005).

Uma depressão sinusal de intensidade moderada surge a partir da sonolência e é responsável por uma bradicardia com a possibilidade do aparecimento de pausas sinusais (< 2,5 s). Uma depressão nodal é igualmente encontrada com o surgimento de bloqueios átrio-ventriculares de primeiro e, mais raramente, de segundo grau. Essas diminuições de atividades sinusais e nodais são, às vezes, explicadas por uma ativação do tônus parassimpático e por uma diminuição do tônus simpático durante o sono lento e isto é mais pronunciado durante o sono profundo, dando origem a importantes bradicardias (Dauvilliers e Billiard 2004).

Variações da Pressão Arterial

A pressão arterial não está sob controle circadiano, mas é mais dependente do estágio de sono. A pressão arterial sistêmica diminui durante o sono lento, principalmente durante o sono lento profundo (diminuição de 10 – 15% em relação à vigília) e isto

está relacionado à vasodilatação dos vasos periféricos. Essa queda da pressão arterial é ausente se houver aumento da atividade simpática, como nos casos de hipertensão ou apnéias obstrutivas do sono e na insuficiência renal crônica.

Há pouca diferença tanto da pressão sistólica quanto da pressão diastólica, durante os estágios não-REM. Durante o sono REM, a pressão arterial, embora instável, não parece ser diferente daquela que ocorre durante a vigília relaxada. Ao despertar, a pressão arterial aumenta rapidamente, independente do estágio do sono que preceda a vigília. O sono não-REM funciona como um sistema cardioprotetor, no qual o fluxo sanguíneo coronariano permanece razoavelmente constante durante o sono, por outro lado, há redução na pressão de perfusão que pode levar à isquemia coronariana, em casos de patologias arteriais.

Deve-se notar que, em indivíduos sadios não há uma mudança notável da pressão arterial pulmonar durante o sono (Dauvilliers e Billiard 2004; Shneerson 2005).

Variações do Débito Cardíaco e Cerebral

O débito cardíaco parece diminuir durante o sono lento, embora este dado ainda seja controverso (Dauvilliers e Billiard 2004). As principais modificações da pressão arterial durante o sono são explicadas pela diminuição da resistência vascular, sem necessidade de diminuição do débito cardíaco.

Durante o sono REM, o débito cardíaco parece ser estável. O aumento da pressão arterial devido às atividades fásicas, pode ser explicado pelo aumento da resistência vascular.

O fluxo sanguíneo cerebral aumenta quando há um aumento das atividades fásicas do sono REM, principalmente nas estruturas do tronco cerebral. O fluxo sanguíneo cerebral aumenta se há um desenvolvimento de hipercapnia, pois isso provoca vasodilatação cerebral e um aumento do débito cardíaco. Também há um aumento desse fluxo durante a vigília na presença de hipóxia, entretanto, esta resposta é ausente nos estágios 3 e 4 do sono não-REM. Como resultado, há um potencial para redução no suprimento de oxigênio ao cérebro, em situações de hipóxia, tais como durante as

apnéias obstrutiva e central do sono. Isso pode influenciar a função medular e, assim, modificar as vias aéreas superiores e o controle dos músculos da parede do tórax. O fluxo sanguíneo cutâneo aumenta devido à vasodilatação, mas provavelmente há uma pequena mudança na distribuição do fluxo sanguíneo visceral durante o estágio não-REM (Shneerson 2005).

2.6 Distúrbios do Sono e Patologias Ligadas ao Sono

Há muitos distúrbios do sono, por si só, sendo necessário distinguir os distúrbios do sono relacionados ao meio ambiente (vizinhos barulhentos, temperatura ambiente inadequada, cama inadequada, dentre outros), ou ocasionados por uma outra doença infecciosa (infecções bacterianas, virais ou parasitárias), respiratória (apnéias e hipoventilação, asma e bronquites crônicas, apnéias obstrutivas (*Obstructive sleep apnoeas*, OSA, dentre outras), cardiovascular (disritmias cardíacas, hipertensão, angina e infarto do miocárdio, insuficiência cardíaca etc), neurológica (esclerose múltipla, traumatismo craniano, epilepsia etc) ou psiquiátrica (distúrbios de humor, distúrbios obsessivos compulsivos, alcoolismo, esquizofrenia etc) (Shneerson 2005). As doenças do sono podem afetar os sistemas de regulação, como é o caso dos distúrbios do ritmo circadiano.

2.7 Fisiologia da Variabilidade da Frequência Cardíaca

2.7.1 O Sinal da Variabilidade da Frequência Cardíaca

A importância do acometimento do sistema nervoso autônomo (SNA) nas doenças cardiovasculares tem sido objeto de intensas investigações (Gomes 2001). A avaliação funcional do controle do SNA sobre o coração pode ser feita por meio de testes autonômicos, nos quais se observa a resposta reflexa fisiológica à aplicação de um estímulo quantificável, fisiológico ou farmacológico, como a respiração, o exercício e a injeção de atropina e fenilefrina. Alternativamente, informações sobre o controle autonômico cardíaco podem ser obtidas pela observação da variação intrínseca da frequência car-

díaca (FC), tanto em registros curtos, de dois a cinco minutos de repouso, como em traçados prolongados, de 24 h, durante as atividades habituais. O sinal da variabilidade da frequência cardíaca (VFC) consiste de uma série dos intervalos de tempo entre as ondas R do eletrocardiograma (ECG) - série de intervalos RR. Pode-se também considerar a série de valores de frequência cardíaca instantânea (tacograma) que nada mais é do que uma série de intervalos RR inversos. A análise da VFC é uma técnica não-invasiva que parte do princípio que, em condições normais, o ritmo cardíaco modifica-se em resposta a estímulos diversos, como exercício e estresse mental, ou mesmo em condições de repouso, flutuando em torno de uma média. Tal variabilidade relaciona-se, predominantemente, às alterações contínuas do balanço simpato-vagal, em resposta a mecanismos de controle cardiovascular (Gomes 2001; Junqueira 1998; Vermeiren 1996). A parte superior da Figura 2.4 mostra uma série de intervalos RR com duração aproximada de 24 h de um adulto normal. Pode-se notar que a FC cardíaca é menor (intervalos RR maiores) durante o sono, período aproximadamente compreendido entre os batimentos 30.000 e 55.000. A parte inferior desta Figura mostra uma seção de 5 min que foi extraída da série de 24 h.

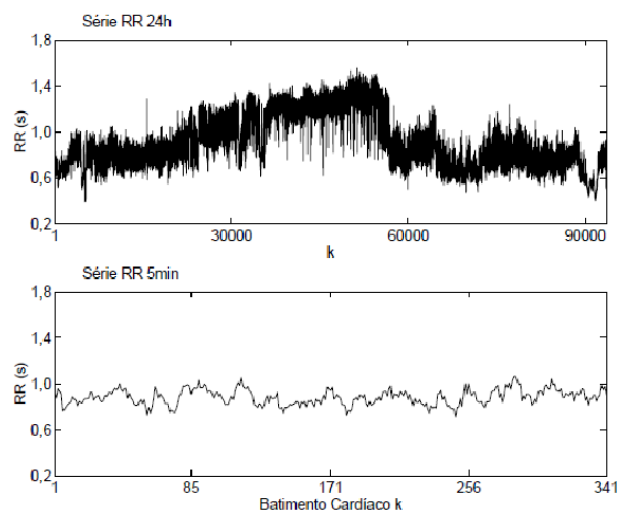


Figura 2.4: (a) Série de intervalos RR de um indivíduo controle. (b) Segmento supostamente estacionário de 5min. Fonte: (Gomes 2001).

A importância clínica da VFC tornou-se aparente no final dos anos 80, quando a

VFC foi considerada um forte e independente preditor de mortalidade após infarto agudo do miocárdio (Task-Force 1996). A VFC tem sido objeto de diversos estudos, e tem se mostrado como uma ferramenta não-invasiva capaz de fornecer diagnósticos úteis (Baselli et al. 1988; Baselli et al. 2001; Baumert et al. 2007; Figueiredo et al. 1996; Guzzetti et al. 1990; Guzzetti et al. 1991; Porta et al. 2007; Ribeiro 1996)

Importantes relações entre o SNA e a mortalidade devida a problemas no sistema cardiovascular, incluindo a morte súbita cardíaca, já são muito conhecidas (Cacoub 2006; Lombardi 2000; Lombardi et al. 2001). Além disso, tem sido verificada a importância da VFC como um indicador quantitativo da atividade autonômica na fisiopatologia cardiovascular (Gomes 2001). Aliada à facilidade de obtenção do sinal de VFC por técnica não-invasiva, verifica-se que valiosas informações sobre os mecanismos de regulação da atividade cardíaca estão imersos na complexa série de intervalos entre batimentos cardíacos (Task-Force 1996; Gomes 2001).

Recentemente, vários estudos vêm se preocupando com a relação do sono com o sistema cardiovascular. Assim, várias ferramentas de análise, principalmente as de análise não-linear, vêm sendo construídas e aplicadas com o intuito de classificar os estágios do sono, via análise de sinais de VFC (Bunde et al. 2000; Mendez et al. 2009; Mendez e Milano 2010). Esses estudos ainda são inconclusivos.

As variações na FC podem ser avaliadas por alguns métodos, sendo os do *domínio do tempo* os mais simples de serem utilizados (Task-Force 1996). Ainda não há um consenso sobre o melhor índice de avaliação da VFC disponível para uso clínico (Gomes 2001). Os métodos de aferição da VFC são aplicados nos domínios do tempo e da frequência. Os principais índices do domínio do tempo são:

- SDNN: desvio padrão de todos os intervalos RR normais (intervalos NN);
- SDANN: desvio padrão das médias dos intervalos RR em todos os segmentos de 5 min de um registro de 24 h;
- RMSSD: raiz quadrada da média da soma dos quadrados das diferenças entre intervalos RR adjacentes;

- pNN50: proporção ou número de intervalos RR que são maiores que 50 ms (para humanos) de um registro de 24h.

No domínio da frequência tem-se a análise da função densidade espectral de potência, DEP, que fornece informações básicas de como a potência se distribui em função da frequência. Métodos para o cálculo da DEP podem ser classificados como *métodos paramétricos* e *não-paramétricos* (Task-Force 1996). A análise espectral pode ser realizada pelo emprego de técnicas clássicas baseadas na Transformada de Fourier ou por meio de técnicas fundamentadas na modelagem paramétrica de séries temporais que se baseiam no ajuste de um modelo auto-regressivo de média móvel (ARMA) ou suas variações à série temporal sob análise (Kay 1987). Os principais índices do domínio da frequência são:

- TotPow: Variância dos intervalos RR;
- LF: Potência na faixa de baixas frequências;
- HF: Potência na faixa de altas frequências;
- LF/HF: Razão LF/HF.

Métodos não-analíticos provenientes da dinâmica não-linear têm possibilitado novas abordagens para o estudo e entendimento das características do comportamento da FC segundo (Huikuri et al. 1999), estes métodos de medida da VFC se diferem dos tradicionais no sentido de que não são destinados a avaliar a magnitude da variabilidade, em vez disso, estimam as propriedades de correlação e complexidade da VFC que não são explicáveis pelas técnicas tradicionais. No entanto, a aplicabilidade dos métodos baseados na dinâmica não-linear ainda precisa ser verificada por meio de mais estudos que permitam estabelecer padrões que possam ser usados na clínica médica.

Alguns métodos não-lineares mais recentes, discutidos no próximo capítulo, foram usados neste trabalho para o estudo da VFC e de sinais polissonográficos.

2.7.2 Controle Neural da Frequência Cardíaca

A regulação da frequência cardíaca (FC), da pressão arterial (PA) e de outras variáveis cardiovasculares é mediada, principalmente, pelo sistema nervoso autônomo (SNA), junto com os mecanismos homeostáticos intrínsecos e extrínsecos. Complexas alterações são observadas nestas variáveis e a análise de sua variabilidade tornou-se uma importante ferramenta não-invasiva para estudar as interações simpato-vagais em diversas condições fisiopatológicas. A análise da variabilidade da frequência cardíaca representa uma ferramenta promissora para tais estudos. Técnicas de análise de sistemas dinâmicos não-lineares têm se mostrado ferramentas promissoras no estudo da variabilidade da frequência cardíaca. O presente trabalho visa a estudar e implementar técnicas de dinâmica não-linear e avaliar seu potencial para análise e, possível classificação de cardiopatias.

Este estudo dará ênfase à disfunção autonômica cardíaca (DAC), que será descrita no Capítulo 5. Para maiores detalhes sobre a cardiodinâmica e sobre as funções fisiológicas da variabilidade cardíaca que é mediada pelo SNA, por centros nervosos periféricos e por um complexo sistema de receptores, o leitor interessado poderá consultar as seguintes referências: (Guyton e Hall 2006; Berne et al. 2004; Aires 1999; Gomes 2001; Oliveira 2007).

Ferramentas de Dinâmica Não-Linear

Aplicadas à Análise e Classificação de

Sinais Fisiológicos

“Une cause très petite, qui nous échape, détermine un effet considérable que nous ne pouvons pas ne pas voir, ... même que les lois naturelles n’auraient plus de secret pour nous, nous ne pourrions connaître la situation initiale qu’approximativement. . . il peut arriver que de petites différences dans le conditions initiales en engendrent de très grandes dans les phénomènes finaux.”

H. Poincaré - 1908¹

3.1 Introdução

A investigação de dinâmica não-linear (DNL) e a introdução de índices para quantificar a complexidade da dinâmica caótica têm contribuído para o desenvolvimento de técnicas para a análise da regulação da variação cardíaca, da pressão sanguínea (PS), da respiração (RES) e de mecanismos ligados ao sistema nervoso autônomo (SNA) como um todo. O conhecimento e o desenvolvimento destas ferramentas vêm trazendo inovações no campo da dinâmica cardiovascular, neural e respiratória.

¹Uma pequena causa, que escapa à nossa atenção, determina um considerável efeito que não podemos deixar de notar ... mesmo que as leis naturais não fossem mais um segredo para nós, poderíamos conhecer somente aproximadamente a situação inicial ... Pode ser que pequenas diferenças nas condições iniciais produzam grandes diferenças nos fenômenos finais.

Durante as últimas décadas, métodos derivados da DNL tiveram sucesso em aplicações em diversos campos do conhecimento, como na física, astrofísica, química, economia, biologia e medicina. Entretanto, o impacto das métricas não-lineares determinísticas para melhor compreensão da fisiologia só em parte foi explorado até a data.

Inicialmente, a teoria da DNL, desenvolvida durante os anos de 1970 e 1980, gerou interesse entre muitos pesquisadores para explorar o comportamento caótico em sistemas biológicos e aplicar os seus achados à biologia e à medicina. Fenômenos não-lineares estão, certamente, envolvidos no sistema biológico humano, devido a complexas interações entre variáveis hemodinâmicas, eletrofisiológicas e humorais. Em particular, a literatura vem mostrando evidências de não-linearidades no sistema cardiovascular, em sinais respiratórios e em sinais de eletroencefalograma (EEG) do sono (Baselli et al. 2002; Cerutti e Signorini 2002; Gautama, Mandic e Van Hulle 2004; Gomes 2001; Kaplan e Cohen 1990; McClintock e Stefanovska 2002; Perc 2005; Souza 2001). Sistemas não-lineares podem ser modelados a partir de representações matemáticas baseadas, por exemplo, em modelos polinomiais, racionais e neurais (modelos NARMAX - *nonlinear autorregressive moving average model with exogenous variables*, por exemplo). Dentre as várias técnicas que vem sendo aplicadas, na literatura, podem ser citadas: imersão, entropia aproximada e de *Shannon*, gráficos de Recorrência, mapas de primeiro retorno, modelos lineares e não-lineares, análise de dados sub-rogados (*surrogate data*), dinâmica simbólica, cálculo de índices como a dimensão de correlação, expoente de *Lyapunov*, *power-law scaling analysis*, análise de flutuações com tendências excluídas (*Detrended Fluctuation Analysis*, DFA) e diversas ferramentas desenvolvidas no estudo do caos determinístico (Baumert et al. 2007; Gomes et al. 2000; Gomes 2001; Kleiger, Stein e Bigger 2005; Oliveira 2002; Oliveira 2007; Oliveira et al. 2002; Oliveira e Aguirre 2006; Oliveira et al. 2008; Porta et al. 2007; Rabarimanantsoa-Jamous 2008; Ribeiro et al. 2002b; Schreiber 1999; Souza-Neto et al. 2004; Sousa et al. 2006; Heikki et al. 2003; Wessel et al. 2006; Wessel et al. 2007b; Wessel et al. 2007c). Em (Eke et al. 2002) e em (Voss e Caminal 2008) há boas revisões sobre técnicas não-lineares aplicadas na caracterização da complexidade de séries temporais de sinais fisiológicos, com ênfase

em sinais de variabilidade da frequência cardíaca (VFC).

Um trabalho pioneiro realizado por (Glass e Mackey 1998) introduziu abordagens não-lineares para a análise do ritmo cardíaco. Bifurcações de duplicação de período foram previstas teoricamente e observadas experimentalmente em células embrionárias de coração de frangos.

A análise da VFC tornou-se uma importante ferramenta clínica, familiar aos cardiologistas (Lombardi 2000). As primeiras abordagens das análises da VFC baseadas em dinâmica não-linear foram realizadas por (Goldberger e West 1987). Mais tarde, os autores de (Goldberger e Greenwald 1988; Lombardi 2000) relataram que pacientes propensos ao risco elevado de morte súbita cardíaca mostraram evidência de dinâmica não-linear na frequência cardíaca (FC), incluindo variações espectrais bruscas nas oscilações de baixa frequência (*low frequency*, LF). Posteriormente, os mesmos autores sugeriram que uma perda de variabilidade fisiológica pode ocorrer sob certas condições patológicas tais como uma redução na dinâmica da FC antes da morte súbita e no envelhecimento (Goldberger 1991).

Os autores de (Babloyantz e Destexhe 1988) realizaram a primeira análise multivariada não-linear da VFC. Com a ajuda de vários métodos independentes para quantificar DNL's, como o retrato de fases, a seção de *Poincaré*, dimensão de correlação, expoente de *Lyapunov* e entropia de *Kolmogorov*, os eletrocardiogramas (ECG's) de quatro pacientes sadios foram estudados qualitativa e quantitativamente. Os autores demonstraram que a variabilidade implícita nos intervalos entre os batimentos cardíacos não é aleatória, mas exhibe pequenas variações governadas por leis determinísticas. Técnicas de reconstrução em espaço de fases (ou espaço de estados) e análises dimensionais foram aplicadas na análise em sinais de FC obtidos em doze pacientes sadios (Chaffin e Reed 1991).

Para estimar a complexidade da dinâmica cardiovascular, em (Pincus 1991) foram modificados os conceitos de dimensão de correlação original e a entropia de *Kolmogorov* (Grassberger e Procaccia 1983), criando o conceito de entropia aproximada (ApEn). Esta técnica foi, posteriormente, modificada e denominada entropia amostral (SampEn) por (Richman e Moorman 2000), reduzindo a polarização do método original.

Em (Novak e Nadeau 1993) foram fornecidas evidências de um acoplamento não-linear entre os sistemas respiratório e cardiovascular. A fim de investigar as interações e acoplamentos entre a FC e a RES e entre a FC e a PS, uma variedade de métodos de DNL tem sido desenvolvida e aplicada (Baumert et al. 2007; Faes, Nollo e Chon 2008; Nollo et al. 2001; Nollo et al. 2005). Os autores de (Peng e Goldberger 1995) aplicaram a DFA para quantificar a estrutura fractal da FC, que foi validada, posteriormente por (Mäkikallio e Huikuri 1999).

Os autores de (Kurths e Wessel 1995) introduziram a dinâmica simbólica (DS) para análise da VFC, sugerindo ser um bom indicador de risco de morte súbita cardíaca, baseado em abordagens multivariadas (Voss e Malik 1998). O método baseado em DS foi desenvolvido por (Porta e Cerutti 2001) para aplicação em séries de FC de curta duração (Guzzetti 2005). A descoberta da natureza multifractal da dinâmica da FC por (Ivanov e Struzik 1999) mostrou que a modulação dos batimentos cardíacos é ainda mais complexa do que se suspeitava anteriormente, necessitando de múltiplas escalas de expoentes para sua caracterização.

Várias tentativas foram feitas para empregar abordagens não-lineares para a modelagem do sistema cardiovascular (Baselli, Porta e Pagani 2006; Oliveira 2002; Oliveira 2007; Oliveira et al. 2002; Oliveira e Aguirre 2006; Oliveira et al. 2008; Gomes et al. 2000). Entretanto, ainda há muito para ser estudado e aperfeiçoado em técnicas de análise da VFC, principalmente, no que concerne às técnicas não-lineares que são pouco utilizadas como auxílio na clínica médica, pois não há uma padronização dos protocolos médicos, uma vez que a avaliação clínica da função autonômica cardíaca tem se restringido a técnicas estocásticas. Muitas questões ainda são controversas, como a investigação de caos determinístico em séries de VFC. O periódico "Chaos" (2009) dedicou um volume inteiro (19) à discussões sobre esta questão.

Em (Rabarimanantsoa-Jamous 2008) foram feitas análises de não-linearidades utilizando ferramentas, como entropia de *Shannon* e gráficos de Recorrência para análise de assincronismos na ventilação não-invasiva. Os autores de (Penzel et al. 2007) fizeram uma análise da dinâmica implícita nos sistemas cardiovascular e respiratório durante um sono normal e um sono patológico, verificando traços de não-linearidades

nas dinâmicas destes sistemas e sincronismo e relações de “causa e efeito” no sistema cardiorespiratório, e em (Souza 2001) foram feitas análises e modelagens não-lineares de sinais respiratórios e cardiovasculares.

Alguns estudos vêm aplicando máquinas de vetores de suporte (do inglês, *support vector machines*, SVM's) e modelos autoregressivos (AR) para classificar os estágios do sono. Muitos trabalhos seguem as seguintes etapas durante o processo de classificação: (i) selecionam características (tais como os coeficientes de um modelo, coeficientes da DFT (*Discrete Fourier Transform*) ou coeficientes das *wavelets*, dentre outras características); (ii) reduzem a dimensão do espaço de características, quando for o caso (usando PCA (*Principal Component Analysis*), por exemplo) e (iii) por fim, treinam um classificador (um SVM, uma RNA (rede neural artificial), dentre outros) (Agarwal e Gotman 2001; Barcaro et al. 2004; Caffarel et al. 2006; Chisci et al. 2010; Crisler et al. 2008; Estévez et al. 2002; Gerla et al. 2009; Kaplan et al. 2001; Malinowska et al. 2009; Moser 2009; Olbrich, Achermann e Meier 2003; Olbrich e Achermann 2005; Pardey et al. 1996; Roberts e Tarassenko 1992b; Roberts e Tarassenko 1992a; Pardey, Roberts e Tarassenko 1996; Sinha 2008; Sitnikova et al. 2009; Uchida, Atsumi e Kojima 1994; Ventouras et al. 2005; Vural e Yildiz 2010; Zoubek et al. 2007). Em termos de pré-processamento, na literatura, alguns autores que separam as bandas do EEG usando um filtro de *Butterworth* de segunda ou quarta ordem, tiram os artefatos visualmente e usam janelas de 4 ou 5 s (supostamente estacionárias para o EEG), em vez de separar as janelas de cada estágio individualmente. Entretanto, ainda há muito para ser estudado e investigado nesta direção a fim de que se possa obter uma técnica de classificação automática dos estágios do sono.

Os mecanismos fisiopatológicos ligados aos sono ainda são pouco conhecidos. Na literatura, diversos trabalhos utilizaram ferramentas não-lineares, como entropia, dimensão fractal, análise com tendências excluídas (DFA) para a análise da duração do sono, da fragmentação do sono, do aparecimento de micro-despertares (*micro-arousals*, MA) durante o sono, de patologias ligadas ao sono, dentre outras variáveis (Bonnet e Arand 2007; Danker-Hopfe et al. 2009; Haba-Rubio, Ibanez e Sforza 2004; Morrel et al. 2000; Smurra et al. 2001; Stepanski 2007; Stradling et al. 1999; Susmakova e Kra-

kovska 2007; Wesensen, Balkin e Belenky 1999). Técnicas de identificação de sistemas têm se mostrado ferramentas promissoras para a análise e modelagem dos diversos sistemas fisiológicos (Olbrich, Achermann e Meier 2003; Pardey, Roberts e Tarassenko 1996; Oliveira 2002; Oliveira 2007; Oliveira et al. 2002; Oliveira e Aguirre 2006; Oliveira et al. 2008). Técnicas de identificação de sistemas e estimação de parâmetros foram, inicialmente, investigadas, neste trabalho para a classificação dos estágios do sono.

Ferramentas automáticas para classificação dos estágios do sono vêm sendo elaboradas desde a década de 90 (Pardey et al. 1996; Roberts e Tarassenko 1992b; Roberts e Tarassenko 1992a; Pardey, Roberts e Tarassenko 1996), mas ainda há muito a ser elucidado e investigado neste campo de pesquisa, uma vez que as técnicas ainda são inconclusivas ou insatisfatórias para a aplicação na clínica médica e muitos resultados ainda são controversos. Técnicas de análise espectral que foram amplamente aplicadas possuem algumas limitações, como a própria questão da estacionariedade que influencia a robustez de muitas ferramentas utilizadas no domínio da frequência.

Assim, uma das primeiras questões colocadas nesta tese a fim de objetivar um caráter inovador e inédito foi: *É possível classificar os estágios do sono por meio de sinais fisiológicos coletados sem a utilização de eletrodos?*

Ainda no contexto de reconhecimento de padrões para a classificação de sinais fisiológicos, surgiu a seguinte questão, no que concerne à análise da VFC: *É possível caracterizar a disfunção autonômica em pacientes chagásicos por meio da classificação de séries simbólicas da VFC?* Logo, os principais problemas colocados nesta tese, se resumem a problemas de classificação de sinais fisiológicos.

Este capítulo enfoca a importância da análise por meio da DNL no estudo de sinais fisiológicos, com ênfase na análise da variabilidade cardiovascular e em sinais de EEG do sono, para explorar as características dinâmicas e estruturais da regulação cardiovascular e neural e sua relevância clínica. A seguir, será apresentada uma introdução a alguns conceitos relativos aos sistemas dinâmicos e que têm relação com o trabalho desenvolvido nesta tese. São apresentadas as idéias básicas da análise de estabilidade e de análise topológica, mais especificamente, de dinâmica simbólica. São também apresentados os conceitos de entropia. Por fim, são apresentados conceitos gerais de

identificação de sistemas e uma breve revisão sobre máquinas de vetores de suporte, que representam os principais classificadores utilizados nesta tese.

3.2 Análise Topológica

Esta seção apresenta diferentes conceitos e ferramentas da teoria de análise topológica, que permitiu uma caracterização fina do comportamento de sistemas dinâmicos, muito empregados na análise de sistemas com comportamento caótico. De maneira mais específica, são apresentados os conceitos básicos referentes à dinâmica simbólica e algumas investigações preliminares acerca do comportamento dinâmico de séries simbólicas de modelos conhecidos, tais como o mapa de primeiro retorno e sinais aleatórios.

3.2.1 Dinâmica Simbólica

Segundo (Bai-Lin 1989), a teoria da dinâmica simbólica fornece a mais rigorosa maneira de se definir o movimento caótico nos sistemas dinâmicos. Trata-se de uma abordagem algébrica introduzida por (Morse e Hedlund 1938).

Considere um conjunto finito A de *símbolos* que se denomina *alfabeto*. Os elementos de A são também chamados de *letras* que podem, tipicamente, ser denotados por a, b, c, \dots , ou por dígitos como $0, 1, 2, \dots$, quando for mais adequado. Expansões decimais, por exemplo, usam o alfabeto $A = \{0, 1, \dots, 9\}$ (Lind e Marcus 1995).

Na prática, a dinâmica simbólica corresponde à descrição de um processo de evolução, isto é, ela descreve o essencial da evolução dos sistemas dinâmicos por meio de suas sequências simbólicas. Uma descrição completa da evolução de uma aplicação exigiria a realização da aplicação para todas as condições iniciais possíveis. Não obstante, é possível realizar uma descrição da evolução discretizada de forma a não se considerarem os valores tomados sobre o conjunto de dados analisados, mas se ocupar de reter somente as características essenciais do comportamento dinâmico, como a periodicidade, por exemplo.

Para mais detalhes sobre o assunto, o leitor interessado poderá recorrer às referências citadas ao longo do texto. Outras sugestões são: os livros dos autores (Lind e Marcus 1995) e (Bai-Lin 1989) dedicados à uma introdução à dinâmica simbólica; uma abordagem matemática sobre o assunto pode ser encontrada em (Morse e Hedlund 1938); a análise topológica por meio da dinâmica simbólica pode ser vista em (Gilmore 1998; Letellier 2000; Godelle e Letellier 2000; Letellier 2008); a partição do espaço de estados e a análise estatística das sequências simbólicas são tratadas em (Godelle e Letellier 2000; Letellier 2008) e algumas aplicações da técnica podem ser encontradas em (Godelle e Letellier 2000; Letellier 2008; Kurths e Wessel 1995; Voss e Malik 1998; Voss e Caminal 2008; Oliveira 2007).

Esta seção apresenta conceitos básicos de dinâmica simbólica, como a partição do espaço de fases e a codificação de órbitas periódicas, a fim de proporcionar a compreensão de sua aplicação neste trabalho.

Partição

A primeira etapa de uma caracterização topológica consiste em particionar um atrator em diferentes zonas com topologias diferentes ou, de maneira equivalente, com dinâmicas diferentes. A partição de um atrator em diferentes regiões é determinada com a ajuda de uma aplicação do mapa de primeiro retorno à uma seção de Poincaré.

Uma possível simplificação para o estudo de sistemas dinâmicos contínuos seria o estudo de um sistema em tempos discretos, de forma que o estado do sistema só é analisado em tempos determinados. A dinâmica simbólica pode ser vista como uma simplificação mais geral do estudo do sistema, em que não só o tempo é discretizado como também o espaço (Lind e Marcus 1995). Sua idéia básica consiste em dividir o espaço de estados do sistema em um número finito de regiões disjuntas, associando um símbolo a cada uma destas regiões. Dessa forma, o estudo do sistema se resume em analisar a variação destes símbolos ao longo do tempo.

Codificação

Como dito anteriormente, uma dinâmica simbólica aplicada à descrição da evolução de uma trajetória em um atrator pode ser construída a partir da seção de Poincaré desse atrator e de uma de suas coordenadas, utilizada para se obter um mapa de primeiro retorno. Os pontos fixos do mapa de primeiro retorno definem a partição do atrator em vários ramos, cujo número pode depender do valor do parâmetro de bifurcação.

A cada ramo deve corresponder um símbolo: toda trajetória é então caracterizada por uma sequência de símbolos que descreve a sucessão de ramos visitados. Etiketando-se cada intervalo monotônico do mapa de primeiro retorno com um símbolo, toda trajetória em um atrator pode ser caracterizada de maneira única por uma sequência infinita de símbolos. Dessa forma, séries temporais diferentes podem apresentar uma mesma sequência de símbolos. De fato, o que pode parecer um inconveniente é na verdade uma vantagem, pois isto permite a elaboração de uma classificação de diferentes sequências numéricas e, portanto, dos comportamentos dinâmicos (Bai-Lin 1989) (Oliveira 2007).

Os ramos crescentes são associados à preservação da ordem das intersecções com a seção de Poincaré. Inversamente, os ramos decrescentes correspondem a uma inversão de suas intersecções. Assim, por convenção, os números pares são associados aos ramos crescentes e os números ímpares, aos ramos decrescentes.

Com a partição da seção de Poincaré e a atribuição dos ramos respectivos, cada intersecção de uma órbita com a seção de Poincaré corresponde a uma sequência bi-infinita, composta, por exemplo de 0 e de 1. Isso se escreve no caso de uma aplicação unimodal:

$$S = \cdots s_{-m} \cdots s_{-1} s_0 \bullet s_1 s_2 \cdots s_n \cdots \quad (3.1)$$

em que s_n é o n ésimo ponto futuro, s_m é o código do m -ésimo ponto do passado e o ponto presente é indicado por s_0 . O ponto \bullet divide a sequência bi-infinita em duas sequências infinitas, ou seja, a sequência passada $\cdots s_{-m} \cdots s_{-1} s_0 \bullet$ e a sequência futura $\bullet s_1 s_2 \cdots$ (Oliveira 2007).

As sequências associadas às órbitas de período p repetem indefinidamente a mesma série. Nesse caso, p símbolos são usados na codificação da órbita periódica. Na seção de Poincaré, uma órbita de período p tem p pontos periódicos.

A fim de se codificar de maneira unívoca os pontos periódicos, as regras seguintes podem ser adotadas:

1. cada ponto periódico é codificado por uma sequência de p símbolos, cujo primeiro é o símbolo do ramo sobre o qual o ponto está localizado. Em seguida, vêm os símbolos associados aos pontos periódicos visitados segundo as iterações sucessivas;
2. essas sequências são ordenadas segundo uma regra específica chamada *Ordem Unimodal* (Oliveira 2007);
3. entre essas sequências, a mais forte, segundo as regras de ordenação, é escolhida para codificar a ordem que liga os p pontos periódicos na seção de Poincaré.

Quando cada sequência simbólica possível é associada a uma órbita de um atrator, a dinâmica é dita *completa*. Mas, segundo o parâmetro de bifurcação, todas as órbitas podem ser criadas ou destruídas pelas bifurcações, por exemplo, por bifurcações de duplicação de período e bifurcações sela-nó.

Exemplo 3.2.1 *Por exemplo, um comportamento aperiódico pode ser gerado por um simples mapa, como o mapa logístico, descrito por:*

$$x_{k+1} = f(x_k) = \mu x_k(1 - x_k). \tag{3.2}$$

para $\mu = [0,4]$ e $I = [0,1]$, em que μ é o parâmetro de bifurcação e I representa intervalo de mapeamento da função logística.

Para realizar uma análise simbólica deste mapa, divide-se seu espaço de estados em q regiões disjuntas separadas pelo ponto crítico C (ponto de máximo). Conseqüentemente, a evolução temporal do sistema dinâmico é traduzida em uma sequência de símbolos rotulando elementos da partição visitados por uma órbita. No caso do mapa logístico a partição é dada pelo ponto

crítico C que separa o mapa em dois ramos monotônicos e q é, então igual a 2. A região de crescimento é associada ao símbolo 0 e a região de decrescimento ao símbolo 1, conforme ilustrado na Figura 3.1. Então, uma trajetória caótica é descrita por uma sequência σ_n de símbolos definida de acordo com:

$$\sigma_n = \begin{cases} 0, & \text{se } x_n < x_c, \\ 1, & \text{se } x_n > x_c, \end{cases}$$

onde x_c é a coordenada do ponto crítico C . Tal mapa é dito unimodal uma vez que um único ponto crítico é envolvido (Godelle e Letellier 2000; Letellier 2008).

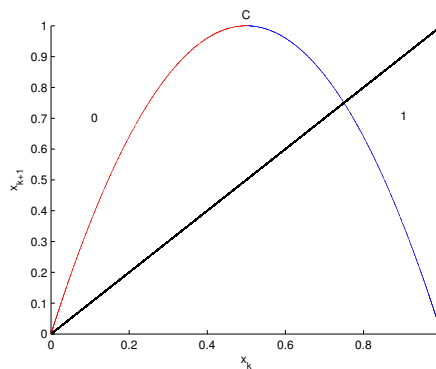


Figura 3.1: Partição topológica do mapa de primeiro retorno da função logística para $\mu = 3,9999$ e $N = 10000$ pontos. O ponto crítico está localizado no ponto de máximo da parábola, isto é, $x_c = 0,5$ para qualquer valor de μ

Esse tipo de partição do espaço de estados é denominada *partição topológica*, uma vez que associa diferentes símbolos para regiões com características distintas.

Exemplo 3.2.2 No caso do mapa logístico, a região crescente é uma região de preservação de ordem, pois dados dois pontos $x < y$ dessa região, $f(x) < f(y)$. Já a região decrescente é denominada região de inversão de ordem, pois se $x < y$ então $f(x) > f(y)$. Dessa forma, a evolução temporal do sistema é codificada em uma sequência de símbolos binários. Por exemplo, para $\mu = 3,2$ e $x_0 = x_c = 0,5$, a seguinte sequência é gerada: $S(x_0) = 111111 \dots$; para $\mu = 3,5545$ e $x_0 = x_c = 0,5$, a seguinte sequência é gerada: $S(x_0) = 110111 \dots$. A Figura 3.2 ilustra partições topológicas do mapa de primeiro retorno para os referidos valores de μ .

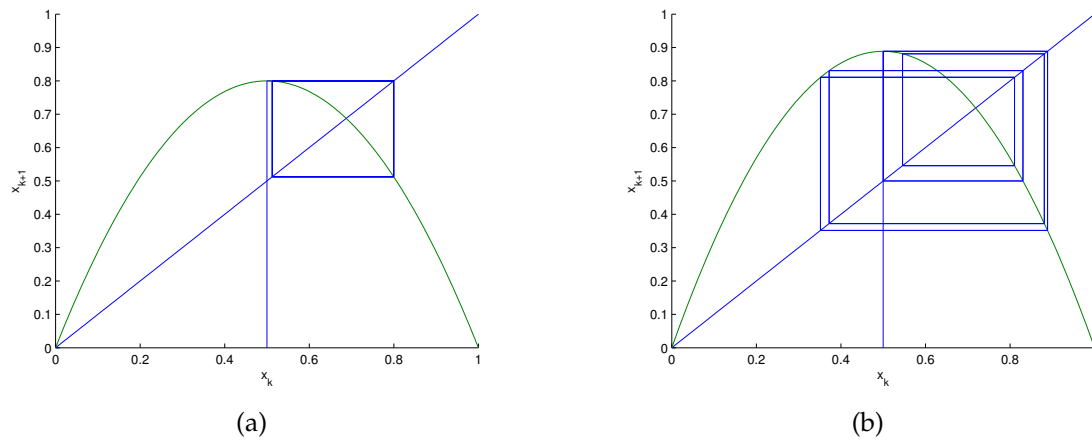


Figura 3.2: Partição topológica do mapa de primeiro retorno da função logística para (a) $\mu = 3,2$, (b) $\mu = 3,5545$ e $N = 10000$ pontos. O ponto crítico está localizado no ponto de máximo da parábola, isto é, $x_C = 0,5$ (*kneading sequence*) para qualquer valor de μ .

O mapa logístico atua na sequência como um deslocamento (*deslocamento de Jacques-Bernoulli* (Monteiro 2006)):

$$S(f^t(x)) = \sigma^t S(x) = s_t s_{t+1} s_{t+2} \dots \tag{3.3}$$

A partir dessa sequência é possível realizar várias análises sobre o comportamento do sistema.

Exemplo 3.2.3 *Por exemplo, é possível encontrar todas as órbitas periódicas para um dado valor do parâmetro de controle μ . Para isso, inicialmente é necessário definir uma regra de ordenação das sequências, de forma a preservar a ordenação dos pontos, ou seja, se $S(x) < S(y)$, então, $x < y$ e se $x < y$, então $S(x) = < S(y)$. Dessa maneira, dadas duas sequências $S(x) = s_0 s_1 s_2 \dots$ e $S(y) = t_0 t_1 t_2 \dots$, caso $s_0 < t_0$ então $x < y$ ou se $s_0 > t_0$ então $x > y$. Caso $s_0 = t_0$, deve-se comparar o segundo símbolo das sequências. Porém, caso $s_0 = t_0 = 1$, a ordem dos pontos é invertida (o símbolo 1 está associado a região de inversão de ordem), dessa forma se $s_1 < t_1$ então $x > y$ (e $x < y$ se $s_1 > t_1$). Considerando em seguida o caso em que os dois primeiros símbolos sejam iguais e, por exemplo, $s_2 < t_2$. Se os dois primeiros símbolos forem 00, então não há inversão de ordem e $x < y$. Porém, caso os primeiros símbolos sejam 10 ou 01, há inversão de ordem implicando $x > y$. Finalmente, para o caso dos símbolos serem 11, ocorreram duas*

inversões de ordem, que acarreta na preservação da ordem, portanto $x < y$.

Como regra geral, para se ordenarem sequências que possuam n primeiros símbolos iguais, deve-se contar o número de símbolos iguais a 1 e, caso esse número seja par, utiliza-se a ordenação natural do primeiro símbolo que diverja. Caso contrário, ou seja, o número de símbolos 1 seja ímpar, a ordenação é realizada de forma inversa para o primeiro símbolo que diverja (se $s_{n+1} < t_{n+1}$ $x > y$) (Tufillaro 1999).

Essas regras de ordenação implicam em consequências importantes. Define-se a sequência iniciada por $x = f(C)$ (onde C é o ponto crítico), isto é, $K(f) = S(f(C))$ como *kneading sequence*. Existe um teorema que diz que se $K(f)$ for não-periódica, dada uma sequência s , se $\sigma^i(s) < K(f)$ para todo $i \geq 0$, então existe um ponto $x \in I$ de forma que $S(x) = s$. Isto significa que é possível construir uma sequência aleatória s e, se essa sequência preceder $K(f)$, então existe um ponto no intervalo I correspondente a sequência s (Godelle e Letellier 2000; Letellier 2008).

