

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E
INFORMÁTICA INDUSTRIAL

SOLIVAN ARANTES VALENTE

**RECONSTRUÇÃO DE IMAGENS DE ULTRASSOM USANDO
ESPARSIDADE – MÉTODOS ITERATIVOS RÁPIDOS**

TESE DE DOUTORADO

CURITIBA

2017

SOLIVAN ARANTES VALENTE

**RECONSTRUÇÃO DE IMAGENS DE ULTRASSOM USANDO
ESPARSIDADE – MÉTODOS ITERATIVOS RÁPIDOS**

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Doutor em Ciências” – Área de Concentração: Engenharia Biomédica.

Orientador: Prof. Dr. Marcelo Victor Wüst Zibetti

Co-orientador: Prof. Dr. Daniel Rodrigues Pipa

CURITIBA

2017

Dados Internacionais de Catalogação na Publicação

V154r
2017 Valente, Solivan Arantes
Reconstrução de imagens de ultrassom usando esparsidade :
métodos iterativos rápidos / Solivan Arantes Valente.-- 2017.
116 f. : il. ; 30 cm

Texto em português com resumo em inglês
Disponível também via World Wide Web
Tese (Doutorado) - Universidade Tecnológica Federal do Pa-
raná. Programa de Pós-graduação em Engenharia Elétrica e In-
formática Industrial, Curitiba, 2017
Bibliografia: f. 110-116

1. Ultrassonografia. 2. Diagnóstico por ultrassom. 3. Algorit-
mos computacionais. 4. Ultrassonografia – Imagem – Avaliação.
5. Engenharia elétrica – Teses. I. Zibetti, Marcelo Victor Wüst. II.
Pipa, Daniel Rodrigues. III. Universidade Tecnológica Federal do
Paraná. Programa de Pós-Graduação em Engenharia Elétrica e
Informática Industrial. IV. Título.

CDD: Ed. 22 – 621.3

Biblioteca Central da UTFPR, Câmpus Curitiba

TERMO DE APROVAÇÃO DE TESE Nº 158

A Tese de Doutorado intitulada “**Reconstrução de Imagens de Ultrassom Usando Esparsidade – Métodos Iterativos Rápidos**”, defendida em sessão pública pelo(a) candidato(a) **Solivan Arantes Valente**, no dia 23 de agosto de 2017, foi julgada para a obtenção do título de Doutor em Ciências, área de concentração Engenharia Biomédica, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial.

BANCA EXAMINADORA:

Prof(a). Dr(a). Daniel Rodrigues Pipa - Presidente – (UTFPR)

Prof(a). Dr(a). Hugo Vieira Neto – (UTFPR)

Prof(a). Dr(a). Priscila Savulski Ferreira – (UTFPR)

Prof(a). Dr(a). Elias Salomão Helou Neto – (USP)

Prof(a). Dr(a). Diego Barreto Haddad - (CEFET-RJ)

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 23 de agosto de 2017.

Dedico este singelo trabalho aos meus pais, Solerte e Lindamir (*in memoriam*) que me ensinaram as primeiras letras e os valores e princípios que me definem. O esforço aqui empregado não se equipara à sua abnegação e dedicação aos filhos, mesmo em períodos de dificuldade. Minha eterna gratidão e reconhecimento pelo contínuo, incansável e incondicional apoio aos meus estudos, que me possibilitou chegar até aqui. Com muito amor, do seu filho Solivan.

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus pelo dom da vida, por sua misericórdia e graça, e por seu cuidado para comigo ao longo de toda a minha jornada.

Sou também muito grato à minha amada esposa Tânia, por seu incansável apoio, compreensão e amor dedicados a mim nos últimos 23 anos. Sem a sua companhia e incentivo este trabalho não teria sido completado.

Agradeço de modo muito especial ao meu caríssimo orientador, Prof. Dr. Marcelo Zibetti, por sua grande dedicação, apoio, incentivo e orientação ao longo desses anos. Que o seu exemplo inspire outros professores, como certamente tem inspirado a mim.

Meus igualmente sinceros agradecimentos ao meu co-orientador, Prof. Dr. Daniel Pipa, pelo apoio, *insights* e orientação.

Expresso também minha gratidão aos professores do CPGEI, em especial ao Prof. Dr. Joaquim Maia, e aos membros do grupo de pesquisa LIPRO, pela colaboração e compartilhamento de seus trabalhos e conhecimento.

Meu muitíssimo obrigado aos membros da banca avaliadora e a todos que, de modo direto ou indireto, contribuíram para a lapidação deste trabalho, que materializa um sonho pessoal.

“I do not know what I may appear to the world, but to myself I seem to have been only like a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.” Isaac Newton (1643 – 1727)

RESUMO

VALENTE, Solivan. RECONSTRUÇÃO DE IMAGENS DE ULTRASSOM USANDO ESPARSIDADE – MÉTODOS ITERATIVOS RÁPIDOS. 117 f. Tese de Doutorado – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2017.

Este trabalho contribui para a busca de métodos rápidos para reconstrução esparsa em ultrassonografia. O objetivo é alcançado em três etapas: a validação de um modelo discreto de aquisição, uma avaliação comparativa de algoritmos adequados ao problema e uma proposição de aceleração para um dos métodos de melhor desempenho. A estratégia de validação do modelo consiste em reconstruções a partir de dados sintéticos de resultado conhecido e subsequente validação com dados reais, coletados por uma plataforma de pesquisa em ultrassom com um *phantom* de uso profissional. As reconstruções são realizadas por um conjunto selecionado de algoritmos iterativos de otimização convexa, que têm seus parâmetros, resultados e desempenhos analisados. O trabalho propõe a aceleração do método ADMM (*Alternating Direction Method of Multipliers*) que está entre os de melhor desempenho em termos de custo computacional, e que pode dobrar sua velocidade inicial de convergência com a modificação proposta. Como a aceleração também pode ser utilizada em outras aplicações do ADMM, a modificação proposta é validada em quatro casos de estudo, sendo dois em ultrassonografia e dois em imageamento por ressonância magnética.

Palavras-chave: Ultrassom, reconstrução de imagens, otimização convexa, métodos iterativos.

ABSTRACT

VALENTE, Solivan. ULTRASONIC IMAGE RECONSTRUCTION USING SPARSITY – FAST ITERATIVE METHODS. 117 f. Tese de Doutorado – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2017.

This study contributes to the search for fast iterative methods for ultrasonic sparse image reconstruction. The goal is achieved in three steps: the validation of a discrete acquisition model, a comparative evaluation of algorithms suitable to the problem and an acceleration proposal for one of the best performing methods. The model validation strategy consists of image reconstructions from synthetic data with previously known results, and subsequent validation with real data, collected by an ultrasound research platform with a professional phantom. The reconstructions are performed by a selected set of iterative algorithms of convex optimization, which have their parameters, results and performances analyzed. This study proposes the acceleration of the ADMM (Alternating Direction Method of Multipliers), which is among the best performing methods in terms of computational cost, and which can have its initial convergence speed doubled by the proposed modification. Since the acceleration can also be used in other applications of ADMM, the proposed modification is validated in four cases of study: two in ultrasonography and two in magnetic resonance imaging.

Keywords: Ultrasound, image reconstruction, convex optimization, iterative methods.

LISTA DE FIGURAS

FIGURA 1	– Curvas das PDF Gaussiana e Laplaciana e das distâncias associadas.	33
FIGURA 2	– Forma geral da Curva L	36
FIGURA 3	– Formação do feixe por focagem dinâmica em ultrassonografia	39
FIGURA 4	– Representação esquemática das estruturas dos vetores e da matriz \mathbf{H}	43
FIGURA 5	– Representação visual dos elementos da matriz \mathbf{H}	44
FIGURA 6	– Conjunto de dados \mathbf{g} – refletor pontual único no centro da ROI	45
FIGURA 7	– Conjunto de dados \mathbf{g} – refletores em todos os pontos da ROI	45
FIGURA 8	– Forma geral da Curva L com as normas ℓ_1 e ℓ_2	60
FIGURA 9	– Curva L – Conjunto de dados sintético	61
FIGURA 10	– Curva L – Conjunto de dados real	61
FIGURA 11	– Avaliação dos parâmetros α e η do OMFISTA – geral	63
FIGURA 12	– Avaliação dos parâmetros α e η do OMFISTA – detalhe	63
FIGURA 13	– Equipamentos de ultrassom	65
FIGURA 14	– Vista esquemática dos <i>targets</i> dentro do <i>phantom</i>	66
FIGURA 15	– Vista ampliada dos <i>targets</i> na grade de <i>pixels</i> da imagem de referência ...	67
FIGURA 16	– Atribuição de valores aos <i>pixels</i> da imagem de referência para o <i>Target 1</i> .	67
FIGURA 17	– Conjunto de dados sintético $\mathbf{g}_{sint} = \mathbf{H}\mathbf{f}_{sint} + \mathbf{e}$	68
FIGURA 18	– Conjunto de dados real \mathbf{g}_{real}	68
FIGURA 19	– Curva de razões da métrica API (<i>Array Performance Indicator</i>)	70
FIGURA 20	– Imagens reconstruídas a partir de conjunto de dados sintético	71
FIGURA 21	– Erros de estimação $\ \hat{\mathbf{f}}_k - \mathbf{f}^*\ _2 \times$ iterações – Caso sintético	72
FIGURA 22	– Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ iterações – Caso sintético	73
FIGURA 23	– Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ tempo – Caso sintético	74
FIGURA 24	– Imagens reconstruídas a partir do conjunto de dados real	75
FIGURA 25	– Erros de estimação $\ \hat{\mathbf{f}}_k - \mathbf{f}^*\ _2 \times$ iterações – Caso real	76
FIGURA 26	– Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ iterações – Caso real	77
FIGURA 27	– Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ tempo – Caso real	78
FIGURA 28	– Imagens reconstruídas a partir de outro conjunto de dados real, com outra ROI	79
FIGURA 29	– Erros de estimação $\ \hat{\mathbf{f}}_k - \mathbf{f}^*\ _2 \times$ iterações – Outro caso real	80
FIGURA 30	– Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ iterações – Outro caso real	81
FIGURA 31	– Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ tempo – Outro caso real	81
FIGURA 32	– Conjunto de dados real	82
FIGURA 33	– Outro caso real	82
FIGURA 34	– Estimativa do fator β para os casos 1 e 2 (Ultrassom).	93
FIGURA 35	– Estimativa de α e δ para o caso 1 (Ultrassom).	94
FIGURA 36	– Caso 1 (Ultrassom) - Custo \times iterações (esquerda) e tempo (direita).	95
FIGURA 37	– Caso 2 (Ultrassom) - Custo \times iterações (esquerda) e tempo (direita).	95
FIGURA 38	– Casos 1 e 2 (Ultrassom) - Erros de estimação \times iterações.	96
FIGURA 39	– Caso 1 (Ultrassom) - Imagens reconstruídas.	97
FIGURA 40	– Caso 2 (Ultrassom) - Imagens reconstruídas.	97
FIGURA 41	– Estimativa do fator β para o caso 3 (MRI).	99

FIGURA 42	– Caso 3 (MRI) - Custo \times iterações (esquerda) e tempo (direita).	100
FIGURA 43	– Caso 3 (MRI) - Erros de estimação \times iterações.	101
FIGURA 44	– Caso 3 (MRI) - Imagens reconstruídas.	102
FIGURA 45	– Estimativa do fator β para o caso 4 (MRI TV).	103
FIGURA 46	– Caso 4 (MRI TV) - Custo \times iterações (esquerda) e tempo (direita).	104
FIGURA 47	– Caso 4 (MRI TV) - Erros de estimação \times iterações.	105
FIGURA 48	– Caso 4 (MRI TV) - Imagens reconstruídas.	106

LISTA DE TABELAS

TABELA 1	–	Parâmetros principais do sistema de imageamento por ultrassom	64
TABELA 2	–	Valores da API para imagens reconstruídas com 30 iterações – Caso sintético	72
TABELA 3	–	Valores da API para imagens reconstruídas com 30 iterações – Caso real	76
TABELA 4	–	Valores da API para imagens reconstruídas com 30 iterações – Outro caso real	80
TABELA 5	–	Parâmetros β , α e δ para os casos 1 e 2	94
TABELA 6	–	Valores API para as imagens de ultrassom reconstruídas.	98
TABELA 7	–	Parâmetros β , α e δ para o caso 3	100
TABELA 8	–	Parâmetros β , α e δ para o caso 4	104
TABELA 9	–	Modificação do parâmetro de penalidade na primeira iteração do ADMM	106

LISTA DE QUADROS

QUADRO 1	–	Algoritmo FISTA	50
QUADRO 2	–	Algoritmo MFISTA	51
QUADRO 3	–	Algoritmo OMFISTA	51
QUADRO 4	–	Algoritmo NLCG	53
QUADRO 5	–	Algoritmo IRLS	54
QUADRO 6	–	Algoritmo ADMM	56
QUADRO 7	–	Algoritmo ADMM acelerado	90

