

Investigação Sobre Aplicação de Amostragem Compressiva a Sinais de Áudio

Rubem José Vasconcelos de Medeiros

Dissertação de Mestrado submetida à Coordenação do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande - Campus de Campina Grande como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências no Domínio da Engenharia Elétrica.

Área de Concentração: Automação e Processamento Digital de
Sinais

João Maques de Carvalho, Dr.

Edmar Candeia Gurjão, D. Sc.

Orientadores

Campina Grande, Paraíba, Brasil

©Rubem José Vasconcelos de Medeiros, Maio de 2010

Investigação Sobre Aplicação de Amostragem
Compressiva a Sinais de Áudio

Rubem José Vasconcelos de Medeiros

Dissertação de Mestrado apresentada em Maio de 2010

João Maques de Carvalho, Dr.
Edmar Candeia Gurjão, D. Sc.
Orientadores

Campina Grande, Paraíba, Brasil, Maio de 2010

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

M488i

Medeiros, Rubem José Vasconcelos

Investigação sobre aplicação de amostragem compressiva a sinais de áudio / Rubem José Vasconcelos Medeiros. — Campina Grande, 2010.

42 f : il. color.

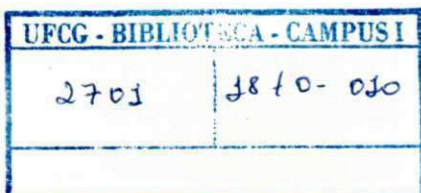
Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

Referências.

Orientadora: Prof. Dr. João Marques de Carvalho.

1. Amostragem Compressiva
2. Compressão de Áudio
3. Amostragem I. Título.

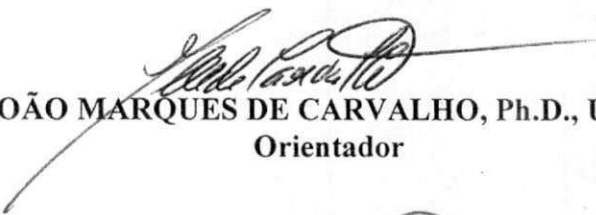
CDU 519.243 (043)



**INVESTIGAÇÃO SOBRE APLICAÇÃO DE AMOSTRAGEM COMPRESSIVA
A SINAIS DE ÁUDIO**

RUBEM JOSÉ VASCONCELOS DE MEDEIROS

Dissertação Aprovada em 18.06.2010


JOÃO MARQUES DE CARVALHO, Ph.D., UFCG
Orientador


EDMAR CANDEIA GURJÃO, D.Sc., UFCG
Orientador


FRANCISCO MADEIRO BERNARDINO JUNIOR, D.Sc., UNICAP
Componente da Banca


FRANCISCO MARCOS DE ASSIS, Dr., UFCG
Componente da Banca


BENEDITO GUIMARÃES AGUIAR NETO, Dr.-Ing., UFCG
Componente da Banca

CAMPINA GRANDE - PB
JUNHO - 2010

Dedicatória

Dedico este trabalho a minha família.

Resumo

Amostragem Compressiva é um recente campo de estudo que deu origem a um conjunto de ferramentas matemáticas que exploram esparsidade de um sinal em algum domínio para realizar amostragem abaixo da taxa de Nyquist sem distorção ou com distorção controlada. Este trabalho propõe o uso desta nova técnica para a compressão de sinais de áudio antes mesmo de passar para o domínio digital. Foram comparados e avaliados diferentes tipos de base de esparsidade baseadas em transformadas e de amostragem sob a medida objetiva de avaliação da qualidade perceptual PEAQ. O sistema proposto é baseado na Transformada Discreta do Cosseno Modificada como domínio de esparsidade e a reconstrução do sinal por minimização da norma l_1 . O sistema proposto tem como características baixa complexidade e universalidade no codificador, que pode ser implementado por *hardware*, tornando o mesmo atrativo para aplicações com limitação de *hardware*. Tal sistema foi comparado com outros mecanismos de compressão.

Palavras-chave: Amostragem Compressiva, Compressão de Áudio, Amostragem.

Abstract

Compressive sampling is a new framework that exploits sparsity of a signal in a transform domain to perform sampling below the Nyquist rate. In this work, we propose and study a new technique for lossy audio compression of real, non-sparse audio signals. The method is based on the Modified Discrete Cosine Transform as a sparse basis and on l_1 norm optimization for signal reconstruction. We evaluate audio compression perceptual quality with the objective measure PEAQ. The proposed method has low computational complexity, universality and can be implemented in hardware on the encoder side, making it attractive to applications with hardware limitations as wireless sensor networks.

Keywords: Compressed Sensing, Audio Compression, Sampling

Índice

1	Introdução	1
1.1	Objetivos	3
1.2	Organização do texto	3
2	Conceitos e Definições	4
2.1	Introdução	4
2.2	Representações de Sinais Discretos	4
2.3	Base Vetorial	5
2.4	Representação de um Vetor em uma Base	6
2.5	Normas de um vetor	7
2.6	Ortogonalidade e Normalidade de uma Matriz	8
2.7	Esparsidade de um Vetor	8
2.8	Incoerência entre Bases	10
2.9	Amostragem Discreta	12
2.10	Amostragem de Shannon/Nyquist	13
2.11	Conclusão	14
3	Amostragem Compressiva	15
3.1	Introdução	15
3.2	Amostragem Aleatória	15
3.3	Reconstrução	16
3.4	Teoremas da Amostragem	17
3.5	Conclusão	18
4	Revisão Bibliográfica	20
4.1	Introdução	20
4.2	Amostragem Clássica e Amostragem Compressiva	20
4.3	Aplicações de Amostragem Compressiva em Áreas Afins	21
4.4	Aplicações de Amostragem Compressiva em Voz e Áudio	23
4.5	Investigação de Base de Esparsidade em Áudio	23

4.6	Medidas de Avaliação Objetiva de Qualidade Perceptual de Áudio	24
5	Amostragem Compressiva Aplicada a Sinais de Áudio	25
5.1	Investigação de Base de Esparsidade para Áudio	25
5.2	Investigação de Técnicas de Reconstrução	30
5.3	Proposta de Sistema de Amostragem e Compressão	31
5.3.1	Codificador	31
5.3.2	Decodificador	31
5.3.3	Propriedades	32
5.4	Qualidade Perceptual do Sistema Proposto	33
5.5	Conclusão	34
6	Considerações Finais	37
6.1	Trabalhos futuros	38
	Referências Bibliográficas	39

Lista de Tabelas

5.1	<i>Interpretação do valor da Qualidade Perceptual PEAQ</i>	27
-----	--	----

Lista de Figuras

2.1	Representação Gráfica de $y[n]$	5
2.2	Representação Gráfica de $y[n]$ do exemplo 5	9
2.3	Representação Gráfica de $Y[n]$ do exemplo 5	10
5.1	Qualidade Perceptual versus Razão de Esparsidade S/N com N=128. Comparação entre as bases de esparsidade DFT, DCT e MDCT.	27
5.2	Gráfico da Qualidade Perceptual versus Razão de Esparsidade S/N com N=256. Comparação entre as bases de esparsidade DFT, DCT e MDCT.	28
5.3	Gráfico da Qualidade Perceptual versus a Razão de Esparsidade S/N com N=512. Comparação entre as bases de esparsidade DCT e MDCT.	28
5.4	Gráfico da Qualidade Perceptual versus Razão de Esparsidade S/N para a MDCT. Comparação entre tamanhos de bloco	29
5.5	Capacidade de Reconstrução do sinal por minimização l_1	30
5.6	Qualidade Perceptual do Sistema para N=128	33
5.7	Qualidade Perceptual do Sistema para N=256	34
5.8	Qualidade Perceptual do Sistema para N=512	35
5.9	Qualidade Perceptual do Sistema para N=128, 256 e 512	35
5.10	Comparação entre CODECs comuns e ACS	36

Capítulo 1

Introdução

Em 1949, Claude Shannon provou um teorema no qual estabeleceu que um sinal periódico e limitado em banda, se amostrado em uma taxa maior ou igual à sua maior frequência, pode ser recuperado a partir das amostras por um filtro passa-baixas ideal [1]. Essa taxa de amostragem ficou conhecida como a taxa de Nyquist, em homenagem a Harry Nyquist, por ter este proposto a mesma afirmação no seu estudo sobre transmissões telegráficas publicado em 1928 [2]. O teorema da amostragem de Shannon/Nyquist tem sido a base para a amostragem de todos os tipos de sinais, tais como áudio, vídeo, rádio frequência, ressonância magnética e sinais biométricos.

Com o avanço das tecnologias digitais, um importante conjunto de ferramentas tornou-se possível: o Processamento Digital de Sinais (do inglês *Digital Signal Processing* ou DSP). A partir deste ponto, qualquer tipo de sinal poderia ser processado digitalmente. Geralmente, a primeira etapa de um sistema de DSP é um conversor analógico-digital¹ (do inglês *Analog-to-Digital Converter* ou ADC), que é regido pelo teorema da amostragem de Shannon/Nyquist.

Entretanto, para algumas aplicações, o ADC se mostra inadequado. Por exemplo, a área de comunicações em banda larga pode chegar a trabalhar com frequências da ordem de GHz. Não existe ainda ADC que trabalhe em frequências tão elevadas e, portanto, para tal faixa de frequências, não se pode aplicar diretamente as técnicas de DSP sem antes demodular o sinal para uma frequência suportável pelo ADC. Outro cenário de inadequação é quando há limitação de *hardware* ou de banda de comunicação. O ADC converte muitos dados, que precisam ser armazenados em memórias para posterior processamento. Porém, por questões de limitação de banda ou memória, a quantidade de amostras transmitidas pode ser menor que a obtida na amostragem. O ideal seria converter os dados analógicos diretamente para a informação útil à aplicação. Um cenário de limitação de *hardware* e

¹As técnicas DSP podem ser utilizadas diretamente em sinais digitais que não foram provenientes de um ADC.

de banda ocorre no caso de redes de sensores sem fio (do inglês *Wireless Sensor Array* ou WSA) para medição de áudio. Tais redes são utilizadas, por exemplo, para monitorar eventos em ecossistemas para estudos biológicos nos quais a presença física humana poderia interferir nas medições como a chegada de aves migratórias, primeiros cantos de acasalamento de anfíbios, sensoreamento biológico [3].

O áudio digital está presente já há algum tempo no cotidiano das pessoas, através dos diversos tipos de mídias digitais. Por trás de toda essa facilidade, existe muita pesquisa e desenvolvimento, tanto por parte da academia como pela indústria. Um das linhas de pesquisa sobre áudio se concentra na busca da redução da quantidade de bits de informação utilizada no armazenamento ou transmissão do áudio, mantendo-se alguns parâmetros controláveis, como sua qualidade perceptual. A compressão de áudio é o ato de reduzir o espaço ocupado por dados num determinado dispositivo ou, de forma análoga, a redução da banda necessária para a sua transmissão por um canal de comunicação. Uma compressão pode ser realizada com ou sem perda de informação². Diz-se que a compressão é sem perda quando a informação obtida após a descompressão é igual à original. Já a compressão com perda permite que o sinal descomprimido apresente algum erro em relação ao original.

Por outro lado, os estudos em amostragem de sinais estão em contínuo progresso. Procurou-se sempre aumentar a taxa de amostragem, a resolução e o barateamento do circuito de amostragem. De uma das linhas de pesquisa em amostragem, surgiu um novo e vasto campo de estudos chamado de Amostragem Compressiva (do inglês *Compressed Sampling*, *Compressed Sensing* ou *Compressive Sensing* ou CS), cuja idéia central é unir as etapas de amostragem e compressão em uma única etapa. A principal conclusão dos estudos nessa área é que é possível fazer uma amostragem a uma taxa menor que a taxa de Nyquist, sem perda de informação. Tal afirmação não viola o teorema de Shannon/Nyquist, apenas amplia os horizontes para uma teoria da amostragem mais geral. Esse campo de estudo vem gerando repercussões não só na área de amostragem de sinais, mas também em outras áreas do conhecimento, a exemplo de processamento de sinais, estatística, teoria da informação, teoria da codificação e ciência da computação. Da Amostragem Compressiva, surgiram importantes ferramentas matemáticas e, mais recentemente, arquiteturas de projeto que estão propiciando novas aplicações.

No entanto, pouca pesquisa vem sendo realizada na aplicação dessas novas técnicas a sinais de áudio. O foco desta dissertação é justamente o estudo da aplicação de CS para tais sinais.

²Encontra-se também na literatura o termo “compactação”, usado no sentido de compressão sem perdas, e “compressão”, usado no sentido de compressão com perdas.

1.1 Objetivos

O principal objetivo deste trabalho é apresentar um estudo sobre a utilização das técnicas de Amostragem Compressiva aplicada à compressão de sinais de áudio. São avaliadas bases de esparsidade para áudio e técnicas de reconstrução do sinal. Por fim, propõe-se um sistema de compressão de áudio e avalia-se a qualidade perceptual do mesmo.

1.2 Organização do texto

A dissertação está organizada da seguinte forma. O Capítulo 2 apresenta uma revisão de alguns conceitos e definições chaves para uma melhor compreensão de Amostragem Compressiva. No Capítulo 3, há uma descrição da técnica de Amostragem Compressiva. No Capítulo 4 mostra-se a revisão bibliográfica realizada para o desenvolvimento desta dissertação. O capítulo 5 apresenta os resultados da investigação da aplicação de CS para compressão de áudio. Por último, no Capítulo 6, são apresentadas as considerações finais.

