

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

DENNER ROBERT RODRIGUES GUILHON

COMPRESSÃO DE SINAIS DE ELETROCARDIOGRAMA
UTILIZANDO ANÁLISE DE COMPONENTES
INDEPENDENTES

São Luís - MA

2006

DENNER ROBERT RODRIGUES GUILHON

COMPRESSÃO DE SINAIS DE ELETROCARDIOGRAMA
UTILIZANDO ANÁLISE DE COMPONENTES
INDEPENDENTES

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

Orientador: Allan Kardec Duailibe Barros Filho

Universidade Federal do Maranhão

São Luís - MA

2006

Guilhon, Denner Robert Rodrigues

Compressão de sinais de eletrocardiograma utilizando análise de componentes independentes / Denner Robert Rodrigues Guilhon. - São Luís, 2006.

Dissertação (Mestrado) - Programa de Pós-Graduação em Engenharia de Eletricidade, Universidade Federal do Maranhão.

1.Eletrocardiograma-Algoritmo. 2. Análise de componentes independentes. 3.KLT. I. Título.

CDU 004.421:616.12-073.97

DENNER ROBERT RODRIGUES GUILHON

COMPRESSÃO DE SINAIS DE ELETROCARDIOGRAMA
UTILIZANDO ANÁLISE DE COMPONENTES
INDEPENDENTES

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

Aprovada em 24 de fevereiro de 2006

BANCA EXAMINADORA

Allan Kardec Duailibe Barros Filho

Universidade Federal do Maranhão

Hani Camille Yehia

Universidade Federal de Minas Gerais

João Viana da Fonseca Neto

Universidade Federal do Maranhão

*Ao meu pai, que à sua maneira sempre torceu
pelo sucesso de cada um de seus filhos.*

*À minha mãe e aos meus irmãos, pelo
companheirismo em todos os momentos.*

*À minha esposa, Elizângela, pelo apoio
incondicional.*

Resumo

A demanda contínua de por sistemas de processamento de eletrocardiogramas de alto desempenho e baixo custo tem exigido a elaboração de técnicas de compressão de ECG cada vez mais eficientes e confiáveis. O objetivo deste trabalho é avaliar o desempenho de um algoritmo baseado em análise de componentes independentes (ICA) para a compressão de eletrocardiogramas (ECGs). Para cada um dos sinais de ECG utilizados foram obtidos, através de ICA, subespaços vetoriais construídos a partir de suas funções base, pois o sinal pode ser expresso como uma combinação linear destas.

O sinal de ECG foi subdividido em m janelas de comprimento fixo, e cada uma delas foi projetada no subespaço, resultando em um vetor \mathbf{w} de coeficientes para cada janela. Um processo de quantização simples foi executado para os m vetores \mathbf{w} , segundo níveis de quantização definidos, cada um gerando diferentes níveis de erro de reconstrução.

Foi observado que o armazenamento dos coeficientes implica na utilização de um menor espaço em memória em comparação àquele utilizado pelas janelas correspondentes do sinal de eletrocardiograma. A medida tradicionalmente utilizada de erro de reconstrução, diferença média quadrática percentual (PRD), foi empregada para a avaliação do algoritmo. Os resultados foram comparados àqueles obtidos utilizando a transformada de Karhunen Loève (KLT).

PALAVRAS-CHAVE: análise de componentes independentes, eletrocardiograma, compressão, KLT.

Abstract

The continuing demand for high performance and low cost electrocardiogram processing systems have required the elaboration of even more efficient and reliable ECG compression techniques. The objective of this work is to evaluate the performance of an electrocardiogram (ECG) compression algorithm based on independent components analysis (ICA). To each of the ECG signal we processed, using ICA, vectorial subspaces composed of its basis functions were obtained, for the signal can be expressed as a linear combination of them.

The ECG signal was subdivided into m fixed length windows, and each of them was projected in the subspace, resulting in a vector \mathbf{w} of coefficients for each window. A simple quantization process was performed over the m vectors \mathbf{w} , according to defined levels of quantization, each one generating different levels of reconstruction error.

It was observed that the storage of the coefficients implies the use of less space in memory in comparison to that one used by the corresponding windows of the electrocardiogram signal. The reconstruction error measure traditionally used, the percent root mean-square difference (PRD), was used into the evaluation of the algorithm. The results had been compared with those obtained using the Karhunen Loéve transform (KLT).

KEYWORDS: independent component analysis, electrocardiogram, KLT.

Agradecimentos

Mais do que a todos os outros, agradeço a Deus por ter me permitido chegar ao fim desta etapa da minha vida.

Agradeço também aos meus pais e irmãos por terem sempre e invariavelmente me incentivado a perseverar nos estudos, ainda que algumas vezes eu tenha discordado.

Sou muito grato à minha inabalável esposa, Elizângela, que esteve sempre ao meu lado, mesmo nas piores horas.

Expresso minha gratidão aos colegas do Laboratório de Processamento da Informação Biológica - PIB e dos corredores: Ana Bárbara, André Cavalcante, Carla Borba, Carlos Magno, Ewaldo Santana, Fausto Lucena, Glenda Salgado, Ivan Júnior, Jaciani Pereira, Lúcio Campos, Raniere Machado, Renata Feques, Ricardo Robson, entre outros. Também agradeço a todos os professores, em especial ao Prof. Eugênio Medeiros, por sua colaboração paciente, e ao Prof. Allan Kardec Barros, por não desanimar diante da inércia dos alunos.

“O problema não é como ter pensamentos novos e inovadores, mas como tirar os velhos da cabeça.”

Dee Hock

Sumário

Lista de Figuras		viii
Lista de Tabelas		x
Lista de Abreviações		xi
1 Introdução		1
2 Aspectos Básicos da Teoria da Informação		3
2.1 Fontes de Informação e Entropia		3
2.2 Teoria da Codificação de Fontes		5
3 Análise de Componentes Independentes		8
3.1 Histórico e Motivação		8
3.2 Definições		9
3.3 Estimação de Componentes Independentes		11
3.3.1 Estimação através de Maximização de Não-Gaussianidade		11
3.3.2 Negentropia como Medida de Não-Gaussianidade		13
4 Eletrocardiogramas		14
4.1 Definições e Forma de Aquisição		14
4.2 Compressão de Eletrocardiogramas		17
4.2.1 Transformada de Karhunen-Loève		19
5 Compressão utilizando Análise de Componentes Independentes		22
5.1 Método		22

	vii
5.1.1 Prova	23
5.1.2 Estimação das Funções Bases através de ICA	24
5.1.3 Estimação dos Coeficientes da Projeção	25
5.2 Material Utilizado	27
5.3 Resultados	27
5.4 Discussão	29
6 Conclusões	35
Referências	37

Lista de Figuras

2.1	Diagrama de blocos de um sistema de comunicações.	3
3.1	Segmento de um sinal de eletrocardiograma como combinação linear de suas características.	9
3.2	Segmentos de sinal de eletrocardiograma dispostos como combinação linear das funções base do sinal. (a) Segmentos do ECG. (b) Componentes independentes. (c) Funções base.	10
4.1	Triângulo de Einthoven e as derivações DI, DII e DIII.	14
4.2	Esquema do terminal central de Wilson.	15
4.3	Esquema das derivações aumentadas de Goldberger.	15
4.4	Esquema das derivações precordiais.	16
4.5	Exemplos de registros das 12 derivações: (a) derivações bipolares e aumentadas. (b) derivações precordiais.	17
4.6	Estrutura do eletrocardiograma normal.	17
5.1	Diagrama de blocos para o algoritmo de compressão proposto: (a) fase de treinamento: estimação das funções base. (b) fase de estimação dos coeficientes.	22
5.2	Combinador linear de múltiplas entradas.	26
5.3	PRD após a reconstrução de 15 registros do MIT-BIH NSRD, segundo taxas de compressão fixas, obtidas utilizando PCA (KLT) e ICA. Observar que os traçados de PCA e ICA têm a mesma cor (vermelho, azul, verde ou preto) para taxas de compressão iguais.	30

5.4	PRD após a reconstrução de 15 registros do MIT-BIH SVDB, segundo taxas de compressão fixas, obtidas utilizando PCA (KLT) e ICA. Observar que os traçados de PCA e ICA têm a mesma cor (vermelho, azul, verde ou preto) para taxas de compressão iguais.	31
5.5	PRD após a reconstrução de 15 registros do MIT-BIH VFDB, segundo taxas de compressão fixas, obtidas utilizando PCA (KLT) e ICA. Observar que os traçados de PCA e ICA têm a mesma cor (vermelho, azul, verde ou preto) para taxas de compressão iguais.	32
5.6	Resultados do algoritmo proposto para as 600 primeiras amostras para o registro 16265. (a) Sinal original. (b) Sinal obtido após a reconstrução, com CR = 2.4:1 e PRD = 3.26%. (c) Erro de reconstrução.	33
5.7	Resultados do algoritmo proposto para as 600 primeiras amostras para o registro SVDB 809. (a) Sinal original. (b) Sinal obtido após a reconstrução, com CR = 2:1 e PRD = 4.85%. (c) Erro de reconstrução.	33

Lista de Tabelas

- 4.1 Comparação entre alguns esquemas de Compressão de ECG 19
- 5.1 Relação de registros utilizados das base de dados. 28
- 5.2 Taxa de amostragem e resolução das bases de dados. 28

Lista de Abreviações

AZTEC	amplitude zone-time epoch coding
BIH	Beth Israel Hospital
CORTES	coordinate reduction time encoding system
CR	compression ratio
DPCM	differential pulse code modulation
ECG	eletrocardiograma
ICA	independent component analysis
KLT	Karhunen-Loève transform
MIT	Massachusetts Institute of Technology
MSE	mean-square error
MOS	mean opinion score
NSRD	Normal Sinus Rhythm Database
PCA	principal components analysis
PRD	percent root mean-square difference
SAPA	scan-along polygonal approximation
SVDB	Supraventricular Arrhythmia Database
TP	turning point
VFDB	Malignant Ventricular Arrhythmia Database
WDD	weighted diagnostic distortion

1 Introdução

O eletrocardiograma é usado para determinar as condições do coração através de medidas de sua atividade elétrica. Os sinais eletrocardiográficos dos pacientes podem ser armazenados com os propósitos de diagnóstico, comparação e avaliação futura. O conjunto de inúmeros desses registros compõem imensas bases de dados que, por motivos de eficiência de armazenamento e transmissão, exigem métodos de compressão de dados efetivos.

Em ciência da computação e teoria da informação, a compressão de dados é o processo de conversão de uma mensagem em uma representação que consuma menos bits (ou outra medida de informação) que a representação original. A compressão é possível devido à presença de redundâncias estatísticas nos dados a serem comprimidos.

A compressão de dados é importante pois reduz o consumo de recursos, como espaço em disco ou largura de banda da conexão. Entretanto, requer capacidade de processamento, o que pode se mostrar dispendioso. O projeto de esquemas de compressão inclui compromissos entre vários fatores como a capacidade de compressão, quantidade de distorção introduzida e recursos computacionais exigidos.

Os esquemas de compressão de eletrocardiogramas são classificados essencialmente em: compressão direta de dados, métodos de transformação e técnicas de extração de parâmetros. Trabalhos recentes sobre compressão de eletrocardiogramas utilizam principalmente os métodos de transformação, tais como a transformação de Karhunen-Loève [OLMOS et al (1996)] e wavelets [BLANCO-VELASCO et al(2004), MIAOU & CHAO (2005), RAJOUB (2002)].

Entretanto, tais trabalhos não exibem com clareza os resultados dos algoritmos propostos, pois fazem uso de técnicas adicionais de codificação (DPCM, Huffman e quantização vetorial, por exemplo) [JALALEDDINE et al (1990)]. Isso faz com que tanto a complexidade dos sistemas necessários para reproduzir tais taxas de compressão seja aumentada quanto nos torna incapazes de definir a eficiência dos algoritmos propostos sem o uso dessa codificação final.

Neste trabalho propomos um novo método de compressão onde é criado, para um dado sinal de ECG, um subespaço vetorial através do qual podemos expressar completamente o sinal. Para a estimação desse subespaço é utilizada a análise de componentes independentes. O sinal é então dividido em segmentos, para os quais são calculados coeficientes de tal forma que todas esses segmentos sejam descritas como combinações lineares do subespaço vetorial encontrado.

Com essa técnica, obtemos uma representação menos redundante do sinal de ECG, o que nos permite uma codificação mais eficiente. Diferentemente de outros trabalhos, optou-se por mostrar os resultados do algoritmo sem o beneficiamento de técnicas adicionais, sendo utilizado apenas quantização simples.

Este trabalho está organizado da seguinte forma. O Capítulo 2 apresenta os elementos essenciais ao trabalho no que se refere à Teoria da Informação, tais como fontes, entropia e codificação. No Capítulo 3 são apresentadas as bases de análise de componentes independentes, sua definição e forma de estimação de componentes. O Capítulo 4 trata da definição de eletrocardiograma, sua forma de aquisição e uma breve discussão sobre compressão de ECGs. No Capítulo 5 é descrito o método proposto, assim como sua fundamentação, material utilizado para teste, apresentação e discussão dos resultados obtidos. Finalmente, o Capítulo 6 mostra as considerações finais sobre o trabalho, assim como propostas de trabalhos futuros.