LISTA DE SIGLAS

ADMM	<i>Alternating Direction Method of Multipliers</i>
API	<i>Array Performance Indicator</i>
BPDN	<i>Basis Pursuit Denoising</i>
CG	<i>Conjugate Gradient</i>
CS	<i>Compressed Sensing</i>
DY	<i>Método de cálculo do parâmetro β proposto por Dai e Yuan</i>
FISTA	<i>Fast Iterative Shrinkage-Thresholding Algorithm</i>
GCV	<i>Generalized Cross-Validation</i>
GSVD	<i>Generalized Singular Value Decomposition</i>
IPB	<i>Inverse Problem-Based</i>
IRLS	<i>Iteratively Re-weighted Least Squares</i>
IRLS-OLS	<i>Iteratively Re-weighted Least Squares with Optimal Line Search</i>
IRLS-CG	<i>Iteratively Re-weighted Least Squares with Conjugate Gradient</i>
IRLS-CG-OLS	<i>Iteratively Re-weighted Least Squares with Conjugate Gradient and Optimal Line Search</i>
ISTA	<i>Iterative Shrinkage-Thresholding Algorithm</i>
JMAP	<i>Joint Maximum a Posteriori</i>
LASSO	<i>Least Absolute Shrinkage and Selection Operator</i>
LIPRO	<i>Lineup for Image Processing and Reconstruction</i>
LTI	<i>Linear Time-Invariant</i>
MAP	<i>Maximum a Posteriori</i>
MFISTA	<i>Monotone Fast Iterative Shrinkage-Thresholding Algorithm</i>
ML	<i>Maximum Likelihood</i>
MRI	<i>Magnetic Resonance Imaging</i>
NLCG	<i>Nonlinear Conjugate Gradient</i>
NLCG-OLS	<i>Nonlinear Conjugate Gradient with Optimal Line Search</i>
OLS	<i>Optimal Line Search</i>
OMFISTA	<i>Over-Relaxation of Monotone Fast Iterative Shrinkage-Thresholding Algorithm</i>
OMFISTA-OLS	<i>Over-Relaxation of Monotone Fast Iterative Shrinkage-Thresholding Algorithm with Optimal Line Search</i>
PDF	<i>Probability Density Function</i>
PSF	<i>Point Spread Function</i>
ROI	<i>Region Of Interest</i>
SNR	<i>Signal-to-Noise Ratio</i>
SSF	<i>Separable Surrogate Functionals</i>
SVD	<i>Singular Value Decomposition</i>
TGC	<i>Time-Gain Compensation</i>
TV	<i>Total Variation</i>

LISTA DE SÍMBOLOS

\mathbf{f}	Vetor da imagem no espaço \mathbb{R}^m
\mathbf{g}	Vetor de dados no espaço \mathbb{R}^n
\mathbf{H}	Matriz de aquisição no espaço $\mathbb{R}^{n \times m}$
$\Psi(\mathbf{f})$	Função custo convexa
$R(\mathbf{f})$	Termo de regularização da função custo
ℓ_1	Norma L_1
ℓ_2	Norma Euclidiana
λ	Parâmetro do termo de regularização da função custo
κ	Fração do parâmetro λ máximo
c	Parâmetro dos algoritmos da família ISTA
ρ	Parâmetro de penalidade do algoritmo ADMM
ξ	Parâmetro de regularização de Tikhonov
α	Fator de multiplicação para o parâmetro de regularização de Tikhonov
β	Fator de multiplicação para o parâmetro de penalidade do algoritmo ADMM no termo quadrático
δ	Fator de multiplicação para o parâmetro de penalidade do algoritmo ADMM no operador proximal
Λ	Comprimento de onda de ultrassom

SUMÁRIO

1 INTRODUÇÃO	17
1.1 OBJETIVOS	18
1.1.1 Objetivo Geral	18
1.1.2 Objetivos Específicos	18
1.2 CONTRIBUIÇÕES	19
1.3 ESTRUTURA DA TESE	19
2 RECONSTRUÇÃO DE IMAGENS	21
2.1 ASPECTOS FUNDAMENTAIS	21
2.2 RECONSTRUÇÃO POR PROBLEMAS INVERSOS	22
2.2.1 Métodos Diretos	24
2.2.2 Função Custo, Normas e Esparsidade	25
2.2.3 Métodos Iterativos	27
2.3 REGULARIZAÇÃO DE PROBLEMAS MAL-POSTOS	28
2.3.1 Regularização de Tikhonov	29
2.3.2 Estimadores de Máxima Verossimilhança e <i>Maximum a Posteriori</i>	32
2.3.3 Métodos de Escolha do Parâmetro de Regularização	35
2.4 RECONSTRUÇÃO DE IMAGENS EM ULTRASSONOGRRAFIA	37
2.4.1 Reconstrução por <i>Beamforming</i>	37
2.4.2 Modelo Matemático de Aquisição	38
3 ALGORITMOS ITERATIVOS DE RECONSTRUÇÃO	47
3.1 BUSCA EM LINHA (<i>LINE SEARCH</i>)	48
3.2 ALGORITMOS DA FAMÍLIA ISTA	48
3.3 GRADIENTE CONJUGADO NÃO LINEAR	52
3.4 QUADRADOS MÍNIMOS ITERATIVAMENTE REPONDERADOS	53
3.5 MULTIPLICADORES DE LAGRANGE COM DIREÇÕES ALTERNADAS	55
4 ESTUDO DE CASO EM ULTRASSONOGRRAFIA ESPARSA	58
4.1 AVALIAÇÃO DOS PARÂMETROS DOS ALGORITMOS	59
4.1.1 Parâmetro λ – Ajuste da Esparsidade	59
4.1.2 Parâmetro β – NLCG	60
4.1.3 Parâmetros α e η – OMFISTA	62
4.1.4 Parâmetro ρ – ADMM	62
4.2 DESCRIÇÃO DOS EXPERIMENTOS – MATERIAIS E MÉTODOS	64
4.2.1 Conjunto de Dados Sintético	65
4.2.2 Conjunto de Dados Real	67
4.3 RESULTADOS DOS EXPERIMENTOS	68
4.3.1 Resultados para o Conjunto de Dados Sintético	70
4.3.2 Resultados para o Conjunto de Dados Real	74
4.3.3 Resultados para Outro Conjunto de Dados Real	77
4.3.4 Sobre o Critério de Parada	79
4.4 CONCLUSÕES SOBRE OS EXPERIMENTOS	81
5 PROPOSTA DE ACELERAÇÃO DO MÉTODO ADMM	84

5.1	GENERALIZAÇÃO DO TERMO DE REGULARIZAÇÃO	84
5.1.1	O Operador Proximal	85
5.1.2	Versão Proximal do ADMM	85
5.1.3	Parâmetro de Penalidade do ADMM	87
5.1.4	Versão Proximal do FISTA	88
5.2	ACELERAÇÃO DA PRIMEIRA ITERAÇÃO DO ADMM	89
5.3	EXPERIMENTOS EM ULTRASSONOGRRAFIA ESPARSA	92
5.3.1	Descrição dos Experimentos – Materiais e Métodos	92
5.3.2	Estimativa dos Parâmetros	93
5.3.3	Resultados da Aceleração em Ultrassonografia Esparsa	94
5.4	EXPERIMENTO EM RESSONÂNCIA MAGNÉTICA NO DOMÍNIO WAVELET .	98
5.4.1	Descrição do Experimento – Materiais e Métodos	98
5.4.2	Estimativa dos Parâmetros	99
5.4.3	Resultados da Aceleração em MRI	100
5.5	EXPERIMENTO EM RESSONÂNCIA MAGNÉTICA COM <i>TOTAL VARIATION</i> .	102
5.5.1	Descrição do Experimento – Materiais e Métodos	103
5.5.2	Estimativa dos Parâmetros	103
5.5.3	Resultados da Aceleração em MRI TV	104
5.6	CONCLUSÕES SOBRE A ACELERAÇÃO DO ADMM	105
6	CONCLUSÃO E TRABALHOS FUTUROS	108
	REFERÊNCIAS	110

1 INTRODUÇÃO

As imagens são, sem dúvida, onipresentes nas sociedades modernas. Sejam estáticas ou em movimento, artísticas ou técnicas, acadêmicas ou comerciais, profissionais ou de entretenimento, o seu uso é tão intenso que já há algumas décadas diz-se que vivemos na “civilização da imagem” (JOLY, 2006). Entre outros motivos, é possível que a razão esteja ligada à forma como entendemos e interpretamos o mundo que nos cerca; segundo o filósofo grego Aristóteles, as imagens são requeridas em todo processo de pensamento: “*Ora, no que concerne à alma pensante, as imagens tomam o lugar das percepções diretas [...] a alma nunca pensa sem uma imagem mental.*” (MANGUEL, 2001).

Por serem tão significativas para o ser humano, o estudo das imagens remonta à Grécia antiga. No início do Século XX, pesquisadores como Charles Sanders Peirce criaram uma Ciência dedicada ao estudo específico de signos e ícones, a Semiótica. Embora a Semiótica não seja limitada às imagens, ela deu origem em meados do mesmo século a um ramo específico de estudo da comunicação visual, a chamada Semiologia da Imagem (JOLY, 2006).

Nas telecomunicações, nos mais variados tipos de mídia, as imagens são um meio primário de veiculação de mensagens, publicidade, notícias etc. Em inúmeros campos da Ciência e da Tecnologia, as imagens são um veículo fundamental para a investigação de fenômenos e para o desenvolvimento de ferramentas para controlá-los. A ciência dos materiais e a pesquisa biomédica, por exemplo, não poderiam avançar sem as imagens microscópicas; as reservas de petróleo, essenciais para a economia, são em geral encontradas por imageamento sísmico; o diagnóstico médico, em muitos casos, apoia-se em imagens obtidas por sistemas de imageamento dedicados a este propósito; o controle da produção agrícola e o uso militar são também exemplos do uso extensivo de sistemas de imageamento, entre muitos outros (BARRETT; MYERS, 2004).

Dada a importância das imagens, é natural que se busque o desenvolvimento e o aperfeiçoamento de sistemas e equipamentos capazes de representar características ou propriedades de objetos e ambientes por meio de imagens. Esses sistemas são chamados

de *sistemas de imageamento*. Na realidade, é virtualmente impossível que esses sistemas representem perfeitamente *o objeto* observado; eles, em geral, permitem-nos apenas aprender algo sobre ele e caracterizá-lo por um conjunto finito de parâmetros que o descrevem aproximadamente (BARRETT; MYERS, 2004). Entretanto, mesmo com essas limitações, as imagens obtidas são de grande valia para, por exemplo, auxiliar um profissional da saúde a diagnosticar uma patologia ou para indicar a localização de uma falha em um equipamento industrial.

Dentre os diversos sistemas de imageamento existentes, e dentre as mais variadas técnicas de formação das imagens, este trabalho está focado em sistemas de ultrassom cujas imagens são formadas por técnicas de reconstrução baseadas em problemas inversos. Pela característica de esparsidade¹ que as imagens de ultrassonografia podem apresentar, voltamos nossa atenção para técnicas de otimização convexa, com funções custo que combinam as normas ℓ_2 e ℓ_1 . A explicação desses termos e técnicas, bem como a razão de tais escolhas, são descritas ao longo dos capítulos que se seguem.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

O objetivo geral deste trabalho é contribuir com métodos iterativos rápidos de reconstrução esparsa em ultrassonografia, por meio da validação do modelo de aquisição, da avaliação comparativa de métodos adequados ao problema e da proposição de melhoria de um método selecionado.

1.1.2 OBJETIVOS ESPECÍFICOS

- a) Descrever o modelamento discreto do sistema de aquisição em ultrassonografia bidimensional;
- b) Avaliar o modelo do sistema por meio de reconstruções com dados sintéticos de resultado conhecido, e por meio de dados reais para os quais são conhecidas as reconstruções pela técnica tradicional de *beamforming*;
- c) Avaliar o desempenho de um conjunto de algoritmos iterativos adequados à solução do problema e identificar o(s) mais apropriado(s); e

¹Imagens *esparsas* são aquelas em que os elementos visuais são pequenas regiões com *pixels* concentrados, que representam objetos imageados de natureza pontual ou de pequena seção transversal. Parte significativa dos *pixels* da imagem reconstruída é, portanto, nula e eles são representados na imagem como pontos negros.

- d) Propor e validar a aceleração de um dos métodos de melhor desempenho dentre os avaliados.

1.2 CONTRIBUIÇÕES

Este trabalho apresenta as seguintes contribuições:

- 1) Proposição de uma estratégia de validação do modelo do sistema de aquisição em ultrassonografia bidimensional;
- 2) Utilização da métrica *Array Performance Indicator* (API) como auxílio ao método da Curva L na escolha de um parâmetro de regularização adequado para reconstruções esparsas;
- 3) Identificação de métodos iterativos rápidos para a reconstrução de imagens esparsas em ultrassonografia; e
- 4) Proposição e validação de uma forma de aceleração do algoritmo *Alternating Direction Method of Multipliers* (ADMM).

As três primeiras contribuições estão no artigo (VALENTE et al., 2017), publicado² no *Journal Sensors* do *Multidisciplinary Digital Publishing Institute* (MDPI). A quarta contribuição está em um artigo em processo de submissão ao *Journal Transactions on Computational Imaging*, da *Signal Processing Society* do *Institute of Electrical and Electronics Engineers* (IEEE).

1.3 ESTRUTURA DA TESE

O Capítulo 2 apresenta os conceitos fundamentais da reconstrução de imagens por problemas inversos, abordando de forma sucinta o problema de regularização. O modelo discreto do sistema de aquisição é apresentado, e o método tradicional de imageamento em ultrassonografia (o *beamforming*) é brevemente descrito.

O Capítulo 3 enquadra o problema da reconstrução de imagens como um caso particular de otimização convexa, e apresenta os métodos avaliados neste estudo. Esse capítulo também aborda o conceito de Busca em Linha (*Line Search*), utilizado nas variantes dos métodos considerados identificadas com o sufixo “OLS”.

²Artigo disponível *online* em: <http://www.mdpi.com/1424-8220/17/3/533>

No Capítulo 4 o modelo de aquisição é validado com base em um conjunto de dados sintéticos e, em seguida, em dois conjuntos de dados reais coletados por meio de uma plataforma de testes de ultrassom. O capítulo descreve os objetos imageados, a parametrização dos algoritmos e os resultados obtidos.