Capítulo 2

Conceitos e Definições

2.1 Introdução

Neste capítulo será feita uma revisão de alguns conceitos e definições essenciais para o bom entendimento de Amostragem Compressiva. Cada conceito ou definição é apresentado com certo direcionamento aos conceitos de Amostragem Compressiva. São revisados os seguintes tópicos:

- Representações de Sinais Discretos
- Base Vetorial
- Representação de um Vetor em uma Base
- Normas de um Vetor
- Ortogonalidade e Normalidade de uma Matriz
- Esparsidade de um Vetor
- Incoerência entre Bases
- Amostragem Discreta
- Amostragem de Shannon/Nyquist

2.2 Representações de Sinais Discretos

Um sinal discreto de dimensão N pode ser representado de diversas maneiras, dentre elas estão uma sequência numérica, uma função, um vetor e um gráfico. Cada representação ressalta aspectos diferentes do sinal, permitindo, assim, a observação do mesmo objeto por várias ópticas.

Exemplo 1 *Seqüência Numérica:* $\{1, 0, -1, 0, 1, 0, -1, 0\}$

Representação funcional: $y[n] = \text{sen}(\pi n/2)$, com $1 \leq n \leq 8$, $n \in \mathbb{N}$

Representação vetorial: $\mathbf{y} = [1 \ 0 \ -1 \ 0 \ 1 \ 0 \ -1 \ 0]^T$

Representação gráfica:

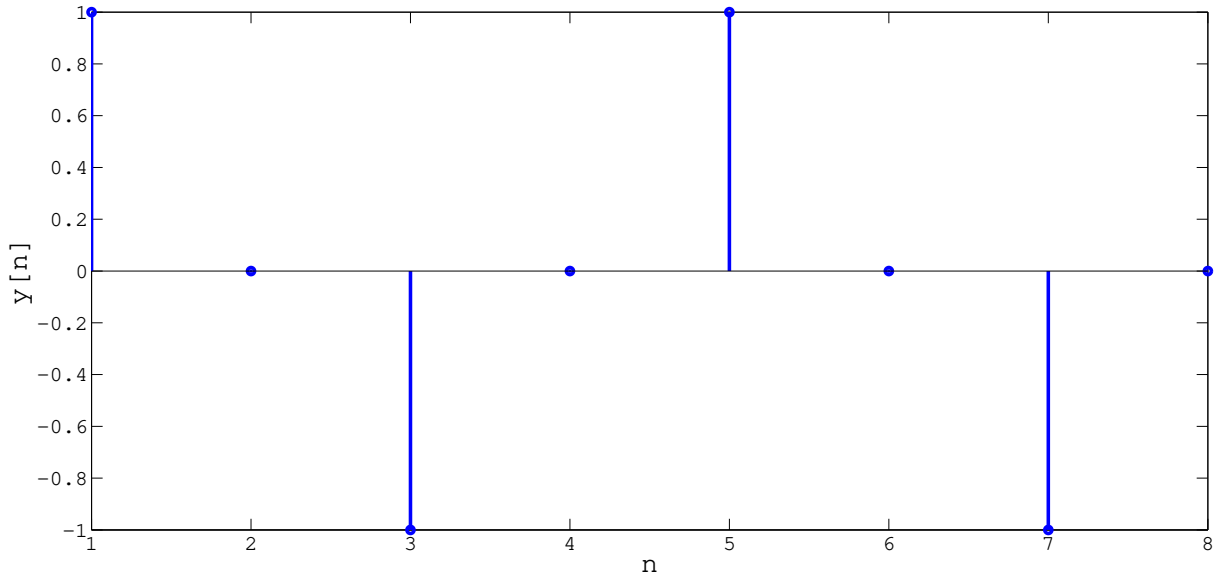


Figura 2.1: Representação Gráfica de $y[n]$.

2.3 Base Vetorial

Uma base vetorial Φ é o conjunto de vetores $\{\phi_1, \phi_2, \dots, \phi_N\}$, cada um com N componentes e linearmente independentes, que geram qualquer vetor de um espaço vetorial V de dimensão N . Da mesma forma que vetores, as bases podem ter diferentes representações.

Exemplo 2 *Considere os vetores ϕ_i :*

$$\phi_1 = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$$

$$\phi_2 = [0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$$

$$\phi_3 = [0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0]^T$$

$$\phi_4 = [0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0]^T$$

$$\phi_5 = [0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0]^T$$

$$\phi_6 = [0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0]^T$$

$$\phi_7 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0]^T$$

$$\phi_8 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1]^T$$

O conjunto $\Phi = \{\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7, \phi_8\}$ é uma base que gera o espaço vetorial \mathbb{R}^8 . Esta base também pode ser representada através de uma matriz, em que a coluna i

corresponde aos valores do vetor ϕ_i .

$$\Phi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

2.4 Representação de um Vetor em uma Base

O vetor \mathbf{x} na equação a seguir é a representação de \mathbf{y} na base $\Psi = \{\psi_1, \psi_2, \dots, \psi_N\}$:

$$y[k] = \sum_{i=1}^N x[i] \psi_i[k] \quad (2.1)$$

Ou na forma matricial $\mathbf{y} = \Psi \mathbf{x}$:

$$\begin{bmatrix} y[0] \\ y[1] \\ \dots \\ y[N-1] \end{bmatrix} = \begin{bmatrix} \psi_1[0] & \psi_2[0] & \dots & \psi_N[0] \\ \psi_1[1] & \psi_2[1] & \dots & \psi_N[1] \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1[N-1] & \psi_2[N-1] & \dots & \psi_N[N-1] \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ \dots \\ x[N-1] \end{bmatrix}$$

Exemplo 3 Vamos representar o vetor \mathbf{y} do exemplo 1 na base de Fourier $\Psi = \{\psi_1, \psi_2, \dots, \psi_8\}$, em que $\psi_i[k] = \frac{1}{\sqrt{N}} e^{2\pi j k i / N}$. Assim $\psi_i = (\psi_i[0], \psi_i[1], \psi_i[2], \psi_i[3], \psi_i[4], \psi_i[5], \psi_i[6], \psi_i[7])$

Fazendo $A = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}j$ tem-se

$$\begin{aligned} \psi_1 &= \frac{1}{\sqrt{8}} [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]^T & \psi_5 &= \frac{1}{\sqrt{8}} [1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 1 \ -1]^T \\ \psi_2 &= \frac{1}{\sqrt{8}} [1 \ \bar{A} \ -j \ -A \ -1 \ -\bar{A} \ j \ A]^T & \psi_6 &= \frac{1}{\sqrt{8}} [1 \ -\bar{A} \ -j \ A \ -1 \ -\bar{A} \ j \ -A]^T \\ \psi_3 &= \frac{1}{\sqrt{8}} [1 \ -j \ -1 \ j \ 1 \ -j \ -1 \ j]^T & \psi_7 &= \frac{1}{\sqrt{8}} [1 \ j \ -1 \ -j \ 1 \ j \ -1 \ -j]^T \\ \psi_4 &= \frac{1}{\sqrt{8}} [1 \ -A \ j \ \bar{A} \ -1 \ A \ -j \ -\bar{A}]^T & \psi_8 &= \frac{1}{\sqrt{8}} [1 \ A \ j \ -\bar{A} \ -1 \ -A \ -j \ \bar{A}]^T \end{aligned}$$

De forma matricial temos

$$\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \bar{A} & -j & -A & -1 & -\bar{A} & j & A \\ 1 & -j & -1 & j & 1 & -j & -1 & j \\ 1 & -A & j & \bar{A} & -1 & A & -j & -\bar{A} \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -\bar{A} & -j & A & -1 & \bar{A} & j & -A \\ 1 & j & -1 & -j & 1 & j & -1 & -j \\ 1 & A & j & -\bar{A} & -1 & -A & -j & \bar{A} \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ x[2] \\ x[3] \\ x[4] \\ x[5] \\ x[6] \\ x[7] \end{bmatrix}$$

Donde pode-se obter $\mathbf{x} = [0 \ 0 \ \sqrt{2} \ 0 \ 0 \ 0 \ \sqrt{2} \ 0]^t$.

2.5 Normas de um vetor

Existem algumas normas (métricas ou medidas) que podem ser aplicadas a um vetor \mathbf{x} :

Norma l_0 :

Considere a função $I[k] = \begin{cases} 0, & \text{se } x[k] = 0 \\ 1, & \text{se } x[k] \neq 0 \end{cases}$

define-se a norma l_0 como $\|\mathbf{x}\|_0 = \sum_{i=0}^{N-1} I[i]$

Em outras palavras, a norma l_0 é o número de elementos não-nulos de \mathbf{x} .

Norma l_p :

$$\|\mathbf{x}\|_p = \left(\sum_{i=0}^{N-1} |x[i]|^p \right)^{1/p}, \text{ sendo } p \geq 1$$

As normas l_p mais utilizadas são quando $p = 1$ e $p = 2$

$\|\mathbf{x}\|_1 = \sum_{i=0}^{N-1} |x[i]|$. Em outras palavras, a norma l_1 é o somatório dos valores absolutos dos elementos de \mathbf{x} .

$$\|\mathbf{x}\|_2 = \left(\sum_{i=0}^{N-1} |x[i]|^2 \right)^{1/2}. \text{ A norma } l_2 \text{ é também chamada de norma Euclidiana.}$$

Norma Infinita:

$\|\mathbf{x}\|_\infty = \max(|\mathbf{x}|)$, ou seja, o elemento de maior módulo de \mathbf{x} .

Exemplo 4 Vamos calcular as normas dos vetores a) \mathbf{x} e b) \mathbf{y} do exemplo 3.

$$a) \mathbf{x} = \begin{bmatrix} 0 & 0 & \sqrt{2} & 0 & 0 & 0 & \sqrt{2} & 0 \end{bmatrix}^T$$

$$\|\mathbf{x}\|_0 = 0 + 0 + 1 + 0 + 0 + 0 + 1 + 0 = 2$$

$$\|\mathbf{x}\|_1 = 0 + 0 + \sqrt{2} + 0 + 0 + 0 + \sqrt{2} + 0 = 2\sqrt{2}$$

$$\|\mathbf{x}\|_2 = \sqrt{0^2 + 0^2 + (\sqrt{2})^2 + 0^2 + 0^2 + 0^2 + (\sqrt{2})^2 + 0^2} = 2$$

$$\|\mathbf{x}\|_\infty = \max(0, 0, \sqrt{2}, 0, 0, 0, \sqrt{2}, 0) = \sqrt{2}$$

$$b) \mathbf{x} = \begin{bmatrix} 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}^T$$

$$\|\mathbf{x}\|_0 = 1 + 0 + 1 + 0 + 1 + 0 + 1 + 0 = 4$$

$$\|\mathbf{x}\|_1 = 1 + 0 + 1 + 0 + 1 + 0 + 1 + 0 = 4$$

$$\|\mathbf{x}\|_2 = \sqrt{1^2 + 0^2 + (-1)^2 + 0^2 + 1^2 + 0^2 + (-1)^2 + 0^2} = 2$$

$$\|\mathbf{x}\|_\infty = \max(1, 0, 1, 0, 1, 0, 1, 0) = 1$$

2.6 Ortogonalidade e Normalidade de uma Matriz

A matriz $\Psi = \begin{bmatrix} \psi_1 & \psi_2 & \dots & \psi_N \end{bmatrix}$ será:

- normal se $\|\psi_i\|_2 = 1 \forall i \in \{1, 2, \dots, N\}$
- ortogonal se $\psi_i \psi_j^t = 0 \forall i, j \in \{1, 2, \dots, N\}$
- ortonormal se for normal e ortogonal

2.7 Esparsidade de um Vetor

Definição 1 A esparsidade de um vetor \mathbf{x} é dada por $K(\mathbf{x}) = \|\mathbf{x}\|_0$.

Como um vetor pode ser representado em bases diferentes, sua esparsidade pode mudar, conforme a base utilizada. No exemplo 3, vimos que \mathbf{x} é a representação de \mathbf{y} na base Ψ . No exemplo 4, vimos que $K(\mathbf{y}) = 4$ e $K(\mathbf{x}) = 2$. Desta forma, ao representar um vetor em uma base diferente, podemos modificar a esparsidade de sua representação.

Definição 2 A base Ψ é dita base de esparsidade de um vetor \mathbf{y} de dimensão N se $K(\Psi\mathbf{y}) \ll N$

Exemplo 5 As figuras 2.2 e 2.3 apresentam os gráficos da função discreta $y[n] = 0.3\cos(2\pi 4n/N) + 0.3\cos(2\pi 7n/N) + 0.6\cos(2\pi 13n/N)$ e sua DFT $Y[n]$ (ou a sua representação na base de Fourier)

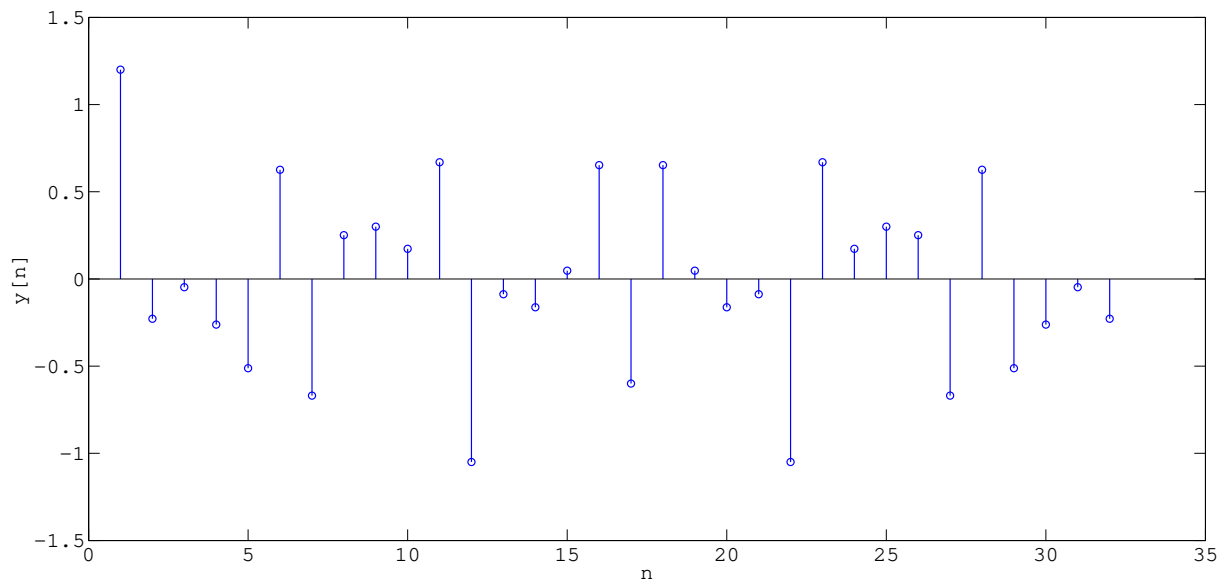


Figura 2.2: Representação Gráfica de $y[n]$ do exemplo 5

Nota-se que $y[n]$ possui 30 componentes não-nulas, ou seja, $K(y) = 30$. Já $Y[n]$ possui apenas 6 componentes não-nulas, $K(Y) = 6$. A base de Fourier tem um poder de diminuir K para vetores periódicos, ou seja, é uma base de esparsidade para vetores periódicos.