2 Aspectos Básicos da Teoria da Informação

A Teoria da Informação abrange a transmissão e o armazenamento de dados, tratando da análise dos sistemas de comunicação [ASH (1990)], tradicionalmente representado pelo diagrama de blocos mostrado na Figura 2.1. A transmissão de dados através de um canal ruidoso é um problema clássico da teoria da informação.

O principal resultado dessa teoria é o fato de que a utilização de estratégias de codificação e decodificação apropriadas tornam possível a comunicação em canais não-confiáveis a taxas menores que a (mas arbitrariamente próximas da) capacidade do canal, com pequenas probabilidades de erro [WIKIPEDIA (2005d)].

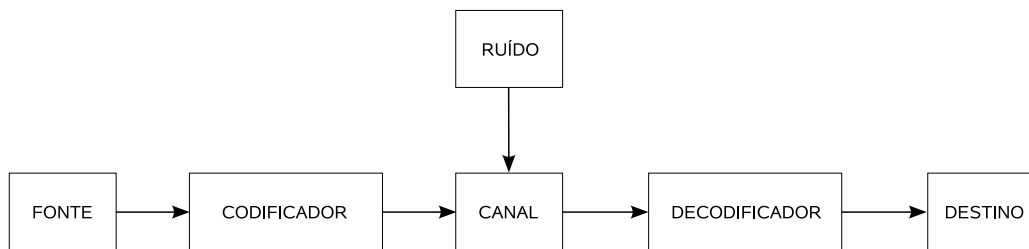


Figura 2.1: Diagrama de blocos de um sistema de comunicações.

No esquema da Figura 2.1, a fonte produz as informações a serem transmitidas, enquanto o codificador associa a cada mensagem gerada pela fonte seqüências binárias mais adequadas para a transmissão através do canal. O decodificador, por sua vez, opera na saída do canal, com o objetivo de extrair a mensagem original da mensagem codificada. Porém, a qualidade desta transmissão é geralmente afetada pela presença de ruído.

2.1 Fontes de Informação e Entropia

Uma fonte de informação pode ser representada em termos de modelos probabilísticos que emitem eventos ou variáveis aleatórias [VITERBI & OMURA (1979)]. Uma fonte discreta sem memória é caracterizada por sua saída, a variável aleatória S , que toma

elementos de um alfabeto finito de comprimento N

$$\vartheta = (s_1, s_2, \dots, s_N), \quad (2.1)$$

com probabilidades

$$P(S = s_k) = p_k \quad k = 1, 2, \dots, N. \quad (2.2)$$

A fonte é dita sem memória, pois a cada unidade de tempo T_s emite uma variável aleatória que é independente de todas as saídas anteriores e posteriores da fonte. Portanto, se a qualquer tempo a saída da fonte é $S = s_k$, então a quantidade de informação, em bits, contida na saída da fonte é dada por

$$I(s_k) = \log_2 \left(\frac{1}{p_k} \right). \quad (2.3)$$

A quantidade de símbolos necessários, por unidade de tempo, para representar a fonte totalmente e, em seguida, reconstruir a sua seqüência de saída, ou seja, a quantidade média de informação por símbolo da fonte, é chamada entropia da fonte [HAYKIN (2001), PAPOULIS & PILLAI (2002), VITERBI & OMURA (1979)], dada por

$$H(\vartheta) = \sum_{k=1}^N P(s_k) I(s_k) \quad (2.4a)$$

$$= \sum_{k=1}^N p_k \log_2 \left(\frac{1}{p_k} \right). \quad (2.4b)$$

A entropia $H(\vartheta)$, de acordo com (2.4b), depende apenas das probabilidades dos símbolos do alfabeto ϑ da fonte. Quanto mais aleatória, ou seja, quanto mais imprevisível e não estruturada for a variável, maior a sua entropia. A entropia da fonte está limitada a $0 \leq H(\vartheta) \leq \log_2 N$, onde $H(\vartheta) = 0$ ocorre quando não existe incerteza sobre o símbolo emitido pela fonte, e $H(\vartheta) = \log_2 N$ quando os símbolos emitidos pela fonte são equiprováveis.

Na discussão de conceitos de teoria da informação, é bastante útil considerar blocos de símbolos em vez de símbolos individuais. Cada bloco consiste de n símbolos sucessivos emitidos pela fonte [HAYKIN (2001)]. O alfabeto da fonte estendida pode ser representado por ϑ^N . Devido símbolos da fonte discreta sem memória serem estatisticamente independentes, a probabilidade da fonte de símbolos ϑ^N é igual ao produto das probabilidades das N fontes de símbolos ϑ que constituem ϑ^N . Portanto, podemos escrever

$$H(\vartheta^N) = NH(\vartheta). \quad (2.5)$$

2.2 Teoria da Codificação de Fontes

A codificação de fontes é uma aplicação direta da Teoria da Informação, pois um dos problemas em comunicações é a representação eficiente dos dados gerados por uma fonte discreta. O processo através do qual essa representação é obtida é chamado codificação de fonte. Um código é o mapeamento da saída da fonte, cujas palavras pertencem ao alfabeto ϑ da fonte, em palavras código, de um alfabeto ψ de código.

Quando a compressão é utilizada para aplicações de transmissão de dados, o objetivo é velocidade. Velocidade de transmissão depende do número de bits enviados, do tempo necessário para codificador gerar código e para o decodificador recuperar o sinal original. Quando a aplicação é de armazenamento de dados, a preocupação é a quantidade de memória utilizada após a compressão.

As técnicas de compressão de dados têm sido utilizadas em várias áreas de comunicação, tais como fal e imagem. Tais métodos são tradicionalmente classificados em três categorias principais [JALALEDDINE et al (1990)]:

- a) Compressão direta de dados;
- b) Métodos de transformação;
- c) Técnicas de extração de parâmetros.

Compressão de dados através de transformação ou de métodos diretos contêm dados transformados ou dados reais do sinal original, podendo ser reconstruídos através de processo inverso. A compressão direta de dados baseia sua detecção de redundâncias na análise direta das amostras sinal.

Já os métodos de transformação utilizam principalmente a análise das distribuições espectrais e de energia para a detecção de redundâncias. Por outro lado, a extração de parâmetros é um processo irreversível com o qual uma característica ou parâmetro particular do sinal é extraído. Os parâmetros extraídos (e.g. medidas da distribuição de probabilidade) são então utilizados para classificação baseada em conhecimento *a priori* das características do sinal.

A compressão de dados pode ser também classificada em: sem perdas, onde os dados podem ser exatamente reconstruídos; com perdas, em que são calculados quantos bits

são necessários para reconstruir os dados dentro de um nível estabelecido de fidelidade. Entre as técnicas de compressão sem perdas estão a codificação de Huffman e a codificação aritmética. As com perdas incluem JPEG, MP3, MPEG, etc.

O propósito do par codificador-decodificador de fonte na Figura 2.1 é reduzir a saída da fonte à representação mínima, ou seja, a codificação mais eficiente. Para se alcançar essa representação, é necessário conhecer a estatística da fonte, de forma a criar um código unicamente decodificável [ASH (1990), HAYKIN (2001), VITERBI & OMURA (1979)]. Essa representação será ótima quando sua redundância for mínima.

Dada uma fonte discreta sem memória, cujo alfabeto ϑ possua probabilidades p_k , onde $k = 1, \dots, N$. Considerando que para cada símbolo s_k gerado é atribuída uma palavra código de comprimento l_k bits, define-se o comprimento médio de palavra código como

$$\bar{L} = \sum_{k=1}^N P(s_k) l_k \quad (2.6a)$$

$$= \sum_{k=1}^N p_k l_k. \quad (2.6b)$$

Fisicamente, \bar{L} representa o número médio de bits por símbolo da fonte usado no processo de codificação [HAYKIN (2001)]. Assim, a redundância de uma codificação pode ser definida como [LELEWER & HIRSCHBERG (1987)]

$$\mathfrak{R} = \sum_k p_k l_k - \sum_k p_k \log_2 \left(\frac{1}{p_k} \right), \quad (2.7)$$

o que, de acordo com (2.4b) e (2.6a), equivale a

$$\mathfrak{R} = \bar{L} - H(\vartheta). \quad (2.8)$$

Assim, a redundância mede a diferença entre o comprimento médio de palavra código e a quantidade média de informação. Se um código tem o mínimo comprimento médio de palavra código \bar{L} , para uma dada distribuição de probabilidades, então esse código tem redundância mínima. O menor valor possível de \bar{L} pode ser determinado através do primeiro teorema de Shannon, o Teorema da Codificação de Fontes. Esse teorema afirma que, dada uma fonte discreta sem memória de entropia $H(\vartheta)$, o comprimento médio de palavra código \bar{L} , para qualquer esquema de codificação sem distorção, é limitado por [SHANNON (1948)]

$$\bar{L} \geq H(\vartheta). \quad (2.9)$$

Logo, a entropia da fonte $H(\vartheta)$ é a taxa mínima, ou ainda, o menor comprimento de código pelo qual uma seqüências de uma fonte digital estocástica pode ser transmitida perfeitamente [PAPOULIS & PILLAI (2002), VITERBI & OMURA (1979)].

A quantidade de compressão obtida por um esquema de codificação pode ser medida pela a taxa de compressão (do inglês, *compression ratio*), definido por

$$CR = \frac{\bar{M}}{\bar{L}} \quad (2.10)$$

sendo que \bar{M} é o comprimento médio da mensagem e \bar{L} , conforme (2.6a), é comprimento médio de palavra código. Esse valor mostra a comprimento da mensagem original em relação à mensagem codificada.

3 Análise de Componentes Independentes

3.1 Histórico e Motivação

O problema de separação de fontes é antigo na engenharia elétrica. Existem muitos algoritmos, dependendo da natureza da mistura dos sinais. O problema de separação cega de fontes é maior, pois sem o conhecimento dos sinais que foram misturados, é impossível propor o pré-processamento apropriado que os separe de forma ótima. A estrutura geral de análise de componentes independentes foi introduzida por Jean Herault e Christian Jutten em 1986, sendo mais claramente proposta por Pierre Comon em 1994.

Em 1995 Tony Bell e Terry Sejnowski introduziram um algoritmo rápido e eficiente de ICA baseado no infomax, um princípio criado por Ralph Linsker em 1992. Em 1997, Shun-ichi Amari melhorou este algoritmo usando o gradiente natural, o que foi descoberto de forma independente por Jean Francois Cardoso. Entretanto, o algoritmo original, com não-linearidades sigmoidais, era apenas apropriado para fontes super-Gaussianas. Te-Won Lee, juntamente com Mark Girolami, desenvolveu uma versão ampliada eficiente do algoritmo de ICA baseado no infomax, própria para sinais não-Gaussianos em geral [WIKIPEDIA (2005c)].

Muitas abordagens diferentes foram utilizadas para a separação cega de fontes, o que inclui máxima verossimilhança, método Bussgang baseado em cumulantes e negentropia. Todas elas são relacionadas à estrutura do infomax. Por isso, um grande número de pesquisadores que se dedicaram ao estudo de ICA, vindos de várias áreas, convergiram para um conjunto comum de princípios e de algoritmos. Para um histórico mais completo, ver [COMON (1994)].

Considere, por exemplo, os registros elétricos das atividades do cérebro, dados por um eletroencefalograma (EEG), consistindo dos potenciais elétricos em diferentes pontos do escalpo. Esses potenciais são supostamente gerados pela mistura de componentes implícitas da atividade cerebral. Essa situação é semelhante ao problema de *cocktail party*, pois desejamos encontrar as componentes originais da atividade cerebral, mas podemos apenas observar as

misturas das componentes. ICA pode revelar informações interessantes sobre a atividade cerebral a partir das suas componentes independentes

Outra aplicação importante de ICA é a extração de características. Um problema fundamental em processamento de sinais é encontrar uma representação apropriada para imagem, áudio e outros formatos de dados quando da compressão e da remoção de ruídos. Representação de dados pode ser feita através de transformações lineares, logo é bastante útil estimar tal transformação dos próprios dados, pois nesse caso a transformação seria adaptada ao tipo de dado processado [HYVÄRINEN & OJA (2000), HYVÄRINEN et al (2001)].

3.2 Definições

Considere que sejam observados n segmentos aleatórios x_1, \dots, x_n de um eletrocardiograma, modeladas como combinação linear de n funções base

$$\mathbf{x}_n = \varphi_1 s_{j1} + \varphi_2 s_{j2} + \dots + \varphi_n s_{jn} \quad j = 1, \dots, n \quad (3.1)$$

e que cada segmento \mathbf{x}_n , assim como cada componente independente \mathbf{s}_n 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 [WIKIPEDIA (2005a)]. As Figuras 3.1 e 3.2 ilustram o proposto acima.