O Capítulo 5 apresenta uma proposta de aceleração do algoritmo ADMM. A proposição é avaliada em quatro experimentos, sendo dois em ultrassonografia esparsa e dois em ressonância magnética.

Por fim, o Capítulo 6 apresenta as conclusões gerais e algumas sugestões de trabalhos futuros.

2 RECONSTRUÇÃO DE IMAGENS

2.1 ASPECTOS FUNDAMENTAIS

As imagens obtidas nos sistemas de imageamento são resultado da aplicação de ferramentas matemáticas conhecidas como *métodos de reconstrução*. Esses métodos são formas de manipular os dados de sensores, sondas ou outros dispositivos, capazes de fazer a aquisição de sinais provenientes do objeto imageado. Esses sinais são então tratados matematicamente, dando origem a imagens que nos mostram certas propriedades do objeto.

Dependendo da natureza do objeto imageado e das características dele que desejamos observar, os sistemas de imageamento podem ser construídos de muitas maneiras, utilizando sinais físicos de várias naturezas. Algumas taxonomias destes sistemas são apresentadas por Barrett e Myers (2004) e incluem classificações:

- Quanto ao tipo de radiação ou campo usado para formar a imagem;
- Quanto à propriedade do objeto representada na imagem;
- Quanto ao mecanismo de formação da imagem ou do conjunto de dados;
- Quanto ao modo de formação da imagem (direto ou indireto); e
- Quanto ao modo de captura do sinal da fonte da imagem (passivo ou ativo).

Apesar de distintos em relação a um ou mais desses critérios, todos os sistemas de imageamento podem ser modelados como uma operação de mapeamento (linear ou não linear) entre dois espaços vetoriais. Tanto a imagem que representa o objeto quanto os dados coletados são representados por vetores nesses espaços. Neste trabalho limitamos nossa atenção às representações por meio de vetores discretos em espaços euclidianos de dimensão finita, embora a representação em espaços de Hilbert de dimensão infinita também seja possível.

2.2 RECONSTRUÇÃO POR PROBLEMAS INVERSOS

Com o desenvolvimento crescente da capacidade de tratamento numérico de dados, proporcionado pelas unidades centrais de processamento (CPU¹) e unidades gráficas de processamento (GPU²), o tratamento dos dados discretos provenientes de sistemas de aquisição e a subsequente reconstrução das imagens têm se apoiado fortemente nos princípios da álgebra matricial.

A técnica de reconstrução por problemas inversos (*Inverse Problem-Based reconstruction*, IPB) baseia-se na resolução de um sistema de equações no qual estão presentes os dados coletados e o modelo de aquisição do sistema de imageamento. Mais detalhes sobre a construção do modelo e sobre outras aplicações de imageamento pela abordagem IPB são apresentados na Seção 2.4.2.

O conjunto de equações é usualmente colocado sob a forma de uma equação matricial do tipo

$$\mathbf{g} = \mathbf{H}\mathbf{f}, \quad (1)$$

em que o vetor coluna de incógnitas $\mathbf{f} \in \mathbb{R}^m$ representa a imagem desejada, a matriz $\mathbf{H} \in \mathbb{R}^{n \times m}$ representa o modelo discreto de aquisição, e os dados lidos por sensores estão armazenados no vetor coluna $\mathbf{g} \in \mathbb{R}^n$.

Apesar de os dados coletados em sistemas físicos serem de natureza real, podemos em certos casos estender os domínios destas variáveis para o conjunto de números complexos. Cabe notar que a imagem, comumente representada por uma matriz bidimensional de elementos (*pixels*), é aqui expressa por um vetor coluna \mathbf{f} , obtido pela concatenação de todas as colunas da imagem procurada. Após a resolução do problema, manipulamos o vetor \mathbf{f} encontrado de modo a recompor a imagem bidimensional para visualização. O termo *problema inverso* tem origem no problema matemático envolvido na busca do(s) vetor(es) \mathbf{f} que satisfaz(em) a Equação (1), o que de algum modo envolve a inversão da matriz \mathbf{H} .

Dado que o objeto de interesse é contínuo, mas tanto os dados coletados quanto a própria imagem são representados por vetores discretos, a aquisição representada por \mathbf{H} possui intrinsecamente uma componente de erro associada ao processo de amostragem. Essa perda de informação, somada à presença de ruído no mecanismo de aquisição, pode ser representada por

¹CPU, *Central Processing Unit*.

²GPU, *Graphic Processing Unit*.

um termo adicional $\eta \in \mathbb{R}^n$ na Equação (1):

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \eta . \quad (2)$$

Portanto, o vetor \mathbf{f} é apenas uma representação aproximada do objeto imageado, porque:

- O vetor η é de natureza aleatória, de modo que os valores exatos de seus elementos variam a cada nova aquisição e não podem ser previstos de modo determinístico;
- O ruído, que faz parte da composição de η , está presente em todos os sistemas de aquisição reais e não pode ser evitado; e
- Os dados coletados \mathbf{g} não representam o objeto imageado em si, mas a observação de alguma propriedade deste objeto, de acordo com a concepção do sistema de aquisição e dos sinais associados.

Formalmente, a incompletude das informações presentes nos vetores \mathbf{f} e \mathbf{g} está relacionada aos quatro subespaços da matriz de aquisição \mathbf{H} (BARRETT; MYERS, 2004). Supondo que \mathbf{H} seja uma matriz de posto Q , e que sua decomposição em valores singulares (*Singular Value Decomposition*, SVD) seja expressa por $\mathbf{H} = \mathbf{V}\mathbf{S}\mathbf{U}^\dagger$, sendo \mathbf{U}^\dagger a matriz adjunta³ de \mathbf{U} , o seu domínio \mathbb{U} de dimensão m é dividido em dois subespaços:

Espaço Mensurável: de dimensão $Q \leq m$, recebe este nome porque apenas os Q primeiros valores singulares da matriz \mathbf{S} são não nulos ($\sigma_1 \dots \sigma_Q$). Assim, no processo de aquisição são capturados elementos até a dimensão Q , de modo que nesse espaço estão as informações sobre o objeto imageado visíveis ao modelo de aquisição \mathbf{H} . O espaço mensurável corresponde ao *range* (ou imagem) da matriz adjunta \mathbf{H}^\dagger , $R(\mathbf{H}^\dagger)$.

Espaço Nulo: de dimensão $m - Q$, $N(\mathbf{H})$ recebe este nome porque os valores singulares $\sigma_{Q+1} \dots \sigma_m$ de \mathbf{S} são todos nulos. No processo de aquisição, as informações nesse subespaço sobre o objeto imageado são anuladas porque não são visíveis ao modelo \mathbf{H} .

O domínio da matriz adjunta \mathbf{H}^\dagger , espaço \mathbb{V} de dimensão n , é também dividido em dois subespaços, relacionados ao vetor de dados capturados \mathbf{g} . São eles:

³A matriz adjunta também é chamada de *matriz transposta conjugada*. Para uma matriz $\mathbf{A} = [a_{ij}]$, a sua adjunta \mathbf{A}^\dagger é definida como $[\mathbf{A}^\dagger]_{ij} = \bar{a}_{ji}$, sendo \bar{a} o conjugado complexo de a (MEYER, 2000).

Espaço Consistente: de dimensão $Q \leq n$, recebe este nome porque nesse subespaço todos os elementos correspondem a objetos imageados \mathbf{f} em \mathbb{U} , isto é, são dados consistentes com os respectivos objetos que representam. O espaço consistente corresponde ao *range* de \mathbf{H} , $R(\mathbf{H})$.

Espaço Inconsistente: de dimensão $n - Q$, recebe este nome porque nesse subespaço todos os elementos representam informações que não vieram dos objetos imageados \mathbf{f} em \mathbb{U} ; são provenientes de ruído, erros de discretização, erros de modelamento ou outras fontes externas ao processo de aquisição modelado por \mathbf{H} . O espaço inconsistente corresponde ao espaço nulo de \mathbf{H}^\dagger , $N(\mathbf{H}^\dagger)$.

A resolução do problema inverso, isto é, a estimação de vetores \mathbf{f} que satisfaçam a Equação (1), pode ser realizada por métodos diretos ou iterativos.

2.2.1 MÉTODOS DIRETOS

Os métodos diretos (ou algoritmos determinísticos) são aqueles em que o vetor \mathbf{f} é estimado a partir dos dados \mathbf{g} e da matriz de aquisição \mathbf{H} por uma expressão fechada.

A primeira alternativa é observar que, se a matriz \mathbf{H} possuir inversa \mathbf{H}^{-1} (ou pseudo-inversa \mathbf{H}^+ no caso geral em que $m \neq n$), a Equação (1) dá origem a um vetor estimado $\hat{\mathbf{f}}_{inv}$, dado por

$$\hat{\mathbf{f}}_{inv} = \mathbf{H}^{-1} \mathbf{g}. \quad (3)$$

Entretanto, o termo η apresentado na Equação (2) modifica essa estimativa:

$$\hat{\mathbf{f}}_{inv} = \mathbf{H}^{-1}(\mathbf{H}\mathbf{f} + \eta) = \mathbf{f} + \mathbf{H}^{-1}\eta, \quad (4)$$

o que indica que, se \mathbf{H} for mal-condicionada⁴, o processo de inversão pode amplificar consideravelmente o erro presente no processo de aquisição. Por esse motivo, salvos os casos em que \mathbf{H} é bem condicionada, a solução pela inversão simples tem em geral pouca utilidade prática.

Outra alternativa é estimar \mathbf{f} pela matriz transposta \mathbf{H}^T (ou pela adjunta \mathbf{H}^\dagger no caso de modelos complexos):

$$\hat{\mathbf{f}}_{transp} = \mathbf{H}^T \mathbf{g}. \quad (5)$$

⁴Uma matriz não singular \mathbf{A} é dita *mal-condicionada* (*ill-conditioned*) se uma mudança relativamente pequena em \mathbf{A} puder causar uma mudança relativa grande na sua inversa \mathbf{A}^{-1} . O grau de mal-condicionamento é medido pelo *número de condicionamento* (*condition number*) $\kappa = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$, em que $\|\cdot\|$ é uma norma matricial. Para o caso da norma ℓ_2 , temos $\kappa = \sigma_1/\sigma_n = \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 \geq 1$, sendo σ_1 e σ_n o maior e o menor valor singular de \mathbf{A} , respectivamente (MEYER, 2000).

Formalmente, essa não é uma solução para a Equação (1), mas é a projeção do espaço consistente do mapeamento \mathbf{H} em direção ao seu domínio \mathbb{U} , onde encontra-se o vetor \mathbf{f} procurado. A limitação desta alternativa, porém, fica clara quando observamos que a operação de transposição não inverte os valores singulares da matriz \mathbf{H} , de modo que a solução obtida não tem as componentes em escala apropriada.

A estimação direta de \mathbf{f} tem ainda uma solução baseada no método dos Quadrados Mínimos (*Least Squares*). Há um grande interesse neste método porque o teorema de Gauss-Markov nos garante que a solução por Quadrados Mínimos é o estimador linear não tendencioso (*unbiased*) de variância mínima para o vetor \mathbf{f} da Equação (2). Essa garantia é válida para a suposição de que as componentes de η não são correlacionadas entre si, possuem a mesma variância (não necessariamente conhecida) e têm média igual a zero (MEYER, 2000).

A aplicação desse método à reconstrução da imagem é dada pela minimização da norma quadrática do erro de estimação

$$\Psi(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2, \quad (6)$$

que, sendo uma função convexa, tem seu mínimo em relação a \mathbf{f} facilmente encontrado quando igualamos o seu gradiente $\nabla\Psi(\mathbf{f}) = -2\mathbf{H}^T(\mathbf{g} - \mathbf{H}\mathbf{f})$ a zero, obtendo

$$\hat{\mathbf{f}}_{LS} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{g}. \quad (7)$$

Entretanto, a solução por Quadrados Mínimos também é susceptível à instabilidade devida ao mal condicionamento da matriz \mathbf{H} , e também pode não ser a solução única para o sistema de equações. O tratamento desta questão é abordado com mais detalhes na Seção 2.3. Além disso, dependendo da natureza da imagem \mathbf{f} procurada, a minimização da norma quadrática do erro pode não ser a mais interessante do ponto de vista prático, em função do uso que se fará da imagem reconstruída.

2.2.2 FUNÇÃO CUSTO, NORMAS E ESPARSIDADE

A função $\Psi(\mathbf{f})$ na Equação (6) é comumente designada de *função custo*. Trata-se de uma função escalar multivariável, que define uma hipersuperfície de desempenho e que representa o critério (e.g. erro de estimação) a ser minimizado, de modo que a obtenção da imagem estimada $\hat{\mathbf{f}}$ resume-se fundamentalmente à solução do problema

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi(\mathbf{f}), \quad (8)$$

em que $\arg \min_{\mathbf{f}}$ indica a minimização em relação a \mathbf{f} .

As funções custo podem ter um ou mais termos, dependendo das formas de restrição e regularização utilizadas no tratamento do problema, conforme abordado na Seção 2.3. Comum a todas elas, no entanto, é o fato de que sua formulação em geral caracteriza uma medida de distância, de modo que sua amplitude indica o quanto a solução obtida está próxima daquela aceitável para uma dada aplicação.

Relacionado à noção de distância, está o conceito de *norma*. De acordo com Barrett e Myers (2004), a norma $\|\mathbf{r}\|$ de um vetor \mathbf{r} é um escalar que deve satisfazer três condições:

- a) $\|\mathbf{r}\| \geq 0$, com igualdade se, e somente se, $\mathbf{r} = \mathbf{0}$;
- b) $\|\lambda\mathbf{r}\| = |\lambda| \cdot \|\mathbf{r}\|$, sendo λ um número arbitrário, real ou complexo; e
- c) $\|\mathbf{r}_1 + \mathbf{r}_2\| \leq \|\mathbf{r}_1\| + \|\mathbf{r}_2\|$, condição conhecida como *desigualdade triangular*.