Se um vetor \mathbf{y} , pertencente a um espaço vetorial V de dimensão N , pode ser representado em uma base na qual ele é muito esparsos (muitas componentes nulas), então ele pertence a um subespaço W de V de dimensão $M < N$. Se esta base leva um conjunto de vetores \mathbf{x}_i sempre a este subespaço W , podemos representar os vetores \mathbf{y}_i de dimensão N como vetores \mathbf{x}_i de dimensão M , sem perda de qualquer informação sobre \mathbf{y}_i . Esta redução de dimensão é chamada de compressão sem perdas. Desta forma percebe-se que o conceito de esparsidade está ligado ao conceito de compressibilidade. Quanto maior a esparsidade de um vetor (ou seja, quanto menor valor de K), maior a sua compressibilidade.

Alguns vetores que representam sinais físicos, como áudio ou imagem, quando representados em uma base apropriada, não são naturalmente esparsos. Porém possuem a maioria de suas componentes próximas de zero. Definimos tais sinais como quase esparsos ou compressíveis. Formalmente, considere o vetor \mathbf{x}_S obtido mantendo-se os S maiores elementos de \mathbf{x} e tornando igual a zero os demais, isto é, $K(\mathbf{x}_S) = S$. Desta forma, podemos desprezar componentes de um sinal, com o compromisso de não causar

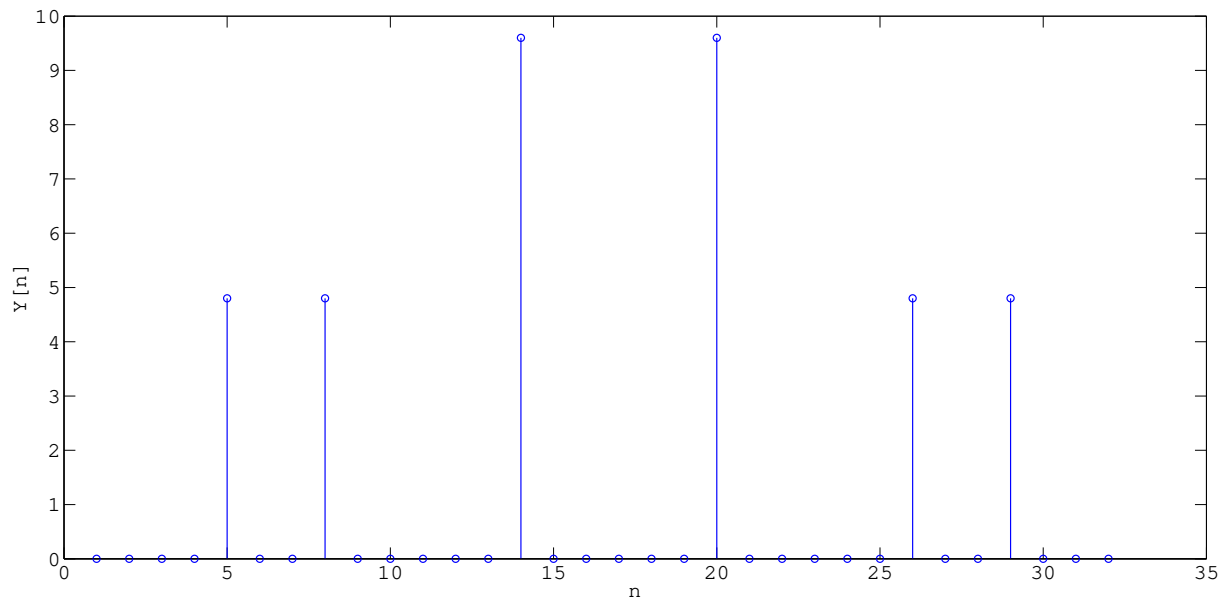


Figura 2.3: Representação Gráfica de $Y[n]$ do exemplo 5

perda de informação perceptível na qualidade subjetiva do sinal. Se aplicarmos a mesma redução de dimensão descrita no parágrafo anterior, chamamos tal redução de compressão com perdas, uma aproximação dependendo da aplicação. Em compressão de imagens, por exemplo, representa-se a imagem numa base wavelet [4], e desta retira-se a maioria dos coeficientes (mais de 90%). Ainda assim, é possível uma qualidade aceitável de imagem. Esse é o princípio no qual se baseia o padrão de compressão de imagens JPEG2000 [5].

2.8 Incoerência entre Bases

Considere duas bases ortonormais $\Psi, \Phi \in \mathbb{R}^{\sqrt{N} \times \sqrt{N}}$. A coerência entre as bases é dada por:

$$\mu(\Psi, \Phi) = \sqrt{N} \max\{\Phi^T \Psi\}. \quad (2.2)$$

Em outras palavras, a coerência é o maior valor entre todos os produtos internos (dois a dois) entre os elementos de Ψ e Φ .

Exemplo 6 Considere a base Φ do exemplo 2 e a base Ψ do exemplo 3. A coerência entre as bases é calculada a seguir.

$$\mu(\Psi, \Phi) = \sqrt{8} \max\{\Phi^T \Psi\}$$

$$\mu(\Psi, \Phi) = \sqrt{8} \max \left\{ \begin{array}{l} \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \frac{1}{\sqrt{8}} \left[\begin{array}{cccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \bar{A} & -j & -A & -1 & -\bar{A} & j & A \\ 1 & -j & -1 & j & 1 & -j & -1 & j \\ 1 & -A & j & \bar{A} & -1 & A & -j & -\bar{A} \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -\bar{A} & -j & A & -1 & \bar{A} & j & -A \\ 1 & j & -1 & -j & 1 & j & -1 & -j \\ 1 & A & j & -\bar{A} & -1 & -A & -j & \bar{A} \end{array} \right] \end{array} \right\}$$

$$\mu(\Psi, \Phi) = \sqrt{8} \frac{1}{\sqrt{8}} = 1$$

No exemplo 6, pode-se considerar que a base Ψ é a representação do sinal no domínio do tempo e a base Φ é a representação de um sinal no domínio da frequência. A coerência $\mu = 1$ significa que as representações no tempo e na frequência são minimamente coerentes (ou maximamente incoerentes).

Exemplo 7 Considere novamente a base Φ do exemplo 2. A coerência de Φ com ela mesma é dada pela expressão:

$$\mu(\Psi, \Phi) = \sqrt{8} \max \{ \Phi^t \Phi \}$$

$$\mu(\Psi, \Phi) = \sqrt{8} \max \left\{ \left(\begin{array}{l} \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \end{array} \right\}$$

$$\mu(\Psi, \Phi) = \sqrt{8} \cdot (1) = \sqrt{8}$$

Percebe-se, então, que a coerência entre uma base com ela mesma foi $\sqrt{8}$ (\sqrt{N}). Esta coerência é máxima, ou seja, as bases são minimamente incoerentes.

Assim como no exemplo 6, existem vários pares de bases que são bastante incoerentes entre si, por exemplo as wavelets e noiselets [6]. Por outro lado, existem algumas bases que tem baixa coerência com quaisquer outras bases. Como exemplo, temos a base gaussiana, em que os vetores da base são variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com distribuição gaussiana. Esta base tem baixa coerência com qualquer outra base.

2.9 Amostragem Discreta

Considere um vetor (que representa um sinal discreto) $\mathbf{y} = [y[0] \ y[1] \ \dots \ y[N-1]]^T$, para o qual as componentes $y[i]$ são desconhecidas. Considere ainda a matriz $M \times N$ ortonormal de amostragem (ou de medição) Φ . O processo de amostragem é expresso por:

$$\mathbf{y}_a = \Phi \mathbf{y}. \quad (2.3)$$

Para cada matriz Φ , existe uma forma de reconstrução do sinal. Na amostragem tradicional de sinais, Φ é a matriz identidade ($N \times N$). Desta forma o sinal é amostrado no domínio do tempo de forma linear e uniforme. No entanto a matriz Φ pode assumir outros valores, configurando outros tipos de amostragem, e necessitando outros tipos de reconstrução.

Exemplo 8 *Seja o vetor $\mathbf{y} = [1 \ 0 \ -1 \ 0 \ 1 \ 0 \ -1 \ 0]^T$ com a matriz de amostragem Φ sendo a) identidade, b) Fourier e c) Gaussiana.*

$$a) \mathbf{y}_a = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

$$b) \mathbf{y}_a = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \bar{A} & -j & -A & -1 & -\bar{A} & j & A \\ 1 & -j & -1 & j & 1 & -j & -1 & j \\ 1 & -A & j & \bar{A} & -1 & A & -j & -\bar{A} \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -\bar{A} & -j & A & -1 & \bar{A} & j & -A \\ 1 & j & -1 & -j & 1 & j & -1 & -j \\ 1 & A & j & -\bar{A} & -1 & -A & -j & \bar{A} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \sqrt{2} \\ 0 \\ 0 \\ 0 \\ \sqrt{2} \\ 0 \end{bmatrix}$$

$$c) \mathbf{y}_a = \begin{bmatrix} -0,1611 & -0,2300 & 0,5382 & -0,3047 & 0,2537 & -0,3863 & 0,4766 & -0,3129 \\ -0,6204 & -0,3852 & 0,0196 & 0,2928 & 0,0786 & 0,4842 & -0,1137 & -0,3564 \\ 0,0467 & 0,1090 & -0,0823 & 0,4536 & 0,8500 & -0,1703 & -0,0470 & 0,1403 \\ 0,1072 & -0,2819 & -0,1937 & -0,5718 & 0,3713 & 0,5033 & 0,1890 & 0,3431 \\ -0,4270 & 0,0614 & -0,0381 & -0,4716 & 0,1464 & -0,3349 & -0,6710 & 0,0788 \\ 0,4436 & -0,7883 & -0,0542 & 0,1050 & -0,0391 & -0,2405 & -0,3219 & -0,0695 \\ 0,4429 & 0,2804 & 0,3199 & -0,1648 & 0,2008 & 0,3579 & -0,3576 & -0,5472 \\ -0,0140 & -0,0594 & 0,7476 & 0,1676 & -0,0762 & 0,1895 & -0,2008 & 0,5720 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

$$\mathbf{y}_a = \begin{bmatrix} -0,9222 & -0,4476 & 1,0260 & 0,4832 & 0,4284 & 0,7805 & 0,6814 & -0,6370 \end{bmatrix}^T$$

A matriz de amostragem linear reproduz o sinal no mesmo domínio em que estava e portanto não altera sua esparsidade. Já a matriz de amostragem de Fourier altera a esparsidade, tornando a representação do sinal mais esparsa, concentrando a informação do sinal em apenas duas componentes. Já a matriz Gaussiana faz o oposto da matriz de Fourier. Ela tem a propriedade de tornar o sinal menos esparsa possível, espalhando assim a informação do sinal por todas as componentes.

2.10 Amostragem de Shannon/Nyquist

O teorema da amostragem de Shannon/Nyquist [1] relaciona a dimensão da matriz de amostragem com a maior frequência presente no vetor a ser amostrado. Suponha que as dimensões da matriz de amostragem e do vetor a ser amostrado sejam iguais às da seção anterior e seja f a maior frequência discreta do vetor a ser amostrado. Assim, a matriz de amostragem é da forma $[\psi_1 \ \psi_2 \ \dots \ \psi_M]^T$, em que $\psi_k[n] = \delta[n - Tk]$ e $T = N/M$ é a taxa de amostragem. O teorema afirma que para a perfeita reconstrução do sinal, é preciso que

$$M \geq 2f. \tag{2.4}$$

Exemplo 9 Seja o vetor \mathbf{y} do exemplo 1, $\mathbf{y} = [1 \ 0 \ -1 \ 0 \ 1 \ 0 \ -1 \ 0]^T$. No exemplo 3 foi visto que sua representação na base de Fourier (no domínio da frequência) é $\mathbf{x} = [0 \ 0 \ \sqrt{2} \ 0 \ 0 \ 0 \ \sqrt{2} \ 0]^T$.

Desta representação, podemos observar que sua maior (e única) frequência é $f = 2$. Pelo teorema de Shannon/Nyquist devemos ter que $M \geq 2(2) \Rightarrow M \geq 4$. Adotemos o limiar $M = 4$ e $T = N/M = 8/4 = 2$. A amostragem fica:

$$\mathbf{y}_a = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = [1 \quad -1 \quad 1 \quad -1]^T$$

2.11 Conclusão

Neste capítulo foi feita uma breve revisão sobre conceitos chaves para o bom entendimento de Amostragem Compressiva. Foi dado um enfoque prático, com exemplos direcionados, já introduzindo conceitos do capítulo 3: Amostragem Compressiva.