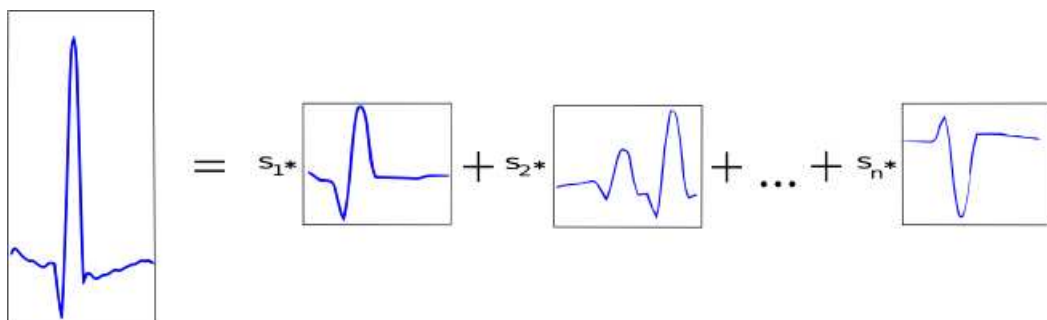


Figura 3.1: Segmento de um sinal de eletrocardiograma como combinação linear de suas características.

Sem perda de generalidade, supomos que tanto as variáveis originadas dos segmentos do ECG quanto aquelas das componentes independentes têm média zero. Por conveniência, será usada a notação vetorial em vez de somas, como aquelas vistas em (3.1),

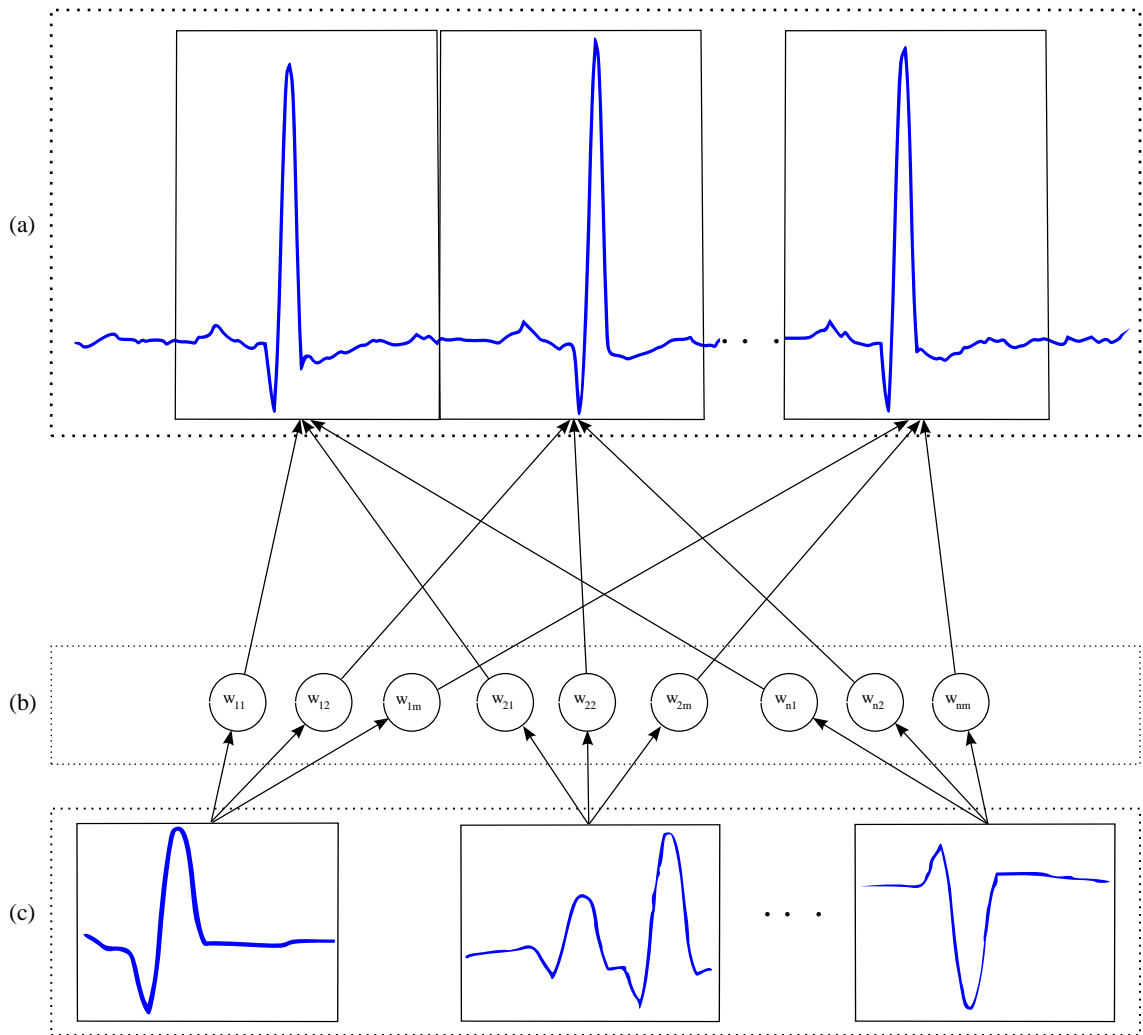


Figura 3.2: Segmentos de sinal de eletrocardiograma dispostos como combinação linear das funções base do sinal. (a) Segmentos do ECG. (b) Componentes independentes. (c) Funções base.

utilizando letras minúsculas e maiúsculas, ambas em **negrito**, para representar, respectivamente, vetores coluna e matrizes. Dessa maneira, podemos reescrever (3.1) das seguintes formas

$$\mathbf{X} = \mathbf{\Phi}\mathbf{S}. \quad (3.2)$$

O modelo estatístico de (3.2) é chamado de modelo de análise de componentes independentes. É preciso estimar tanto a matriz de componentes independentes \mathbf{S} quanto a matriz de funções base $\mathbf{\Phi}$, que também é desconhecida, pois tudo o que se observa são os segmentos do eletrocardiograma, \mathbf{X} .

Para tanto, é preciso fazer suposições tão gerais quanto possível. Portanto, supomos que [HYVÄRINEN et al (2001)]:

- a) As componentes s_n são estatisticamente independentes;

- b) Elas têm distribuições não-gaussianas;
- c) Por motivos de simplicidade, a matriz Φ seja quadrada.

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

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

Ambas derivam do fato de \mathbf{S} e Φ serem desconhecidas. No item 1, existe ambigüidade já que qualquer escalar α_n multiplicando uma das fontes s_n pode ser cancelado dividindo-se a coluna φ_n correspondente pelo mesmo escalar α_i , ou seja

$$\mathbf{x}_n = \sum_n \left(\frac{1}{\alpha_n} \varphi_n \right) (\mathbf{s}_{jn} \alpha_n). \quad (3.3)$$

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 (3.1), denominando qualquer componente como a primeira.

3.3 Estimação de Componentes Independentes

3.3.1 Estimação através de Maximização de Não-Gaussianidade

A não-gaussianidade é um elemento chave para a estimação do modelo de ICA, pois a matriz Φ não é identificável quando as componentes independentes têm distribuição gaussiana. Assumimos que \mathbf{x} seja um dos segmentos de ECG, conforme a Figura 3.2, distribuído de acordo com o modelo de ICA em (3.2) e que todas as componentes independentes s têm distribuições iguais. Para estimar as componentes independentes, basta encontrar as combinações lineares corretas de x_i , de modo que

$$\mathbf{s} = \Phi^{-1} \mathbf{x}. \quad (3.4)$$

Assim, podemos expressar uma combinação linear de x_i por

$$y = \mathbf{b}^T \mathbf{x} \quad (3.5a)$$

$$= \sum_i b_i x_i \quad (3.5b)$$

$$= \mathbf{b}^T \Phi \mathbf{s}, \quad (3.5c)$$

onde \mathbf{b} deve ser determinado. A partir de (3.5c) podemos observar que y é uma combinação linear de s_i , com coeficientes dados por $\mathbf{q} = \mathbf{b}^T \Phi$. Logo obtemos

$$y = \mathbf{q}^T \mathbf{s} \quad (3.6a)$$

$$= \sum_i q_i s_i. \quad (3.6b)$$

Se \mathbf{b} corresponder a uma das linhas da inversa de Φ , então 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. Não é possível determinar \mathbf{b} exatamente, mas podemos estimar seu valor com boa aproximação.

Uma forma de determinar \mathbf{b} é variar os coeficientes em \mathbf{q} e então verificar como a distribuição de $y = \mathbf{q}^T \mathbf{s}$ muda. Já que, conforme o Teorema do Limite Central [PAPOULIS & PILLAI (2002)], a soma de duas variáveis aleatórias independentes é mais gaussiana que as variáveis originais [PAPOULIS & PILLAI (2002)], $y = \mathbf{q}^T \mathbf{s}$ normalmente é mais gaussiana que qualquer uma das s_i e menos gaussiana quando se iguala a uma das s_i . Nesse caso, apenas um dos elementos q_i de \mathbf{q} é diferente de zero [HYVÄRINEN et al (2001)].

Como, na prática, os valores de \mathbf{q} são desconhecidos e sabemos, através de (3.5a) e (3.6a), que

$$\mathbf{b}^T \mathbf{x} = \mathbf{q}^T \mathbf{s}, \quad (3.7)$$

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 $\mathbf{q} = \Phi^T \mathbf{s}$, vetor esse que possui apenas uma de suas componentes diferente de zero. Isso significa que y em (3.5a) é igual a uma das componentes independentes. Logo, a maximização da não-gaussianidade de $\mathbf{b}^T \mathbf{x}$ permite encontrar uma das componentes.

3.3.2 Negentropia como Medida de Não-Gaussianidade

Uma importante medida de não-gaussianidade é a negentropia. A definição de entropia em (2.4a) pode ser generalizada para vetores e variáveis aleatórias contínuas, vindo a ser chamada entropia diferencial. Tomando um vetor aleatório \mathbf{y} cuja função densidade de probabilidade é $f(\mathbf{y})$, temos a entropia diferencial dada por

$$H(\mathbf{y}) = - \int f(\mathbf{y}) \log f(\mathbf{y}). \quad (3.8)$$

Como um dos resultados fundamentais da Teoria da Informação, sabe-se que uma variáveis gaussianas tem a maior entropia entre todas as variáveis aleatórias de igual variância [HYVÄRINEN et al (2001), PAPOULIS & PILLAI (2002)]. Isso quer dizer que uma versão modificada da entropia diferencial pode ser usada como medida de não-gaussianidade. Essa medida é chamada negentropia, definida por

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}), \quad (3.9)$$

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 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. Logo

$$J(\mathbf{y}) \approx \frac{1}{12} E \{y^3\}^2 + \frac{1}{48} \text{kurt}(y)^2 \quad (3.10)$$

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

$$\text{kurt}(y) = E \{y^4\} - 3(E \{y^2\})^2. \quad (3.11)$$

A kurtose é zero para variáveis gaussianas e maior que zero para a maioria das variáveis aleatórias não-gaussianas.

4 Eletrocardiogramas

4.1 Definições e Forma de Aquisição

O eletrocardiograma (ECG) é um teste não-invasivo usado para determinar as condições do coração através das medidas de sua atividade elétrica [KULICK (2005)]. O ECG é construído utilizando as medidas dos potenciais elétricos presentes nos tecidos cardíacos. Estas medidas são obtidas por meio de derivações, circuitos formados por dois eletrodos ligados aos pólos positivo e negativo de um galvanômetro.

As derivações DI, DII e DIII (também chamadas de derivações bipolares) foram introduzidas por Einthoven, que imaginou o coração no centro de um triângulo equilátero cujos vértices estariam representados pelo braço direito (R), braço esquerdo (L) e perna esquerda (F) [UNIFESP VIRTUAL (2003)], conforme mostrado na Figura 4.1. Essa orientação foi baseada na Segunda Lei de Kirchoff que diz que num circuito fechado, a soma das diferenças de potencial é igual a zero.

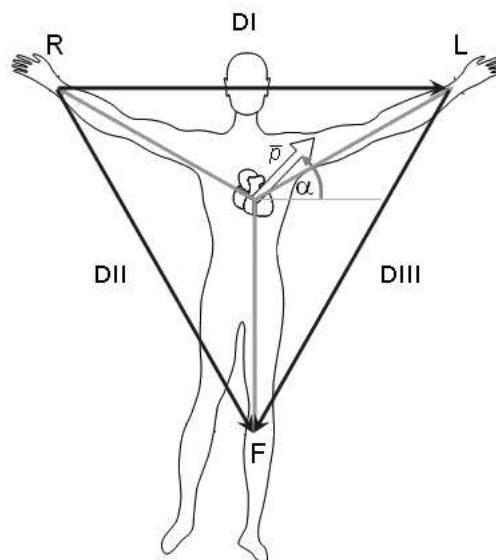


Figura 4.1: Triângulo de Einthoven e as derivações DI, DII e DIII.

