

UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Enio Aguiar de Araújo

Codificação Eficiente de Sinais de Eletrocardiograma

Sã Luís - MA

2010

Enio Aguiar de Araújo

Codificação Eficiente de Sinais de Eletrocardiograma

Dissertação apresentada ao Programa de Pós Graduação em Engenharia Elétrica da UFMA, como requisito para a obtenção do grau de MESTRE em Engenharia Elétrica.

Orientador: Allan Kardec Duailibe Barros Filho

Universidade Federal do Maranhão

Sã Luís - MA

2010

Araújo, Enio Aguiar de

Codificação eficiente de sinais de eletrocardiograma / Enio Aguiar de Araújo. – São Luís, 2010.
73 f.

Orientador: Prof. Dr. Allan Kardec Duailibe Barros Filho
Dissertação (Mestrado) – Programa de Pós Graduação em Engenharia de Eletricidade, Universidade Federal do Maranhão, 2010.

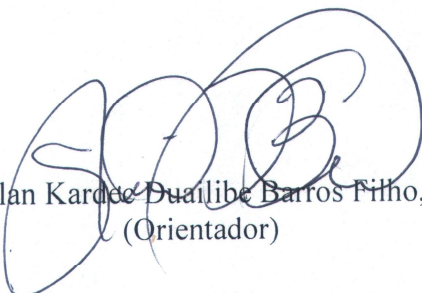
1. Eletrocardiograma – processamento de sinais. 2. Sinais digitais – codificação eficiente. I. Título.

CDU 621.391:616.12-073.97

CODIFICAÇÃO EFICIENTE DE SINAIS DE ELETROCARDIOGRAMA

Enio Aguiar de Araújo

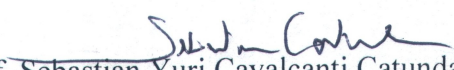
Dissertação aprovada em 28 de maio de 2010.



Prof. Allan Kardec Dualibe Barros Filho, Ph.D.
(Orientador)



Prof. Dráulio Barros de Araújo, Dr.
(Membro da Banca Examinadora)



Prof. Sebastian Yuri Cavalcanti Catunda, Dr.
(Membro da Banca Examinadora)

À comunidade científica.

Resumo

Tipicamente, em processamento digital de sinais de eletrocardiografia, são utilizadas transformações lineares visando tornar os sinais mais tratáveis de acordo com a aplicação. Para aplicações como classificação ou compressão de dados, normalmente tem-se como objetivo dessas transformações, reduzir a redundância presente nesses sinais, o que aumenta o potencial dessas aplicações. Existem diversos métodos usualmente utilizados para essa tarefa, como a transformada de Fourier, a transformada de *Wavelets*, e análise de componentes principais. Todos esses métodos têm alguma limitação, seja a utilização de um espaço predefinido e ortogonal ou utilizar apenas estatística de segunda ordem. Neste trabalho propomos a utilização do método de análise de componentes independentes para a codificação de sinais de ECG utilizando como base o conceito da neurociência de codificação eficiente. Dois resultados importantes foram obtidos, o espaço de funções bases gerado pelo método proposto se diferencia dos espaços de transformações utilizados usualmente, e, em média, o método proposto teve maior capacidade de reduzir a redundância dos sinais. Concluimos que os métodos tradicionais podem não explorar ao máximo o potencial de codificação de sinais de ECG devido às suas limitações, e que ICA pode ser um método plausível para melhorarmos os sistemas atualmente utilizados.

PALAVRAS-CHAVE: Eletrocardiograma, Codificação Eficiente, Codificação, Redução de Redundância

Abstract

Typically, in the digital processing of electrocardiography signal, linear transformations are used to turn the signals more tractable in accordance to some application. For applications such as classification or data compression, it usually aimed to reduce the redundancy present in the signals, increasing the potential of the applications. There are various methods usually used for the task, the Fourier transform, the wavelet transform and principal component analysis. All those methods have any sort of limitation, being the use of a predefined space, orthogonal spaces or the limitations to second order statistics. In this work we propose the use of the independent component analysis method for the encoding of the ECG signals, using as theoretical basis the neuroscience concept of efficient coding. Two important results were found, the basis functions space generated by the proposed method is different from the spaces seen on the usual methods, and, on average, the method can reduce the redundancy of the signal. We concluded that the traditional methods might not exploit the coding potential of ECG signals due to their limitations, and also that ICA might be a reliable method for improving the performance comparing to the current systems.

KEYWORDS: Electrocardiogram, Efficient Coding, Coding, Redundancy Reduction

Agradecimentos

A Deus.

À Humanidade.

*Ao que te ferir numa face, oferece-lhe
também a outra. E ao que te tirar a capa,
não impeças de levar também a túnica.*

*Dá a todo o que te pedir; e ao que tomar
o que é teu, não lho reclames.*

*E como vós quereis que os homens vos
façam, da mesma maneira fazei-lhes vós
também.*

Lucas 6:29-31

Sumário

Lista de Figuras	8
Lista de Tabelas	12
Lista de Abreviações	13
1 Introdução	14
2 Eletrocardiograma	17
2.1 Fisiologia do Coração	17
2.2 Morfologia do ECG	19
2.3 Alterações no ECG	21
2.3.1 Fibrilação Atrial	21
2.3.2 Morte Súbita Cardíaca	23
3 Wavelets	26
4 Análise de Componentes Principais	33
4.1 PCA por maximização de variância	34
5 Análise de Componentes Independentes	37

5.1	Definições	38
5.2	Estimação das componentes Independentes	40
5.2.1	Estimação através de Maximização de Não-Gaussianidade	40
5.2.2	Negentropia como Medida de Não-Gaussianidade	42
5.2.3	Um Algoritmo de ponto fixo para ICA	44
6	Teoria da Informação e Codificação Eficiente	46
6.1	Entropia	46
6.2	Codificação Eficiente	48
7	Material e Métodos	53
8	Resultados	56
8.1	Estimação das Funções Bases	56
8.2	Análise da Codificação	59
9	Discussão	65
9.1	Funções bases	65
9.2	Codificação	66
10	Conclusão	69
	Referências	71

Lista de Figuras

- 2.1 Representação de um coração humano mostrando partes que participam dos processos eletrofisiológicos. 18
- 2.2 Exemplo de um sinal de ECG capturado no momento de um batimento cardíaco com a nomenclatura usual das ondas em cada parte do batimento. 20
- 2.3 Exemplo de sinal de ECG de uma paciente com fibrilação atrial 22
- 2.4 Exemplo de sinal de ECG de uma paciente com fibrilação atrial, no momento da ocorrência de um *flutter* atrial 23
- 2.5 Exemplo de sinal de ECG de um paciente acometido por morte súbita cardíaca durante a gravação do sinal. A parada cardíaca aconteceu alguns minutos após o início da arritmia observada a partir de 6 s 25
- 2.6 Exemplo de sinal de ECG de uma pessoa acometida por morte súbita cardíaca durante a gravação do sinal, mostrando o momento da parada cardíaca 25
- 2.7 Exemplo de sinal de ECG de uma pessoa acometida por morte súbita cardíaca durante a gravação do sinal, alguns minutos após o evento. 25

3.1	(a) Resposta em frequência dos filtros passa-banda que codificam o sinal em duas bandas de frequência. (b) Esquema mostrando a entrada e as saídas do processo contendo as duas etapas, a filtragem e a subamostragem em cada banda, a qual é representada pela seta apontando para baixo com o número dois significando a subamostragem por dois.	29
3.2	Ilustração comparativa das transformadas de Fourier e <i>wavelets</i> . Cada retângulo é a representação de um vetor base do espaço, no caso de Fourier, à esquerda, cada função base é uma senóide com um frequência determinada, ocupando assim todo o eixo de tempo. No caso das <i>wavelets</i> a medida que se limita a localização no tempo do vetor base, ele ocupa mais espaço na frequência.	30
3.3	Banco de filtros demonstrando os níveis de codificação na transformada discreta de wavelets. O conjunto de coeficientes $[\mathbf{a}_j, \mathbf{d}_j, \mathbf{d}_{j-1}, \dots, \mathbf{d}_1]$ resultantes contem toda informação do sinal	31
4.1	Em (a) a distribuição conjunta de duas variáveis aleatórias correlacionadas. O processo de análise de componentes principais encontra os vetores bases ortogonais que descorrelacionam as variáveis, ou seja, quando projetas nesse novo espaço as variáveis estarão descorrelacionadas.	34

6.1	Esquema comparativo da codificação de um sinal sonoro (a) em duas bases diferentes (b), no caso uma de <i>wavelets</i> (acima) e Fourier (abaixo). O estímulo é representado pelos coeficientes da combinação linear das funções bases. Utilizamos as distribuições de probabilidade dos coeficientes w_n resultantes da representação para caraterizar a eficiência; quanto mais esparsa o código gerado, mais eficiente a codificação, pois menor é a entropia da representação.	50
7.1	O processo de codificação é feito em duas partes, primeiro alguns trechos do sinal são utilizados para encontrar os vetores bases e em seguida utiliza-se essa base para projetar qualquer outro bloco desse mesmo sinal.	54
8.1	Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente saudável e (b) algumas funções bases geradas pela codificação eficiente do sinal	56
8.2	Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente saudável e (b) algumas funções bases geradas pela codificação eficiente do sinal	57
8.3	Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente com fibrilação atrial e (b) algumas funções bases geradas pela codificação eficiente do sinal	57
8.4	Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente com fibrilação atrial e (b) algumas funções bases geradas pela codificação eficiente do sinal	58

8.5	Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente acometido por morte súbita cardíaca e (b) algumas funções bases geradas pela codificação eficiente do sinal	58
8.6	Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente acometido por morte súbita cardíaca e (b) algumas funções bases geradas pela codificação eficiente do sinal	58
8.7	Entropia média da codificação de todos os sinais do grupo de pessoas saudáveis para todos os métodos testados PCA, ICA e <i>Wavelets</i> bem como do sinal original sem ser codificado	59
8.8	Entropia média da codificação de todos os sinais do grupo de pacientes com fibrilação atrial para todos os métodos testados PCA, ICA e <i>Wavelets</i> bem como do sinal original sem ser codificado	60
8.9	Entropia média da codificação de todos os sinais do grupo de pacientes que sofreram morte súbita cardíaca para todos os métodos testados PCA, ICA e <i>Wavelets</i> bem como do sinal original sem ser codificado	61
8.10	Erro médio resultante da quantização dos coeficientes das várias representações para os sinais do grupo de pacientes suadáveis	62
8.11	Erro médio resultante da quantização dos coeficientes das várias representações para os sinais dos pacientes com fibrilação atrial	62
8.12	Erro médio resultante da quantização dos coeficientes das várias representações para os sinais do grupo de pacientes que sofreu morte súbita cardíaca	63

Lista de Tabelas

7.1	Informações sobre a aquisição dos sinais de ECG das bases de dados utilizadas.	54
8.1	Comparação da Entropia da codificação para os três métodos testados <i>Waveletes</i> PCA e ICA, sobre as três bases de dados, normal, fibrilação atrial e morte súbita utilizando todos os sinais. Para o cálculo da distribuição de probabilidade foi utilizada quantização uniforme de 8 bits.	61
8.2	Comparação do erro gerado pela quantização dos coeficientes para vários valores de quantização, para cada método, sobre os sinais de indivíduos saudáveis.	63
8.3	Comparação do erro gerado pela quantização dos coeficientes para vários valores de quantização, para cada método, para os sinais de pacientes com fibrilação atrial	64
8.4	Comparação do erro gerado pela quantização dos coeficientes para vários valores de quantização, para cada método, para os sinais de morte súbita .	64

Lista de Abreviações

ICA	Análise de Componententes Independentes
PCA	Análise de Componententes Principais
CWT	Transformada Contínua de Wavelets
DWT	Transformada Discreta de Wavelets
ECG	Eletrocardiograma
EEG	Eletroencefalograma

1 Introdução

Uns dos principais avanços no campo da medicina de diagnóstico deve-se à aquisição e processamento de informações medidas do corpo humano por dispositivos elétricos, como a Eletrocardiografia (ECG), Eletroencefalografia (EEG), Eletromiografia (EMG). Inicialmente toda a análise dessas informações era feita por um médico ou especialista utilizando, simplesmente, a observação do traçado desses sinais. Porém, o desenvolvimento da tecnologia de processamento digital da informação permitiu que a análise desses sinais pudesse ser feita matematicamente, utilizando-se um computador, aumentando a precisão da análise

Normalmente, no processamento digital de sinais, é importante, e às vezes essencial, que o sinal discreto que foi digitalizado passe por um processo de transformação antes de ser analisado ou utilizado em qualquer aplicação. Essa transformação consiste em criar um espaço vetorial no qual o sinal original vai ser representado, sendo que essa nova representação deve apresentar propriedades que o sinal original não possuía e que podem ser importantes de acordo com a aplicação.

Sabe-se pela Teoria da Informação que a entropia, ou certos momentos estatísticos de um sinal, são propriedades que influenciam fortemente o resultado de aplicações como filtragem, compressão ou classificação desses sinais. No caso da compressão, por exemplo, a entropia define o limite mínimo até o qual o sinal pode ser comprimido. Dentre as transformações mais utilizadas podemos destacar a transformada de Fourier, a transformada de *wavelets* as técnicas de análise de

componentes principais (*principal component analysis* - PCA) e análise de componentes independentes (*independent component analysis* - ICA).

Especificamente para o processamento de sinais de ECG a transformada de *wavelets* e PCA já foram aplicados obtendo-se resultados positivos em filtragem [1],[2] e compressão [1],[3],[4], [5] dos sinais. Porém, sabe-se que o primeiro método, assim como a transformada de Fourier, possui dois fatores limitantes: utiliza-se de funções que devem ser preestabelecidas antes do processamento, além de formarem um espaço obrigatoriamente ortogonal, apesar de permitir um processamento mais rápido. PCA por outro lado, usa informações do próprio sinal para gerar a base vetorial, porém, também é limitado à uma base ortogonal.

Por outro lado, a transformação utilizando ICA utiliza estatística de alta ordem do sinal para gerar uma base vetorial, sem restrição de ortogonalidade. Além do mais, este processo está intimamente ligado ao conceito da neurociência de Codificação Eficiente, que afirma que o cérebro deve representar as informações sensoriais com a menor quantidade de impulsos nervosos[6], [7], [8]. Isso implica que, no caso prático de codificação de sinais, essa transformação gera um novo sinal com menor redundância [9].

Dessa forma, propomos neste trabalho avaliar os efeitos da codificação eficiente em sinais de ECG utilizando ICA através do exame das propriedades da representação gerada, e com isso inferir sobre a possibilidade de aplicar esse método em situações práticas como compressão e classificação, bem como comparar com os demais métodos que são normalmente usados na maioria das aplicações, de compressão, filtragem e classificação.

O trabalho está organizado da seguinte forma: no capítulo 2 é feito um breve embasamento sobre o funcionamento do coração, sobre eletrocardiografia, além de falar um pouco sobre as doenças das quais serão utilizados aquisições de ECG nesse trabalho.

Nos capítulos 3, 4 e 5, temos uma fundamentação teórica sobre os métodos matemáticos utilizados no trabalho, a saber: *Wavelets*, análise de componentes principais e análise de componentes independentes, respectivamente. O capítulo 6 é dedicado à explicar sobre codificação eficiente, a teoria utilizada como base para o desenvolvimento dos resultados desse trabalho. No capítulo 7 são apresentados os resultados, no capítulo 8 as discussões e finalmente, no capítulo 9 a conclusão do trabalho com algumas sugestões de continuidade do trabalho.

2 Eletrocardiograma

Eletrocardiograma é o gráfico construído pelo dispositivo que mede a diferença de potencial na superfície entre dois pontos na superfície da pele gerada pela atividade coração. O funcionamento do coração é o resultado de conjunto seqüencial de descargas elétricas em seus músculos, as quais se propagam até a superfície da pele onde são medidas por eletrodos implantados sobre ela. Este exame, tipicamente não invasivo, permite aos médicos acessarem o funcionamento do coração e diagnosticar uma série de patologias, cardíacas ou não, que alteram o funcionamento normal deste órgão, e podem ser notadas através de sua atividade elétrica.

2.1 Fisiologia do Coração

Para entender como o sinal de ECG é gerado pelo coração é preciso entender como esse órgão funciona. O coração é o órgão responsável por manter a circulação sanguínea no corpo humano. Podemos dividir suas partes em três grupos, as cavidades, as válvulas, e o sistema de condução elétrica.