Capítulo 3

Amostragem Compressiva

3.1 Introdução

Pelo sucesso da área de compressão de dados já se percebeu que em várias aplicações, muitos - até mesmo a maioria - dos dados amostrados não são necessários para representar a informação de interesse, ou seja, parte da informação é irrelevante. Um exemplo claro é o fato de uma imagem de tamanho da ordem de Megabytes poder ter uma representação mais eficiente, no formato JPEG, por dados de tamanho da ordem de Kilobytes. Outro exemplo é uma gravação de áudio de tamanho de 30 MB em formato WAV poder ser representada por 3MB em formato MP3.

Ambas as representações, JPEG e MP3, apresentam perda de qualidade perceptual, porém ainda se mostram útil para várias aplicações. Logo, a informação contida no sinal é, nestes casos, menor que a informação contida nos dados produzidos pela amostragem, na taxa de Nyquist. Desta forma, a idéia central da Amostragem Compressiva [7], [8], [9] e [6] é combinar as etapas de aquisição de dados e compressão, com a intenção de adquirir apenas amostras suficientes que representem a informação de interesse presente no sinal. A Amostragem Compressiva traduz dados analógicos ou digitais diretamente para informação útil através de um processo de amostragem com a restrição do sinal ser esparso (ou compressível, se distorções forem permitidas) em alguma base. Os sinais naturais, em sua grande maioria, são compressíveis, quando representados em uma base adequada [10].

3.2 Amostragem Aleatória

O teorema de Shannon/Nyquist é preciso e correto. Sem nenhuma informação adicional sobre o sinal, não é possível realizar uma amostragem a uma taxa menor do que duas vezes sua maior frequência, sem a ocorrência de distorções. A idéia de Candès, Romberg

e Tao foi incluir o conhecimento prévio sobre a esparsidade dos sinais a serem amostrados, já que são de fato esparsos em alguma base. Formalmente, assume-se que o sinal a ser amostrado tem alta esparsidade (valor de S pequeno) em alguma base $\Psi \in \mathbb{R}^{N \times N}$. E, em outra base $\Phi \in \mathbb{R}^{M \times N}$, realiza-se a amostragem:

$$\mathbf{y}_a = \Phi \mathbf{y}. \quad (3.1)$$

O método não-adaptativo de amostragem proposto por Candès, Romberg e Tao [7] usa aleatoriedade como mecanismo de amostragem. O método foi em parte baseado na idéia de que se Ψ for uma base de esparsidade (sessão 2.7) e tiver baixa coerência (sessão 2.8) com Φ , então a representação do sinal na base Φ não será esparsa [11]. Em outras palavras, a informação total contida em \mathbf{y} será uniformemente espalhada por todos os coeficientes de \mathbf{y}_a , de forma que cada coeficiente carregará parte da informação global sobre o sinal. Como existia redundância no sinal e a representação do sinal na base Φ gera coeficientes que carregam a mesma quantidade de informação, pode-se inferir que amostrando-se apenas $M < N$ coeficientes, é possível reconstruir a informação completa sobre o sinal.

A incoerência entre as bases de amostragem e de esparsidade é uma condição necessária para o sucesso da técnica de Amostragem Compressiva. A escolha de um mecanismo de amostragem adequado é essencial. Devido ao fato de bases aleatórias como a base Gaussiana ou de Bernoulli (seção 2.8) terem baixa coerência com qualquer base de esparsidade, a amostragem aleatória se mostra um mecanismo universal para a aquisição dos dados.

3.3 Reconstrução

Para a reconstrução de \mathbf{y} a partir de \mathbf{y}_a , não é possível simplesmente solucionar o sistema linear $\mathbf{y}_a = \Phi \mathbf{y}$. É um sistema indeterminado, ou seja, existem mais incógnitas do que equações. Porém Chen e Huo [12] provaram que entre as possíveis soluções, aquela que tem menor norma l_0 é solução única se o posto de Kruskal [13]¹ de Φ for no mínimo $2S$, sendo $S = K(\mathbf{y})$. Assim esta solução única pode ser encontrada solucionando o problema de minimização [8]:

$$\min_{\tilde{\mathbf{x}} \in \mathbb{R}^n} \|\tilde{\mathbf{x}}\|_0 \text{ sujeito a } \Phi \Psi \tilde{\mathbf{x}} = \mathbf{y}_a \quad (3.2)$$

Apesar de esta minimização produzir uma solução correta, este é um problema combinatorial, NP-completo [7] [8], isto é, é preciso testar o universo completo das possíveis soluções para afirmar qual delas é a correta. Este procedimento é computacionalmente impraticável com os recursos computacionais atuais. Alguns algoritmos de complexidade

¹Posto de Kruskal é o número de colunas linearmente independentes

polinomial foram propostos na literatura, destacando-se duas abordagens: a substituição de otimização de norma l_0 pela norma l_1 [7] [8], um problema que pode ser reformulado como de programação linear e o algoritmo *orthogonal matching pursuit* [14] [15], que utiliza o método guloso para achar a solução mais esparsa. Estes algoritmos fornecem a solução exata somente se algumas condições foram satisfeitas para cada caso.

Na solução por otimização da norma l_0 , Φ precisa atender a dois critérios para a perfeita reconstrução do sinal: satisfazer o teorema da amostragem para este caso e satisfazer o Princípio da Isometria Restritiva [7] [16] [17] (do inglês *Restricted Isometry Principle* - RIP).

Para introduzir como a otimização l_1 depende do RIP, considere uma matriz de amostragem Φ ($N \times M$) e Φ_T ($N \times T$) uma submatriz obtida extraíndo T colunas de Φ . Assim, δ_S , ou constante de restrição a S , é o menor valor que satisfaz

$$(1 - \delta_S)\|\mathbf{y}\|_2^2 \leq \|\Phi_S \mathbf{y}\|_2^2 \leq (1 + \delta_S)\|\mathbf{y}\|_2^2 \quad (3.3)$$

para todas as possíveis combinações de Φ_S , em que S é a esparsidade de \mathbf{y} . Esta propriedade basicamente estabelece que qualquer conjunto de S colunas de Φ se comporta como uma matriz ortonormal. Considere as condições

$$\delta_{2S} < 1 \quad (3.4)$$

e

$$\delta_{2S} < 1 - \sqrt{2} \quad (3.5)$$

Se a condição 3.4 é atendida, então a otimização da norma l_0 fornece uma solução única e igual a \mathbf{y} . Se a condição 3.5 é atendida, então a solução da otimização da norma l_1 é igual à solução da otimização da norma l_0 .

Assim, para que a solução do problema de reconstrução do sinal amostrado por Amostragem Compressiva seja única e correta, a base de amostragem deve obedecer ao RIP. Felizmente, as bases Gaussianas e de Bernoulli obedecem ao RIP na grande maioria dos casos.

3.4 Teoremas da Amostragem

Além da condição de RIP, para a perfeita reconstrução do sinal, outra condição deve ser satisfeita. É o teorema da amostragem, específico para cada tipo de base de amostragem Φ e base de esparsidade Ψ . Genericamente:

$$M \geq C\mu^2(\Psi, \Phi)S \ln(N), \quad (3.6)$$

em que M é o número mínimo de amostras para a perfeita reconstrução, C é uma constante a ser determinada experimentalmente, $\mu^2(\Psi, \Phi)$ é a coerência entre a base de amostragem Φ e a base de esparsidade Ψ , S e N são esparsidade e a dimensão do sinal a ser amostrado, respectivamente.

O papel da coerência é bastante claro. Quanto menor a coerência, menor o valor de M . Para sinais periódicos, se Φ for construída a partir de colunas escolhidas aleatoriamente da matriz de Fourier [7], então a equação se reduz a

$$M \geq CS \ln(N). \quad (3.7)$$

No entanto, para que o teorema sirva para quaisquer bases de esparsidade, ou seja, para qualquer tipo de sinal, a base de amostragem deve ser aleatória como já mencionado. Se Φ for uma matriz Gaussiana ou de Bernoulli [6], [18], a equação se reduz a

$$M \geq CS \ln(N/S). \quad (3.8)$$

3.5 Conclusão

A Amostragem Compressiva estabelece um novo paradigma para a amostragem de sinais. Este afirma que o número de medidas necessárias para a perfeita reconstrução do sinal pode ser menor que o que a Taxa de Nyquist. Porém tal afirmação não é uma violação do teorema porque ele se aplica quando não se tem nenhuma informação adicional sobre o sinal. Em AC, considera-se que o nível de esparsidade do sinal é conhecido. Outra importante observação é que a esparsidade (ou, equivalentemente, a compressibilidade, se erros forem tolerados) tem o papel mais importante na equação que rege o número de medidas necessárias para a perfeita reconstrução do sinal: o número mínimo de amostras é diretamente proporcional à esparsidade. Já a dimensão do vetor tem um papel apenas secundário: o número mínimo de amostras é proporcional ao logaritmo da razão entre a esparsidade e a dimensão.

O número de medidas para a perfeita reconstrução do vetor pode chegar a ser muito menor (da ordem de centena de milhares) do que o limite de Nyquist/Shannon para sinais amostrados de forma uniforme e depende de quão boa seja a base de esparsidade, ou seja, é preciso encontrar uma base de esparsidade adequada para o sinal para que este tipo de amostragem seja superior.

Observa-se, ainda, que, na reconstrução do vetor, a otimização é uma operação não-linear, que pode ser reformulada através de um modelo de programação linear. Existem

alguns tipos de algoritmos amplamente utilizados para este problema, porém aplicações de reconstrução do sinal em tempo real para altas dimensões ainda são um desafio em virtude do grau de complexidade computacional dessas soluções.

Portanto, a aplicação de CS a certo tipo de sinal se concentra basicamente no problema de identificar uma base de esparsidade adequada ao sinal de interesse e, para sinais não-esparsos, mas compressíveis, estimar o nível de compressibilidade tolerável para que o sinal reconstruído ainda seja considerado de interesse.

Capítulo 4

Revisão Bibliográfica

4.1 Introdução

Neste capítulo é realizada uma revisão bibliográfica compreendendo os seguintes tópicos:

- Amostragem Clássica e Amostragem Compressiva
- Aplicações de Amostragem Compressiva em Áreas Afins
- Aplicações de Amostragem Compressiva em Áudio
- Investigação de Base de Esparsidade em Áudio
- Medidas de Avaliação Objetiva de Qualidade Perceptual de Áudio

4.2 Amostragem Clássica e Amostragem Compressiva

A teoria clássica da amostragem de sinais foi inicialmente proposta por Nyquist [2] em 1929, que estabelecia que para a perfeita reconstrução de um sinal uniformemente amostrado, sua amostragem deveria acontecer a uma taxa duas vezes maior do que a banda que ele ocupa no domínio de Fourier. A prova do teorema da amostragem proposto por ele foi apenas dada por Shannon [1] em 1949. Desde então, apenas extensões desse teorema foram propostas, e resumidas em um tutorial por Jerri [19].

Com uma abordagem diferente da teoria da amostragem clássica, a teoria da Amostragem Compressiva (AC) foi inicialmente proposta por Candès, Romberg e Tao [7] e estendida por Donoho [8]. Esses autores propuseram uma amostragem de sinais esparsos através de bases incoerentes. Desta forma estabeleceram a possibilidade de novos teoremas da amostragem e formas de reconstrução.

No campo da amostragem, Candès e Tao [9] propuseram uma base de amostragem aleatória, através de matrizes que obedecem certos critérios, como forma de Amostragem Compressiva independente da natureza do sinal, ou seja, universal. No campo da reconstrução, surgiram várias propostas de algoritmos para a reconstrução do sinal amostrado por AC. Candès e Tao [18] e Donoho e Elady [20] propuseram a reconstrução pela minimização de norma l_1 , um problema que pode ser reformulado para um problema equivalente de programação linear. Mallat e Zhang [21] propuseram a reconstrução por *Matching Pursuit*, Chen, Donoho e Saunders [22] por *Basis Pursuit*, Tropp por *Basis Pursuit* [14] e por *Convex Relaxation* [15]. Mais de duzentos artigos tratam variantes e extensões da teoria de AC.

4.3 Aplicações de Amostragem Compressiva em Áreas Afins

O primeiro protótipo que utiliza técnicas de AC foi desenvolvido por Duarte et al. [23]. Esse protótipo consiste em um arranjo de micro-espelhos distribuídos de forma pseudo-aleatória que reflete uma imagem para uma lente que focaliza as imagens refletidas em um sensor de silício (operação equivalente a uma combinação linear dos pixels que cada espelho reflete); em seguida um conversor ADC lê a tensão do sensor e o converte em dados digitais, gerando uma medida. Para cada arranjo pseudo-aleatório de espelhos, gera-se uma medida; um conjunto finito de medidas representam a imagem comprimida e quanto mais medidas, maior a resolução. Um Processador Digital de Sinais (DSP) transforma, então, o conjunto de medidas em uma imagem por um algoritmo de otimização convexa. Esta aplicação possibilita a construção de sensores simples, pequenos e baratos, interessante em situações que se trabalha em espectros não-visíveis, nas quais o sensor é caro.

A partir das técnicas de AC, foram propostas arquiteturas de conversores analógico-informação (A2I), que são conversores que transformam o sinal analógico em informação, integrando as etapas de amostragem e compressão. Kirolos et al. [24] propuseram uma arquitetura baseada em uma demodulação aleatória. Nela o sinal analógico é multiplicado por uma sequência pseudo-aleatória, integrado e quantizado, e por fim, reconstruído. Essa arquitetura foi simulada como prova de conceito em nível de transistor para sinais de banda larga esparsos no domínio de Fourier, na qual foi possível uma perfeita reconstrução do sinal para uma taxa de amostragem seis vezes menor do que a taxa de Nyquist. Tropp et al. [25] propuseram outra arquitetura de A2I, baseada em filtragem aleatória. Essa arquitetura é baseada em um filtro FIR de tamanho fixo e coeficientes variáveis e pseudo-aleatórios e é projetada para sinais esparsos ou compressíveis no domínio do tempo, de

Fourier e Wavelet [4], com sua reconstrução por um algoritmo de *Orthogonal Matching Pursuit*. Resultados numéricos indicaram a eficiência do método para a amostragem de sinais, contudo ainda é um método experimental que necessita de maiores investigações.