Wilson introduziu o chamado terminal central (T), cujo potencial é considerado

nulo em relação a todas as regiões do coração. Os potenciais dos pontos L, R e F são medidos com referência ao terminal. O terminal de Wilson é obtido unindo-se os vértices do triângulo de Einthoven a um terminal central através de resistências iguais de $5\text{k}\Omega$, como ilustra a Figura 4.2. Três derivações adicionais são obtidas medindo-se o potencial entre o eletrodo de cada membro e o terminal central de Wilson. Em 1942, Goldberger observou que esses sinais poderiam ser aumentados omitindo-se a resistência do terminal central de Wilson [MALMIVUO & PLONSEY (1995)], que é conectado ao eletrodo de medida. As derivações assim obtidas são chamadas aVR, aVL e aVF, mostradas na Figura 4.3.

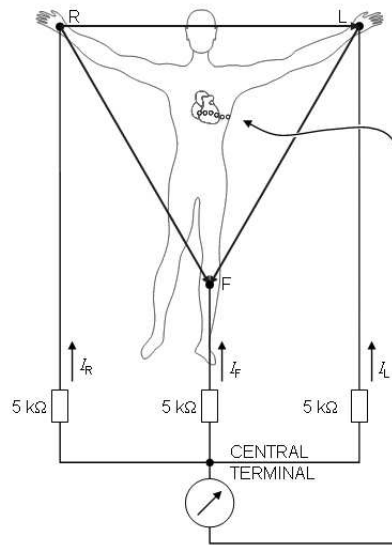


Figura 4.2: Esquema do terminal central de Wilson.

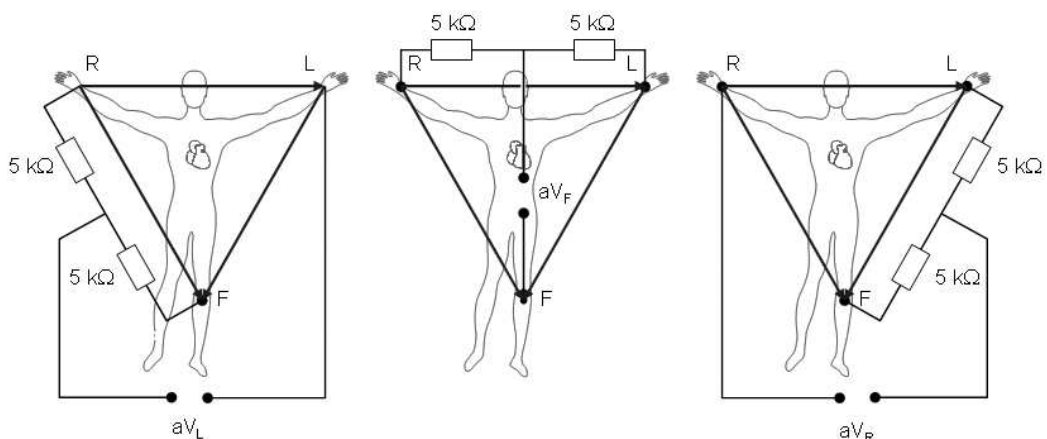


Figura 4.3: Esquema das derivações aumentadas de Goldberger.

Para medir os potenciais próximos ao coração, Wilson introduziu as derivações precordiais. Estas derivações são obtidas unindo-se o terminal de Wilson, onde o eletrodo

negativo é colocado, e posicionando o eletrodo explorador, positivo, sucessivamente sobre seis posições da superfície torácica, conforme a Figura 4.4: quarto espaço intercostal, à direita do esterno (V1); quarto espaço intercostal, à esquerda do esterno (V2); a meio caminho entre os pontos V2 e V4 (V3); quinto espaço intercostal esquerdo, na linha claviclar média (V4); quinto espaço intercostal esquerdo, na linha axilar anterior (V5); quinto espaço intercostal esquerdo, na linha axilar média (V6) [UNIFESP VIRTUAL (2003)].

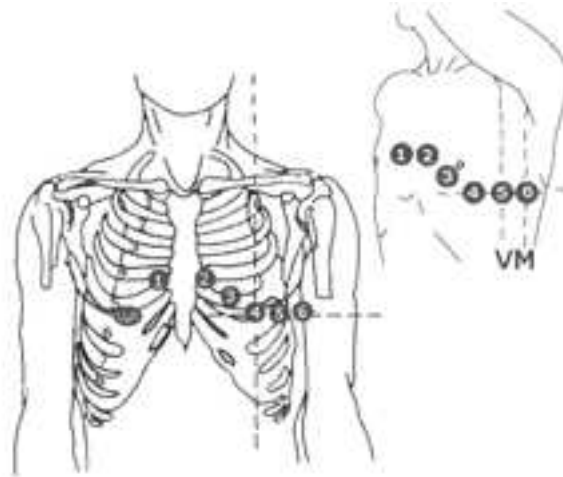


Figura 4.4: Esquema das derivações precordiais.

Portanto, as doze derivações registram informações de regiões particulares do coração. As derivações inferiores (DII, DIII e aVF) registram a atividade elétrica do topo do ventrículo esquerdo, enquanto as derivações laterais (DI, aVL, V5 e V6) registram a atividade elétrica da parede frontal do ventrículo esquerdo. Já as derivações anteriores (V1 a V6) representam a parede anterior do coração, ou seja, a parede frontal do ventrículo esquerdo. A derivação aVR raramente é usada para fins de diagnóstico, mas indica se as derivações foram dispostas corretamente no paciente.

A Figuras 4.5(a) e 4.5(b) mostra exemplos dos registros obtidos a partir de cada uma das derivações, onde as diferenças entre os registros mostram, a partir de cada derivação, a seqüência de polarização e despolarização dos átrios e ventrículos [KLABUNDE (2005)]. A Figura 4.6 apresenta um registro eletrocardiográfico típico.

Através do ECG, é possível diagnosticar [THE BETTER HEALTH CHANNEL (2003)]: aumento do coração; defeitos cardíacos congênitos; arritmias; posição anormal do coração; inflamações (pericardite e miocardite), entre outras.

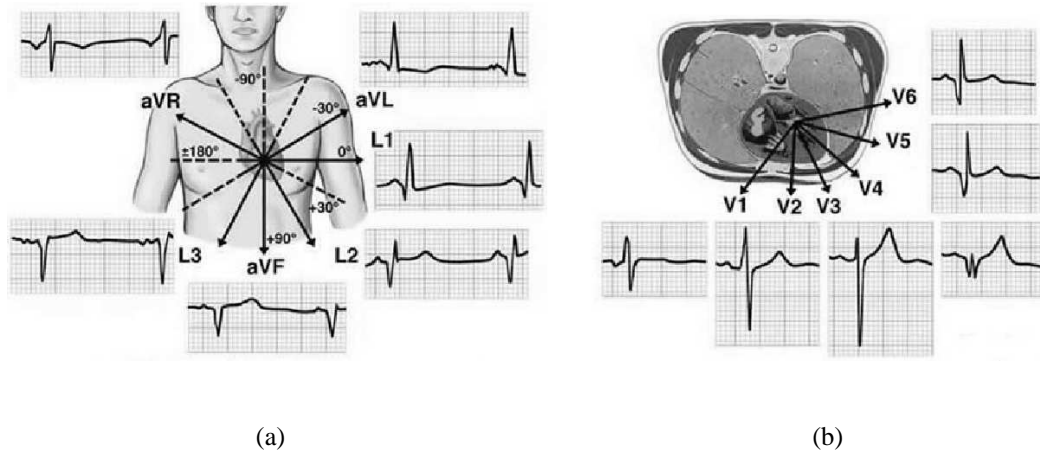


Figura 4.5: Exemplos de registros das 12 derivações: (a) derivações bipolares e aumentadas. (b) derivações precordiais.

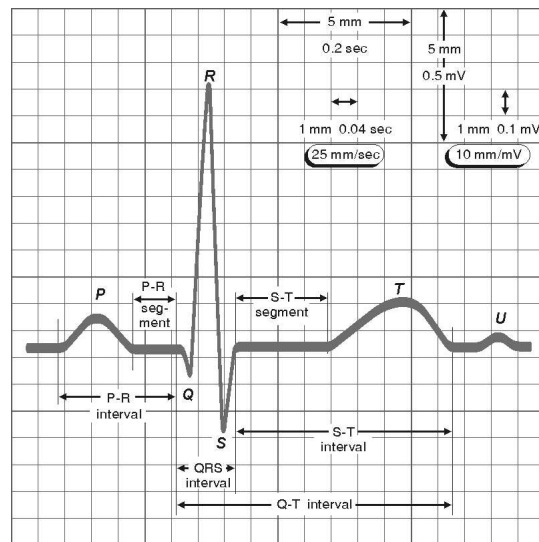


Figura 4.6: Estrutura do eletrocardiograma normal.

4.2 Compressão de Eletrocardiogramas

O uso contínuo de sistemas computadorizados para o processamento de eletrocardiogramas, assim como a necessidade de melhor desempenho e menor custo, têm cada vez mais exigido técnicas de compressão de ECG confiáveis, precisas e eficientes. A importância prática da compressão de ECG evidencia-se diante:

1. Do aumento da capacidade de armazenamento de ECGs como base de dados para verificação posterior;
2. Da possibilidade de transmitir ECG em tempo real, ou *off-line* para centros de

interpretação;

3. Da funcionalidade crescente dos monitores de ECG ambulatoriais.

O objetivo principal da compressão de eletrocardiograma, assim como de qualquer outra técnica de compressão, é a máxima redução do volume dos dados, mantendo as características morfológicas significativas do sinal após da sua reconstrução. Essa compressão de dados, além da diminuição de custo, proporciona a diminuição do peso e tamanho do equipamento

As técnicas utilizadas na compressão de eletrocardiograma podem ser classificadas, conforme a seção 2.2, como: compressão direta de dados (AZTEC, DPCM, TP, CORTES, SAPA, FAN), métodos de transformação (Fourier, Walsh, KLT) e técnicas de extração de parâmetros [JALALEDDINE et al (1990)].

Os métodos diretos de compressão de dados baseiam-se na utilização de algoritmos de predição e interpolação. Essas técnicas tentam reduzir as redundâncias presentes nos dados. As técnicas de transformação, de forma geral, envolvem pré-processamento do sinal de entrada por meio de transformações lineares ortogonais e a codificação apropriada do sinal transformado. Para reconstruir o sinal, é executado o processo inverso, o que resulta em um determinado grau de erro. As técnicas de extração de parâmetros são processos irreversíveis nos quais características particulares do sinal são extraídas para serem utilizadas em uma classificação baseada no conhecimento *a priori* das características do sinal.

As técnicas compressão direta de dados têm se mostrado mais eficientes em seu desempenho quando comparadas àquelas de transformação, tanto em relação ao tempo de processamento quanto a taxa de compressão. Alguns trabalhos, como por exemplo [BLANCO-VELASCO et al(2004), MIAOU & CHAO (2005) , OLMOS et al (1996), RAJOURB (2002)], com o intuito de melhorar seus resultados, fazem uso de métodos de codificação adicionais, como DPCM e Huffman.

Os esquemas de compressão de ECG que envolvem perdas normalmente reduzem o volume de dados de forma significativa. Entretanto, tais esquemas envolvem a perda de informações possivelmente úteis para diagnósticos. Assim, alguns trabalhos preferem abordar a compressão sem perdas, de forma a preservar fielmente as informações essenciais ao diagnóstico, ainda que dessa forma se tornem menos eficientes [MIAOU & CHAO (2005)].

Além da comparação visual, a maioria esquemas de compressão tem utilizado a diferença média quadrática percentual (PRD) como forma de avaliação do sinal reconstruído. O PRD é definido por

$$\text{PRD} = \sqrt{\frac{\sum_{i=1}^n [ecg_{orig}(i) - ecg_{rec}(i)]^2}{\sum_{i=1}^n ecg_{orig}^2(i)}} * 100 \quad (4.1)$$

onde ecg_{orig} é o sinal de ECG original, enquanto ecg_{rec} é o sinal recuperado após a descompressão. A Tabela 4.1 mostra a comparação de alguns métodos de compressão de ECG em relação ao CR, definido em (2.10), e o PRD.

Tabela 4.1: Comparação entre alguns esquemas de Compressão de ECG

Método	CR	PRD (%)
AZTEC	10.0	28.0
CORTES	4.8	7.0
DPCM	2.5	-
FAN/SAPA	3.0	4.0
Fourier	7.4	7
KLT	3.0	-
TP	2.0	5.3

Entretanto, o PRD é irrelevante do ponto de vista do diagnóstico, pois não revela se o algoritmo é capaz ou não de preservar características significativas do ECG [JALALEDDINE et al (1990)]. Por esse motivo outras formas de avaliação foram utilizadas, como o WDD (*weighted diagnostic distortion*) e o teste MOS (*mean opinion score*), conforme [ZIGEL et al (2000)].