Uma generalização desse teorema diz que, dadas duas sequências s e t , se existe um número $y \in I$ de forma que $S(y) = t$, caso exista um valor $m \geq 0$ de forma que $\sigma^n(s) < \sigma^m(t)$ para todo $n \geq 0$, então existe um valor $x \in I$ de forma que $S(x) = s$. A partir dessa generalização, dado $K(f)$, é possível determinar todas as órbitas periódicas possíveis para o sistema em função do parâmetro μ , assim como é feito na análise do diagrama de bifurcação.

Além disso, é possível realizar uma análise estatística das sequências de símbolos geradas, de forma a se estimar a probabilidade de ocorrência de diferentes sequências de tamanhos limitados. Assim, por meio desta análise, é possível encontrar sequências limitadas que caracterizem uma determinada dinâmica ou diferenciar entre tipos distintos de dinâmicas presentes em um sistema em diferentes condições.

(Godelle e Letellier 2000) realizam essa análise estatística das sequências de símbolos de um sistema dinâmico para diferenciar dinâmicas determinísticas e comportamentos estocásticos. Para isso foi realizada uma partição do espaço de estados do sistema denominada *partição estatística*. Esse tipo de partição é geralmente utilizado quando o sinal analisado é contaminado por ruído, impedindo a definição de uma partição

topológica (Letellier 2008).

Exemplo 3.2.4 A Figura 3.3 apresenta partições estatísticas do mapa de primeiro retorno para vários valores do parâmetro de bifurcação (μ). Observa-se, nesta figura, que a partição estatística coincide com a partição topológica para $\mu = 3,999$.

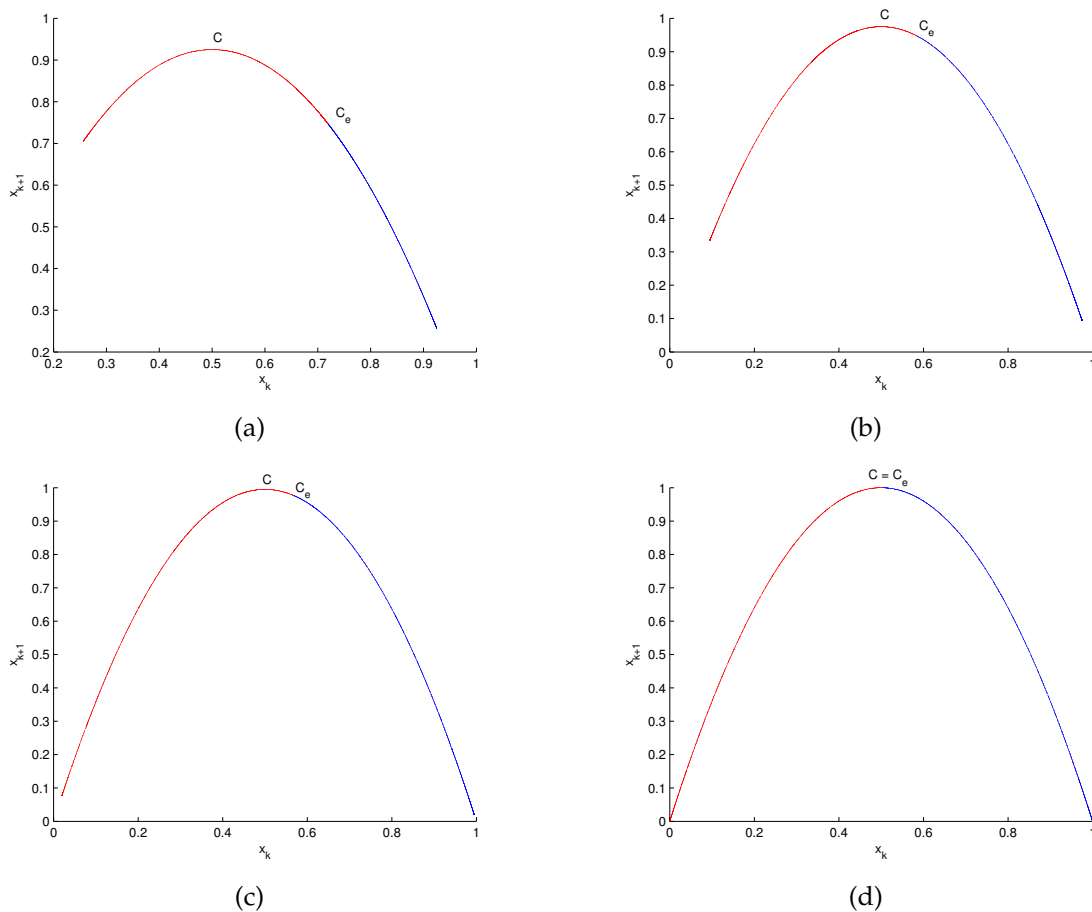


Figura 3.3: Partição estatística do mapa de primeiro retorno da função logística para (a) $\mu = 3,7$, (b) $\mu = 3,9$, (c) $\mu = 3,98$ e (d) $\mu = 3,999$; $N = 10000$ pontos. O ponto crítico está localizado no ponto de máximo da parábola, isto é, $x_C = 0,5$ para qualquer valor de μ . A partição topológica (C) coincide com a estatística (C_e) em $\mu = 3,999$.

Para realizar essa partição, escolhe-se um número q de regiões disjuntas que representarão o espaço de estados do sistema e, então, o sistema é particionado de forma que todos os símbolos $\sigma_i \in \{0,1, \dots, q - 1\}$ sejam equiprováveis. Por exemplo, para o

mapa logístico (Equação 3.2) e $q = 2$, escolhe-se um ponto de separação de forma que os símbolos 0 e 1 sejam equiprováveis.

Esse tipo de análise é realizada em uma *seção de Poincaré* do sistema de forma que cada intersecção de uma trajetória com o plano é convertida em um dos q símbolos.

A partir dessa sequência de símbolos, vários níveis de investigação podem ser realizados. Sequências de tamanho $n = 1$ são equiprováveis, sendo que, cada uma das q sequências possui probabilidade de ocorrência igual a $1/q$. Para $n = 2$, existem q^2 sequências possíveis, onde, por exemplo, P_{13} é a probabilidade de ocorrência do símbolo 13. Quanto maior o valor de n melhor a descrição do sistema obtida, porém, para que as frequências de ocorrência sejam suficientemente bem definidas, sugere-se que n seja limitado por N de forma que $1/q^n \approx 1\%$ dos pontos da seção de *Poincaré* (Godelle e Letellier 2000; Letellier 2008).

Dessa forma, definido um valor de n , pode-se gerar um histograma descrevendo as probabilidades de ocorrência de cada uma das sequências. Para isso, as q^n sequências possíveis são convertidas para números na base q e ordenadas de acordo com sua ordem natural. Por exemplo, dado $q = 2$, $\sigma_i = 0,1$, e $n = 3$, a sequência 011 é atribuída ao número 3.

A partir deste histograma é possível realizar diversos tipos de análise da dinâmica do sistema. Por exemplo, para o mapa logístico (Equação 3.2), quando todas as possíveis q^n sequências são equiprováveis (para $q = 2$ e $n = 6$ de forma que $1/2^6 \approx 0,01 \times 10000$), a dinâmica simbólica é dita *completa* e o valor do limiar de separação utilizado na partição estatística é igual ao ponto crítico utilizado na partição topológica. Essa análise é válida para outros mapas, como o mapa cúbico assimétrico e o mapa quártico.

Exemplo 3.2.5 *Dessa forma, esse tipo de análise pode ser utilizada para encontrar o valor do parâmetro de controle do sistema associado a uma dinâmica completa. Porém, esse histograma uniforme também pode ser associado ao ruído branco. Sendo assim, para se distinguir entre uma dinâmica determinística completa e um comportamento estocástico, deve-se investigar a dinâmica do sistema para diferentes valores de q , de modo que o histograma será uniforme apenas quando q for igual ao número de regiões disjuntas definidas pela partição topológica. Já no caso*

do comportamento estocástico, seu histograma será uniforme para qualquer valor de q (somente válido para ruído branco). Isso pode ser visto na Figura 3.4.

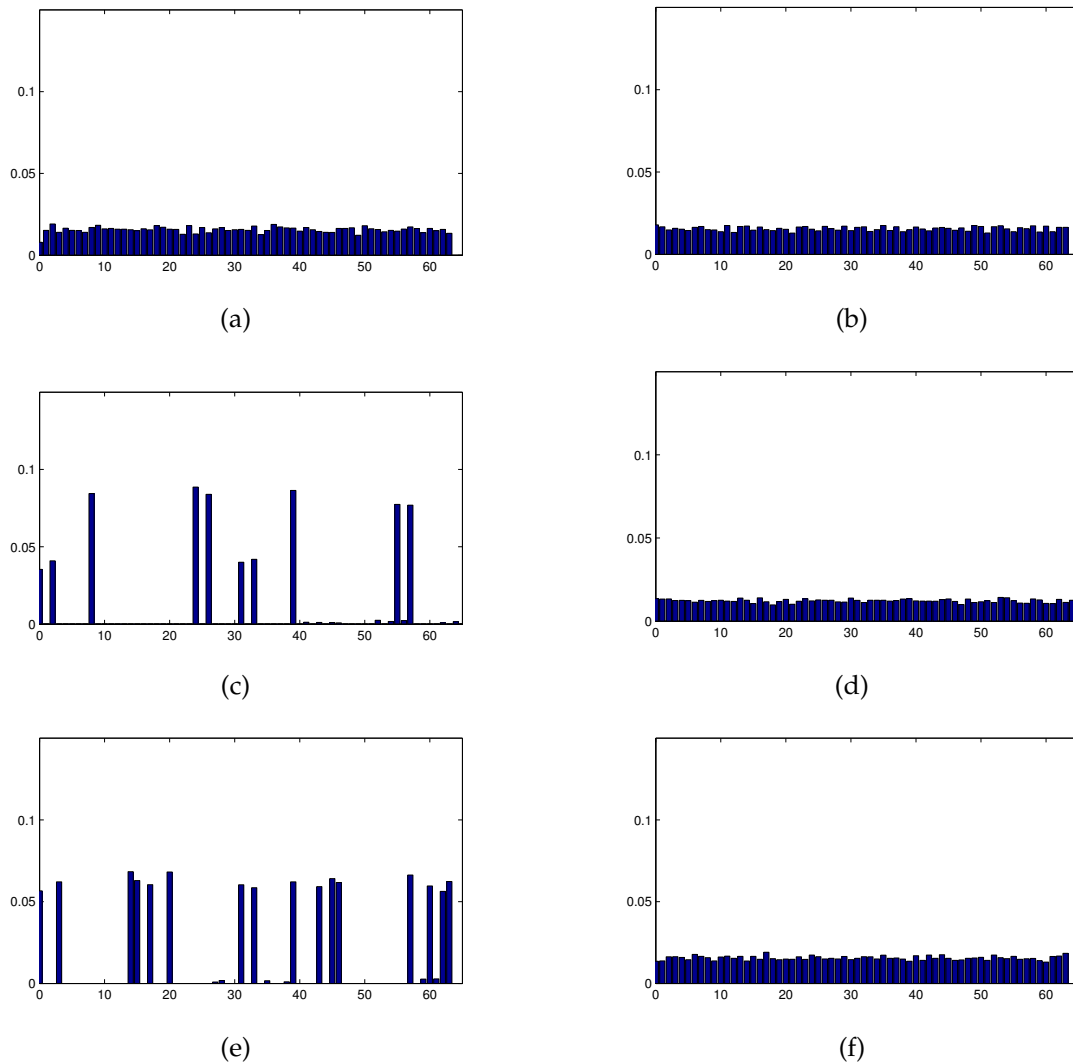


Figura 3.4: Histogramas das sequências simbólicas para uma série com dinâmica caótica obtida por meio do mapa logístico com $\mu = 3,9999$ (a) $q = 2, n = 6$; (c) $q = 3, n = 4$; (e) $q = 4, n = 3$) e para uma série com dinâmica estocástica (variável aleatória ou ruído branco) (b) $q = 2, n = 6$; (d) $q = 3, n = 4$; (f) $q = 4, n = 3$).

Exemplo 3.2.6 A Figura 3.1 mostra o mapa de primeiro retorno do mapa logístico para $\mu = 3,9999$, a partir do qual foram gerados 10000 pontos. Pode-se verificar que para este valor de μ a dinâmica está completamente desenvolvida, o que não ocorre para outros valores de μ , como pode ser visto nos histogramas da Figura 3.5:

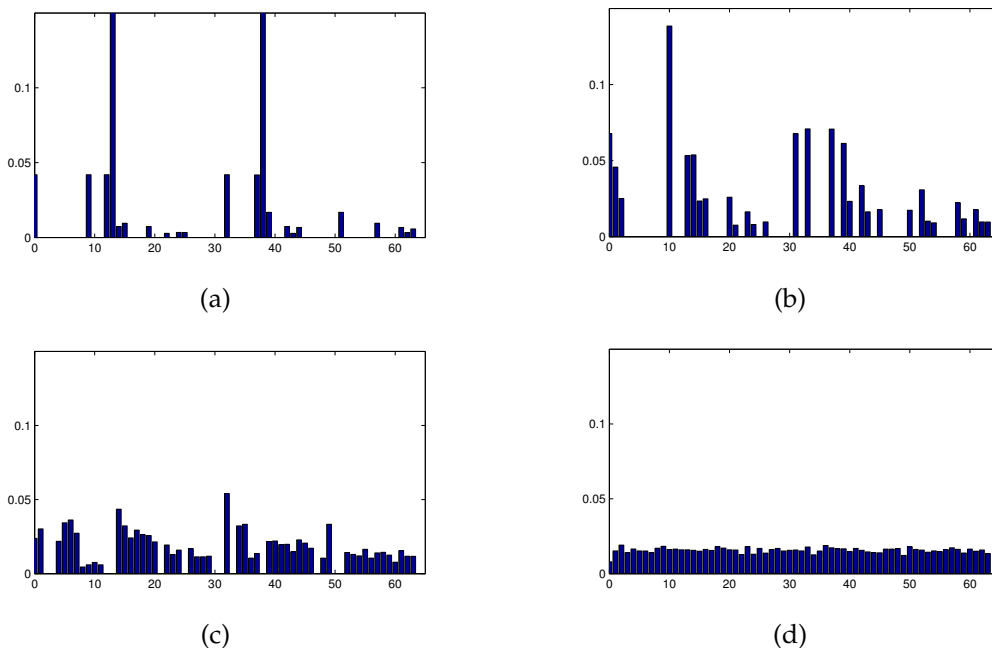


Figura 3.5: Histogramas das sequências simbólicas. Os cálculos foram feitos a partir de 10000 pontos, com $q = 2$ e $n = 6$: (a) $\mu = 3,70$; (b) $\mu = 3,90$; (c) $\mu = 3,98$; (d) $\mu = 3,9999$.

Vários níveis de investigação podem ser usados de acordo com o comprimento n das sequências de símbolos consideradas. No n -ésimo nível, q^n probabilidades são usadas; quanto maior o nível, melhor a descrição da dinâmica. Mas, para assegurar uma estatística suficiente, a escolha de n é limitada pelo comprimento N , em que N é o comprimento da série gerada pelas iterações de um mapa ou um número de intersecções de um fluxo com uma secção de Poincaré. Dessa forma, n pode ser determinado de acordo com a aproximação: $1/q^n \approx 1\%$ (Letellier 2008).

Exemplo 3.2.7 Assim, para $N = 10000$, a escolha de n pode ser, por exemplo, $q = 2$, $n = 6 \Rightarrow q^n = 64$, ou $q = 3$, $n = 4 \Rightarrow q^n = 81$, ou $q = 4$, $n = 3 \Rightarrow q^n = 64$, como ilustrado nos histogramas da Figura 3.4.

Ao se investigar os histogramas construídos para os valores de μ variando de 3,70 a 3,9999 (Figura 3.5), observa-se que surgem sequências diferentes, quando o valor de μ é aumentado.

A Figura 3.5d mostra que, além de conter todas as sequências possíveis, o histograma

é plano, isto é, as sequências são equiprováveis. Uma dinâmica simbólica completa em Σ_q é associada com um histograma plano quando a estatística da sequência simbólica é realizada com o número de símbolos induzidos pela partição topológica. Entretanto, um histograma plano pode ser também associado com ruído branco (como mostra a Figura 3.4 b,d,f), para o qual todas as sequências de qualquer comprimento são, também, equiprováveis. Sendo assim, é importante ter um critério para distinguir um comportamento determinístico, caracterizado por uma dinâmica simbólica completa de um ruído branco, principalmente no caso de dados experimentais dos quais as equações são desconhecidas.

A saída está na investigação da dinâmica por meio da variação do número de símbolos q com a qual a trajetória é codificada. O histograma somente é plano quando a partição adotada corresponde à partição topológica. Portanto, o comportamento determinístico caracterizado por uma dinâmica simbólica completa pode ser distinguido de um ruído branco, pois no caso do ruído branco o histograma permanece plano para qualquer escolha de q e n (como pode ser visto nas Figuras 3.4 b,d,f). De fato, isso pode ser verificado.

Exemplo 3.2.8 *A Figura 3.5 d mostra um histograma plano quando foram escolhidos $q = 2$ e $n = 6$, sabidamente a partição topológica, e com $\mu = 3,9999$, para o qual a dinâmica simbólica é completa. Agora, se mantido o valor de μ e variando-se os valores de q e n para, por exemplo, $q = 3$ e $n = 4$ ou $q = 4$ e $n = 3$, o histograma não permanece plano (Figuras 3.4 c,e), o que indica que a partição não é a correta. Fazendo o mesmo para o caso do ruído branco, ou seja, mudando a partição para $q = 3$ e $n = 4$ ou $q = 4$ e $n = 3$, verifica-se que o histograma permanece plano (Figuras 3.4 d,f) como esperado, insensível à partição.*

Outro ponto importante, como será visto na próxima seção, a *entropia de Shannon* pode ser calculada a partir de sequências simbólicas obtidas por meio da dinâmica simbólica. O valor da entropia de Shannon atingirá o seu valor máximo, isto é, o valor correto para a série analisada quando a partição for realizada corretamente. Assim sendo, a entropia de Shannon pode ser usada para determinar a partição correta nos casos de sistemas reais que não apresentam uma partição bem definida, como o mapa

logístico (Oliveira 2007).

3.2.2 Entropia

O conceito de *entropia* foi inicialmente utilizado na termodinâmica em meados do século XIX. No século passado, foi aplicado na mecânica estatística e, em 1948, utilizado pelo matemático Claud E. Shannon na Teoria da Informação (Shannon 1948; Papoulis 1991; Monteiro 2006).

Esta seção apresenta duas formas de calcular entropia, a *entropia de Shannon* e a *entropia aproximada*.

Entropia Informacional de Shannon

O significado da palavra *informação* não difere radicalmente do significado do senso-comum. Uma mensagem é dita informativa se tal mensagem disser algo que antes era pouco esperado ou desconhecido. Isso significa que mais informação é transportada pelas mensagens menos prováveis. Essa ideia tem uma implicação importante: se é possível medir a probabilidade de ocorrência de uma mensagem, então, pode-se medir seu conteúdo de informação de forma similar. De fato, a *Teoria da Informação* é baseada em probabilidade; mais especificamente, na frequência relativa com que cada mensagem é emitida por uma fonte.

Em 1928, Hartley apresentou uma fórmula para medir informação quantitativamente. Seu objetivo era aplicar essa fórmula para comparar as capacidades de transmissão de vários sistemas elétricos de telecomunicações (telégrafo, telefone e rádio). Hartley sugeriu que informação surge da seleção sucessiva de símbolos pertencentes a um dado vocabulário. Se há N símbolos disponíveis em cada seleção, então a quantidade de informação H associada a uma seleção é igual a $\log N$, isto é, $H = \log N$.

Posteriormente, surgiram os trabalhos de Wiener e Shannon que originaram a Teoria da Informação. Shannon tratou o problema num sentido estatístico mais geral do que Hartley havia feito. A medida logarítmica de Hartley envolve o número de alternativas possíveis. A medida da quantidade de informação proposta por Shannon depende

da probabilidade de ocorrência de cada mensagem possível, e não do número de mensagens possíveis.

Suponha que um evento ocorra e que seja transmitida uma mensagem relatando sua ocorrência. Define-se a quantidade de informação h contida nessa mensagem como o logaritmo da razão entre a probabilidade q do evento após a mensagem ser recebida e a probabilidade p no receptor do evento antes da mensagem ser recebida. Num canal sem ruído, o receptor tem certeza que a mensagem recebida está correta. Portanto, a probabilidade do evento após a mensagem ser recebida é $p = 1$ e a informação recebida h torna-se o logaritmo do inverso da probabilidade p , ou seja:

$$h = \log(1/p) = -\log p. \tag{3.4}$$

Essa fórmula condiz com a noção de que a quantidade de informação aumenta quando a probabilidade de uma mensagem tende a zero. No outro extremo, tem-se que, quando $p = 1$, isto é, quando se sabe exatamente o que será transmitido, então a quantidade de informação recebida é nula ($h = 0$).

Espera-se que o conteúdo de informação associado a duas mensagens independentes, com probabilidades de ocorrência p_1 e p_2 , seja dado pela soma dos conteúdos de informação dessas mensagens. Ou seja, espera-se que $h(p_1 p_2) = h(p_1) + h(p_2)$. Essa propriedade de aditividade ajuda a reforçar a escolha de h como uma função logarítmica da probabilidade, pois $\log(p_1 p_2) = \log(p_1) + \log(p_2)$.

Para o caso em que há N mensagens alternativas, cada qual com probabilidade de ocorrência p_j ($\sum_j p_j = 1$) e quantidade de informação h_j , então a quantidade média de informação H obtida por mensagem recebida é dada pela média ponderada dos h_j , isto é, $H = \sum_j p_j h_j$, ou:

$$H = - \sum_j^N p_j \log_2 p_j. \tag{3.5}$$

Shannon chamou H de *entropia informacional*, que é medida em bits por mensagem (ou bits por símbolo). Essa expressão coincide com a fórmula de Hartley somente

quando todas as N mensagens são equiprováveis. Nessa situação, as probabilidades p_j são iguais ao inverso do número N de alternativas, ou seja, $p_j = 1/N$ e, assim, $H = \log_2 N$, que é seu valor máximo.

A fórmula da entropia informacional é idêntica à fórmula da entropia termodinâmica escrita segundo a interpretação da Mecânica Estatística. Tal interpretação foi concebida por Boltzmann.

Entropia Aproximada

O cálculo da entropia aproximada tem por objetivo estimar a entropia de Kolmogorov-Sinai (Kolmogorov 1958). Na literatura, diz-se que “uma entropia de Kolmogorov-Sinai (K) não-nula positiva e finita é condição suficiente para a ocorrência de caos determinístico” (Fiedler-Ferrara e Prado 1994). A entropia de Kolmogorov-Sinai mede a taxa média de criação de informação em um sistema dinâmico ou, de forma equivalente, a perda de informação de um observador sobre o sistema. Pode-se dizer que a entropia quantifica a regularidade ou a complexidade de séries temporais oriundas de sistemas dinâmicos não-lineares.

Para se definir a entropia de Kolmogorov-Sinai, considere um ponto $y(t) = y_1(t), \dots, y_d(t)$ evoluindo em uma órbita de um espaço de fases de dimensão d . Divida este espaço em caixas de dimensão ϵ^d e observe o sistema a intervalos uniformes de tempo τ . O parâmetro τ pode ser interpretado como o passo usado na reconstrução de Takens (Takens 1981). Define-se a probabilidade conjunta $P(i_1, i_2, i_3, \dots, i_b)$ como a probabilidade de observar o ponto $y(t)$ na caixa i_1 no tempo $t = 0$, na caixa i_2 em $t = \tau$, ... e i_b em $t = b\tau$. A entropia de Kolmogorov-Sinai é definida como:

$$K = - \lim_{\tau \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{b \rightarrow \infty} \frac{1}{b\tau} \sum_{i_1 \dots i_b} P(i_1, \dots, i_b) \ln P(i_1, \dots, i_b). \quad (3.6)$$

Observe que para pontos fixos, por exemplo, a probabilidade de se observar o sistema no ponto fixo é igual à unidade, enquanto que a probabilidade de se observar o sistema em outros pontos é nula. Neste caso, a quantidade de informação que se ganha ao se observar a trajetória do sistema é nula; pode-se verificar que $k = 0$. Por outro lado,

quando as probabilidades conjuntas são iguais, a entropia atinge o seu valor máximo.

O cálculo de K baseado em sua definição é de difícil implementação. O método da entropia aproximada simplifica o cálculo de K , não obstante exige um tempo computacional elevado. Este método foi desenvolvido por (Pincus 1991), como uma variante do algoritmo de Grassberger e Procaccia (1983) (Grassberger e Procaccia 1983) para o cálculo da entropia de Kolmogorov-Sinai. Para a definição da entropia aproximada, considera-se uma variável observada do sistema $y(k)$ constituída por N amostras e um espaço de fases reconstruído com dimensão de imersão d_e . A entropia aproximada é definida por:

$$\begin{aligned}
 EnAp(N, d_e, \epsilon) &= \frac{1}{N - d_e} \sum_{n=0}^{N-d_e-1} \ln P[|y(n+j) - y(m+j)| \leq \epsilon |y(n+k) - y(m+k)| \leq \epsilon] \\
 \forall k &= 0, 1, \dots, j-1, m, j \in N
 \end{aligned} \tag{3.7}$$

3.3 Máquinas de Vetores de Suporte

Máquinas de Vetores de Suporte (*Support Vector Machine, SVM*) são modelos de aprendizado de máquina (redes neurais artificiais alimentadas adiante), propostos por Vapnik, muito utilizados para resolver problemas de classificação de padrões (Haykin 2008). Neste contexto, a ideia principal de uma SVM é construir um hiperplano como superfície de decisão de tal forma que a margem de separação entre exemplos positivos e negativos seja máxima. Assim, o SVM consiste em um mapeamento não-linear de um vetor de entrada para um *espaço de características* de alta dimensionalidade, que é oculto da entrada e da saída. O hiperplano ótimo separa as características descobertas após o mapeamento. A Figura 3.6 ilustra esta transformação.

A SVM representa uma implementação do *método de minimização estrutural de risco*, que é um princípio estatístico indutivo baseado no fato de que a taxa de erro de uma máquina de aprendizagem sobre dados de teste (taxa de erro de generalização) é limitada pela soma da taxa de erro de treinamento e por um termo que depende da *dimensão*

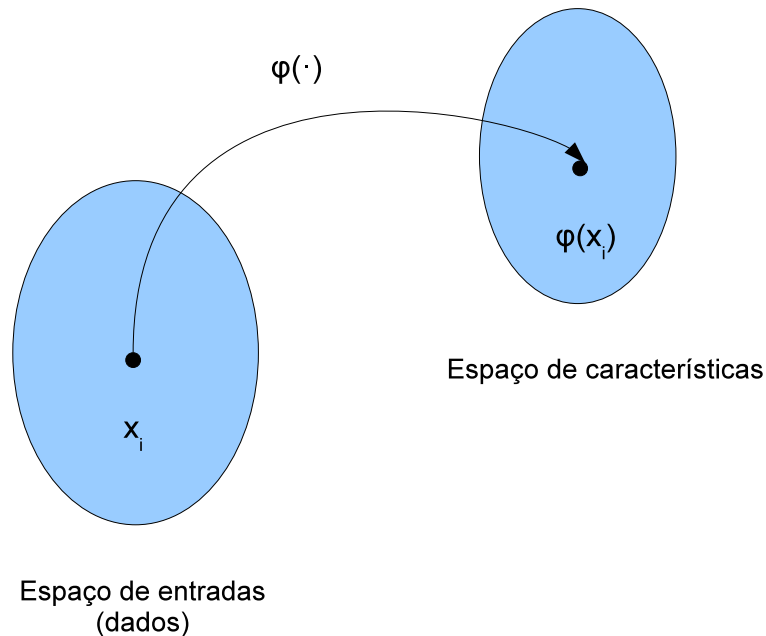


Figura 3.6: Mapa não-linear $\varphi(\cdot)$ do espaço de entrada para o espaço de características.

de Vapnik-Chervonenkis (V-C); no caso de padrões separáveis, um SVM produz um valor de zero para o primeiro termo e minimiza o segundo termo. Conseqüentemente, o SVM pode fornecer um bom desempenho de generalização em problemas de classificação de padrões, apesar do fato de que ele não incorpora conhecimento do domínio do problema. Este atributo é único dos SVM's. A seguir, a teoria das SVM's será detalhada de forma matemática. As equações mostradas, nesta Seção, foram baseadas no livro de (Haykin 2008).

Considere uma amostra de treinamento $\{(x_i, d_i)\}_{i=1}^N$, em que x_i é o padrão de entrada para o i -ésimo exemplo e d_i é a resposta desejada correspondente (saída-alvo). Para começar, assume-se que o padrão (classe) representado pelo subconjunto $d_i = +1$ e o padrão representado pelo subconjunto $d_i = -1$ são "linearmente separáveis". A equação de uma superfície de decisão na forma de um hiperplano que realiza esta separação é:

$$\mathbf{w}^T \mathbf{x} + b = 0, \quad (3.8)$$

em que \mathbf{x} é um vetor de entrada, \mathbf{w} é um vetor peso ajustável e b é um *viés*. Pode-se assim escrever:

$$\begin{aligned} \mathbf{w}^T \mathbf{x}_i + b &\geq 0 \quad \text{para } d_i = +1 \\ \mathbf{w}^T \mathbf{x}_i + b &\leq 0 \quad \text{para } d_i = -1 \end{aligned} \quad (3.9)$$

Para um dado vetor peso de \mathbf{w} e *viés* b , a separação entre o hiperplano definido na equação 3.8 e o ponto de dado mais próximo é denominada a *margem de separação*, representada por ρ . O objetivo de um SVM é encontrar o hiperplano particular para o qual a margem de separação ρ é máxima. Sob esta condição, a superfície de decisão é referida como o *hiperplano ótimo*. A Figura 3.7 ilustra a construção geométrica de um hiperplano ótimo para um espaço de entrada bidimensional.

Considere que \mathbf{w}_o e b_o representem os valores ótimos do vetor peso e do *viés*, respectivamente. Consequentemente, o *hiperplano ótimo*, representando uma superfície de decisão linear multidimensional no espaço de entrada, é definido por

$$\mathbf{w}_o^T \mathbf{x} + b_o = 0, \quad (3.10)$$

o que é a equação 3.8 reescrita. A função discriminante

$$g(\mathbf{x}) = \mathbf{w}_o^T \mathbf{x} + b_o = 0, \quad (3.11)$$

fornece uma medida algébrica da *distância* de \mathbf{x} até o hiperplano (Duda, Hart e Stork 2006). Talvez a maneira mais fácil de visualizar isto seja expressar \mathbf{x} como

$$\mathbf{x} = \mathbf{x}_p + r \frac{\mathbf{w}_o}{\|\mathbf{w}_o\|}$$

em que \mathbf{x}_p é a projeção normal de \mathbf{x} sobre o hiperplano ótimo e r é a distância algébrica desejada; r é positiva se \mathbf{x} estiver no lado positivo do hiperplano ótimo e negativa se \mathbf{x}

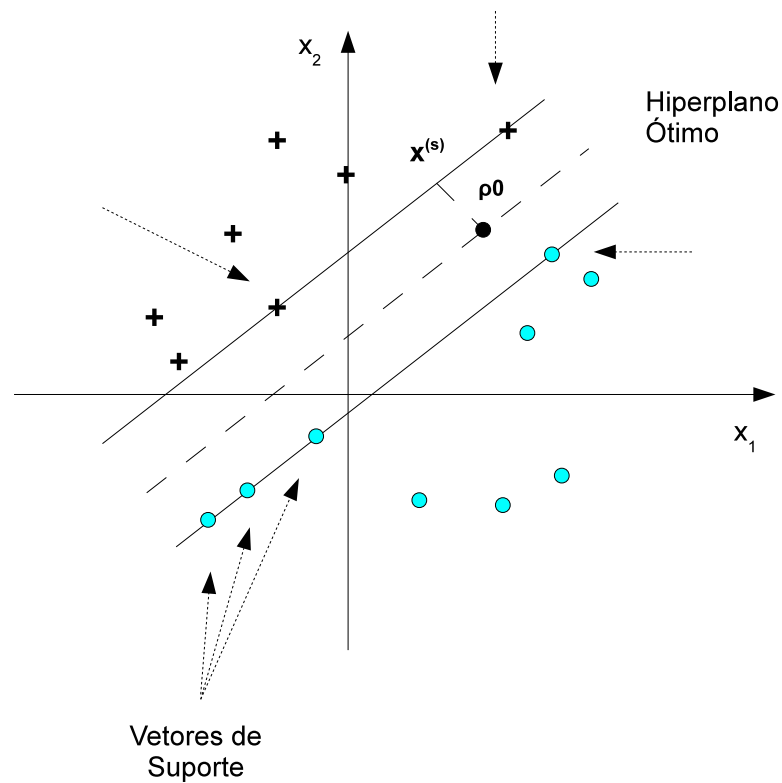


Figura 3.7: Ilustração da idéia de um hiperplano ótimo para padrões linearmente separáveis. As setas indicam os vetores de suporte.

estiver no lado negativo. Como, por definição, $g(\mathbf{x}_p) = 0$, resulta em

$$g(\mathbf{x}) = \mathbf{w}_0^T \mathbf{x} + b_0 = r \|\mathbf{w}_0\|$$

ou

$$r = \frac{g(\mathbf{x})}{\|\mathbf{w}_0\|}. \quad (3.12)$$

Em particular, a distância da origem (por exemplo, $\mathbf{x}=\mathbf{0}$) até o hiperplano ótimo

é dada por $b_o/\|\mathbf{w}_o\|$. Se $b_o > 0$, a origem está no lado positivo do hiperplano ótimo; $b_o < 0$, ela está no lado negativo. Se $b_o = 0$, o hiperplano ótimo passa pela origem. Uma interpretação geométrica desses resultados algébricos é dada na Figura 3.8.

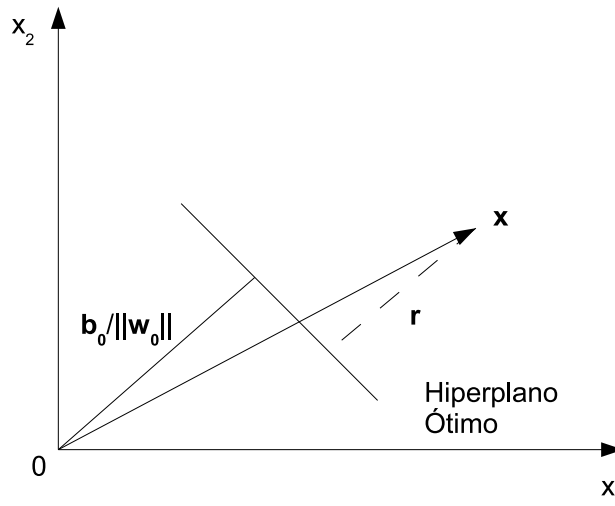


Figura 3.8: Interpretação geométrica das distâncias algébricas de pontos até o hiperplano ótimo para um caso bidimensional.

A questão a resolver é encontrar os parâmetros \mathbf{w}_o e b_o para o hiperplano ótimo, dado o conjunto de treinamento $\mathfrak{I} = \{(x_i, d_i)\}$. Com base nos resultados retratados na Figura 3.8, observa-se que o par (\mathbf{w}_o, b_o) deve satisfazer a restrição:

$$\begin{aligned} \mathbf{w}_o^T \mathbf{x}_i + b_o &\geq 0 \quad \text{para } d_i = +1 \\ \mathbf{w}_o^T \mathbf{x}_i + b_o &\leq 0 \quad \text{para } d_i = -1 \end{aligned} \tag{3.13}$$

Note que se a equação 3.9 for válida, isto é, os padrões forem linearmente separáveis, pode-se sempre escalar \mathbf{w}_o e b_o de modo que a equação 3.13 seja válida; esta operação de escalamento não afeta a equação 3.10.

Os pontos de dados particulares (x_i, d_i) para os quais a primeira ou a segunda linha da equação 3.13 é satisfeita com o sinal de igualdade são chamados de *vetores de suporte*²,

²Os *vetores de suporte* são amostras do conjunto de treinamento que definem um hiperplano ótimo de separação entre as classes (Duda, Hart e Stork 2006).

por isso o nome “máquina de vetor de suporte”. Estes vetores desempenham um papel proeminente na operação desta classe de máquinas de aprendizagem. Em termos conceituais, os vetores de suporte são aqueles pontos de dados que se encontram mais próximos da superfície de decisão e são, portanto, os mais difíceis de classificar (Haykin 2008). Dessa forma, têm uma influência direta na localização ótima da superfície de decisão.

Considere um vetor de suporte \mathbf{x}^s para o qual $d^s = +1$. Então, por definição, tem-se

$$g(\mathbf{x}^s) = \mathbf{w}_o^T \mathbf{x}^s \mp b_o \mp 1 \quad \text{para } d^s = \mp 1 \quad (3.14)$$

Da equação 3.12 a *distância algébrica* do vetor de suporte \mathbf{x}^s até o hiperplano ótimo é

$$r = \frac{g(\mathbf{x}^s)}{\|\mathbf{w}_o\|}$$

$$= \begin{cases} \frac{1}{\|\mathbf{w}_o\|}, & \text{se } d^s = +1 \\ \frac{-1}{\|\mathbf{w}_o\|}, & \text{se } d^s = -1 \end{cases}$$

em que o sinal positivo indica que \mathbf{x}^s se encontra no lado positivo do hiperplano ótimo e o sinal negativo indica que \mathbf{x}^s está no lado negativo do hiperplano ótimo. Considere que ρ represente o valor ótimo da *margem de separação* entre as duas classes que constituem o conjunto de treinamento \mathfrak{S} . Então, da equação 3.15 resulta que

$$\rho = 2r$$

$$= \frac{2}{\|\mathbf{w}_o\|} \quad (3.15)$$

Esta equação afirma que maximizar a margem de separação entre classes é equivalente a minimizar a norma euclidiana do vetor peso \mathbf{w} .

Em resumo, o hiperplano ótimo definido na equação 3.10 é *único* no sentido de que o vetor peso \mathbf{w}_o fornece a máxima separação possível entre exemplos positivos e negativos. Esta condição ótima é alcançada minimizando-se a norma euclidiana do

vetor peso \mathbf{w} .

3.3.1 Técnicas de Otimização para a determinação do Hiperplano Ótimo

Existem várias técnicas de otimização que podem ser utilizadas para se encontrar o hiperplano ótimo. Neste trabalho, será descrita a *otimização quadrática*. Assim, tem-se como objetivo, o desenvolvimento de um procedimento eficiente do ponto de vista computacional para, utilizando a amostra de treinamento $\mathfrak{J} = \{(\mathbf{x}_i, d_i)\}$, encontrar o hiperplano ótimo, sujeito à restrição

$$d_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \quad \text{para } i = 1, 2, \dots, N \quad (3.16)$$

Esta restrição combina as duas linhas da equação 3.13 com \mathbf{w} no lugar de \mathbf{w}_0 .

O *problema primordial* de otimização consiste em encontrar os valores ótimos do vetor peso \mathbf{w} e do viés b , dada uma amostra de treinamento $\{(\mathbf{x}_i, d_i)\}$, de tal maneira que o vetor peso minimize a *função de custo* dada por:

$$\Phi(\mathbf{w}) = \frac{1}{2} \mathbf{w}^T \mathbf{w}, \quad (3.17)$$

em que $1/2$ é um fator de escala.

Conseqüentemente, pode-se resolver o problema de otimização restrito usando o *método dos multiplicadores de Lagrange* (Haykin 2008).

Primeiramente, contrói-se a *função lagrangiana*:

$$J(\mathbf{w}, b, \alpha) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{i=1}^N \alpha_i [d_i(\mathbf{w}^T \mathbf{x}_i + b) - 1], \quad (3.18)$$

em que as variáveis auxiliares não-negativas α_i são chamadas de *multiplicadores de Lagrange*. A solução do problema de otimização restrito é determinada pelo *ponto de sela* da função lagrangiana $J(\mathbf{w}, b, \alpha)$, que deve ser *minimizada* em relação a \mathbf{w} e a b e *maximizada* em relação a α . Assim, diferenciando $J(\mathbf{w}, b, \alpha)$ em relação a \mathbf{w} e a b e

igualando os resultados a zero, obtêm-se as duas *condições de otimização*:

$$\begin{aligned}\frac{\partial J(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} &= \mathbf{0} \\ \frac{\partial J(\mathbf{w}, b, \alpha)}{\partial b} &= 0\end{aligned}\quad (3.19)$$

Resolvendo ambas as condições de otimização, obtêm-se, respectivamente:

$$\begin{aligned}\mathbf{w} &= \sum_{i=1}^N \alpha_i d_i \mathbf{x}_i \\ \sum_{i=1}^N \alpha_i d_i &= 0\end{aligned}\quad (3.20)$$

O vetor solução \mathbf{w} é definido em termos de uma expansão que envolve os N exemplos de treinamento. É importante notar que no ponto de sela, para cada multiplicador de Lagrange α_i , o produto daquele multiplicador pela sua restrição correspondente desaparece, como mostrado por

$$\alpha_i [d_i(\mathbf{w}^T \mathbf{x}_i + b) - 1] = 0 \quad \text{para } i = 1, 2, \dots, N \quad (3.21)$$

Dessa forma, apenas aqueles multiplicadores que satisfazem exatamente a equação 3.21 podem assumir valores *não-nulos*. Esta propriedade resulta das *condições de Kuhn-Tucker* da teoria da otimização (Haykin 2008).