Na figura 2.1 vemos um modelo de coração com uma indicação das partes principais para o entendimento dos processos estudados neste trabalho. Este órgão é dotado de quatro cavidades musculares capazes de se contrair periodicamente bombeando o sangue para as artérias e veias. As cavidades superiores, os átrios esquerdo e direito, são responsáveis por receber os sangue, venoso e arterial respectivamente, e bombeá-los para as cavidades inferiores, os ventrículos. Tais cavidades são responsáveis por bombear o

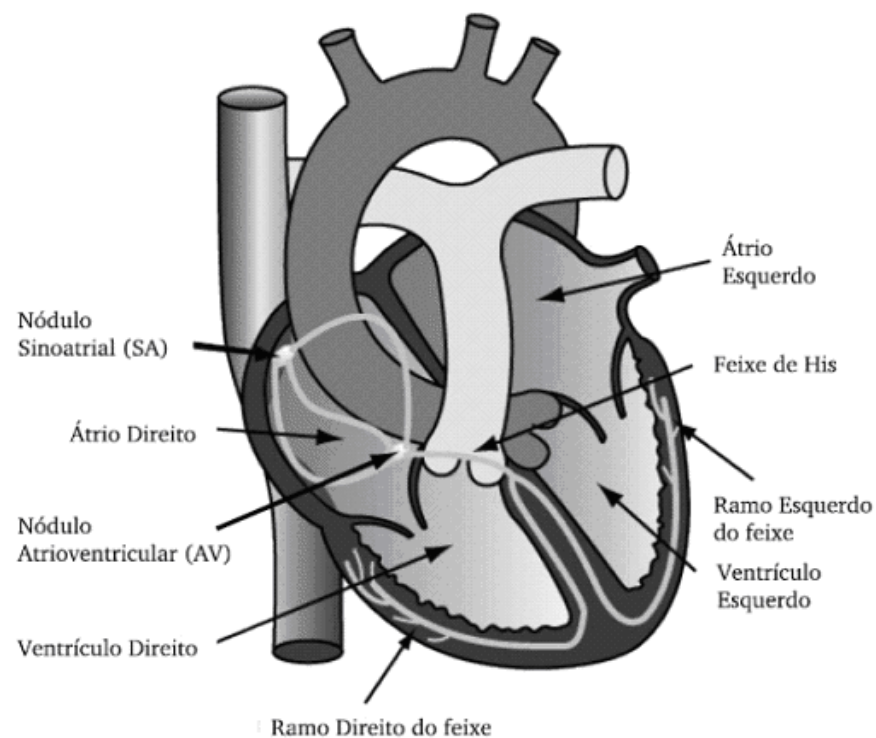


Figura 2.1: Representação de um coração humano mostrando partes que participam dos processos eletrofisiológicos.

sangue recebido dos átrios para o corpo, No caso, o ventrículo esquerdo bombeia o sangue, para ser distribuído para o corpo inteiro e suprir as células, enquanto que o ventrículo direito bombeia o sangue, já utilizado pelo corpo, para os pulmões, para ser abastecido com oxigênio.

As válvulas servem para impedir o sangue de circular no sentido contrário. Existem as válvulas átrio-ventriculares, que impedem o sangue que chegou ao ventrículo, a partir do átrio, faça o caminho inverso. E as válvulas semilunares que impedem o refluxo de sangue das artérias para os ventrículos.

O sistema de condução elétrica, por sua vez, é o responsável por estimular os músculos cardíacos a se contraírem realizando assim o batimento, bem como manter o

ritmo cardíaco na frequência adequada. É formado por células musculares diferenciadas e adaptadas especialmente com a capacidade de gerar um impulso elétrico, polarizando e despolarizando sua membrana de forma periódica. Também são capazes de conduzir e propagar um impulso elétrico gerado em uma célula vizinha.

Este sistema pode ser dividido, seguindo a sequência da condução elétrica, nas seguintes partes: nodo sinotrial, vias internodais, nodo atrioventricular, feixe de His e fibras de Purkinje, como mostrados na Figura 2.1. A diferença fundamental dessas partes é a frequência de polarização e despolarização, ou seja, da geração do impulso elétrico. Dessa forma, as células mais rápidas, localizadas no nodo sinoatrial originam um impulso elétrico que se propaga por todo o coração nessa sequência, estimulando a musculatura e gerando o batimento cardíaco.

Cada uma das etapas por onde o impulso elétrico se propaga, e a conseqüente descarga elétrica do músculo que se contrai, se propagam até a superfície da pele de onde pode ser captada por eletrodos, possibilitando assim a construção do eletrocardiograma.

2.2 Morfologia do ECG

Na Figura 2.2 é mostrado um trecho de um eletrocardiograma de um indivíduo saudável, mostrando a nomenclatura usual de cada trecho deste sinal. Cada trecho do sinal é o reflexo de uma certa atividade elétrica do coração, as quais são descritas da seguinte forma:

- **Onda P** Primeira etapa do batimento, conseqüência da atividade elétrica atrial.

O batimento originado no nodo sinoatrial se propaga pelos feixes internodais até o nodo atrioventricular. Durante sua propagação estimula as células atriais, que

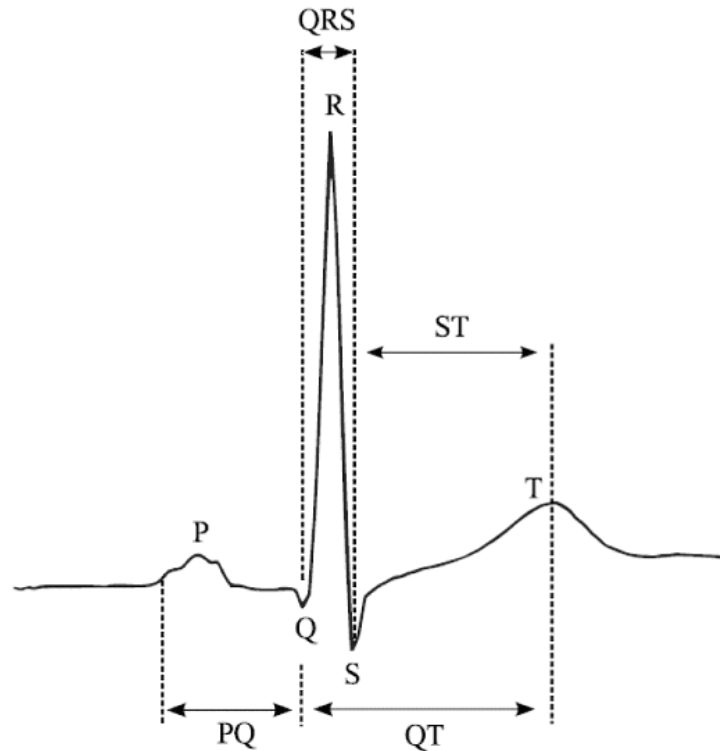


Figura 2.2: Exemplo de um sinal de ECG capturado no momento de um batimento cardíaco com a nomenclatura usual das ondas em cada parte do batimento.

despolarizam gerando atividade elétrica vista na onda P, e conseqüentemente, contraindo o músculo atrial.

- **Complexo QRS:** Resultado da despolarização ventricular. Após passar pelos átrios o impulso elétrico chega ao nodo atrioventricular onde sofre um atraso, percebido no intervalo PQ, então segue pelo feixe de His e logo em seguida para as fibras de Purkinje, que o músculo ventricular causando sua rápida despolarização, resultadando, no ECG o complexo QRS. É importante lembrar que a repolarização atrial ocorre juntamente com o complexo QRS, o que a torna praticamente indiscernível, uma vez que tem um nível de tensão bem mais baixo.
- **Onda T:** Resultado da repolarização dos ventrículos, que ocorre algum tempo após

sua despolarização, tempo esse percebido pelo intervalo ST.

É importante ressaltar que esse padrão de funcionamento de um indivíduo saudável se mantém aproximadamente constante para cada batimento, a única diferença ocorre no intervalo entre a onda T e o próximo batimento cardíaco, ou seja, todos os batimentos tem o mesmo formato.

2.3 Alterações no ECG

A análise do ECG realizada neste trabalho foi feita tanto para sinais de ECG de pacientes saudáveis, quanto sinais que apresentam alguma anomalia devido a alguma enfermidade cardíaca que se expresse no ECG. As formas como as doenças se expressam no ECG são bem variadas, desde pequenas deformidades em uma das ondas até formas de ondas de batimento totalmente irregulares e diferenciados de um batimento normal. Além disso, as doenças podem se expressar tanto como deformações que se repetem da mesma forma a cada batimento ou em episódios localizados. Escolhemos neste trabalho utilizar dois grupos de sinais com enfermidades diferentes, além dos sinais normais: fibrilação atrial e morte súbita cardíaca.

2.3.1 Fibrilação Atrial

Das arritmias cardíacas classificadas até hoje, a Fibrilação Atrial é considerada a mais comum. Aproximadamente 4% da população mundial acima de 60 anos, já experimentou um episódio dessa arritmia, [10]; acima de 80 anos, a prevalência dessa arritmia chega a 8%. Além disso, Essa patologia é uma das maiores causas de ataques

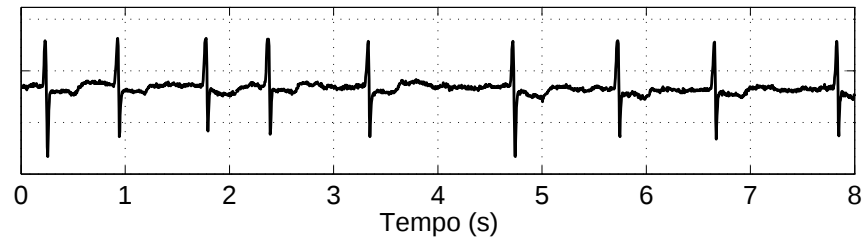


Figura 2.3: Exemplo de sinal de ECG de uma paciente com fibrilação atrial

cardíacos, pacientes com Fibrilação Atrial tem de 2 a 7 vezes mais chances de ter um ataque cardíaco [11].

Essa doença é caracterizada pela progressiva fibrose dos átrios, causada pela dilatação dos mesmos. Esse processo tem sua origem em quase qualquer processo que aumente a pressão intra cardíaca. O resultado do processo de fibrose causa uma completa desordem no processo natural de propagação do impulso elétrico pelos átrios. Normalmente o impulso elétrico percorre todo seu caminho através do coração de forma ordenada, começando no nodo sinoatrial até os ventrículos. No caso da fibrilação, ao invés de começar pelo nodo sinoatrial vários impulsos descoordenados são gerados na superfície dos átrios. Sem organização, o nodo atrioventricular passa a gerar impulsos elétricos de forma irregular e normalmente acelerada quando comparada o ritmo normal de um indivíduo em repouso. Conseqüentemente, essas mudanças no padrão elétrico do batimento são refletidas no sinal de eletrocardiograma. A alteração mais clara é a ausência da onda P, juntamente com um ritmo bastante irregular do batimento, como mostrado na Figura 2.3.

Outra situação que pode acontecer em pessoas com fibrilação atrial é o chamado *flutter* atrial. Nesse estágio ao invés de completamente desordenada a atividade elétrica atrial, há o surgimento de batimentos fora do nodo sinoatrial, mas de forma bem acelerada chegando a 500 batimentos por minuto. Desse modo, os batimentos

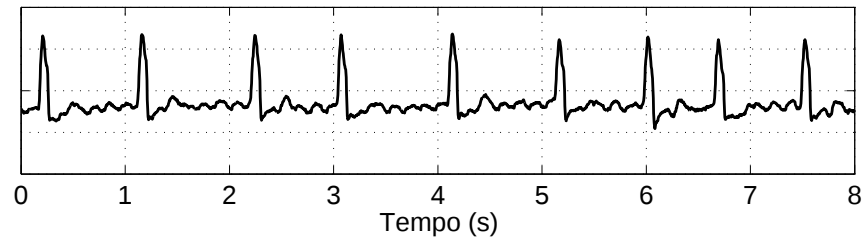


Figura 2.4: Exemplo de sinal de ECG de uma paciente com fibrilação atrial, no momento da ocorrência de um *flutter* atrial

não conseguem ser transmitidos para o ventrículo fazendo com que ambos batam em frequências diferentes, sendo os átrios bem mais acelerados. No ECG, o *flutter* atrial é bem claro, pois esse batimento do átrio faz surgir ondas parecidas com a onda P, consecutivas ao longo de todo o sinal, como podemos ver na figura 2.4.

Começando em geral assintomática, existe uma certa dificuldade em diagnosticar essa arritmia. Ainda, os primeiros sintomas que se expressam, como palpitações, e às vezes fraqueza, são normalmente ignorados pelo paciente. Além disso, uma pesquisa mostra [12] que mesmo combinando softwares classificadores de patologias cardíacas com as primeiras análises de eletrocardiograma, por técnicos ou enfermeiras, não se tem um diagnóstico suficientemente acurado na identificação da fibrilação atrial.

2.3.2 Morte Súbita Cardíaca

Morte súbita cardíaca é a morte resultante de uma parada brusca e inesperada das funções do coração. Geralmente ocorre pouco tempo depois da aparição de sintomas ou pode ocorrer sem nenhum sintoma aparente prévio. Praticamente qualquer enfermidade cardíaca pode preceder e de alguma forma ser causadora desta condição. A causa mais comum, contudo, é a isquemia, ou falta de suprimento sanguíneo no miocárdio,

gerado pela oclusão das artérias coronárias. A morte súbita cardíaca está associada também a outros fatores como danos ou deformações no miocárdio, sejam congênitos ou causados por outras doenças que afetam o coração, até mesmo uma parada cardíaca anterior. Outro motivo, que não está associado a nenhuma enfermidade cardíaca anterior, é a grande descarga de adrenalina em atletas, por exemplo, causado por um esforço exagerado do organismo, e conseqüentemente do coração. Contudo, em grande parte dos casos, em que os pacientes estão sobre monitoramento médico, é possível o retorno dessa condição. Porém, o paciente ressuscitado pode sofrer danos graves no miocárdio devido à isquemia temporária, aumentando seriamente o risco de uma segunda parada cardíaca [13].

Existem muitas pesquisas tentando relacionar as condições anteriores do paciente com a súbita parada cardíaca, para tentar prever e até mesmo evitá-la [14]. Uma das tentativas para isso é utilizando eletrocardiografia. Por se tratar, em geral, de pacientes com condições distintas, porém severas, os sinais captados de ECG se apresentam com fortes alterações, que se tornam ainda mais severas algum tempo antes do fenômeno, como podemos ver, por exemplo, no sinal da figura 2.5. Na figura 2.6 vemos um exemplo de um sinal de ECG no momento da parada cardíaca. Na figura 2.7 vemos o sinal do mesmo paciente da figura 2.6 alguns minutos após o retorno da atividade cardíaca. Percebe que leva um tempo para a atividade cardíaca voltar ao normal.

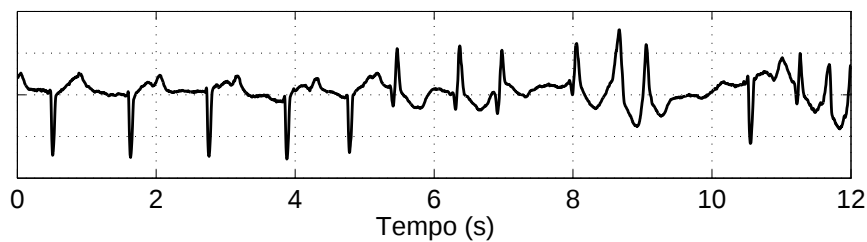


Figura 2.5: Exemplo de sinal de ECG de um paciente acometido por morte súbita cardíaca durante a gravação do sinal. A parada cardíaca aconteceu alguns minutos após o início da arritmia observada a partir de 6 s

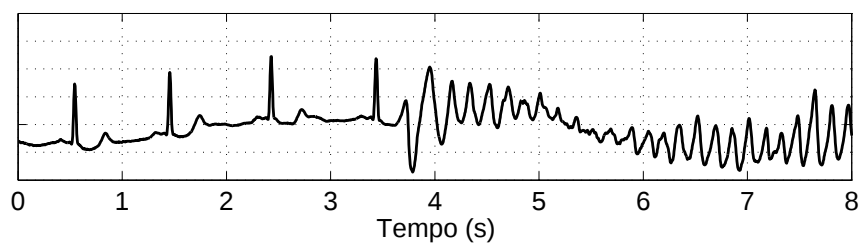


Figura 2.6: Exemplo de sinal de ECG de uma pessoa acometida por morte súbita cardíaca durante a gravação do sinal, mostrando o momento da parada cardíaca

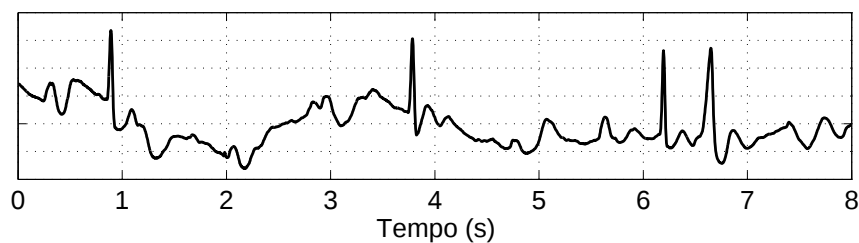


Figura 2.7: Exemplo de sinal de ECG de uma pessoa acometida por morte súbita cardíaca durante a gravação do sinal, alguns minutos após o evento.