4.2.1 Transformada de Karhunen-Loève

Em estatística, análise de componentes principais (PCA) é uma técnica que pode ser utilizada para simplificar um conjunto de dados. Formalmente, é uma transformação linear que escolhe um novo sistema de coordenadas para os dados de tal forma que a maior variância entre as projeções dos dados venha a recair sobre o primeiro eixo (chamado primeiro eixo principal), a segunda maior variância no segundo eixo e assim por diante.

O PCA também é chamado de transformada de Karhunen-Loève. Essa transformada tem a distinção de ser a transformação linear ótima para manter o subespaço que tem a maior variância. Entretanto, o custo computacional dessa transformação é muito grande. Ao contrário de outras transformações lineares, KLT não tem um número fixo de funções base.

Considere a combinação linear dos elementos de \mathbf{x}

$$y_1 = \sum_{k=1}^n w_{k1} x_k = \mathbf{w}_1^T \mathbf{x}. \quad (4.2)$$

Os termos w_{11}, \dots, w_{n1} são coeficientes escalares do vetor n -dimensional \mathbf{w}_1 . O fator y_1 é chamado de primeira componente principal de \mathbf{x} , caso a variância de y_1 seja a maior possível. Já que a variância depende tanto da norma quanto da orientação do vetor de coeficientes \mathbf{w}_1 e cresce sem limites à medida que a norma cresce, é imposta a restrição de que a norma de \mathbf{w}_1 seja igual a 1. Assim, procuramos o vetor de coeficientes \mathbf{w}_1 que maximize o critério

$$J_1^{PCA}(\mathbf{w}_1) = E[y_1^2] = E[(\mathbf{w}_1^T \mathbf{x})^2] \quad (4.3a)$$

$$= \mathbf{w}_1^T E[\mathbf{x}\mathbf{x}^T] \mathbf{w}_1 \quad (4.3b)$$

$$= \mathbf{w}_1^T \mathbf{C}_x \mathbf{w}_1, \quad (4.3c)$$

sendo que $\|\mathbf{w}_1\| = 1$ e definindo

$$\mathbf{C}_x = E[\mathbf{x}\mathbf{x}^T]. \quad (4.4)$$

A solução para o problema de PCA é dada em termos dos autovetores $\mathbf{v}_1, \dots, \mathbf{v}_n$ de \mathbf{C}_x , ordenados de maneira que os autovalores d_1, \dots, d_n satisfaçam $d_1 \geq d_2 \geq \dots \geq d_n$. Assim a solução de (4.3c) é dada por

$$\mathbf{w}_1 = \mathbf{v}_1. \quad (4.5)$$

Assim, a primeira componente principal de \mathbf{x} é $y_1 = \mathbf{v}_1^T \mathbf{x}$. O critério definido em (4.3) pode ser generalizado para m componentes, sendo m qualquer número entre 1 e n . Igualmente, para a m -ésima componente principal $y_m = \mathbf{w}_m^T \mathbf{x}$, temos como solução $\mathbf{w}_m = \mathbf{v}_m$ e, portanto, $y_m = \mathbf{v}_m^T \mathbf{x}$.

Seja um conjunto de m funções base ortonormais $\mathbf{w}_1, \dots, \mathbf{w}_m$ tal que o erro médio quadrático entre \mathbf{x} e sua projeção seja mínimo. Então

$$J_{MSE}^{PCA} = E \left[\left\| \mathbf{x} - \sum_{i=1}^m (\mathbf{w}_i^T \mathbf{x}) \mathbf{w}_i \right\|^2 \right]. \quad (4.6)$$

Devido à ortogonalidade dos vetores \mathbf{w}_i , (4.6) pode ser reescrito como

$$J_{MSE}^{PCA} = E [\|\mathbf{x}\|^2] - E \left[\sum_{i=1}^m (\mathbf{w}_i^T \mathbf{x})^2 \right] \quad (4.7a)$$

$$= \mathbf{tr}(\mathbf{C}_x) - \sum_{i=1}^m \mathbf{w}_i^T \mathbf{C}_x \mathbf{w}_i \quad (4.7b)$$

$$= \sum_{i=m+1}^n d_i, \quad (4.7c)$$

ou seja, o valor do MSE é igual à soma dos autovalores descartados $\mathbf{v}_{m+1}, \dots, \mathbf{v}_n$.

Uma importante aplicação de KLT é a compressão de dados, onde o vetor \mathbf{x} é o sinal original que é aproximado pela expansão truncada

$$\hat{\mathbf{x}} = \sum_{i=1}^m y_i \mathbf{e}_i. \quad (4.8)$$

A partir de (4.7c) sabemos que o erro diminui à medida que mais termos são incluídos (4.8), até que se torne nulo quando $m = n$ ou todas as componentes principais são incluídas. O problema consiste em escolher m de forma que haja um compromisso entre erro e taxa de compressão.

5 Compressão utilizando Análise de Componentes Independentes

5.1 Método

Suponha que um sinal de ECG $e(t)$ possa ser dividido em m janelas de comprimento fixo n . Suponha também que é possível encontrar, através de treinamento, um subespaço vetorial $\Phi(t) = [\varphi_1(t), \dots, \varphi_i(t)]$, definindo as colunas $\varphi_i(t)$ como funções base de $e(t)$. Seja a projeção da m -ésima janela de $e(t)$ em $\Phi(t)$ definida por [GUILHON (2005)]

$$\mathbf{e}_m(t) = \mathbf{w}_m^T \Phi(t), \quad (5.1)$$

sendo \mathbf{w}_m os vetores de coeficientes da projeção. Por motivo de simplicidade, a variável independente t será omitida daqui em diante. Assim, os coeficientes \mathbf{w}_m podem ser utilizados para encontrar uma versão estimada de \mathbf{e}_m , ou seja, $\hat{\mathbf{e}}_m$. Esses coeficientes devem ser calculados de forma que $[\mathbf{e}_m - \hat{\mathbf{e}}_m]^2, \forall m$, seja minimizado. Esses passos são ilustrados na Figura 5.1

Os coeficientes \mathbf{w}_m são, então, quantizados usando tantos níveis quantos necessários para manter um nível desejado de precisão. O sinal pode ser reconstruído, conforme (5.1), a partir de Φ e \mathbf{w}_m .

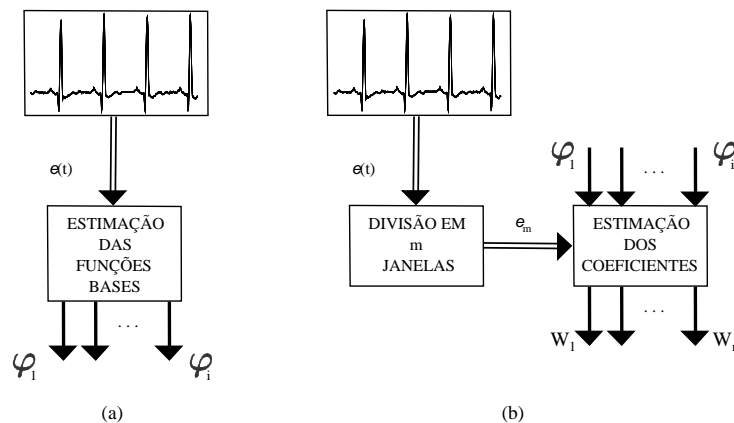


Figura 5.1: Diagrama de blocos para o algoritmo de compresso proposto: (a) fase de treinamento: estimaco das funes base. (b) fase de estimaco dos coeficientes.

5.1.1 Prova

Seja a informação mútua das variáveis aleatórias $\mathbf{e}_1, \dots, \mathbf{e}_m$ definida por

$$I(\mathbf{e}_1, \dots, \mathbf{e}_m) = \sum_{i=1}^m H(\mathbf{e}_i) - H(\mathbf{e}_1, \dots, \mathbf{e}_m), \quad (5.2)$$

sendo $H(\mathbf{e}_1, \dots, \mathbf{e}_m)$ a entropia conjunta de $\mathbf{e}_1, \dots, \mathbf{e}_m$. A informação mútua mede a dependência entre as variáveis. Sabendo que não é possível perder informação [HAYKIN (2001)], podemos afirmar que

$$I(\mathbf{e}_1, \dots, \mathbf{e}_m) \geq 0. \quad (5.3)$$

Substituindo (5.3) em (5.2), obtemos

$$\sum_{i=1}^m H(\mathbf{e}_i) \geq H(\mathbf{e}_1, \dots, \mathbf{e}_m). \quad (5.4)$$

Da mesma forma, seja a informação mútua das variáveis aleatórias $\mathbf{w}_1, \dots, \mathbf{w}_m$ definida por

$$I(\mathbf{w}_1, \dots, \mathbf{w}_m) = \sum_{i=1}^m H(\mathbf{w}_i) - H(\mathbf{w}_1, \dots, \mathbf{w}_m), \quad (5.5)$$

sendo $H(\mathbf{w}_1, \dots, \mathbf{w}_m)$ a entropia conjunta de $\mathbf{w}_1, \dots, \mathbf{w}_m$. Supondo que $\mathbf{w}_1, \dots, \mathbf{w}_m$ sejam independentes, podemos dizer que [HYVÄRINEN et al (2001)]

$$I(\mathbf{w}_1, \dots, \mathbf{w}_m) = 0. \quad (5.6)$$

Igualmente, substituindo (5.6) em (5.5), obtemos

$$\sum_{i=1}^m H(\mathbf{w}_i) = H(\mathbf{w}_1, \dots, \mathbf{w}_m). \quad (5.7)$$

Dada a transformação linear

$$\mathbf{e}_m = \Phi \mathbf{w}_m, \quad (5.8)$$

tem-se que [PAPOULIS & PILLAI (2002)]

$$H(\mathbf{e}_1, \dots, \mathbf{e}_m) = H(\mathbf{w}_1, \dots, \mathbf{w}_m). \quad (5.9)$$

Logo, a partir (5.4), (5.7) e (5.9), obtemos

$$\sum_{i=1}^m H(\mathbf{e}_i) \geq H(\mathbf{w}_1, \dots, \mathbf{w}_m) \quad (5.10a)$$

$$\implies \sum_{i=1}^m H(\mathbf{e}_i) \geq \sum_{i=1}^m H(\mathbf{w}_i). \quad (5.10b)$$

Assim, a partir de (2.9) e (5.10b), podemos concluir que

$$\sum_{i=1}^m \bar{L}_{min}(\mathbf{e}_i) \geq \sum_{i=1}^m \bar{L}_{min}(\mathbf{w}_i), \quad (5.11)$$

ou seja, o comprimento mínimo total de código necessário para representar as m janelas \mathbf{e}_m é maior que aquele necessário para representar os m coeficientes \mathbf{w}_m . Isso significa que, em média, os coeficientes da projeção em (5.1) ocupam menos espaço em memória que o sinal de ECG $e(t)$.

5.1.2 Estimação das Funções Bases através de ICA

Supondo que temos k variáveis aleatórias $\epsilon_1, \dots, \epsilon_k$ (segmentos aleatórios, de comprimento n , de um sinal de ECG), modelados como uma combinação linear de n variáveis aleatórias $\mathbf{a}_1, \dots, \mathbf{a}_n$, tais que

$$\epsilon_i = \varphi_{i1}\mathbf{a}_1 + \varphi_{i2}\mathbf{a}_2 + \dots + \varphi_{in}\mathbf{a}_n \quad i = 1, \dots, k, \quad (5.12)$$

sendo que φ_{in} são coeficientes reais. Definimos \mathfrak{E} , Φ e \mathbf{A} como

$$\mathfrak{E} = [\epsilon_1, \dots, \epsilon_k] \quad (5.13)$$

$$\Phi = \begin{bmatrix} \varphi_{11} & \dots & \varphi_{1n} \\ \vdots & & \vdots \\ \varphi_{k1} & \dots & \varphi_{kn} \end{bmatrix} \quad (5.14)$$

$$\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n]. \quad (5.15)$$

Usando (5.13)-(5.15), podemos reescrever (5.12) como

$$\mathfrak{E} = \Phi^T \mathbf{A}. \quad (5.16)$$

Através de análise de componentes independentes, podemos determinar Φ de forma que os \mathbf{a}_i sejam mutuamente estatisticamente independentes. Para tanto, o algoritmo utilizado foi o *FastICA*, descrito em [HYVÄRINEN et al (2001)] pelos seguintes passos:

1. Centralizar os dados para tornar sua média nula;

2. Branquear os dados, resultando em \mathbf{z} ;
3. Escolher p , o número de componentes independentes a estimar;
4. Escolher os valores iniciais para \mathbf{b}_i , $i = 1, \dots, p$, cada um de norma unitária. Ortogonalizar a matriz \mathbf{B} como no passo 6 abaixo;
5. Para cada $i = 1, \dots, p$, seja

$$\mathbf{b}_i \leftarrow E \{ \mathbf{z}g(\mathbf{b}_i^T \mathbf{z}) \} - E \{ g'(\mathbf{b}_i^T \mathbf{z}) \} \mathbf{w},$$

sendo g definido;

6. Ortogonalizar simetricamente a matriz $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_p)^T$ fazendo

$$\mathbf{B} \leftarrow (\mathbf{B}\mathbf{B}^T)^{-1/2}\mathbf{B}.$$

7. Caso o algoritmo não convirja, repetir o passo 5.

Para o algoritmo proposto, foi utilizado $p = n$. Após a estimação de \mathbf{B} , podemos facilmente obter Φ .