Dentre as normas existentes, as mais comumente utilizadas nas funções custo em reconstrução de imagens são as normas ℓ_2 (Euclidiana), ℓ_1 , ℓ_p e ℓ_∞ . Considerando um vetor \mathbf{r} pertencente a um espaço vetorial de dimensão n , real ou complexo, essas normas são assim definidas:

$$\begin{aligned}
 \|\mathbf{r}\|_2 &= \left[\sum_{k=1}^n |r_k|^2 \right]^{1/2} && \text{norma } \ell_2 \text{ (Euclidiana)} \\
 \|\mathbf{r}\|_1 &= \sum_{k=1}^n |r_k| && \text{norma } \ell_1 \\
 \|\mathbf{r}\|_p &= \left[\sum_{k=1}^n |r_k|^p \right]^{1/p} && \text{norma } \ell_p \\
 \|\mathbf{r}\|_\infty &= \lim_{p \rightarrow \infty} \|\mathbf{r}\|_p = \max_{k=1 \dots n} |r_k| && \text{norma } \ell_\infty.
 \end{aligned} \tag{9}$$

É importante ressaltar que ℓ_p é uma norma apenas para $p \geq 1$, com valores inteiros ou não. Para $p < 1$, a expressão acima não satisfaz a condição de desigualdade triangular e, por isso, não é formalmente uma norma (ELAD, 2010).

As normas são medidas de distância distintas e a conveniência do uso de cada tipo está relacionada à geometria considerada mais adequada para a solução do problema de minimização. Quando minimizamos uma distância calculada pela norma Euclidiana, por exemplo, encontramos um vetor cuja raiz quadrada da soma dos quadrados de suas componentes é mínima; essa abordagem é interessante se todos os elementos do vetor minimizado forem igualmente preponderantes. Entretanto, certas imagens possuem a característica de possuírem muitos elementos nulos ou de valor relativo muito inferior aos demais (conhecidas como

imagens esparsas), de modo que a minimização da norma ℓ_2 não resultará em uma imagem com essa característica, pois tende a não privilegiar nenhum dos seus elementos em particular. Além disso, a norma ℓ_2 não é robusta a erros impulsivos que podem estar presentes dos dados \mathbf{g} .

No casos em que temos um conhecimento *a priori* sobre a imagem procurada, e supomos que ela deva ser esparsa, é conveniente adicionar um termo à função custo com a norma ℓ_p com $1 \leq p < 2$, uma vez que a solução encontrada tenderá a ter maior esparsidade. Um inconveniente de usarmos a norma ℓ_p , no entanto, é o fato de ela manter a função custo convexa (como a norma ℓ_2), mas não *estritamente* convexa, de modo que não temos mais a garantia de que a solução mínima seja única (ELAD, 2010).

A utilização de normas não quadráticas na função custo equivale à representação de um sistema de equações *não linear*, para o qual não há solução na forma fechada (por um método direto), mas apenas soluções obtidas de modo iterativo.

2.2.3 MÉTODOS ITERATIVOS

Os métodos iterativos de descida são utilizados para buscar um ponto de mínimo da hipersuperfície de desempenho definida pela função custo $\Psi(\mathbf{f})$, partindo de uma solução inicial \mathbf{f}_0 e atualizando, a cada iteração, um novo vetor $\hat{\mathbf{f}}$ que resulte num custo cada vez menor⁵, até que um determinado critério de parada do algoritmo seja alcançado. Há uma extensa literatura a respeito dos algoritmos iterativos de otimização. Nos trabalhos de Kelley (1999) e de Byrne (2014), por exemplo, diversos métodos são apresentados, juntamente com uma extensa bibliografia de referência.

Segundo apontam Carcreff et al. (2014), o problema de reconstrução em ultrassonografia, foco deste trabalho, tem como característica imagens esparsas, para as quais a otimização com o uso da norma ℓ_1 tem apresentado resultados promissores. Isso foi mostrado no trabalho de O'Brien et al. (1994), e em (VALENTE et al., 2017).

Por esse motivo, os algoritmos apresentados no Capítulo 3 foram selecionados entre aqueles que reconhecidamente apresentam bons resultados para reconstruções esparsas, em muitas aplicações de processamento de sinais e também na reconstrução de imagens. Convidamos o(a) leitor(a) a consultar o trabalho de Elad (2010) e as referências citadas em cada método para maiores detalhes.

⁵Embora o objetivo final seja descer até o ponto de mínimo da hipersuperfície de custo, alguns métodos permitem que sejam atingidos pontos de custo superior em iterações intermediárias para possibilitar a descida mais rápida nas iterações seguintes.

2.3 REGULARIZAÇÃO DE PROBLEMAS MAL-POSTOS

Em função da quantidade de equações, do número de incógnitas e das características da matriz de coeficientes, um sistema de equações pode não possuir uma solução única e estável; nessas situações, o sistema é dito *mal-posto*. Esse conceito foi inicialmente apresentado pelo matemático francês Jacques Hadamard (1865-1963) no início do século XX. Ele classificou como mal-posto um problema cuja solução não é única ou que não seja uma função contínua dos dados, isto é, que pode sofrer grandes variações face a pequenas perturbações nos dados (HANSEN, 1998).

Diz-se que um problema é *bem-posto* quando ele não se enquadra em nenhuma das três condições propostas por Hadamard, detalhadas a seguir no contexto da Equação (1).

1ª Condição de Hadamard: o problema não tem solução.

Nesse caso, não existe um vetor \mathbf{f} que satisfaça a igualdade $\mathbf{g} = \mathbf{H}\mathbf{f}$, isto é, nenhum vetor do espaço \mathbb{U} pode produzir um vetor no espaço \mathbb{V} por meio do mapeamento \mathbf{H} . Isso significa que o vetor \mathbf{g} não está no espaço consistente (*range*) de \mathbf{H} . Essa situação pode ocorrer quando há mais equações do que incógnitas ($n > m$) ou quando há componentes nos dados que são provenientes de fontes não modeladas (e.g. ruído).

2ª Condição de Hadamard: o problema tem infinitas soluções.

Nesse caso, há infinitos vetores \mathbf{f} que satisfazem a igualdade $\mathbf{g} = \mathbf{H}\mathbf{f}$. Isso significa que a matriz \mathbf{H} não representa um mapeamento bijetor (um para um), isto é, vários vetores \mathbf{f} no espaço \mathbb{U} podem produzir o mesmo vetor \mathbf{g} no espaço \mathbb{V} . Essa condição pode ocorrer quando há mais incógnitas do que equações ($n < m$), ou quando o posto Q de \mathbf{H} é menor que m , ou seja, quando \mathbf{H} possui valores singulares nulos.

3ª Condição de Hadamard: a solução existe, mas é instável.

Essa condição pode ocorrer quando a matriz \mathbf{H} é mal-condicionada, isto é, quando seus menores valores singulares não-nulos são muito menores que seu maior valor singular. A instabilidade da solução do sistema de equações é causada pela inversão de \mathbf{H} , uma vez que a consequente inversão dos valores singulares pequenos de \mathbf{H} pode resultar em valores singulares elevados na inversa de \mathbf{H} , o que provoca uma grande amplificação do ruído presente nos dados \mathbf{g} , conforme indicado na Equação (4).

Cada condição de Hadamard pode ser contornada pela adoção de uma estratégia

adequada, com vistas a tornar o problema bem-posto. No caso da primeira condição, quando não há solução exata, buscamos uma solução aproximada através da *minimização* de uma função custo adequada ao problema. Na segunda condição, em que há infinitas soluções possíveis, estabelecemos um critério de *restrição* ao qual a solução encontrada deve sujeitar-se, o que limita o número de soluções possíveis. Na terceira condição, precisamos tratar de modo especial os valores singulares da matriz \mathbf{H} próximos a zero, o que é feito por meio de *regularização*.

Tanto a restrição quanto a regularização podem ser implementadas pela adição de novos termos à função custo. No caso da restrição, o termo adicional penaliza o espaço nulo da matriz \mathbf{H} , mas não trata os seus valores singulares muito próximos a zero. Por isso, essa técnica resolve as duas primeiras condições de Hadamard (o problema passa a ter solução e ela é única), mas não resolve a terceira condição (a solução obtida ainda pode ser instável). No caso da regularização, no entanto, o termo adicional penaliza os termos fortemente atenuados no espaço mensurável de \mathbf{H} em função dos seus pequenos valores singulares. A regularização resolve as duas últimas condições de Hadamard, mas necessita de um critério para a escolha do *parâmetro de regularização*, que determina a intensidade da regularização aplicada.

Em geral, a reconstrução de imagens em ultrassonografia por problemas inversos apresenta a primeira e a terceira condições de Hadamard. A primeira, porque os dados provenientes dos transdutores de ultrassom podem gerar vetores \mathbf{g} com muito mais elementos do que o vetor \mathbf{f} (imagem); a terceira, porque a matriz \mathbf{H} obtida pelo modelamento do sistema de aquisição é, em geral, mal-condicionada. Por esses motivos, a regularização tende a ser a abordagem mais apropriada para a reconstrução dessas imagens.

2.3.1 REGULARIZAÇÃO DE TIKHONOV

Este método de regularização foi desenvolvido independentemente por Andrey Tikhonov e David L. Phillips e é também conhecido por outros nomes na literatura, tais como *ridge regression* (na Estatística) ou método de Phillips-Twomey (HANSEN, 1998).

A ideia central da regularização de Tikhonov é a incorporação de informação *a priori* sobre a imagem a ser reconstruída, por meio da adição de uma função de penalidade (*prior*) à função custo quadrática da Equação (6). A penalidade pode assumir duas formas distintas:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi_{Tik}(\mathbf{f}) = \arg \min_{\mathbf{f}} [\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \xi^2 \|\mathbf{L}\mathbf{f}\|_2^2] , \quad (10)$$

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi_{Tik}(\mathbf{f}) = \arg \min_{\mathbf{f}} [\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \xi^2 R(\mathbf{f})] , \quad (11)$$

sendo $\xi \in \mathbb{R}$ o parâmetro de regularização, $\mathbf{L} \in \mathbb{R}^{p \times m}$ uma matriz que realiza uma transformação linear sobre \mathbf{f} , com p sendo uma dimensão arbitrária, e $R : \mathbb{R}^m \rightarrow \mathbb{R}$.

Em ambos os casos, a função custo possui um termo adicional que depende do parâmetro de regularização ξ . Na primeira forma, o termo de penalidade é uma função quadrática do vetor $\mathbf{L}\mathbf{f}$. Na segunda, a função $R(\mathbf{f})$ torna a regularização mais genérica. A primeira forma é conhecida como *Regularização de Tikhonov*, enquanto a segunda é chamada de *Regularização de Tikhonov Generalizada* (VOGEL, 2002).

O parâmetro de regularização ξ controla o equilíbrio entre os dois termos da função custo, de modo que valores tendendo a zero diminuem o efeito de regularização, tornando a imagem reconstruída mais “ruidosa” (com menor filtragem); valores maiores para ξ dão maior peso ao termo de regularização, aumentando a filtragem, e tornando a imagem final mais “suave” (HANSEN, 1998). Duas formas de determinar um valor adequado para o parâmetro de regularização são discutidas na Seção 2.3.3.

A Equação (10) tem solução fechada, dada por

$$\hat{\mathbf{f}}_{Tik} = (\mathbf{H}^T \mathbf{H} + \xi^2 \mathbf{L}^T \mathbf{L})^{-1} \mathbf{H}^T \mathbf{g}, \quad (12)$$

que é a versão regularizada da solução de Quadrados Mínimos da Equação (7). Essa forma fechada nos garante uma solução única, desde que os espaços nulos das matrizes \mathbf{H} e \mathbf{L} sejam distintos, isto é, $N(\mathbf{H}) \cap N(\mathbf{L}) = \emptyset$. Entretanto, há outras formas de se chegar à mesma solução que não envolvem a inversão da matriz $(\mathbf{H}^T \mathbf{H} + \xi^2 \mathbf{L}^T \mathbf{L})$, tais como fatorações via SVD ou GSVD (*Generalized Singular Value Decomposition*), além dos métodos iterativos (HANSEN, 1998).

Se tomarmos a definição da SVD da matriz de aquisição, $\mathbf{H} = \mathbf{V}\mathbf{S}\mathbf{U}^\dagger$ em conjunto com a Equação (12), a solução fechada para o caso em que \mathbf{L} é uma matriz identidade $\mathbf{L} = \mathbf{I}_m \in \mathbb{R}^{m \times m}$, tem a forma

$$\hat{\mathbf{f}}_{Tik} = \sum_{i=1}^Q \left(\frac{\sigma_i^2}{\sigma_i^2 + \xi^2} \right) \frac{\mathbf{v}_i^T \mathbf{g}}{\sigma_i} \mathbf{u}_i, \quad (13)$$

em que \mathbf{v}_i são as colunas ortonormais da matriz \mathbf{V} (vetores singulares à esquerda), \mathbf{u}_i são as colunas ortonormais de \mathbf{U} (vetores singulares à direita), \mathbf{g} é o vetor de dados e σ_i são os valores singulares da matriz \mathbf{H} de posto Q , ordenados em forma decrescente $\sigma_1 \geq \dots \geq \sigma_Q > \sigma_{Q+1} = \dots = \sigma_p = 0$, para $p = \min(m, n)$ (BOVIK, 2000).