3 Wavelets

A idéia de se representar uma informação temporal em outros domínios tem sua história marcada pela contribuição de Fourier no início do século XIX. A idéia desenvolvida por Fourier foi de que qualquer informação temporal pode ser representada como uma combinação de componentes senoidais. Desta fica possível conhecer os fenômenos harmônicos que ocorrem no sinal analisado através dos coeficientes de ponderação das senóides. Matematicamente conhecemos a representação de Fourier para uma função temporal $f(t)$ da seguinte forma:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(w) e^{iwt} dw \quad (3.1)$$

Assim, $\hat{f}(w)$ descreve a função $f(t)$ em termos de senóides complexas. A teoria de Fourier também demonstra que podemos calcular $\hat{f}(w)$ ou seja, a amplitude de cada senóide complexa e^{iwt} , o que chamamos de Transformada de Fourier, da seguinte forma:

$$\hat{f}(w) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) e^{-iwt} dt \quad (3.2)$$

Essa ferramenta matemática se tornou útil em diversas áreas. Na geologia, por exemplo, pode ser usada para distinguir fenômenos sismológicos; em processamento de áudio, como equalizadores, por exemplo, que alteram as amplitudes das componentes senoidais isoladamente. Em processamento de sinais biomédicos teve uma alta relevância uma vez que esses sinais costumam apresentar componentes de frequência bem definidas. Através da análise de Fourier da Eletroencefalografia, por exemplo, percebe-se uma clara

variação das frequências do sinal medido de acordo com o estado do indivíduo.

Contudo, a análise de Fourier apresenta uma limitação, ela nos permite saber apenas quais as frequências que estão contidas no sinal analisado, mas não em que momentos essas frequências acontecem. A solução imediata para esse problema seria calcular a transformada em janelas. Porém, é importante perceber que, a medida que diminuimos o tamanho da janela, a sua transformada de Fourier ocupa maior "espaço" do domínio da frequência; no limite, a transformada de Fourier de um impulso, ocupado todo o espectro de frequências. Ou seja, não se pode calcular exatamente o momento e a frequência de um evento em um sinal qualquer. Esse princípio da incerteza, análogo ao princípio da incerteza primeiramente postulado na mecânica quântica por Heisenberg, cria um limite para a descrição simultânea de um sinal no tempo e na frequência.

Nesse contexto surgiu a idéia de se utilizar ao invés de senóides como base para a representação do sinal, funções com localizações específicas de tempo e frequência, de modo a descrever com mais precisão sinais mais complexos com eventos de frequências variadas em momentos variados. As *Wavelets* seriam portanto funções que descrevem uma forma de onda com uma localização específica no tempo e na frequência [15], sendo que para mapear o espaço de tempo-frequência bastaria deslocar a *wavelet* no tempo ou escaloná-la, deslocando-a assim no eixo de frequência. A formalização dessas idéias resultou na idéia da Transformada de *Wavelets*.

Formalmente uma *wavelet* pode ser definida como:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad (3.3)$$

em que u é o parâmetro que desloca a *wavelet* no tempo e s de escalonamento

da *wavelet*, alterando sua posição no espectro de frequência. A transformada de *wavelet* de uma função $f(t)$ na posição u e na escala s é definida pela correlação da *wavelet* com a função $f(t)$, ou seja:

$$Wf(u, s) = \int_{-\infty}^{+\infty} f(t)\psi_{u,s}(t)dt = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right)dt \quad (3.4)$$

Essa é uma transformação de uma variável uni-dimensional contínua em uma variável bi-dimensional também contínua, por isso denominada Transformada de *Wavelet* Contínua (*Continuous Wavelet Transform* - CWT). Apesar da capacidade de representar uma gama de sinais mais variados essa transformação é redundante, uma vez que existe uma sobreposição entre as infinitas versões escalonadas e transladadas das *wavelets*. Além disso, computacionalmente essa representação aumenta grandemente a quantidade de espaço físico necessário para descrever a mesma função, além de dificultar a reconstrução do sinal a partir de sua transformação, devido à redundância.

Esses problemas foram solucionados com o desenvolvimento da Transformada de *Wavelets* Discreta (do inglês *Discrete Wavelet Transform* - DWT). Além de eliminar a redundância entre as *wavelets*, a base passou a ser ortogonal. Ou seja, pode ser tratada analiticamente. A origem dessa técnica, por outro lado, está relacionada com outras linhas de estudo em processamento digital de sinais, a teoria de Codificação por Sub-bandas e Análise em Multiresolução.

A codificação por sub-bandas consiste em utilizar um banco de filtros passa-banda com frequências de corte distintas, que cubram todo o espectro de frequências. Nota-se a semelhança com a transformada de *wavelets* contínua uma vez que wavelets são basicamente filtros passa-banda, pois tem uma localização específica na frequência. A

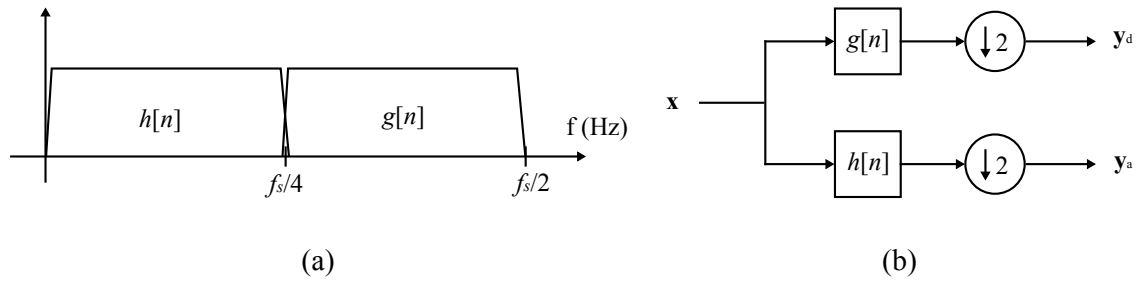


Figura 3.1: (a) Resposta em frequência dos filtros passa-banda que codificam o sinal em duas bandas de frequência. (b) Esquema mostrando a entrada e as saídas do processo contendo as duas etapas, a filtragem e a subamostragem em cada banda, a qual é representada pela seta apontando para baixo com o número dois significando a subamostragem por dois.

diferença aqui é que os filtros são digitais e sem sobreposição em suas bandas de frequência, de modo que temos uma base ortogonal.

Dividindo um sinal discreto $\mathbf{x}$ em duas bandas de frequência, com iguais larguras de banda, por exemplo, como nos filtros $g[n]$ e $h[n]$ na figura 3.1a, teremos dois vetores com mesmo tamanho do sinal original. Contudo sabe-se pelo teorema da amostragem de Nyquist que um sinal cuja frequência de amostragem é f_s contem todas as frequências até $f_s/2$; além disso, da teoria de amostragem de sinais sabe-se também que um sinal subamostrado para uma frequência f_s/n , onde $n \in \mathbf{I}_+$ tem seu espectro de alta frequência no intervalo $[f_s/n; 2f_s/n]$ refletido para a região de baixa frequência no intervalo $[0; f_s/n]$. Dessa forma a subamostragem dos dois sinais filtrados não perde nenhuma informação sobre o sinal original.

Podemos expressar essas operações da seguinte forma:

$$d[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-2k] \longrightarrow \mathbf{d} = \mathbf{x} * \mathbf{g} \quad (3.5)$$

$$a[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-2k] \longrightarrow \mathbf{a} = \mathbf{x} * \mathbf{h} \quad (3.6)$$

sendo $h[n]$ e $g[n]$ são os filtros passa-baixa e passa-alta respectivamente. As respostas $\mathbf{y}_d$ e

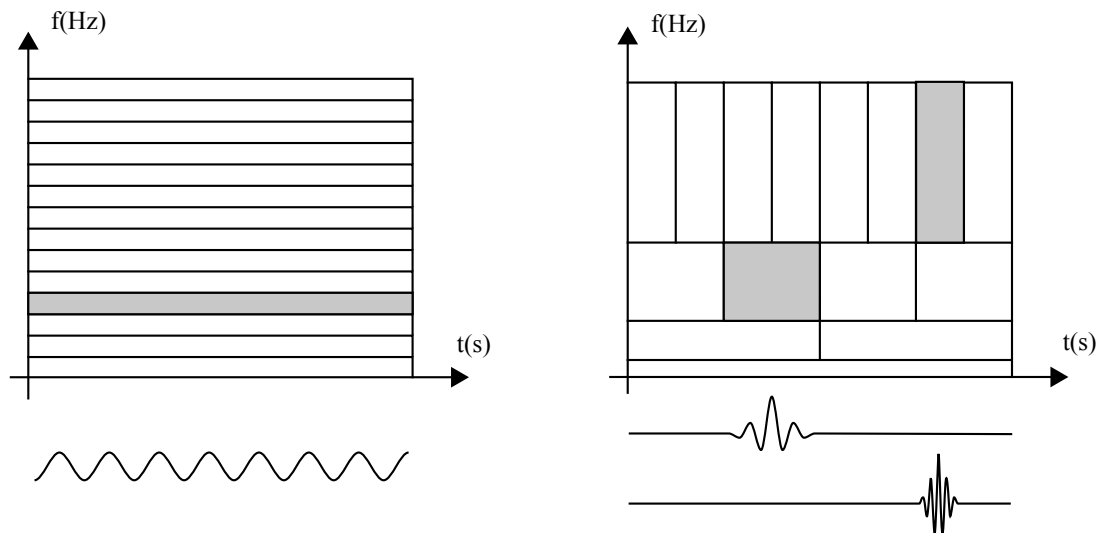


Figura 3.2: Ilustração comparativa das transformadas de Fourier e *wavelets*. Cada retângulo é a representação de um vetor base do espaço, no caso de Fourier, à esquerda, cada função base é uma senóide com um frequência determinada, ocupando assim todo o eixo de tempo. No caso das *wavelets* a medida que se limita a localização no tempo do vetor base, ele ocupa mais espaço na frequência.

y_a resultantes da decomposição de x são denominados coeficientes de detalhe e coeficientes de aproximação respectivamente.

Da mesma forma que dividimos o sinal discreto em duas bandas, podemos dividir cada uma das bandas individuais utilizando os mesmos filtros. O uso de um mesmo filtro para analisar diversas resoluções de um sinal, através de sub ou superamostragem é conhecido como análise em Multiresolução.

Com isso, está construída a base para a transformada discreta de *wavelets*. A idéia é conseguir, a partir de uma *wavelet* $\psi(t)$, gerar o filtro $g[n]$ que permita que a análise de multiresolução com os filtros de bancos gerem um espaço ortogonal no qual os coeficientes da filtragem descrevem todo o sinal.

Comparativamente à transformada de Fourier, podemos expressar graficamente a DWT como na figura 3.2 onde cada nível de decomposição do banco

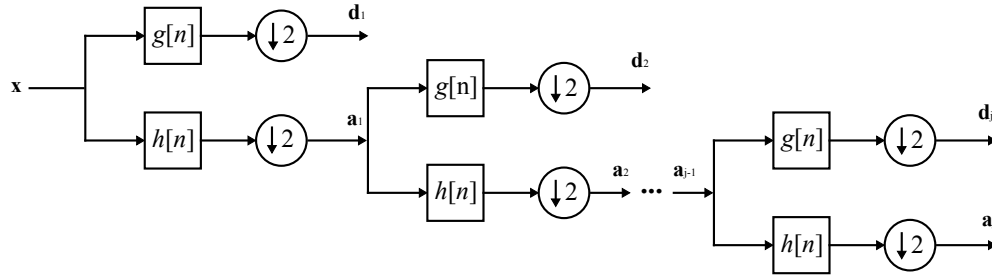


Figura 3.3: Banco de filtros demonstrando os níveis de codificação na transformada discreta de wavelets. O conjunto de coeficientes $[a_j, d_j, d_{j-1}, \dots, d_1]$ resultantes contem toda informação do sinal

de filtros na figura 3.3 é associado à uma *wavelet*. Cada retângulo representa a extensão do filtro no tempo e na frequência, assim, ocupando todo o espectro, a representação do sinal é completa. No caso de Fourier a decomposição só permite saber quais frequências compõem o sinal, mas não em que momento; no caso da DWT sabemos a frequência e em que momento ela tem mais energia ao longo do sinal.

Na prática já existem diversos filtros de *wavelets* $g[n]$ que geram um espaço ortogonal, ou seja, que permitem a utilização da DWT. Apesar de terem a mesma característica no módulo da amplitude, ou seja, de filtro passa-banda, tem estruturas temporais diferentes, o que faz com que capturem diferentes informações para um mesmo sinal. *Wavelets* muito utilizadas são Haar, Daubechies, Symlets, Coiflets, entre outras.

De forma geral, cada nível de decomposição da DWT pode ser descrito matematicamente da seguinte maneira:

$$d_{j+1}[n] = \sum_{k=-\infty}^{\infty} a_j[k]g[n - 2k] \quad (3.7)$$

$$a_{j+1}[n] = \sum_{k=-\infty}^{\infty} a_j[k]h[n - 2k] \quad (3.8)$$

sendo que a_0 é o sinal x .

A quantidade de níveis de decomposição utilizada é limitada é determinada

por dois fatores. A quantidade de pontos do sinal original, que não pode ser decimado infinitamente, já que se trata de um sinal discreto finito. A outra forma é heurística, é necessária uma quantidade de níveis que permita a captura de informações relevantes do sinal, a partir do momento em que as divisões não diferenciam mais os coeficientes, ou seja, os coeficientes $\mathbf{a}_j$ e $\mathbf{d}_j$ tendem a se tornar parecido, não é necessário mais um outro nível.

Em processamento de sinais de biomédicos, e especialmente sinais de ECG, a DWT tem diversas aplicações, desde a remoção de ruídos como o da rede ou ruídos de baixa frequência, à compressão [3] [16] e classificação de sinais [2] [17]. Sua aplicabilidade se dá devido à capacidade que esta transformação possui de capturar as informações relevantes dos sinais de ECG ao longo do espaço de tempo-frequência. Note a semelhança da própria divisão em bandas da DWT com a estrutura do ECG, que é composto por oscilações rápidas localizadas no tempo, e com alta frequência, no caso do complexo QRS, bem como oscilações mais lentas, como a onda P. Ainda, a sequência dos batimentos é praticamente harmônica, ou seja, composta de uma única frequência que se espalha por todo o sinal, além dos próprios ruídos de baixa frequência como o de respiração, que pode estar espalhado temporalmente por todo o sinal, sendo também quase harmônicos.

4 Análise de Componentes Principais

A Análise de componentes principais, ou PCA, é um método estatístico que visa construir, a partir de um conjunto de variáveis, uma base para representar esse conjunto, com menor redundância. Para medir a redundância o método utiliza a correlação entre as amostras, ou variáveis, do conjunto de dados. O fato de usar somente estatística de segunda ordem pode ser proveitoso quando o conjunto é formado por variáveis aleatórias gaussianas correlacionadas, nesse caso PCA consegue retirar toda a redundância entre as variáveis [18].

Analizando o conjunto de dados como um conjunto de variáveis aleatórias de dimensão q , $\mathbf{x} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$. Considerando também que sejam variáveis gaussianas correlacionadas com média zero, ou seja, centradas no eixo de coordenadas, o problema de PCA pode ser entendido como uma operação em duas partes. Primeiro uma busca de um novo eixo de coordenadas ortogonal, que torne as variáveis descorrelacionadas, através somente de uma rotação, ou seja, uma transformação linear. Segundo, a transformação deve identificar qual a direção de maior variância da variável, ou seja, onde o sinal “guarda” mais informação. Para o caso de duas variáveis com dimensão $q = 2$, pode-se ver o problema resolvido na figura 4.1

A representação criada por PCA descorrelaciona as variáveis. Quando a distribuição das variáveis é gaussiana, como no caso da figura, 4.1 as variáveis se tornam não só descorrelacionadas como também estatisticamente independentes.