Lustig, Donoho e Pauly [26] aplicam técnicas de AC em medições de ressonância magnética (RM). Exploram a esparsidade implícita nestas medições, ou em transformações de esparsidade das mesmas, para acelerar a aquisição das imagens ou o aumento da resolução espacial das mesmas. A RM tem limitações físicas (provenientes do equipamento de medição) e fisiológicas (provenientes do corpo humano) que inviabilizam a aceleração do procedimento. Desta forma, essa aplicação pode viabilizar, mesmo com tais limitações, a produção de imagens de RM de forma mais rápida, podendo chegar até na produção de vídeos com a evolução da técnica.

Candès et al. [27] propuseram uma aplicação de AC na construção de códigos corretores de erros. Como as medições são aleatórias, cada medida carrega informação sobre o sinal como um todo. Desta forma, aumentando o número de medições, o número de medidas fica redundante, e, por conseguinte, se alguma medida for corrompida, as demais podem ainda reconstruir o sinal perfeitamente. Este tipo de técnica é robusta, tanto para erros aleatórios, quanto erros em rajadas. Candès e Randall [28] analisaram o problema não apenas no caso de erros corrompendo o sinal no canal, mas também erros de quantização desse mesmo sinal.

Duarte et al. [29] e Drori [30] citam a possibilidade de utilizar AC para realizar criptografia, pela característica de a amostragem ser aleatória. De forma prática dois indivíduos que queiram se comunicar de forma criptografada compartilhariam uma chave única, a matriz de amostragem (ou uma semente para gerá-la). O transmissor enviaria a mensagem criptografada com a matriz previamente compartilhada e o receptor reconstruiria a mensagem com sua matriz. Porém Rachlin e Baron [31] realizaram um estudo formal desta criptografia. Eles provam que esta criptografia não é perfeita, pois a informação mútua entre a mensagem original e a criptografada não é nula pela propriedade de RIP, que mantém a norma 2 dos símbolos. Pelo lado computacional, eles provam que só é possível reconstruir o sinal original com a matriz original, o que torna esta criptografia uma criptografia computacionalmente segura caso o indivíduo que quer interceptar a mensagem possua recursos computacionais limitados. Este método apresenta a vantagem que, em uma única aplicação, pode-se criptografar e comprimir, ou criptografar e proteger contra erros a mensagem.

Bobin, Starck, e Ottensamer [32] aplicam técnicas de AC para capturar imagens em astronomia. Estas imagens, captadas de satélites, são naturalmente esparsas, ideais para a aplicação de AC. Atualmente se gasta muita energia (escassa em satélites), para enviar para a Terra as imagens capturadas. Os autores afirmam que através de AC, é possível

um aumento de 30% na resolução espacial em relação às técnicas da Agência Espacial Européia.

4.4 Aplicações de Amostragem Compressiva em Voz e Áudio

Na área de voz e áudio, foram encontrados na literatura apenas dois artigos. Griffin e Tsakalides [33] comparam o desempenho de AC com o uso da DCT e da DWT como bases esparsas para sinais de voz, de música, de canto de pássaros e de um impulso. Concluíram que a DCT representa de forma mais esparsa os sinais, com exceção do impulso. Propuseram e simularam ainda um sistema de localização espacial de um objeto que transmite sinal de áudio esparsa, baseada numa rede de sensores. Propuseram ainda um algoritmo de estimação do sinal baseado na cooperação dos sensores. Como parâmetro de qualidade, apresentam a medida relação sinal-distorção, que para a aplicação apresentada mostrou-se suficiente, já que pretendeu-se apenas estimar a localização de um objeto. Já Carmi, Kanevsky e Ramabhadran [34] propuseram um algoritmo de compressão com perdas de voz através de técnicas de AC com base de esparsidade a DFT e um filtro de Kalman na saída. Apresentam resultados apenas por comparação gráfica entre um único sinal original e o mesmo comprimido a 65% do tamanho original. Não fornecem medidas objetivas ou subjetivas da qualidade da reconstrução do sinal através da técnica proposta.

4.5 Investigação de Base de Esparsidade em Áudio

Bases de esparsidade para áudio vêm sendo estudadas para aplicações em separação cega de fontes. Em sua *Habilitation à Diriger des Recherches*, Gribonval [35] faz uma revisão sobre separação cega de fontes de áudio e, principalmente, trata sobre representação esparsa de sinais de áudio e algoritmos para aproximação esparsa. Outros autores, a exemplo de Blumensath e Davies [36], Plumbley, Abdallah, Blumensath, Jafari, Nesbit, Vincent, Wang [37] e Davies e Daudet propõe diversas bases para a representação esparsa de áudio. Além disso, os padrões de compressão de áudio utilizam em uma etapa a MDCT para compressão de áudio [38], o que pode indicar ser uma boa base de esparsidade.

4.6 Medidas de Avaliação Objetiva de Qualidade Perceptual de Áudio

De modo a avaliar a qualidade perceptual de um sinal de áudio comprimido, idealmente se realiza testes subjetivos com pessoas. Porém tais testes consomem muito tempo e às vezes são impraticáveis. Alguns métodos de avaliação perceptual foram propostos [39], [40], [41] e [42]. Um esforço da *International Telecommunication Union* analisou os diversos métodos de análise objetiva de compressão e elaborou um método padrão de avaliação chamado *Perceptual Evaluation of Audio Quality* (PEAQ), considerando vários parâmetros de qualidade. Este é atualmente o método mais aceito na comunidade acadêmica, relatado por Thiede et al. [43]. O PEAQ simula as propriedades perceptuais do ouvido humano através por vários modelos e reúne as saídas de cada modelo em uma saída única, como se o objeto de teste tivesse sido submetido a testes subjetivos com pessoas. O resultado do método revela uma escala de qualidade perceptual que fica entre 0 (excelente) e -5 (muito ruim).

Capítulo 5

Amostragem Compressiva Aplicada a Sinais de Áudio

De modo a atingir os objetivos da dissertação, foram elaboradas rotinas no ambiente de cálculo numérico Matlab. Cada rotina foi elaborada em formato de experimento numérico, afim de verificar cada problema apresentado:

- Investigação de bases esparsidade para áudio
- Investigação de técnicas de reconstrução
- Proposição de um sistema de compressão de áudio com base nas investigações
- Avaliação da qualidade perceptual do sistema por parâmetro objetivo

Para cada avaliação foi elaborado um algoritmo que, em seguida, foi codificado para o ambiente Matlab. Os arquivos de áudio para testes foram retirados do Material de avaliação de Qualidade de Som da EBU (*European Broadcasting Union*) [44]. Trechos de 10 segundos dos seguintes arquivos foram testados:

- música pop: ABBA (Stereo).wav
- música erudita: Piano (Schubert) (Stereo).wav, Clarinet (arpeggio;melodious phrase) (Stereo).wav, Orchestra (R. Strauss) (Stereo).wav
- fala: Female Speech (English) (Mono).wav, Male Speech (English) (Mono).wav

5.1 Investigação de Base de Esparsidade para Áudio

No trabalho publicado por Griffin e Tsakalides [33], as bases de esparsidade analisadas foram as matrizes da Transformada Discreta do Cosseno (DCT) e da Transformada

Wavelet Discreta (DWT). Eles concluíram que a DCT se mostrou a base mais apropriada à maioria dos sinais testados com relação à medida sinal-distorção. O trabalho publicado por Carmi, Kanevsky e Ramabhadran [34] utilizou a Transformada Discreta de Fourier (DFT) como base de esparsidade.

Com base nos trabalhos citados e com a intenção de encontrar uma base completa baseada em transformada, as bases DCT e DFT foram escolhidas para avaliação. A Transformada Discreta do Cosseno Modificada (MDCT) [45] também foi escolhida, por ser a transformada utilizada nos padrões de compressão de áudio [38].

As bases foram analisadas através de gráficos comparativos, gerados segundo o algoritmo:

- Entrada: arquivos de áudio para teste
- Para cada arquivo de áudio faça
 - Para Tamanho de janela $N = 128$, $N = 256$ e $N = 512$ faça
 - * Para esparsidade $S = 1$ a $S = N$, faça
 1. Divida o arquivo em blocos N amostras
 2. Calcule cada transformada para cada bloco
 3. Torne os $N - S$ menores coeficientes de cada bloco iguais a zero
 4. Calcule a transformada inversa para cada bloco
 - * Reconstrua o sinal juntando os blocos reconstruídos,
 - Calcule a diferença objetiva, pelo método de PEAQ, entre o arquivo de áudio original e o reconstruído,
- Saída: Gráficos da média da diferença objetiva dos arquivos de áudio em função da taxa de compressão S/N

Para o caso da MDCT, ainda é preciso um estágio de soma entre janelas por ser uma transformada superposta (do inglês *lapped transform*). A saída do algoritmo pode ser observada nos gráficos das Figuras 5.1, 5.2, 5.3 e 5.4. A análise para $N = 512$ foi feita apenas para a DCT e MDCT por já ter concluído das demais bases que a DFT não é uma boa base de esparsidade para sinais de áudio.

As grandezas presentes nos gráficos são:

Qualidade Perceptual : Medida objetiva da qualidade perceptual do áudio reconstruído em comparação com o áudio original, através do método PEAQ [43]. A interpretação da medida se encontra na tabela 5.1.

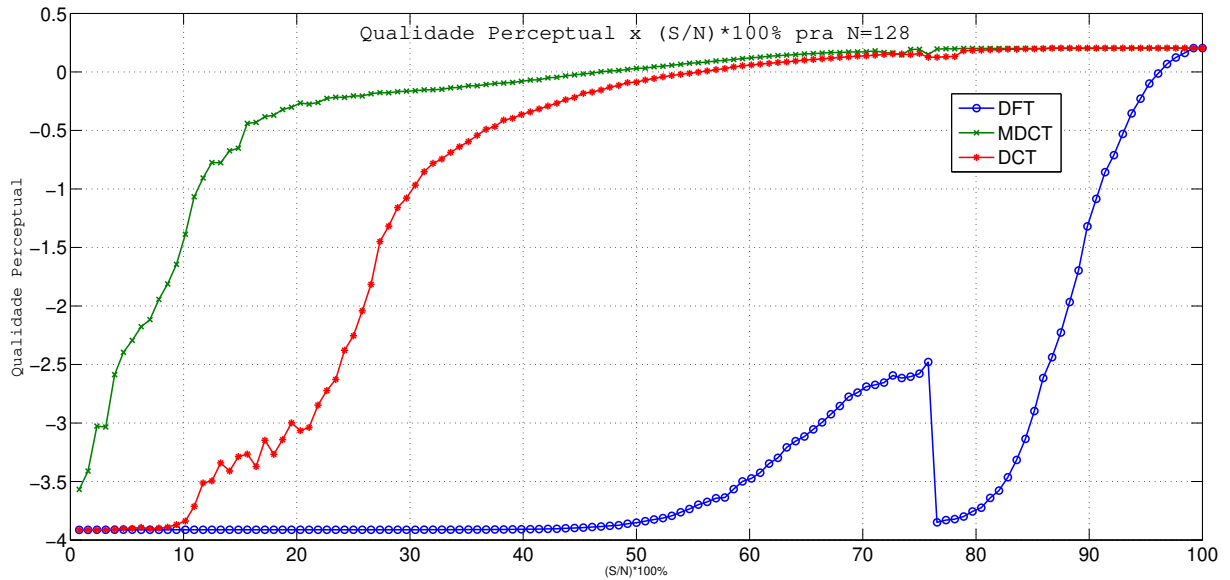


Figura 5.1: Qualidade Perceptual versus Razão de Esparsidade S/N com N=128. Comparação entre as bases de esparsidade DFT, DCT e MDCT.

Qualidade Perceptual	Interpretação
-5 a -4	Diferença perceptível e muito desagradável
-4 a -3	Diferença perceptível e desagradável
-3 a -2	Diferença perceptível e levemente desagradável
-2 a -1	Diferença perceptível mas não desagradável
-1 a 0	Diferença imperceptível

Tabela 5.1: Interpretação do valor da Qualidade Perceptual PEAQ

Razão de Compressão : Relação entre a dimensão do sinal comprimido M e a dimensão do sinal original N. Corresponde à subamostragem para o caso analógico.