O termo entre parênteses na Equação (13) é conhecido como *peso* (*weight*) ou *fator de filtro* (*filter factor*) e nos mostra o princípio fundamental da regularização de Tikhonov: sem esse peso, a i -ésima componente da solução, por ser dividida por σ_i , resulta num valor muito

elevado para pequenos valores singulares, provocando o efeito indesejado de amplificação do ruído⁶. Com a inclusão do peso, a inversão dos valores singulares elevados é preservada (pois o peso é próximo da unidade), enquanto a inversão dos valores singulares muito próximos a zero é atenuada com o quadrado de ξ .

Neste caso em que $\mathbf{L} = \mathbf{I}_m$, o termo de regularização da Equação (10) representa uma medida da energia de \mathbf{f} ; a sua inclusão na função custo evita que as componentes da imagem assumam valores muito elevados, como acontece na solução não regularizada da Equação (7). Nesse caso, o papel de ξ é proporcionar um balanço entre a fidelidade aos dados (medida pelo erro de estimação, o primeiro termo da função custo) e a energia da solução.

Com a escolha $\mathbf{L} = \mathbf{I}_m$ a regularização de Tikhonov não nos permite tentar reconstruir componentes da imagem \mathbf{f} que não estejam nos dados \mathbf{g} . Entretanto, para $\mathbf{L} \neq \mathbf{I}_m$ há várias opções de escolha para a matriz \mathbf{L} que permitem a inclusão de informação *a priori* sobre a imagem, de modo que a reconstrução tenderá a ter um resultado mais próximo do desejado. Por exemplo, a matriz \mathbf{L} pode representar a operação de diferenças finitas sobre o vetor \mathbf{f} , de modo que o termo $\|\mathbf{L}\mathbf{f}\|_2^2$ represente uma medida da variação ou da “aspereza” (*roughness*) da imagem. Nesses casos, a inclusão da matriz \mathbf{L} força uma solução que limita a energia das componentes de alta frequência, e representa a suposição de que a imagem deve ser suave. É possível, ainda, que a função custo tenha vários termos na forma $\|\mathbf{L}_i\mathbf{f}\|_2^2$ de modo a criar penalidades derivativas ponderadas de ordens múltiplas (BOVIK, 2000).

Também é possível calcular a solução da Equação (10) de forma fechada para $\mathbf{L} \neq \mathbf{I}_m$ com o uso da GSVD do par (\mathbf{H}, \mathbf{L}) . Nesse caso, o fator de filtro é modificado, utilizando os valores singulares *generalizados* γ_i no lugar dos valores singulares σ_i (HANSEN, 1998). Notamos ainda que valores de ξ superiores ao maior valor singular de \mathbf{H} (para $\mathbf{L} = \mathbf{I}_m$) ou superiores ao maior valor singular generalizado do par (\mathbf{H}, \mathbf{L}) (para $\mathbf{L} \neq \mathbf{I}_m$) tendem a provocar regularização excessiva, pois os fatores de filtro tendem rapidamente a zero.

No caso da regularização de Tikhonov Generalizada, o termo $R(\mathbf{f})$ da Equação (11) pode assumir diversas formas, sendo a mais simples a norma ℓ_p com $p \neq 2$:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi_{Tik}^{gen}(\mathbf{f}) = \arg \min_{\mathbf{f}} [\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \xi^2 \|\mathbf{L}\mathbf{f}\|_p^p] . \quad (14)$$

Determinar a solução que minimiza a função custo $\Psi_{Tik}^{gen}(\mathbf{f})$ com $p \neq 2$, equivale a resolver um sistema de equações não linear, de modo que a solução deve ser encontrada por um

⁶Uma descrição mais detalhada sobre a causa da amplificação do ruído pode ser encontrada em Hansen (1998). Ela consiste na decomposição explícita do vetor $\mathbf{g} = \mathbf{g}_{SR} + \mathbf{e}$ na Equação (13), sendo \mathbf{g}_{SR} a componente dos dados sem ruído e \mathbf{e} o vetor de erros e de ruído aditivo.

método iterativo, já que não há solução fechada. Por outro lado, a solução tende a preservar mais adequadamente certas características da imagem do que a regularização quadrática (e.g. esparsidade), e tende a eliminar melhor o ruído. Um caso bastante estudado na literatura corresponde ao uso de $p = 1$, com a matriz \mathbf{L} sendo uma aproximação discreta do operador de gradiente; esse caso de regularização é conhecido como *Total Variation* (TV) (MONTERO, 2005).

2.3.2 ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA E *MAXIMUM A POSTERIORI*

Outra forma de interpretar os termos da função custo vem da observação das propriedades estatísticas (ou estocásticas) do ruído e da própria imagem procurada. Enquanto a abordagem determinística nos dá uma boa intuição sobre a geometria do problema de otimização por meio das características das normas ℓ_2 e ℓ_1 , por exemplo, também podemos construir a função custo com vistas a maximizar certas probabilidades de interesse.

Seja $pr(\mathbf{g}, \mathbf{f})$ a probabilidade de o vetor \mathbf{f} ter dado origem aos dados \mathbf{g} . Uma vez que temos os dados coletados \mathbf{g} e buscamos a imagem \mathbf{f} , é razoável que o vetor estimado seja tal que maximize essa probabilidade:

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} pr(\mathbf{g}, \mathbf{f}) . \quad (15)$$

Essa formulação é chamada de *Máxima Verossimilhança* (*Maximum Likelihood*, ML) e a probabilidade pode ser definida em função do erro de estimação $\mathbf{g} - \mathbf{Hf}$. Se assumirmos, por exemplo, que o erro tem comportamento Gaussiano de média nula, variância constante e com componentes independentes e identicamente distribuídas (i.i.d.⁷), a função densidade de probabilidade (*Probability Density Function*, PDF) tem a forma

$$pr(\mathbf{g}, \mathbf{f}) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{\|\mathbf{g} - \mathbf{Hf}\|_2^2}{2\sigma^2}} \quad (\text{Erro Gaussiano}) , \quad (16)$$

sendo n a dimensão do vetor $\mathbf{g} - \mathbf{Hf}$ e σ o desvio padrão da distribuição. Se assumirmos, por outro lado, que o erro tem comportamento Laplaciano de média zero e com componentes i.i.d., a PDF será dada por

$$pr(\mathbf{g}, \mathbf{f}) = \frac{1}{(2\sigma)^n} e^{-\frac{\|\mathbf{g} - \mathbf{Hf}\|_1}{\sigma}} \quad (\text{Erro Laplaciano}) . \quad (17)$$

A função custo a ser minimizada é então obtida como uma medida de distância

⁷Da designação em Inglês *independent and identically distributed*.

associada às PDF. Dada a natureza exponencial dessas funções, uma medida de distância conveniente é obtida pelo negativo do logaritmo natural, $-\ln[(pr(\mathbf{g}, \mathbf{f}))]$. A Figura 1 ilustra as curvas de densidade de probabilidade e as medidas de distância associadas a ambas.

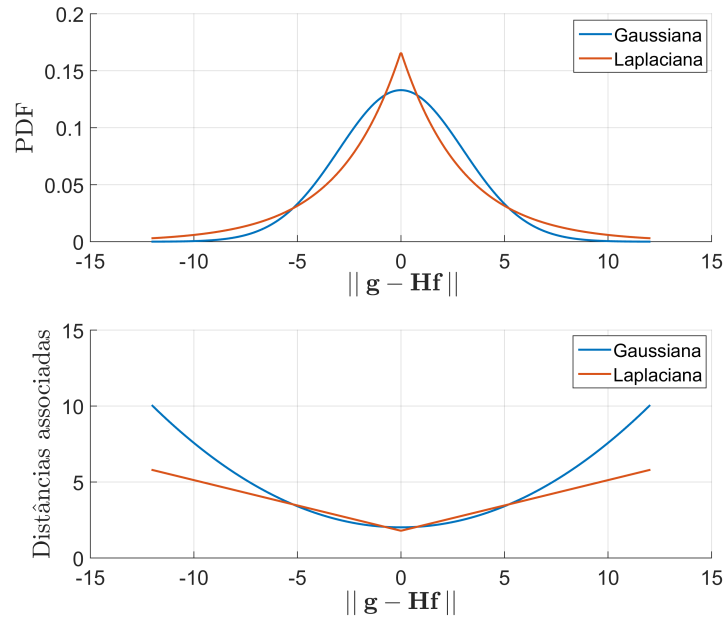


Figura 1: Curvas das PDF Gaussiana e Laplaciana e das distâncias associadas.

Fonte: O autor (2017)

As funções custo assumem então as formas

$$\Psi_G(\mathbf{f}) = \frac{\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2}{c_1} - \ln(c_2), \quad (18)$$

$$\Psi_L(\mathbf{f}) = \frac{\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_1}{c_3} - \ln(c_4), \quad (19)$$

sendo c_1 , c_2 , c_3 e c_4 constantes. Essas equações são essencialmente funções custo não regularizadas e as suas minimizações representam as soluções ótimas para sistemas com distribuições de erro Gaussianas e Laplacianas, respectivamente. Porém, pela discussão anterior, não temos garantia de que essas soluções sejam únicas ou estáveis, de modo que é necessário acrescentar termos de regularização às funções custo.

Essa questão é resolvida pela abordagem conhecida como *Maximum a Posteriori* (MAP). Nessa formulação, os vetores \mathbf{g} e \mathbf{f} são considerados variáveis aleatórias e utilizamos a Regra de Bayes para acrescentar o termo de regularização. Partimos da premissa de que a solução estimada maximiza a probabilidade condicional $pr(\mathbf{f} | \mathbf{g})$ de \mathbf{f} dado \mathbf{g} ,

$$\hat{\mathbf{f}}_{MAP} = \arg \max_{\mathbf{f}} pr(\mathbf{f} | \mathbf{g}), \quad (20)$$

que, pela Regra de Bayes, torna-se

$$\begin{aligned}\hat{\mathbf{f}}_{MAP} &= \arg \max_{\mathbf{f}} \frac{pr(\mathbf{g} | \mathbf{f}) pr(\mathbf{f})}{pr(\mathbf{g})}, \\ \hat{\mathbf{f}}_{MAP} &= \arg \max_{\mathbf{f}} pr(\mathbf{g} | \mathbf{f}) pr(\mathbf{f}).\end{aligned}\quad (21)$$

Tomando a mesma medida de distância do ML, obtemos

$$\hat{\mathbf{f}}_{MAP} = \arg \min_{\mathbf{f}} \{ -\ln [pr(\mathbf{g} | \mathbf{f})] - \ln [pr(\mathbf{f})] \}, \quad (22)$$

em que os termos $pr(\mathbf{g} | \mathbf{f})$ e $pr(\mathbf{f})$ representam, respectivamente, os modelos estatísticos dos dados e da imagem. Se supusermos que ambos têm comportamento Gaussiano, a função custo a minimizar será

$$\Psi_{MAP}(\mathbf{f}) = \frac{\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2}{2\sigma^2} - \ln(c_5) + \frac{\|\mathbf{L}\mathbf{f}\|_2^2}{\beta^2} - \ln(c_6), \quad (23)$$

sendo c_5 e c_6 duas constantes, σ o desvio padrão do erro de estimação $\mathbf{g} - \mathbf{H}\mathbf{f}$ e β o desvio padrão do modelo da imagem. O vetor estimado é dado por

$$\hat{\mathbf{f}}_{MAP} = \arg \min_{\mathbf{f}} [\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \xi^2 \|\mathbf{L}\mathbf{f}\|_2^2], \quad (24)$$

que é igual à Equação (10), salvo que o parâmetro de regularização tem um valor particular, dado pela relação entre a variância do erro e a variância da imagem

$$\xi^2 = \frac{\sigma^2}{\beta^2}. \quad (25)$$

Outras pressuposições quanto aos modelos estatísticos dos termos de erro e de regularização da Equação (22) levam a diferentes termos na Equação (24), sendo de particular interesse para a reconstrução de imagens o modelo Laplaciano, que leva a funções custo que utilizam a norma ℓ_1 .

Quando os parâmetros σ e β dessas distribuições são desconhecidos, eles podem ser incluídos como incógnitas no processo de otimização, de modo que a Equação (20) torna-se

$$\hat{\mathbf{f}}_{JMAP} = \arg \max_{\mathbf{f}, \sigma, \beta} pr(\mathbf{f}, \sigma, \beta | \mathbf{g}). \quad (26)$$

Nessa formulação, conhecida como *Joint Maximum a Posteriori* (JMAP), σ e β são chamados de *hiperparâmetros*, e pela Regra de Bayes obtemos

$$\hat{\mathbf{f}}_{JMAP} = \arg \max_{\mathbf{f}, \sigma, \beta} pr(\mathbf{g} | \mathbf{f}, \sigma) pr(\sigma) pr(\mathbf{f} | \beta) pr(\beta). \quad (27)$$

O termo $pr(\mathbf{f} | \beta)$ é designado *prior* porque representa uma suposição *a priori* sobre o comportamento da imagem (e.g. “suavidade”), uma vez que é a função densidade de probabilidade de \mathbf{f} para um dado desvio padrão β . Os termos $pr(\beta)$ e $pr(\sigma)$ são designados *hyperpriors* porque representam a suposição sobre o comportamento estatístico dos parâmetros do modelo. Note que β e σ não precisam ser conhecidos *a priori*, mas precisamos supor que conhecemos seu comportamento do ponto de vista estatístico. A função custo explícita para o JMAP depende das escolhas dos modelos para os quatro termos à direita da Equação (27).

2.3.3 MÉTODOS DE ESCOLHA DO PARÂMETRO DE REGULARIZAÇÃO

Um dos problemas da regularização é o afastamento da solução obtida daquela que melhor caracteriza a fidelidade aos dados. Embora a noção de qualidade da imagem reconstruída seja relativa, e possa depender do parâmetro de regularização empregado, algum método quantitativo precisa ser utilizado para estabelecer um valor ou faixa de valores para esse parâmetro.