5.1.3 Estimação dos Coeficientes da Projeção

Para a estimação dos coeficientes \mathbf{w}_m é utilizado um combinador linear, cuja estrutura é mostrada na Figura 5.2. Os termos φ_i , \mathbf{e}_m , $\hat{\mathbf{e}}_m$ e \mathbf{w}_m correspondem, respectivamente, ao vetor de entrada, ao sinal desejado, ao sinal estimado e aos coeficientes da estimação. O sinal estimado $\hat{\mathbf{e}}_m$ pode ser expresso por

$$\hat{\mathbf{e}}_m = \sum_{j=0}^n w_{jm} \varphi_{ji} \quad (5.17a)$$

$$= \mathbf{w}_m^T \varphi_i \quad (5.17b)$$

$$= \varphi_i^T \mathbf{w}_m. \quad (5.17c)$$

Definindo o erro de estimação como sendo a diferença entre o sinal desejado e o sinal estimado, e considerando (5.17a)-(5.17c), obtemos

$$\boldsymbol{\varepsilon}_i = \mathbf{e}_i - \hat{\mathbf{e}}_i \quad (5.18a)$$

$$= \mathbf{e}_i - \varphi_i^T \mathbf{w} \quad (5.18b)$$

$$= \mathbf{e}_i - \mathbf{w}^T \varphi_i. \quad (5.18c)$$

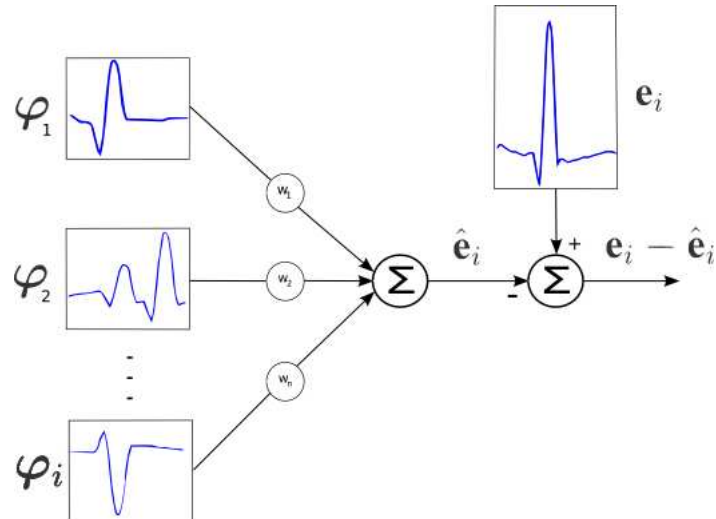


Figura 5.2: Combinador linear de múltiplas entradas.

Supondo ϵ_m , \mathbf{e}_m e φ_i como estatisticamente estacionários, tomamos o valor esperado de ϵ_m^2 , ou seja

$$E[\epsilon_m^2] = E[(\mathbf{e}_m - \varphi_i^T \mathbf{w}_m)^2] \quad (5.19a)$$

$$= E[(\mathbf{e}_m^2 - \mathbf{w}_m^T \varphi_i \varphi_i^T \mathbf{w}_m - 2\mathbf{e}_m \varphi_i^T \mathbf{w}_m)^2] \quad (5.19b)$$

$$= E[\mathbf{e}_m^2] - \mathbf{w}_m^T E[\varphi_i \varphi_i^T] \mathbf{w}_m - 2E[\mathbf{e}_m \varphi_i^T] \mathbf{w}_m. \quad (5.19c)$$

Podemos definir

$$\mathbf{R} = E[\varphi_i \varphi_i^T] \quad (5.20)$$

$$\mathbf{P} = E[\mathbf{e}_m \varphi_i], \quad (5.21)$$

sendo \mathbf{R} a matriz de autocorrelação de entrada e \mathbf{P} é a matriz de correlação cruzada entre a entrada e o sinal desejado. Assim, designando o erro médio quadrático em (5.19c) por ξ e reescrevendo essa equação a partir de (5.20) e (5.21)

$$\text{MSE} \triangleq \xi = E[\epsilon_m^2] \quad (5.22a)$$

$$= E[\mathbf{e}_m^2] - \mathbf{w}_m^T \mathbf{R} \mathbf{w}_m - 2\mathbf{P}^T \mathbf{w}_m. \quad (5.22b)$$

Desejamos encontrar a matriz de coeficientes $\mathbf{w}_m = [w_1, \dots, w_n]^T$ que resulte em um valor mínimo para o MSE. Uma forma de encontrar tais valores é através dos métodos gradientes. O gradiente do MSE, definido por $\nabla(\xi)$ pode ser obtido pela diferenciação de

(5.22b) [HAYKIN (1996), PRINCIPE et al (2000), WIDROW & STEARNS (1985)], ou seja

$$\nabla(\xi) \triangleq \frac{\partial \xi}{\partial \mathbf{w}_m} = \left[\frac{\partial \xi}{\partial w_1} \quad \frac{\partial \xi}{\partial w_2} \quad \dots \quad \frac{\partial \xi}{\partial w_n} \right]^T \quad (5.23a)$$

$$= 2\mathbf{R}\mathbf{w}_m - 2\mathbf{P}. \quad (5.23b)$$

O erro médio quadrático mínimo é obtido quando a matriz de coeficientes tem valor ótimo \mathbf{w}_m^* , onde o gradiente $\nabla(\xi)$ é igual zero:

$$\nabla(\xi) = 0 = 2\mathbf{R}\mathbf{w}_m^* - 2\mathbf{P} \quad (5.24a)$$

$$\mathbf{w}_m^* = \mathbf{R}^{-1}\mathbf{P}. \quad (5.24b)$$

Para tanto, assumimos que \mathbf{R} seja não-singular. A expressão (5.24b) é chamada de equação de Wiener-Hopf.

5.2 Material Utilizado

Foram utilizados os 15 primeiros registros das bases MIT-BIH Normal Sinus Rhythm Database (NSRD), da MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB) e da MIT-BIH Supraventricular Arrhythmia Database (SVDB) [9], conforme mostrado na Tabela 5.1.

A primeira base de dados inclui registros de pacientes que não apresentaram arritmia significativa, a segunda mostra exemplos de arritmias supraventriculares e a última, pacientes que tiveram episódios de taquicardia ventricular, *flutter* e fibrilação ventricular. Os registros da NSRD, SVDB e VFDB foram digitalizados utilizando conforme a Tabela 5.2.

5.3 Resultados

O algoritmo proposto foi testado para todos os 45 registros mostrados na Tabela 5.1. Para cada registro, os trinta minutos iniciais foram utilizados para a estimação das funções base, enquanto os dois minutos iniciais foram comprimidos usando o algoritmo [GUILHON (2005)].

Então foram calculados os valores de PRD, conforme definidos em (4.1), para valores fixos de CR, sendo este aqui redefinido como

$$CR = \frac{N_O S_O}{N_R S_R} \quad (5.25)$$

Tabela 5.1: Relação de registros utilizados das base de dados.

Registro	NSRD	SVDB	VFDB
1	16265	800	418
2	16272	801	419
3	16273	802	420
4	16420	803	421
5	16483	804	422
6	16539	805	423
7	16773	806	424
8	16786	807	425
9	16795	808	426
10	17052	809	427
11	17453	810	428
12	18177	811	429
13	18184	812	430
14	19088	820	602
15	19090	821	605

Tabela 5.2: Taxa de amostragem e resolução das bases de dados.

	NSRD	SVDB	VFDB
Amostragem (a/s)	128	128	250
Resolução (bits)	12	10	12

onde N e S representam o número de bits por amostra e a quantidade de amostras, respectivamente. Os índices O e R , por sua vez, indicam o sinal original e o sinal reconstruído. Já que S_O e S_R têm o mesmo valor, pois não são suprimidas amostras, CR depende apenas dos comprimentos das palavras código, ou seja

$$CR = \frac{N_O}{N_R} \quad (5.26)$$

Os valores de PRD do método proposto são mostrados nas Figuras 5.3 - 5.3, onde são comparados aos obtidos utilizando o método KLT [OLMOS et al (1996)]. Cada uma das

figuras exibe quatro pares de traçado (em vermelho, azul, verde e preto), correspondentes a quatro níveis diferentes de compressão, especificados nas legendas.

As Figuras 5.6 e 5.7 mostram os sinais originais, as suas reconstruções após o processo de compressão e os respectivos erros de reconstrução, para os registros NSRD - 16265 e SVDB - 809.

5.4 Discussão

Os resultados encontrados se referem aos registros listados na Tabela 5.1. Foram utilizados 30 minutos de cada registro para encontrar, através de ICA, um subespaço vetorial que caracterize o sinal, e então 2 minutos desse mesmo registro foram divididos em janelas de igual comprimento. Essas janelas foram projetadas no subespaço gerado, resultando em um conjunto de coeficientes dessas projeções.

Esses coeficientes foram quantizados, armazenados e então usados para reconstruir o sinal. As Figuras 5.3 - 5.3 mostram a avaliação do erro de reconstrução do algoritmo para níveis de quantização fixos. Observa-se que os valores de PRD tornaram-se menores à medida que a taxa de compressão diminuía. Isso evidencia a relação entre as perdas de quantização e o erro na reconstrução do sinal. Conseqüentemente, a obtenção de erros pequenos nos limita a CRs menores, como reflexo de um maior número de níveis de quantização.

O método proposto atua na da remoção de redundâncias. Quanto menos redundante os dados se tornam, mais eficiente é o código que o representa [HAYKIN (2001)]. O algoritmo de análise de componentes independentes remove a informação mútua do sinal de ECG de forma a encontrar o subespaço vetorial onde as projeções das componentes desse sinal sejam mutuamente independentes.

Portanto, qualquer sinal projetado nesse subespaço tem componentes (coeficientes) também independentes. Codificar tais sinais, dado que suas componentes são independentes, como um todo ou a partir de suas componentes individuais resulta no mesmo comprimento de código [HYVÄRINEN et al (2001)].

Os resultados alcançados neste trabalho se referem apenas à redução do número de bits necessários para representar os dados, devido à sua característica menos redundante.

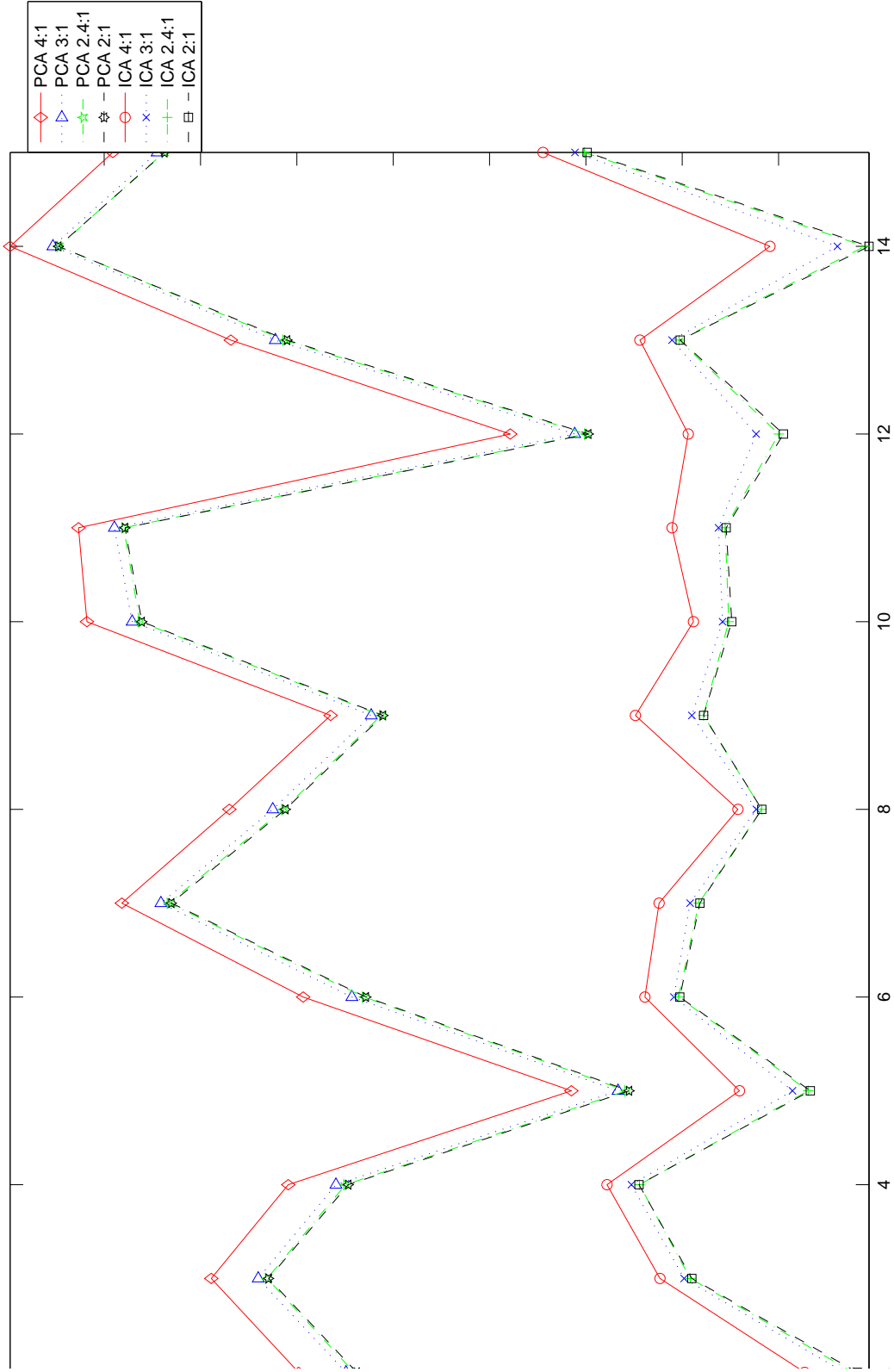


Figura 5.3: PRD após a reconstrução de 15 registros do MIT-BIH NSRD, segundo taxas de compressão fixas, obtidas utilizando PCA (KLT) e ICA. Observar que os traçados de PCA e ICA têm a mesma cor (vermelho, azul, verde ou preto) para taxas de compressão iguais.