Razão de Esparsidade : Relação entre a esparsidade S e a dimensão N do sinal original para cada janela de análise

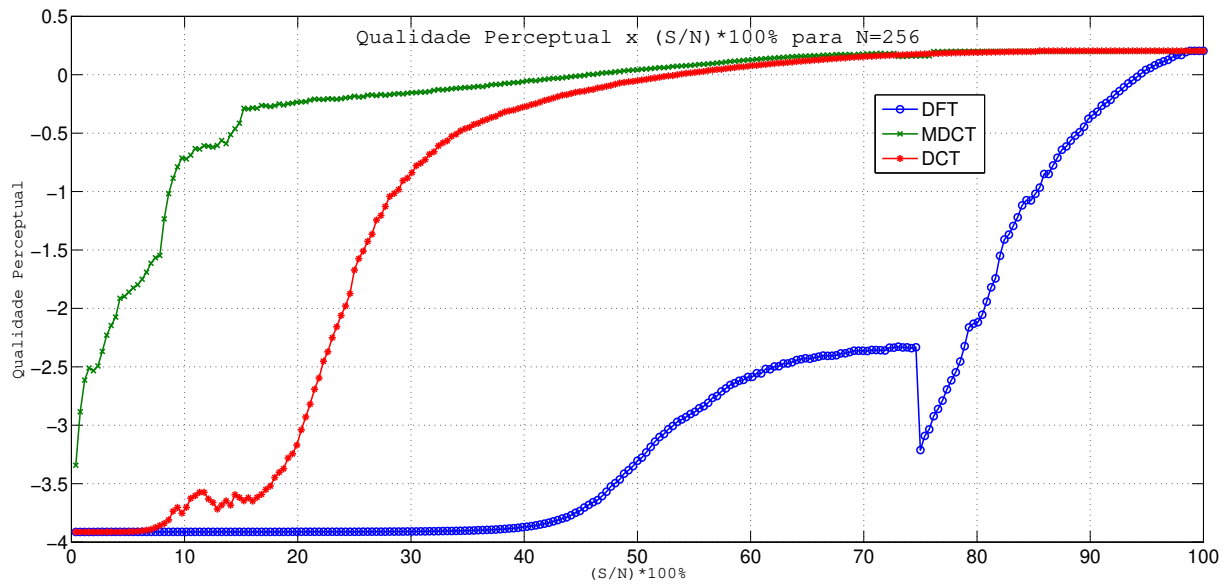


Figura 5.2: Gráfico da Qualidade Perceptual versus Razão de Esparsidade S/N com N=256. Comparação entre as bases de esparsidade DFT, DCT e MDCT.

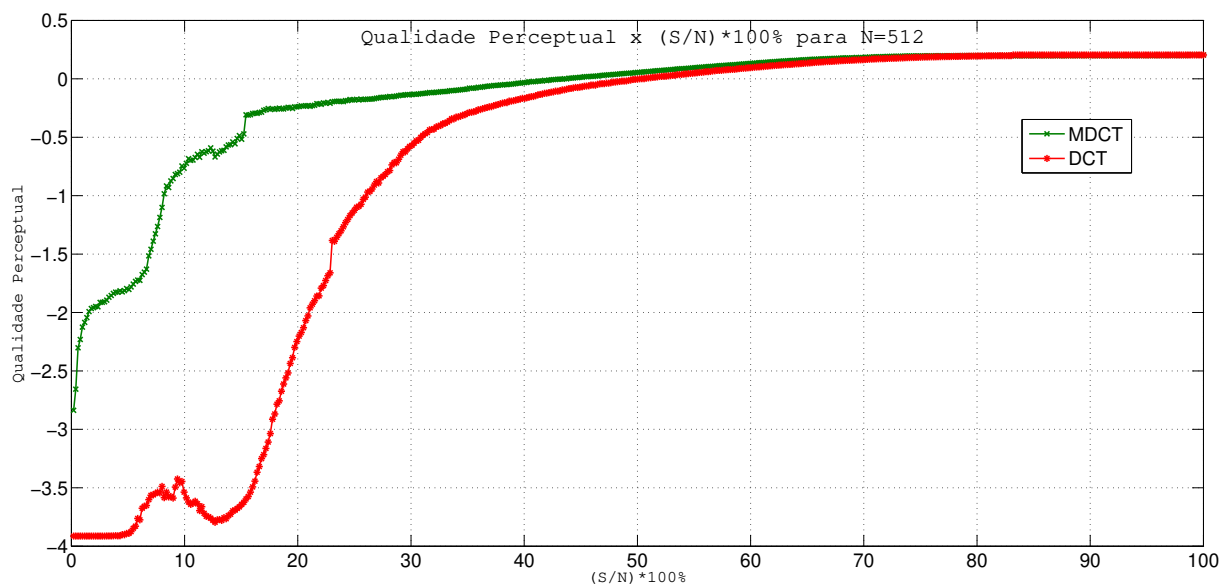


Figura 5.3: Gráfico da Qualidade Perceptual versus a Razão de Esparsidade S/N com N=512. Comparação entre as bases de esparsidade DCT e MDCT.

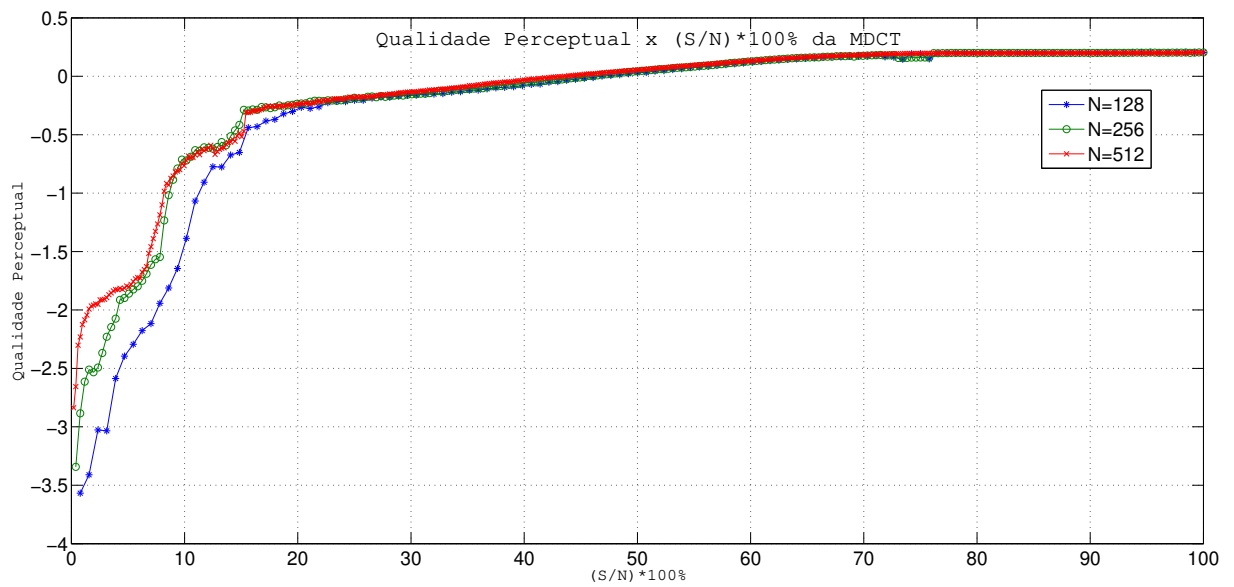


Figura 5.4: Gráfico da Qualidade Perceptual versus Razão de Esparsidade S/N para a MDCT. Comparação entre tamanhos de bloco

Interpretando os gráficos, pode-se observar que a transformada que apresentou melhor desempenho foi a MDCT, independente do número de amostras por janela. A DCT apresentou desempenho intermediário e a DFT não apresentou desempenho satisfatório. Com base nisso pode-se concluir que a MDCT é a transformada mais adequada como base de esparsidade para áudio. O gráfico da figura 5.4 mostra que existe um pequeno ganho de qualidade com o aumento da dimensão da janela, mas apenas para taxas de esparsidade baixas. Para taxas de esparsidade maiores que 20%, o desempenho é equivalente.

5.2 Investigação de Técnicas de Reconstrução

Apesar de existir uma equação que relaciona o número mínimo de medidas com a dimensão e a esparsidade do sinal, apresentada no capítulo 3, existe uma constante C que só pode ser determinada empiricamente. A investigação empírica tem um aspecto prático para futuros projetos, de modo a saber exatamente quanto é possível se alcançar de taxa de amostragem para se reconstruir determinada razão de esparsidade. Investigou-se a técnica de reconstrução por minimização l_1 para dimensões do sinal de 64, 128, 256 e 512. Os resultados encontram-se no gráfico da figura 5.5. Considerou-se que uma perfeita reconstrução do sinal é aquela que difere por um erro médio quadrático menor que 10^{-6} .

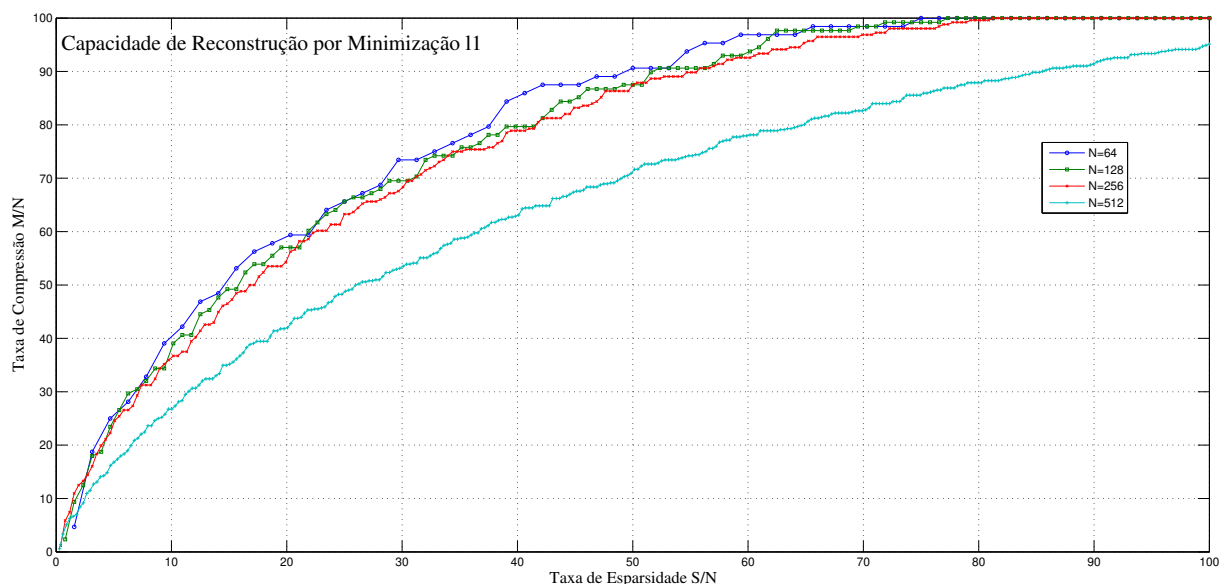


Figura 5.5: Capacidade de Reconstrução do sinal por minimização l_1

Observou-se que a curva se comporta aproximadamente como a equação prevê. A relação entre razão de esparsidade e razão de compressão de fato é aproximadamente logarítmica, quanto menor a razão de esparsidade, muito menor é a razão de compressão. Observou-se também que há uma pequena melhora de desempenho da reconstrução com o aumento da dimensão do sinal.

5.3 Proposta de Sistema de Amostragem e Compressão

Um sistema de compressão de áudio é composto essencialmente por um codificador e um decodificador. Diferentemente de um sistema de compressão clássico, na nossa proposta, o projeto do decodificador pode ser realizado sem nenhum conhecimento sobre o codificador, pela propriedade de universalidade da amostragem de Amostragem Compressiva. A Amostragem Compressiva foi concebida inicialmente para sinais discretos e finitos no tempo. No entanto, do ponto de vista do sistema de amostragem, o sinal de áudio é infinito no tempo e portanto, a amostragem deve ser realizada por janelas, e conseqüentemente, também a reconstrução.

5.3.1 Codificador

A Amostragem Compressiva realiza a amostragem e compressão do sinal em uma mesma operação. Embora a teoria de Amostragem Compressiva tenha sido concebida inicialmente para sinais discretos, ela foi posteriormente estendida para sinais analógicos [46]. Um aspecto importante da teoria de Amostragem Compressiva é que não é preciso amostrar acima da taxa de Nyquist e em seguida realizar a compressão. As duas etapas podem ser combinadas e realizadas por *hardware* barato e de fácil concepção. Algumas arquiteturas foram propostas, simuladas e até implementadas fisicamente [46], [25], [24]. Independente da arquitetura de amostragem escolhida para a implementação por *hardware*, esta pode ser traduzida para o domínio discreto por uma matriz de amostragem equivalente. Este trabalho abordou o problema da compressão sob este ponto de vista.

A matriz de amostragem escolhida foi a matriz de Bernoulli, em que cada coluna da matriz é uma seqüência pseudo-aleatória de valores +1 e -1. Esta matriz apresentou resultados equivalentes à matriz Gaussiana, porém foi escolhida por ser mais fácil de implementar em *hardware*. Para diferentes aplicações, outro tipo de matriz pode ser usada de modo a facilitar a concepção do circuito de amostragem.

Alternativamente, o codificador também pode ser concebido exclusivamente por *software* como parte integrante de um codec. Neste caso, a operação de amostragem/compressão se resume a uma multiplicação de uma matriz de amostragem por um vetor, no caso uma janela do arquivo. Ao contrário de codecs clássicos, onde o codificador tem uma complexidade computacional maior que o decodificador, este codificador tem uma complexidade muito menor que o decodificador.

5.3.2 Decodificador

O decodificador irá se compor basicamente de três blocos funcionais: *Basis Pursuit*, transformação inversa e superposição de janelas. O bloco de *Basis Pursuit* realiza a seguinte

operação:

$$\min_{\tilde{\mathbf{x}} \in \mathbb{R}^n} \|\tilde{\mathbf{x}}\|_1 \text{ sujeito a } \Phi\Psi\tilde{\mathbf{x}} = \mathbf{y}_a \quad (5.1)$$

Em outras palavras, *Basis Pursuit* busca o vetor esparso $\tilde{\mathbf{x}}$ que apresenta menor norma l_1 e que, multiplicado por $\Phi\Psi$, resulta no vetor da medição. Ψ é a base de esparsidade (MDCT) e Φ é a base de amostragem (Bernoulli). Esta operação pode recuperar a janela do áudio de forma exata ou parcialmente correta, dependendo da esparsidade e do número de medições para cada janela [45].

O *Basis Pursuit* encontra o sinal $\tilde{\mathbf{x}}$ na base de esparsidade, ou seja, está no domínio da transformada. No entanto, o sinal de interesse está no domínio do tempo. Assim, é preciso uma operação de transformação inversa de modo a recuperar o sinal no domínio original. Um último bloco é necessário, quando a MDCT é adotada como base de esparsidade. A MDCT é uma transformada superposta, assim, é preciso ainda uma operação de superposição e soma de janelas para reconstruir o sinal.