Para preservar a coerência de notação, a partir deste ponto passamos a utilizar o símbolo λ para identificar o parâmetro de regularização da função custo, e não mais ξ . Reservamos o símbolo ξ para um uso específico associado à regularização de Tikhonov, detalhado no Capítulo 5.

Método da Curva L

O método gráfico da Curva L foi inicialmente proposto por Lawson e Hanson (1974) e popularizado por Hansen (1992). Esse método, bastante conveniente para a determinação do parâmetro de regularização, consiste em plotar a curva da norma do termo de regularização em função da norma de erro de estimação, ou resíduo, em escala logarítmica. A Curva L mostra claramente o compromisso entre a minimização desses dois termos, que é o equilíbrio procurado por qualquer método de regularização (HANSEN, 1998).

No caso de usarmos a norma ℓ_2 em ambos os termos da função custo, como na regularização de Tikhonov, o eixo das abscissas é dado por $\log_{10} \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}_\lambda\|_2^2$ e o eixo das ordenadas por $\log_{10} \|\mathbf{L}\hat{\mathbf{f}}_\lambda\|_2^2$, sendo $\hat{\mathbf{f}}_\lambda$ a imagem estimada para um dado valor de λ , conforme ilustra a Figura 2.

Hansen (1998) demonstra que os extremos da Curva L correspondem aos valores limite do parâmetro de regularização λ , sendo o lado esquerdo correspondente a $\lambda \rightarrow 0$ (solução não regularizada) e o lado direito a $\lambda \rightarrow \infty$ (solução com regularização excessiva). Desejamos

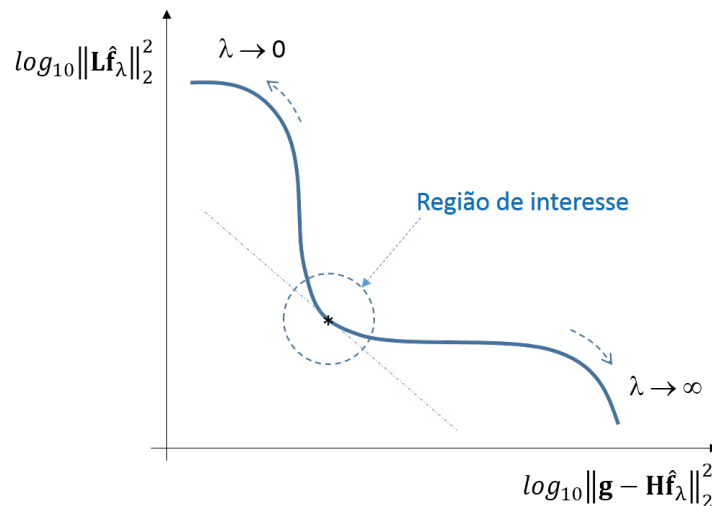


Figura 2: Forma geral da Curva L

Fonte: Adaptada de Hansen (1998)

identificar o ponto de maior curvatura na região central da curva, que corresponde a um parâmetro λ que representa um compromisso razoável entre os dois termos minimizados.

A escala *log-log* é utilizada com o propósito de evidenciar a região de máxima curvatura, ressaltando o formato de “L” que dá origem ao nome do método, e facilitando a identificação do parâmetro de regularização associado. Entretanto, a determinação do ponto exato de inflexão pode não ser muito clara em certos casos. Entre as alternativas apresentadas na literatura, Hansen (1998) propõe o critério da maior curvatura positiva, enquanto Reginska (1996) propõe o critério do ponto tangente com uma reta de inclinação negativa, ambos em escala *log-log*. Zibetti et al. (2008) propõem o mesmo critério de Reginska, porém com a Curva L traçada na escala *sqr-sqr* (raiz quadrada).

Segundo Montero (2005), a principal limitação do método da Curva L é o fato de que a forma de “L” não é garantida para todos os métodos de regularização.

Método da Validação Cruzada Generalizada (*Generalized Cross-Validation, GCV*)

O Método GCV foi inicialmente proposto por Golub et al. (1979) e baseia-se na minimização da função

$$\text{GCV}(\lambda) = \frac{\|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}_\lambda\|_2^2}{[\text{tr}(\mathbf{I} - \mathbf{A}_\lambda)]^2}, \quad (28)$$

em que $\hat{\mathbf{f}}_\lambda$ é o vetor \mathbf{f} estimado (por um método direto ou iterativo) para um dado valor do

parâmetro de regularização λ . O operador $tr(\cdot)$ indica o traço de uma matriz, e \mathbf{A}_λ é dada por

$$\mathbf{A}_\lambda = \mathbf{A}\mathbf{A}^\#, \quad (29)$$

sendo $\mathbf{A}^\#$ a matriz que relaciona o vetor estimado com os dados, $\hat{\mathbf{f}}_\lambda = \mathbf{A}^\# \mathbf{g}$.

Segundo Vogel (2002), a função $GCV(\lambda)$ é um estimador do erro preditivo e, sendo uma função de λ , a sua minimização aponta o valor do parâmetro de regularização que melhor discrimina a informação sobre a imagem a partir do ruído presente nos dados coletados no processo de aquisição. Entretanto, o método GCV não converge para o valor ótimo de λ quando o ruído está correlacionado com o sinal e, além disso, é um método computacionalmente mais custoso que o método da Curva L (LAVARELLO et al., 2006).

2.4 RECONSTRUÇÃO DE IMAGENS EM ULTRASSONOGRRAFIA

2.4.1 RECONSTRUÇÃO POR *BEAMFORMING*

A reconstrução de imagens bidimensionais em sistemas de ultrassonografia (conhecidas como imagens no “Modo B”⁸) utiliza tradicionalmente uma técnica conhecida como *beamforming*, que se caracteriza pela captura de sinais (ecos de ultrassom) por elementos transdutores e pela estimativa da resposta acústica de uma região de interesse (*Region Of Interest*, ROI).

Os elementos transdutores são excitados com sinais elétricos e dão origem a pulsos acústicos que “iluminam” a região de interesse; os mesmos elementos são usados para receber os sinais ultrassônicos de eco correspondentes. Após serem recebidos pelos elementos do transdutor, os sinais são tratados de forma sequencial com:

- introdução de atrasos de tempo variáveis para a formação do feixe de interesse;
- demodulação, para detecção da envoltória das ondas de ultrassom capturadas;
- compensação de ganho variável com o tempo (*Time-Gain Compensation*, TGC), que aplica ganhos maiores para sinais que chegam mais tarde ao transdutor e que, por isso, estão mais atenuados;
- compressão logarítmica, para melhor evidenciar os sinais com amplitudes distintas; e

⁸A designação “Modo B” vem da palavra *bright* no idioma Inglês, que indica que cada ponto da imagem tem um brilho (em escala de cinzas) proporcional à intensidade de reflexão da onda ultrassônica.

- conversão analógico-digital, para a geração de sinais digitais apropriados ao armazenamento e à exibição em monitores de vídeo.

Muito do desenvolvimento da qualidade de imagem em ultrassonografia resultou da introdução da técnica de formação digital do feixe. Com o uso de elementos transdutores agrupados em conjuntos lineares (*Linear Phased Arrays*), a distância focal efetiva e a abertura de feixe dos transdutores podem ser ajustadas dinamicamente, enquanto os sinais estão sendo *adquiridos*. Vem daí o nome *beamforming* (“formação de feixe”) (WEBB, 2002).

Esse processo é essencialmente o inverso de outra técnica utilizada na *transmissão* dos pulsos de ultrassom, conhecida como *Dynamic Focusing* (“focalização dinâmica”), ilustrada na Figura 3. Nesse exemplo, um *Linear Phased Array* com cinco elementos transdutores é estimulado com pulsos de tensão aplicados em diferentes instantes de tempo. Em (a), os pulsos são aplicados simultaneamente aos elementos 1 e 5; após pequenos intervalos, são aplicados aos elementos 4, 2 e 3, respectivamente. Essa sequência de pulsos induz a formação de uma frente de onda curva, com a distância focal sendo ajustada pela duração dos atrasos de tempo. Em (b), observamos que um atraso de tempo menor entre os pulsos provoca um aumento da distância focal, enquanto em (c) e (d) percebemos que atrasos de tempo assimétricos em relação ao *array* podem ser usados para controlar a direção do feixe emitido, controlando assim a iluminação do ambiente imageado pelas ondas de ultrassom. A focalização dinâmica consiste nesse processo, associado ao aumento progressivo do número de elementos ativados durante a transmissão, de modo a otimizar a resolução lateral da imagem para regiões mais profundas do ambiente imageado (WEBB, 2002).

2.4.2 MODELO MATEMÁTICO DE AQUISIÇÃO

Embora de uso amplamente difundido e em constante aprimoramento pelos fabricantes de equipamentos de ultrassonografia, a técnica de *beamforming* não lança mão de um modelo do sistema físico de imageamento. Como vimos na seção anterior, ela trata o sinal com compensações de atraso temporal, demodulação, controle de ganho e compressão apropriadas, sem construir modelos matemáticos dos transdutores, nem do comportamento do meio imageado quanto à propagação da onda ultrassônica.

A utilização desses modelos pode oferecer condições de reconstruir imagens de melhor qualidade por tirar proveito de componentes do sinal que são ignoradas pelo *beamforming* convencional. A melhoria da qualidade proporcionada por essa abordagem foi mostrada, por exemplo, por Lavarello et al. (2006), Lingvall e Olofsson (2007) e Ellis et al. (2010), enquanto

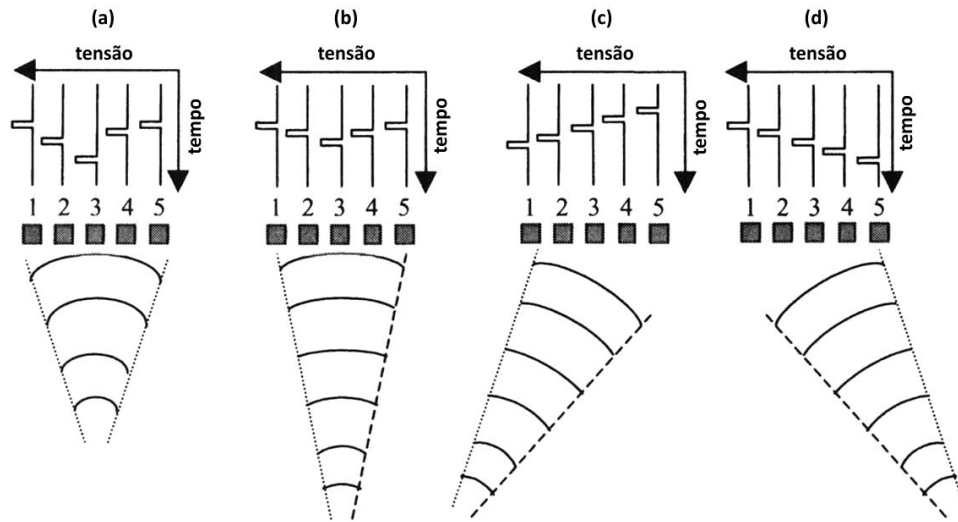


Figura 3: Formação do feixe por focagem dinâmica em ultrassonografia

Fonte: Adaptada de Webb (2002)

Zanin et al. (2011) mostraram que uma melhora significativa nas resoluções axial e lateral em relação ao *beamforming* podem ser alcançadas.

O uso de modelos de aquisição para a reconstrução de imagens é um campo fértil de pesquisa. Esse conceito aplica-se em áreas como a Super-resolução (DENG et al., 2016; EFTEKHARI; WAKIN, 2013; ZIBETTI et al., 2008), a Tomografia Computadorizada (*Computed Tomography*, CT) (ZHANG et al., 2016; ZIBETTI et al., 2015a; NATTERER; WUBBELING, 2001), a Tomografia Fotoacústica (*Photoacoustic Tomography*, PAT) (HUANG et al., 2013; HUANG, 2012; HUANG et al., 2012), e o Imageamento por Ressonância Magnética (*Magnetic Resonance Imaging*, MRI) (CHAUFFERT et al., 2016; PENG et al., 2016), apenas para citar alguns exemplos.

Entretanto, imagens melhores têm um custo, porque os algoritmos iterativos usados na abordagem por problemas inversos em geral requerem um grande esforço computacional. Nessa abordagem, dois aspectos têm grande relevância. Em primeiro lugar, um modelo de aquisição satisfatório e *priors* adequados são cruciais para uma boa reconstrução. Em segundo lugar, algoritmos rápidos são importantes para reduzir o tempo total de reconstrução. É com base nesses dois motivos que traçamos os objetivos desse estudo, apresentados na Seção 1.1.

O modelo representado pela matriz \mathbf{H} tem um papel chave na reconstrução. O modelamento linear apresentado a seguir para a construção da matriz \mathbf{H} é baseado em trabalhos de pesquisadores da comunidade de ultrassom, como os de Stepanishen (1971a, 1971b, 1971c, 1981) e outros mais recentes de Christensen (1988), de Jensen et al. (1993), de Hedrick et al.

(1995), de York e Kim (1999), de Anderson e Trahey (2000), de Demirli e Saniie (2003), de Lingvall e Olofsson (2007) e de Azhari (2010).

Esse tipo de modelamento tem sido investigado para a ultrassonografia de uso geral, como nos trabalhos de Shen e Ebbini (1996) e de Lingvall e Olofsson (2007), e também para a ultrassonografia médica, como nas publicações de Lavarello et al. (2006) e de Viola et al. (2008).

Nos experimentos apresentados no Capítulo 4, a matriz \mathbf{H} é construída com auxílio do programa de uso livre *Field II* (JENSEN et al., 1996; JENSEN; SVENDSEN, 1992), rodando em ambiente MATLAB[®] (The MathWorks Inc., Natick, USA).