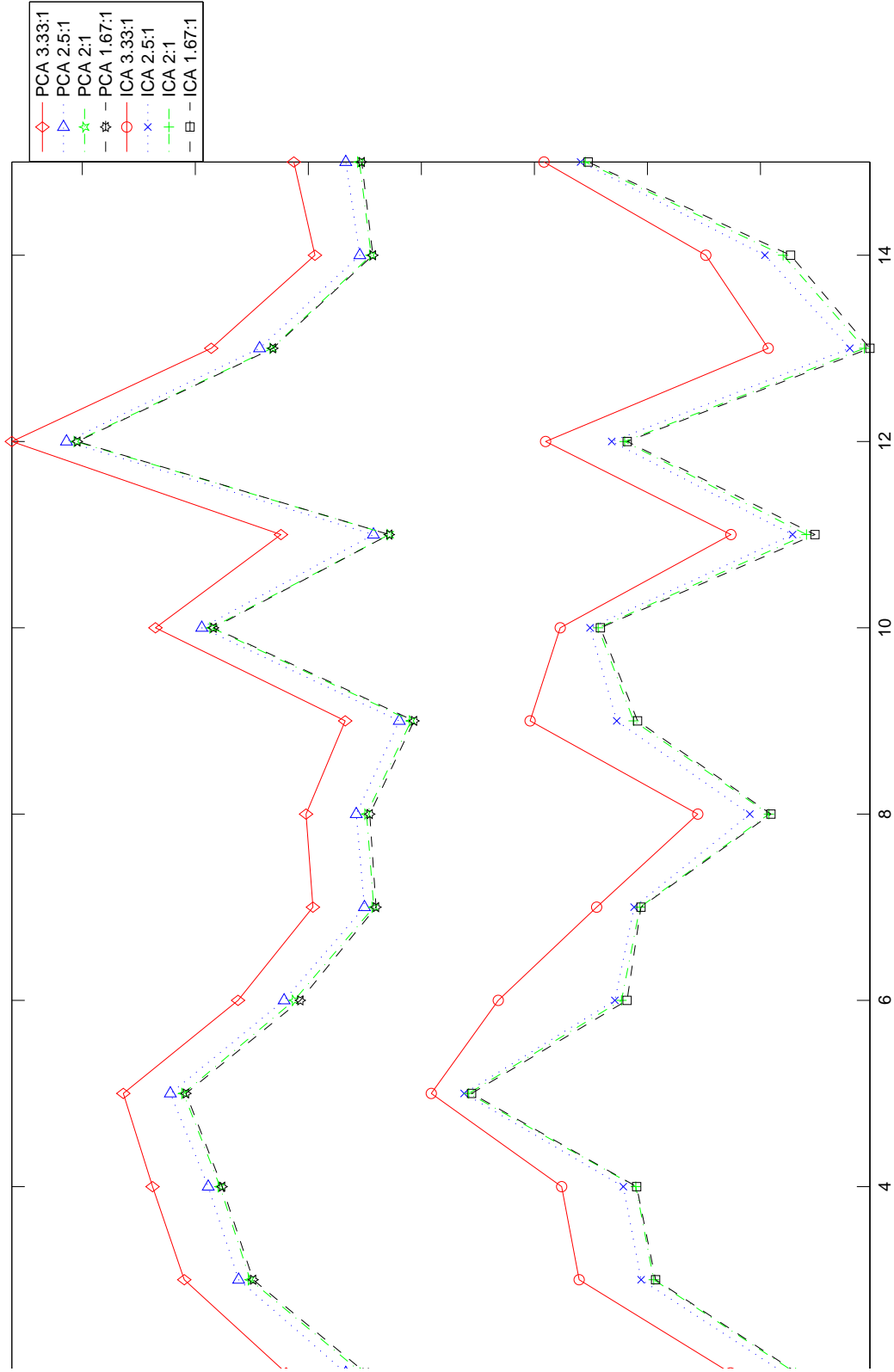


Figura 5.4: PRD após a reconstrução de 15 registros do MIT-BIH SVDB, segundo taxas de compressão fixas, obtidas utilizando PCA (KLT) e ICA. Observar que os traçados de PCA e ICA têm a mesma cor (vermelho, azul, verde ou preto) para taxas de compressão iguais.

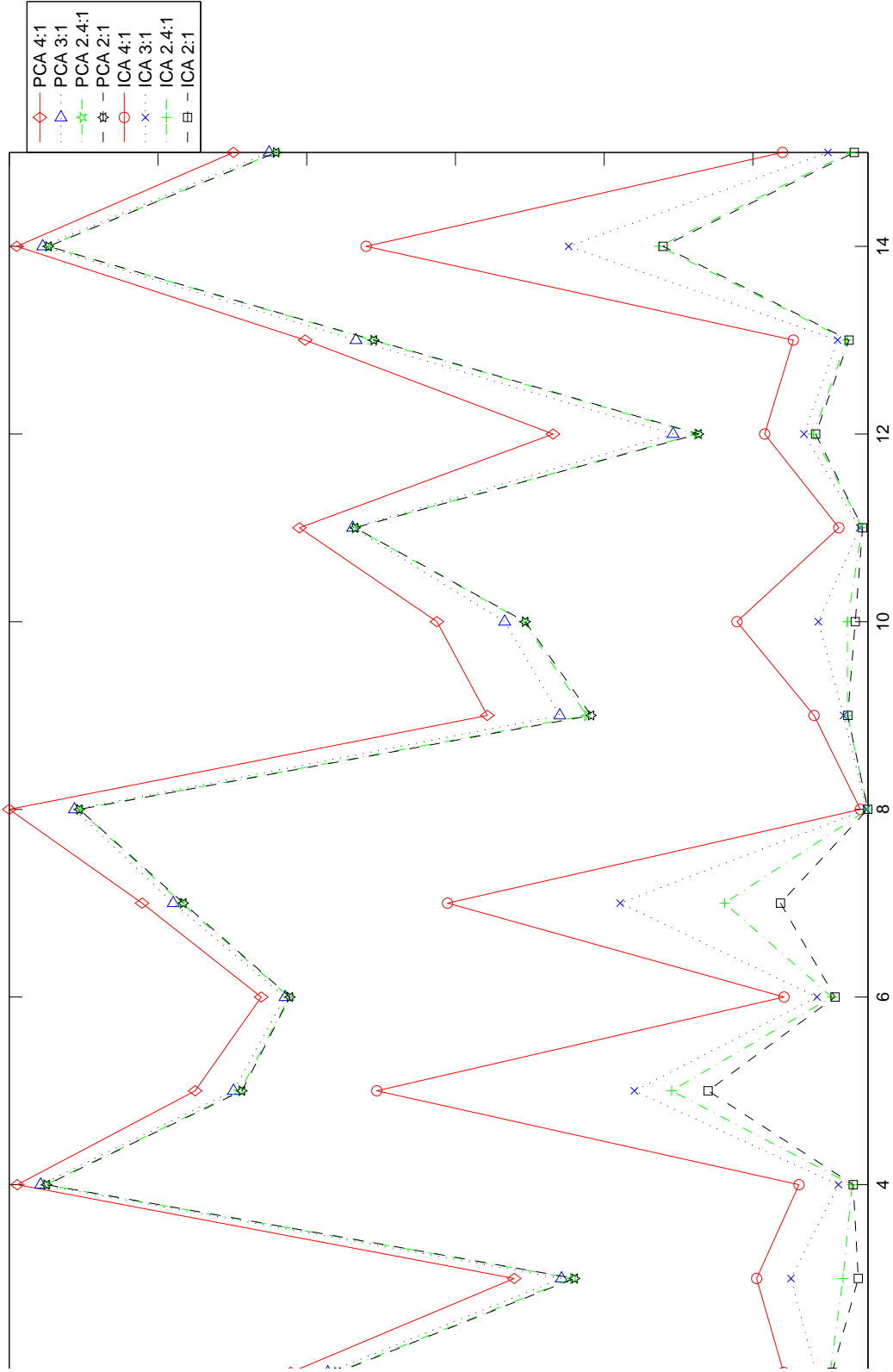


Figura 5.5: PRD após a reconstrução de 15 registros do MIT-BIH VFDB, segundo taxas de compressão fixas, obtidas utilizando PCA (KLT) e ICA. Observar que os traçados de PCA e ICA têm a mesma cor (vermelho, azul, verde ou preto) para taxas de compressão iguais.

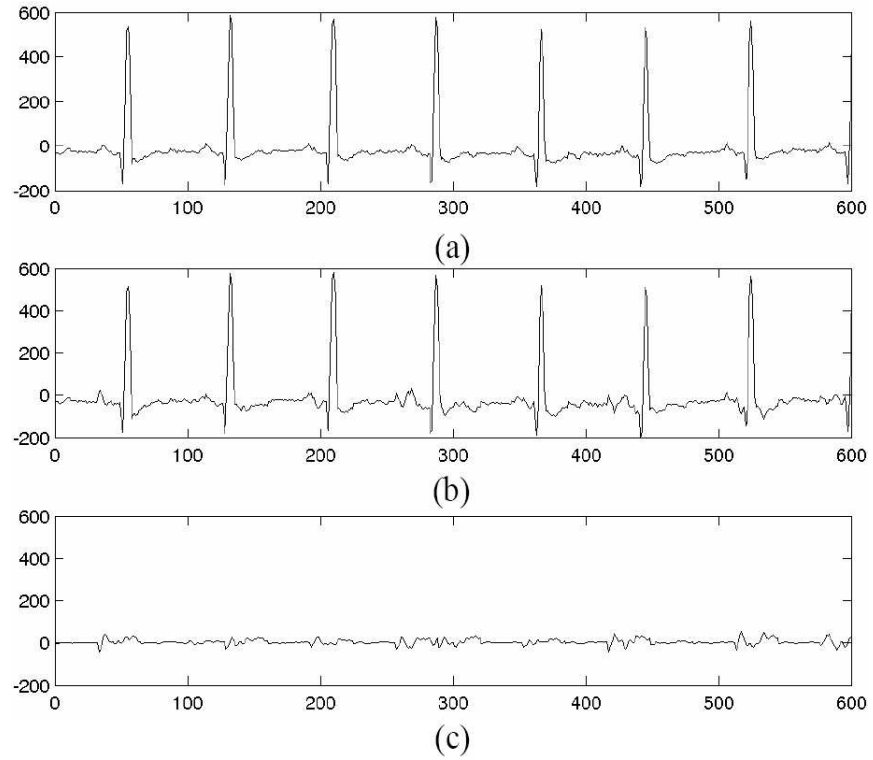


Figura 5.6: Resultados do algoritmo proposto para as 600 primeiras amostras para o registro 16265. (a) Sinal original. (b) Sinal obtido após a reconstrução, com $CR = 2.4:1$ e $PRD = 3.26\%$. (c) Erro de reconstrução.