Os multiplicadores de Lagrange, neste caso, fornecem a solução ótima e este problema de otimização é, então, chamado de *problema dual* que minimiza a função de custo, ou seja:

$$\Phi(\mathbf{w}_o) = J(\mathbf{w}, b, \alpha) = \min_{\mathbf{w}} J(\mathbf{w}, b, \alpha)$$

Expandindo a equação 3.18 e fazendo a função objetivo $J(\mathbf{w}, b, \alpha) = Q(\alpha)$, tem-se:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j \mathbf{x}_i^T \mathbf{x}_j. \quad (3.22)$$

Assim, pode-se formular o problema dual de otimização que consistem em encontrar os multiplicadores de Lagrange $\{\alpha_i\}_{i=1}^N$, dada a amostra de treinamento $\{(\mathbf{x}_i, d_i)\}$, que maximizem a função objetivo, dada pela equação 3.22, sujeita às restrições:

$$\begin{aligned} \sum_{i=1}^N \alpha_i d_i &= 0 \\ \alpha_i &\geq 0 \quad \text{para } i = 1, 2, \dots, N \end{aligned}$$

Havendo calculado os multiplicadores de Lagrange, pode-se calcular o vetor peso \mathbf{w} e o viés b utilizando as equações 3.20 e 3.14.

3.3.2 Hiperplano ótimo para padrões não-separáveis

Para padrões não-separáveis, introduzem-se as *variáveis de folga*, que são um conjunto de variáveis escalares não-negativas, $\{\xi_i\}_{i=1}^N$, na definição do hiperplano de separação, como definido a seguir:

$$d_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \text{para } i = 1, 2, \dots, N \quad (3.23)$$

Essas variáveis medem o desvio de um ponto de dado da condição ideal de separabilidade (Haykin 2008).

Assim, deve-se encontrar um hiperplano de separação para o qual o erro de classificação, como média sobre o conjunto de treinamento, é minimizado (Haykin 2008). Pode-se fazer isto, minimizando o funcional

$$\Phi(\mathbf{w}, \xi) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i, \quad (3.24)$$

em que o parâmetro C é um parâmetro positivo, denominado *parâmetro de regularização*, que controla o compromisso entre a complexidade do SVM e o número de pontos não-separáveis. Esse parâmetro deve ser especificado pelo usuário.

Assim, o problema primordial, para o caso não-separável, pode ser formulado das seguintes maneiras:

1. Deve-se encontrar os valores ótimos do vetor peso \mathbf{w} e do viés b e as variáveis soltas ξ_i , dada uma amostra de treinamento $\{(\mathbf{x}_i, d_i)\}$, que minimizem o funcional de custo $\Phi(\mathbf{w}, \xi)$ e satisfaçam às restrições:

$$\Phi(\mathbf{w}, \xi) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i$$

$$\begin{aligned} d_i(\mathbf{w}^T \mathbf{x}_i + b) &\geq 1 - \xi, \quad \text{para } i = 1, 2, \dots, N \\ \xi_i &\geq 0 \quad \forall \quad i \end{aligned} \quad (3.25)$$

2. Deve-se encontrar os multiplicadores de Lagrange $\{\alpha_i\}_{i=1}^N$, dada uma amostra de treinamento $\{(\mathbf{x}_i, d_i)\}$, que maximizem a função objetivo $Q(\alpha)$ e e satisfaçam às restrições:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j d_i d_j \mathbf{x}_i^T \mathbf{x}_j$$

$$\sum_{i=1}^N \alpha_i d_i = 0 \quad \text{para } i = 1, 2, \dots, N \quad (3.26)$$

O problema de otimização para padrões não-separáveis se reduz ao caso linearmente separável, fazendo $\xi_i = 0 \forall i$ nas equações 3.23 e 3.24.

3.3.3 Núcleo do produto interno

Uma noção que é central à construção do algoritmo de aprendizagem por *vetor de suporte* é o núcleo (*kernel*) do produto interno entre um vetor de suporte \mathbf{x}_i e o vetor \mathbf{x} retirado do espaço de entrada.

Considere que \mathbf{x} represente um vetor retirado do *espaço de entrada*, que é assumido como tendo *dimensão* m_0 . Considere que $\{\varphi_j(\mathbf{x})\}_{j=1}^{m_1}$ represente um conjunto de transformações não-lineares do espaço de entrada para o *espaço de características*: m_1 é a dimensão do espaço de características. Assume-se que $\varphi_j(\mathbf{x})$ é definido *a priori* $\forall j$. Dado este conjunto de transformações não-lineares, podemos definir um hiperplano atuando como superfície de decisão como segue:

$$\sum_{j=1}^{m_1} w_j \varphi_j(\mathbf{x}) + b = 0, \quad (3.27)$$

em que $\{w_j\}_{j=1}^{m_1}$ representa um conjunto de pesos lineares conectando o espaço de características com o espaço de saída e b é o *viés*.

Pode-se simplificar o desenvolvimento escrevendo

$$\sum_{j=1}^{m_1} w_j \varphi_j(\mathbf{x}) = 0, \quad (3.28)$$

onde foi assumido que $\varphi_0(\mathbf{x}) = 1 \forall \mathbf{x}$, de modo que w_0 represente o *viés* b . Esta equação define a superfície de decisão calculada no espaço de características em termos dos pesos lineares do SVM. A quantidade $\varphi_j(\mathbf{x})$ representa a entrada fornecida ao peso w_j através do espaço de características. Definindo o vetor:

$$\varphi(\mathbf{x}) = [\varphi_0(\mathbf{x}), \varphi_1(\mathbf{x}), \dots, \varphi_{m_1}(\mathbf{x})]^T, \quad (3.29)$$

onde, por definição, tem-se:

$$\varphi_0(\mathbf{x}) = 0 \quad \forall \mathbf{x} \quad (3.30)$$

Na verdade, o vetor $\varphi(\mathbf{x})$ representa a “imagem” induzida no espaço de características pelo vetor de entrada \mathbf{x} , como ilustrado na Figura 3.6. Assim, em termos desta imagem, pode-se definir a superfície de decisão na forma compacta:

$$\mathbf{w}^T \varphi(\mathbf{x}) = 0. \quad (3.31)$$

Adaptando a equação 3.20 à presente situação envolvendo um espaço de características no qual procura-se a separabilidade “linear” de características, pode-se escrever

$$\mathbf{w} = \sum_{i=1}^N \alpha_i d_i \varphi(\mathbf{x}_i), \quad (3.32)$$

em que o *vetor de características* $\varphi(\mathbf{x}_i)$ corresponde ao padrão de entrada \mathbf{x}_i no i -ésimo exemplo. Dessa forma, substituindo a equação 3.32 em 3.31, pode-se definir a superfície de decisão calculada no espaço de características como:

$$\mathbf{w} = \sum_{i=1}^N \alpha_i d_i \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}). \quad (3.33)$$

O termo $\varphi(\mathbf{x}_i)^T \varphi(\mathbf{x})$ representa o produto interno de dois vetores induzidos no espaço de características pelo vetor de entrada \mathbf{x} e o padrão de entrada \mathbf{x}_i relativo ao i -ésimo exemplo. Pode-se, então, introduzir o *núcleo do produto interno* ou *kernel* representado por $K(\mathbf{x}, \mathbf{x}_i)$ e definido por

$$\begin{aligned} K(\mathbf{x}, \mathbf{x}_i) &= \varphi(\mathbf{x})^T \varphi(\mathbf{x}_i) \\ &= \sum_{j=0}^{m_1} \varphi_j(\mathbf{x}) \varphi_j(\mathbf{x}_i) \quad \text{para } i = 1, 2, \dots, N \end{aligned} \quad (3.34)$$

Desta definição, pode-se observar que o *kernel* é uma *função simétrica* de seus argumentos, como mostrado por

$$K(\mathbf{x}, \mathbf{x}_i) = K(\mathbf{x}_i, \mathbf{x}) \quad \forall i \quad (3.35)$$

O mais interessante é que se pode usar o *kernel* $K(\mathbf{x}, \mathbf{x}_i)$ para construir o hiperplano ótimo no espaço de características sem ter que se considerar o próprio espaço de características de forma explícita. Isto pode ser visto substituindo-se a equação 3.34 em 3.33 que resulta em uma nova equação para definição do hiperplano:

$$\sum_{i=1}^N \alpha_i d_i K(\mathbf{x}, \mathbf{x}_i) = 0. \quad (3.36)$$

Assim, a arquitetura do SVM pode ser ilustrada na Figura 3.9 por meio de funções

kernel.

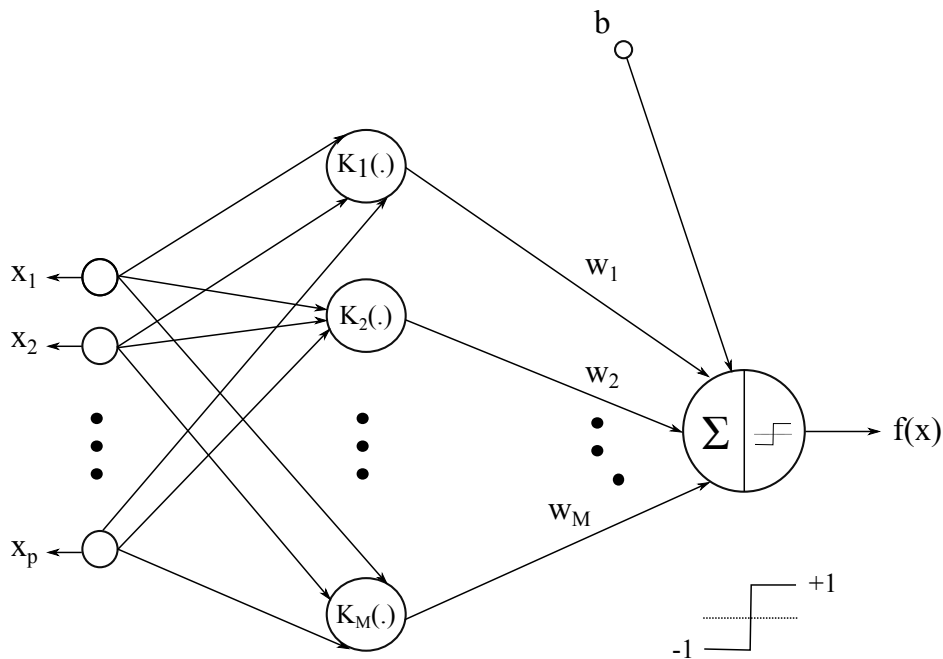


Figura 3.9: Arquitetura do SVM, em que $K(\cdot)$ é a função *kernel*, M é o número de vetores de suporte e P é a dimensão da entrada.

As funções *kernel* podem ser construídas de diversas maneiras, mas é desejável que elas satisfaçam ao *Teorema de Mercer* (Haykin 2008). Os tipos de *kernel* utilizados nessa tese são:

- *Kernel* Linear

$$K(\mathbf{x}, \mathbf{x}_i) = \mathbf{x}^T \mathbf{x}_i, \quad (3.37)$$

em que \mathbf{x} é o padrão de entrada.

- *Kernel* RBF (Rede de Função de Base Radial)

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2}\right), \quad (3.38)$$

em que o parâmetro σ^2 é especificado *a priori* pelo usuário.

Classificação da Variabilidade da Frequência Cardíaca por meio de Sequências Simbólicas

“(..) C’est le temps que tu as perdu pour ta rose qui fait ta rose si importante. (...) L’essentiel est invisible pour les yeux. (...)”

Citation remaniée d’Antoine de Saint Exupéry dans *Le Petit Prince*¹

Como primeiras contribuições desta tese, geradas pelos principais resultados obtidos neste Capítulo, pode-se citar os trabalhos publicados no XXII Congresso Brasileiro de Engenharia Biomédica - CBEB 2010 e no XXIII Congresso Brasileiro de Engenharia Biomédica - CBEB 2012 intitulados: *Dinâmica Simbólica Aplicada à Análise e Classificação da Disfunção Autônoma da Variabilidade da Frequência Cardíaca na Doença de Chagas* (Rocha, Lemos e Aguirre 2010) e *Investigação de Características Baseadas em Dinâmica Simbólica e Padrões Estatísticos Ordinais na Classificação da Variabilidade da Frequência Cardíaca* (Sady e Aguirre 2012), respectivamente.

4.1 Introdução

A variabilidade da frequência cardíaca (VFC) reflete as interações complexas do sistema cardiovascular. Patologias cardiovasculares frequentemente se manifestam

¹Foi o tempo que perdeste com tua rosa que fez tua rosa tão importante. O essencial é invisível aos olhos.

por meio de mudanças na VFC. Conseqüentemente, a habilidade para classificar (ou distinguir entre) padrões de VFC fisiológicos e patológicos é importante para o desenvolvimento de novas ferramentas de diagnóstico. Entretanto, o sucesso na classificação de séries de intervalos RR depende da avaliação de características relevantes (Asl, Setarehdan e Mohebbi 2008; Gudmundsson e Sigurdsson 2012; Parlitz et al. 2012).

A dinâmica simbólica (DS) vem sendo utilizada, na literatura, para a análise (Kurths e Wessel 1995) e classificação da VFC (Parlitz et al. 2012). Os autores em (Kurths e Wessel 1995) utilizaram ferramentas baseadas em DS na análise dos sinais de VFC com o objetivo de discriminar entre comportamentos dinâmicos associados a indivíduos saudáveis e indivíduos com alguma patologia ou síndrome relacionada ao sistema cardiovascular. Outras características baseadas em sequências de símbolos, tais como os padrões estatísticos ordinais (PEO's), foram recentemente aplicadas na classificação do ritmo cardíaco (Parlitz et al. 2012). Entretanto, mais estudos são necessários para a melhoria no desempenho destas técnicas na identificação de diferentes padrões da VFC para a distinção entre diversas classes. Diante desse panorama, retoma-se o problema de classificação inicialmente investigado em (Rocha, Lemos e Aguirre 2010), no qual também utilizou-se apenas a DS para fazer a distinção de padrões de VFC em pacientes com doença de Chagas, conseguindo diferenciar entre duas classes: pacientes controle e chagásicos. Entretanto, nesse estudo não foi possível a distinção entre três classes (controle, chagásicos cardiopatas e chagásicos não cardiopatas). Assim, o uso de novas características baseadas em DS e PEO's aliadas aos índices clássicos tempo-frequenciais parecem ser bons indicadores para a classificação da VFC em pacientes chagásicos.

Neste trabalho propõe-se investigar o uso de características baseadas em DS e em PEO's, associadas aos índices clássicos tempo-frequenciais, na classificação de padrões da VFC de duas e três classes de pacientes chagásicos utilizando um Discriminante Linear, uma SVM e uma SVM multiclases.

4.1.1 Disfunção Autonômica

A permanente influência reguladora exercida pelo sistema nervoso sobre o funcionamento dos diversos órgãos e sistemas que integram o organismo é essencial para que o mesmo tenha assegurado o seu equilíbrio fisiológico interno e, adequadamente, exerça suas interações com o meio ambiente circundante. Essa regulação, automática e involuntária, é efetuada pela divisão autonômica do sistema nervoso, representada pelo sistema nervoso autônomo (SNA), por meio das suas duas subdivisões anátomo-funcionais: o sistema nervoso simpático e o sistema nervoso parassimpático.

O controle neuro-autonômico do sistema cardiovascular, em particular do coração, constitui importante processo homeostático orgânico, refletindo a extraordinária capacidade de adaptação fisiológica que este sistema deve ter, de momento a momento, com vistas ao atendimento das necessidades metabólicas do organismo em diversas circunstâncias funcionais. Tal peculiar capacidade de adaptação do sistema cardiovascular depende de complexa interação entre os seus diferentes componentes e o sistema nervoso, a qual tem por base uma intrincada estrutura nervosa que envolve um conjunto de vias e centros neurais centrais e periféricos, e grande multiplicidade de neurotransmissores (Berne et al. 2004; Kandel, Schwarts e Jessel 2000).

Por disfunção autonômica cardíaca (DAC), ou disautonomia cardíaca, entende-se um distúrbio funcional, de natureza primária ou secundária, resultante de alterações puramente funcionais ou orgânicas localizadas em um ou em ambos os componentes do SNA, em qualquer das suas estruturas - vias aferentes, centros encefálicos e vias eferentes. Esse distúrbio representa importante e comum condição fisiopatológica e tem significativa implicação clínica, terapêutica e prognóstica, refletindo-se até mesmo sobre a morbi-mortalidade cardiovascular. O reconhecimento e a valorização da DAC constitui-se um fator essencial para a compreensão e melhor controle de muitas condições clínicas e funcionais que interferem na inervação autonômica do coração. Assim, deve-se entender por função autonômica cardíaca, o estado do controle simpático e parassimpático exercido sobre o automatismo sinusal, e não sobre o conjunto das demais propriedades do coração (Junqueira 1998).

A DAC manifesta-se por alteração isolada ou combinada, absoluta ou relativa, das atividades simpática e parassimpática, e está associada com muitos dos processos patológicos e distúrbios funcionais do próprio coração e de outros órgãos, e com condições de natureza sistêmica que comprometem difusamente o SNA. O substrato patológico da DAC inclui alterações inflamatórias, isquêmicas, degenerativas, tóxicas e metabólicas, agudas e crônicas, da inervação intrínseca do coração, de estruturas aferentes, eferentes e centrais integrantes dos reflexos cardiovasculares, assim como de estruturas do sistema nervoso central (SNC) e periférico ligados ao SNA. Neoplasias do SNC e alterações degenerativas ou metabólicas sistêmicas também podem resultar em DAC, pelo envolvimento de componentes do SNA. As disfunções simpática e parassimpática podem ser os fatores causais de diferentes condições clínicas e funcionais, podendo, também, ser consequências secundárias de várias outras. Como exemplos de condições que podem induzir secundariamente DAC aguda, incluem-se, o botulismo, a intoxicação alcoólica e a intoxicação por diferentes substâncias, tais como metais pesados, solventes orgânicos e drogas antineoplásicas (Junqueira 1998). Entre as condições que podem determinar alterações secundárias crônicas da função autonômica cardíaca incluem-se, a doença de Chagas, o *diabetes mellitus*, a insuficiência cardíaca, o infarto do miocárdio, a insuficiência coronariana crônica, a hipertensão arterial, a insuficiência renal crônica e as doenças cerebrais e síndromes neurológicas (Cacoub 2006; Angelis et al. 2007; Lombardi et al. 2001; Parisotto et al. 2008; Ribeiro et al. 2005; Schroeder et al. 2003).

A justificativa para a avaliação clínica da função autonômica cardíaca tem por base o fato de que o estado de equilíbrio entre as atividades simpática e parassimpática exercidas sobre o coração investe-se de importante significado funcional em diversas condições clínicas e funcionais, e pode ser determinante de manifestações cardiovasculares e outras. Pode ter também relevantes implicações terapêuticas, determinando a escolha de drogas com efeitos autonômicos, e ser fator prognóstico e influenciador independente da morbi-mortalidade cardiovascular.

No contexto clínico, a caracterização da função autonômica cardíaca baseia-se usualmente na avaliação da variabilidade espontânea ou induzida da frequência cardíaca

(FC), a qual traduz as influências tônicas e fásicas das atividades simpática e parassimpática exercidas sobre o nó sinusal.

4.1.2 Disautonomia Parassimpática e Disfunção Ventricular Esquerda

A grande maioria dos estudos de função autonômica avaliou o controle parassimpático cardíaco, indicando que esta divisão do SNA é acometida intensa e precocemente na doença de Chagas. Para os autores de (Marin-Neto et al. 1980; Marin-Neto et al. 1986), as evidências de acometimento simpático são discretas e menos pronunciadas do que o acometimento vagal. Esses autores observaram menor aumento da frequência cardíaca tardia após inclinação passiva em chagásicos, quando comparados à indivíduos normais, indicando acometimento simpático. Porém, os autores deste grupo não encontraram, entre os chagásicos, alterações significativas na resposta simpática ao esforço ou na resposta pressórica a manobras como o esforço isométrico ou ao teste da inclinação passiva.

A existência de comprometimento do sistema nervoso simpático, cardíaco e vascular, precoce e significativo, foi postulado principalmente pelos autores de (Iosa, Caeiro e Palmero 1980; Iosa et al. 1989; Guzzetti et al. 1990; Guzzetti et al. 1991) baseando-se em:

- provas funcionais autonômicas,
- menor aumento da norepinefrina circulante em relação a pacientes com insuficiência cardíaca de outras etiologias, e
- ausência de aumento do componente espectral de baixa frequência, à análise da VFC, durante teste postural ativo e esforço isométrico, sugerindo acometimento simpático.

Os resultados divergentes podem estar relacionados à maior complexidade estrutural do SNA simpático, que dificultaria o reconhecimento precoce do acometimento. Além disto, reconhece-se a existência de limitações metodológicas significativas, relacionadas a dificuldades na avaliação quantitativa do controle simpático cardiovas-

cular. Neste contexto, a análise da VFC por meio da dinâmica não-linear (DNL) tem se mostrado ferramenta particularmente útil no reconhecimento tanto de disfunções simpáticas como vagais. Os estudos em (Moraes et al. 2000) feitos por meio do mapa de primeiro retorno tridimensional indicaram presença de alterações apenas nos índices “vagais” do mapa de primeiro retorno, sem alterações dos índices “simpáticos”. Tais achados sugerem que a disfunção autonômica na doença de Chagas é realmente predominantemente vagal. Confirmando esta conjectura, os estudos de (Ribeiro et al. 2001; Ribeiro et al. 2005) mostraram que há uma aparente disautonomia parassimpática precedente à disfunção ventricular esquerda. Em contrapartida, (Simões et al. 2000) encontraram evidências de desnervação simpática em chagásicos utilizando-se da técnica da metaiodobenzilguanidina.

Por outro lado, o estudo da resposta ao esforço tem mostrado que vários pacientes com doença de Chagas apresentam insuficiência cronotrópica significativa: postula-se que esta resposta anormal ao exercício se relacione à disfunção simpática. Estudos realizados em (Rocha et al. 2001), entretanto, foram incapazes de identificar alterações específicas da função ventricular esquerda ou da VFC em pacientes com insuficiência cronotrópica quando comparados àqueles cronotropicamente suficientes.

A disfunção sistólica global do ventrículo esquerdo é o principal indicador de prognóstico desfavorável na doença de Chagas. Pacientes com depressão da fração de ejeção podem morrer por progressão da insuficiência cardíaca, mas também de morte súbita ou por acidente vascular cerebral, uma vez que a disfunção ventricular esquerda é fator predisponente também no aparecimento de arritmias potencialmente fatais e fenômenos trombo-embólicos. De um modo geral, cerca de 50% dos pacientes com disfunção sistólica global do ventrículo esquerdo são assintomáticos e, na prática clínica, a ecocardiografia é o método diagnóstico mais comumente utilizado no reconhecimento da disfunção ventricular esquerda. Entretanto, a disponibilidade limitada de recursos e a elevada frequência de pacientes com função ventricular esquerda normal impossibilitam a utilização deste método em toda a população de chagásicos (Ribeiro et al. 2001; Ribeiro et al. 2005).

A dosagem do *peptídeo natriurético tipo B* (BNP) tem sido preconizada como método

de escolha para o rastreamento da disfunção sistólica do ventrículo esquerdo, tanto no contexto do cuidado primário como na sala de urgência. Evidências experimentais iniciais sugerem que os peptídeos natriuréticos participam ativamente na fisiopatologia da doença de Chagas. Estudos recentes mostraram que chagásicos com depressão significativa da fração de ejeção do ventrículo esquerdo apresentam níveis elevados de BNP (Ribeiro et al. 2002a).

O impacto da utilização rotineira da dosagem do BNP na prática clínica ainda é desconhecido. Inicialmente, em decorrência do fato de que o método ainda não está comercialmente disponível, a um custo razoável. Estudos iniciais em (Ribeiro et al. 2001; Ribeiro et al. 2005) sugerem que a radiografia de tórax não acrescenta dados diagnósticos significativos ao eletrocardiograma e que poderia ser substituído, com vantagens pelo BNP. Esta nova estratégia (“regra de predição clínica”) precisa ser formalmente testada em novas amostras para se revelar efetiva ou não. Por fim, o valor prognóstico da elevação do BNP, estabelecido em outras condições clínicas, ainda não está validado na doença de Chagas.

4.1.3 Disfunção Autonômica e Prognóstico por meio de Técnicas de Análise da Variabilidade da Frequência Cardíaca

A disfunção autonômica, uma anormalidade característica da doença de Chagas, tem sido implicada como fundamental na patogênese desta doença. A análise da VFC é uma ferramenta estabelecida para o estudo do controle autonômico cardíaco, reconhecendo pacientes de alto risco para morte em diferentes condições clínicas. Em uma série de estudos realizados nos últimos anos, (Ribeiro et al. 2001; Ribeiro et al. 2002b; Ribeiro et al. 2003; Ribeiro et al. 2005) descreveram o comportamento da VFC nas diferentes formas clínicas da doença de Chagas, durante manobras curtas, ao *Holter* de 24 horas e utilizando-se de novas técnicas não-lineares. As conclusões obtidas nestes estudos, até o momento, podem ser assim sumariadas:

- “Existem evidências consistentes de alteração do controle vagal do coração, avaliado de diferentes formas, em pacientes chagásicos sem cardiopatia aparente,

independente da presença de dilatação ventricular esquerda ou diminuição da fração de encurtamento do ventrículo esquerdo.

- Índices convencionais de curta duração ao Holter de 24 horas, de origem vagal, estão reduzidos em todas as formas clínicas da doença, indicando que a lesão autonômica é primária, em contraposição a afirmação de outros autores (Davila, Inglessis e Mazzei 1998).
- Índices lineares globais ao Holter de 24 horas não se alteram na doença de Chagas, ao contrário de outras cardiopatias; entretanto, índices não-lineares globais estão alterados, indicando perda da dinâmica fractal, também independentemente da depressão da função ventricular esquerda.
- Como o BNP está elevado apenas em pacientes com depressão ventricular esquerda significativa, os dados de (Ribeiro 1996) comprovam que não é a ativação neurohumoral a causa da disfunção autonômica na doença de Chagas: pacientes com fração de ejeção normais podem ter índices alterados da VFC, mesmo com valores normais de BNP.
- O papel dos anticorpos contra receptores autonômicos está sendo investigado.”

Assim, boa parte dos questionamentos feitos em estudos anteriores acerca da disfunção autonômica na doença de Chagas não é mais objeto de dúvida, tendo sido esclarecido em estudos ou análises prévias. Entretanto, duas questões permanecem ainda não respondidas: o papel da disfunção simpática na patogênese da doença de Chagas e o impacto prognóstico da disfunção autonômica.

Em outras condições clínicas, como após o infarto agudo do miocárdio e na insuficiência cardíaca, a diminuição da influência autonômica cardíaca, avaliada por meio de testes autonômicos simples ou por meio da análise da VFC, é um marcador forte e independente de risco aumentado de morte, principalmente por arritmias cardíacas (Task-Force 1996; Figueiredo et al. 1996). Na doença de Chagas, postula-se um possível papel da disfunção autonômica na gênese da morte súbita (Ribeiro 1996), possivelmente pela ausência do controle simpático-vagal preciso em situações de mudança

brusca das condições cardiocirculatórias, como na emoção e no exercício (Sousa et al. 2006). Evidências preliminares de que a redução da VFC pode ter importância prognóstica na cardiopatia chagásica foram relatadas recentemente (Ribeiro et al. 2008). Assim a definição da importância de uma ferramenta de auxílio no prognóstico da disfunção autonômica na doença de Chagas é um dos objetivos maiores deste estudo.

4.1.4 Variabilidade da Frequência Cardíaca e Dinâmica Simbólica

A importância do acometimento do SNA nas doenças cardiovasculares tem sido objeto de intensas investigações. A avaliação funcional do controle do SNA sobre o coração pode ser feita por meio de testes autonômicos, nos quais se observa a resposta reflexa fisiológica à aplicação de um estímulo quantificável, fisiológico ou farmacológico, como a respiração, o exercício e a injeção de atropina e fenilefrina. Alternativamente, informações sobre o controle autonômico cardíaco podem ser obtidas pela observação da variação intrínseca da FC, tanto em registros curtos, de dois a cinco minutos de repouso, como em traçados prolongados, de 24 horas, durante as atividades habituais. O sinal da variabilidade da frequência cardíaca (VFC) consiste de uma série dos intervalos de tempo normais entre as ondas R do eletrocardiograma (ECG) - série de intervalos RR. Pode-se também considerar a série de valores de frequência cardíaca instantânea (tacograma) que nada mais é do que uma série de intervalos RR inversos. A análise da VFC é uma técnica não-invasiva que parte do princípio que, em condições normais, o ritmo cardíaco modifica-se em resposta a estímulos diversos, como exercício e estresse mental, ou mesmo em condições de repouso, flutuando em torno de uma média. Tal variabilidade relaciona-se, predominantemente, às alterações contínuas do balanço simpato-vagal, em resposta a mecanismos de controle cardiovascular (Gomes 2001; Junqueira 1998; Vermeiren 1996).

Uma possível compressão de informação para o estudo de sistemas dinâmicos contínuos seria o estudo de um sistema em tempos discretos, de forma que o estado do sistema só é analisado em tempos determinados. A dinâmica simbólica pode ser vista como uma codificação do sistema, em que não só o tempo é discretizado como

também o espaço (Lind e Marcus 1995). Sua idéia básica consiste em dividir o espaço de estados ou uma seção de Poincaré do sistema em um número finito de regiões disjuntas, associando um símbolo a cada uma destas. Dessa forma, o estudo do sistema se resume em analisar a variação destes símbolos ao longo do tempo.

O mapa de primeiro retorno do tacograma pode ser visto como uma analogia a uma uma seção de Poincaré do sistema analisado (Janson e Anishchenko 1998). Dessa forma, diversos trabalhos propostos na literatura utilizam ferramentas baseadas em dinâmica simbólica na análise dos sinais de variabilidade da frequência cardíaca com o objetivo de discriminar entre comportamentos dinâmicos associados a indivíduos saudáveis e indivíduos com alguma patologia ou síndrome relacionada ao sistema cardiovascular (Kurths e Wessel 1995; Wessel et al. 2006; Wessel et al. 2007b; Voss e Caminal 2008).

Os autores de (Kurths e Wessel 1995) podem ser considerados os primeiros a estudarem a VFC utilizando ferramentas baseadas no conceito de dinâmica simbólica. Nesse trabalho, essas ferramentas são utilizadas para inferir sobre o risco de morte súbita cardíaca em indivíduos que sofreram infarto do miocárdio por meio da análise dos intervalos RR. A morte súbita cardíaca é causada, em muitos casos, por arritmias ventriculares. Os autores de (Kleiger, Stein e Bigger 2005) associam uma diminuição da VFC a um aumento no risco de morte súbita em indivíduos que sofreram infarto agudo do miocárdio. Dessa maneira, o trabalho proposto em (Kurths e Wessel 1995) tenta associar essa diminuição a uma mudança estrutural na dinâmica dos intervalos entre batimentos (entre os intervalos RR). Essa mudança estrutural é medida por meio de uma análise estatística de sequências de símbolos gerados pela partição estatística do mapa de primeiro retorno de tacogramas. Além disso, foram utilizadas técnicas tradicionais no domínio do tempo e da frequência com o objetivo de comparar os resultados obtidos e construir uma técnica híbrida com alto rendimento.

Em (Kurths e Wessel 1995) foram analisados tacogramas de indivíduos referentes a três grupos distintos: indivíduos saudáveis e indivíduos que sofreram infarto do miocárdio com baixo e alto risco de morte súbita. Cada tacograma foi extraído de sinais de ECG obtidos durante 30 a 60 minutos. Com relação a análise estatística baseada na

dinâmica simbólica, os mapas de primeiro retorno dos tacogramas foram codificados em 4 símbolos por meio de duas partições distintas, definidas nas transformações seguintes:

$$s_i = \begin{cases} 0, & t_i > (1+a)\mu, \\ 1, & \mu < t_i \leq (1+a)\mu, \\ 2, & (1-a)\mu < t_i \leq \mu, \\ 3, & t_i \leq (1-a)\mu, \end{cases}$$

sendo que t_i é o valor do intervalo temporal correspondente ao evento i , μ é o intervalo RR médio e a é um parâmetro definido empiricamente como 0,1.

$$\tilde{s}_i = \begin{cases} 0, & \Delta_{t_i} > 1,5\sigma_{\Delta}, \\ 1, & 0 < \Delta_{t_i} \leq 1,5\sigma_{\Delta}, \\ 2, & -1,5\sigma_{\Delta} < \Delta_{t_i} \leq 0, \\ 3, & \Delta_{t_i} \leq -1,5\sigma_{\Delta}, \end{cases}$$

onde $\Delta_{t_i} = t_{i+1} - t_i$ e σ_{Δ} é o desvio padrão de Δ_{t_i} .

Em seguida, calculou-se a frequência de cada uma das $4^3 = 64$ sequências de símbolos possíveis de comprimento 3. Esse comprimento foi definido empiricamente, como um bom compromisso entre a inclusão de alguma dinâmica e a credibilidade no cálculo das frequências de cada sequência. A partir dessas frequências, calculou-se o número de palavras proibidas presentes na dinâmica, definido como o número de palavras com probabilidade de ocorrência menor que 0,001. Além disso, três medidas de entropia foram calculadas a partir do histograma normalizado das palavras; utilizou-se a *Entropia de Shannon*, a *Entropia de Renyi* e uma terceira medida de entropia denominada *Entropia Renormalizada*. Finalmente, as medidas de entropia e número de palavras proibidas foram utilizadas, juntamente com as medidas derivadas de técnicas no domínio do tempo e da frequência, para inferir sobre o risco de morte súbita cardíaca.

As técnicas do domínio do tempo utilizadas são baseadas em estatísticas extraídas dos tacogramas, tais como média de intervalos NN (intervalos de batimentos normais) medidas a cada 1 e 5 minutos e proporção das diferenças entre intervalos NN con-

secutivos maiores que 50 e 100 milissegundos. No domínio da frequência, o espectro de densidade de potência foi determinado e a potência de quatro faixas de frequência foram analisadas, assim como a razão da faixa LF e todo o espectro e a razão LF/HF (LF = 0,04 ··· 0,15 Hz e HF = 0,15 ··· 0,40 Hz), de forma que cada faixa analisada está relacionada a diferentes fenômenos fisiológicos. As técnicas baseadas em dinâmica não-linear apresentaram a melhor taxa de detecção quando comparadas às técnicas no domínio do tempo e da frequência, quando comparadas separadamente. As melhores taxas de detecção foram observadas utilizando-se as técnicas não-lineares, número de palavras proibidas e medidas de entropia baseadas na Entropia de *Renyi* e na Entropia Renormalizada, juntamente com as técnicas no domínio da frequência.

A análise da VFC parece ser um método promissor para permitir o reconhecimento de pacientes de risco, passíveis de intervenções terapêuticas. Além disso, ainda há necessidade de estratégias de reconhecimento de pacientes com pior prognóstico e de identificação daqueles com disfunção ventricular esquerda, pois a importância prognóstica da disfunção autonômica na doença de Chagas é desconhecida (Ribeiro 1996). Assim, o estudo da VFC na doença de Chagas poderá esclarecer questões controversas sobre papel da disfunção simpática nesta patologia e contribuir como possível fator preditor no prognóstico e diagnóstico desta cardiopatia. Diante desse panorama, o presente trabalho pretende verificar o desempenho de algumas ferramentas de análise em problemas de classificação. Espera-se, portanto, que este trabalho possa contribuir com estudos futuros para o desenvolvimento de protocolos médicos que possam ser padronizados e aplicados na clínica médica para o diagnóstico e prognóstico de doenças cardiovasculares que estão intimamente relacionadas com a variação do ritmo cardíaco.

Este Capítulo está organizado da seguinte maneira: a Seção 4.3 (*Materiais e Métodos*) descreve o protocolo experimental, o pré-processamento do sinal utilizado e a extração de características para treinar os classificadores baseados em SVM e em Discriminante Linear. Os resultados são apresentados na Seção 4.4 (*Resultados*) e discutidos na Seção 4.5 (*Discussão*) e na Seção 4.6 (*Conclusões do Capítulo*) as principais conclusões são fornecidas. O Capítulo 5 também possui a mesma estrutura. A escolha dessa organi-

zação se justifica pelo fato de que há dois problemas de classificação advindos de bases de dados diferentes, logo, parte da metodologia para a resolução desses problemas também é diferente. Assim, optou-se pela divisão da metodologia/protocolo, resultados e discussão em dois Capítulos a fim de deixar o texto mais claro para o leitor.

4.2 Objetivos

- O objetivo inicial deste Capítulo consiste, então, em distinguir entre os registros dos pacientes do grupo controle, dos indivíduos chagásicos cardiopatas e dos indivíduos chagásicos não-cardiopatas. Assim, a metodologia desenvolvida no trabalho proposto foi elaborada com o intuito de classificar esses três grupos.
- Como descrito neste texto, o diagnóstico da Doença de Chagas requer uma série de exames laboratoriais, tais como exames sorológicos dos pacientes, técnicas de cintilografia e dosagem de BNP, que são necessários para o diagnóstico e para a evolução desta patologia. Assim, a construção uma ferramenta de auxílio no diagnóstico da Doença de Chagas e no prognóstico da cardiopatia chagásica, baseada na análise da VFC por meio dos índices tempo-frequenciais, da dinâmica simbólica e de padrões estatísticos ordinais, teria grande relevância na clínica médica. Além disso, essa seria uma técnica não-invasiva de baixo custo.

4.3 Material e Métodos

4.3.1 Protocolo Experimental

O banco de dados utilizado é constituído de 66 registros *Holter* de 24h obtidos no Hospital das Clínicas da UFMG (HC-UFMG). Foram consideradas as recomendações da Organização Mundial de Saúde e da Declaração de Helsinque de 1975, assim como a Resolução 196/96 sobre Pesquisa envolvendo Seres Humanos do Ministério da Saúde, no sentido de se salvaguardarem os direitos e o bem-estar das pessoas estudadas. Os exames propostos foram realizados apenas após a obtenção do consentimento, por escrito,

dos indivíduos que aceitaram participar do estudo. Durante a realização destes, foram tomadas todas as medidas possíveis para se reduzir qualquer desconforto inerente aos procedimentos. Ao fim da avaliação, todos os participantes do estudo receberam, por escrito, os resultados dos exames realizados, assim como as recomendações e encaminhamentos apropriados. Os procedimentos, em questão, foram aprovados pelo Comitê de Ética em Pesquisa do Hospital das Clínicas da UFMG em 14 de maio de 1997, como projeto do prof. Antônio Luiz Pinho Ribeiro intitulado “Disfunção autonômica na Doença de Chagas - mecanismos e implicações prognósticas” (Ribeiro 1996).

Seleção dos pacientes

Foram selecionados os pacientes, chagásicos ou não, na faixa de 15 a 55 anos, que não se enquadraram nos critérios de exclusão (Ribeiro 1996; Gomes 2001), listados a seguir, e cuja avaliação inicial não sugeria a presença de outras patologias cardíacas ou sistêmicas significativas.

“Para a realização do recrutamento, os pacientes atendidos no referido ambulatório foram submetidos à anamnese e exame físico padronizados, exames laboratoriais, eletrocardiograma e telerradiografia de tórax. A avaliação clínica foi direcionada para a detecção de evidências da cardiopatia ou de patologias subjacentes que tornassem o paciente inabilitado para a realização dos procedimentos seguintes. Todos os pacientes foram submetidos a eletrocardiograma (ECG) de 12 derivações, que analisado de forma cega e padronizada, calculando-se a FC, medindo-se a duração da onda P, do complexo QRS e dos intervalos PR e QT, corrigindo-se o último pela FC, utilizando-se a *fórmula de Bazet*. Avaliou-se especificamente as alterações da morfologia e do ritmo, classificando-as conforme os *Crítérios de diagnóstico eletrocardiográfico en la cardiopatia chagásica crónica*². A telerradiografia de tórax em PA e perfil esquerdo foi realizada conforme a rotina do setor de radiologia, no dia do recrutamento, com análise realizada de forma cega, completa e padronizada. Foram efetuados hemograma completo, dosagem bioquímica de creatinina, uréia, glicemia de jejum, potássio e sódio, dosagem de TSH e exames sorológicos para infecção pelo *Trypanosoma cruzi* por pelo menos três

²WHO - Expert Committee on the Control of Chagas Disease (1991): Control of Chagas Disease, 1-95.

métodos diferentes, realizados no Laboratório Central do HC-UFMG.