5.3.3 Propriedades

A Amostragem Compressiva goza de algumas propriedades que são naturalmente herdadas pelo sistema proposto e podem se apresentar como vantagens, conforme a aplicação. Por exemplo:

Amostragem Aleatória A amostragem se dá pela multiplicação por uma matriz pseudo-aleatória que pode ser gerada por uma semente. Tal propriedade pode ser aplicada à criptografia, quando a semente é encarada como uma chave criptográfica. Assim, o codificador só pode reconstruir o sinal se tiver conhecimento da semente e gerar a matriz de amostragem localmente. Qualquer outra matriz de amostragem não reconstruirá o sinal.

Robustez a Erros A perda de amostras do sinal comprimido não gera a perda total do sinal reconstruído. O decodificador encara a perda de amostras apenas como uma redução do número de amostras do sinal, e assim, apenas consegue reconstruir um número menor de maiores componentes. Assim, a perda causa uma queda da qualidade do sinal como um todo, e não de algum ponto específico.

Universalidade Embora a Amostragem Compressiva não seja adaptativa, ela é universal. O projeto do codificador leva em conta a esparsidade do sinal, e não sua banda de frequência. Assim, sinais com o mesmo nível de esparsidade podem ser amostrados com o mesmo codificador, independente da natureza do sinal. Sinais como voz, áudio, música pop, música instrumental podem ser amostrados da mesma forma,

e reconstruídos com bases de esparsidade diferentes, adaptadas para cada tipo de sinal.

Complexidade Reversa O codificador é extremamente simples. Isso possibilita a aplicação de compressão onde existe limitação de *hardware*, como em redes de sensores sem fio [47]. Neste cenário, os módulos sensores devem ser extremamente simples, porém a central de fusão que reconstrói o sinal é concebida com alto poder de processamento.

5.4 Qualidade Perceptual do Sistema Proposto

Como último resultado foi analisada a qualidade perceptual do sistema utilizando como base de esparsidade a MDCT, como base de amostragem a matriz de Bernoulli e como técnica de reconstrução a minimização da norma l_1 . Foram feitas análises para o tamanho de janela $N=128$, 256 e 512. Idealmente, a análise subjetiva da qualidade perceptual seria mais precisa. No entanto, devido ao grande número de casos que precisariam ser testados, a análise só foi possível por medida objetiva de qualidade perceptual pelo método de PEAQ. As figuras 5.6, 5.7 e 5.8 mostram gráficos relacionando a razão de compressão com a qualidade perceptual. A figura 5.9 mostra a comparação com diferentes tamanhos de janela.

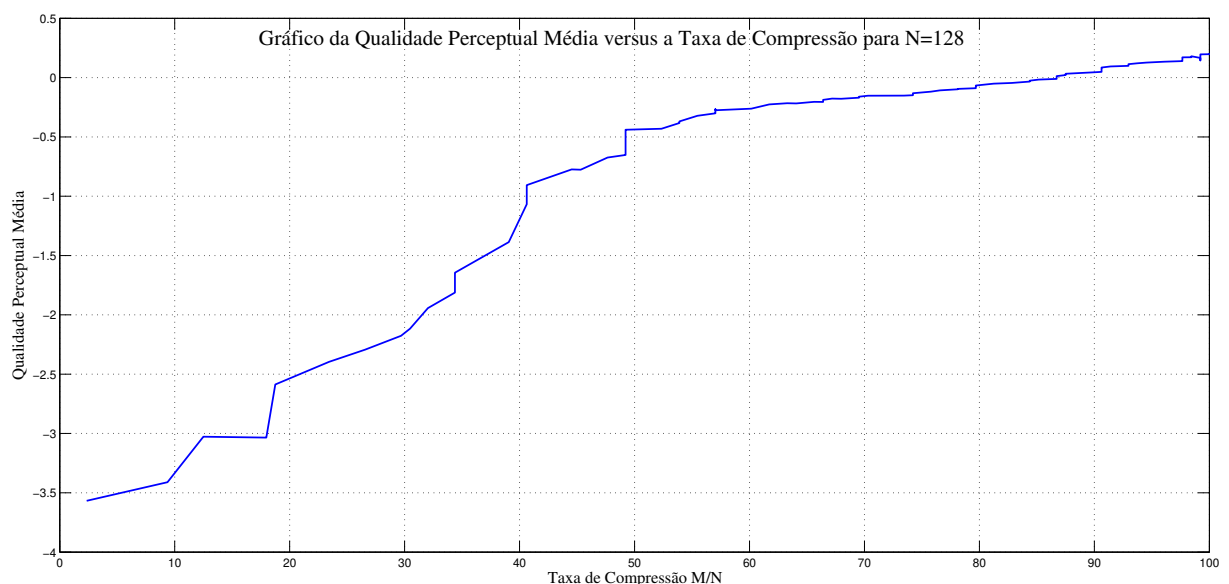


Figura 5.6: Qualidade Perceptual do Sistema para $N=128$

Apesar do codificador proposto não ser convencional, por fazer a compressão durante a amostragem, convém contextualizá-lo com os CODECs mais comuns disponíveis. A comparação é encontrada na figura 5.10. Pode-se observar que para valores de razão de compressão menores que 0,13, o sistema proposto batizado de ACS (Audio Compressed

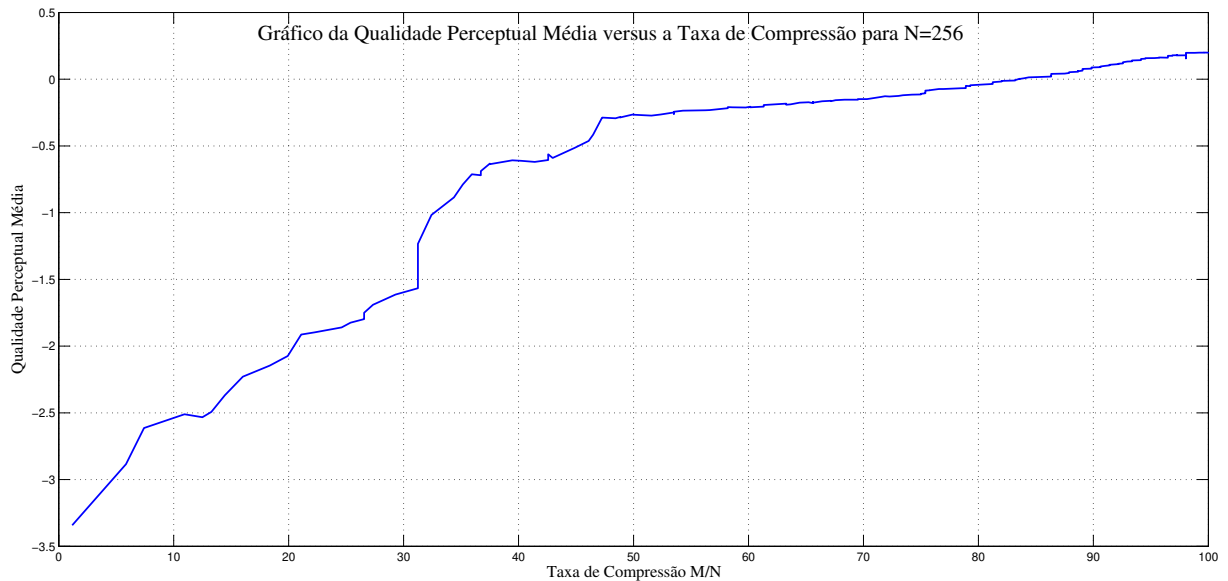


Figura 5.7: Qualidade Perceptual do Sistema para N=256

Sensing) se mostra melhor que os demais CODECs. Para valores maiores, os demais CODECs superam.

Tal comparação é apenas a título de contextualização, pois o sistema proposto não é apenas um CODEC, mas sim um sistema de amostragem que, com uma taxa de amostragem abaixo da Taxa de Nyquist, é capaz de obter certo nível de qualidade controlável, sem perda de informação de frequência, mas sim com perdas de informação sobre o sinal como um todo, dentro de uma mesma janela, além de ter as propriedades já apresentadas. A razão de compressão corresponde na prática à relação entre taxa do conversor ADC utilizado em ACS e a taxa de Nyquist.

5.5 Conclusão

Neste capítulo foram apresentados os resultados das investigações do tema da dissertação. Foram apresentados gráficos comparativos entre os objetos comparados e gráficos para futuras referências de projeto. Com base nas investigações, concluiu-se que a MDCT como base de esparsidade, a matriz de Bernoulli como base de amostragem e a minimização l_1 para reconstrução do sinal formam um conjunto de ferramentas apropriados para Amostragem Compressiva de sinais de áudio.