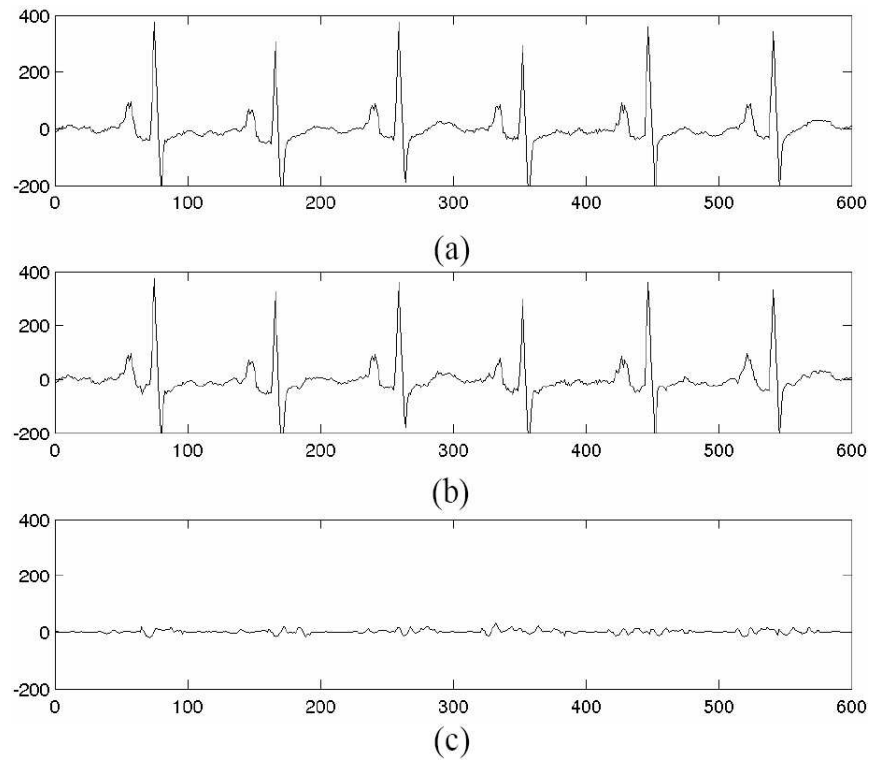


Figura 5.7: Resultados do algoritmo proposto para as 600 primeiras amostras para o registro SVDB 809. (a) Sinal original. (b) Sinal obtido após a reconstrução, com $CR = 2:1$ e $PRD = 4.85\%$. (c) Erro de reconstrução.

Nenhuma estratégia adicional foi empregada para melhorar os resultados [GUILHON (2005)], ao contrário dos trabalhos anteriormente mencionados [BLANCO-VELASCO et al(2004), MIAOU & CHAO (2005) , OLMOS et al (1996), RAJOURB (2002)], tendo sido usado apenas um processo de quantização simples. Assim, podemos avaliar a eficiência do algoritmo, o que não pode ser feito facilmente nos referidos trabalhos, já que estes não explicitam os resultados dos seus respectivos algoritmos sem o tratamento adicional.

Para efeito de comparação, o KLT [OLMOS et al (1996)] foi selecionado e simulado, obtendo resultados menos eficientes que o método proposto sob condições similares. Isso é mostrado nas Figuras 5.3 - 5.3, de onde podemos observar que os erros obtidos usando a estratégia baseada em ICA são até 20 vezes menores do que aqueles encontrados usando KLT.

6 Conclusões

Este trabalho propõe um método de compressão de eletrocardiogramas usando análise de componentes independentes. O método proposto gera uma nova representação do sinal, menos redundante e, por isso, de código mais eficiente. Isso torna possível a diminuição do espaço exigido para o seu armazenamento. Foram mostrados inicialmente os aspectos teóricos sobre os quais o trabalho se fundamenta, como teoria da informação, codificação eficiente, compressão, análise de componentes independentes e compressão de eletrocardiogramas.

Foram discutidos os resultados da compressão de 2 minutos de eletrocardiograma de 45 pacientes de três bases de dados diferentes (NSRD, SVDB e VFDB). Para 4 taxas de compressão fixas, foram calculados os PRDs de cada um dos 45 registros. Estes foram comparados àqueles PRDs obtidos quando da utilização do KLT como forma de construir novas representações dos sinais de ECG. Verificou-se que o método proposto gerou erros até vinte vezes menores que o KLT. Isso se reflete no erro de reconstrução do sinal comprimido utilizando ICA, não exibindo diferenças visuais significativas entre este e o sinal original. Isto indica as características morfológicas do eletrocardiograma foram preservadas.

A partir do que foi proposto no presente trabalho, maiores avanços podem ser vislumbrados no que se refere tanto à compressão de ECG quanto a de outras formas de dados. O algoritmo aqui mostrado pode ser adaptado para diferentes usos. Pode-se, por exemplo:

- Utilizar ICA como pré-processamento para um método de compressão clássico de ECG, buscando assim maiores CRs;
- Criar um critério de eliminação de componentes, com o intuito de reduzir o número de coeficientes a serem armazenados, melhorando a compressão;
- Submeter os resultados do algoritmo à análise de um cardiologista para avaliação MOS;
- Avaliar o desempenho do algoritmo para diferentes tipos de dados, como áudio, imagens, vídeos, etc;

- Adaptar partes do algoritmo para finalidades diversas, como reconhecimento de padrões, por exemplo.

Este trabalho deu origem ao artigo publicado nos anais do *2005 IEEE Signal Processing Society Workshop*, sob o título *ECG Data Compression by Independent Component Analysis*, conforme [GUILHON (2005)].

Referências

- [ASH (1990)] ASH, Robert B. **Information Theory**. Nova York: Dover Publications Inc. 1990. 352 p.
- [BLANCO-VELASCO et al(2004)] BLANCO-VELASCO, M.; CRUZ-ROLDÁN, F; GODINO-LLORENTE, J.I.; BARNER, K.E. ECG compression with retrieved quality guaranteed. **Electronics Letters**. 11 Nov. 2004. vol. 40, n. 23, pp 1466 – 1467.
- [CAVALCANTE et al (2006)] CAVALCANTE, André B.; MANDIC, Danilo P.; RUTKOWSKI, Tomasz; BARROS, Allan K. Speech Enhancement Based on the Response Features of Facilitated EI Neurons. **Proceedings of 6th International Conference on Independent Component Analysis and Blind Source Separation**, Março 2006, Charleston, South Carolina, USA.
- [COMON (1994)] COMON, Pierre. Independent component analysis, a new concept? **Signal Processing**. 36, 1994, 28-314.
- [DECCACHE & CASTRO (1993)] DECCACHE, Waldemar; CASTRO, Maria do C.V. **Eletrocardiograma: Semiótica e Clínica**. 1.ed. Rio de Janeiro: Revinter, 1993. 392 p.
- [GOLDBERGER (2000)] GOLDBERGER, Ary L.; AMARAL, Luís A.N.; GLASS, Leon; HAUSDORFF Jeffrey M.; IVANOV, Plamen Ch.; MARK Roger G.; MIETUS, Joseph E.; MOODY George B.; PENG, Chung-Kang; STANLEY H.Eugene. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220 [Circulation Electronic Pages; <http://circ.ahajournals.org/cgi/content/full/101/23/e215>]; 2000 (June 13).
- [GUILHON (2005)] GUILHON, Denner; BARROS, Allan K.; MEDEIROS, Eugênio. ECG Data Compression By Independent Component Analysis. In: 2005 IEEE International Workshop on Machine Learning For Signal Processing, 2005, Mystic, EUA. **Proceedings of the 2005 IEEE Signal Processing Society Workshop**, 2005. pp. 189-193.
- [HAYKIN (1996)] HAYKIN, Simon. **Adaptive Filter Theory**. 3.ed. Englewood Cliffs: Prentice-Hall. 1996. 989 p.
- [HAYKIN (2001)] HAYKIN, Simon. **Communication Systems**. 4.ed. Nova York: John Wiley & Sons. 2001. 816 p.
- [HYVÄRINEN & OJA (2000)] HYVÄRINEN, Aapo; OJA, Erkki. Independent Component Analysis: Algorithms and Applications. **Neural Networks**, 13(4-5):411-430, 2000.
- [HYVÄRINEN et al (2001)] HYVÄRINEN, Aapo; KARHUNEN, Juha; OJA, Erkki. **Independent Component Analysis**. Nova York: John Wiley & Sons. 2001. 481p.
- [JALALEDDINE et al (1990)] JALALEDDINE, Sateh M.S.; HUTCHENS, Chriswell G.; STRATTAN, Robert D; COBERLY, WilliamA. ECG Data Compression Techniques - A unified approach. **IEEE Transactions on Biomedical Engineering**. Abril 1990, vol. 37, no. 4, pp. 329–343.
- [KLABUNDE (2005)] KLABUNDE, Richard E. **Cardiovascular Physiology Concepts**. Disponível em <<http://www.cvphysiology.com/Arrhythmias/A009.htm>>. Acesso em: 29 de nov. de 2005.
- [KULICK (2005)] KULICK, Daniel L. **Electrocardiogram (ECG or EKG)**. Disponível em: <http://www.medicinenet.com/electrocardiogram_ecg_or_ekg/article.html>. Acesso em: 12 de dez. de 2005.

- [LELEWER & HIRSCHBERG (1987)] LELEWER, Debra A.; HIRSCHBERG, Daniel S. Data compression. **ACM Comput. Surv.** Set 1987, 19, 3, pp 261-296.
- [MALMIVUO & PLONSEY (1995)] MALMIVUO, Jaakko; PLONSEY, Robert. **Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields**. New York: Oxford University Press. 1995. Disponível em <<http://butler.cc.tut.fi/malmivuo/bem/bembook/>>. Acesso em: 14 de dez. de 2005.
- [MIAOU & CHAO (2005)] MIAOU, Shaou-Gang; CHAO, Shu-Nien. Wavelet-Based Lossy-to-Lossless ECG Compression in a Unified Vector Quantization Framework. **IEEE Transactions on Biomedical Engineering**. Março 2005. vol. 52, no. 3, pp. 539-543.
- [OLMOS et al (1996)] OLMOS, Salvador; MILLAN, Mar; GARCIA, José; LAGUNA, Pablo. ECG data compression with the Karhunen-Loève transform. **Computers in Cardiology**. 8-11 Set. 1996. pp. 253-256.
- [PAPOULIS & PILLAI (2002)] PAPOULIS, Athanasios; PILLAI, S. Unnikrishna. **Probability, Random Variables and Stochastic Processes**. 4 e.d. Nova York: McGraw-Hill. 2002. 852 p.
- [PRINCIPE et al (2000)] PRINCIPE, José C.; EULIANO, Neil R.; LEFEBVRE, W. Curt. **Neural and Adaptive Systems**. Nova York: John Wiley & Sons. 2000. 672 p.
- [RAJOUB (2002)] RAJOUB, B.A. An efficient coding algorithm for the compression of ECG signals using the wavelet transform. **IEEE Transactions on Biomedical Engineering**. Abril 2002. vol. 49, n. 4, pp 355 - 362.
- [SHANNON (1948)] SHANNON, Claude E. A Mathematical Theory of Communication. **The Bell System Technical Journal**. 1948, Vol. 27, p. 379-423, 623-656, Julho e Outubro.
- [THE BETTER HEALTH CHANNEL (2003)] THE BETTER HEALTH CHANNEL, **Electrocardiogram**. Disponível em <<http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Electrocardiogram?OpenDocument>>. Acesso em: 22 de dez. de 2005.
- [UNIFESP VIRTUAL (2003)] UNIFESP VIRTUAL. **Eletrocardiograma: Módulo 3**. Disponível em: <<http://www.virtual.epm.br/material/tis/curr-bio/trab2003/g5/menu.html>>. Acesso em: 11 de dez. de 2005.
- [VITERBI & OMURA (1979)] VITERBI, Andrew J.; OMURA, Jim K. **Principles of Digital Communication and Coding**. Tokyo: McGraw-Hill. 1979. 560 p.
- [WIDROW & STEARNS (1985)] WIDROW, Bernard; STEARNS, Samuel D. **Adaptive Signal Processing**. Nova Jersey: Prentice-Hall. 1985. 528 p.
- [WIKIPEDIA (2005a)] WIKIPEDIA. **Basis Functions**. Disponível em <http://en.wikipedia.org/wiki/Basis_functions>. Acesso em: 15 de jan. de 2006.
- [WIKIPEDIA (2005b)] WIKIPEDIA. **Electrocardiogram - Wikipedia, The Free Encyclopedia**. Disponível em <<http://en.wikipedia.org/wiki/Electrocardiogram>>. Acesso em: 14 de dez. de 2005.
- [WIKIPEDIA (2005c)] WIKIPEDIA. **Independent component analysis**. Disponível em <http://en.wikipedia.org/wiki/Independent_component_analysis>. Acesso em: 22 de dez. de 2005.
- [WIKIPEDIA (2005d)] WIKIPEDIA. **Information theory**. Disponível em <http://en.wikipedia.org/wiki/Information_theory>. Acesso em: 02 de jan. de 2006.
- [WIKIPEDIA (2005e)] WIKIPEDIA. **Principal components analysis**. Disponível em <http://en.wikipedia.org/wiki/Principal_components_analysis>. Acesso em: 03 de jan. de 2006.
- [ZIGEL et al (2000)] ZIGEL, Yaniv; COHEN, Arnon; KATZ, Amos. The Weighted Diagnostic Distortion (WDD) Measure for ECG Signal Compression. **IEEE Transactions on Biomedical Engineering**. Nov. 2000. vol. 47, no. 11, pp. 1422-1430.