Foram considerados chagásicos os indivíduos que apresentavam positividade sorológica para *Trypanosoma cruzi* por meio de duas ou mais técnicas diferentes, incluindo as reações de imunofluorescência indireta (IFI), hemaglutinação indireta (HAI), de fixação de complemento (RFC) e ELISA, em exames realizados no Laboratório Central do HC-UFMG ou no Laboratório de Parasitologia da Fundação Ezequiel Dias. Foram considerados não-chagásicos indivíduos que não apresentaram reação sorológica positiva, quando realizadas pelo menos três técnicas distintas, no Laboratório Central do HC-UFMG ou no Laboratório de Parasitologia da Fundação Ezequiel Dias.

Foram considerados critérios de exclusão:

1. "Impossibilidade ou ausência de disponibilidade para a realização dos exames.
2. Hipertensão arterial sistêmica, definida operacionalmente como:
 - pressão arterial medida durante o exame físico superior a 160/95 mmHg, ou
 - pressão arterial medida durante o exame físico entre 140-159/90-94 mmHg, associado à história de hipertensão arterial sistêmica, ou quarta bulha ao exame físico, ou sobrecarga ventricular esquerda provável ao ECG pelo critério de Romhilt-Estes, ou evidências de dilatação aórtica à radiografia de tórax.
3. História compatível com doença arterial coronariana, conforme anamnese.
4. Episódio prévio sugestivo de doença reumática aguda.
5. Diabetes mellitus ou tolerância reduzida à glicose, conforme definido pelo *National Diabetes Data Group*.
6. Disfunção tireoidiana, manifesta por níveis anormais de hormônio estimulante da tireóide (TSH) e da tiroxina (T4) livre.
7. Insuficiência renal, definida pelo aumento dos níveis de creatinina e uréia.

8. Doença pulmonar obstrutiva crônica, conforme presença de história, exame físico, ECG e alterações radiológicas sugestivas.
9. Distúrbios hidroeletrólíticos: níveis séricos anormais de potássio e sódio.
10. Anemia significativa, definida como hemoglobina menor que 10 g/dl.
11. Alcoolismo definido como consumo médio semanal acima de 420 g de etanol (média diária acima de 60 g de etanol).
12. Qualquer outra doença sistêmica significativa, crônica ou aguda.
13. Gravidez, definida por critérios laboratoriais.”

Classificação dos pacientes em subgrupos

Os pacientes que não preencheram os critérios de exclusão (Ribeiro 1996; Gomes 2001) e não foram excluídos durante a avaliação inicial, por meio dos métodos clínicos e laboratoriais citados acima, foram classificados em três grupos, com vistas ao estudo das respostas cardiovasculares aos métodos de avaliação específica e do impacto prognóstico das anormalidades encontradas:

- grupo 1: composto por registros de indivíduos saudáveis (grupo controle);
- grupo 2: composto por registro de indivíduos chagásicos com fração de ejeção > 50% ($FE > 50\%$), ou seja, indivíduos chagásicos sem cardiopatia aparente.
- grupo 3: composto por registro de indivíduos chagásicos com fração de ejeção < 50% ($FE < 50\%$), ou seja, indivíduos chagásicos cardiopatas.

4.3.2 Aquisição das séries e Pré-processamento

As séries de intervalos RR foram obtidas a partir da realização de eletrocardiografia dinâmica (sistema *Holter*). As gravações foram realizadas pelo período nominal de 24h, com o paciente mantendo as suas atividades habituais. Os dados foram considerados satisfatórios quando apresentaram pelo menos 18h de traçado de boa qualidade e

mais de 85% de batimentos cardíacos sinusais. As gravações foram analisadas no sistema Holter DMI/Burdick 8400 por meio de técnicas semi-automáticas que permitem a detecção de batimentos normais, artefatos e ectopias. O sistema gera um arquivo com extensão *ooc* com o registro de dados do paciente. Os registros foram convertidos da extensão *ooc* para o formato de arquivo com extensão *.txt*, segundo metodologia desenvolvida em (Gomes 2001).

Foram selecionadas 22 séries de cada grupo. Os intervalos RR foram extraídos de cada um dos registros de ECG utilizando-se um algoritmo automático de localização dos complexos QRS, por meio do sistema *Holter*. O sistema Holter marca os batimentos ectópicos com -1. Em seguida foram removidos os batimentos ectópicos e foi feita uma interpolação linear entre os intervalos RR adjacentes a fim de estimar os valores da série associados às possíveis ectopias.

4.3.3 Extração de Características

Nesta seção serão introduzidas as características extraídas dos sinais de VFC, derivadas dos índices tempo-frequenciais, da dinâmica simbólica e de padrões estatísticos ordinais. Foram escolhidas 14 características para treinar o classificador. A escolha foi empírica e baseada na literatura (motivos históricos). Cada índice selecionado era introduzido na matriz de características e avaliava-se a taxa de acerto durante o treinamento. Os índices com maiores taxas de acertos eram selecionados. Além disso, os índices tempo-frequenciais foram utilizados em diversos trabalhos da literatura em problemas de classificação de sinais de ECG (Parlitz et al. 2012). Inicialmente, trabalhou-se apenas com características relacionadas com a dinâmica simbólica e, posteriormente, utilizou-se também características relacionadas com os índices tempo-frequenciais e com padrões estatísticos ordinais.

Extração de Características por meio de Índices Tempo-Frequenciais

Os índices do domínio do tempo são baseados em métodos estatísticos derivados dos intervalos RR, bem como da diferença entre eles. Esses parâmetros podem ser

calculados para janelas (épocas) de 5 min e de 24 h, representando a variabilidade de curta e longa duração, respectivamente (Parlitz et al. 2012; Task-Force 1996). Os índices temporais utilizados, neste trabalho, foram:

- mNN: média dos intervalos RR;
- SDNN: desvio padrão dos intervalos RR;
- SDANN: desvio padrão das médias dos intervalos RR normais de segmentos de 5 min;
- RMSSD: raiz quadrada da média da soma das diferenças quadráticas entre intervalos RR adjacentes;
- pNN50: porcentagem de intervalos RR cuja diferença com o intervalo anterior exceda 50 ms;
- MSD: média das diferenças sucessivas de intervalos RR adjacentes.

Os índices do domínio da frequência baseiam-se em componentes periódicas das séries de VFC. Existem diferentes técnicas para a análise espectral, dentre elas: métodos baseados na transformada rápida de Fourier, em modelos autorregressivos e na decomposição *wavelet*. A densidade espectral de potência (PSD) é, então, estimada em faixas de altas (0,15 – 0,4 Hz) e baixas (0,04 – 0,15 Hz) frequências (Task-Force 1996). A PSD na faixa de altas frequências (HF) representa a modulação da atividade vagal, enquanto a PSD na faixa de baixas frequências (LF) representa tanto atividade vagal quanto simpática. A razão entre altas e baixas frequências (HF/LF) representa o balanço simpato-vagal. Os índices frequenciais usados neste trabalho foram:

- TotPow: Variância dos intervalos RR;
- LF: Potência na faixa de baixas frequências;
- HF: Potência na faixa de altas frequências;
- LF/HF: Razão LF/HF.

Extração de Características por meio da Dinâmica Simbólica

Conforme descrito no Capítulo 3, para se analisarem comportamentos dinâmicos de sistemas por meio de ferramentas de dinâmica não-linear baseadas no conceito de dinâmica simbólica deve-se, inicialmente, definir uma partição do seu espaço de estados ou, no caso específico da VFC, de uma “seção de Poincaré do sistema”.

O mapa de primeiro retorno de um tacograma, geralmente, é formado por uma nuvem de pontos orientados ao longo da identidade x_1 ($RR_{i+1} = RR_i$). A dispersão de pontos perpendiculares a essa reta reflete a variabilidade do sinal de longa duração e é quantificada por $SD1$, definido como o desvio padrão ao longo da reta x_2 perpendicular a x_1 . Por outro lado, a dispersão ao longo de x_1 reflete a variabilidade de curta duração e é quantificada pelo desvio padrão ao longo desta reta $SD2$ (Kurths e Wessel 1995; Letellier 2008), conforme ilustrado na Figura 4.1.

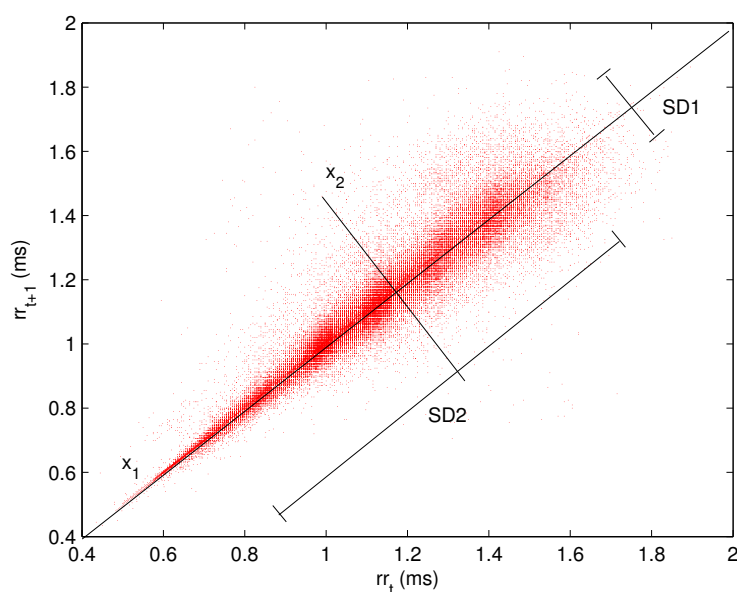


Figura 4.1: Mapa de primeiro retorno de um tacograma e as variabilidades dos pontos de curta e longa duração

Foram definidas, neste trabalho, duas partições estatísticas do mapa de primeiro retorno dos tacogramas. Primeiramente, definiu-se uma partição estatística com valores de limiar perpendiculares à reta x_1 , de forma que os símbolos resultantes descrevem

variações da dinâmica de sinais de VFC de longa duração. Essa partição pode ser definida por:

$$s_i = \begin{cases} 0, & RR_i \leq b_0, \\ 1, & RR_i \leq b_1, \\ 2, & RR_i \leq b_2, \\ \vdots & \vdots \\ q-1, & RR_i \leq b_{q-1}, \end{cases}$$

sendo que os parâmetros b_k , $k = 0, \dots, q-1$ são definidos de forma que os símbolos resultantes sejam equiprováveis. Esse tipo de partição equiprovável é definida como uma partição estatística do mapa de primeiro retorno (Godelle e Letellier 2000; Letellier 2008). O parâmetro q define o número de partições, ou seja, o tamanho do alfabeto. A Figura 4.2 ilustra esta partição.

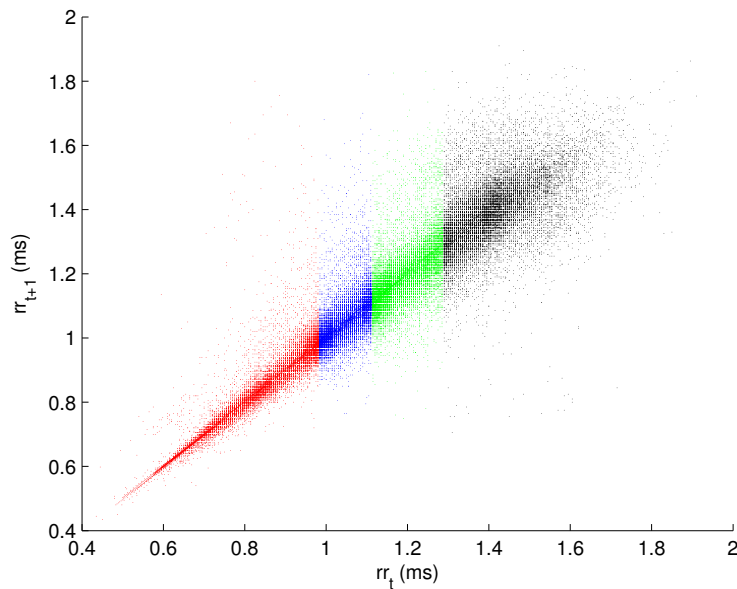


Figura 4.2: Partição estatística do mapa de primeiro retorno definida ao longo da reta perpendicular à reta x_1 .

Posteriormente, definiu-se outra partição estatística do mapa de primeiro retorno dos tacogramas com valores de limiar paralelos à reta x_1 , de forma que os símbolos

resultantes descrevam variações da dinâmica de sinais de VFC de curta duração. Esta partição está definida pela equação, abaixo:

$$s_i = \begin{cases} 0, & |RR_i - RR_{i+1}| \leq b_0, \\ 1, & |RR_i - RR_{i+1}| \leq b_1, \\ 2, & |RR_i - RR_{i+1}| \leq b_2, \\ \vdots & \vdots \\ q-1, & |RR_i - t_{i+1}| \leq b_{q-1}, \end{cases}$$

Da mesma maneira, os parâmetros b_k , $k = 0, \dots, q-1$ são definidos de forma que os símbolos resultantes sejam equiprováveis. A Figura 4.3 ilustra essa partição.

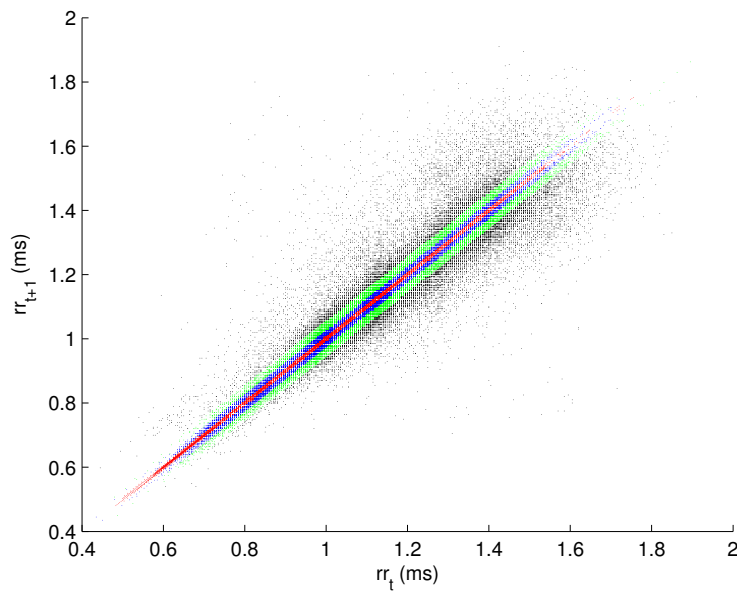


Figura 4.3: Partição estatística do mapa de primeiro retorno definida ao longo da reta paralela à reta x_1 .

Foram feitos experimentos, descritos nos Testes 1 e 2, com ambas as partições e verificou-se que os resultados obtidos com a primeira partição foram melhores, como era de se esperar, posto que os tacogramas obtidos do sistema *Holter* de 24h refletem a VFC de longa duração. Assim, optou-se por esta partição.

Uma vez definida a partição dos mapas de primeiro retorno, cada ponto do mapa é

convertido no símbolo correspondente, formando uma série de valores discretizados. A partir dessa série geram-se sequências de símbolos (S_i) de tamanho finito n por meio de um algoritmo de janelas deslizantes (deslocamento de Jacques-Bernoulli) (Monteiro 2006).

Dessa forma, definido-se um valor de n , pode-se estimar a função densidade de probabilidade (PDF) a partir da elaboração de um histograma de frequências relativas de cada uma das sequências. Para isso, as q^n sequências possíveis são convertidas para números na base q e ordenadas de acordo com sua ordem natural. Por exemplo, dado $q = 2$, $s_i = \{0,1\}$, e $n = 3$, a sequência 011 é atribuída ao número 3.

Segundo (Letellier 2008), os valores de q e n devem ser escolhidos de forma a garantir que as estatísticas permaneçam bem definidas, dessa forma, para um dado valor de q , deve-se limitar o valor de n de forma que $q^n \approx 1\%$ do número total de pontos do mapa de primeiro retorno (N). Como o valor médio de pontos, ou seja, o número total de cada série RR ($N + 1$), é aproximadamente $N = 70000$, para cada valor de q , variando de 2 a 8, limitou-se o valor máximo de n correspondente de forma que $q^n \approx 700$. Dessa forma, os dois experimentos realizados neste trabalho, foram executados para cada combinação de valores de q e n . Os melhores resultados obtidos foram para valores de $q = 3$ e $n = 6$, que são os apresentados neste trabalho.

Finalmente, a partir desse histograma é possível realizar a distinção entre comportamentos dinâmicos distintos, de forma que um histograma uniforme caracteriza um comportamento estocástico e um histograma contendo valores com alta probabilidade de ocorrência indica a presença de um processo aleatório com estrutura diferenciada; no último caso, o histograma pode ser o resultado de uma lei determinística. Sendo assim, a partir do histograma é possível inferir sobre a complexidade do sinal.

Deve-se, então, definir uma grandeza capaz de quantificar o nível de complexidade presente no histograma das sequências de símbolos. Para isso, utiliza-se a *Entropia de Shannon*, definida na equação, abaixo:

$$H_{ds} = - \sum_{i=0}^{q^n-1} p(s_i) \log_2(p(s_i)), \quad (4.1)$$

sendo $p(s_i)$ a frequência de ocorrência da palavra (ou sequência) de tamanho n relacionada ao valor i na base q . Essa grandeza assume valor nulo quando uma palavra tem frequência de ocorrência igual a um e possui valor máximo, $\log_2(1/q^n)$, quando todas as palavras são equiprováveis.

A Figura 4.4 esquematiza as etapas do processo de extração de características via DS.

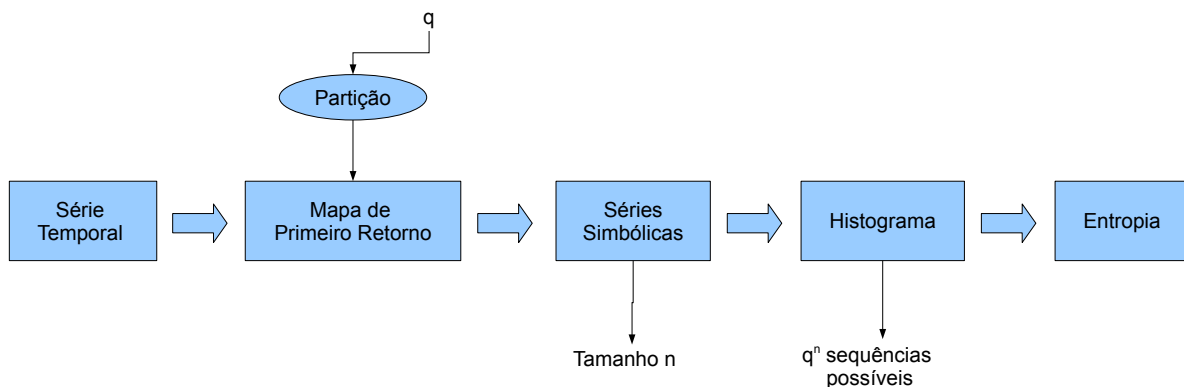


Figura 4.4: Etapas do processo de extração de características via DS.

Posteriormente, foram selecionadas as palavras ou sequências proibidas. As palavras proibidas (nPP) são as palavras que não ocorreram (ou não foram verificadas) no registro de dados utilizado. Porém, como o sistema é corrompido por ruído, as palavras proibidas são definidas como palavras com baixa frequência de ocorrência (com probabilidade de ocorrência menor que 0,1%) (Kurths e Wessel 1995; Gilmore 1998). Logo, é interessante também tentar classificar os sinais usando essa informação, uma vez que tanto o número de palavras proibidas quanto a entropia, caracterizam a “riqueza” do comportamento dinâmico do sinal. Assim, dois ou mais comportamentos dinâmicos diferentes poderiam ser discriminados com base nessas informações.

Extração de Características por meio de Padrões Estatísticos Ordinais

Os PEO's descrevem relações entre pequenos segmentos de comprimento l de uma dada série temporal. Assim como na DS a série é caracterizada por uma sequência simbólica. Padrões ordinais possuem diversas aplicações, como na análise da comple-

xidade de séries temporais e em problemas de classificação (Parlitz et al. 2012).

Dada uma série temporal x_1, x_2, \dots, x_N , a primeira etapa da abordagem via PEO's consiste na reordenação de amostras consecutivas $x_n, x_{n+1}, x_{n+2}, \dots, x_{n+q-1}$ em relação a sua amplitude de tal maneira que

$$x_{\pi(n)} \leq x_{\pi(n+1)} \leq x_{\pi(n+2)} \leq \dots \leq x_{\pi(n+q-1)} \tag{4.2}$$

computando, então, o índice de permutação do padrão π . Um único índice pode ser atribuído a cada padrão ordinal, interpretando a subsequência como uma permutação que é caracterizada por um índice de permutação. Assim como na DS, na abordagem por meio de PEO's, cada sequência temporal pode ser transformada numa sequência de permutação de índices (π_i) que pode ser formalmente considerada como uma sequência de símbolos com um alfabeto finito de tamanho $n!$. A Figura 4.5 ilustra a ideia básica dos PEO's.

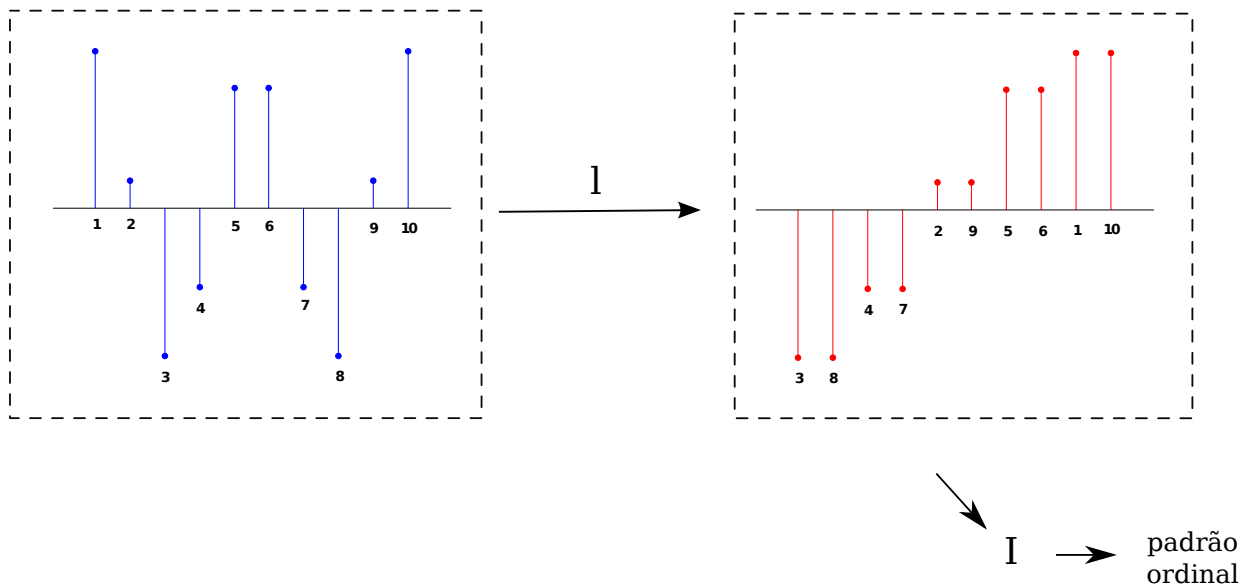


Figura 4.5: PEO's representados a partir da permutação das amostras do sinal, segundo sua amplitude.

O conceito de padrões ordinais pode ser estendido considerando não somente amostras consecutivas mas também subsequências com amostras $x_n, x_{n+\tau}, x_{n+2\tau}, \dots, x_{n+(q-1)\tau}$ que são separadas no tempo por um atraso τ , múltiplo do tempo de amostragem (T_s)

[8]. Para n diferentes palavras, há $n! = n \times (n-1) \times (n-2) \times \dots \times 2 \times 1$ padrões ordinais (π) possíveis que são também chamados de *permutações* (Bandt 2005). As probabilidades de ocorrência de padrões específicos com índices de permutação I para um dado atraso τ e comprimento n são usadas para caracterizar os intervalos RR, ou seja, o número de padrões π na série representa uma medida da complexidade do sinal. Assim, calcula-se a entropia de Shannon das frequências de ocorrência dos símbolos (padrões) gerados pelos PEO's. Esta entropia é também chamada de entropia de permutação e é dada pela equação:

$$H_{perm} = \sum_{\pi=1}^{n!} p(\pi) \log_2 p(\pi), \quad (4.3)$$

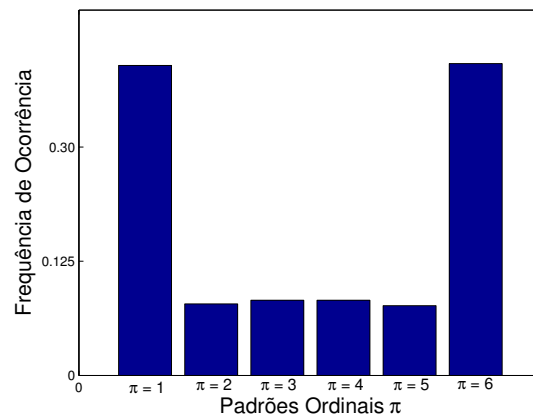
A entropia de permutação refere-se à estrutura de ordem local da série temporal. O menor valor possível de H_{perm} é zero e o maior valor possível é $\log_2 n!$ que ocorre quando todas as permutações são equiprováveis. Isso acontece para o ruído branco. A Figura 4.6 mostra diferentes padrões gerados para a função seno e para uma série aleatória (ruído branco). Nota-se ainda que os padrões gerados para uma série aleatória apresentam distribuição uniforme. Nessa Figura, pode-se observar que para $n = 3$ tem-se 6 padrões ordinais ($\pi = 6$).

Para este trabalho, escolheu-se valores de $n = 4$ e $\tau = 3$ e de $n = 4$ e $\tau = 5$ para gerar os padrões ordinais das 66 séries de intervalos RR e calculou-se as respectivas entropias H_{perm1} e H_{perm2} . A lista completa de características é dada na Tabela 5.4.

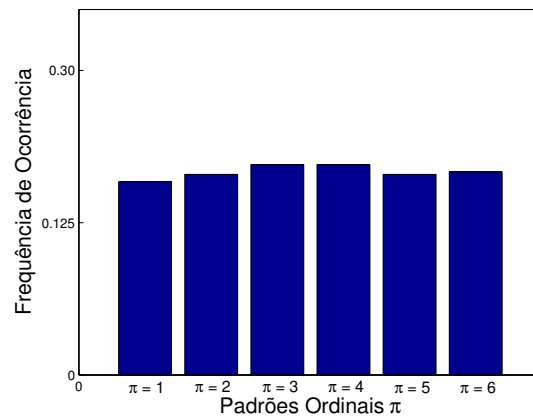
4.3.4 Classificação

Classificação baseada em DS

Inicialmente foi feita uma classificação utilizando apenas características relacionadas com a dinâmica simbólica: entropia (H_{ds}) e número de palavras proibidas (nPP). Após a extração das características de cada classe, foram feitos dois Testes com o intuito de discriminar entre duas (controle e chagásicos) e três classes (controle, chagásicos não-cardiopatas e chagásicos cardiopatas).



(a)



(b)

Figura 4.6: PEO's para séries temporais. Padrões para $n = 3$. Frequência de ocorrência dos padrões π para a função seno (a) e para o ruído branco (b).

Teste 1 - Análise de Discriminante Linear de Fisher

Foram selecionadas 22 séries de cada grupo que foram divididas em 2 subconjuntos, um conjunto de treinamento e um conjunto de validação, ambos com 11 registros. Vale notar que no problema, em questão, os dados estão “equilibrados”, ou seja, as classes estão balanceadas. Caso isso não ocorresse, algumas providências deveriam ser tomadas para garantir o balanceamento de classes, tais como: a abordagem baseada no pré-processamento dos dados (*e.g.* a reamostragem aleatória dos dados no espaço de entrada) e as abordagens baseadas em adaptação de algoritmos de aprendizado (*e.g.* abordagem baseada em reconhecimento e abordagem sensível ao custo) (Castro e

Tabela 4.1: Lista de Características

Índice	Características	Natureza
1	mNN	
2	SDNN	
3	SDANN	
4	RMSSD	
5	pNN50	Tempo-Frequenciais
6	MSD	
7	TotPow	
8	LF	
9	HF	
10	LF/HF	
11	H_{ds}	Dinâmica Simbólica
12	nPP	
13	H_{pem1}	Padrões Estatísticos Ordinais
14	H_{pem2}	

Braga 2011). O balanceamento via reamostragem aleatória foi adotado para o problema de classificação que será abordado no Capítulo 5.

Nesse Teste, foi utilizado um Discriminante Linear de Fisher para a classificação dos grupos. Esse classificador está descrito em (Horta 2008). Primeiramente, foi feita a classificação entre os *três* grupos separadamente, ou seja, as 33 séries de validação foram classificadas, uma a uma, utilizando o referido classificador. Posteriormente, as 33 séries foram classificadas em apenas *dois* grupos: controle e chagásicos, quer cardiopatas, quer não.

Teste 2 - Máquina de Vetor de Suporte

O experimento realizado no Teste 1 foi repetido, utilizando-se uma Máquina de Vetor de Suporte (*Support Vector Machine, SVM*) para discriminar o grupo controle (grupo 1) do grupo de chagásicos (grupos 2 e 3), logo, a classificação foi feita entre *duas* classes: *classe controle* e *classe chagásicos*. Esse classificador está descrito no Capítulo 3 e em (Acy e Guzelis 2004; Chisci et al. 2010; Haykin 2008; Horta 2008). Os parâmetros da SVM foram definidos como: *kernel* linear e constante de regularização $C = 1$. As características utilizadas para a classificação dos grupos foram tanto a entropia quanto

o número de palavras proibidas. Foram selecionadas as mesmas 22 séries de cada grupo, do experimento anterior, porém os grupos de treinamento e validação foram selecionados aleatoriamente.

Classificação baseada em DS, PEO's e índices tempo-frequenciais

Posteriormente, foi feita uma classificação utilizando características relacionadas com a DS, com os PEO's e com os índices convencionais do tempo e da frequência. Foi utilizada uma Máquina de Vetor de Suporte (SVM) multiclases com kernel RBF (*Radial Basis Function*) [3]. Detalhes sobre esse classificador foram descritos nos Capítulos 3 e 5. As características listadas na Tabela 4.1 foram usadas como entrada do classificador. Os grupos de treinamento e de teste foram escolhidos aleatoriamente, sendo 20% das amostras utilizadas para o treinamento e as 80% restantes para o teste. Foram feitas validações cruzadas para a escolha dos parâmetros C (parâmetro de regularização) e σ do SVM. Foram feitas duas classificações: a primeira para distinguir duas classes (controle e chagásicos) e a segunda para distinguir entre três classes distintas (controle, chagásicos não-cardiopatas e chagásicos cardiopatas).

O desempenho do classificador é avaliado por meio da taxa de acerto (A_{cc}):

$$A_{cc} = \frac{n_c}{M} \times 100\%, \quad (4.4)$$

em que n_c indica o número de épocas corretamente classificada e M representa o total de épocas.

Análise de Variância (ANOVA simples - one-way ANOVA)

O teste ANOVA foi aplicado neste trabalho, com o intuito de fazer uma análise estatística das entropias de cada grupo para verificar se, em média, haveria diferenças entre as entropias dos três grupos. Para maiores detalhes sobre essa técnica, o leitor interessado pode consultar as referências (Bolfarine e Sandoval 2000; Casella e Berger 1990; Magalhães e Lima 2002) e o Apêndice B.

4.4 Resultados

4.4.1 Análise da VFC por meio da Dinâmica Simbólica

A Figura 4.7 ilustra os histogramas normalizados característicos relacionados a um indivíduo do grupo controle, a um indivíduo chagásico não-cardiopata e a um indivíduo chagásico cardiopata. Foram analisadas sequências de símbolos de tamanho $n = 6$.

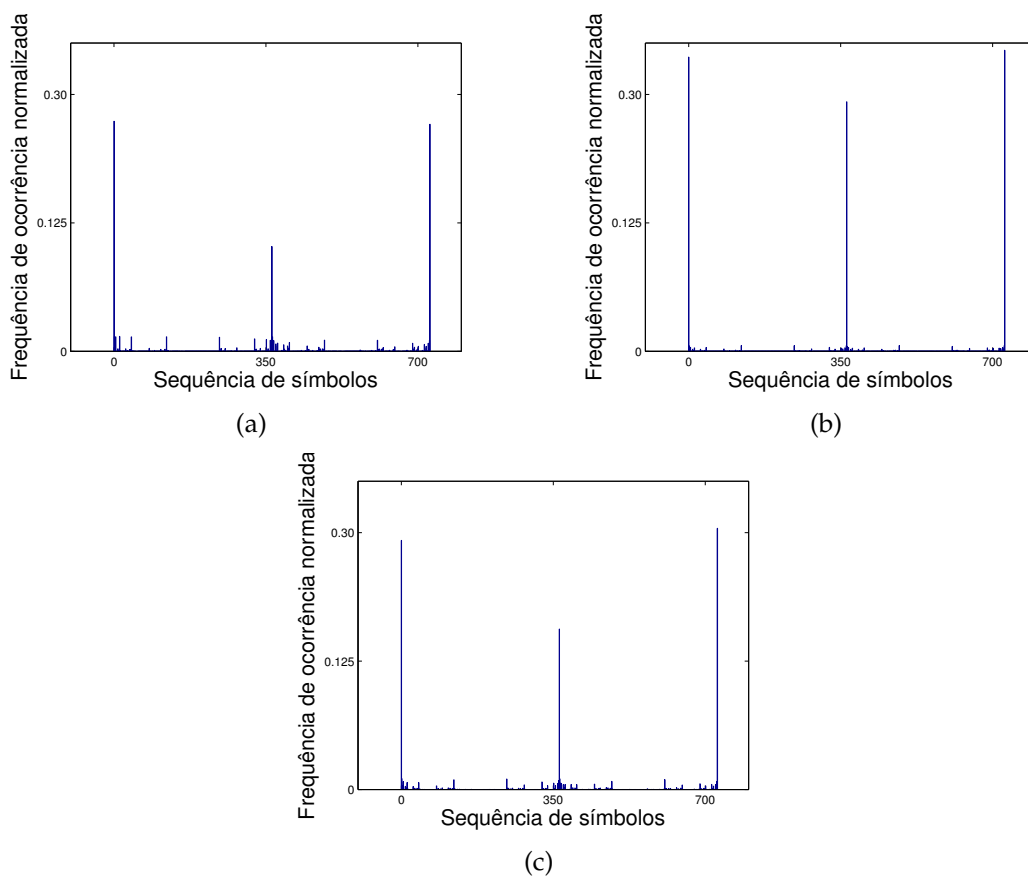


Figura 4.7: Histogramas normalizados de (a) um indivíduo do grupo controle; (b) um indivíduo do grupo 2; (c) um indivíduo do grupo 3.

4.4.2 Análise da VFC por meio de Testes Estatísticos - ANOVA simples

Em seguida, foram calculados os valores da entropia para cada um dos grupos. A Figura 4.8 mostra que os valores médios de entropia são maiores para o grupo controle do que para os outros grupos e que a diferença entre esse grupo comparada aos outros dois foi significativa ($p = 2,4055 \times 10^{-6}$), para um nível de significância $\alpha = 0,05$.

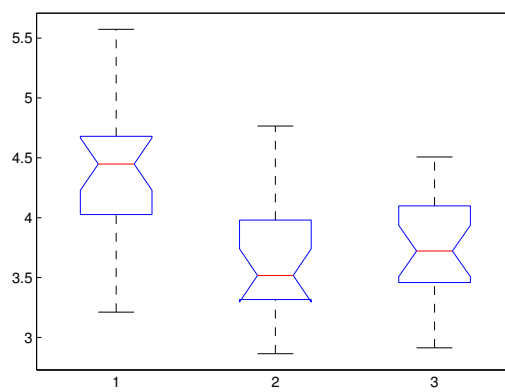


Figura 4.8: Gráfico *boxplot* representando os valores de entropia médios encontrados para os três grupos. A reta central representa a mediana das entropias, as bordas da caixa são os percentis 25% e 75% e as retas nas extremidades são os valores máximos e mínimos de entropia.

Foi aplicado o teste estatístico *Lilliefors test* nas distribuições das entropias de cada grupo e verificou-se que a hipótese nula de gaussianidade dos dados não foi rejeitada, com nível de significância $\alpha = 0,05$. Assim, pôde-se aplicar o teste ANOVA às populações de cada grupo.

4.4.3 Resultados obtidos na Classificação

Uma vez comprovado que é possível realizar a distinção dos comportamentos dinâmicos do grupo controle comparado aos grupos 2 e 3, foram desenvolvidos critérios de extração de características que, posteriormente, foram utilizados para treinar os classificadores descritos em (Acyr e Guzelis 2004; Chisci et al. 2010; Haykin 2008; Horta

2008).

Análise do Teste 1

Na primeira parte do Teste 1, a taxa de acerto foi de 55%. Ao separar as 33 séries em dois grupos, a taxa de acerto foi de 86%.

Análise do Teste 2

Aplicando-se a metodologia descrita no Teste 2, conseguiu-se classificar com 88% de acerto, os indivíduos do grupo controle ao compará-los com grupos restantes (grupos 2 e 3 pertencentes a um mesmo grupo - grupos de chagásicos). O desempenho do classificador (SVM Linear) pode ser observado na Tabela 4.2.

Tabela 4.2: Desempenho do SVM

Exatidão	0,88
Sensibilidade	0,75
Especificidade	0,92

Análise da VFC por meio da DS, PEO's e índices tempo-frequenciais

Com o intuito de motivar o uso de vetores de características e classificação não-linear usando SVM, analisaram-se as médias das características de cada grupo. Aplicando-se um teste de normalidade (*Lilietest*) observou-se que a maioria das características não apresentou distribuição Gaussiana, assim, optou-se por um teste não paramétrico (Teste de *Wilcoxon*) para comparar as médias de cada grupo. O teste foi aplicado comparando-se os grupos dois a dois, ou seja, foi feita a comparação entre as médias das características do grupo controle (grupo 1) e grupo 2, dos grupos 1 e 3 e dos grupos 2 e 3. A Tabela 4.3 mostra o resultado do teste de hipótese. Como pode ser observado, embora a maioria das características seja estatisticamente significativa ($p \leq 0,05$), as séries de VFC não podem ser classificadas corretamente, por meio da análise de um simples limiar.

Tabela 4.3: p – valores (p) para a comparação das médias das características entre os grupos. Os índices i representam os índices listados na Tabela 4.1 em ordem crescente de p . Os números subscritos em p representam os grupos que estão sendo comparados. Os valores em negrito indicam que as características relacionadas à DS e aos PEO’s são significativas.

i	p_{12}	i	p_{13}	i	p_{23}
14	0,0000	8	0,0000	10	0,0000
9	0,0000	14	0,0000	5	0,0000
5	0,0000	7	0,0002	4	0,0000
4	0,0000	10	0,0008	12	0,0000
6	0,0000	1	0,0014	9	0,0000
13	0,0000	9	0,0127	6	0,0000
11	0,0000	2	0,0177	13	0,0001
8	0,0001	6	0,0647	11	0,0001
12	0,0002	5	0,2621	14	0,0223
10	0,0014	11	0,3243	1	0,0531
7	0,0748	12	0,6439	8	0,1294
2	0,0952	13	0,7082	7	0,1753
1	0,4355	3	0,7092	2	0,4881
3	0,8381	4	0,7126	3	0,9703

A Tabela 4.4 mostra os principais resultados da classificação para duas classes (controle e chagásicos) e para três classes (controle, chagásicos não-cardiopatas e chagásicos cardiopatas).

Tabela 4.4: Desempenho do SVM multiclases para 2 e 3 classes

Classes	Taxa de Acerto (%)	Características
2	84	1–14
3	70	1–14

A Tabela 4.5 mostra os principais resultados da classificação para 2 e 3 classes com a retirada de características associadas à DS e aos PEO’s.

Tabela 4.5: Desempenho do SVM multiclases para 2 e 3 classes para grupos de características específicos

Classes	Taxa de Acerto (%)	Características
2	77	1–12
3	61	1–12
2	82	1–10, 13 e 14
3	66	1–10, 13 e 14

4.5 Discussão

Análise por meio da DS

Analisando-se a Figura 4.7, percebe-se que o histograma do indivíduo do grupo controle indica frequências de ocorrência não nulas para um número maior de palavras, sugerindo que os indivíduos do grupo controle apresentam um comportamento mais complexo (mais estocástico) do que os indivíduos chagásicos. Isso pode ser comprovado na Figura 4.8, na qual se observa que a mediana das entropias do grupo controle é maior do que nos outros dois, confirmando essa conjectura. Percebe-se, ainda, na Figura 4.7 que a variedade de palavras, ou seja, de padrões de VFC, é maior para o grupo controle.