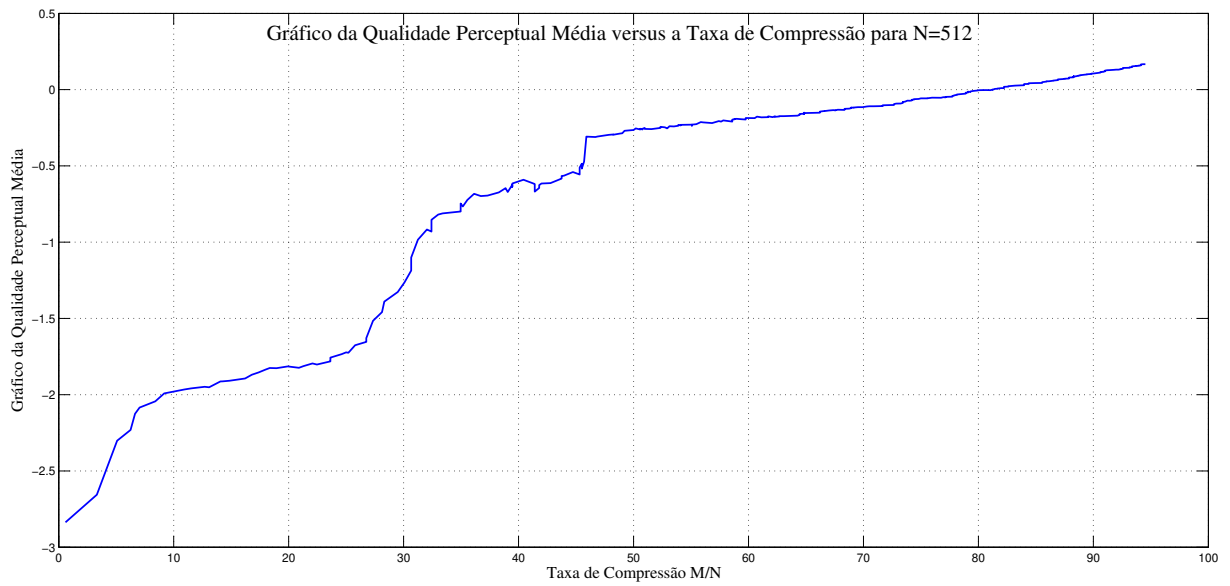


Figura 5.8: Qualidade Perceptual do Sistema para N=512

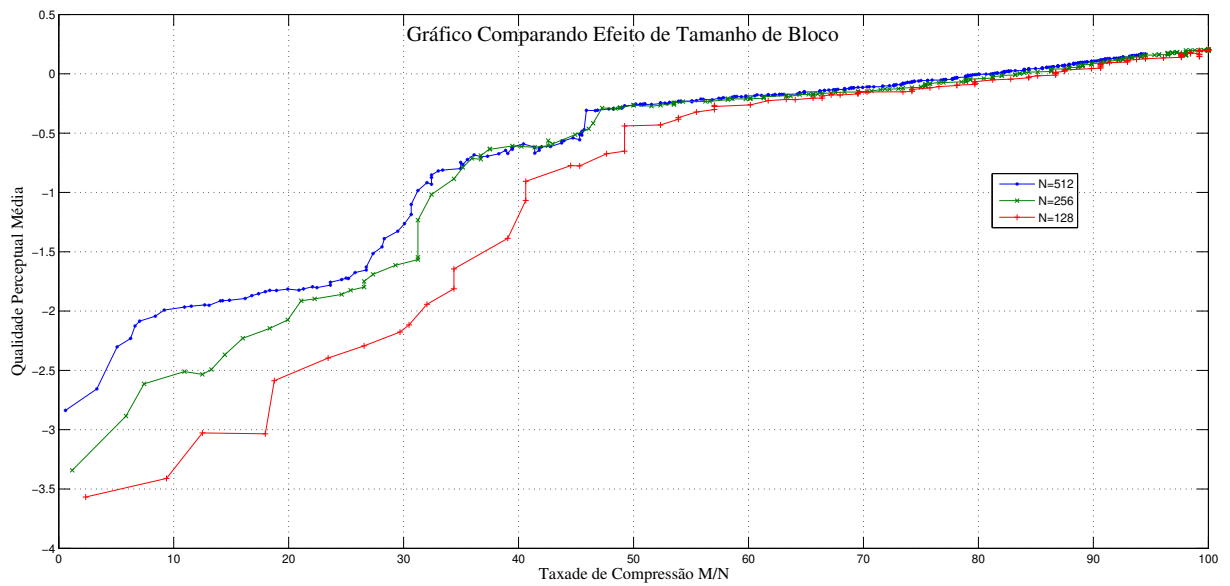


Figura 5.9: Qualidade Perceptual do Sistema para N=128, 256 e 512

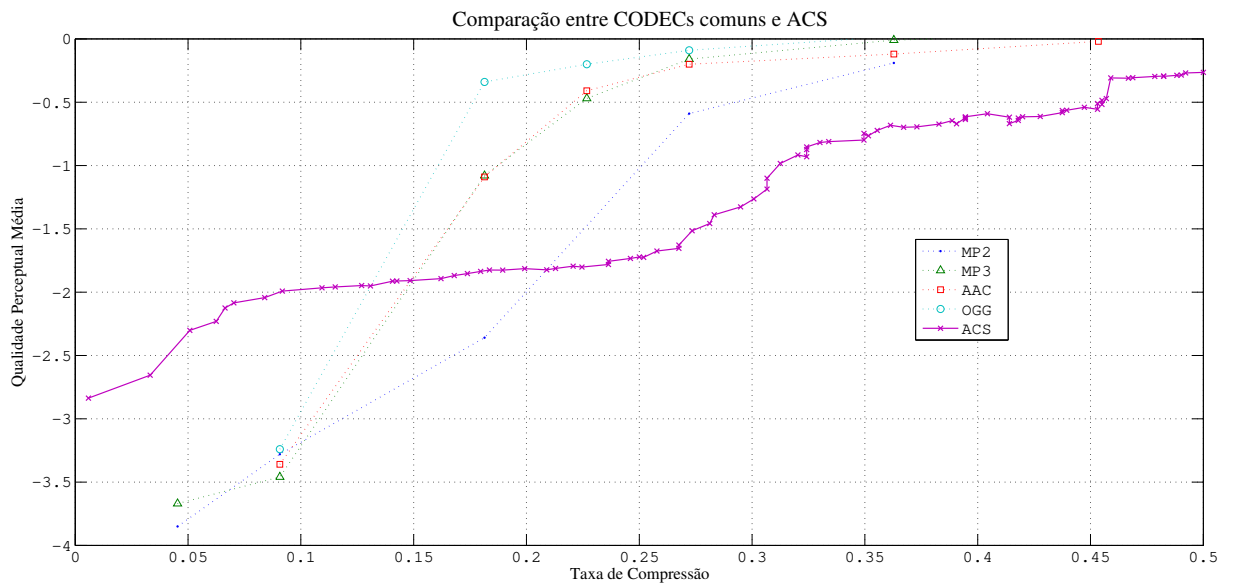


Figura 5.10: Comparação entre CODECs comuns e ACS

Capítulo 6

Considerações Finais

Este trabalho visou estudar a viabilidade de aplicação de Amostragem Compressiva na compressão de sinais de áudio e a realização de testes de avaliação perceptual de qualidade, de modo a avaliar o desempenho do sistema proposto.

Foram avaliadas bases de esparsidade que melhor se adaptam a sinais de áudio. As matrizes das transformadas DFT, DCT e MDCT foram analisadas. Também foi avaliada uma técnica de reconstrução do sinal amostrado por Amostragem Compressiva. Destas avaliações, foi proposto um sistema de compressão de áudio baseado na Amostragem Compressiva. Por fim, foi realizada uma avaliação da qualidade perceptual do mesmo por métodos objetivos.

Dentre as bases de esparsidade avaliadas, a Transformada Discreta do Cosseno Modificada (MDCT) se mostrou mais apropriada para aplicação de áudio. A base de amostragem de Bernoulli e Gaussiana apresentaram resultados equivalentes. A técnica de reconstrução por minimização da norma l_1 se mostrou satisfatória para a aplicação desejada, desprezando tempo de execução do algoritmo.

Apesar de obter resultados que evidenciam a viabilidade da aplicação de Amostragem Compressiva para áudio, por ser uma técnica recente e ainda pouco estudada, a qualidade perceptual é aquém dos codecs de áudio de estado da arte. Porém ainda se mostrou superior que os CODECS para baixas taxas de amostragem. A pesquisa da aplicação de Amostragem Compressiva a áudio tem resultados, mas ainda está na sua fase inicial. Com os testes realizados do sistema proposto, mostrou-se então a viabilidade real da aplicação das ferramentas de amostragem compressiva, beneficiando-se de suas propriedades intrínsecas como redução da taxa de amostragem, criptografia, robustez ao ruído, universalidade.

6.1 Trabalhos futuros

De modo a obter uma visão mais precisa da qualidade perceptual do sistema proposto, testes subjetivos¹ com um número reduzido de casos são aconselhados.

A propriedade de transferir praticamente toda a complexidade para o decodificador torna o sistema atraente para aplicações em redes de sensores sem fio pois os sensores demandam baixo consumo na parte do codificador, e não há problema em ter custo alto na central de fusão de dados da rede, além de oferecer codificação de canal sem custo adicional.

Para o aperfeiçoamento do sistema, pode-se ainda investigar bases de esparsidades que não se baseiam em transformadas, como por exemplo as bases supercompletas. Ainda pode-se investigar algoritmos de reconstrução que, com menos amostras, resultem em melhores aproximações para o sinal de áudio, utilizando o modelo psicoacústico e técnicas tradicionais de compressão de áudio aliadas à Amostragem Compressiva. Isto aumentaria o tempo de reconstrução do sinal, que também precisa ser otimizado.

O algoritmo de reconstrução é lento para altos valores de tamanho de janela. Novos algoritmos precisam ser estudados, possivelmente utilizando técnicas de computação paralela como GPU ou GRID.

¹Testes subjetivos não foram realizados neste trabalho pelo grande número de casos que precisavam ser testados que necessitariam de muitos recursos humanos (pessoas de ouvido treinado) e materiais (estúdio de áudio) indisponíveis no laboratório LAPS

Referências Bibliográficas

- [1] SHANNON, C. E. Communication in the presence of noise. *Proceedings of the Institute of Radio Engineers*, v. 37, p. 10–21, Janeiro 1949.
- [2] NYQUIST, H. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, v. 47, p. 617–644, Janeiro 1928.
- [3] PORTER, J. et al. Evaluating low bitrate scalable audio quality using advanced version of peaq and energy equalization approach. *BioScience*, 2005.
- [4] MALLAT, S. *A Wavelet Tour of Signal Processing, Second Edition (Wavelet Analysis & Its Applications)*. [S.l.]: Academic Press, 1999. Hardcover. ISBN 012466606X.
- [5] TAUBMAN, D. S.; MARCELLIN, M. W. *JPEG2000 : Image Compression Fundamentals, Standards and Practice*. [S.l.]: Kluwer Academic Publishers, Boston, 2002.
- [6] CANDÈS, E. J.; ROMBERG, J. Sparsity and incoherence in compressive sampling. *Inverse Problems*, v. 23, p. 969–985, 2007.
- [7] CANDÈS, E. J.; ROMBERG, J.; TAO, T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, v. 52, n. 2, p. 489–509, 2006.
- [8] DONOHO, D. L. Compressed sensing. *IEEE Transactions on Information Theory*, v. 52, n. 4, p. 1289–1306, Abril 2006.
- [9] CANDÈS, E. J.; TAO, T. Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Transactions on Information Theory*, v. 52, n. 12, p. 5406–5425, 2006.
- [10] CANDÈS, E. J.; WAKIN, M. An introduction to compressive sampling: A sensing/sampling paradigm that goes against the common knowledge in data acquisition. *IEEE Signal Processing Magazine*, v. 25, n. 2, p. 21–30, Março 2008.
- [11] DONOHO, D. L.; HUO, X. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, v. 47, p. 2845–2862, 2001.

- [12] CHEN, J.; HUO, X. Theoretical results on sparse representations of multiple-measurement vectors. *IEEE Transactions on Signal Processing*, v. 54, n. 12, p. 4634–4643, Dezembro 2006.
- [13] KRUSKAL, J. B. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra and Its Applications*, v. 18, n. 2, p. 95–138, 1977.
- [14] TROPP, J. A.; GILBERT, A. C.; STRAUSS, M. J. Algorithms for simultaneous sparse approximation. part i: Greedy pursuit. *Signal Processing*, v. 86, n. 3, p. 572–588, Março 2006.
- [15] TROPP, J. A.; GILBERT, A. C.; STRAUSS, M. J. Algorithms for simultaneous sparse approximation. part ii: Convex relaxation. *Signal Processing*, v. 86, n. 3, p. 589–602, Março 2006. Sparse Approximations in Signal and Image Processing.
- [16] CANDÈS, E. J. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathématique*, v. 346, n. 9-10, p. 589–592, Maio 2008.
- [17] CANDÈS, E. J.; ROMBERG, J.; TAO, T. Stable signal recovery from incomplete and inaccurate measurements. *Communications On Pure and Applied Mathematics*, v. 59, n. 8, p. 1207–1223, Março 2006.
- [18] CANDÈS, E. J.; TAO, T. Decoding by linear programming. *IEEE Trans. Inform. Theory*, v. 51, n. 12, p. 4203–4215, Dezembro 2005.
- [19] JERRI, A. J. The shannon sampling theorem - its various extensions and applications: A tutorial review. *Proceedings of the IEEE*, v. 65, n. 11, p. 1565–1596, Novembro 1977.
- [20] DONOHO, D. L.; ELAD, M. Maximal sparsity representation via l_1 minimization. In: *Proceedings of National Academy of Science*. [S.l.: s.n.], 2003. v. 100, p. 2197–2202.
- [21] MALLAT, S.; ZHANG, Z. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, v. 41, n. 12, p. 3397–3415, 1993.
- [22] CHEN, S. S. et al. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, v. 20, p. 33–61, 1998.
- [23] DUARTE, M. F. et al. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, v. 25, n. 2, p. 83–91, Março 2008.
- [24] KIROLOS, S. et al. Analog-to-information conversion via random demodulation. In: *Design, Applications, Integration and Software, 2006 IEEE Dallas/CAS Workshop on*. [S.l.: s.n.], 2006. p. 71–74.

- [25] TROPP, J. et al. Random filters for compressive sampling and reconstruction. In: *IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP 2006*. [S.l.: s.n.], 2006. v. 3, p. III-872-III-875. ISSN 1520-6149.
- [26] LUSTIG, M.; DONOHO, D.; PAULY, J. M. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine*, Magnetic Resonance Systems Research Laboratory, Department of Electrical Engineering, Stanford University, Stanford, California; Statistics Department, Stanford University, Stanford, California, v. 58(6), p. 1182-1195, Dezembro 2007.
- [27] CANDÈS, E. J. et al. Error correction via linear programming. *Annual IEEE Symposium on Foundations of Computer Science*, IEEE Computer Society, Los Alamitos, CA, USA, p. 295-308, 2005.
- [28] CANDÈS, E. J.; RANDALL, P. Highly robust error correction by convex programming. *IEEE Transactions on Information Theory*, v. 54, n. 7, p. 2829-2840, Julho 2008.
- [29] DUARTE, M. F. et al. Joint sparsity models for distributed compressed sensing. *Online Proceedings of the Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, Novembro 2005.
- [30] DRORI, I. Compressed video sensing. *BMVA Symposium on 3D Video - Analysis, Display, and Applications*, 2008.
- [31] RACHLIN, Y.; BARON, D. The secrecy of compressive sensing measurements. *Proceedings of the 46th Allerton Conference on Communication, Control, and Computing, Monticello, IL*, Setembro 2008.
- [32] BOBIN, J.; STARCK, J.-L.; OTTENSAMER, R. Compressed sensing in astronomy. *IEEE Journal of Selected Topics in Signal Processing*, v. 2, n. 5, p. 718-726, Outubro 2008.
- [33] GRIFFIN, A.; TSAKALIDES, P. Compressed sensing of audio signals using multiple sensors. *Lausanne, Switzerland: Proceedings of the 16th European Signal Processing Conference (EUSIPCO'08)*, Agosto 2008.
- [34] CARMI, A.; KANEVSKY, D.; RAMABHADRAN, B. *Lossy Speech Compression Via Compressed Sensing-Based Kalman Filtering*. [S.l.], Junho 2009.
- [35] GRIBONVAL, R. *Sur quelques problèmes mathématiques de modélisation parcimonieuse*. [S.l.], Outubro 2007.

- [36] BLUMENSATH, T.; DAVIES, M. Sparse and shift-invariant representations of music. In: *IEEE Transactions on Audio, Speech and Language Processing*. [S.l.: s.n.], 2006. p. 50–57.
- [37] PLUMBLEY, M. D. et al. *Musical Audio Analysis Using Sparse Representations*. 2006.
- [38] NOLL, P. Mpeg digital audio coding. *IEEE Signal Processing Magazine*, v. 14, p. 59–81, Setembro 1997.
- [39] BRANDENBURG, K. Evaluation of quality of audio encoding at low bit rates. *82nd AES Convention of the Audio Engineering Society*, p. 2433, 1987.
- [40] BEERENDS, J. G.; STEMERDINK, J. A. Perceval: Perceptual evaluation of the quality of audio signals. *Journal of Audio Engineering Society*, p. 21–31, Fevereiro 1992.
- [41] THIEDE, T.; KABOT, E. New perceptual quality measure for the bitrate reduced audio. *100th Convention of the Audio Engineering Society*, p. 4280, 1996.
- [42] BEERENDS, J. G.; STEMERDINK, J. A. A perceptual audio quality measure based on psychoacoustic sound representation. *Journal of Audio Engineering Society*, p. 963–978, dezembro 1992.
- [43] THIEDE, T. et al. Peaq - the itu standard for objective measurement of perceived audio quality. *Journal of the Audio Engineering Society*, v. 48, p. 3–29, 2000.
- [44] SOUND Quality Assessment Material recordings for subjective tests. Disponível em <http://tech.ebu.ch/publications/sqamcd>, Último Acesso em abril 2010.
- [45] PRINCEN, J.; BRADLEY, A. Analysis/synthesis filter bank design bases on time domain aliasing cancellation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-34, p. 1153–1161, Outubro 1986.
- [46] CANDÈS, E. J.; TAO, T. Blind multi-band signal reconstruction: Compressed sensing for analog signals. *IEEE Transactions on Information Theory*, v. 51, n. 12, p. 4203–4215, Dezembro 2005.
- [47] SRIVASTAVA, M.; CULLER, D.; ESTRIN, D. Overview of wireless sensor networks. *IEEE Computer Society, Special Issue in Sensor Networks*, Agosto 2004.