Transmissão do Pulso Ultrassônico

Seja $u_k(t)$ o pulso elétrico aplicado ao k -ésimo elemento do transdutor, $h_k^{ef}(t)$ a sua resposta ao impulso *eletro-acústica* direta, e $h_k^{f-SIR}(\mathbf{r}, t)$ a sua resposta ao impulso *espacial* direta, que modela o meio físico entre o emissor e um ponto \mathbf{r} na região imageada, no instante t . A *pressão acústica* $p_k(\mathbf{r}, t)$ no ponto \mathbf{r} associada ao sinal proveniente desse elemento, é dada por

$$p_k(\mathbf{r}, t) = h_k^{f-SIR}(\mathbf{r}, t) * h_k^{ef}(t) * u_k(t), \quad (30)$$

em que $*$ denota uma convolução temporal.

Como em muitas aplicações, consideramos que a resposta ao impulso eletro-acústica é modelada por cosseno modulado em amplitude por um pulso gaussiano (ZANIN et al., 2011):

$$h^{ef}(t) = e^{-\alpha \cdot BW \cdot t^2} \cos(2\pi f_c t), \quad (31)$$

em que BW representa a *largura de banda fracionária* (a razão entre a largura de banda do transdutor e sua frequência central f_c), e α é um fator de atenuação. Nesse estudo, consideramos que todos os elementos do transdutor compartilham a mesma resposta impulsional eletro-acústica direta.

A resposta impulsional espacial direta é dada por

$$h_k^{f-SIR}(\mathbf{r}, t) = \int_S \frac{\delta(t - \frac{|\mathbf{r}|}{v})}{2\pi|\mathbf{r}|} dS, \quad (32)$$

em que $\delta(t - \frac{|\mathbf{r}|}{v})$ é uma função Delta de Dirac deslocada, que faz com que a contribuição do k -ésimo elemento ao longo da área S seja somada apenas quando o ponto \mathbf{r} for excitado pela onda ultrassônica correspondente, que viaja a uma velocidade constante v . Essa resposta impulsional espacial modela o decréscimo do campo acústico à medida que a onda viaja através de um meio

linear homogêneo *não atenuante* (JENSEN, 1999). Se considerarmos a atenuação, a Equação (32) é modificada para

$$h_k^{f-SIR}(\mathbf{r}, t) = \int_T \int_S a\left(t - \tau, |\mathbf{r}|\right) \frac{\delta\left(\tau - \frac{|\mathbf{r}|}{v}\right)}{2\pi|\mathbf{r}|} dS d\tau, \quad (33)$$

em que a é a *resposta impulsional de atenuação*, que supomos ser a mesma em todo o meio (JENSEN et al., 1993). Nesse estudo, o programa *Field II* é ajustado para considerar atenuações dependentes e independentes da frequência. Como o modelo pretende representar um meio de propagação específico, as taxas de atenuação são obtidas a partir das especificações do *phantom* profissional de ultrassom utilizado para a coleta do conjunto de dados. Os detalhes são apresentados no Capítulo 4.

Assumindo que o meio de propagação tem comportamento linear, a contribuição de todos os K elementos do transdutor para a pressão acústica no ponto \mathbf{r} ao longo do tempo é dada pela soma de todos os $p_k(\mathbf{r}, t)$:

$$p(\mathbf{r}, t) = \sum_{k=1}^K h_k^{f-SIR}(\mathbf{r}, t) * h_k^{ef}(t) * u_k(t). \quad (34)$$

Reflexão do Pulso Ultrassônico (Eco)

Cada ponto da região excitada apresenta uma resposta distinta, que varia de acordo com a impedância acústica dessa porção particular do objeto imageado. Na verdade, as *diferenças* de impedância no meio de propagação são responsáveis pela reflexão parcial ou total das ondas incidentes, e elas ocorrem essencialmente em função de variações de densidade, diferentes tipo de materiais, ou nas fronteiras entre estruturas e órgãos (no caso de imagens médicas).

A intensidade de reflexão no ponto \mathbf{r} é quantificada por $f(\mathbf{r}) \in [-1, 1]$, para a qual um valor 0 indica que não há reflexão e -1 ou $+1$ indicam reflexão total, com ou sem inversão de fase, respectivamente. Essa quantidade também é conhecida como *coeficiente de reflexão em amplitude* (*amplitude reflection coefficient*), e difere do *coeficiente de reflexão em intensidade* (*intensity reflection coefficient*), que relaciona as *energias* incidente e refletida no ponto observado (HEDRICK et al., 1995).

Assim, o ponto \mathbf{r} pode ser considerado um emissor acústico, no qual as ondas têm amplitude inicial dada por $p(\mathbf{r}, t)f(\mathbf{r})$, e iniciam seu retorno em direção a todos os K elementos do transdutor.

Seja $h_n^{b-SIR}(\mathbf{r}, t)$ a resposta ao impulso espacial *reversa*, que modela o meio entre o emissor em \mathbf{r} e o n -ésimo elemento do transdutor, no instante t . Seja também $h_n^{eb}(t)$ a resposta

ao impulso eletro-acústica reversa desse n -ésimo elemento. De acordo com Zanin et al. (2011), é razoável considerar que todos os elementos do transdutor compartilham as mesmas respostas impulsivas eletro-acústicas diretas e reversas. Também é plausível assumir que as respostas impulsivas espaciais diretas e reversas sejam iguais, dado que o meio é supostamente linear. Por isso, ambas as premissas são adotadas nesse estudo.

Deste modo, podemos estimar o sinal elétrico $g_n(\mathbf{r}, t)$ gerado por esse elemento receptor ao longo do tempo, associado a um ponto particular \mathbf{r} , por:

$$g_n(\mathbf{r}, t) = h_n^{eb}(t) * h_n^{b-SIR}(\mathbf{r}, t) * p(\mathbf{r}, t) f(\mathbf{r}). \quad (35)$$

Com base nas Equações (34) e (35), obtemos a resposta impulsional total $h_n(\mathbf{r}, t)$ para o n -ésimo elemento receptor do transdutor

$$h_n(\mathbf{r}, t) = h_n^{eb}(t) * h_n^{b-SIR}(\mathbf{r}, t) * \left[\sum_{k=1}^K h_k^{f-SIR}(\mathbf{r}, t) * h_k^{ef}(t) * u_k(t) \right], \quad (36)$$

e a Equação (35) torna-se

$$g_n(\mathbf{r}, t) = h_n(\mathbf{r}, t) f(\mathbf{r}). \quad (37)$$

Integrando e discretizando a Equação (37) ao longo da região de interesse (ROI), obtemos

$$g_n[t_i] = \sum_{\mathbf{r} \in ROI} h_n[\mathbf{r}, t_i] f[\mathbf{r}] + e_n[t_i], \quad (38)$$

em que t_i são os instantes de amostragem, $\mathbf{r} \in ROI$ são pontos no espaço, e e_n representa os erros de discretização e o ruído presentes em g_n .

Os erros implícitos nesse procedimento de amostragem são bem conhecidos na literatura geral de processamento de sinais e no contexto da ciência das imagens (ver, por exemplo, os trabalhos de Marks (1991) e de Barrett e Myers (2004)). Nesse estudo, utilizamos uma frequência de amostragem quatro vezes maior que a frequência do pulso ultrassônico para a discretização no tempo e uma resolução de um comprimento de onda para a discretização da ROI. Enquanto o critério de Nyquist garante que não haverá perda de dados na amostragem temporal, nós de fato perdemos a informação espacial dos interstícios, isto é, das regiões entre os pontos mapeados (também chamados de *pontos fora da grade*).

Modelo do Sistema

Podemos exprimir a Equação (38) na forma matricial, com uma expressão para cada elemento do transdutor, escrevendo $\mathbf{g}_n = \mathbf{H}_n \mathbf{f} + \mathbf{e}_n$. Ou, de modo conveniente, podemos combinar todos os K elementos do transdutor em uma única equação matricial, concatenando

vetores e matrizes apropriadamente:

$$\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_K \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_K \end{bmatrix} \mathbf{f} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_K \end{bmatrix},$$

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{e}. \quad (39)$$

Esse sistema linear representa o processo de aquisição. A Figura 4 mostra uma representação esquemática das estruturas internas dos vetores \mathbf{g} , \mathbf{f} e \mathbf{e} , e da matriz \mathbf{H} .

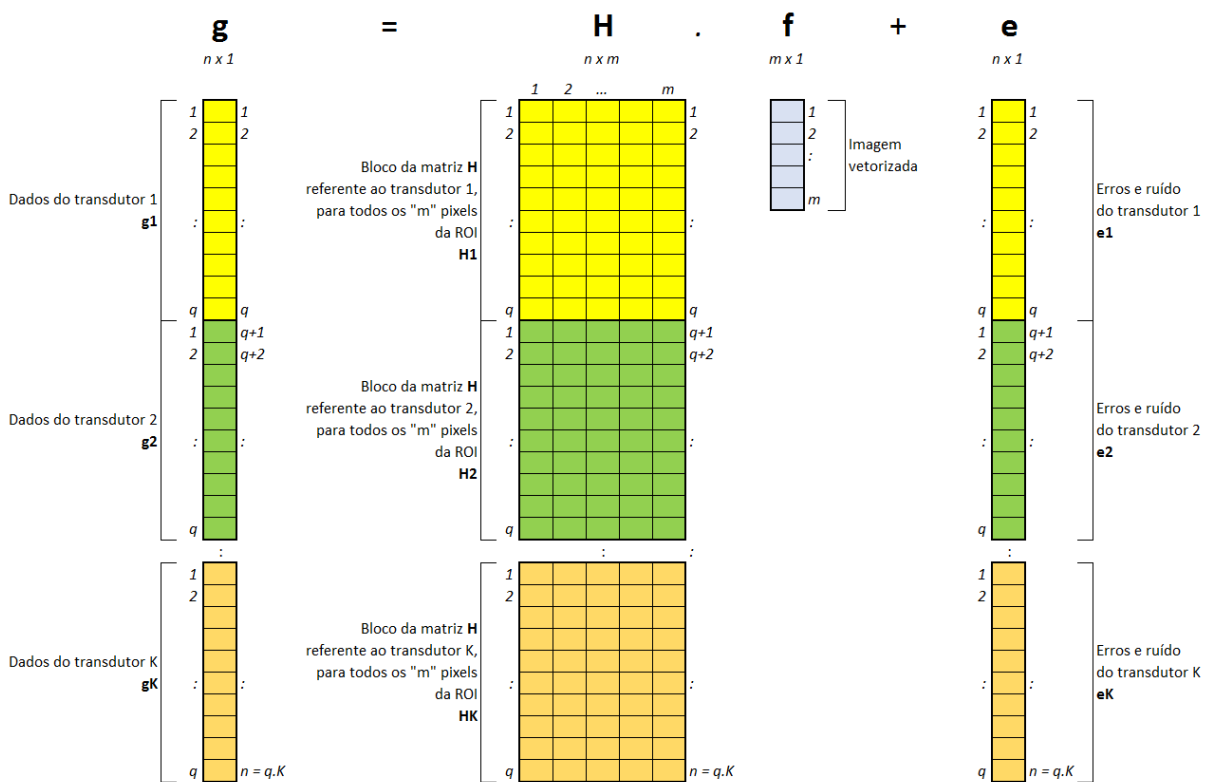


Figura 4: Representação esquemática das estruturas dos vetores e da matriz \mathbf{H}

Fonte: O autor (2017)

A distribuição dos elementos da matriz \mathbf{H} depende da escolha das posições dos elementos do *array* no sistema de coordenadas usado para localizar os pontos \mathbf{r} na ROI. Nesse estudo, adotamos uma distribuição *simétrica*, com o conjunto de transdutores posicionado de modo paralelo à borda superior da ROI. Metade dos elementos estão à esquerda do eixo central da ROI, e metade à direita. Ao usar esse posicionamento simétrico, é possível simplificar o cálculo de \mathbf{H} ; realizamos os cálculos apenas para um lado da ROI (e.g. o esquerdo), e então espelhamos os elementos de \mathbf{H} para obter aqueles que correspondem ao lado oposto. Essa

estratégia reduz à metade o tempo de construção do modelo.

Como detalhado no Capítulo 4, usamos $K = 64$ elementos do transdutor. A Figura 5 ilustra visualmente a estrutura interna da matriz \mathbf{H} . Em (a), as três faixas horizontais representam as sub-matrizes \mathbf{H}_1 a \mathbf{H}_3 ; em (b) estão representadas as sub-matrizes \mathbf{H}_{31} a \mathbf{H}_{33} e em (c), as sub-matrizes \mathbf{H}_{62} a \mathbf{H}_{64} . As regiões negras na figura indicam elementos nulos, enquanto os pontos brancos representam elementos não nulos, o que evidencia a esparsidade da matriz de modelamento.