Ao analisar o resultado obtido no Teste 1 (55%), optou-se por fazer a distinção entre indivíduos do grupo controle e indivíduos chagásicos, em vez de separar os três grupos (controle, chagásicos cardiopatas e chagásicos não-cardiopatas), obtendo-se, assim, melhores resultados tanto no Teste 1 (86%) quanto no Teste 2 (88%, Tabela 4.2); ou seja, com a metodologia proposta foi possível fazer uma distinção entre indivíduos saudáveis e indivíduos chagásicos.

Análise por meio de DS, PEO's e índices tempo-frequenciais

Os autores de (Parlitz et al. 2012) utilizaram características baseadas no domínio do tempo e da frequência, em DS e em PEO's para a classificação de séries de 24h (longa duração) de batimentos cardíacos de 15 pacientes com insuficiência cardíaca congestiva (ICC). A classificação foi feita com o intuito de distinguir entre duas classes: pacientes

sadios e pacientes com ICC. Esses autores obtiveram taxas de acerto que variaram entre 73 a 100% no teste com validação cruzada via *leave-one-out* utilizando uma SVM em um espaço bidimensional, ou seja, eles utilizaram apenas grupos de duas características para a classificação. Entretanto, ao comparar com os resultados obtidos neste trabalho (Tabela 4.4), pode-se verificar que houve uma taxa de acerto de 84% para duas classes e de 70% para três classes. Esse percentual de acerto (70%) para a distinção de *três classes* ainda não foi encontrado na literatura, até o conhecimento dos autores, em termos de classificação da VFC.

Em (Parlitz et al. 2012) foi sugerido que características baseadas em DS e PEO's combinadas com outros índices da VFC em um espaço de características de alta dimensão poderia melhorar o desempenho na classificação de sinais cardiovasculares, entretanto, isto não foi realizado pelos referidos autores. Neste trabalho foram apresentados resultados em \mathbb{R}^{14} . Entretanto, estudos (Asl, Setarehdan e Mohebbi 2008; Gudmundsson e Sigurdsson 2012) sugerem que a redução da dimensão do espaço de características pode diminuir o erro de classificação, assim mais estudos são necessários na busca de características relevantes por meio de técnicas de seleção de características. Isto pode ser observado na Tabela 4.3. Nota-se, nessa Tabela, que as características relacionadas à DS e aos PEO's são significativas para distinguir os três grupos (embora ao comparar os grupos 1 e 2 apenas a característica 14 mostrou-se significativa). Isto pode ser comprovado, ao comparar com os resultados listados na Tabela 4.5 em que se verifica que houve uma piora no desempenho da SVM ao retirar características ligadas à DS (características 11 e 12) ou aos PEO's (características 13 e 14). Ao comparar este estudo com o estudo feito anteriormente pelos autores deste trabalho (Rocha, Lemos e Aguirre 2010), confirma-se a conjectura de que algumas características são reduntantes, pois no estudo anterior obteve-se uma maior taxa de acerto ($A_{cc} = 88\%$) (Tabela 4.2) ao utilizar apenas características relacionadas à DS para distinguir entre duas classes. No entanto, ao comparar o estudo anterior ($A_{cc} = 55\%$) (Seção 4.4.3) com o presente trabalho, nota-se que houve uma melhora significativa ($A_{cc} = 77\%$) na classificação de três classes ao utilizar todas as características listadas na Tabela 5.4. Este resultado também confirma a conjectura de que estas características usadas em conjunto são úteis para a distinção

de diversos padrões da VFC, tanto fisio quanto patológicos.

4.6 Conclusões do Capítulo

Os resultados da primeira etapa deste Capítulo, que utilizaram SVM com *kernel* linear e apenas características relacionadas com a DS, mostraram que foi possível definir um procedimento adequado para a classificação da disfunção autonômica da VFC por meio da dinâmica simbólica, atendendo assim, aos objetivos propostos na Seção 4.2. Foi possível a distinção entre o grupo controle e o grupo de chagásicos (grupos 2 e 3) utilizando valores estimados de entropia e número de palavras proibidas. Porém não foi possível discriminar os pacientes chagásicos não-cardiopatas (grupo 2) e chagásicos cardiopatas (grupo 3). Logo, mais estudos são necessários para buscar ferramentas que sejam mais eficientes para a classificação desses grupos.

Na segunda etapa, que utilizou SVM multiclases e as 14 características listadas na Tabela 4.1, os resultados mostraram que foi possível a distinção entre 2 e 3 classes de indivíduos sadios e de indivíduos portadores da doença de Chagas. Assim, pode-se sugerir que o uso de novas características baseadas em dinâmica simbólica e padrões estatísticos ordinais aliadas aos índices clássicos tempo-frequenciais é potencialmente adequado para a classificação da VFC em pacientes chagásicos.

Classificação Automática dos Estágios do Sono por meio de Sinais Fisiológicos Disponíveis a um Ventilador

*“O caminho da ciência e dos saberes é o caminho da multiplicidade.
(...) Não há fim para as coisas que podem ser conhecidas e sabidas.
O mundo dos saberes é um mundo de somas sem fim.”*

Concerto para corpo e alma, de Rubem Alves

A principal contribuição desta tese, gerada pelos principais resultados obtidos neste capítulo, foi aceita na revista *Computers in Biology and Medicine*, com trabalho intitulado: “Automatic Sleep Staging from Ventilator Signals in Non Invasive Ventilation” (Sady et al. 2013).

5.1 Introdução

A ventilação não invasiva (VNI) é um tratamento reconhecido para pacientes com insuficiência respiratória crônica. Este tratamento é, geralmente, aplicado durante o sono e tem sido cada vez mais usado em domicílio (Elliott et al. 1992; Barbé et al. 1996). O objetivo principal da VNI é controlar a hipoventilação durante o sono, melhorando, assim, a qualidade do sono e, conseqüentemente, melhorando o bem-estar do paciente.

Entretanto, é bastante difícil a confirmação deste benefício, principalmente porque a VNI pode induzir eventos respiratórios indesejáveis de interrupção do sono (interações paciente-ventilador (Rabarimanantsoa et al. 2007; Achour et al. 2007), desconforto devido à máscara, etc). O sono pode ser influenciado também por outras condições físicas externas ou por razões clínicas. A fim de determinar objetivamente se o paciente dorme suficientemente bem ou para entender porque às vezes o uso de VNI não é benéfico, a avaliação regular da qualidade do sono durante a VNI é uma passo crucial (Elliott 2011).

A polissonografia (PSG) é considerada como o *padrão ouro* para avaliar o sono, mas trata-se de um procedimento demorado que requer pessoal especializado e que é mais frequentemente realizado em um laboratório do sono com gravações supervisionadas. O paciente passa a noite no hospital com vários sensores conectados ao escalpo e em outras partes do corpo. Vários sinais fisiológicos são registrados, tais como, a atividade neurocortical (eletroencefalograma, EEG), a atividade ocular (eletrooculograma, EOG), eletromiogramas (EMG), o eletrocardiograma (ECG), a oximetria de pulso (SpO₂), dentre outros sinais cardio-respiratórios. São necessários cuidados especiais na colocação dos eletrodos e a aquisição dos dados deve ser monitorada durante a noite. Embora seja possível realizar medições em domicílio, muitas vezes sua qualidade não é suficiente para que tais medições despertem interesse prático (Portier et al. 2000; Gagnadoux et al. 2002). Intervenção manual especializada também é necessária para uma classificação visual do sono de acordo com regras padrões e especificações (Iber et al. 2007a). Todos esses fatores tornam a PSG um procedimento complexo e caro, limitando o monitoramento do sono a clínicas e hospitais. Eletrodos e sensores usados para registrar esses sinais podem afetar a qualidade e a estrutura do sono. Além disso, uma maneira mais simples de avaliar o sono a reduzidos custos é, portanto, de valor inestimável, principalmente, para o tratamento em domicílio de pacientes ventilados.

Há vários estudos, na literatura, que descrevem métodos não invasivos para o registro de uma noite de sono completa (PSG completa) que poderiam, em princípio, ser aplicados em domicílio (Redmond e Heneghan 2006; Harper, Schechtman e Kluge 1987; Karlen, Mattiussi e Floreano 2009; Várady et al. 2002; Chung et al. 2009; Marcos

et al. 2009; Terrill et al. 2010). Sensores de cama, como o balistograma, são sistemas de medição sem contato bem conhecidos utilizados para medir variações na frequência cardíaca e na respiração (Kortelainen et al. 2010; Mendez et al. 2010). Características relacionadas com o sistema cardiovascular parecem ser promissoras para a classificação dos estágios do sono. Estágios do sono de adultos foram estimados utilizando informação cardiorrespiratória obtida a partir de um colchão de ar – um sistema de medição não-invasiva (Watanabe e Watanabe 2004). Distúrbios do sono em crianças foram diagnosticados usando *pulse time transit* (PTT) em (Smith et al. 1999; Pépin et al. 2005), indicando, assim, que esse sinal contém informação útil sobre a estrutura do sono. A detecção do estágio REM, que pode ser entendida como um problema de classificação binário, foi realizada usando apenas sinais respiratórios em (Chung et al. 2009). Em outros trabalhos, sinais respiratórios foram usados para a detecção de eventos de apneia (Várady et al. 2002; Marcos et al. 2009), patologia que está intimamente relacionada aos distúrbios do sono. A actigrafia também tem sido amplamente utilizada como um método de monitorização portátil para avaliar a qualidade do sono, observando se o paciente está tranquilo durante o sono (Kortelainen et al. 2010). Entretanto, em nenhum desses estudos, foi possível identificar mais de três estágios diferentes do sono e nenhuma informação suficiente foi, então, disponível para avaliar a estrutura do sono. Além disso, a maioria desses métodos requer sensores adicionais.

Ventiladores modernos são utilizados na VNI comumente para a medição e o registro de dados no circuito ventilatório (pelo menos o fluxo de ar e a pressão). Vários modelos estão também equipados com um oxímetro de pulso para estudar o sistema cardio-respiratório do paciente. É importante investigar se esses dados podem permitir a classificação do sono e, conseqüentemente, avaliar o sono de pacientes ventilados em domicílio, pois, em caso afirmativo, seria possível obter uma classificação mais prática e de baixo custo. Além disso, essa avaliação do sono poderia ser potencialmente muito menos perturbadora se comparada com a PSG, uma vez que o paciente já está ventilado e nenhum sensor adicional seria necessário, com exceção de um oxímetro de pulso que já é utilizado em alguns casos.

O objetivo principal deste Capítulo é descrever o desenvolvimento e a avaliação

de desempenho de uma classificação técnica do sono em pacientes sob VNI que usa apenas sinais disponíveis a um ventilador. O ponto chave é considerar a amplitude do sinal derivado do pletismograma de pulso (PPG). Esse tipo de sinal ainda não foi completamente explorado, apesar de seu aparente potencial como mencionado em (Chua et al. 2007). A fim de ser capaz de avaliar com precisão a estrutura do sono, três a cinco estágios do sono serão classificados. O estagiamento do sono (classificação) é realizado por meio de uma máquina de vetor de suporte (SVM) multiclases. A técnica descrita neste Capítulo poderá ser eventualmente integrada a ventiladores, permitindo, então, o monitoramento contínuo do sono em domicílio para pacientes tratados via VNI.

Este Capítulo está organizado da seguinte maneira: a Seção 5.2 (*Metodologia*) descreve o protocolo e o pré-processamento dos três sinais utilizados para treinar o classificador baseado em SVM's, que é descrito na Seção 5.3 (*Classificador*). Os resultados são apresentados na Seção 5.4 (*Resultados*) e discutidos na Seção 5.5 (*Discussão*) e na Seção 5.6 (*Conclusões do Capítulo*) as principais conclusões são fornecidas.

5.2 Materiais e Métodos: Dados e Processamento dos Sinais

Os dados utilizados, neste trabalho, foram coletados durante um estudo observacional na Unidade Pulmonar e Respiratória do Hospital da Universidade (CHU) de Rouen (França) (Naeck 2011)¹. O objetivo desse estudo era avaliar as interações paciente-ventilador e mudanças na estrutura do sono durante as primeiras noites após o início da VNI. Alguns detalhes dos treze pacientes que sofrem de insuficiência respiratória crônica, selecionados, a partir do referido estudo, serão descritos na sequência. Dados de um paciente (P_5), originalmente incluído em (Naeck 2011), não foram utilizados no presente estudo, já que esse paciente desistiu do tratamento no meio da coleta de dados, logo, não foram coletados todos os dados necessários para análises, neste trabalho, para o referido paciente.

5.2.1 Pacientes

Características clínicas de treze pacientes são listadas na Tabela 5.1. Sete pacientes ($P_3, P_4, P_7, P_8, P_{11}, P_{12}, P_{13}$) apresentavam síndrome de obesidade hipoventilatória (SOH) associada a uma síndrome de apneia/hipoapneia obstrutiva do sono (SAHOS) de severidade variável (Índice de Apneia/Hipoapneia (IAH) variando entre 24 a 142 eventos/hora). Quatro pacientes (P_2, P_6, P_{10}, P_{14}) apresentavam esclerose lateral amiotrófica (ELA) dos quais os três primeiros tinham uma manifestação periférica e SAHOS, o quarto, uma manifestação bulbar. Dois outros pacientes apresentavam patologias restritivas: um paciente (P_1) sofria de seqüela de tuberculose e o outro (P_9) de cifoescoliose (deformação torácica) associada a SAHOS.

Todos os pacientes realizaram três PSG's completas. A primeira PSG foi realizada sob respiração espontânea (noite 1) e as outras duas (noite 2 e noite 15, ou seja, duas semanas mais tarde) foram feitas sob VNI. Apenas os registros de duas PSG's foram considerados no presente trabalho devido ao interesse em investigar o sono de pacientes

¹ClinicalTrials.gov identifier: NCT01255111.

Tabela 5.1: Características clínicas dos treze pacientes (6 mulheres e 7 homens) incluídas no estudo de (Naeck 2011).

Características clínicas	média \pm σ	n
Idade (anos)	65.4 \pm 10.9	–
Índice de Massa Corpórea (kg.m ⁻²)	32.8 \pm 8.62	–
Índice de Apnea/Hipoapnea (por hora)	53.4 \pm 41.6	–
SAHOS + SOH	–	7
SAHOS + ELA	–	3
ELA	–	1
Sequela de Tuberculose	–	1
SAHOS + Deformação Torácica	–	1

sob VNI (Naeck 2011). A PSG referente à noite 15 não foi registrada para o o paciente P_{13} .

5.2.2 Protocolo Experimental

O estudo observacional realizado por (Naeck 2011) foi autorizado como parte de um protocolo de cuidados de rotina pelo Comitê de Proteção de Pessoas (*CPP Nord-Ouest 1*), autorização datada em 16 de outubro de 2009. A execução habitual de uma assistência ao *hospital semana*² quase não foi modificada. A única diferença introduzida por este protocolo foi a instalação de sensores neurofisiológicos (polissonografia) acoplados à poligrafia habitual; as restrições ligadas a este exame são insignificantes para os pacientes. Os procedimentos adotados, neste protocolo, obtiveram o consentimento oral de cada paciente. A adaptação clínica à VNI foi efetuada sob ventilação espontânea e, posteriormente, sob ventilação mecânica não invasiva, pela equipe médica do Serviço de Pneumologia do Hospital de Bois-Guillaume – CHU de Rouen. Os pacientes estiveram sob a vigilância de um interno do hospital semana, por enfermeiros especializados em VNI, por fisioterapeutas e pelo médico sênior do setor. A colocação dos eletrodos e, conseqüentemente, as polissonografias foram efetuadas pela própria autora do estudo

²O hospital semana (ou semana de internação hospitalar) é utilizado pelos serviços médicos e cirúrgicos para que qualquer assistência ao paciente seja realizada durante vários dias (exatamente cinco dias e quatro noites), geralmente, entre segunda e sexta-feira. Esse procedimento, adotado na França, se assemelha ao hospital dia, no Brasil, entretando, varia-se o período de internação do paciente.

(Naeck 2011) com o auxílio dos demais profissionais. Posteriormente, a vigilância do paciente foi efetuada pelo enfermeiro responsável do laboratório do sono (dedicado ao diagnóstico, assim como o monitoramento dos pacientes afetados por distúrbios respiratórios do sono).

Cada paciente foi hospitalizado, de segunda à noite a quarta após o meio-dia, para efetuar a introdução à VNI. O paciente retornava 15 dias após sua primeira noite de hospitalização para uma noite de controle. O paciente foi admitido ao laboratório do sono segunda, aproximadamente às 20 h, para um registro polissonográfico em ventilação espontânea em ar ambiente. Essa noite de diagnóstico (N_1) permitiu ao paciente se familiarizar com o seu ambiente e com a presença de sensores. O paciente foi instalado em sala individual e a equipe estava equipada com diferentes sensores necessários ao registro dos sinais polissonográficos. Uma gazometria arterial foi realizada em ventilação espontânea em ar ambiente. As luzes eram sistematicamente apagadas entre 22:30 h e 23 h e o registro dos dados podia começar. Na terça, pela manhã (no dia seguinte), o paciente era acordado por volta de 5:30 h para efetuar uma gazometria de fim de noite em ventilação espontânea em ar ambiente. O paciente era, em seguida, instalado em um quarto do hospital semana em que o material de ventilação (um ventilador e uma interface) lhe era apresentado durante o dia. A Tabela 5.2 resume os procedimentos adotados durante as noites sob VNI.

Tabela 5.2: Programação do procedimento de ventilação não invasiva

	Segunda (N_1)	Terça (N_2)	Quarta	Quinta	Sexta	Segunda (N_{15})
Dia	–	P	P	P	P	–
Noite	Labo do sono	Labo do sono	H	H	–	Labo do sono
Exames	PSG em VE	PSG em VNI	\emptyset \emptyset	\emptyset \emptyset	\emptyset \emptyset	PSG em VNI

PSG: polissonografia; VE: ventilação espontânea; VNI: ventilação não invasiva; H: hospital semana; P: Início da VNI, adaptação às configurações no hospital.

Todos os pacientes foram equipados por um ventilador VPAP IV (*Resmed, Austrália*) configurado em modo espontâneo/temporário com uma frequência de segurança de 8

ciclos por minuto. Todos os pacientes possuíam, também, uma máscara facial (*Quattro ou Mirage, ResMed, Austrália*). Os parâmetros do ventilador foram configurados pela equipe médica. Os pacientes efetuaram, então, uma hora de ventilação no fim da manhã ou no início da tarde, antes da coleta de uma nova gazometria arterial. Em função dos resultados, os parâmetros do ventilador puderam ser modificados uma primeira vez.

Na terça à noite (N_2), o paciente retornava ao laboratório do sono para um registro polissonográfico sob VNI e oximetria. Diferentes sensores foram colocados à partir de 20 h; as luzes foram, então, apagadas às 23 h. Na quarta de manhã, o paciente acordava por volta das 5:30 h. Uma gazometria arterial foi efetuada sob VNI e o paciente retornava ao seu quarto do hospital semana. De acordo com os resultados da gazometria do sangue, as configurações do ventilador puderam ser novamente modificadas. As noites de quarta e quinta foram efetuadas no hospital semana sob controle de uma oximetria segundo o caso. O paciente retornava a seu domicílio na quinta à tarde ou na sexta. Cada paciente era novamente hospitalizado no laboratório do sono na segunda-feira, 15 dias após a primeira noite de gravação (N_{15}). Foi verificado que os pacientes utilizaram adequadamente o ventilador. Um registro polissonográfico sob VNI e oximetria (N_2) foi, então, realizado (a duração média dos registros, entre as noites N_2 e N_{15} , foi de 6 h). As configurações da VNI e dos débitos de oxigênio adicional puderam ser modificadas no hospital semana entre as noites N_2 e N_{15} . De acordo com o caráter não intervencional e de cuidados de rotina do protocolo, em questão, os dados obtidos por meio das polissonografias não foram utilizados para a configuração da VNI. Esses dados, tanto neurofisiológicos quanto ventilatórios, foram lidos e interpretados por um neurologista especializado nesta área.

Para cada paciente, um caderno de observações foi preenchido para cada noite (Naeck 2011). Os cadernos e os registros polissonográficos foram anônimos. Esse caderno foi dividido em duas partes: a primeira parte continha as características gerais do paciente, sua patologia e os dados funcionais e gazométricos que indicavam o estabelecimento de uma VNI em domicílio e a segunda parte concernia ao monitoramento do paciente durante a semana de hospitalização e da noite de registro 15 dias depois. A gazometria

arterial realizada em cada noite e em cada manhã, assim como as configurações do ventilador efetuadas em cada noite, foram listadas nesse caderno. O questionário de qualidade do sono do *St Mary's Hospital* foi preenchido todas as manhãs (Naeck 2011); o questionário de auto-avaliação do sono convida o paciente a escrever seus hábitos, ao qual se acrescenta a escala de sonolência de *Epworth* que foi preenchida pelo paciente na noite N_1 e na noite N_{15} (Naeck 2011).

5.2.3 Registro de Dados

Polissonografia

Os dados foram registrados por meio de um sistema de aquisição CID102-L8D (CIDELEC SA, França). Todos os sinais registrados e os sensores correspondentes estão listados na Tabela 5.3 com as respectivas frequências de amostragem.

Um exame polissonográfico corresponde ao registro de vários sinais simultaneamente:

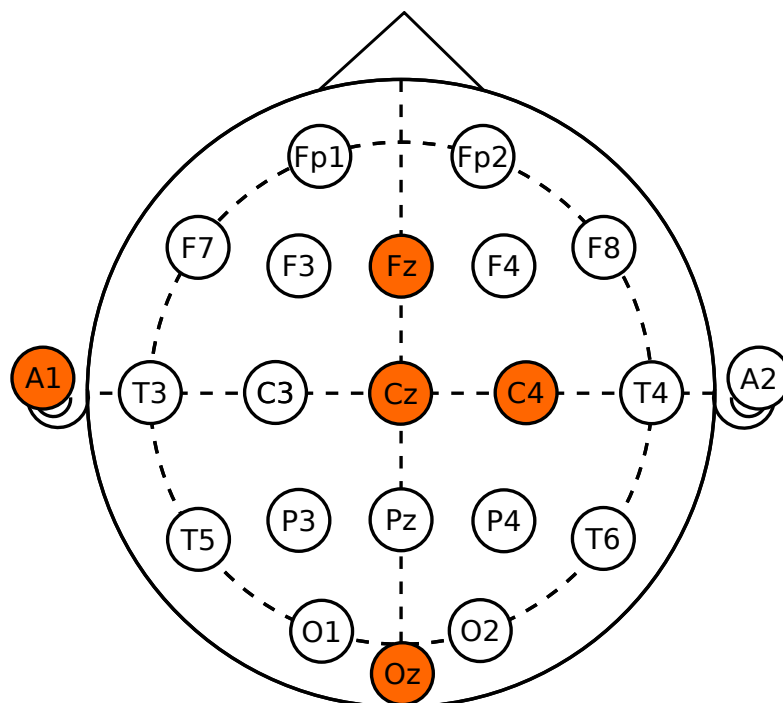


Figura 5.1: Montagem da eletroencefalografia. A colocação dos eletrodos está circulado em laranja. Fonte: Figura modificada de (Wikipedia 2013a).

Tabela 5.3: Sinais registrados durante cada polissonografia nas noites N_1 , N_2 e N_{15} . Com exceção da pneumotacografia e da Tosca 500, todos os sensores são da marca CIDEELEC.

Sinais	Sensores	Frequência de Amostragem (Hz)
Poligrafias Ventilatórias		
Mobilidade torácico-abdominal	Faixas emborrachadas	8
Posição do paciente	Sensor de posição	1
Movimentos corporais	Actímetro	8
Sons traqueais	Microfone	16
Frequência de pulso	Oxímetro de pulso	64
Saturação de O_2	Oxímetro de pulso	1
$PtcCO_2$	Eletrodo auricular, Tosca 500, RADIOMETER	1
Fluxo de ar	Pneumotacógrafo de <i>Fleish n° 2</i>	128
Pressão nasal	Sensor de pressão diferencial	256
Pressão do ar	Sensor de pressão diferencial	128
Neurofisiológicas		
EEG	Eletrodos	128
EOG	Eletrodos	128
EMG (mentoniano + pernas)	Eletrodos	128
ECG	Eletrodos	128

$PtcCO_2$ = pressão parcial de CO_2 medida por via transcutânea.

- sinais que permitem a codificação dos diferentes estágios do sono: EEG, EOG e EMG dos músculos mentonianos. As três derivações utilizadas para o registro dos sinais são as seguintes: $F_z - C_z$, $C_z - O_z$, $C_4 - A_1$ e o eletrodo terra é colocado na frente (o posicionamento dos eletrodos está especificado na Figura 5.1);
- a atividade dos músculos tibiais das duas pernas;
- a atividade elétrica cardíaca (ECG) por intermédio de dois eletrodos paraesternais;
- a atividade respiratória: uma cânula nasal foi inserida para registro da vazão respiratória de ar do paciente na primeira noite em ar ambiente e em ventilação espontânea. Ao longo da segunda noite e da 15ª noite, um pneumotacógrafo foi instalado na saída do ventilador permitindo coletar as variações do débito. Um

sensor de pressão foi instalado na máscara via orifício específico. Um circuito simples foi utilizado com uma máscara facial de vazamentos intencionais calibrados, permitindo assegurar a evacuação do gás expirado e evitar a reinalação do CO_2 (Naeck 2011);

- sinais do actímetro (sensor posicionado no dedo) e mudanças de posição na cama (sensor colocado no abdômen/tórax);
- a $Pt\text{CO}_2$ correspondendo à medida transcutânea do CO_2 por um sensor colocado no lóbulo da orelha externa conectado ao aparelho de medida, TOSCA 500 (Radiometer, França). Após estabelecimento do eletroencefalograma, a validade dos sinais foi testada por diferentes testes de oclusão e de rotação ocular, de deglutição e de mobilização dos membros inferiores. O ventilador utilizado (VPAP IV, ResMed) e os demais sensores estão representados na Figura 5.2. A posição dos sensores polissono e poligráficos está representada no desenho esquemático da Figura 5.3.

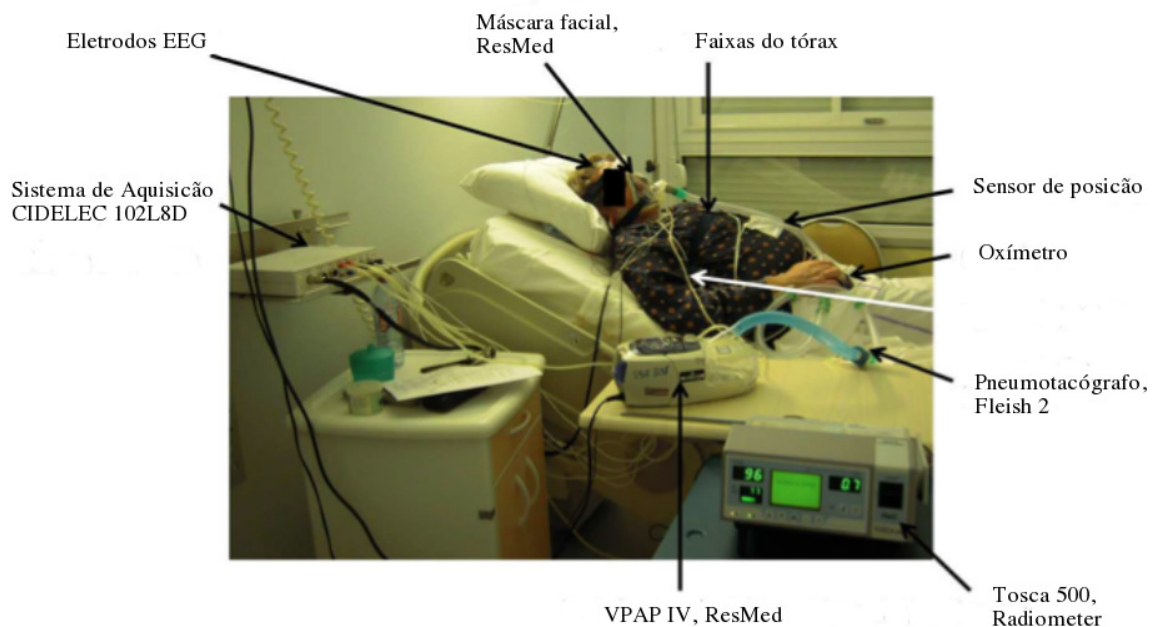


Figura 5.2: Polissonografia sob ventilação não invasiva: exemplo de um paciente do estudo no laboratório do sono (Adaptado de (Naeck 2011)).

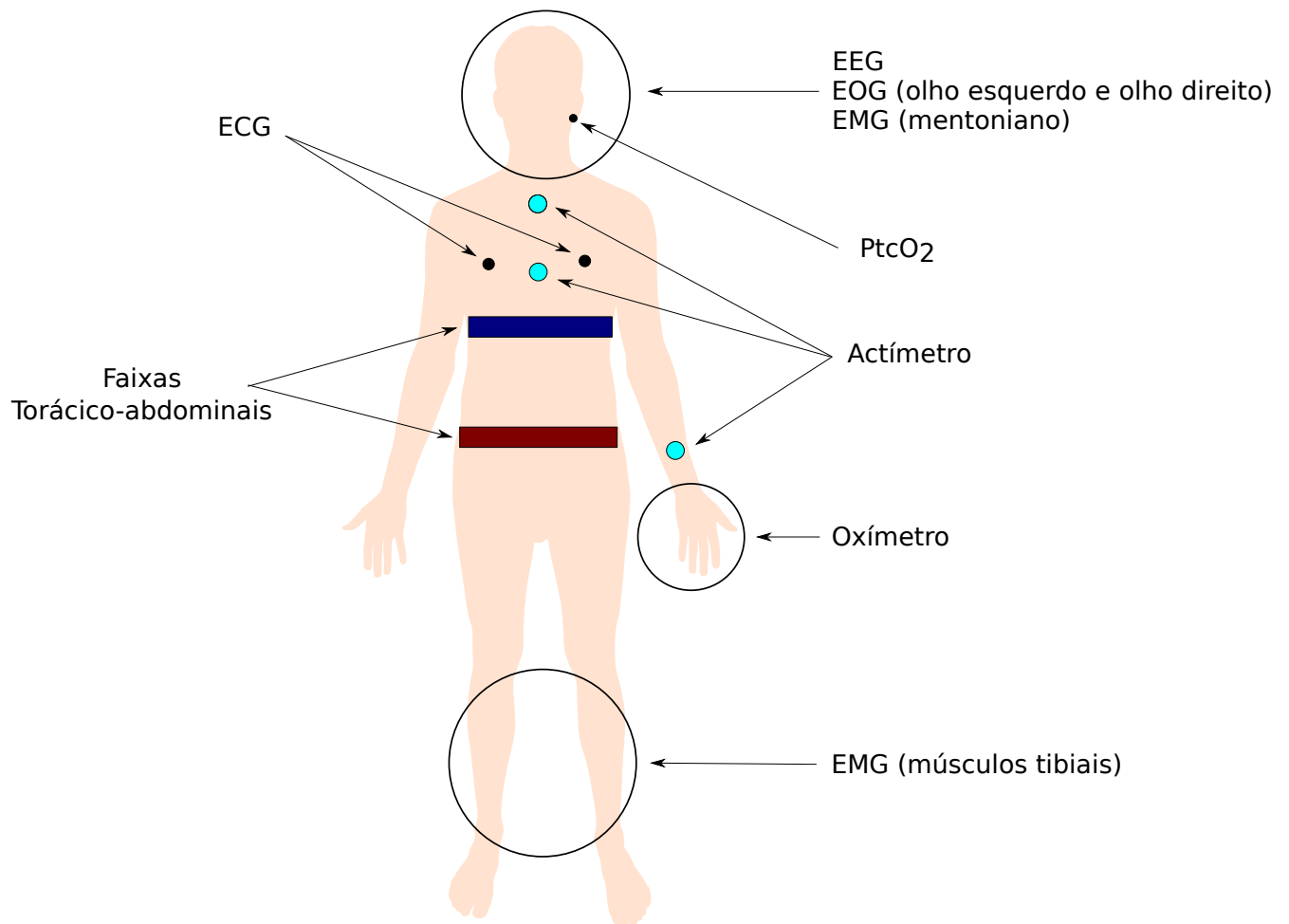


Figura 5.3: Desenho esquemático da posição dos sensores durante uma polissonografia completa. Fonte: Figura modificada de (Wikipedia 2013b).

Como mencionado, o registro de sinais inclui sinais eletrofisiológicos (EEG, EMG, EOG e ECG) e dados fisiológicos ventilatórios (pressão, fluxo de ar, saturação de oxigênio, dentre outros). O sistema de aquisição usado foi o CID102-L8D (CIDELEC SA, França). Cada conjunto de dados foi analisado por um neurologista e o sono foi codificado em épocas de 30 s de acordo com as diretrizes da Associação Americana de Medicina do Sono (AAMS) (Iber et al. 2007a). Cinco estágios foram distinguidos: Vigília (W), estágio REM (*Rapid Eye Movements*) e estágios N_1 , N_2 e N_3 (*slow wave sleep*).

De todos os sinais registrados na PSG, somente três foram utilizados:

- i) o fotopletismograma de pulso (PPG);
- ii) a saturação de oxigênio da hemoglobina (SpO_2) e

iii) a vazão de ar.

Os dois primeiros foram medidos por meio de um sensor colocado no dedo do paciente. A frequência de amostragem do PPG foi de 64 Hz e a da SpO₂ foi de 1 Hz (Tabela 5.3). A vazão foi medida por meio de um pneumotacógrafo inserido no circuito ventilatório e foi amostrada a 128 Hz (Tabela 5.3). Embora o ventilador tenha um pneumotacógrafo interno, como a maioria dos ventiladores modernos, suas medições não são facilmente acessíveis. Ao contrário do sinal de saturação de oxigênio (SpO₂), os sinais de PPG e vazão requerem um pré-processamento como o descrito nas próximas duas subseções.

5.2.4 Pré-processamento do Fotopletismograma

Todas as *características* relacionadas à atividade cardíaca foram determinadas pela análise do sinal de fotopletismograma (o sinal de ECG não foi usado neste trabalho). Apesar de o uso do PPG na determinação de batimentos cardíacos para a classificação do sono ter sido sugerido em (Chua et al. 2007), evitando a utilização de medidas obtidas via sinais de ECG e os eletrodos necessários, isto ainda não foi implementado. O sinal de PPG é o *signal-chave* utilizado, neste trabalho, que permite a classificação dos estágios do sono em pacientes sob VNI, possivelmente em domicílio e sem a necessidade de qualquer eletrodo.

O sinal de PPG foi, primeiramente, pré-processado aplicando um filtro *Butterworth* passa-banda de terceira ordem com frequências de corte de 0,667 Hz and 5 Hz, que correspondem, respectivamente, a 40 e 300 batimentos por minuto. O sinal filtrado foi, então, utilizado para detectar os batimentos cardíacos e para calcular os intervalos entre batimentos. A frequência de amostragem relativamente baixa do sinal de PPG (64 Hz) induz algum ruído de *quantização* nos intervalos de batimentos cardíacos, uma vez que qualquer intervalo calculado deve ser um múltiplo inteiro do período de amostragem do PPG, isto é, 15,625 ms no presente caso. A fim de se obter um menor erro de quantização para os intervalos de batimentos cardíacos, o sinal de PPG filtrado foi interpolado. A transformada de *Fourier* do sinal filtrado foi *zero-padded* e, em seguida,

transformada de volta para o domínio do tempo de tal forma que o novo sinal foi amostrado a 256 Hz. Este tipo de interpolação não altera o conteúdo espectral do sinal original.

Os batimentos cardíacos foram detectados observando os cruzamentos por zero (*zero-crossings*) do sinal de PPG filtrado e interpolado (Figura 5.4). Somente os cruzamentos por zero mudando de valores do negativo para o positivo foram considerados. Isto mostrou ser numericamente mais eficiente do que o análogo procedimento para a detecção de picos do sinal de PPG. A detecção das posições dos batimentos cardíacos foram, então, utilizadas para calcular o tacograma. A amplitude de oscilações do PPG foi calculada usando a diferença entre o máximo e o mínimo do sinal de PPG original (não-filtrado, não-interpolado) em cada batimento detectado (Figura 5.4) produzindo, então, uma série temporal das das flutuações das oscilações da PPG. Esse procedimento resultou em duas séries temporais:

- i) uma correspondente ao tacograma (obtida a partir do sinal de PPG processado), e
- ii) outra contendo a *amplitude* de oscilações do PPG para cada batimento cardíaco (obtida a partir do sinal de PPG original).

5.2.5 Pré-processamento da Vazão

O sinal de vazão foi filtrado com um filtro *Butterworth* passa-banda de terceira ordem com frequências de corte de 0,0833 Hz e 0,667 Hz, que correspondem a 5 e 40 respirações por minuto. Uma série temporal foi calculada a partir do tempo de duração dos ciclos respiratórios e uma outra calculada por meio das amplitudes da vazão foram geradas usando o mesmo método unidirecional de cruzamentos por zero como o usado com o PPG (Figura 5.4). As amplitudes de oscilações da vazão também são raramente utilizadas, já que tais medições requerem o uso de uma máscara durante toda a noite e isto pode perturbar o sono. No entanto, uma vez que os pacientes estão sob VNI, são necessariamente equipados com uma máscara; este procedimento representa o caso do presente trabalho, sem custo extra.

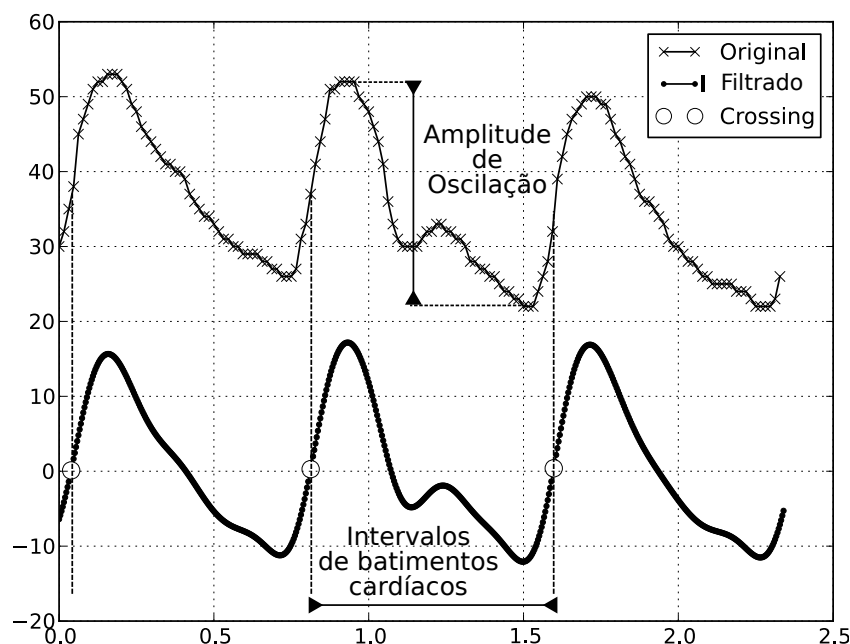


Figura 5.4: As amplitudes de oscilações do PPG foram calculadas usando os máximos e mínimos do sinal de PPG original (gráfico superior). Os intervalos de batimentos cardíacos foram calculados usando cruzamentos por zero no sinal de PPG filtrado e interpolado (gráfico inferior).

5.3 Métodos: Arquitetura do Classificador e Extração de Características

O objetivo principal da fase de classificação é desenvolver, a partir dos dados, um algoritmo matemático, referido como o *classificador*, que deve ser capaz de classificar os estágios do sono. Em vez de tentar construir o classificador diretamente a partir dos registros dos três sinais, sabe-se que é mais conveniente determinar algumas *características-chave* a partir dos dados e, então, treinar um classificador a partir de tal informação (Duda, Hart e Stork 2006). Exemplos de características incluem: coeficientes de modelos (Lainscsek et al. 2012), coeficientes da transformada discreta de *Fourier*, coeficientes das *wavelets* e assim por diante. No desenvolvimento de classificadores, a escolha de quais características usar é essencial, e está longe de ser trivial. O resultado desse passo importante será descrito abaixo, antes de discutir o classificador.

5.3.1 Extração de Características

Após uma minuciosa investigação, a ser detalhada a seguir, quatorze características foram escolhidas para treinar o classificador. Inicialmente, a forma de onda dos sinais escolhidos (PPG, S_pO_2 e vazão) foi analisada de tal maneira em que se observaram variações nas amplitudes desses sinais, segundo o estágio do sono. Exemplo dessa variação da amplitude pode ser notado na Figura 5.5 na qual se observa uma forte diminuição na amplitude do sinal de PPG para o estágio REM. Assim, observando as variações de amplitudes do sinal de PPG, notou-se que a dinâmica desse sinal está correlacionada com a dinâmica do sono.