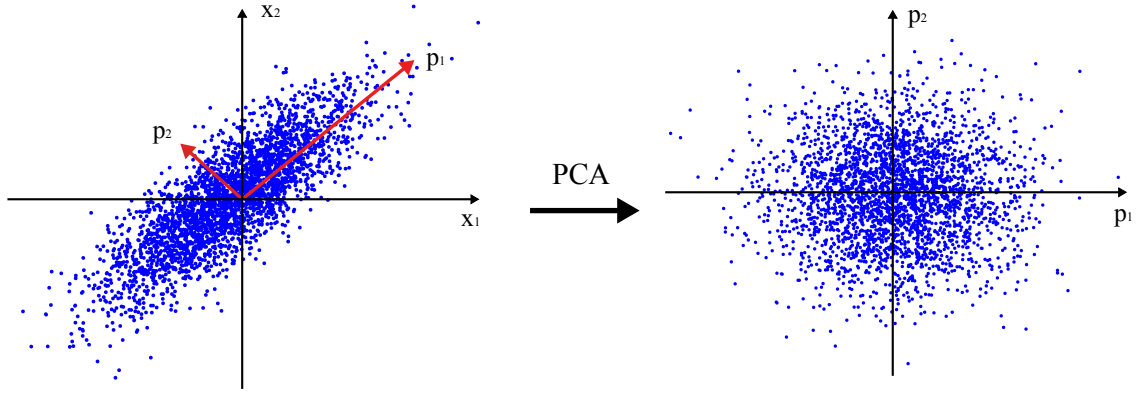


Figura 4.1: Em (a) a distribuição conjunta de duas variáveis aleatórias correlacionadas. O processo de análise de componentes principais encontra os vetores bases ortogonais que descorrelacionam as variáveis, ou seja, quando projetas nesse novo espaço as variáveis estarão descorrelacionadas.

4.1 PCA por maximização de variância

PCA pode ser calculado diretamente pela maximização da variância.

Matematicamente, consideremos a combinação linear:

$$\mathbf{y}_1 = \sum_{k=1}^n w_{k1} \mathbf{x}_k = \mathbf{w}_1^T \mathbf{x} \quad (4.1)$$

onde $\mathbf{x} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$ são as amostras analisadas, $\mathbf{w}_1 = [\mathbf{w}_{11} \ \dots \ \mathbf{w}_{n1}]$ é o vetor de coeficientes. Dizemos que $\mathbf{y}_1$ é a primeira componente principal quando sua variância é máxima. A variância depende tanto da direção quanto da norma desse vetor peso, crescendo indefinidamente com a norma. Por isso, para encontrar a primeira componente principal, basta maximizarmos a variância de $\mathbf{y}_1$ através de uma função de custo, mantendo a norma de $\mathbf{w}_i$ constante e igual a 1:

$$J^{PCA}(\mathbf{w}_1) = E[\mathbf{y}_1^2] = E[(\mathbf{w}_1^T \mathbf{x})^2] = \mathbf{w}_1^T E[\mathbf{x}\mathbf{x}^T] \mathbf{w}_1 = \mathbf{w}_1^T \mathbf{C}_x \mathbf{w}_1 \quad (4.2)$$

sendo que:

$$\|\mathbf{w}_1\| = 1 \quad (4.3)$$

A maximização dessa função de custo resulta na primeira componente principal. Sabe-se da álgebra que a solução para a maximização de $\mathbf{w}_1^T \mathbf{C}_x \mathbf{w}_1$ é dada pelos vetores $\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_n$ que são os auto-vetores da matriz de correlação $\mathbf{C}_x$. A ordem dos auto-vetores é tal que os auto-valores correspondentes $d_1, d_2, \dots, d_n$ satisfazem $d_1 \geq d_2 \geq \dots \geq d_n$. A solução para a maximização de $J^{PCA}(\mathbf{w}_1)$ é:

$$\mathbf{w}_1 = \mathbf{e}_1 \quad (4.4)$$

de onde resulta que a primeira componente principal é de x é $\mathbf{y}_1 = \mathbf{e}_1^T \mathbf{x}$

A busca de uma nova componente $\mathbf{y}_m$ é feita da mesma forma, porém com a restrição de que a próxima componente sempre seja ortogonal, descorrelacionada, às componentes principais já estimadas $\mathbf{y}_k, \mathbf{y}_{k-1}, \dots, \mathbf{y}_1$, ou seja:

$$E[\mathbf{y}_m \mathbf{y}_k] = 0, \ k < m \quad (4.5)$$

de onde resulta que:

$$E[\mathbf{y}_m \mathbf{y}_k] = E[(\mathbf{w}_m^T \mathbf{x})(\mathbf{w}_k^T \mathbf{x})] = \mathbf{w}_m^T \mathbf{C}_x \mathbf{w}_k = 0 \quad (4.6)$$

Dessa forma, para a segunda componente principal temos:

$$\mathbf{w}_2^T \mathbf{C}_x \mathbf{w}_1 = 0. \quad (4.7)$$

Como sabemos que $\mathbf{w}_1 = \mathbf{e}_1$, e $\mathbf{C}_x \mathbf{e}_1 = d_1 \mathbf{e}_1$, temos:

$$d_1 \mathbf{w}_2^T \mathbf{e}_1 = 0 \quad (4.8)$$

A solução para a segunda componente principal, de maior variância, ortogonal ao primeiro auto-vetor é o segundo autovetor, portanto

$$\mathbf{w}_2 = \mathbf{e}_2 \quad (4.9)$$

Dessa forma podemos inferir:

$$\mathbf{w}_k = \mathbf{e}_k \quad (4.10)$$

sendo a k -ésima componente principal:

$$\mathbf{y}_k = \mathbf{e}_k^T \mathbf{x} \quad (4.11)$$

Dessa forma, para o cálculo das componentes principais e dos vetores $\mathbf{w}$ que descorrelacionam essas componentes $\mathbf{y}$ basta se decompor a matriz de autocorrelação dos dados de entrada, em seus autovetores e autovalores. Esse resultado faz de PCA um método rápido em comparação com outros métodos estatísticos como ICA e útil quando se trata de variáveis aleatórias gaussianas.

5 Análise de Componentes Independentes

Análise de componentes independentes foi primeiramente proposto como um método para separação cega de fontes, misturadas linearmente. Como motivação para o método temos o clássico problema das misturas de áudio captadas por múltiplos microfones, conhecido em inglês *cocktail party problem*. Imagine um ambiente com vários interlocutores falando ao mesmo tempo, três, por exemplo, s_1 , s_2 e s_3 , sendo que existem também três microfones, distando diferentemente dos interlocutores, mas captando as três vozes misturadas, x_1 , x_2 e x_3 . Podemos modelar este sistema da seguinte forma:

$$\begin{aligned}x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \\x_3(t) &= a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t)\end{aligned}\tag{5.1}$$

O objetivo é a partir apenas das misturas gravadas recuperar os sinais originais das falas dos interlocutores $s_1(t)$, $s_2(t)$ e $s_3(t)$. Se tivéssemos de posse dos coeficientes a_{mn} , que chamamos de matriz de mistura A , resolveríamos simplesmente invertendo a matriz. Porém no *cocktail party problem* temos posse apenas das misturas. Para simplificar o modelo ignoramos certas distorções de fase devido a distância dos microfones ou outras fontes de ruído. A solução para o problema pode ser alcançada simplesmente pela suposição de que as fontes originais são **estatisticamente independentes** em cada momento t . Essa condição foi suficiente para desenvolver um algoritmo que estime simultaneamente a matriz de mistura A e as fontes $\mathbf{s}$, ou componentes independentes

[18].

O desenvolvimento de ICA, e certamente outros algoritmos de separação cega de fontes, permitiram diversos avanços no processamento de sinais biológicos, como por exemplo, de eletroencefalografia (EEG). A aquisição desse sinal segue um princípio parecido com o exemplo anterior: temos vários eletrodos espalhados pelo escalpo captando sinais elétricos com fontes em localizações diferenciadas no cérebro, que podem ser estatisticamente independentes.

Outra aplicação importante, a qual será utilizada neste trabalho, é a de extração de características. Apesar da abordagem um pouco diferente, o princípio de separação de fontes também pode ser empregado. A idéia consiste em considerar um único sinal que pode ser modelado em janelas isoladas como combinação linear de várias características. Consideramos que cada coluna da matriz A corresponde a uma característica e que cada sinal $s_n(t)$ é a seqüência de coeficientes que pondera uma das características na matriz A para gerar um determinado trecho do sinal original. Nesse caso a matriz funciona como uma base vetorial para modelar trechos do sinal original. Portanto, ICA gera uma transformação linear, assim como Fourier, PCA e *wavelets*, para codificar um sinal qualquer, com a vantagem de gerar uma base com menos restrições do que as demais, como a da ortogonalidade. Vamos ver que isso tem importantes implicações em codificação de sinais digitais.

5.1 Definições

Considere que sejam observados n misturas lineares $\mathbf{x}_1, \dots, \mathbf{x}_n$ de um sinal, modeladas como combinação linear de m vetores de uma base vetorial A , ou simplesmente

funções bases:

$$\mathbf{x}_j = s_{1j}\mathbf{a}_1 + s_{2j}\mathbf{a}_2 + \cdots + s_{mj}\mathbf{a}_n \quad j = 1, \dots, n \quad (5.2)$$

e que cada amostra $\mathbf{x}_j$, assim como cada componente independente $\mathbf{s}_j$ seja uma variável aleatória. Em análise funcional e suas aplicações, um espaço de funções pode ser visto como um espaço vetorial de infinitas dimensões cujos vetores bases são funções e não vetores. Isso significa que cada função no espaço de funções pode ser representada como combinação linear das funções base.

Sem perda de generalidade, supomos que tanto as variáveis originadas das amostras da imagem quanto aquelas das componentes independentes têm média zero. Por conveniência, será usada a notação vetorial em vez de somas. Dessa maneira, podemos reescrever (5.2) da seguinte forma:

$$\mathbf{x} = \mathbf{A}\mathbf{s}. \quad (5.3)$$

Esse modelo descreve os dados observados pelo processo de mistura das componentes independentes $\mathbf{s}_n$, que não podem ser observadas diretamente. É preciso estimar tanto as componentes independentes $\mathbf{s}$ quanto a matriz de funções bases, ou matriz de características, $\mathbf{A}$, que também é desconhecida, pois tudo o que se observa são as amostras do sinal, $\mathbf{x}$.

Para tanto, é preciso fazer algumas suposições tão gerais quanto possível. Portanto, supomos que [18]:

1. As componentes $\mathbf{s}_n$ são estatisticamente independentes;
2. Elas têm distribuições não-gaussianas;
3. Por motivos de simplicidade, a matriz $\mathbf{A}$ seja quadrada.

O modelo de ICA apresenta algumas ambigüidades em relação às componentes independentes:

1. Não se pode determinar suas variâncias (energias);
2. Não se pode determinar a sua ordem.

Ambas derivam do fato de $\mathbf{s}$ e $\mathbf{A}$ serem desconhecidos. No item 1, existe ambigüidade já que qualquer escalar α_n multiplicando uma das fontes $\mathbf{s}_n$ pode ser cancelado dividindo-se a coluna $\mathbf{a}_n$ correspondente pelo mesmo escalar α_n , ou seja:

$$\mathbf{x}_n = \sum_n \left(\frac{1}{\alpha_n} \mathbf{a}_n \right) (s_{jn} \alpha_n). \quad (5.4)$$

Daí também ocorre a ambigüidade de sinal, pois é possível multiplicar uma componente por -1 . Já no item 2, a ambigüidade ocorre devido à possibilidade de se alterar livremente a ordem dos termos em (5.2), denominando qualquer componente como a primeira.

5.2 Estimação das componentes Independentes

5.2.1 Estimação através de Maximização de Não-Gaussianidade

A não-gaussianidade é um elemento chave para a estimação do modelo de ICA, pois a matriz $\mathbf{A}$ não é identificável quando as componentes independentes têm distribuições gaussianas [3]. Assumimos que $\mathbf{x}$ seja uma das amostra do sinal, distribuído de acordo com o modelo de ICA em (5.3) e que todas as componentes independentes $\mathbf{s}_n$ têm distribuições iguais. Para estimar as componentes independentes, basta encontrar as

combinações lineares corretas de $\mathbf{x}_j$, de modo que:

$$\mathbf{s} = \mathbf{A}^{-1}\mathbf{x}. \quad (5.5)$$

Suponha, então, uma combinação linear qualquer de vários $\mathbf{x}_i$ dada por

$$\mathbf{y} = \mathbf{b}^T \mathbf{x} \quad (5.6)$$

como $\mathbf{x} = \mathbf{A}\mathbf{s}$, pode-se escrever:

$$\mathbf{y} = \sum_i \mathbf{b}_i \mathbf{x}_i \quad (5.7a)$$

$$= \mathbf{b}^T \mathbf{A}\mathbf{s}, \quad (5.7b)$$

onde $\mathbf{b}$ deve ser determinado. A partir de (5.7b) podemos observar que $\mathbf{y}$ é uma combinação linear de $\mathbf{s}_i$, com coeficientes dados por $\mathbf{q} = \mathbf{b}^T \mathbf{A}$. Logo obtemos:

$$\mathbf{y} = \mathbf{q}^T \mathbf{s} \quad (5.8a)$$

$$= \sum_i \mathbf{q}_i \mathbf{s}_i. \quad (5.8b)$$

Se $\mathbf{b}$ corresponder a uma das linhas da inversa de $\mathbf{A}$, então $\mathbf{y}$ será uma das componentes independentes e, nesse caso, apenas um dos elementos de $\mathbf{q}$ será igual a 1, enquanto todos os outros serão iguais a zero.

Lembre-se, porém, que apenas o vetor de mistura $\mathbf{x}$ é conhecido, e, portanto o vetor $\mathbf{b}$ não pode ser determinado exatamente. Precisa-se, encontrar um estimador que forneça uma boa aproximação de $\mathbf{b}$.

Uma forma de determinar $\mathbf{b}$ é variar os coeficientes em $\mathbf{q}$ e então verificar como a distribuição de $\mathbf{y} = \mathbf{q}^T \mathbf{s}$ muda. Já que, conforme o Teorema do Limite Central [19], a soma de duas variáveis aleatórias independentes é mais gaussiana que as variáveis originais

dado que tenham a mesma distribuição de probabilidade [19], $\mathbf{y} = \mathbf{q}^T \mathbf{s}$ normalmente é mais gaussiana que qualquer uma das $\mathbf{s}_i$ e menos gaussiana quando se iguala a uma das $\mathbf{s}_i$. Nesse caso, apenas um dos elementos $\mathbf{q}_i$ de $\mathbf{q}$ é diferente de zero [18].

Como, na prática, os valores de $\mathbf{q}$ são desconhecidos e sabemos, através de (5.6) e (5.8a), que

$$\mathbf{b}^T \mathbf{x} = \mathbf{q}^T \mathbf{s}, \quad (5.9)$$

podemos variar $\mathbf{b}$ e observar a distribuição de $\mathbf{b}^T \mathbf{x}$. Portanto, podemos tomar, como $\mathbf{b}$, um vetor que maximiza a não-gaussianidade de $\mathbf{b}^T \mathbf{x}$, sendo que esse vetor necessariamente corresponde a $\mathbf{q} = \mathbf{A}^T \mathbf{s}$, vetor esse que possui apenas uma de suas componentes diferente de zero. Isso significa que $\mathbf{y}$ em (5.6) é igual a uma das componentes independentes. Logo, a maximização da não-gaussianidade de $\mathbf{b}^T \mathbf{x}$ permite encontrar uma das componentes. Resta, então, apenas encontrar uma medida para gaussianidade da distribuição de $\mathbf{y}$.

5.2.2 Negentropia como Medida de Não-Gaussianidade

A negentropia é uma importante medida de não-gaussianidade. A definição de entropia [20] [19] pode ser generalizada para vetores e variáveis aleatórias contínuas, vindo a ser chamada entropia diferencial. A entropia de uma variável aleatória está relacionada à quantidade de informação que essa variável contém. A entropia será maior o quanto mais imprevisível for a variável. Tomando um vetor aleatório $\mathbf{y}$ cuja função densidade de probabilidade é $f(\mathbf{y})$, temos a entropia diferencial dada por:

$$H(\mathbf{y}) = - \int f(\mathbf{y}) \log f(\mathbf{y}). \quad (5.10)$$

Como um dos resultados fundamentais da Teoria da Informação, sabe-se que

variáveis gaussianas tem a maior entropia entre todas as variáveis aleatórias de igual variância [18] [19]. Tipicamente uma versão modificada da entropia diferencial é usada como medida de não-gaussianidade. Essa medida é chamada negentropia, definida por

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}), \quad (5.11)$$

sendo $\mathbf{y}_{gauss}$ uma variável aleatória de mesma matriz de covariância que $\mathbf{y}$. A negentropia sempre é não-negativa, tem valor igual a zero se e somente se $\mathbf{y}$ tem distribuição gaussiana e é invariante para transformações lineares inversíveis.