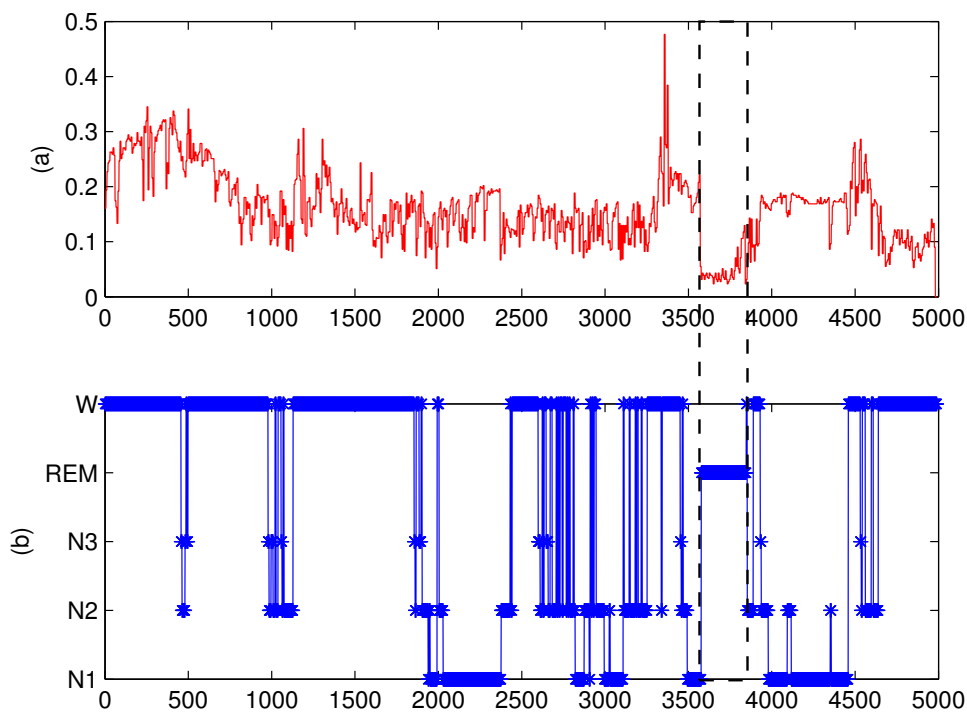


Figura 5.5: (a) Mediana da amplitude de oscilações do pletismograma. (b) Hipnograma da noite N_2 do paciente P_1 .

Decidiu-se, então, por calcular a mediana e o intervalo interquartil das amplitudes de oscilações dos três sinais. Os valores de mediana e intervalos interquartis foram utilizados em vez da média e do desvio padrão devido a sua menor sensibilidade aos

outliers (maior robustez).

Essas características foram divididas em quatro grupos, dependendo dos dados a partir dos quais elas foram estimadas (Tabela 5.4). Quatro características foram estimadas a partir do sinal de PPG:

- i) VFC_{med} , a mediana dos intervalos de batimentos cardíacos;
- ii) VFC_{sd} , o intervalo interquartil dos intervalos de batimentos cardíacos;
- iii) amp_{med} , a mediana das amplitudes de oscilações do PPG e
- iv) amp_{sd} , o intervalo interquartil das amplitudes de oscilações do PPG.

Quatro outras características foram estimadas a partir da vazão:

- i) tot_{med} , a mediana dos intervalos de ciclos respiratórios (tot);
- ii) tot_{sd} , o intervalo interquartil dos tot's;
- iii) $ampf_{med}$, a mediana das amplitudes de oscilações da vazão e
- iv) $ampf_{sd}$, o intervalo interquartil das amplitudes de oscilações da vazão.

Duas características foram estimadas a partir do sinal de SpO_2 , do qual não foi necessário um pré-processamento, como mencionado na Seção 5.2.3:

- i) sat_{med} , a mediana dos valores de SpO_2 e
- ii) sat_{sd} , o intervalo interquartil dos valores de SpO_2 .

Por fim, quatro características seguintes foram obtidas a partir de "outras" fontes:

- i) idade do paciente;
- ii) índice de massa corpórea (IMC);
- iii) pressão expiratória positiva (PEP);
- iv) pressão inspiratória positiva (PIP).

Os três primeiros grupos de características (Tabela 5.4) foram calculados a partir das três séries temporais listadas na Seção 5.2.3 para cada época. O último grupo, de quatro características, está relacionado com o paciente e foi tomado como constante ao longo de cada noite de PSG, mas podem variar de paciente para paciente ou de uma PSG para a outra.

Tabela 5.4: Lista de características calculadas para cada época de 30 s.

Índice	Característica	Fonte
1	VFC_{med}	PPG
2	VFC_{sd}	
3	amp_{med}	
4	amp_{sd}	
5	tot_{med}	Vazão
6	tot_{sd}	
7	$ampf_{med}$	
8	$ampf_{sd}$	
9	sat_{med}	Saturação de O_2
10	sat_{sd}	
11	idade	Outras
12	IMC	
13	PEP	
14	PIP	

Características advindas de coeficientes de modelos, tais como modelos NAR (*Non-linear Auto-Regressive Moving Average*) e LSP (*Line Spectral Pairs*), técnica muito utilizada para reconhecimento de voz (McLoughlin 2008; Turk e Arslan 2006; Choi, Kim e Lee 2000), também foram calculadas/extraídas para os três sinais. As características eram, então, utilizadas para treinar uma SVM e aquelas com as melhores taxas de acerto durante o treinamento eram escolhidas. A cada treinamento, acrescentava-se ou retirava-se uma característica (ou um grupo de características; exemplo: características relacionadas com a respiração ou com variáveis cardiológicas, e assim por diante) por vez e avaliava-se a taxa de acerto. Na realidade, é isso o que o método *wrapper* faz e que, para o presente trabalho, foi feito “manualmente”. Esse procedimento foi feito para cada nova característica que era extraída dos sinais, assim, observou-se que as características extraídas de coeficientes de modelos não melhoravam o desempenho

do classificador, logo, não foram “selecionadas” para este estudo. Esses resultados sugerem que o método de extração de características via modelos não era um bom método para ser utilizado, neste trabalho, ou não foi um bom método para extrair características relevantes desses sinais. Essa é uma limitação desse trabalho, já que mais estudos são necessários na busca de técnicas de *seleção de características* mais robustas. A biblioteca LIBSVM (Chang e Lin 2011) fornece um método de seleção de características baseado no *F-score* e essa técnica também foi, inicialmente, utilizada nesse trabalho, entretanto, verificou-se que o desempenho do classificador foi melhor utilizando as 14 características listadas na Tabela 5.4 do que utilizando o *F-score* para fazer um *ranking* dessas características e escolher as com menor valores de *p-valor*.

5.3.2 Classificador

Os classificadores (SVM's) foram treinados a partir dos dados que foram divididos em épocas de 30 s, seguindo a prática atual de classificação do sono (Iber et al. 2007a). Cada época tem seu próprio conjunto de catorze características (cujos índices são relatados na Tabela 5.4) e um estágio do sono (uma classe rotulada) correspondente atribuído por um neurologista. Após o treinamento, o classificador determina automaticamente o estágio do sono correspondente ao conjunto de quatorze características associado com cada época (Figura 5.6). O procedimento seguinte é baseado em uma SVM multiclases com um kernel Gaussiano (*Radial Basis Function*, RBF) e o treinamento foi realizado utilizando a biblioteca de código aberto – LIBSVM (Chang e Lin 2011). Detalhes sobre esquemas baseados em classificação via SVM podem ser encontrados em (Duda, Hart e Stork 2006; Haykin 1999).

Máquina de Vetor de Suporte

As SVM's fornecem um bom desempenho de generalização em problemas de reconhecimento de padrões, no caso de padrões separáveis (Haykin 1999). Assume-se que um conjunto de treinamento é composto por pares *exemplo-rótulo* (\mathbf{x}_i, y_i) , $i = 1, \dots, \ell$, em que ℓ é o número de épocas usadas no treinamento. Vetores $\mathbf{x}_i \in \mathbb{R}^n$ são compostos

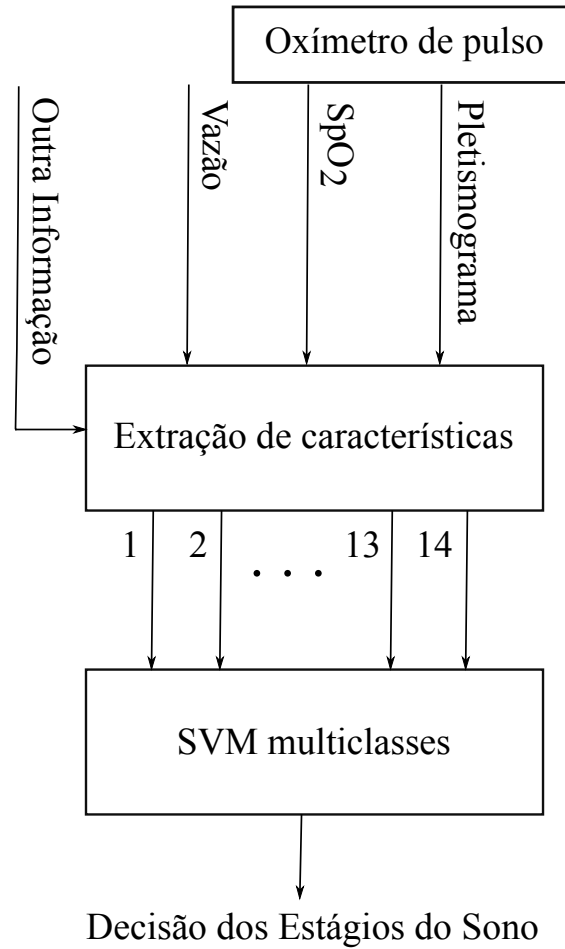


Figura 5.6: Esboço do processo de classificação.

de n características (Tabela 5.4) e $y_i \in \{1, -1\}$ é a classe rotulada (classe do sono) de \mathbf{x}_i .

Um vetor de entrada de dados $\mathbf{x} \in \mathbb{R}^n$ pode ser representado em um espaço N -dimensional ($N > n$) usando uma função vetorial $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^N$ em que a classificação pode ser realizada por meio de um hiperplano. A SVM produz uma *superfície de decisão* não-linear (fronteira de classificação não-linear) no espaço original \mathbb{R}^n por meio da construção de um hiperplano linear no espaço transformado \mathbb{R}^N (Debnath, Takahide e Takahashi 2004). Isto é conseguido tomando-se a função sinal

$$f(\mathbf{x}) = \mathbf{w} \cdot \phi(\mathbf{x}_i) + b. \tag{5.1}$$

Assim, para padrões linearmente separáveis, $\exists (\mathbf{w}, b)$ tais que as desigualdades

$$\begin{aligned} f(\mathbf{x}_i) &\geq +1 \quad \text{se } y_i = 1 \\ f(\mathbf{x}_i) &\leq -1 \quad \text{se } y_i = -1 \end{aligned} \quad (5.2)$$

sejam válidas para todos os elementos do conjunto de treinamento (Cortes e Vapnik 1995; Vapnik 1999). Os vetores que satisfazem as *igualdades* são chamados de *vetores de suporte*. Claramente, por definição, uma SVM é um classificador binário. O problema de classificação multiclass será mencionado concisamente. Os cálculos são simplificados pela formulação do funcional a ser minimizado em relação ao vetor de pesos \mathbf{w} conforme (Haykin 1999)

$$\Phi(\mathbf{w}, \epsilon) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^{\ell} \epsilon_i, \quad (5.3)$$

em que $C > 0$ é o *parâmetro de regularização* e $\epsilon_i \geq 0$ são as *variáveis de folga*. Após alguma álgebra, a função f pode ser escrita como (Haykin 1999)

$$f(\mathbf{x}) = \sum_{i=1}^{\ell} y_i \alpha_i K(\mathbf{x}, \mathbf{x}_i), \quad (5.4)$$

em que \mathbf{x}_i é o i -ésimo vetor de treinamento, $\alpha_i \geq 0$ são os *multiplicadores de Lagrange* e $K : (\mathbb{R}^N \times \mathbb{R}^N) \rightarrow \mathbb{R}$ é a *função kernel* definida como o produto interno $K(\mathbf{x}, \mathbf{x}_i) = \phi(\mathbf{x}) \cdot \phi(\mathbf{x}_i)$.

SVM Multiclass

A abordagem “um contra um” (Liu, Hao e Tsang 2008) é implementada em (Chang e Lin 2011) para problemas de classificação *multiclass*. O número de classificadores binários necessários para a classificação entre k classes é $k(k-1)/2$. Cada classificador (binário) é treinado utilizando dados de apenas duas classes diferentes. Cada vetor de dados \mathbf{x} é submetido ao conjunto de classificadores binários. No final, um vetor é designado para pertencer à classe em que recebeu o número máximo de votos. No caso de empate, escolhe-se a classe que aparece primeiro na matriz de armazenamento (Chang e Lin 2011).

Dados Desbalanceados

Devido à natureza do sono, o problema de classificação, em questão, é desbalanceado. O tempo de permanência de um paciente em um estágio do sono pode variar significativamente de um estágio para outro e o número de épocas de uma classe para outra pode ser muito diferente. Assim, o treinamento com dados desbalanceados pode resultar em classificadores polarizados. A abordagem escolhida para lidar com este problema é a *subamostragem aleatória* (Castro e Braga 2011), com a qual garante-se que cada classe (estágio de sono) tenha o mesmo número de ocorrências nos dados de treinamento. Isso foi feito para todos os conjuntos de dados de treinamento utilizados neste trabalho.

Seleção do Modelo - Kernel RBF e Validação-Cruzada

A próxima etapa consiste na escolha do núcleo (*kernel*) e dos parâmetros da SVM. Em geral, o núcleo RBF é uma primeira escolha razoável

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^{\ell} \alpha_i \exp \left\{ \frac{|\mathbf{x} - \mathbf{x}_i|^2}{\sigma^2} \right\} \right). \quad (5.5)$$

Os principais parâmetros a serem escolhidos são o par (C, σ) nas equações (5.3) e (5.5), respectivamente. Em vez de fixar um valor *a priori* para esses parâmetros, a LIBSVM foi utilizada para realizar uma busca em um espaço de parâmetros 2D para os valores mais adequados para este problema. Um problema importante é a escolha de valores ótimos para o par de valores (C, σ) . Isso é feito testando vários valores no espaço de parâmetros e fazendo a validação cruzada dos SVM's correspondentes. Aquele par relacionado com o SVM que obtém a melhor taxa de acerto é selecionado. Valores logaritmicamente espaçados foram utilizados, assim $C \in \{2^{-1}, 2^{-0,25}, 2^{0,5}, 2^{1,25}, 2^2, \dots, 2^8\}$ e $\sigma \in \{2^5, 2^{4,25}, 2^{3,5}, 2^{2,75}, 2^2, \dots, 2^{-2}\}$. Uma vez que o melhor par (C, σ) foi encontrado, o classificador final foi treinado usando tal par e todo o conjunto de dados de treinamento.

5.3.3 Classificação Dependente do Paciente

Idealmente, um classificador automático do sono deve ser capaz de trabalhar com dados de um novo paciente sem treinamento adicional, ou seja, ele deve ser *independente do paciente* (IP). O treinamento de tal classificador requer, portanto, dados derivados de um conjunto representativo de pacientes. Desse modo, quando dados de um novo paciente forem usados, desde que os seus padrões de sono sejam, de alguma forma, contabilizados nos dados de treinamento originais, uma classificação correspondente adequada pode ser esperada. Por outro lado, um sistema *dependente do paciente* (DP) baseia-se em um classificador treinado com os dados advindos de um paciente específico. É claro que, esse único paciente é muito menos representativo e, a menos que outros pacientes com padrões de sono semelhantes sejam considerados, *não* se deve esperar, necessariamente, um bom desempenho deste classificador aplicado a qualquer *outro* paciente, ainda que isso possa ocorrer. Apesar de suas limitações, um classificador dependente do paciente também é útil. Primeiro, porque serve como uma prova para a viabilidade da técnica e também porque serve como um limite superior para a taxa de acerto que se espera obter usando uma classificação independente do paciente. Em segundo lugar, um classificador pode ser treinado para um paciente específico cujo sono é monitorado por meio de seu ventilador privado.

Classificadores do sono treinados com dados de um único paciente foram analisados i) com dados de uma única noite e ii) com dados de duas noites, como discutido na sequência.

Classificação de Uma Noite

Uma SVM multiclasse foi treinada para *cada noite de cada paciente* (dois classificadores para cada paciente – com exceção do P_{13} para o qual estão disponíveis os dados correspondentes a uma única noite de sono). Dois cenários foram avaliados: um classificador de *cinco classes* usando os estágios do sono W, REM, N_1 , N_2 e N_3 e um classificador de *três classes* usando os estágios W, REM e não-REM (no qual os estágios N_1 , N_2 e N_3 foram agrupados). Assim, 25 classificadores foram treinados para cinco classes e 25

para três classes.

Cada classificador foi treinado com tipicamente 12% das épocas disponíveis e foi subsequentemente testado com todos os dados. A classificação via SVM foi considerada como sendo correta quando coincidiu com a do neurologista. Assim, para um determinado paciente e um registro de uma noite inteira dividido em M épocas, a *taxa de acerto* é dada por

$$A_{cc} = \frac{n_a}{M} \times 100\%, \quad (5.6)$$

em que n_a indica o número de épocas corretamente classificadas. A_{cc} é uma estimativa da taxa de acerto global do classificador. Para o estudo de classificação de uma noite, no final, houve 25 classificadores com as respectivas A_{cc} . Foram calculados a média e o desvio padrão de cada amostra.

Classificação de Duas Noites

O mesmo procedimento mencionado, no contexto de classificação de uma noite, foi seguido para registros de *duas noites de cada paciente* para os quais foram registradas duas noites (há somente 12). Épocas de ambas as noites foram agrupadas para formar um único conjunto de dados de maior generalidade e um único classificador para cada paciente foi treinado, com 12% das épocas. Testes foram realizados usando todos os dados das duas noites. Os dados de treinamento foram sempre balanceados. Novamente, a taxa de acerto A_{cc} em (5.6) foi calculada para cada paciente. A média e o desvio padrão desses doze valores foram também calculados.

Posteriormente, será vista na Seção 5.3.8 uma justificativa para a utilização de dados de duas noites de um mesmo paciente na classificação.

5.3.4 Classificação Independente do Paciente

Sabendo-se que as características de 11 a 14 (Tabela 5.4) variam de um paciente para o outro, elas são úteis no treinamento do classificador independente do paciente. Para a construção de tal classificador, épocas dos 13 pacientes foram agrupadas para

formar os dados de treinamento e teste. Na verdade, dois classificadores foram treinados: um para discriminar entre três classes e outro para classificar entre os cinco estágios do sono. As 14 características (Tabela. 5.4) foram extraídas para cada paciente separadamente. As características de todos os pacientes foram, então, dispostas em uma matriz de características. Posteriormente, 12 diferentes conjuntos de treinamento balanceados de mesmo tamanho foram escolhidos aleatoriamente a partir da matriz de características. Este número foi escolhido a fim de tornar o desempenho comparável ao caso da classificação de duas noites *dependente do paciente* em que 12 classificadores foram obtidos, um para cada paciente. Todos os dados da matriz de características foram usados para o teste. Depois que a taxa de acerto A_{cc} foi calculada para cada uma dessas 12 realizações, a média e o desvio padrão foram calculados.

As etapas da classificação podem ser resumidas na Figura 5.7

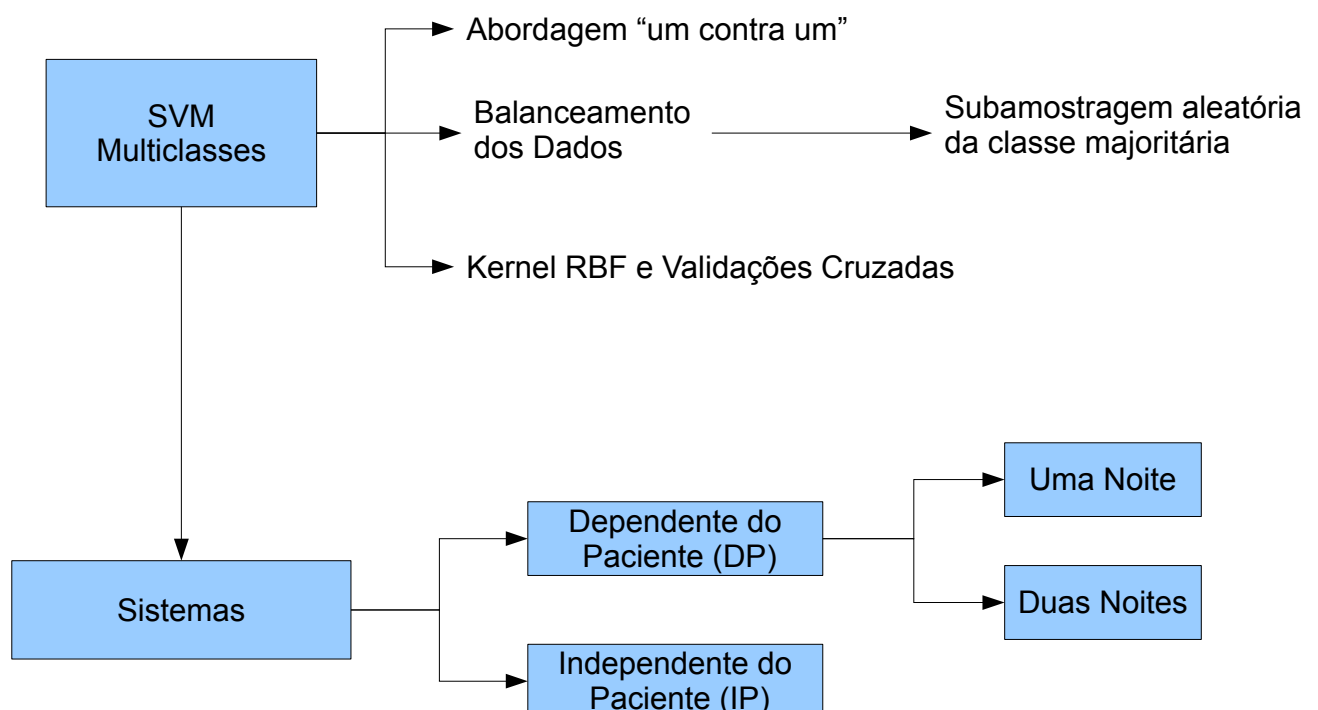


Figura 5.7: Etapas da classificação.

5.3.5 Teste de Pertinência de Características

Dois conjuntos de características utilizados, neste trabalho, não são padrões comumente usados na classificação automática do sono. Há características associadas à i) amplitude de oscilações do PPG e ii) amplitude da vazão. Com o intuito de testar a influência destas características no desempenho do classificador, testes específicos foram realizados nos quais tais características foram excluídas durante o treinamento de diferentes classificadores. O desempenho desses classificadores foi, então, comparado ao obtido com todo o conjunto de características. Três casos foram estudados:

- i) somente as características relacionadas com a amplitude do PPG foram excluídas;
- ii) somente as características relacionadas com a amplitude da vazão foram excluídas e
- iii) as características relacionadas tanto com a amplitude do PPG quanto com a amplitude da vazão foram excluídas.

Para avaliar o impacto de tais mudanças, um teste de diferenças não paramétrico (Teste de *Wilcoxon* pareado) foi aplicado para comparar as taxas de acerto antes e depois da remoção das características.

5.3.6 Efeito do Pós-Processamento no Desempenho do Classificador

Ao analisar os hipnogramas gerados pelos classificadores, que serão apresentados na Seção 5.4, observou-se que os hipnogramas codificados pela neurologista eram menos fragmentados (apresentavam menos flutuações). Essas flutuações são dadas pelo erro de classificação. Assim, com o intuito de diminuir o erro de classificação (melhorar o desempenho do classificador), ou seja, de *suavizar* os hipnogramas estimados, foi proposta uma metodologia de “pós-filtragem” dos classificadores. Inicialmente, foi aplicado um sistema de *vota* o qual classificava uma dada janela com o estágio mais votado nas subdivisões (n_{sub}) desta janela.

Exemplo 5.3.1 *Dada uma janela de 30s do sinal e escolhendo-se $n_{sub} = 6$, tem-se 6 janelas de 5s. Suponha que as três primeiras janelas e a última janela tenham sido classificadas como estágio Vigília e a quarta janela como estágio REM. O filtro vota classificaria a janela de 30s como estágio Vigília, já que este estágio foi o estágio mais votado.*

Os resultados para o *filtro vota* são apresentados na Tabela 5.10 da Seção 5.4.2. Posteriormente, optou-se por um filtro passa-baixas, mas essa metodologia foi abandonada, uma vez que os resultados ficaram piores para este caso.

5.3.7 Análise Prospectiva da Adaptação à Ventilação Não Invasiva

O objetivo da VNI é corrigir anomalias da mecânica ventilatória e/ou do comando ventilatório subjacente, logo, é necessário ajudar o paciente a se adaptar a seu ventilador (ou seja, o paciente se adapta progressivamente após otimização dos ajustes do ventilador) limitando os assincronismos afim de proporcionar um conforto respiratório e uma boa qualidade do sono. Assim, a adaptação à VNI é, portanto, uma etapa crucial do início do tratamento. É nesse momento que o paciente se habitua progressivamente a sua aparelhagem (a sua máscara, a seu ventilador e, mais globalmente, a sua patologia). A avaliação dessa adaptação é efetuada por meio do estudo das interações paciente-ventilador.

No departamento de pneumologia do CHU da Universidade de Rouen, o estabelecimento de uma VNI agendada é executada como parte de uma semana de hospitalização. Nesse departamento, a duração da adaptação é de 4 a 5 dias, frequentemente com uma estabilização e às vezes uma normalização da $PaCO_2$ diurna e noturna (em função das patologias). Os pacientes do protocolo da tese (Naeck 2011) foram, então, hospitalizados de maneira programada e efetuaram três polissonografias, uma em ventilação espontânea (VE) e duas em VNI com quatorze dias de intervalo.

O objetivo desta Seção é tentar identificar um indicador de adaptação ao tratamento ventilatório, ao comparar a 2^a e a 15^a noites sob VNI.

Percebe-se que as características potencialmente variantes no tempo são as características 5 a 8, listadas da Tabela 5.4, que estão relacionadas com a vazão do ventilador.

Assim, com o intuito de fazer uma análise “quantitativa” da adaptação ao ventilador, foram realizados os seguintes procedimentos:

Para cada noite, em questão (noites N_2 e N_{15}), foram feitas duas classificações para cada paciente:

1. *Classificador Pleno* (CP): classificador que utiliza as características de 1 – 10, listadas na Tabela 5.4;
2. *Classificador Independente da Vazão* (CIV): classificador sem as características relacionadas com a vazão (5 – 8).

Cada classificação foi repetida 12 vezes obtendo-se dois vetores de taxas de acerto para cada noite: $Acc_{pleno02}$, Acc_{iv02} , $Acc_{pleno15}$, Acc_{iv15} . Para cada noite, calculou-se a diferença (Δ) entre as taxas de acerto obtidas pelo CP e pelo do CIV, ou seja:

$$\begin{aligned}\Delta_{02} &= Acc_{pleno02} - Acc_{iv02} \\ \Delta_{15} &= Acc_{pleno15} - Acc_{iv15}\end{aligned}\tag{5.7}$$

Por fim, para verificar se havia diferenças significativas entre Δ_{02} e Δ_{15} aplicou-se o Teste de *Wilcoxon* nas distribuições dos Δ 's. Os resultados gerados por este método estão listados na Tabela 5.11.

5.3.8 Estimação de Hipnogramas de Noites sem Polissonografias

Como foi informado na Seção 5.2.2, não foram realizadas polissonografias durante as noites N_3 e N_4 . Foram coletadas somente as poligrafias. Assim, não há classificação realizada pelo neurologista para estas noites.

A partir dos modelos (do classificador) obtidos, em conjunto, para as noites N_2 e N_{15} (caso dependente do paciente, duas noites), estimou-se os hipnogramas das noites N_3 e N_4 , treinando os modelos com os dados das noites N_3 e N_4 . Os hipnogramas estimados são apresentados nas figuras A.1 a A.4 do Apêndice A.

A importância desta técnica está na obtenção de hipnogramas via modelos (SVM's multiclasse), sem a necessidade de uma classificação inicial de um especialista. Isso seria extremamente útil para se arquivar informações de um determinado paciente. Bastaria, por exemplo, que fosse realizada apenas uma polissonografia completa do paciente, já que para noites seguintes (caso o paciente precisasse retornar a um laboratório do sono) os hipnogramas poderiam ser estimados. Isso justifica o uso de dados de duas noites de um mesmo paciente na classificação.

Os hipnogramas estimados têm sido utilizados pelo grupo *Atomosyd* com o intuito de fazer uma análise da fragmentação do sono a partir do cálculo das entropias dos hipnogramas. A análise via entropia vem sendo realizada por (Bounoaire 2012), como parte do trabalho de sua tese.

5.4 Resultados

5.4.1 Classificação

Nas tabelas 5.5 e 5.6 são relatados os resultados para a classificação de cinco e três estágios, respectivamente, para ambos os cenários: caso dependente do paciente (DP) e caso independente do paciente (IP). O hipnograma gerado pelo classificador de cinco estágios para uma noite do caso paciente dependente é mostrado na Figura 5.8. A taxa de acerto, neste caso, é aproximadamente 92%. Pode-se observar que o hipnograma automaticamente codificado é mais fragmentado do que aquele produzido por um especialista do sono. Ao contrário da classificação visual/manual, na qual a continuidade de uma época para outra é subjetivamente tomada em conta, a SVM classifica cada época de forma independente; isso explica a fragmentação da classificação via SVM's.

As tabelas 5.7 e 5.8 mostram os resultados que concernem à classificação automática do sono para cinco e três estágios, respectivamente. O primeiro terço da Tabela corresponde à classificação *sem* características associadas às amplitudes do pletismograma e da vazão (características 3, 4, 7 e 8 na Tabela 5.4). No segundo terço, as características associadas com a amplitude do pletismograma *não* foram utilizadas (características 3 e 4 na Tabela 5.4), mas sim as características associadas à amplitude da vazão. E no último terço, *não* foram utilizadas as características associadas com a amplitude da vazão (características 7 e 8 na Tabela 5.4), mas sim as características associadas à amplitude do pletismograma.

Com o intuito de verificar se havia diferenças significativas entre os resultados das tabelas 5.7 e 5.8 comparadas, respectivamente, com as tabelas 5.5 e 5.6, foi aplicado o teste de *Wilcoxon* pareado. Os resultados são mostrados na Tabela 5.9, que mostra que a maioria das diferenças são estatisticamente significativas ($p \leq 0,05$).

5.4.2 Pós-Processamento do Classificador

A Tabela 5.10 mostra os resultados da classificação com filtro para 3 e 5 estágios do sono, utilizando todas as características listadas na Tabela 5.4.

Tabela 5.5: Resultados para 5 estgios (classes) do sono. \bar{A}_{cc} representa o valor mdio de A_{cc} definido em (5.6). \bar{k} representa a mdia do coeficiente Kappa Cohen. DP: dependente do paciente; IP: independente do paciente e o nmero entre parentesis representa o nmero de noites registradas.

P(N)	$\bar{A}_{cc} \pm \sigma$ (%)	$\bar{k} \pm \sigma$	Caractersticas
DP(1)	$84 \pm 5,0$	$0,78 \pm 0,06$	1-10
DP(2)	$73 \pm 5,2$	$0,64 \pm 0,07$	1-10
IP(2)	$62 \pm 0,85$	$0,51 \pm 0,01$	1-14

Tabela 5.6: Resultados para 3 estgios (classes) do sono. Veja as definies na legenda da Tabela 5.5

P(N)	$\bar{A}_{cc} \pm \sigma$ (%)	$\bar{k} \pm \sigma$	Caractersticas
DP(1)	$91 \pm 2,5$	$0,83 \pm 0,05$	1-10
DP(2)	$84 \pm 3,5$	$0,71 \pm 0,06$	1-10
IP(2)	$78 \pm 0,74$	$0,64 \pm 0,01$	1-14

5.4.3 Adaptao  VNI

A Tabela 5.11 mostra os resultados da aplicao do Teste de *Wilcoxon* s distribues dos Δ 's das noites N_2 e N_{15} de cada paciente.

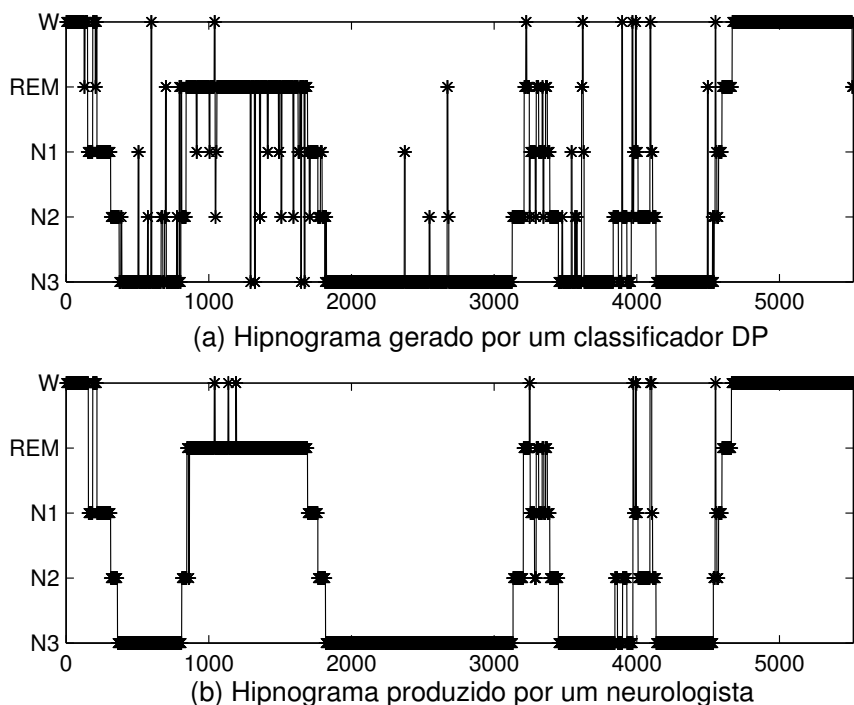


Figura 5.8: Hipnograma produzido por um classificador dependente do paciente (a) Segundo um neurologista para o cenário de uma noite. (b) Caso do paciente P_{12} . Neste caso, $\bar{A}_{cc} \approx 92\%$.

Tabela 5.7: Resultados para 5 estágios (classes) do sono retirando características específicas. Veja legenda da Tabela 5.5 para definições.

P(N)	$\bar{A}_{cc} \pm \sigma(\%)$	$\bar{k} \pm \sigma$	Características
DP(1)	$80 \pm 5,5$	$0,73 \pm 0,07$	1, 2, 5, 6, 9, 10
DP(2)	$67 \pm 5,6$	$0,76 \pm 0,08$	1, 2, 5, 6, 9, 10
IP(2)	$57 \pm 0,96$	$0,44 \pm 0,03$	1, 2, 5, 6, 9–14
DP(1)	$83 \pm 5,8$	$0,76 \pm 0,07$	1, 2, 5–10
DP(2)	$71 \pm 5,6$	$0,61 \pm 0,06$	1, 2, 5–10
IP(2)	$59 \pm 0,68$	$0,47 \pm 0,01$	1, 2, 5–14
DP(1)	$83 \pm 4,2$	$0,76 \pm 0,05$	1 a 6, 9, 10
DP(2)	$72 \pm 6,2$	$0,62 \pm 0,07$	1 a 6, 9, 10
IP(2)	$60 \pm 0,63$	$0,49 \pm 0,01$	1 a 6, 9–14

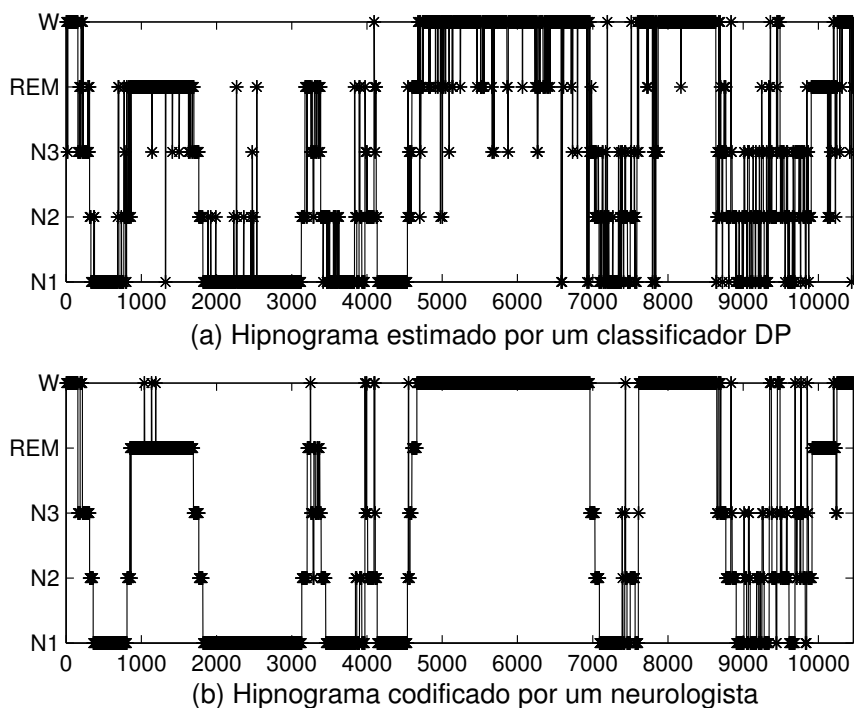


Figura 5.9: Hipnograma produzido por um classificador dependente do paciente (a) Segundo um neurologista para o cenário de duas noites. (b) Caso do paciente P_{12} . Neste caso, $\bar{A}_{cc} \approx 83\%$.

Tabela 5.8: Resultados para 3 estágios (classes) do sono retirando características específicas. Veja legenda da Tabela 5.5 para definições.

P(N)	$\bar{A}_{cc} \pm \sigma(\%)$	$\bar{k} \pm \sigma$	Características
DP(1)	$87 \pm 4,3$	$0,76 \pm 0,08$	1, 2, 5, 6, 9, 10
DP(2)	$79 \pm 4,0$	$0,62 \pm 0,08$	1, 2, 5, 6, 9, 10
IP(2)	$74 \pm 1,4$	$0,56 \pm 0,02$	1, 2, 5, 6, 9–14
DP(1)	$90 \pm 3,3$	$0,80 \pm 0,06$	1, 2, 5–10
DP(2)	$82 \pm 4,3$	$0,67 \pm 0,07$	1, 2, 5–10
IP(2)	$76 \pm 0,83$	$0,60 \pm 0,01$	1, 2, 5–14
DP(1)	$90 \pm 2,9$	$0,81 \pm 0,05$	1 a 6, 9, 10
DP(2)	$83 \pm 3,6$	$0,68 \pm 0,07$	1 a 6, 9, 10
IP(2)	$77 \pm 0,61$	$0,61 \pm 0,01$	1 a 6, 9–14

Tabela 5.9: Teste pareado de *Wilcoxon* para significância estatística dos diferentes desempenhos quando características específicas *não* são utilizadas. O asterisco indica os valores que *não* são estatisticamente significativos.

P(N)	p -valor (5 estágios)	p -valor (3 estágios)
DP(1)	$1,2 \times 10^{-5}$	$1,8 \times 10^{-5}$
DP(2)	$4,88 \times 10^{-4}$	$4,88 \times 10^{-4}$
IP(2)	$4,88 \times 10^{-4}$	$4,88 \times 10^{-4}$
DP(1)	0,0397	0,0068
DP(2)	0,0640*	$9,77 \times 10^{-4}$
IP(2)	$4,88 \times 10^{-4}$	$4,88 \times 10^{-4}$
DP(1)	0,0032	0,0042
DP(2)	0,0522*	0,0640*
IP(2)	$9,77 \times 10^{-4}$	$4,88 \times 10^{-4}$

Tabela 5.10: Resultados para 5 e 3 estágios (classes) do sono para a classificação com filtro. Veja legenda da Tabela 5.5 para definições.

P(N)	$\bar{A}_{cc} \pm \sigma$ (%)	$\bar{k} \pm \sigma$	Características	Classes
DP(1)	$84 \pm 5,0$	$0,78 \pm 0,06$	1-10	5
DP(2)	$70 \pm 4,9$	$0,59 \pm 0,06$	1-10	5
IP(2)	$61 \pm 0,9$	$0,50 \pm 0,01$	1-14	5
DP(1)	$91 \pm 2,5$	$0,83 \pm 0,05$	1-10	3
DP(2)	$81 \pm 5,0$	$0,66 \pm 0,09$	1-10	3
IP(2)	$78 \pm 0,7$	$0,63 \pm 0,01$	1-14	3

Tabela 5.11: Teste de *Wilcoxon* pareado para significância estatística entre os Δ 's das noites N_2 e N_{15} de cada paciente para classificações de 5 e 3 estágios. Os números em negrito indica os valores que *são* estatisticamente significativos.

P_n	p -valor (5 estágios)	p -valor (3 estágios)
P_1	0,2036	0,1763
P_2	0,7910	0,7334
P_3	0,3013	0,0093
P_4	0,7334	0,0034
P_6	0,0425	0,9697
P_7	0,4238	0,0342
P_8	0,3394	0,5693
P_9	0,4697	0,0342
P_{10}	0,2036	0,9697
P_{11}	0,0923	0,2036
P_{12}	0,2334	0,4697
P_{14}	0,1514	0,0771

5.5 Discussão

Sinais cardio-respiratórios foram utilizados em (Redmond e Heneghan 2006) para verificar se as medidas obtidas via sinais de ECG e respiração poderiam fornecer uma classificação dos estágios de vigília, REM e não-REM. O objetivo principal era aumentar o uso de sistemas cardiorrespiratórios ambulatoriais para a detecção de distúrbios respiratórios do sono. Nesse estudo, foi considerado um grupo de 37 pacientes com apneia obstrutiva do sono. Para cada paciente, 27 características foram extraídas de três sinais: séries de intervalos RR, o eletrocardiograma derivado da respiração e o esforço respiratório. Utilizando um Discriminante Quadrático obteve-se uma taxa de acerto de 79% ($\kappa = 0,56$) para o sistema dependente do paciente e de 67% ($\kappa = 0,32$) para o sistema independente do paciente. Os resultados obtidos, no presente trabalho, para um problema de classificação similar, mas para diferentes tipos de pacientes (Tabela 5.6) revelou uma alta taxa de acerto 91% ($\kappa = 0,83$) e 78% ($\kappa = 0,64$), para os respectivos casos.

Medidas cardio-respiratórias foram também utilizadas em (Harper, Schechtman e Kluge 1987) para discriminar entre os estágios de vigília, REM e *sono tranquilo* de 35 lactentes saudáveis. Sete características cardio-respiratórias foram utilizadas, resultando nas seguintes taxas de acerto: 80,0% quando foram utilizadas apenas características baseadas na respiração, 82,0% usando apenas características cardiovasculares e 84,8% quando ambos os conjuntos de características foram empregados. Os sinais de ECG e esforço respiratório (pletismograma de tórax) de seis indivíduos saudáveis foram utilizados em (Karlen, Mattiussi e Floreano 2009) para discriminar somente entre vigília e sono (um problema binário de classificação). As características consistiram no logaritmo da densidade espectral de potência dos sinais de ECG e respiração originais. Utilizando redes neurais artificiais, foram alcançadas as seguintes taxas de acerto: 95,4% para o caso dependente do paciente e de 85,3% para o caso independente do paciente. Comparando com os resultados apresentados nas tabelas 5.5 e 5.6 àqueles em (Harper, Schechtman e Kluge 1987) e (Karlen, Mattiussi e Floreano 2009), deve-se lembrar que, no presente trabalho, os padrões do sono foram classificados em 3 e 5

estágios usando uma técnica *sem eletrodos*.

A inclusão de certas características resultou em uma melhoria estatisticamente significativa na taxa de acerto (Tabela 5.9). A melhoria média foi, aproximadamente de, 1% (resp. 1,7%) quando características relacionadas à vazão foram incluídas; de 1,5% (resp. 2%) quando características relacionadas ao PPG foram incluídas e de 4,5% (resp. 5%) quando ambas as classes de características foram incluídas para os problemas de classificação de três (resp. cinco estágios). A maior vantagem é atingida quando se incluem ambas as características. Esses resultados sugerem que tais características contêm informação a respeito do sono que não pode ser recuperada por meio de outras características e, conseqüentemente, sua inclusão no treinamento pode resultar em classificadores automáticos do sono mais eficientes que não requerem eletrodos.

Ao se compararem os resultados apresentados na Tabela 5.10 com os das tabelas 5.5 e 5.6 verifica-se que as taxas de acerto obtidas tanto para 3 quanto para 5 classes com o *classificador com filtro* foram menores ou aproximadamente iguais às obtidas com o *classificador padrão*. Isso sugere que o filtro aplicado não é um filtro ideal, uma vez que não melhorou o desempenho de classificação. Assim, mais estudos são necessários na busca por um filtro ideal. Algumas sugestões são apresentadas no Cap 6.

Os resultados da Tabela 5.11 sugerem que a maioria dos pacientes está *bem adaptada ao ventilador*, já que na maioria dos casos, as diferenças entre as N_2 e N_{15} não são significativas. Um estudo realizado por (Naeck 2011), usando índices de fragmentação do sono para mesma base de dados, também mostrou que os pacientes estão bem adaptados. Logo, pode-se sugerir que a técnica/análise feita, na seção 5.3.7, seja um bom indicador da adaptação ao ventilador. Entretanto, mais estudos são necessários para confirmar tais resultados de modo mais amplo.

Essa técnica justifica também a classificação de 02 noites de um mesmo paciente, pois como não houve diferenças significativas entre a noite N_2 e a N_{15} , sugere-se que não há problema misturar os dados de duas noites para treinar um classificador, uma vez que isso não alterará o resultado da classificação.

5.6 Conclusões do Capítulo

Os resultados relatados mostram que é possível desenvolver um procedimento apropriado para a classificação automática dos estágios do sono por meio de sinais prontamente *disponíveis a um ventilador*. Isso abre a possibilidade para uma melhor avaliação da qualidade do sono em tratamento sob VNI uma vez que pode permitir a avaliação do sono em domicílio num período, a longo prazo, sem a utilização de eletrodos e nem a colocação de sensores adicionais nos pacientes. No futuro, isso pode ser ainda possível para os próprios ventiladores a fim de proporcionar um relatório do sono juntamente com padrões de conformidade e relatórios de diagnóstico tais como os que as máquinas já produzem. Outra possibilidade para o ventilador seria no uso de informações sobre o sono do paciente para ajustar/configurar as suas definições e, assim, melhorar a qualidade do tratamento.

Conclusões e Propostas de Continuidade

“Ergueu no patamar quatro paredes mágicas. Tijolo com tijolo num desenho lógico.”

Construção, de Chico Buarque

Uma metodologia baseada em técnicas de reconhecimento de padrões foi apresentada, neste trabalho, com o intuito de resolver dois problemas de classificação. O primeiro evoluiu na busca de características extraídas a partir de séries simbólicas para discriminar entre pacientes saudáveis, chagásicos cardiopatas e chagásicos não cardiopatas. O segundo, na busca de uma ferramenta de auxílio na classificação dos estágios dos sonos que visasse a sua aplicação no domicílio do paciente. Para isso, a contribuição/inação do trabalho consistiu na utilização de sinais disponíveis a um ventilador. Em síntese, os resultados obtidos nos capítulos 4 e 5 mostraram que foi possível:

- i) classificar 3 e 5 classes do sono, com taxas de concordância equiparáveis à concordância entre dois neurologistas;
- ii) estimar hipnogramas para noites sem polissonografias. Esses resultados foram utilizados por Bounouaire, 2012 para obter um índice de fragmentação do sono;
- iii) obter um indicador de adaptação à VNI;
- iv) verificar que características relacionadas com o PPG e com a vazão são relevantes para a classificação dos estágios do sono.

Com a obtenção dos índices listados nos itens ii) e iii), pode-se sugerir que foi possível avaliar a qualidade do sono.

Assim, conclui-se que foi possível atender aos objetivos propostos no Capítulo 1, uma vez que foi possível:

- classificar os estágios do sono com sinais coletados sem a utilização de eletrodos, atendendo aos objetivos propostos, neste trabalho.
- classificar pacientes saudáveis e chagásicos a partir de séries simbólicas da VFC.

6.1 Propostas de Continuidade

Acredita-se que as técnicas de reconhecimento de padrões inicialmente utilizadas, neste trabalho, sejam uma aplicação futura para o estudo da doença de Chagas. Um projeto de pós-doutorado vem sendo desenvolvido pela autora, deste trabalho, juntamente com o prof. Luiz Antônio Ribeiro, do HC-UFMG, no qual propõe-se utilizar essas ferramentas aliadas a outras técnicas para a predição de arritmia ventricular maligna na cardiopatia chagásica. Isso seria de grande utilidade, já que ao se fazer um prognóstico da evolução da doença, evitar-se-ia a implantação de cardiodesfibriladores desnecessariamente em pacientes que não correm risco de morte súbita.

De posse da base de dados de sinais polissonográficos, vê-se que outras questões ainda estão abertas e mais análises podem ser feitas, como a classificação de microdespertares. Acredita-se que essa análise propiciaria resultados também inovadores no âmbito da ciência do sono.

Na Seção 5.3.6 foi feita uma tentativa de melhoria no desempenho da classificação ao introduzir-se um classificador com filtro, mas foi verificado que esse filtro não foi o ideal, uma vez que quase não houve mudança nos resultados (mesmas taxas de acerto obtidas com o classificador padrão e com o classificador com filtro ou taxas de acerto menores obtidas com o classificador com filtro). Assim, propõe-se uma nova metodologia ao usar uma característica dinâmica do sistema que estaria relacionada com o estágio do sono do paciente. Por exemplo, ao treinar o classificador seria incluída

como característica o estágio do sono classificado pela neurologista na janela anterior. Isso poderia ter um efeito de filtro diminuindo as fragmentações (flutuações) geradas no hipnograma estimado pelo classificador (efeito de *suavização*). Essas flutuações são devidas aos erros de classificação. Essa ideia está implícita nas cadeias ocultas de Markov (HMM) que pressupõem tanto a informação *a priori* quanto *a posteriori*, ou seja, pressupõem a probabilidade de mudança de um estágio para o outro (estágio transitório); logo, esse classificador poderia apresentar um melhor desempenho, gerando menores erros de classificação (maiores taxas de acerto) e suavizando, assim, o hipnograma estimado. Entretanto, isso pode ser feito com o próprio SVM ao introduzirmos a própria classificação da neurologista como característica em cada janela de dados de treinamento.

Exemplo 6.1.1 *Se a classificação da neurologista fosse vigília para uma dada janela, acrescentaríamos esta característica na janela posterior para treinar o classificador, fazendo o mesmo para as janelas seguintes e assim sucessivamente até varrer todo o sinal. A primeira janela teria as 14 características e 1 característica relacionada com o estágio de vigília (já que, inicialmente, o paciente estava acordado). Supondo que a classificação da neurologista para a segunda janela fosse estágio REM; essa informação seria utilizada como característica da terceira janela, e assim sucessivamente.*

6.1.1 Limitações do Trabalho

Apesar de se terem alcançado os objetivos propostos no Capítulo 1, mais estudos são necessários, diante de algumas limitações do trabalho:

- é necessário que se realizem experimentos com maior número de pacientes para que se tenha uma amostra representativa;
- é necessário que sejam feitas melhorias no classificador, por exemplo, o fornecimento de índices como a sensibilidade e a especificidade;
- construção de um *software* que possa ser adaptado ao sistema de aquisição de dados ou ao ventilador, que gerasse a classificação automática diretamente, base-

ando nas rotinas construídas neste trabalho (melhor automatização das rotinas), ou seja, o produto final seria: conectar o ventilador a um paciente e após uma noite de sono ter-se-ia um hipnograma.

Bibliografia

- Achour, L., C. Letellier, A. Cuvelier, E. Vérin e J.-F. Muir (2007). "Asynchrony and cyclic variability in pressure support noninvasive ventilation," *Computer in Biology and Medicine* 37, pp. 1308–1320.
- Acyr, N. e C. Guzelis (2004). "Automatic spike detection in EEG by a two-stage procedure based on support vector machines". *Computers in Biology and Medicine*.
- Agarwal, R. e J. Gotman (2001). "Computer-assisted sleep staging". *IEEE Transactions on Biomedical Engineering* 48.12, pp. 1412–23.
- Aires, M. M. (1999). *Fisiologia*. 2nd. Guanabara Koogan.
- Angelis, K. de, B. A. Schaan, B. Rodrigues, C. Malfitano e M. C. Irigoyen (2007). "Disfunção Autonômica Cardiovascular no Diabetes Mellitus Experimental". *Arq Bras Endocrinol Metab* 51, pp. 185–194.
- Aserinsky, E. e N. Kleitman (1955). "Two types of ocular motility occurring in sleep". *Journal of Applied Physiology* 8, pp. 1–10.
- Asl, B. M., S. K. Setarehdan e M. Mohebbi (2008). "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal". *Artificial Intelligence in Medicine* 44, pp. 51–64.
- Association, American Sleep Disorders (1992). "EEG arousals : scoring rules and examples, a preliminary report from the Sleep Disorders Atlas Task Force". *Sleep* 15.2, pp. 173–184.
- Babloyantz, A. e A. Destexhe (1988). "Is the normal heart a periodic oscillator?" *Biol. Cybern.* 58, pp. 203–211.
- Bai-Lin, H. (1989). *Elementary Symbolic Dynamics – and Chaos in Dissipative Systems*. Ed. por World Scientific. World Scientific, Singapore.
- Bandt, C. (2005). "Ordinal time series analysis". *Ecological Modelling* 182, pp. 229–238.
- Barbé, Ferran, Maria Antonia Quera-Salva, Jacques de Lattre, Philippe Gajdos e Alvar G.N. Agustí (1996). "Long-term Effects of Nasal Intermittent Positive-Pressure Ventilation on Pulmonary Function and Sleep Architecture in Patients With Neuro-muscular Diseases". *Chest* 110.5, pp. 1179–1183.

- Barcaro, U., E. Bonanni, M. Maestri, L. Murri, L. Parrino e M. G. Terzano (2004). "A general automatic method for the analysis of NREM sleep microstructure". *Sleep Medicine* 5, pp. 567–76.
- Baselli, G., A. Porta e M. Pagani (2006). "Coupling arterial windkessel with peripheral vasomotion: modeling the effects on low-frequency oscillations". *IEEE Transactions on Biomedical Engineering* 53, pp. 53–64.
- Baselli, G., S. Cerutti, M. Livraghi, C Meneghini, M. Pagani e O. Rimoldi (1988). "Causal relationship between heart rate and arterial blood pressure variability signals". *Med Biol Eng Comput* 26, pp. 374–378.
- Baselli, G., A. Porta, S. Cerutti, Caiani, D. Lucini e M. Pagani (2001). "RR-arterial presrue variability relationships". *Autonomic Neuroscience Basic and Clinical* 90, pp. 57–65.
- Baselli, G., E. Caiani, A. Porta, N. Montano, M. G. Signorini e S. Cerutti (2002). "Biomedical Signal Processing and Modeling in Cardiovascular Systems". *Biomedical Engineering* 30, pp. 55–84.
- Baumert, M., N. Wessel, A. Schirdewan, A. Voss e D. Abbot (2007). "Scaling Characteristics of Heart Rate Time Series Before the Onset of Ventricular Tachycardia". *Arq Bras Endocrinol Metab* 35.2, pp. 201–207.
- Berne, R. M., M. N. Levy, B. M. Koeppen e B. A. Stanton (2004). *Fisiologia*. 5th. Rio de Janeiro, RJ.: Elsevier Mosby.
- Biolo, A., Ribeiro A. L. e Clausell N. (2010). "Chagas cardiomyopathy - Where do we stand after a hundred years?" *Prog Cardiovasc Dis* 52, pp. 300–316.
- Blake, H., R.W. Gerard e N Kleitman (1939). "Factors influencing brain potentials during sleep". *Journal of Neurophysiology* 2, pp. 48–60.
- Bolfarine, H. e M. C. Sandoval (2000). *Introdução à Inferência Estatística*. SBM.
- Bonnet, M. H. e D. L. Arand (2007). "EEG Arousal Norms by Age Michael H." *Journal of Clinical Sleep Medicine* 3.3, pp. 271–274.
- Boselli, M., L. Parrino, A. Smerieri e M.G. Terzano (1998). "Effect of age on EEG arousals in normal sleep". *Sleep* 21, pp. 361–367.
- Bounoaire, D. (2012). "Vers une évaluation de l'assistance ventilatoire non invasive synchronisabilité des ventilateurs et qualité du sommeil". Tese de doutorado. Université de Rouen, France.
- Brovelli, A., M. Ding, A. Ledberg, Y. Chen, R. Nakamura e S. L. Bressler (2004). "Beta oscillations in a large-scale sensorimotor cortical network : directional influences revealed by Granger causality". *Proc. Natl. Acad. Sci.* 101, pp. 9849–9854.
- Bunde, A., S. Havlin, J. W. Kantelhardt, T. Penzel, J. H. Peter e K. Voigt (2000). "Correlated and Uncorrelated Regions in Heart-Rate Fluctuations during Sleep". *PHYSICAL REVIEW LETTERS* 85, pp. 3736–3739.
- Cacoub, P. (2006). "Fréquence cardiaque et mortalité cardiovasculaire". *La Revue de Médecine Interne* 27, pp. 727–729.

- Caffarel, J., G.J. Gibson, J.P. Harrison, C.J. Griffiths e M.J. Drinnan (2006). "Comparison of manual sleep staging with automated neural network-based analysis in clinical practice". *Medical Biological Engineering Computing* 44.1-2, pp. 105–10.
- Carskadon, M.A., E.D. Brown e W.C. Dement (1982). "Sleep fragmentation in the elderly: relationship to daytime sleep tendency". *Neurobiology Aging* 3, pp. 321–327.
- Casella, George e Roger L. Berger (1990). *Statistical Inference*. Duxbury Press.
- Castro, C. L. e A. P. Braga (2011). "Aprendizado supervisionado com conjuntos de dados desbalanceados." *Revista Controle & Automação* 22.5, pp. 441–466.
- Cerutti, S. e M. G. Signorini (2002). "Non Linear Advanced Methods for Biological Signal Analysis". In: *Proceedings of the Second Joint EMBS/BMES Conference*. 1: 015101, pp. 23–26.
- Chaffin D. G., Goldberg C. C. e K. L. Reed (1991). "The dimension of chaos in the fetal heart rate". *Am. J. Obstet. Gynecol.* 165, pp. 1425–1429.
- Chang, Chih-Chung e Chih-Jen Lin (2011). "LIBSVM: A library for support vector machines". *ACM Transactions on Intelligent Systems and Technology* 2 (3), 27:1–27:27.
- Chisci, L., A. Mavino, G. Perferi, M. Sciandrone, C. Anile, G. Colicchio e F. Fuggetta (2010). "Real-time epileptic seizure prediction using AR models and support vector machines". *IEEE Transactions on Biomedical Engineering* 57.5, pp. 1124–32.
- Choi, S. H., h. K. Kim e h. S. Lee (2000). "Speech recognition using quantized LSP parameters and their transformations in digital communication". *Speech Communication* 20, pp. 223–233.
- Chua, C-P, J Garvey, S Redmond, C Heneghan e W T McNicholas (2007). "Towards automated sleep state estimation using a Holter-oximeter." In: *Proceedings of the 29th Annual International - Conference of the IEEE EMBS*, pp. 3998–4001.
- Chung, Gih Sung, Byung Hoon Choi, Jin-Seong Lee, Jeong Su Lee, Do-Un Jeong e Kwang Suk Park (2009). "REM sleep estimation only using respiratory dynamics." *Physiological Measurement* 30.12, pp. 1327–40.
- Cohen, Jacob (1960). "A coefficient of agreement for nominal scales". *Educational and Psychological Measurement* 20.1, pp. 37–46.
- Cortes, C e V Vapnik (1995). "Support-Vector Networks". *Machine Learning* 20, pp. 273–297.
- Crisler, S., M. J. Morrissey, A. M. Anch e D. W. Barnett (2008). "Sleep-stage scoring in the rat using a support vector machine". *Journal of Neuroscience Methods* 168.2, pp. 524–34.
- Danker-Hopfe, H., P. Anderer, J. Zeitlhofer, M. Boeck, H. Dorn, G. Gruber, E. Hiller, E. Loretz, D. Moser, S. Parapatics, B. Saletu, A. Schimidt e G. Droffner (2009). "Interrater reliability for sleep scoring according to the Rechtschaffen and Kales and the new AASM standard". *J. Sleep Res* 18, pp. 74–84.
- Dauvilliers, Y. e Billiard (2004). "Aspects du sommeil normal - Normal aspects of sleep". *EMC-Neurologie* 1, pp. 458–480.

- Davila, D. F., G. Inglessis e D. C. A. Mazzei (1998). "Chagas' heart disease and the autonomic nervous system". *Int J Cardiol*.
- Davis, H., P.A. Davis, A.L. Loomis, E.N. Harvey e G. Hobart (1938). "Human brain potentials during the outset of sleep". *Journal of Neurophysiology* 24, pp. 24–38.
- Debnath, R., N. Takahide e H. Takahashi (2004). "A decision based one-against-one method for multi-class support vector machine". *Pattern Analysis and Applications* 7.2, pp. 164–175.
- Dement, W.C. e G. Tononi (1957). "Cyclic variations of EEG during sleep and their relations to eye movements, body motility and dreaming". *EEG Clin. Neurophysiol.* 9.4, pp. 673–690.
- Duda, R. O., P. E. Hart e D. G. Stork (2006). *Pattern Classification*. 2nd. NY.
- Eke, A., P. Herman, L. Kocsis e R. Kozak L. (2002). "Fractal characterization of complexity in temporal physiological signals". *Physiol Meas* 23, R1–R38.
- Elliott, Mark W. (2011). "Non-invasive ventilation during sleep: time to define new tools in the systematic evaluation of the technique". *Thorax* 66.1, pp. 82–84.
- Elliott, M.W., A.K. Simonds, M.P. Carroll, J.A. Wedzicha e M.A. Branthwaite (1992). "Domiciliary nocturnal nasal intermittent positive pressure ventilation in hypercapnic respiratory failure due to chronic obstructive lung disease: effects on sleep and quality of life". *Thorax* 47.5, pp. 342–348.
- Estévez, P.A., C.M. Held, C.A. Holzmann, C.A. Perez, J.P. Pérez, J. Heiss, M. Garrido e P. Peirano (2002). "Polysomnographic pattern recognition for automated classification of sleep-waking states in infants". *Medical Biological Engineering Computing* 40.1, pp. 105–13.
- Faes, L., G. Nollo e K. H. Chon (2008). "Assessment of Granger Causality by Nonlinear Model Identification: Application to Short-term Cardiovascular Variability". *Annals of Biomedical Engineering* 36.3, pp. 381–395.
- Fiedler-Ferrara, N. e C. P. C. Prado (1994). *Caos - Uma Introdução*. USP, São Paulo: Ed. Edgar Blücher.
- Figueiredo, E., A. A. Paola, R. M. F. L. Silva, M. Z. P. Távora e E. E. Martinez (1996). "Variabilidade da frequência cardíaca como preditor de taquicardia ventricular sustentada e morte súbita na cardiopatia chagásica". *Arq Bras Cardiol* 66.5, p. 66.
- Fishgold, H. e P. Mathis (1959). "Obnubilations, coma et stupeurs. Etudes électroencéphalographiques". *Electroenceph. Clin. Neurophysiol.* 1, pp. 1–126.
- Gagnadoux, Frédéric, Nathalie Pelletier-Fleury, Carole Philippe, Dominique Rakotonanahary e Bernard Fleury (2002). "Home Unattended vs Hospital Telemonitored Polysomnography in Suspected Obstructive Sleep Apnea Syndrome". *Chest* 121.3, pp. 753–758.
- Gautama, T., D. P. Mandic e M. M. Van Hulle (2004). "The delay vector variance method for detecting determinism and nonlinearity in time series". *Physica D* 190, pp. 167–176.

- Gerla, V., K. Paul, L. Lhotska e V. Krajca (2009). "Multivariate analysis of full-term neonatal polysomnographic data". *IEEE Transactions on Information Technology in Biomedicine* 13.1, pp. 104–10.
- Gilmore, R. (1998). "Topological analysis of chaotic dynamical systems". *Reviews of Modern Physics* 70.4, pp. 1455–1475.
- Glass, L. e M. C. Mackey (1998). *From clocks to chaos: the rhythms of life*. Princeton, NJ: Princeton: University Press.
- Godelle, J. e C. Letellier (2000). "Symbolic sequence statistical analysis for free liquid jets". *Physical Review E* 62.6, pp. 7973–7981.
- Goldberger, A. L. (1991). "Is the normal heartbeat chaotic or homeostatic?" *News Physiol. Sci.* 6, pp. 87–91.
- Goldberger, A. L. e B. J. West (1987). "Applications of nonlinear dynamics to clinical cardiology". *Ann. N Y Acad. Sci* 504, pp. 195–213.
- Goldberger A. L., Rigney D. R. Mietus J.-Antman E. M. e S. Greenwald (1988). "Non-linear dynamics in sudden cardiac death syndrome: heart rate oscillations and bifurcations". *Experientia* 44, pp. 983–987.
- Gomes, M. E. D. (2001). "Técnicas de Sistemas Dinâmicos Não-Lineares na Análise da Modulação Autônômica da Variabilidade da Frequência Cardíaca". Tese de doutorado. Belo Horizonte, MG: Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Gomes, M. E. D., A. V. P. Souza, H. N. Guimarães e L. A. Aguirre (2000). "Investigation of determinism in heart rate variability". *Chaos* 10.2, pp. 398–410.
- Grassberger, P. e I. Procaccia (1983). "Measuring the strangeness of strange attractors". *Physica D* 9, pp. 189–208.
- Gudmundsson S. and Runarsson, T. P. e S. Sigurdsson (2012). "Test-retest reliability and feature selection in physiological time series classification". *Computer Methods and Programs in Biomedicine* 105, pp. 50–60.
- Guyton, A. C e J. E. Hall (2006). *Fisiologia Médica*. 11th. Philadelphia: Elsevier Saunders.
- Guzzetti, S., D. Iosa, M. Pecis, L. Bonura, M. Prosdocimi e A. Malliani (1990). "Effects of Sympathetic Activation on Heart-Rate-Variability in Chagas Patients". *Journal of the Autonomic Nervous System* 30.Suppl. S, S79–S82.
- Guzzetti, S., D. Iosa, M. Pecis, L. Bonura, M. Prosdocimi e A. Malliani (1991). "Impaired Heart-Rate-Variability in Patients with Chronic Chagas-Disease". *American Heart Journal* 121.6, Part 1, pp. 1727–1734.
- Guzzetti, S. et al. (2005). "Symbolic dynamics of heart rate variability: a probe to investigate cardiac autonomic modulation". *Circulation* 112, pp. 465–470.
- Haba-Rubio, J., V. Ibanez e E. Sforza (2004). "An alternative measure of sleep fragmentation in clinical practice: the sleep fragmentation index". *Sleep Medicine* 5, pp. 577–581.

- Harper, Ronald M, Vicki L Schechtman e Karen A Kluge (1987). "Machine classification of infant sleep state using cardiorespiratory measures". *Electroencephalography and clinical Neurophysiology* 67, pp. 379–387.
- Haykin, S. (1999). *Neural Networks: a comprehensive foundation*. New Jersey: Prentice Hall.
- Haykin, S. (2008). *Redes Neurais - Princípios e Práticas*. Ed. por McMaster University. 2nd. Bookman.
- Heikki, V., Huikuri, H. Timo, Mäkikallio e J. Perkiömäki (2003). "Measurement of heart rate variability by methods based on nonlinear dynamics". *Journal of Electrocardiology* 36.1, pp. 95–99.
- Himanen, S. L. e J. Hasan (2000). "Limitations of Rechtschaffen and Kales". *Sleep Medicine Reviews* 4.2, pp. 149–167.
- Horta, E. G. (2008). "Previsores para a Eficiência da Quimioterapia Neoadjuvante no Câncer de Mama". Diss. de mestrado. Belo Horizonte, MG: Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Huikuri, H. V., T. Makikallio, R. Airakinen Mitrani, A. Castellanos e R. J. Myerburg (1999). "Measurement of heart rate variability: A clinical tool or a research toy?" *J Am Coll Cardiol*, pp. 1878–1883.
- Iber, C., S. Ancoli-Israel, A. Chesson e S. F. Quan (2007a). *The AASM Manual for scoring of sleep and associated events: rules, terminology and technical specifications*. Rel. téc. American Academy of Sleep Medicine.
- Iber, C., S. Ancoli-Israel, A. Chesson e S.F. Quan (2007b). *The AASM Manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. 1st Ed. American Academy of Sleep Medicine.
- Iosa, D., V. Dequattro, D. D. P. Lee, U. Elkayam e H. Palmero (1989). "Plasma Norepinephrine in Chagas Cardioneuromyopathy - A Marker of Progressive Dysaunomia". *American Heart Journal* 117.4, pp. 882–887.
- Iosa, D. J., T. Caeiro e H. Palmero (1980). "Abnormal Hyperventilation Test in Chronic Chagas-Disease". *Journal of the Autonomic Nervous System* 2.1, pp. 87–92.
- Ivanov P. C., Amaral L. A. N. Goldberger A. L. Havlin S. Rosenblum M. G. e Z. H. E. Struzik (1999). "Multifractality in human heartbeat dynamics". *Nat.* 399, pp. 461–465.
- Janson N, Pavlov A Neiman A e V. Anishchenko (1998). "Reconstruction of dynamical and geometrical properties of chaotic attractors from threshold-crossing interspike intervals". *PHYSICAL REVIEW E* 58, pp. 4–7.
- Jasper, H.H. (1958). "The ten twenty electrode system of the International Federation". *Electroencephal Clin Neurophysiol* 10, pp. 371–375.
- Jesus, T. A. (2008). "Detecção de Determinismo e Modelagem Preditiva de Séries Temporais de Consumo de Energia Elétrica". Diss. de mestrado. Belo Horizonte, MG: Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.

- Jouvet, M. (1961). "Telencephalic and rhombencephalic sleep in the cat". In: *The Nature of Sleep Ciba Foundation Symposium*.
- Junqueira, L. F. Jr. (1998). "Disfunção Autonômica". In: *Doenças do Coração - Tratamento e Reabilitação*. Guanabara-Koogan.
- Kales, A., M. Okawa, T. Kobayashi, T. Hori e A. Miyasita (2001). "Proposed supplements and amendments to "A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects", the Rechtschaffen & Kales (1968) standard". *Psychiatry and Clinical Neurosciences* 55, pp. 305–310.
- Kandel, E. R., J.H. Schwarts e T. M. Jessel (2000). *Princípios da Neurociência*. São Paulo: Ed. Manole.
- Kaplan, A., J. Roschke, B. Darkhovsky e J. Fell (2001). "Macrostructural EEG characterization based on nonparametric change point segmentation: application to sleep analysis". *Journal of Neuroscience Methods* 106.1, pp. 81–90.
- Kaplan, D. T e R. J. Cohen (1990). "Is Fibrillation Chaos?" *Circulation* 67, pp. 886–892.
- Karlen, W., C. Mattiussi e D. Floreano (2009). "Sleep and Wake Classification With ECG and Respiratory Effort Signals". *IEEE Transactions on Biomedical Circuits and Systems* 3.2, pp. 71–78.
- Kay, S. M. (1987). *Modern Spectral Estimation: theory and application*. Prentice Hall.
- Kleiger, R. E., P. K. Stein e J. T. Jr. Bigger (2005). "Heart Rate Variability: Measurement and Clinical Utility". *A.N.E.* 10.1, pp. 88–101.
- Kolmogorov, A. N. (1958). "A new metric invariant of transient dynamical systems and automorphisms in lebesgue spaces." *Dokl Akad Nauk SSSR*.
- Kortelainen, Juha M, Martin O Mendez, Anna Maria Bianchi, Matteo Matteucci e Sergio Cerutti (2010). "Sleep Staging Based on Signals Acquired Through Bed Sensor". *IEEE Transactions on Information Technology in Biomedicine* 14.3, pp. 776–785.
- Kurths J, Voss A Saporin P Witt A Kleiner HJ e N. Wessel (1995). "Quantitative analysis of heart rate variability". *Chaos* 5.1, pp. 88–94.
- Lainscsek, C., P. Rowat, L. Schettino, D. Lee, D. Song, C. Letellier e H. Poizner (2012). "Finger tapping movements of Parkinsons disease patients automatically rated using nonlinear delay differential equations". *Chaos* 22. 0131119.
- Letellier, C. (2000). *Systèmes Dynamiques Non Linéaires: vers une description des structures du chaos*. Rouen: Université de Rouen, Notes de cours de l'école doctorale version du 10 mars.
- Letellier, C. (2008). "Symbolic sequence analysis using approximated partition". *Chaos, Solitons and Fractals* 36.1, pp. 88–94.
- Lind, D. e B. Marcus (1995). *An introduction to symbolic dynamics and coding*. 2nd. NY, USA: Cambridge University Press.
- Liu, Bo, Zhifeng Hao e Eric C C Tsang (2008). "Nesting one-against-one algorithm based on SVMs for pattern classification." *IEEE Transactions on Neural Networks* 19.12, pp. 2044–52.

- Lombardi, F. (2000). "Chaos Theory, heart rate variability, and arrhythmic mortality". *Circulation* 101, pp. 8–10.
- Lombardi, F., T. H. Makikallio, R. J. Myerburg e H. V. Huikuri (2001). "Sudden cardiac death: role of heart rate variability to identify patients at risk". *Cardiovascular Research* 50, pp. 210–217.
- Loomis, A.L., E.N. Harvey e G.A. Hobart (1937). "Cerebral states during sleep, a studied by human brain potentials". *Journal of Experimental Psychology* 21, pp. 127–144.
- Lugaresi, E. et al. (1972). "Some periodic phenomena arising during drowsiness and sleep in man". *Electroenceph. Clin. Neurophysiol.* 32, pp. 701–705.
- Magalhães, M. N e A. C. P. Lima (2002). *Noções de Probabilidade e Estatística*. USP, São Paulo: USP.
- Malinowska, U., H. Klekowicz, A. Wakarow, S. Niemcewicz e P. J. Durka (2009). "Fully parametric sleep staging compatible with the classical criteria". *Neuroinformatics* 7.4, pp. 245–53.
- Marcos, J Víctor, Roberto Hornero, Daniel Alvarez, Félix del Campo e Carlos Zamarrón (2009). "Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry." *Medical Engineering & Physics* 31.8, pp. 971–8.
- Marin-Neto, J. A., L. Gallo, J. C. Jr. Manco, A. Rassi e D. S. Amorim (1980). "Mechanisms of tachycardia on standing: studies in normal individuals and in chronic Chagas' heart patients". *Cardiovasc. Res.*
- Marin-Neto, J. A., B. C. Maciel, L. Gallo Junior, L.F. Junqueira Junior e D. S. Amorim (1986). "Effect of parasympathetic impairment on the haemodynamic response to handgrip in Chagas's heart disease". *Br Heart J.*
- McClintock, P. V. E. e A. Stefanovska (2002). "Noise and determinism in cardiovascular dynamics". *Physica A* 314, pp. 69–76.
- McLoughlin, I. V. (2008). "Line spectral pairs". *Signal Processing* 88 (2008) 448–467, pp. 448–467.
- Mendez, M. O., A. M. Bianchi, M. Matteucci, S. Cerutti e T. Penzel (2009). "Sleep apnea screening by autoregressive models from a single ECG lead". *IEEE transactions on bio-medical engineering* 56.12, pp. 2838–50.
- Mendez, Martin O, Matteo Migliorini, Juha M Kortelainen, Domenino Nistico, Edgar Arce-Santana, Sergio Cerutti e Anna M Bianchi (2010). "Evaluation of the sleep quality based on bed sensor signals: Time-variant analysis." In: *32nd Annual International Conference of the IEEE EMBS*, pp. 3994–7.
- Mendez, Martin Oswaldo e Politecnico Milano (2010). "Sleep staging from Heart Rate Variability : time-varying spectral features and Hidden Markov Models Matteo Matteucci Vincenza Castronovo and Luigi Ferini-Strambi Sergio Cerutti and Anna Maria Bianchi". *Biomedical Engineering* 3, pp. 246–263.

- Merica, H. e R. D. Fortune (2005). "Spectral Power Time-courses of Human Sleep EEG Reveal a Striking Discontinuity at 18 Hz Marking the Division between NREM-specific and Wake/REM-specific Fast Frequency Activity". *Cerebral Cortex* 15.7, pp. 877–884.
- Monteiro, L. H. A. (2006). *Sistemas Dinâmicos*. 2nd. São Paulo: Ed. Livraria da Física.
- Montgomery, D. C. e G. C. Runger (2002). *Applied statistics and probability for engineers*. John Wiley & Sons, Inc.
- Moraes, R S, E L Ferlin, C a Polanczyk, L E Rohde, L Zaslavski, J L Gross e J P Ribeiro (2000). "Three-dimensional return map: a new tool for quantification of heart rate variability". *Autonomic neuroscience: basic & clinical* 83.12, pp. 90–9.
- Morrel, M. J., L. Finn, H. Kim, P. E. Peppard, M. S. Badr e T. Young (2000). "Sleep Fragmentation, Awake Blood Pressure, and Sleep-Disordered Breathing in a Population-based Study". *American Journal of Respiratory and Critical Care Medicine* 162, pp. 2091–2096.
- Morse, M. e G. A. Hedlund (1938). "Symbolic Dynamics". *American Journal of Mathematics* 60.4, pp. 815–866.
- Moser, D. et al. (2009). "Sleep Classification According to AASM and Rechtschaffen and Kales: Effects on Sleep Scoring Parameters". *Sleep* 32.2, pp. 139–149.
- Mäkikallio T. H., Hoiber S. Kober L. Torp-Pedersen C. Peng C. K. Goldberger A. L. e H. V. Huikuri (1999). "Fractal analysis of heart rate dynamics as a predictor of mortality in patients with depressed left ventricular function after acute myocardial infarction." *Am. J. Cardiol* 83, pp. 836–839.
- Naeck, R. (2011). "Evaluation de l'adaptation à la ventilation non invasive chez patients atteints d'insuffisance respiratoire chronique." Tese de doutorado. Université de Rouen, France.
- Nicolat, A. et al. (2001). "The restorative effect of naps on perceptual deterioration". *Nat. Neuroscience* 5, pp. 677–681.
- Nollo, G., A. Porta, L. Faes, M. Greco, M. Disertori e F. Ravelli (2001). "Causal linear parametric model for baroreflex gain assessment in patients with recent myocardial infarction". *Am J Physiol Heart Circ Physiol* 280, pp. 1830–1839.
- Nollo, G., A. Porta, L. Faes, M. Greco, M. Disertori e F. Ravelli (2005). "Exploring directionality in spontaneous heart period and systolic pressure variability interactions in humans. Implications in baroreflex gain evaluation". *Am J Physiol Heart Circ Physiol* 288, pp. 1777–1785.
- Novak V., Novak P. de Champlain J. Le Blanc A. R. Martin R. e R Nadeau (1993). "Influence of respiration on heart rate and blood pressure fluctuations". *J. Appl. Physiol.* 74, pp. 617–6.
- Novelli, L., R. Ferri e O. Bruni (2010). "Sleep classification according to AASM and Rechtschaffen and Kales: effects on sleep scoring parameters of children and adolescents". *J. Sleep Res.* (2010) 19, 238–247.

- Olbrich, E. e P. Achermann (2005). "Analysis of oscillatory patterns in the human sleep EEG using a novel detection algorithm". *Journal of Sleep Research* 14.4, pp. 337–46.
- Olbrich, E., P. Achermann e P.F. Meier (2003). "Dynamics of human sleep EEG". *Neurocomputing* 52-54, pp. 857–862.
- Oliveira, D. J. (2002). "Síntese e Análise da Variabilidade da Frequência Cardíaca baseada em Modelos Identificados". Diss. de mestrado. Belo Horizonte, MG: Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Oliveira, D. J. (2007). "Síntese e análise de sinais de Variabilidade da Frequência Cardíaca pela modelagem NARMAX com entrada sintética". Tese de doutorado. Belo Horizonte, MG: Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Oliveira, D. J., M. E. D. Gomes, H. N. Guimarães e L. A. Aguirre (2002). "Síntese de sinais de variabilidade da frequência cardíaca baseada em modelos identificados". *Revista Brasileira de Engenharia Biomédica* 22, pp. 51–62.
- Oliveira, D. J., C. Letellier, M. E. D. Gomes e L. A. Aguirre (2008). "The use of synthetic input sequences in time series modeling". *Physics Letters A* 372, pp. 5276–5282.
- Oliveira D. J., Gomes M. E. D. Guimarães H. N. e L. A. Aguirre (2006). "Síntese de sinais de variabilidade da frequência cardíaca baseada em modelagem não-linear". *Revista Brasileira de Engenharia Biomédica* 22, pp. 51–62.
- Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Process*. 3th. Singapore: McGraw-Hill International Editions.
- Paquereau, J. (2007). "Nouvelles règles de codage visuel chez l'adulte". *Médecine du sommeil* 4.
- Pardey, J., S Roberts e L Tarassenko (1996). "Review A review of parametric EEG analysis modelling techniques for EEG Analysis". *Science C.1*, pp. 2–11.
- Pardey, J., S. Roberts, L. Tarassenko e J. Stradling (1996). "A new approach to the analysis of the human sleep/wakefulness continuum". *Journal of Sleep Research* 5.4, pp. 201–10.
- Parisotto, V. D., E. M. Lima, J. M. P Silva, M. R. Sousa e A. L. P. Ribeiro (2008). "Cardiac dysautonomia in chronic kidney disease". *Journal of Nuclear Cardiology* 15, pp. 246–54.
- Parlitz, U., S. Berg, S. Luther, A. Schirdewan, J. Kurths e N. Wessel (2012). "Classifying cardiac biosignals using ordinal pattern statistics". *Computers in Biology and Medicine* 42, pp. 319–327.
- Peng C. K., Havlin S. Stanley H. E. e A. L. Goldberger (1995). "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series." *Chaos* 5, pp. 82–87.
- Penzel, T., N. Wessel, M. Riedl, J. W. Kantelhardt, S. Rostig, m. Glos, A. Suhrbier, H. Malberg e I. Fietze (2007). "Cardiovascular and respiratory dynamics during normal and pathological sleep". *Chaos* 17.015116, pp. 1–10.