Em contraste às suas qualidades como medida de não-gaussianidade, a negentropia é de difícil estimação. Por isso, é necessária a utilização de aproximações usando, por exemplo, momentos de alta ordem, como:

$$J(\mathbf{y}) \approx \frac{1}{12} E \{ \mathbf{y}^3 \}^2 + \frac{1}{48} \text{kurt}(\mathbf{y})^2 \quad (5.12)$$

sendo $\text{kurt}(\mathbf{y})$, ou seja, a kurtose de $\mathbf{y}$, definida como o momento de quarta ordem da variável aleatória $\mathbf{y}$, expresso por:

$$\text{kurt}(\mathbf{y}) = E \{ \mathbf{y}^4 \} - 3(E \{ \mathbf{y}^2 \})^2. \quad (5.13)$$

A kurtose é zero para variáveis gaussianas e maior que zero para a maioria das variáveis aleatórias não-gaussianas. Tendo essa estimação é possível então desenvolver um algoritmo para estimar as componentes independentes através da maximização dessa função de custo. A estimação pode ser feita de forma adaptativa, por descida de gradiente ou diretamente através de um algoritmo de ponto fixo, o qual usaremos neste trabalho.

5.2.3 Um Algoritmo de ponto fixo para ICA

A forma mais óbvia de se resolver o problema de ICA, seria simplesmente maximizar a função de custo da equação 5.12 porém, como visto em [21] podemos melhorar o desempenho do algoritmo utilizando uma regra de aprendizado mais geral.

Porém, para permitir certas aproximações uma importante etapa no algoritmo é o processo conhecido como branqueamento, que é na verdade uma transformação linear que torna os dados descorrelacionados e com variâncias unitárias. Em outras palavras podemos dizer que esse é o mesmo processo de PCA explicado no capítulo anterior. Basicamente temos que encontrar uma transformação linear $\mathbf{V}$ que descorrelaciona os vetores das amostras $\mathbf{x}$:

$$\mathbf{z} = \mathbf{V}\mathbf{x} = \mathbf{V}\mathbf{A}\mathbf{s} \quad (5.14)$$

de modo que a matriz de correlação de $\mathbf{z}$ seja igual ao vetor unitário: $E[\mathbf{z}\mathbf{z}^T] = \mathbf{1}$. A matriz de mistura, ou matriz de funções bases $\mathbf{A}$ será diferente no fim do processo com os dados branqueados, porém, basta utilizar a inversa da transformação $\mathbf{V}$ para se obter a matriz $\mathbf{A}$ com os dados originais, não branqueados.

O próximo passo é a regra de aprendizado para as componentes independentes. Lembre que maximizando a não-gaussianidade de $\mathbf{b}^T\mathbf{x}$, chegaríamos às componentes, porém agora, com os dados branqueados $\mathbf{z}$, buscamos a combinação $\mathbf{w}^T\mathbf{z}$ que maximize essa não-gaussianidade. Como tínhamos $\mathbf{q} = \mathbf{A}^T\mathbf{x}$, agora com os dados branqueados temos $\mathbf{q} = \mathbf{V}\mathbf{A}^T\mathbf{w}$, e portanto:

$$\|\mathbf{q}\|^2 = (\mathbf{w}^T\mathbf{V}\mathbf{A}) (\mathbf{V}^T\mathbf{A}^T\mathbf{w}) = \|\mathbf{w}\|^2 \quad (5.15)$$

ou seja, a norma de $\mathbf{q}$ é a mesma de $\mathbf{w}$ nesse caso. Como assumimos a princípio que as componentes independentes tem variância unitária, o mesmo ocorre com o vetor $\mathbf{q}$, e

portanto com $\mathbf{w}$. Portanto precisamos maximizar a não gaussianidade dentro da limitação $\|\mathbf{w}\|^2 = 1$.

Para se chegar na regra de adaptação para o vetor $\mathbf{w}$ deve-se encontrar o gradiente da função de custo na equação 5.12 e depois derivar um algoritmo de ponto-fixe. Essas deduções estão detalhadas em [21] e [18]. O algoritmo de ponto fixo final, chamado de *FastICA*, baseado em negentropia, pode ser resumido nos seguintes passos:

1. Branquear o vetor de amostras $\mathbf{x}$ resultando em $\mathbf{z}$
2. Escolher um valor inicial para o vetor $\mathbf{w}$, com norma unitária
3. Fazer $\mathbf{w} \leftarrow \mathbf{E} [\mathbf{z}g(\mathbf{w}^T \mathbf{z})] - \mathbf{E} [g'(\mathbf{w}^T \mathbf{z})] \mathbf{w}$
4. Normalizar $\mathbf{w}$: $\mathbf{w} \leftarrow \mathbf{w} / \|\mathbf{w}\|$
5. Se não convergir, retornar ao passo 3

A função g é uma função não-linear, derivada de uma função G que deve ser uma aproximação da negentropia utilizada na equação 5.12. Algumas funções já foram testadas e provadas serem boas aproximações:

$$G_1(y) = \frac{1}{\alpha_1} \log \cosh \alpha_1 y \quad (5.16)$$

$$g_1(y) = \tanh \alpha_1 y \quad (5.17)$$

$$g'_1(y) = \alpha_1 (1 - \tanh^2(\alpha_1 y)) \quad (5.18)$$

$$G_2(y) = -\exp\left(\frac{-y^2}{2}\right) \quad (5.19)$$

$$g_2(y) = y \exp\left(\frac{-y^2}{2}\right) \quad (5.20)$$

$$g'_2(y) = (1 - y^2) \exp\left(\frac{-y^2}{2}\right) \quad (5.21)$$

Com isso é possível a implementação do algoritmo *FastICA*.

6 Teoria da Informação e Codificação Eficiente

6.1 Entropia

Uma das áreas de conhecimento mais importantes para o processamento digital de informação foi a Teoria da Informação. Utilizando conceitos de outras áreas, Shannon, um dos precursores da Teoria da Informação definiu bases fundamentais para o desenvolvimento do processamento e transmissão da informação digital. Um dos principais conceitos desenvolvidos por ele foi o da Entropia da informação digital. Intuitivamente pode ser entendida, assim como na termodinâmica, como nível de desordem, só que ao invés de se referir às partículas se refere à uma variável aleatória. Na prática, é uma medida do grau de imprevisibilidade da variável.

Um conceito básico para a compreensão da Entropia e sua utilidade é o conceito de **quantidade de informação**. Seja uma variável aleatória discreta, $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]$, não redundante, ou seja, a ocorrência de cada elemento não altera a probabilidade de ocorrência dos demais, cujos elementos podem ter valores dentro de um dicionário $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_l\}$, com probabilidades:

$$P(\lambda_i), \quad \text{sendo } i = 1, 2, \dots, l \quad (6.1)$$

Sendo a quantidade de informação que essa variável “emite” por cada ocorrência de um certo elemento λ_k dada por [20]:

$$I(\lambda_k) = -\log_2 P(\lambda_k) \quad (6.2)$$

Ou seja, quando mais imprevisível forem os valores da variável aleatória, mais informação

ela carrega, pela definição de Shannon.

Porém, para caracterizar a variável aleatória utilizamos o valor médio dessa medida ao longo de todas as possíveis ocorrências. A essa medida chamamos entropia, e é definida como [20]:

$$H(\Lambda_k) = \sum_{k=1}^l P(\lambda_k) I(\lambda_k) = \sum_{k=1}^l P(\lambda_k) \log_2 \frac{1}{P(\lambda_k)} \quad (6.3)$$

Note que a entropia é descrita em termos do dicionário Λ e não da variável aleatória $\mathbf{x}$. Porém dizemos que essa é uma medida que caracteriza a variável aleatória porque, na prática, essas probabilidades $P(\lambda_k)$ são decorrentes da estimação da distribuição de probabilidade de $\mathbf{x}$. A importância dessa medida é bem ampla em processamento digital. Shannon mostrou que a medida da entropia de uma variável é o limite inferior para a taxa de compressão digital. Esse importante resultado é até hoje uma referência para qualquer sistema de compressão de dados.

Porém, a entropia de Shannon é definida para uma variável aleatória não redundante. Dessa forma, caso uma variável contenha redundâncias, ou seja, o conhecimento de um elemento qualquer do vetor aleatório permita inferir com maior probabilidade o valor de outro elemento, o valor da entropia de Shannon não representa esse valor médio da quantidade de informação transmitida. A importância desse resultado é a possibilidade de se utilizar a entropia de Shannon como uma medida de redundância. Uma vez que em uma transformação sem perdas, o valor da entropia possa ser diminuído, significa que a nova representação é menos redundante.

A importância da entropia ultrapassa os limites de aplicações técnicas como o da compressão de dados. Na neurociência, por exemplo, quando se estuda a codificação da informação sensorial realizada pelo cérebro, ou seja, como o cérebro transforma

essas informações em impulsos nervosos, a entropia serve como parâmetro para avaliar a eficiência do código em termos quantidade de informação propagada pelos impulsos nervosos. A teoria da informação contribuiu para o desenvolvimento da teoria que chamamos hoje de **Codificação Eficiente**.

6.2 Codificação Eficiente

Diz-se que os córtices sensoriais se adaptam para representar a informação vinda dos órgãos sensoriais utilizando um código neural com menor entropia possível, e conseqüentemente com menor redundância. Esse conceito de codificação eficiente proposto por Barlow [6] teve profundas conseqüências no campo da neurociência. Partindo desse pressuposto vários pesquisadores começaram a buscar o código que o cérebro deveria gerar ao processar imagens [9] ou sons naturais [22]. Seus resultados foram bastante correlacionados com as medições realizadas em experimentos nos córtices auditivo [22] [23] e visual [24] [25] de animais. Essas descobertas inspiraram sistemas de processamento de dados também mais eficientes, tanto no campo de reconhecimento de padrões [26], quando no caso da compressão de dados [27].

Na prática, seja um estímulo $x(t)$ que deverá ser projetado em um espaço vetorial A de modo a ser eficientemente representado, ou seja, que a entropia da nova representação seja menor que a do próprio sinal. Quando isto é possível podemos afirmar que houve redução de redundância, visto que a informação não se perdeu, dado que se trata de uma transformação linear, e a nova representação tem menor entropia, ou menor quantidade média de informação por elemento.

Por outro lado sabe-se que dentro das variáveis aleatórias de mesma variância,

aquela com maior entropia, a mais imprevisível, é a variável gaussiana. Isso implica que variáveis com menor entropia têm uma distribuição que se distancia de uma distribuição gaussiana. Comumente esse distanciamento tende à uma distribuição esparsa, na qual os valores se concentram em torno de um único ponto, com raros, esparsos, eventos distantes dessa média. No diagrama da Figura 6.1 vemos um exemplo de uma onda sonora $x(t)$ sendo projetada em dois espaços vetoriais distintos, o espaço de Fourier, onde as funções bases são senos e cossenos, e um espaço de *wavelets* de Gabor, que segundo [22] são as funções que codificam os sons no córtex auditivo. Uma codificação eficiente torna a representação mais esparsa, com mostra a distribuição de probabilidade $P(w_1)$ de um dos vetores de coeficientes resultantes da transformação. Enquanto que no caso de Fourier a distribuição de probabilidade da representação se assemelha mais à uma gaussiana, de onde podemos afirmar que tem maior entropia.

É importante notar que no desenvolvimento do algoritmo de ICA, dizemos que a maximização da não-gaussianidade das componentes independentes implica na independência estatística das mesmas, e conseqüentemente reduz a entropia bem como a redundância entre as componentes. Com isso podemos afirmar que ICA é um método adequado para se realizar o processo de codificação eficiente. Assim, se definimos $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_m]$ o conjunto m de amostras de uma mesmo sinal, cada uma com tamanho n , e utilizarmos o mesmo modelo de ICA, na equação 5.3, podemos estimar o espaço $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ que codifica eficientemente o conjunto $\mathbf{x}$. Sendo os coeficientes $\mathbf{s}$ a representação eficiente do conjunto $\mathbf{x}$.

Observe que não necessariamente, o espaço $\mathbf{A}$ gerado a partir de $\mathbf{x}$ somente codifica eficientemente o próprio conjunto. Podemos extrapolar essa idéia e dizer que $\mathbf{A}$ transforma qualquer variável aleatória que pertença à mesma “categoria” do conjunto $\mathbf{x}$,

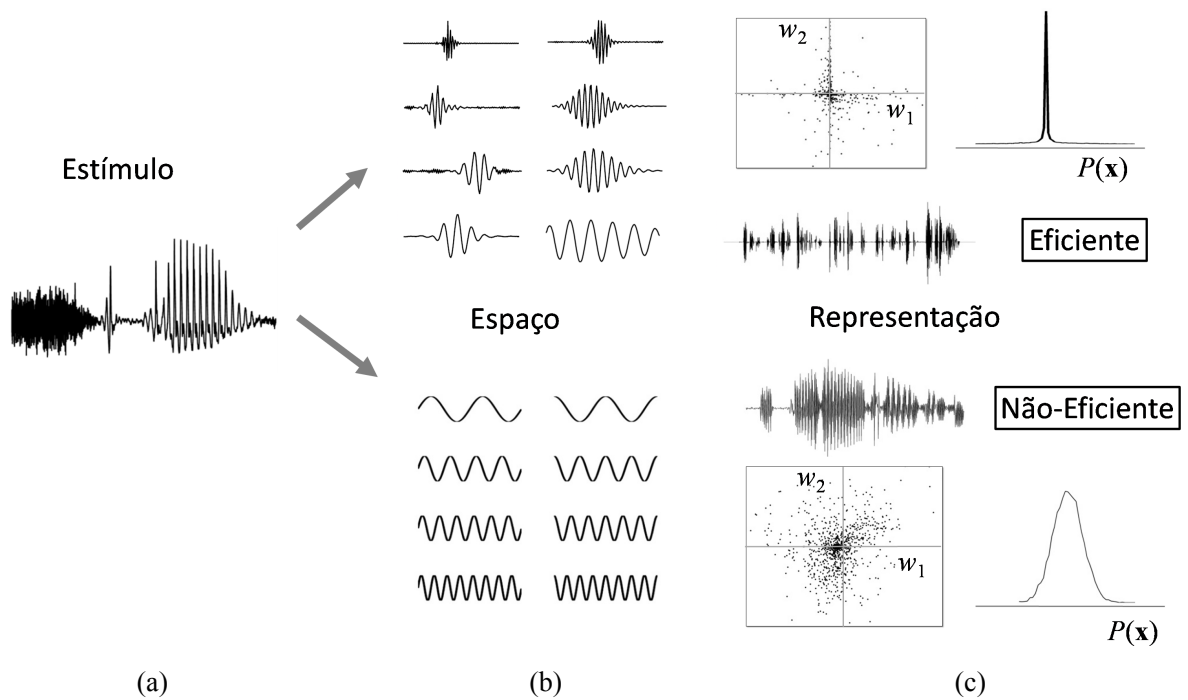


Figura 6.1: Esquema comparativo da codificação de um sinal sonoro (a) em duas bases diferentes (b), no caso uma de *wavelets* (acima) e Fourier (abaixo). O estímulo é representado pelos coeficientes da combinação linear das funções bases. Utilizamos as distribuições de probabilidade dos coeficientes w_n resultantes da representação para caracterizar a eficiência; quanto mais esparsa o código gerado, mais eficiente a codificação, pois menor é a entropia da representação.

ou seja, que possua as mesmas propriedades. Por exemplo, como em [8] podemos utilizar uma base de dados de imagens naturais para estimar as funções bases que serão gerais para codificar eficientemente qualquer outra imagem natural, mesmo fora do conjunto original, desde que possua as mesmas propriedades estatísticas e espaciais que a caracterizem como imagens naturais. Portanto basta aplicar a transformação linear inversa de $\mathbf{A}$ para encontrarmos os coeficientes $\mathbf{w} = [w_1, \dots, w_n]$ que formam a representação de uma variável $\mathbf{x}_j$ qualquer:

$$\mathbf{w}^T = \mathbf{A}^{-1} \mathbf{x}_j^T \quad (6.4)$$

Podemos confirmar a eficiência de ICA para gerar uma representação eficiente utilizando a informação mútua, uma outra medida estatística que permite avaliar a redundância entre duas variáveis aleatórias. Seja a informação mútua entre as variáveis aleatórias $[\mathbf{x}_1, \dots, \mathbf{x}_m]$ definida por:

$$I(\mathbf{x}_1, \dots, \mathbf{x}_m) = \sum_{i=1}^m H(\mathbf{x}_i) - H(\mathbf{x}_1, \dots, \mathbf{x}_m), \quad (6.5)$$

onde $H(\mathbf{x}_1, \dots, \mathbf{x}_m)$ é a entropia conjunta de $\mathbf{x}_1, \dots, \mathbf{x}_m$. Como a informação não pode ser perdida, dizemos que:

$$I(\mathbf{x}_1, \dots, \mathbf{x}_m) \geq 0. \quad (6.6)$$

Substituindo (6.6) em (6.5) resulta em:

$$\sum_{i=1}^m H(\mathbf{x}_i) \geq H(\mathbf{x}_1, \dots, \mathbf{x}_m). \quad (6.7)$$

Da mesma maneira, seja a informação mútua das variáveis aleatórias $\mathbf{w}_1, \dots, \mathbf{w}_m$ be defined as

$$I(\mathbf{w}_1, \dots, \mathbf{w}_m) = \sum_{i=1}^m H(\mathbf{w}_i) - H(\mathbf{w}_1, \dots, \mathbf{w}_m), \quad (6.8)$$

onde $H(\mathbf{w}_1, \dots, \mathbf{w}_m)$ é a entropia conjunta de $\mathbf{w}_1, \dots, \mathbf{w}_m$. Se assumimos que $\mathbf{w}_1, \dots, \mathbf{w}_m$ são independentes [8], podemos dizer que [21]:

$$I(\mathbf{w}_1, \dots, \mathbf{w}_m) = 0. \quad (6.9)$$

Então substituindo (6.9) em (6.8) temos:

$$\sum_{i=1}^m H(\mathbf{w}_i) = H(\mathbf{w}_1, \dots, \mathbf{w}_m). \quad (6.10)$$

Dada a transformação linear:

$$\mathbf{x}_i = \mathbf{A}\mathbf{w}_i, \quad (6.11)$$

temos que [19]

$$H(\mathbf{x}_1, \dots, \mathbf{x}_m) = H(\mathbf{w}_1, \dots, \mathbf{w}_m) \quad (6.12)$$

Então, de (6.7), (6.10) e (6.12) obtemos:

$$\sum_{i=1}^m H(\mathbf{x}_i) \geq H(\mathbf{w}_1, \dots, \mathbf{w}_m) \quad (6.13a)$$

$$\implies \sum_{i=1}^m H(\mathbf{x}_i) \geq \sum_{i=1}^m H(\mathbf{w}_i). \quad (6.13b)$$

A equação 6.13b mostra que a entropia conjunta da representação eficiente gerada através de ICA, ou seja, supondo independência estatística nos coeficientes, será sempre menor que a dos dados originais, A não ser que não haja redundância em $\mathbf{x}$ sendo sua representação eficiente a priori. Esses resultados foram descritos em [28], e são a base para se utilizar a codificação eficiente como uma suporte geral para codificação de vários tipos de sinais, inclusive ECG.

7 Material e Métodos

Neste trabalho, propôs-se analisar o resultado da codificação eficiente de sinais de ECG através do uso de ICA. Foi também comparado o desempenho desta proposta com os métodos tipicamente utilizados para codificação desses sinais, que são PCA e a transformada de *Wavelets*. Um esquema gráfico da codificação eficiente pode ser visto na figura 7.1.

Para isso foram utilizados sinais de ECG disponíveis em uma base de dados gratuita disponibilizada pela *Physionet* [29]. Foram utilizados tanto sinais de pessoas saudáveis quanto de pessoas com algumas patologias que alteram o ECG normal. Um problema é que os sinais têm características diferentes na sua aquisição como frequência de amostragem e algum tipo de filtragem como pre-processamento. Com essas diferenças, para se fazer comparações mais justas, é necessário colocar os sinais na mesma situação em termos de suas propriedades. Para se aproximar, por exemplo, a frequência de amostragem, foi feito um processo de subamostragem nos sinais que apresentam frequência de amostragem mais alta. Na tabela 7.1, são mostradas as informações gerais de cada grupo de sinal, e as modificações realizadas.

A primeira etapa é a estimação do espaço para codificação eficiente. Foi utilizado o *FastICA* sobre cada um dos sinais das bases de dados. Cada sinal foi dividido em 4.000 blocos, tomados aleatoriamente, de 100 pontos cada, para servir de entrada para o algoritmo. Foram geradas conseqüentemente 100 funções bases para cada sinal. O segundo passo é utilizar essas funções bases como base vetorial para a projeção linear dos

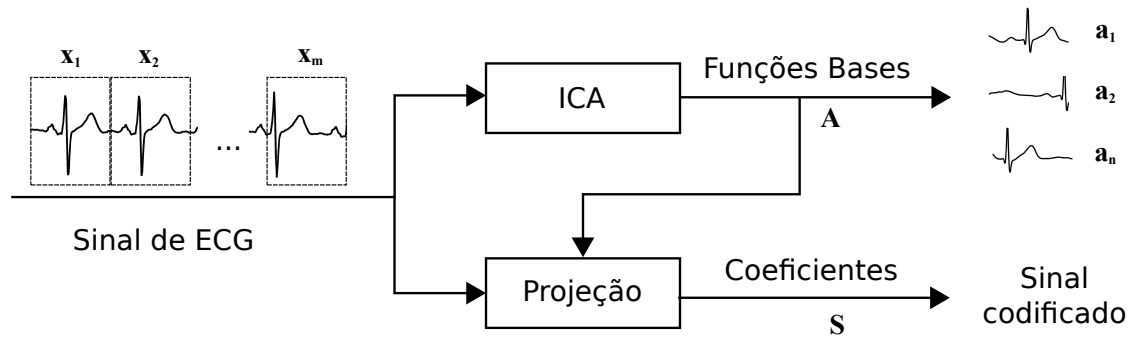


Figura 7.1: O processo de codificação é feito em duas partes, primeiro alguns trechos do sinal são utilizados para encontrar os vetores bases e em seguida utiliza-se essa base para projetar qualquer outro bloco desse mesmo sinal.

Tabela 7.1: Informações sobre a aquisição dos sinais de ECG das bases de dados utilizadas.

Base de dados	Amostragem	Quantização	Subamostragem
MIT-BIH Ritmo Normal	128 Hz	12 bits	-
MIT-BIH Fibrilação Atrial	250 Hz	12 bits	2
Morte Súbita - <i>Holter</i>	250 Hz	12 bits	2

dados e obtenção dos coeficientes. Nessa etapa todo o sinal é projetado, bloco a bloco, sequencialmente, sem sobreposição dos blocos consecutivos. Dessa forma, temos a mesma quantidade de pontos originais do sinal e coeficientes para representá-lo.

A segunda etapa é a estimação da entropia de cada um desses sinais, bem como a comparação com a entropia da codificação utilizando os outros dois métodos, PCA e a DWT. No caso de PCA a codificação é realizada da mesma forma que ICA: estima-se as funções bases depois projeta-se o sinal nas mesmas. No caso de *Wavelets* foi utilizada a DWT com 4 níveis de decomposição e utilizando as *wavelets symlets*, especificamente, *sym2* que são *wavelets* utilizadas para codificação de ECG em outros trabalhos com bons

resultados [30].

Para realizar o cálculo da entropia foi utilizada equação 6.3. Porém, é necessária a estimação da distribuição de probabilidade das variáveis aleatórias, no caso, os coeficientes resultantes da codificação dos sinais. Existem diversas formas para se realizar essa estimação. A forma mais direta é pela divisão da escala do sinal em uma quantidade fixa de intervalos, ou seja, realizar uma quantização linear do sinal, e simplesmente contar a quantidade de valores dentro dessa faixa. Cada faixa, portanto, representaria um valor dentro do dicionário de possibilidades, como definido no capítulo 6. Apesar de não ser o método mais eficiente, é muito usado pois quando se trata com uma quantidade grande de valores, que é o nosso caso com sinais de ECG de longa duração, todas as estimações tendem para a mesma distribuição. [19].

8 Resultados

8.1 Estimação das Funções Bases

Utilizando ICA, foram estimadas 100 funções bases para cada paciente. Nas Figuras de 8.1 a 8.6 podemos ver 6 funções bases selecionadas do grupo de 100, bem como alguns trechos dos sinais originais para comparação. A primeira característica observada nas funções bases para todos os casos é a semelhança das mesmas com relação ao sinal. Pode-se ver claramente que as funções bases se assemelham a batimentos cardíacos com a mesma estrutura dos trechos do sinal. Outra característica fundamental é fato de que as funções bases ocupam todas as localizações temporais, ou seja, em cada uma das funções bases vemos um batimento cujo complexo QRS ocupa uma posição diferente na janela.

Vale ressaltar resultados interessantes de alguns casos particulares. Nos casos, por exemplo, de pacientes que sofreram morte súbita cardíaca, os sinais de ECG contêm diversos trechos com sérias arritmias onde há uma grande alteração na forma do sinal,

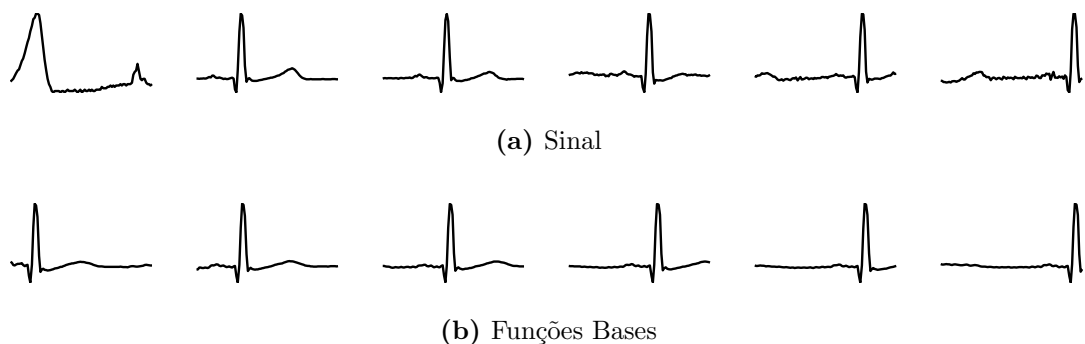


Figura 8.1: Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente saudável e (b) algumas funções bases geradas pela codificação eficiente do sinal

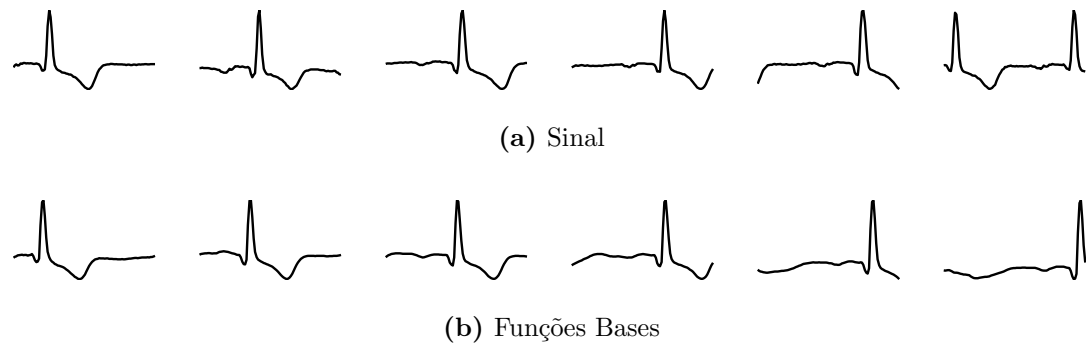


Figura 8.2: Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente saudável e (b) algumas funções bases geradas pela codificação eficiente do sinal

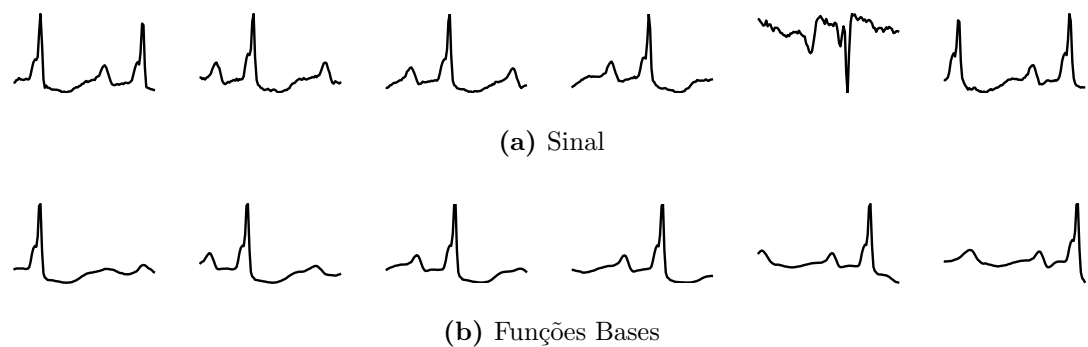


Figura 8.3: Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente com fibrilação atrial e (b) algumas funções bases geradas pela codificação eficiente do sinal

como podemos ver nas figuras 2.5 e 2.6, mas ainda assim, todo o grupo de funções bases tendem a uma forma que lembra o próprio batimento cardíaco 8.5 e 8.6. Já no caso da fibrilação atrial, em que as principais alterações no ECG ocorrem na onda P, sem alterar o restante do sinal, notou-se consistentemente nas funções bases uma redução ou desaparecimento da onda P. Outro resultado interessante é o fato de que as funções bases aparecem como formas de batimento praticamente livre dos ruídos percebidos nos sinais originais.

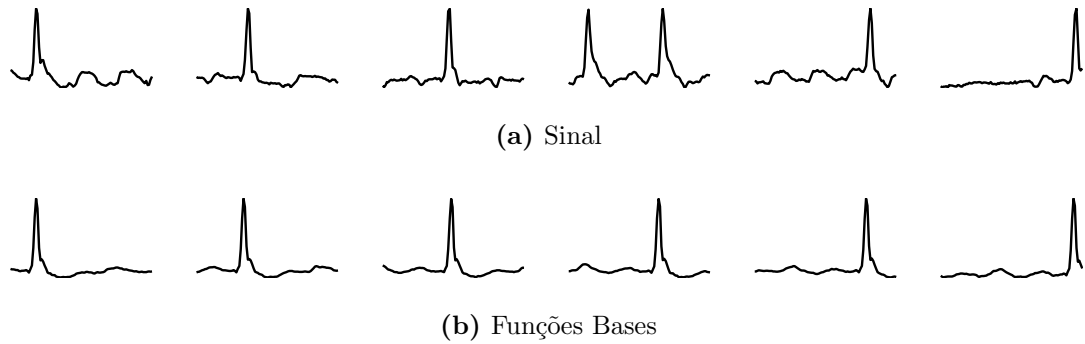


Figura 8.4: Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente com fibrilação atrial e (b) algumas funções bases geradas pela codificação eficiente do sinal

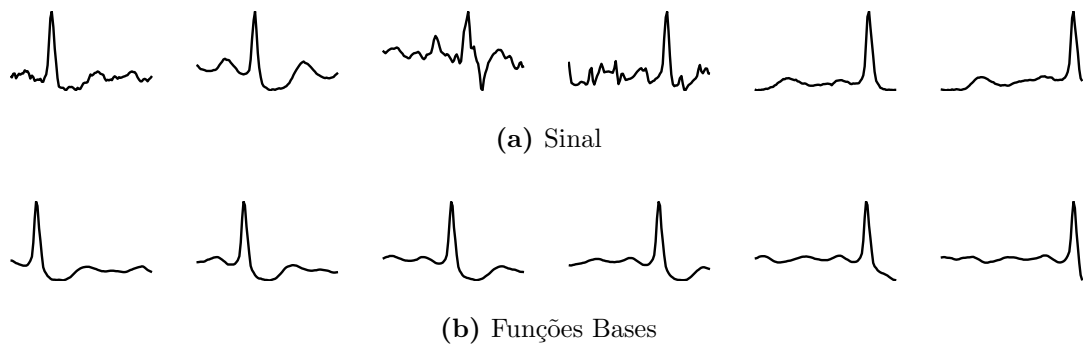


Figura 8.5: Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente acometido por morte súbita cardíaca e (b) algumas funções bases geradas pela codificação eficiente do sinal

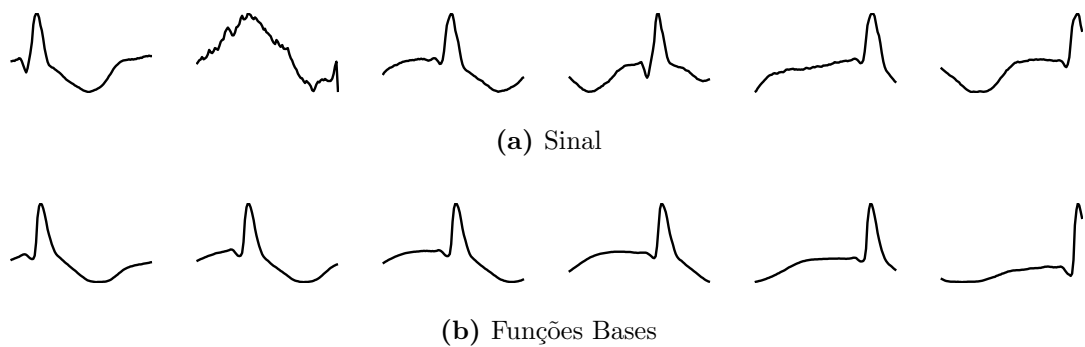


Figura 8.6: Exemplos de (a) trechos aleatórios tomados de um sinal de ECG de um paciente acometido por morte súbita cardíaca e (b) algumas funções bases geradas pela codificação eficiente do sinal

8.2 Análise da Codificação

A segunda parte dos resultados consiste na análise da eficiência da codificação em termos da entropia. Nas figuras 8.7, 8.8 e 8.9, está mostrada a comparação das entropias médias dos três métodos testados, ICA, PCA e *Wavelets*, além da entropia do sinal não codificado. Pode-se ver claramente que todos os métodos testados são capazes de reduzir a entropia, porém o desempenho dos métodos varia com a base de dados.