- Pépin, J.-L., N. Delavie, I. Pin, C. Deschaux, J. Argod, M. Bost e P. Lévy (2005). "Pulse transit time improves detection of sleep respiratory events and microarousals in children". *Chest* 127, pp. 722–730.
- Perc, M. (2005). "Nonlinear time series analysis of the human electrocardiogram". *European Journal of Physics* 26, pp. 757–768.
- Pincus, S. M. (1991). "Approximate entropy as a measure of system complexity". *Proc. Natl Acad. Sci. USA* 88, pp. 2297–2301.
- Porta, A., L. Faes, M. Masé, G. D'Addio, G. D. Pinna, R. Maestri, N. Montano, R. Furlan, S. Guzzetti, G. Nollo e A. Malliani (2007). "An integrated approach based on uniform quantization for the evaluation of complexity of short-term heart period variability: Application to 24 h Holter recordings in healthy and heart failure humans". *Chaos* 17, pp. 1–11.
- Porta A., Guzzetti S. Montano N. Furlan R. Pagani M. Malliani A. e S. Cerutti (2001). "Entropy, entropy rate and pattern classification as tools to typify complexity in short heart period variability series". *IEEE Transactions on Biomedical Engineering* 48, pp. 1282–1291.
- Portier, Florence, Adriana Portmann, Pierre Czernichow, Lionel Vascaut, Etienne Devin, Daniel Benhamou Antoine Cuvelier e Jean-François Muir (2000). "Evaluation of Home versus Laboratory Polysomnography in the Diagnosis of Sleep Apnea Syndrome". *Am. J. Respir. Crit. Care Med.* 162.3, pp. 814–818.
- Rabarimanantsoa, H., L. Achour, C. Letellier, A. Cuvelier e J.-F. Muir (2007). "Recurrence plots and Shannon entropy for a dynamical analysis of asynchronisms in mechanical non-invasive ventilation". *Chaos* 17. 013115.
- Rabarimanantsoa-Jamous, H. (2008). "Qualité des interactions patient-ventilateur en ventilation non invasive nocturne". Tese de doutorado. Université de Rouen, Faculté des Sciences de l'Université de Rouen.
- Rechtschaffen, A. e A. Kales (1968). *A manual of standardized terminology techniques and scoring systems for sleep stages of human subjects*. Brain Information Service/Brain Research Institute. Los Angeles, UCLA.
- Redmond, Stephen J e Conor Heneghan (2006). "Cardiorespiratory-based sleep staging in subjects with obstructive sleep apnea." *IEEE Transactions on Biomedical Engineering* 53.3, pp. 485–96.
- Rheinberger, M. B. e H. H Jasper (1937). "Electrical activity of the cerebral cortex in the unanesthetized cat". *American Journal of Physiology* 119, pp. 186–196.
- Ribeiro, A. L., A. M. Reis, M. V. L. Barros, M. R. Sousa, A. L. L. Rocha, A. A. Perez, J. P. Braga, F. S. Machado e M. O. C. Rocha (2002a). "Brain natriuretic peptide in the diagnosis of systolic left ventricular dysfunction in Chagas disease." *Lancet*.
- Ribeiro, A. L., M. P. Nunes, M. M. Teixeira e M. O. C. Rocha (2012). "Diagnosis and management of Chagas disease and cardiomyopathy". *Nat. Rev. Cardiol.* Pp. 1–14.

- Ribeiro, A. L. P. (1996). "Disfunção autonômica e arritmia ventricular em chagásicos sem cardiopatia aparente". Tese de doutorado. Belo Horizonte, MG: Programa de Pós-Graduação em Medicina Tropical, Universidade Federal de Minas Gerais.
- Ribeiro, A. L. P., R. S. Moraes, J. P. Ribeiro, E. L. Ferlin, R. M. Torres, E. Oliveira e M. O. C. Rocha (2001). "Parasympathetic dysautonomia precedes left ventricular systolic dysfunction in Chagas disease". *American Heart Journal* 141, pp. 260–5.
- Ribeiro, A. L. P., F. Lombardi, M. R. Sousa, M. V. R. Barros, A. Porta, V. C. V. Barros, M. E. D. Gomes, F. S. Machado e M. O. C. Rocha (2002b). "Power-Law Behavior of Heart Rate Variability in Chagas' Disease". *The American Journal of Cardiology* 89, pp. 414–418.
- Ribeiro, A. L. P., G. Schmidt, M. R. Sousa, F. Lombardi, M. E. D. Gomes, A. A. Perez, M. V. L. Barros, F. S. Machado e M. O. C. Rocha (2003). "Heart Rate Turbulence in Chagas Disease". *PACE* 26, pp. 406–410.
- Ribeiro, A. L. P., F. Lombardi, M. R. Sousa e M. O. C. Rocha (2005). "Vagal dysfunction in Chagas disease". *International Journal of Cardiology* 103, pp. 225–226.
- Ribeiro, A. L. P., P. S. Cavalcanti, F. Lombardi, M. C. P. Nunes, M. V. L. Barros e M. O. C. Rocha (2008). "Prognostic Value of Signal-Averaged Electrocardiogram in Chagas Disease". *J Cardiovasc Electrophysiol* 19, pp. 502–509.
- Richman, J. S. e J. R. Moorman (2000). "Physiological time-series analysis using approximate entropy and sample entropy". *Am. J. Physiol. Heart Circ. Physiol.* 278, H2039–H2049.
- Rimbaud, L., P. Passouant e J. Cadilhac (1955). "Participation de l'hippocampe: la régulation des états de veille et de sommeil". *Revue Neurologique* 93, pp. 303–308.
- Roberts, S. e L. Tarassenko (1992a). "Analysis of the sleep EEG using a multilayer network with spatial organisation". *IEE Proceedings F Radar and Signal Processing* 139.6, pp. 420–425.
- Roberts, S. e L. Tarassenko (1992b). "New method of automated sleep quantification". *Medical & Biological Engineering & Computing* 30, pp. 509–517.
- Rocha, A. L. L., A. L. Ribeiro, M. V. L. Barros, Machado F. S., Sousa M. R. e M. O. C. Rocha (2001). "Associação da insuficiência cronotrópica e indicadores de morbidade na cardiopatia chagásica crônica". *Arquivos Brasileiros de Cardiologia*.
- Rocha, C. S. C., A. P. Lemos e L. A. Aguirre (2010). "Dinâmica Simbólica Aplicada à Análise e Classificação da Disfunção Autonômica da Variabilidade da Frequência Cardíaca na Doença de Chagas". *XXII CBEB 2010*, pp. 29–32.
- Rodrigues, G. G., L. A. Aguirre, U. S. Freitas e C. Letellier (2010). "Análise das Interações Paciente-Ventilador Durante Ventilação Mecânica não Invasiva Noturna Utilizando Gráficos de Recorrências Conjuntas". In: *XXII CBEB 2010*, pp. 75–78.
- Sady, C. C. R. e L. A. Aguirre (2012). "Investigação de Características Baseadas em Dinâmica Simbólica e Padrões Estatísticos Ordiniais na Classificação da Variabilidade da Frequência Cardíaca". *XXIII CBEB 2012*.

- Sady, C. C. R., U. S. Freitas, A. Portmann, J. F. Muir, C. Letellier e L. A. Aguirre (2013). "Automatic Sleep Staging From Ventilator Signals in Non Invasive Ventilation". *Comput. in Biol. Med.* <http://dx.doi.org/10.1016/j.compbiomed.2013.04.011>.
- Schreiber, T. (1999). "Interdisciplinary application of nonlinear time series methods". *Phys. Rep.* 308, pp. 1–64.
- Schroeder, E. B., D. Liao, L. E. Chambless, R. J. Prineas, G. W. Evans e G. Heiss (2003). "Hypertension, Blood Pressure, and Heart Rate Variability - The Atherosclerosis Risk in Communities (ARIC) Study". *Hypertension* 42, pp. 1106–1111.
- Shannon C, E. (1948). "A mathematical theory of communication". *The Bell System Technical Journal* 27, pp. 379–423, 623–656.
- Shneerson, J. M. (2005). *Sleep Medicine: A Guide to Sleep and its Disorders*. 2nd. Massachusetts, USA: Blackwell Publishing.
- Silber, M.H. et al. (2007). "The visual scoring of sleep in adults". *Journal of Clinical Sleep Medicine* 3.2, pp. 121–131.
- Simões, M V, Pintya, G Bromberg-Marin, Sarabanda, C M Antloga, Pazin-Filho, B C Maciel e J Marin-Neto (2000). "Relation of regional sympathetic denervation and myocardial perfusion disturbance to wall motion impairment in Chagas' cardiomyopathy". *The American journal of cardiology* 86.9, pp. 975–81.
- Sinha, R. K. (2008). "Artificial Neural Network and Wavelet Based Automated Detection of Sleep Spindles, REM Sleep and Wake States". *Journal of Medical Systems* 32.4, pp. 291–299.
- Sitnikova, E., A. E. Hramov, A. Koronovsky e G. van Luijelaar (2009). "Sleep spindles and spike-wave discharges in EEG: Their generic features, similarities and distinctions disclosed with Fourier transform and continuous wavelet analysis". *Journal of Neuroscience Methods* 180.2, pp. 304–16.
- Smith, R. P., J. Argod, J.-L. Pépin e P. A. Lévy (1999). "Pulse transit time: an appraisal of potential clinical applications". *Thorax* 54, pp. 452–458.
- Smurra, M. V., M. Dury, G. Aubert, D. O. Rodestein e Liistro G. (2001). "Sleep fragmentation: comparison of two definitions of short arousals during sleep in OSAS patients". *Eur Resp J* 17, pp. 723–727.
- Sousa M., R., H. V. Huikuri, F. Lombardi, A. A. Perez, M. E. D Gomes, M. V. L. Barros, V. C. V. Barros, M. O. C. Rocha e A. L. P. Ribeiro (2006). "Abnormalities in Fractal Heart Rate Dynamics in Chagas Disease". *A.N.E.* 11.2, pp. 145–153.
- Souza, A. V. P. (2001). "Modelagem e análise não-lineares de sinais respiratórios e cardiovasculares". Diss. de mestrado. Belo Horizonte, MG: Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais.
- Souza-Neto, J., J. L. A. Carvalho, F. O. Nascimento, A. F. Rocha e L. F. Jr. Junqueira (2004). "Classificação de Pacientes Chagásicos com Escalogramas e Redes Neurais Auto-Organizativas". In: *VIII Simpósio Brasileiro de Redes Neurais - SBRN*. 1: 015101. São Luiz: Anais do Simpósio Brasileiro de Redes Neurais, pp. 1–6.

- Stepanski, E. J. (2007). "Improving the Utility of Interpreting Sleep Fragmentation". *Journal of Clinical Sleep Medicine* 3.3, pp. 275–276.
- Stradling, J. R., D. J. Pitson, L. Bennett, C. Barbour e R. J. O. Davies (1999). "Variation in the Arousal Pattern after Obstructive Events in Obstructive Sleep Apnea". *Am J Respir Crit Care Med* 159, pp. 130–136.
- Susmakova, K. e A. Krakovska (2007). "Classification of Waking, Sleep Onset and Deep Sleep by Single Measures". *Measurement Science Review* 7, pp. 34–38.
- Takens, F. (1981). "Detecting strange attractors in turbulence". In: *Rand, D. a. and Young, L. S., editors, Dynamical systems and turbulence - Lecture Notes in Mathematics*. Springer-Verlag, Berlin, pp. 366–381.
- Task-Force (1996). "Heart rate variability - standards of measurement, physiological interpretation and clinical use". 93, pp. 1043–106.
- Terrill, Philip Ian, Stephen James Wilson, Sadasivam Suresh, David M Cooper e Carolyn Dakin (2010). "Attractor structure discriminates sleep states: recurrence plot analysis applied to infant breathing patterns." *IEEE Transactions on Biomedical Engineering* 57.5, pp. 1108–16.
- Terzano, M.G. e L. Parrino (2000). "Origin and significance of the Cyclic Alternating Pattern". *Sleep Med. Rev.* 4.1, pp. 101–123.
- Terzano, M.G. et al. (1982). "Is the EEG cyclic alternating pattern a true autonomous entity?" *Eur. Neurol.* 21, pp. 324–334.
- Terzano, M.G. et al. (1985). "The cyclic alternating pattern as a physiologic component of normal NREM sleep". *Sleep* 8, pp. 137–145.
- Theiler J., Eubank S. Longtin A. Galdrikian B. e Farmer J. D. (1992). "Testing for nonlinearity in time series: the method of surrogate data". *Physica D* 58, pp. 77–94.
- Tufillaro, Nicholas B. (1999). "Symbolic Dynamics in Mathematics, Physics, and Engineering". *Hewlett-Packard Research Labs at the IMA Industrial Problems Seminar*.
- Turk, O. e L. M. Arslan (2006). "Robust processing techniques for voice conversion". *Computer Speech and Language* 20, pp. 441–467.
- Uchida, S., Y. Atsumi e T. Kojima (1994). "Dynamic relationships between sleep spindles and delta waves during a NREM period". *Brain Research Bulletin* 33.3, pp. 351–355.
- Vapnik, V N (1999). "An overview of statistical learning theory." *IEEE Transactions on Neural Networks* 10.5, pp. 988–99.
- Várady, Peter, Tamás Micsik, Sándor Benedek e Zoltán Benyó (2002). "A novel method for the detection of apnea and hypopnea events in respiration signals." *IEEE Transactions on Biomedical Engineering* 49.9, pp. 936–42.
- Ventouras, E. M., E. Monoyiou, P. Y. Ktonas, T. Paparrigopoulos, D. G. Dikeos, N. K. Uzunoglu e C. R. Soldatos (2005). "Sleep spindle detection using artificial neural networks trained with filtered time-domain EEG: a feasibility study". *Computer Methods and Programs in Biomedicine* 78.3, pp. 191–207.

- Vermeiren, C. (1996). "Analyse et modélisation du système cardio-vasculaire et sa régulation à court terme par le système nerveux autonome". Tese de doutorado. Université de Paris Val de Marne.
- Voss A., Hnatkova K. Wessel N. Kurths J. Sander A. Schirdewan A. Camm A. J. e M. Malik (1998). "Multiparametric analysis of heart rate variability used for risk stratification among survivors of acute myocardial infarction". *Pacing Clin. Electrophysiol* 21, pp. 186–192.
- Voss A., Schulz S. Schroeder R. Baumert M. e P. Caminal (2008). "Methods derived from nonlinear dynamics for analysing heart rate variability". *Phil. Trans. R. Soc. A* 1, pp. 1–22.
- Vural, C. e M. Yildiz (2010). "Determination of Sleep Stage Separation Ability of Features Extracted from EEG Signals Using Principle Component Analysis". *Journal of Medical Systems* 34.1, pp. 83–89.
- Watanabe, T e K Watanabe (2004). "Noncontact method for sleep stages estimation". *IEEE Trans. Biomed. Eng.* 51.10, pp. 1735–48.
- Wesensen, N. J., T. J. Balkin e G. Belenky (1999). "Does sleep fragmentation impact recuperation? A review and reanalysis". *J. Sleep Res* 8, pp. 237–245.
- Wessel, N., H. Malberg, R. Bauernschmitt, A. Schirdewan e J. Kurths (2006). "Nonlinear additive autoregressive model-based analysis of short-term heart rate variability". *Med Biol Eng Comput* 44, pp. 321–330.
- Wessel, N., H. Malberg, R. Bauernschmitt e J. Kurths (2007b). "Nonlinear methods of cardiovascular physics and their clinical applicability". *International Journal of Bifurcation and Chaos* 17, pp. 3325–3371.
- Wessel, N., H. Malberg, R. Bauernschmitt e J. Kurths (2007c). "Nonlinear methods of cardiovascular physics and their clinical applicability". *International Journal of Bifurcation and Chaos* 17.10, pp. 3325–3371.
- Wikipedia (2013a). *21 electrodes of International 10-20 system for EEG.svg*. Acessado em março de 2013. URL: http://commons.wikimedia.org/wiki/File:21_electrodes_of_International_10-20_system_for_EEG.svg.
- Wikipedia (2013b). *Polysomnografie Body en 106.svg*. Acessado em março de 2013. URL: http://commons.wikimedia.org/wiki/File:Polysomnografie_Body_en_106.svg#filelinks.
- Zoubek, L., S. Charbonnier, S. Lesecq, A. Buguet e F. Chapotot (2007). "Feature selection for sleep/wake stages classification using data driven methods". *Biomedical Signal Processing and Control* 2.3, pp. 171–179.

Hipnogramas Estimados

A.1 Estimação de Hipnogramas para noites sem Polissonografia via Modelos

São mostrados, a seguir, os hipnogramas estimados para as noites N_3 e N_4 utilizando os modelos de classificadores obtidos para duas noites do mesmo paciente – caso dependente do paciente DP(2).

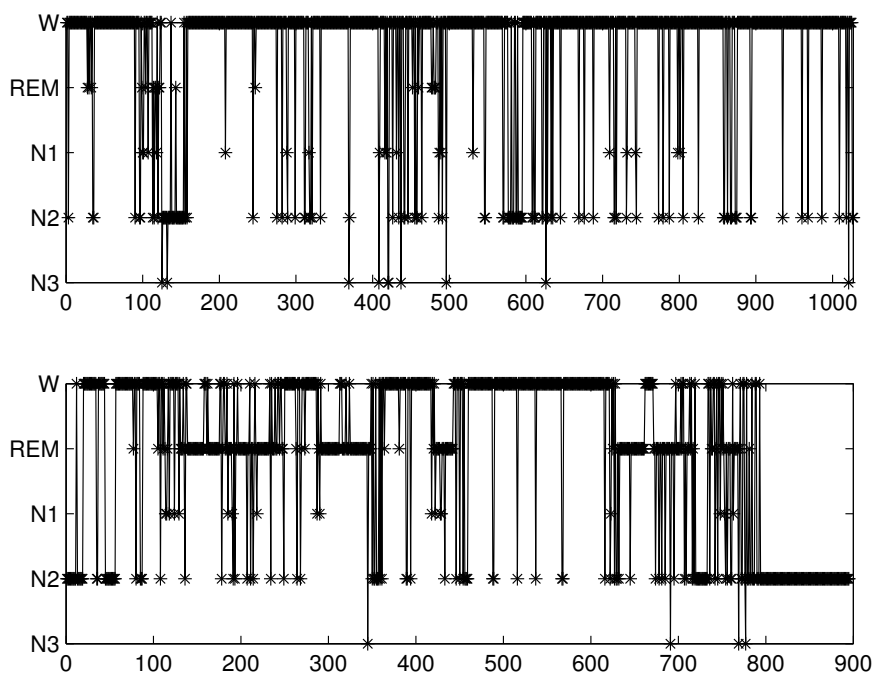
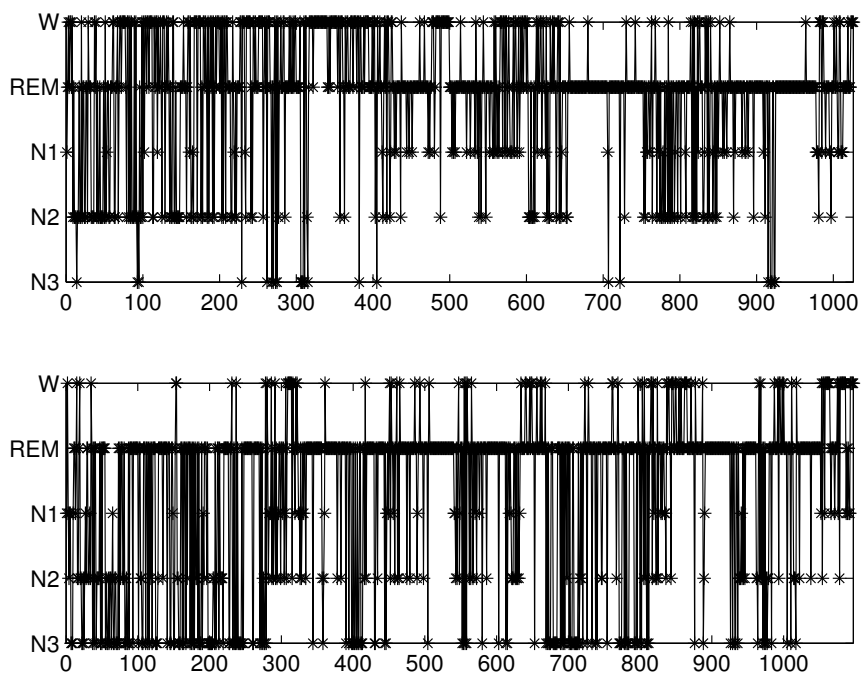
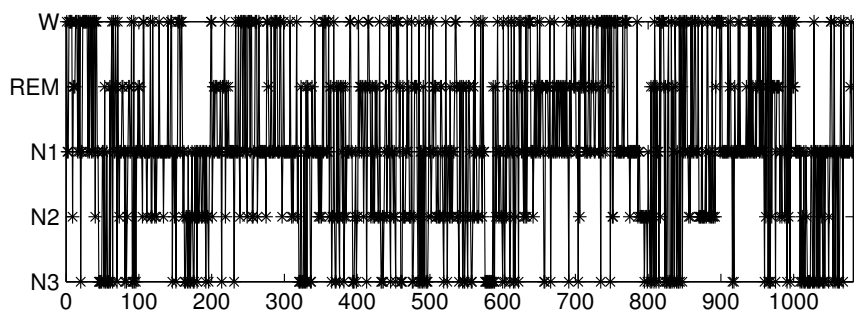
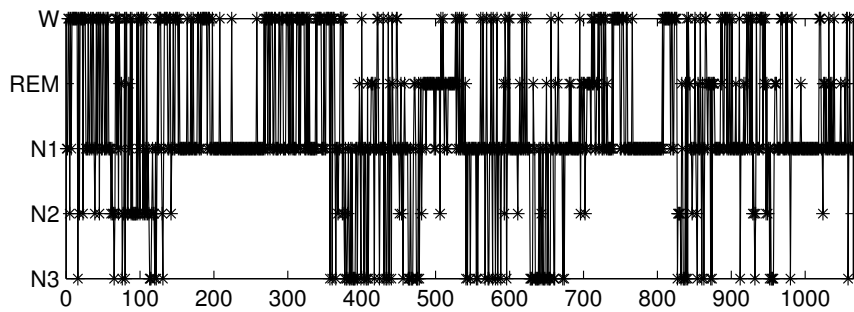
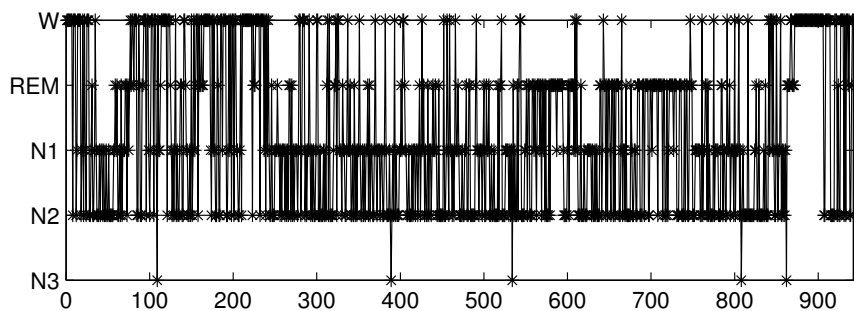
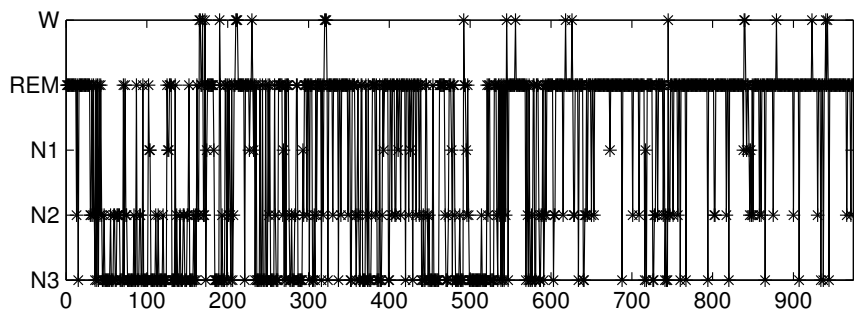
(a) Hipnogramas estimados para o paciente P_1 (b) Hipnogramas estimados para o paciente P_8

Figura A.1: Hipnogramas estimados para o paciente (a) P_1 : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior); (b) P_8 : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior).



(a) Hipnogramas estimados para o paciente P_9



(b) Hipnogramas estimados para o paciente P_{10}

Figura A.2: Hipnogramas estimados para o paciente (a) P_9 : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior); (b) P_{10} : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior).

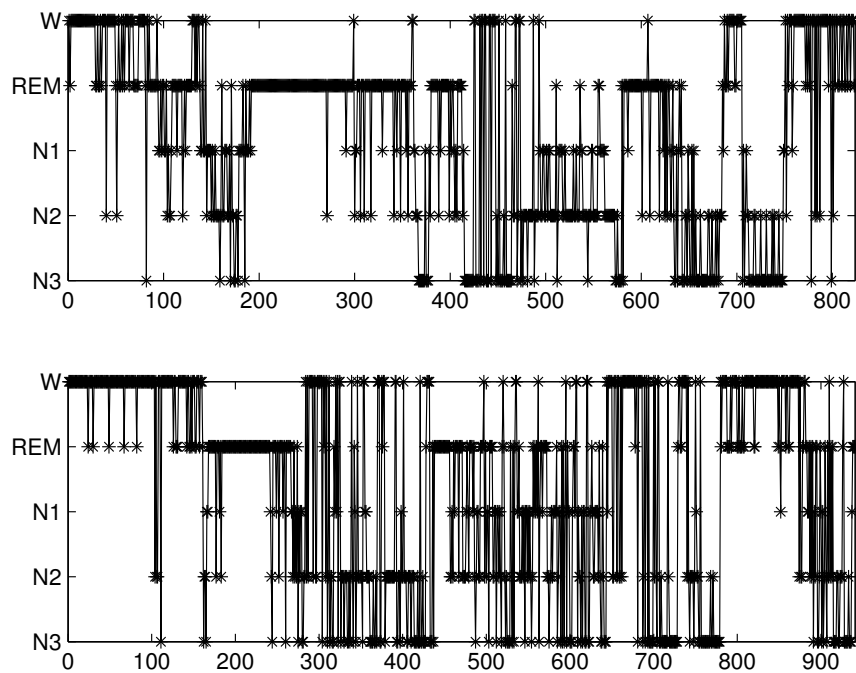
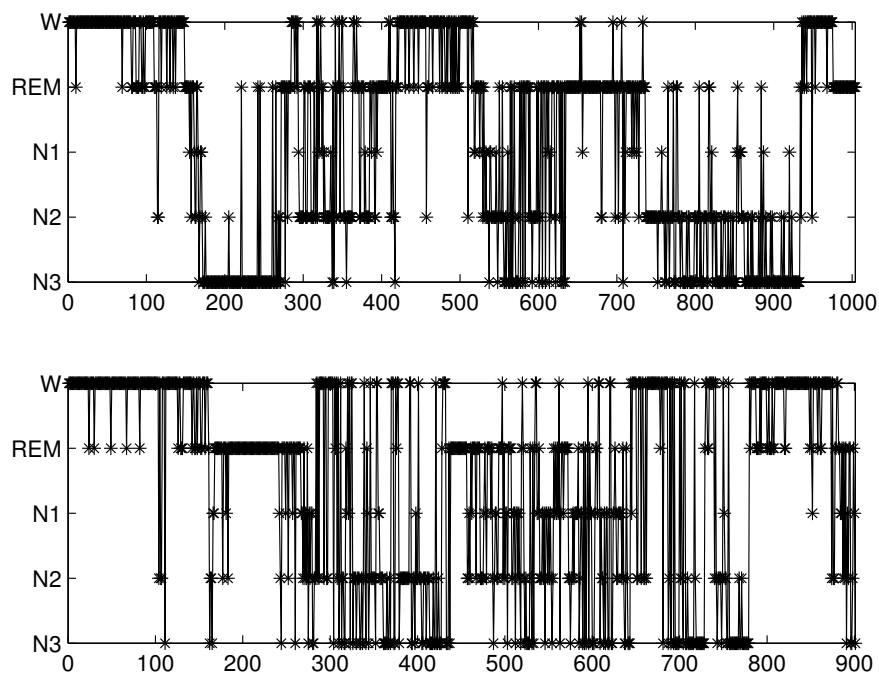
(a) Hipnogramas estimados para o paciente P_{11} (b) Hipnogramas estimados para o paciente P_{12}

Figura A.3: Hipnogramas estimados para o paciente (a) P_{11} : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior); (b) P_{12} : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior).

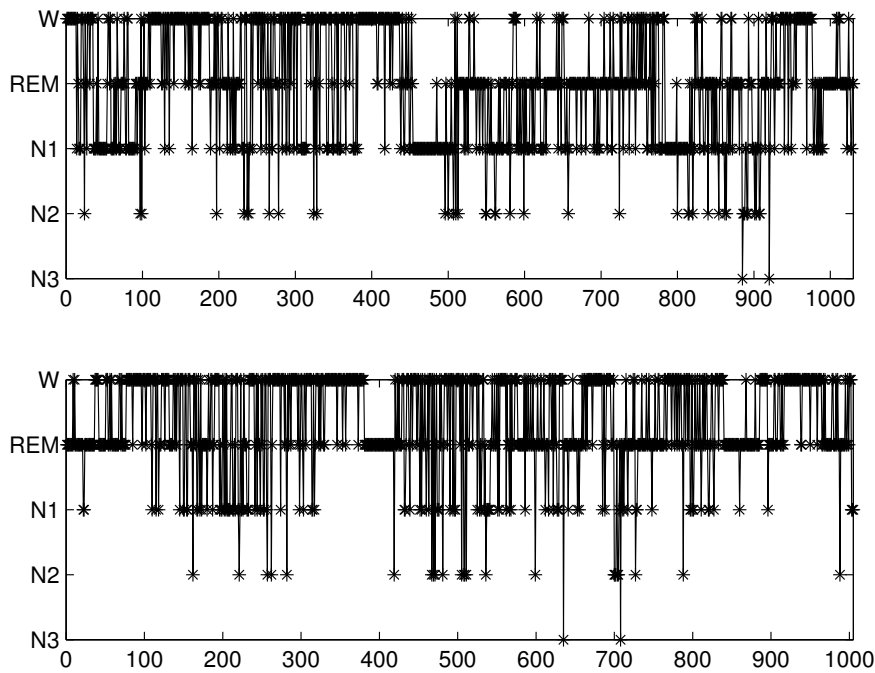
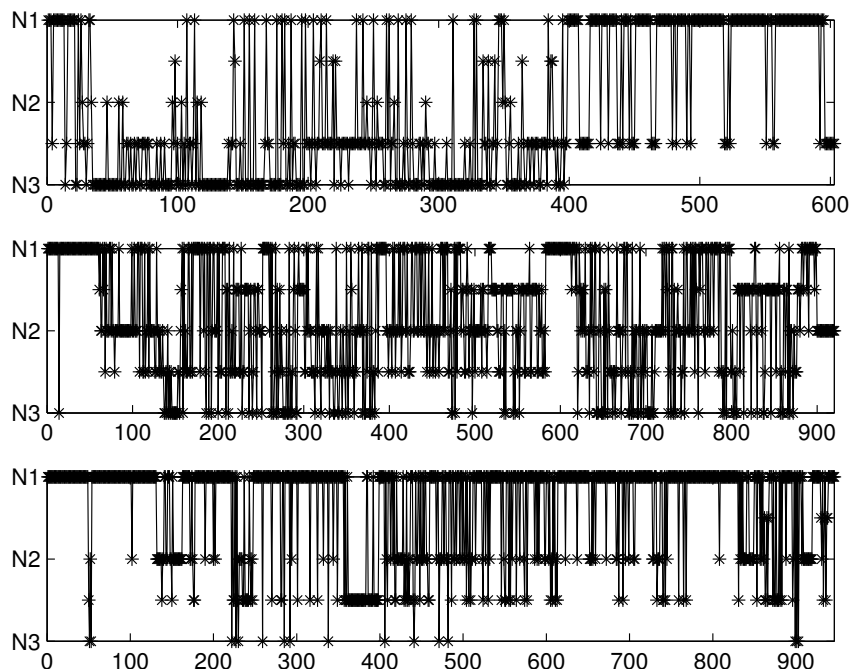
(a) Hipnogramas estimados para o paciente P_{13} (b) Hipnogramas estimados para os pacientes P_2 , P_3 e P_{14}

Figura A.4: Hipnogramas estimados para o paciente (a) P_{13} : noite N_3 (gráfico superior) e noite N_4 (gráfico inferior); (b) P_2 : noite N_3 (gráfico superior), P_3 : noite N_3 (gráfico do meio) e P_{14} : noite N_4 (gráfico inferior).

Testes de Hipóteses

B.1 Análise de Variância (ANOVA)

A análise de variância, comumente referida como ANOVA simples (do inglês, *one-way ANOVA*), é uma das técnicas estatísticas mais amplamente usada. A ANOVA contraria o seu nome no sentido de que a análise não é feita por meio da variância dos dados, mas sim por variações nas médias.

Considere um modelo estatístico, em que cada observação Y pode ser decomposta em duas componentes: *sistemática* e *aleatória*, esta última representando variações individuais e todos os fatores que não são explicados pela parte sistemática. Matematicamente, pode-se escrever (Casella e Berger 1990; Magalhães e Lima 2002):

$$Y = \mu + e. \tag{B.1}$$

Assim, se Y representa a observação associada a uma unidade experimental, a parte sistemática μ pode ser vista como a média populacional, que é fixa, e a parte aleatória e como a informação referente a outros fatores que podem influir nas observações, mas não são incorporados em μ .

Suponha que se queira comparar as médias de K populações, isto é, queira-se testar:

- $H_0 : \mu_1 = \mu_2 = \dots = \mu_K$;
- H_1 : pelo menos uma das médias μ_i é diferente das demais.

Para tanto, obtêm-se K amostras independentes, com m indivíduos em cada uma

delas. Nesta situação, o modelo estatístico para a j -ésima unidade experimental, da i -ésima população é dado por:

$$Y_{ij} : \mu_i + e_{ij}, \quad i = 1, \dots, K; \quad j = 1, \dots, m \quad (\text{B.2})$$

Caso a hipótese H_0 seja verdadeira, tem-se que todas as médias para as K populações serão iguais a um valor comum μ . O modelo pode, então, ser escrito como:

$$Y_{ij} : \mu + e_{ij}^*, \quad i = 1, \dots, K; \quad j = 1, \dots, m \quad (\text{B.3})$$

O Teste ANOVA faz as seguintes suposições (Casella e Berger 1990; Magalhães e Lima 2002):

- Y_{ij} são variáveis aleatórias independentes
- todas as K populações têm variâncias iguais a σ^2 ;
- $Y_{ij} \approx N(\mu_i, \sigma^2)$, $i = 1, \dots, K$ e $j = 1, \dots, m$.

Assim foram usados testes estatísticos a fim de verificar se a distribuição das entropias de cada grupo era gaussiana (H_0).

Para verificar se a hipótese nula H_0 pode ou não ser rejeitada, calcula-se a chamada significância, a qual é dada pela expressão:

$$S = \frac{|Q_D - \mu_H|}{\sigma_H}, \quad (\text{B.4})$$

em que Q_D é o valor da estatística discriminante para a entropia do grupo controle, μ_H e σ_H são a média e o desvio padrão, respectivamente, dos valores da estatística discriminante para as entropias dos grupos 1 e 2. A significância S trata-se portanto de uma medida da distância entre a entropia do grupo controle e as entropias dos grupos 1 e 2 sob a consideração de que a estatística discriminante segue uma distribuição gaussiana.

Como essa hipótese de gaussianidade nem sempre pode ser garantida, uma outra alternativa que se tem em mãos é o cálculo do p -valor (ou probabilidade de significância),

que nada mais é do que a probabilidade de se observar uma significância maior do que S em caso de a hipótese nula H_0 ser verdadeira (Theiler 1992). Em outras palavras, quanto menor o p -valor, mais evidente se torna a rejeição de H_0 . Matematicamente, o p -valor (p) pode ser expresso por:

$$p = \begin{cases} 2i, & \text{se } i \leq 0,5, \\ 2(1 - i), & \text{se } i > 0,5, \end{cases}$$

em que i é o índice de p_i , que por sua vez é o menor i -ésimo percentil inferior da estatística discriminante de todas as entropias dos grupos 1 e 2 em que se encontra o valor de estatística discriminante da entropia do grupo controle. Para fins práticos, a Tabela B.1 apresenta uma classificação dos níveis de evidência de rejeição de H_0 de acordo com os p -valores (Jesus 2008).

Tabela B.1: Evidência de rejeição de H_0 em função do p -valor

p-valor (p)	Evidência
<0,01	evidência muito forte contra H_0
0,01 - 0,05	evidência forte contra H_0
0,05 - 0,1	evidência fraca contra H_0
>0,1	pouca ou nenhuma evidência contra H_0

B.2 Teste de Wilcoxon

O teste de Wilcoxon pareado (do inglês, *Wilcoxon signed-rank test*) é um teste de hipóteses não paramétrico aplicado para o caso de distribuições simétricas contínuas. A partir dessa consideração, a média se iguala à mediana (Montgomery e Runger 2002). Esse teste é utilizado para comparar duas amostras relacionadas, amostras pareadas, ou medidas repetidas em uma única amostra para avaliar se o *rank* das médias da população difere. Tal teste pode ser usado como uma alternativa para o teste t pareado (*Student's t-test*) quando não se pode assumir gaussianidade nos dados.

O teste de Wilcoxon pressupõe as seguintes considerações:

- i os dados são pareados e advêm da mesma população;

- ii cada par é escolhido aleatória e independentemente;
- iii os dados são medidos em uma escala de intervalo que não precisa ser normal.

O teste de Wicoxon considera as seguintes hipóteses:

- H_0 : a mediana das diferenças entre os pares é zero ($\mu = 0$), ou equivalentemente, $\mu = \mu_0$;
- H_0 : a mediana das diferenças não é zero ($\mu \neq 0$), ou equivalentemente, $\mu \neq \mu_0$;

Seja N o tamanho da amostra, ou seja, o número de pares. Seja (X_{1i}, X_{2i}) , para $i = 1, \dots, N$, um conjunto de observações pareadas para duas distribuições contínuas que diferem apenas no que tange às médias. Não é necessário que as distribuições de X_1 e X_2 sejam simétricas, mas assegura-se que a distribuição das diferenças $D_j = X_{2i} - X_{1i}$ seja contínua e simétrica. As etapas do teste são:

- 1 para $i = 1, \dots, N$, calcule $|X_{2i} - X_{1i}|$ e $\text{sgn}(X_{2i} - X_{1i})$, em que sgn é a função sinal;
- 2 exclua pares com $|X_{2i} - X_{1i}| = 0$. Seja N_r o novo tamanho da amostra (sem os pares excluídos);
- 3 ordene os N_r pares restantes em ordem crescente de acordo com o maior valor da diferença $|X_{2i} - X_{1i}|$;
- 4 faça um *rank* dos pares e seja R_i o *rank*;
- 5 calcule o teste estatístico que representa o valor absoluto das somas dos *ranks*:

$$W = \left| \sum_{i=1}^{N_r} [\text{sgn}(x_{2i} - x_{1i}) \cdot R_i] \right| \quad (\text{B.5})$$

6 à medida que N_r aumenta, a distribuição amostral de W converge para uma distribuição normal. Então:

para $N_r \geq 10$, o índice *z-score* pode ser calculado: $z = \frac{W-0.5}{\sigma_W}$, $\sigma_W = \sqrt{\frac{N_r(N_r+1)(2N_r+1)}{6}}$

se $z > z_{\text{critico}}$, então rejeita-se H_0

para $N_r < 10$, compara-se com o valor crítico de uma tabela de referência (Montgomery e Runger 2002)

se $W \geq W_{crítico, N_r}$, então rejeita-se H_0

Alternativamente, o *p-valor* pode ser calculado para a enumeração de todas as possíveis combinações de W e N_r .

B.3 Coeficiente de Kappa Cohen

O coeficiente de Kappa Cohen (k) mede a concordância entre dois avaliadores/observadores (exemplo: neurologista e classificador automático) durante uma classificação. Tal índice foi definido por Jacob Cohen (Cohen 1960) em 1960 e é dado por:

$$k = \frac{P_r - P_a}{1 - P_a} \quad (\text{B.6})$$

em que P_r é a concordância relativa entre os classificadores (avaliadores) e P_a é a probabilidade de uma concordância aleatória. Se os classificadores estiverem totalmente em acordo (concordância), então $k = 1$. Por outro lado, se tais classificadores estiverem totalmente em desacordo (ou em acordo unicamente aleatório), então $k = 0$.