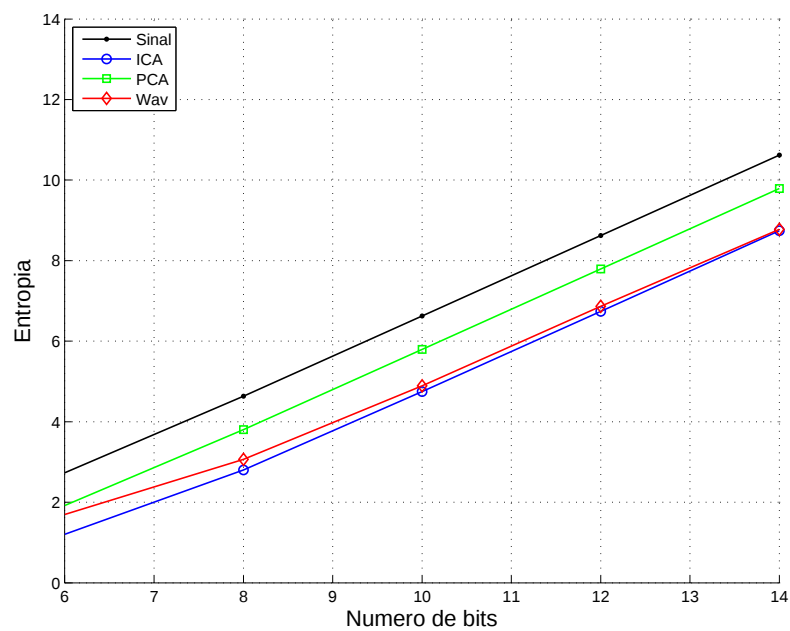


Figura 8.7: Entropia média da codificação de todos os sinais do grupo de pessoas saudáveis para todos os métodos testados PCA, ICA e *Wavelets* bem como do sinal original sem ser codificado

Tanto para os sinais normais, figura 8.7 quanto para os sinais de morte súbita, 8.9 o melhor desempenho foi do método de ICA, seguido por *Wavelets* e PCA, por outro lado no caso da fibrilação atrial a transformada de *Wavelets* obteve um melhor desempenho. Pode-se observar também que a diferença entre os desempenhos se mantém aparentemente constante ao longo dos vários níveis de quantização. Na tabela 8.1, onde está apresentado os valores exatos das entropias e das variâncias da codificação para cada

base de dados. Nota-se que qualquer das codificações utilizadas reduz a entropia com relação ao sinal não codificado porém para as bases de dados Normal e Fibrilação Atrial, ICA alcança menor entropia, tornando o resultado mais estatisticamente significativo.

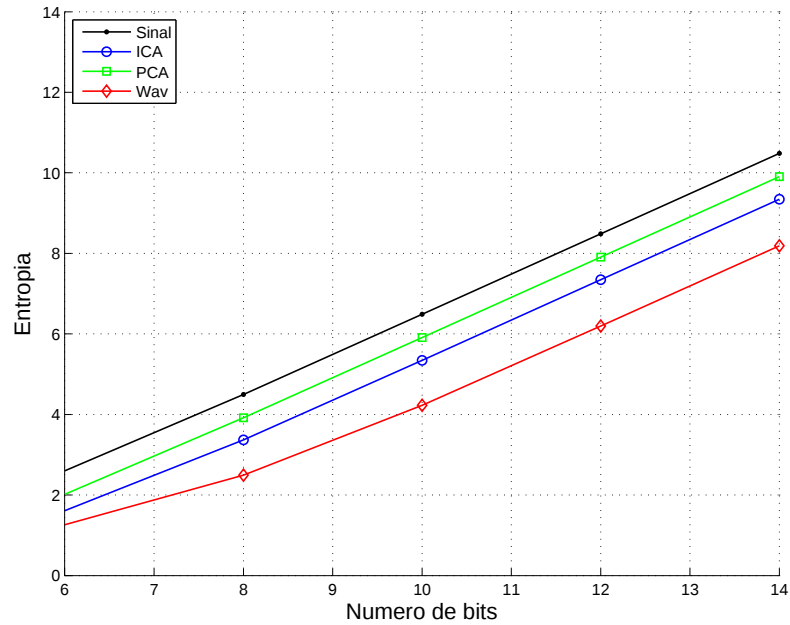


Figura 8.8: Entropia média da codificação de todos os sinais do grupo de pacientes com fibrilação atrial para todos os métodos testados PCA, ICA e *Wavelets* bem como do sinal original sem ser codificado

Nas figuras de 8.10 à 8.12 e nas tabelas de 8.2 à 8.4, é mostrado o erro para cada base de dados para os vários níveis de quantização. Nota-se em todos os casos que os erros gerados pela quantização dos coeficientes de qualquer uma das transformações são maiores que o erro da simples quantização do sinal original. Comparando as transformações, para o grupos normal e morte súbita, percebe-se que o erro para a DWT é o menor, seguido por ICA, porém com um desvio padrão suficientemente grande para ultrapassar as médias das demais representações. Por outro lado, no caso de fibrilação atrial, o erro de ICA é menor e se aproxima do erro de quantização do sinal original, enquanto que o erro de wavelets se torna o maior, e ainda com maior desvio padrão.

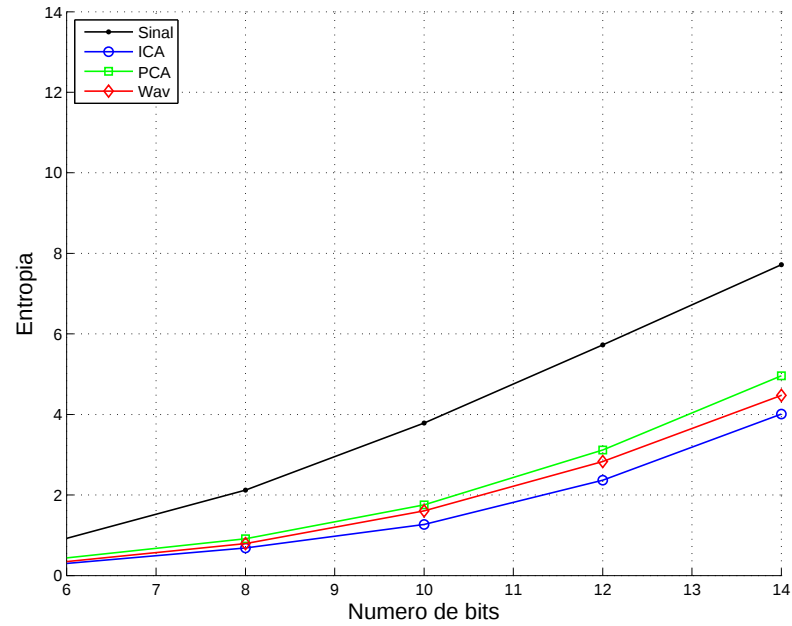


Figura 8.9: Entropia média da codificação de todos os sinais do grupo de pacientes que sofreram morte súbita cardíaca para todos os métodos testados PCA, ICA e *Wavelets* bem como do sinal original sem ser codificado

Tabela 8.1: Comparação da Entropia da codificação para os três métodos testados *Wavelets* PCA e ICA, sobre as três bases de dados, normal, fibrilação atrial e morte súbita utilizando todos os sinais. Para o cálculo da distribuição de probabilidade foi utilizada quantização uniforme de 8 bits.

ENTROPIA				
Base de dados	Não codificado	PCA	<i>Wavelets</i>	ICA
Ritmo Normal	$4,63 \pm 0,42$	$3,80 \pm 0,52$	$3,06 \pm 0,42$	$2,80 \pm 0,23$
Fibrilação Atrial	$4,50 \pm 0,56$	$3,92 \pm 0,37$	$2,49 \pm 0,44$	$3,37 \pm 0,25$
Morte Súbita	$2,11 \pm 1,91$	$0,90 \pm 2,21$	$0,79 \pm 0,53$	$0,68 \pm 1,45$

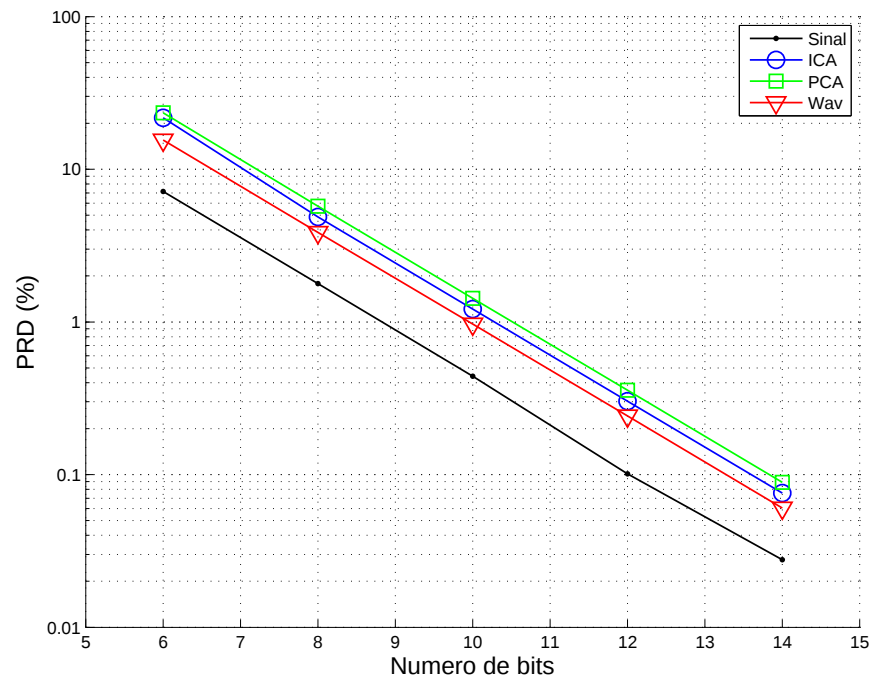


Figura 8.10: Erro médio resultante da quantização dos coeficientes das várias representações para os sinais do grupo de pacientes suadáveis

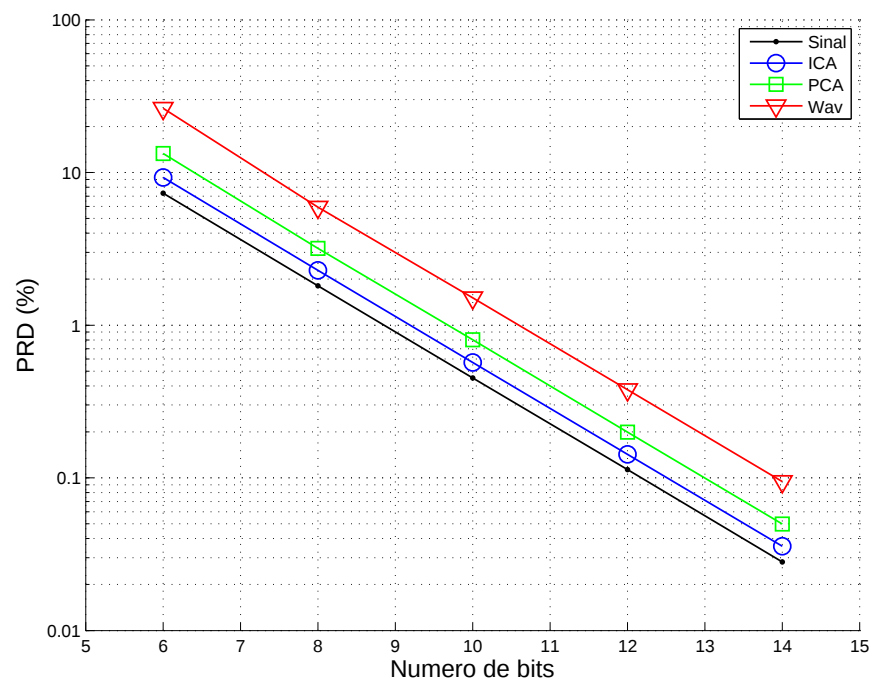


Figura 8.11: Erro médio resultante da quantização dos coeficientes das várias representações para os sinais dos pacientes com fibrilação atrial

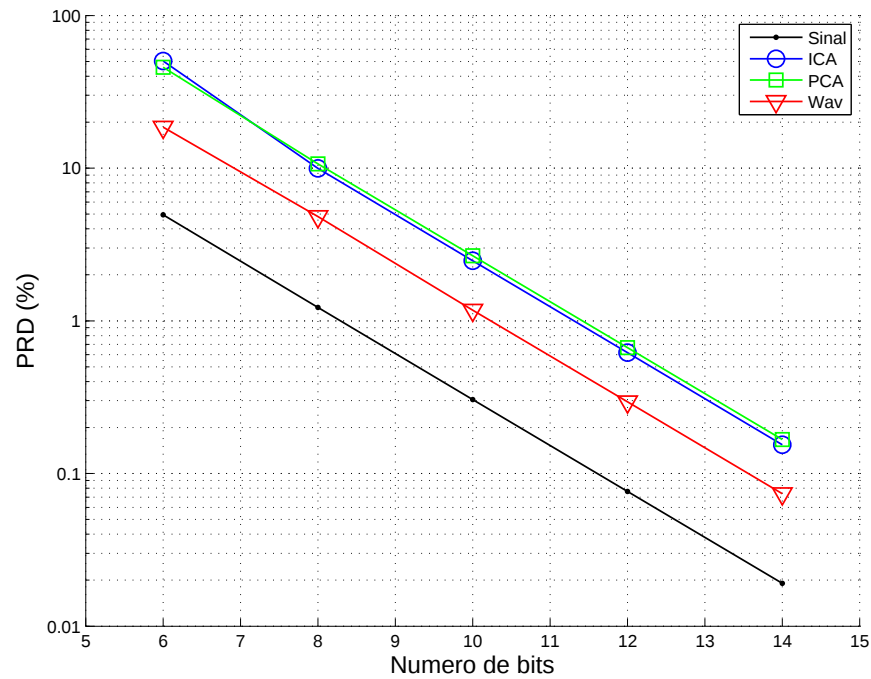


Figura 8.12: Erro médio resultante da quantização dos coeficientes das várias representações para os sinais do grupo de pacientes que sofreu morte súbita cardíaca

Tabela 8.2: Comparação do erro gerado pela quantização dos coeficientes para vários valores de quantização, para cada método, sobre os sinais de indivíduos saudáveis.

ERRO DE QUANTIZAÇÃO				
Quantização	Não codificado	PCA	<i>Wavelets</i>	ICA
6 bits	7,35 ± 3,26	23,17 ± 13,32	16,74 ± 13,11	21,45 ± 12,45
8 bits	1,84 ± 0,82	5,73 ± 3,32	4,31 ± 3,72	5,02 ± 2,56
10 bits	0,45 ± 0,20	1,43 ± 0,83	1,04 ± 0,84	1,25 ± 0,64
12 bits	0,11 ± 0,05	0,36 ± 0,21	0,26 ± 0,21	0,31 ± 0,16
14 bits	0,03 ± 0,01	0,09 ± 0,05	0,06 ± 0,05	0,08 ± 0,04

Tabela 8.3: Comparação do erro gerado pela quantização dos coeficientes para vários valores de quantização, para cada método, para os sinais de pacientes com fibrilação atrial

ERRO DE QUANTIZAÇÃO				
Quantização	Não codificado	PCA	<i>Wavelets</i>	ICA
6 bits	$7,79 \pm 3,26$	$14,39 \pm 13,32$	$25,39 \pm 13,11$	$11,23 \pm 12,45$
8 bits	$1,93 \pm 0,82$	$3,59 \pm 3,32$	$6,81 \pm 3,72$	$2,64 \pm 2,56$
10 bits	$0,47 \pm 0,20$	$0,90 \pm 0,83$	$1,61 \pm 0,84$	$0,66 \pm 0,64$
12 bits	$0,12 \pm 0,05$	$0,22 \pm 0,21$	$0,42 \pm 0,21$	$0,16 \pm 0,16$
14 bits	$0,03 \pm 0,01$	$0,06 \pm 0,05$	$0,10 \pm 0,05$	$0,04 \pm 0,04$

Tabela 8.4: Comparação do erro gerado pela quantização dos coeficientes para vários valores de quantização, para cada método, para os sinais de morte súbita

ERRO DE QUANTIZAÇÃO				
Quantização	Não codificado	PCA	<i>Wavelets</i>	ICA
6 bits	$4,95 \pm 4,51$	$45,84 \pm 19,30$	$18,53 \pm 21,39$	$50,32 \pm 16,68$
8 bits	$1,22 \pm 1,11$	$10,65 \pm 4,46$	$4,82 \pm 4,31$	$9,98 \pm 3,37$
10 bits	$0,30 \pm 0,28$	$2,67 \pm 1,11$	$1,17 \pm 1,06$	$2,47 \pm 0,83$
12 bits	$0,08 \pm 0,07$	$0,67 \pm 0,28$	$0,29 \pm 0,26$	$0,62 \pm 0,21$
14 bits	$0,02 \pm 0,02$	$0,17 \pm 0,07$	$0,07 \pm 0,07$	$0,15 \pm 0,05$

9 Discussão

9.1 Funções bases

A primeira observação interessante da codificação eficiente dos sinais de ECG é que as funções bases tendem a assumir a forma de um batimento cardíaco completo, principalmente no caso dos pacientes normais, como podemos verificar nas figuras de 8.1 a 8.6. Sabe-se que a codificação eficiente de outros tipos de sinais como sons e imagens naturais, apesar de suas estruturas complexas, tende a decompor o sinal em funções bases com uma localização específica em tempo e em frequência, que são normalmente modelados como filtros passa banda e *wavelets* de Gabor, [22] [9] [24]. Uma razão para essa diferença é o fato de que diferente dos sinais naturais, o ECG combina os vários fenômenos que geram o sinal, sempre da mesma forma, não havendo necessidade de que sejam separados para gerar um código eficiente. Intuitivamente, podemos entender da seguinte forma: para um dado trecho contendo um batimento, basta que tenhamos uma única função base e conseguiremos representar o sinal com um único coeficiente positivo e o restante poderá ser zero. Dessa forma através de um único número representamos um trecho completamente. Por outro lado, se dividíssemos as componentes de um batimento em várias partes teríamos que utilizar todas elas para representar cada batimento. Dessa forma podemos supor que ICA gera uma base vetorial ótima para codificação de ECG e que é, de certa forma diferente da transformada de *Wavelets*, a qual divide o sinal em bandas de frequência pode resumir a eficiência do código

Por outro lado, essa constância do sinal de ECG nem sempre se mantém. Em algumas arritmias cardíacas sabemos que a forma do sinal de ECG pode ser completamente alterada, especialmente em casos como a fibrilação atrial. O ECG de pacientes acometidos por essa doença cardíaca apresenta grandes variações que podem ser momentâneas ou não, na forma da onda P, desde sua ausência até o surgimento de ondas P descoordenadas do ritmo cardíaco normal. Isso altera o resultado das funções bases, como vemos, por exemplo, na figura 8.4, em que o sinal apresenta, em alguns momentos, episódios de *flutter* atrial, os quais são quase imperceptíveis nas funções bases, de modo que as funções bases mostram batimentos sem a onda P, o que serviria como indicativo da doença.

De qualquer maneira, o fato de as funções bases se comportarem dessa forma, de acordo com os sinais, sugere a possibilidade de se utilizá-las como ferramentas de diagnóstico de doenças cardíacas. Sendo as funções bases um registro compacto da atividade cardíaca de um dia inteiro essas funções poderiam substituir, em alguns casos, a análise no sinal completo, que normalmente não é feita, ou quando necessária, leva muito tempo. Além disso, sabe-se que sistemas de auxílio diagnóstico utilizando codificação eficiente de sinais biomédicos e imagens médicas tem tido resultados positivos [26].

9.2 Codificação

Quanto à análise da codificação, vemos que o método proposto, apesar de teoricamente buscar o código de menor entropia possível, não obteve o resultado esperado em todas as situações. Apesar de ter superado o desempenho da transformada de *Wavelets* para os sinais de pacientes normais e de morte súbita, teve um desempenho inferior no

caso da fibrilação atrial. Para buscar as causas disso, devemos analisar primeiramente as bases de dados utilizadas.

Essas bases de dados foram escolhidas dentre as bases de dados de longa duração disponibilizadas pela *Physionet* [29]. A primeira dificuldade encontrada na hora de escolher as bases de dados foi o fato de que cada um tem uma origem diferente, ou seja, passou por um processo de aquisição e pre-processamento diferenciados, tornando complicado uma comparação adequada. A base normal, por exemplo, teve sua aquisição feita em 128 Hz, enquanto que as duas em 250 Hz, como mostrado no capítulo 7. Além disso, enquanto que as bases normal e morte súbita não passam por nenhum pre-processamento, até onde é informado pela *Physionet*, os sinais da base de fibrilação atrial passam por um filtro passa-baixas de 40 Hz, ou seja, mesmo tendo feito subamostragem para 125 Hz, ainda existe um “buraco” entre 40 e 62,5 Hz não preenchidos, o que pode comprometer a estimação das funções bases. Além disso, a base de pacientes com ritmo normal apresenta uma menor quantidade de ruídos que as demais, e sabe-se que ruídos podem prejudicar a estimação também, dado que uma das restrições de ICA é a presença de ruído gaussiano na entrada, como mostrado em [18].

Com relação ao erro gerado pela quantização dos coeficientes, nota-se uma relação entre a diminuição da entropia e o aumento do erro, independente dos níveis de quantização. Para os pacientes saudáveis pode-se afirmar que os erros estão próximos se consideramos que os desvios padrões ultrapassam os valores médios dos erros das demais representações. Por outro lado no caso da fibrilação atrial no qual a entropia da representação da DWT foi menor que as demais, o erro associado foi o maior, e ainda com um alto desvio padrão, sendo que para ICA e PCA o desvio padrão foi menor. Já no caso dos sinais de morte súbita, tanto para a entropia quanto para o erro de quantização

os desvios padrões de todas as representações são elevados. É importante notar também que as representações de ICA e PCA, resultaram, em todos os casos, em um desvio padrão menor que o da DWT, tanto para a entropia quanto para o erro. Esse é um resultado esperado, já que esses métodos se adaptam a estatística do sinal representado, de modo que podemos inferir que provavelmente para outras bases de dados de ECG os resultados serão parecidos.

De qualquer forma, a capacidade de se reduzir a entropia a esse ponto pode tornar viável a utilização desse método como uma ferramenta de compressão. De fato, alguns resultados preliminares com compressão de ECG utilizando ICA já foram obtidos anteriormente, [28] porém não foram comparados com outros métodos, nem sujeitos à uma análise de erro de quantização. Na prática existe ainda uma dificuldade para o uso de ICA em compressão, o fato de que para cada sinal é necessária a obtenção de uma nova base vetorial, que deve ser enviada pelo dispositivo que comprime o sinal, para ser decodificado pelo receptor. Uma possibilidade para superar essa dificuldade é através da redução do espaço gerado. Outro trabalho, utilizando ICA para compressão de imagens, obteve bons resultados reduzindo a quantidade de funções bases do espaço. No caso de ECG isso parece ainda mais plausível uma vez que mesmo em sinais com arritmias as funções bases tendem a assumir um único formato.

10 Conclusão

Os resultados apresentados neste trabalho confirmam o potencial da codificação eficiente para análise e codificação de sinais de ECG. A possibilidade de se extrair informações de sinais de longa duração e resumi-los em um espaço vetorial de características é promissora em termos de auxílio diagnóstico. Contudo é fundamental uma opinião médica na avaliação de tais características, bem como a utilização conjunta dessas informações com interpretações eletrofisiológicas que possam corroborar com essa possibilidade.

Apesar de alguns problemas com a codificação de ECG utilizando ICA, sabemos que o *FastICA*, não é o melhor nem o mais geral método para codificação eficiente. Existem outros métodos mais robustos e mais gerais que ainda podem ser testados na codificação eficiente dos sinais de ECG e outros sinais biomédicos como, *Complexity Pursuit* proposto em [31], que é uma generalização do ICA estudado aqui, ou o *Infomax*, proposto em [32] que utiliza maximização da informação mútua para codificar os dados, que já obtiveram também resultados promissores na codificação eficiente de sinais naturais.

Além disso, se faz necessário validar a capacidade de compressão digital de dados do método proposto através de testes mais rigorosos utilizando ferramentas reais de compressão e codificação de digital, como a codificação de Huffman [33], além de outros algoritmos normalmente utilizados para realizar a codificação digital dos coeficientes de *Wavelets* como o SPIHT (*Set partitioning in hierarchical trees* - Particionamento de

conjuntos em árvores de hierarquia) [4] [5].

Por fim se faz necessário também, seja objetivando diagnóstico ou compressão, realizar testes com bases de dados maiores e mais bem documentadas. Para que as comparações sejam mais confiáveis, é ideal que todos os sinais testados, incluindo normais e contendo arritmias, tenham passado pelos mesmos pre-processamentos e mesmo tipo de aquisição. Além disso, testes mostrando o efeito de cada pre-processamento também são importantes.

Referências

- [1] Francisco Castells, Pablo Laguna, Leif Sörnmo, Andreas Bollmann, and José Millet Roig. Principal component analysis in ecg signal processing. *EURASIP J. Appl. Signal Process.*, 2007(1):98–98, 2007.
- [2] Paul S. Addison. Wavelet transforms and the ecg: a review. *Physiol. Meas.*, 26:R155–R199, 2005.
- [3] Leontios J. Hadjileontiadis Robert S. H. Istepanian and Stavros M. Ecg data compression using wavelets and higher order statistics methods. *IEEE Trans. Information Technology In Biomedicine*, 5(2), June 2001.
- [4] Morteza Moazami-Goudarzi Iman Saboori Mohammad Pooyan, Ali Taheri. Wavelet compression of ecg signals using spiht algorithm. *International Journal of Signal Processing* 1;4, 2005.
- [5] Dong Youn Kim Zhitao Lu and William A. Pearlman. Wavelet compression of ecg signals by the set partitioning in hierarchical trees algorithm. *IEEE Transactions on biomedical Engineering*, 47(7), July 2000.
- [6] Horace Barlow. Possible principles underlying transformation of sensory messages. *Sensory Communication*, 1961.
- [7] Anthony J. Bell and Terrence J. Sejnowski. Edges are the independent components of natural scenes. In Michael C. Mozer, Michael I. Jordan, and Thomas Petsche, editors, *Advances in Neural Information Processing Systems*, volume 9, pages 831–837. The MIT Press, 1997.
- [8] Bruno A. Olshausen. Principles of image representation in visual cortex. *The Visual Neurosciences*, pages 1603–1615, 2003.
- [9] Bruno A. Olshausen and David J. Field. Vision and the coding of natural images. *American Scientist*, 88, 2000.
- [10] Michael H. Crawford. *Current Diagnosis & Treatment in Cardiology*. McGraw-Hill/Appleton & Lange, 2nd edition, 2002.
- [11] George A. Mensah Wendy A. Wattigney and Janet B. Croft. Increased atrial fibrillation mortality: United states, 1980-1998. *American Journal of Epidemiology*, 155(9):819–826, 2002.
- [12] F. D. Richard Hobbs Sue Jowett Ellen T. Murray Roger Holder Michael Davies Jonathan Mant, David A. Fitzmaurice and Gregory Y. H. Lip. Accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: analysis of data from screening for atrial fibrillation in the elderly (safe) trial. *British Medical Journal*, pages 335–380, 2007.

- [13] American Heart Association. Sudden cardiac death. Internet, March 2010.
- [14] J. Thomas Bigger Alfred E. Buxton Robert M. Califf Mark Carlson Anne Curtis Jephtha Curtis Eric Fain Bernard J. Gersh Michael R. Gold Ali Haghighi-Mood Stephen C. Hammill Jeff Healey Mark Hlatky Stefan Hohnloser Raymond J. Kim Kerry Lee Daniel Mark Marcus Mianulli Brent Mitchell Eric N. Prystowsky Joseph Smith David Steinhaus Sana M. Al-Khatib, Gillian D. Sanders and Wojciech Zareba. Preventing tomorrows sudden cardiac death today: Part i: Current data on risk stratification for sudden cardiac death. *American Heart Journal*, 2007.
- [15] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, Elsevier, 2nd edition, 1999.
- [16] Michael L. Hilton. Wavelet and wavelet packet compression of electrocardiograms. *IEEE Trans. Biomed. Eng*, 44(5):394–402, May 1997.
- [17] Fayyaz ul Amir Afsar Minhas and Muhammad Arif. Robust electrocardiogram ecg beat classification using discrete wavelet transform. *Physiological Measurement*, April 2008.
- [18] J.; Hyvärinen, A.; Karhunen and E. Oja. *Independent Component Analysis*. John Wiley & Sons, Inc., 1st edition, 2001.
- [19] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 4th edition, 2002.
- [20] A. J.; Viterbi and J. K. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill, Inc., 1st edition, 1979.
- [21] Aapo Hyvärinen and Erkki Oja. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7):1483–1492, 1997.
- [22] Michael S. Lewicki. Efficient coding of natural sounds. *Nature Neuroscience*, 5(4):356–363, 2002.
- [23] E. de Boer and H. R. de Jongh. On cochlear encoding: Potentialities and limitations of the reverse-correlation technique. *J. Acoust. Am.*, 63(1), 1978.
- [24] S. Mardelja. Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.*, 70(11), 1980.
- [25] J. H. van Hateren and A. van der Schaaf. Independent component fillters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond. B*, 1998.
- [26] Allan Kardec Barros Daniel Duarte Costa, Lucio Flavio Campos and Aristofanes Correa Silva. Independent component analysis in breast tissues mammograms images classification using lda and svm. In *Proceedings of the IEEE Engineering in Medicine and Biology Society*, 2007.
- [27] Carlos Magno Sousa, André Borges Cavalcante, Denner Guilhon, and Allan Kardec Barros. Image compression by redundancy reduction. In Mike E. Davies, Christopher J. James, Samer A. Abdallah, and Mark D. Plumbley, editors, *ICA*, volume 4666 of *Lecture Notes in Computer Science*, pages 422–429. Springer, 2007.

- [28] Denner Guilhon, Allan Kardec Barros, and Silvia Comani. Ecg compression by efficient coding. In Mike E. Davies, Christopher J. James, Samer A. Abdallah, and Mark D. Plumbley, editors, *Independent Component Analysis and Signal Separation, 7th International Conference*, volume 4666 of *Lecture Notes in Computer Science*, pages 593–600. Springer, 2007.
- [29] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000. Circulation Electronic Pages: <http://circ.ahajournals.org/cgi/content/full/101/23/e215>.
- [30] Kaïs Ouni Sana Ktata and Nouredine Ellouze. A novel compression algorithm for electrocardiogram signals based on wavelet transform and spiht. *International Journal of Signal Processing*, 5(4), 2009.
- [31] A. Hyvarinen. Complexity pursuit: separating interesting components from time-series. *Neural Computation*, 13(4):883–898, 2001.
- [32] Mark Girolami Te-Won Lee and Terrence J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, 11(2):417–441, 1999.
- [33] David A. Huffman. A method for the construction of minimum-redundancy codes. *Proceedings of the I.R.E.*, 40(9):1098–1101, 1952.