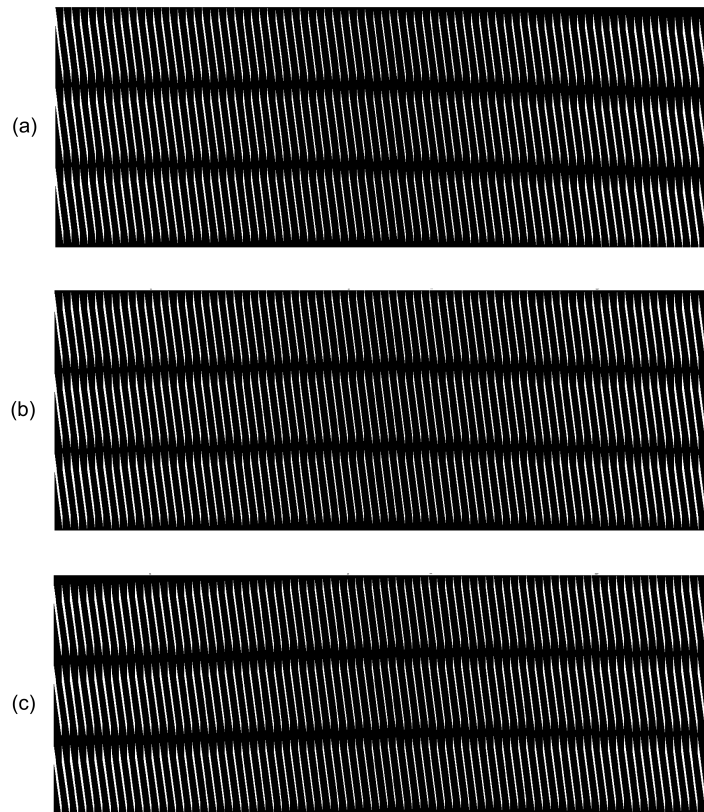


Figura 5: Representação visual dos elementos da matriz \mathbf{H}

Fonte: O autor (2017)

Uma vez que a imagem bidimensional a ser reconstruída é representada pelo vetor unidimensional $\mathbf{f} \in \mathbb{R}^m$, o número de colunas da matriz $\mathbf{H} \in \mathbb{R}^{n \times m}$ é igual ao número de *pixels* da imagem. Como as colunas da imagem bidimensional estão empilhadas em \mathbf{f} , isso explica o padrão periódico de linhas diagonais em cada sub-matriz \mathbf{H}_n .

Observamos agora os conjuntos de dados esperados a partir desse modelo. A Figura 6 mostra a aparência dos dados \mathbf{g} quando há apenas um refletor pontual no centro da ROI, enquanto a Figura 7 mostra \mathbf{g} no caso hipotético em que todos os pontos da ROI possuem refletores. Em ambas as figuras, cada linha vertical representa o conjunto \mathbf{g}_n do n -ésimo

elemento, com o tempo t_i passando de cima para baixo. Na Figura 7 observamos ainda o efeito da atenuação com o passar do tempo t_i , mostrando que níveis mais baixos de sinal são esperados de pontos mais distantes da ROI.

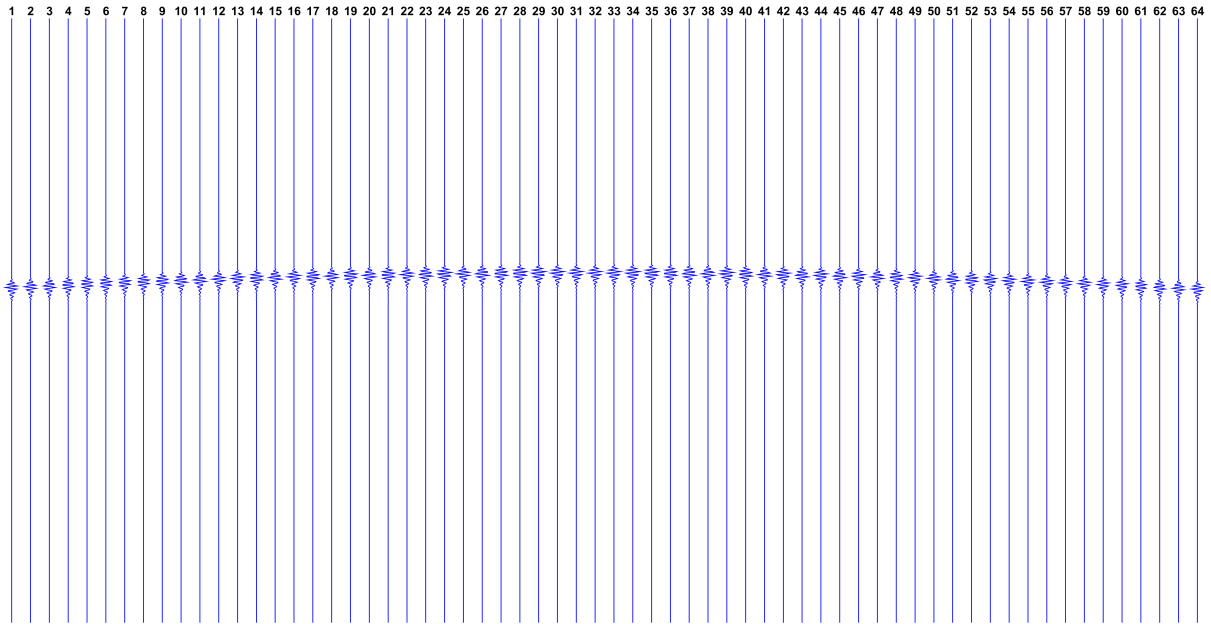


Figura 6: Conjunto de dados g – refletor pontual único no centro da ROI

Fonte: O autor (2017)

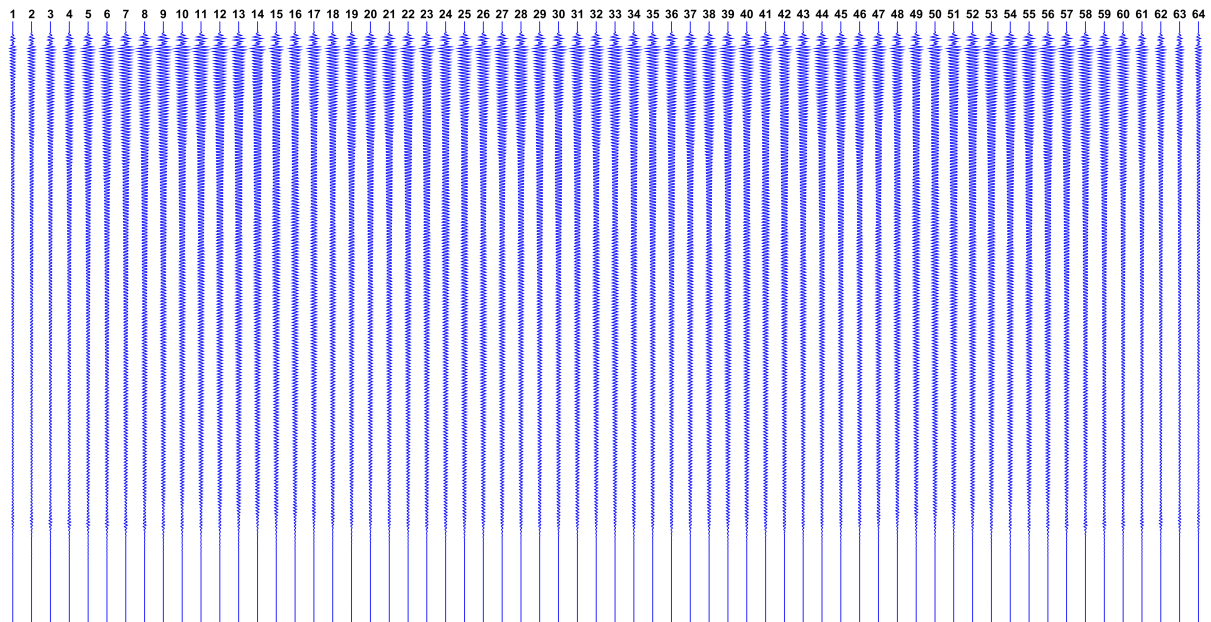


Figura 7: Conjunto de dados g – refletores em todos os pontos da ROI

Fonte: O autor (2017)

Outra possibilidade para a construção e uso do modelo de aquisição é evitar o armazenamento da matriz \mathbf{H} completa, dado que sabemos como calcular cada um de seus elementos. Com o aumento do tamanho das imagens a serem reconstruídas, as dimensões de \mathbf{H} crescem muito rapidamente; isso pode tornar o processo de reconstrução muito lento, ou mesmo torná-lo inviável, dadas as restrições de memória das plataformas computacionais atuais. Por exemplo, no estudo de caso apresentado no Capítulo 4, utilizamos matrizes \mathbf{H} com tamanhos distintos: uma para as seções 4.2.1 e 4.2.2, e outra para a seção 4.3.3. A primeira, utilizada para a reconstrução de imagens de 81×81 *pixels*, possui 50.944×6.561 elementos em formato *double*. A matriz ocupa $2.673.948.672$ *bytes* de memória e é armazenada em um arquivo MATLAB[®] de extensão *.MAT*. Dada a quantidade elevada de elementos nulos, o arquivo ocupa um espaço relativamente reduzido no disco rígido (aproximadamente 405 MB). A segunda matriz \mathbf{H} é utilizada para a reconstrução de imagens de 73×89 *pixels* e possui 58.368×6.497 elementos, também em formato *double*, que ocupam $3.033.735.168$ *bytes* de memória. Há muitos elementos igualmente nulos, e o arquivo *.MAT* que ocupa aproximadamente 547 MB no disco rígido.

Para contornar o problema da grande de ocupação de memória RAM, podemos calcular os elementos de \mathbf{H} *durante* o processo de reconstrução, desde que a plataforma de reconstrução tenha capacidade de processamento suficiente. Nas reconstruções de ultrassom realizadas nos Capítulos 4 e 5, as matrizes \mathbf{H} foram armazenadas de modo integral em memória RAM (*Random Access Memory*) durante o processamento. Por outro lado, as reconstruções de Ressonância Magnética (MRI), realizadas no Capítulo 5, utilizaram matrizes \mathbf{H} implementadas como funções *in-line* no MATLAB[®], de modo que não temos as matrizes verdadeiras armazenadas em memória, mas sim funções que as representam.

3 ALGORITMOS ITERATIVOS DE RECONSTRUÇÃO

Devido à característica da norma ℓ_1 de induzir soluções esparsas quando utilizada como termo de regularização, a solução de reconstrução $\hat{\mathbf{f}}$ para a qual voltamos nossa atenção neste estudo é estimada por

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi(\mathbf{f}) = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \|\mathbf{f}\|_1. \quad (40)$$

A função custo $\Psi(\mathbf{f})$ é conhecida pelos nomes *Basis Pursuit Denoising* (BPDN) e *Least Absolute Shrinkage and Selection Operator* (LASSO). Curiosamente, os pesquisadores que propuseram esses nomes tão distintos trabalham no mesmo departamento de Estatística da Universidade de Stanford (EUA), e desenvolveram a mesma formulação, ao mesmo tempo (ELAD, 2010).

Em função da norma ℓ_1 , a função $\Psi(\mathbf{f})$ é não-quadrática e sua minimização corresponde à solução de um sistema de equações *não linear*, que deve ser resolvido de forma iterativa. De acordo com Elad (2010), ainda que (40) seja um problema *convexo*, a presença da norma ℓ_1 impossibilita a garantia de que a solução seja única, de modo que o problema *não é estritamente convexo*.

Neste trabalho, voltamos nossa atenção para um conjunto restrito de métodos iterativos voltados à minimização de funções custo não lineares convexas, resumidas nas seções subsequentes deste capítulo. Entretanto, salientamos que o estudo de algoritmos para a otimização ℓ_2 - ℓ_1 é um campo de pesquisa intenso e esse estudo não tem a intenção de avaliar os avanços mais recentes na área. Alguns dos resultados mais atuais foram publicados, por exemplo, por Byrd et al. (2016, 2015b, 2015a), por Muckley et al. (2015) e por Lee et al. (2012), além de outros citados nas referências desses trabalhos. Estudos específicos para a área de ultrassom também têm sido desenvolvidos, como por exemplo o de Wu et al. (2015).

3.1 BUSCA EM LINHA (*LINE SEARCH*)

Os métodos iterativos para otimização convexa são usualmente baseados em alguma forma de gradiente, e o seu desempenho pode apresentar uma grande dependência da intensidade com que o algoritmo avança em uma dada direção de busca. O progresso em direção ao mínimo é ajustado por um parâmetro que chamamos de *passo* (*stepsize*). Em alguns métodos o passo é fixo, enquanto em outros um valor mais adequado (ótimo, próximo do ótimo ou que atenda algum critério de redução pré-estabelecido) é calculado a cada iteração. Nesses casos, um procedimento de otimização é realizado a cada iteração do método e, em geral, representa um sobrecusto computacional pequeno com o aumento das dimensões do problema.

O cálculo do passo recebe o nome de *busca em linha*¹ (*line search*), visto que consiste em uma minimização ao longo da linha definida pela direção de busca. Dadas a solução \mathbf{f}_k e a direção de busca \mathbf{d}_k na k -ésima iteração, o procedimento de busca em linha consiste em uma minimização unidimensional, uma vez que avaliamos o custo apenas em função do passo. O valor ótimo α_k para esse escalar é então dado por

$$\alpha_k = \arg \min_{\alpha} \Psi(\mathbf{f}_k + \alpha \mathbf{d}_k) . \quad (41)$$

Embora muitos tipos de busca em linha possam ser aplicados à minimização ℓ_2 - ℓ_1 , nem sempre temos a garantia de atingir o passo ótimo em um tempo razoável. É por esse motivo que alguns pesquisadores têm proposto recentemente procedimentos de busca em linha específicos para o caso ℓ_2 - ℓ_1 , que têm se mostrado mais eficientes que métodos mais genéricos.

Alguns dos operadores de busca em linha ℓ_2 - ℓ_1 são os propostos por Osher e Li (2009), por Wen et al. (2010) e por Zibetti et al. (2015b, 2015a). Alguns dos algoritmos avaliados nesse estudo utilizam a busca em linha, e todos eles são identificados com o sufixo “OLS”, uma sigla para o termo *Optimal Line Search* (busca em linha ótima). Por homogeneidade, adotamos o operador proposto por Zibetti et al. (2015b) em todos os experimentos realizados nesse trabalho.

3.2 ALGORITMOS DA FAMÍLIA ISTA

Os algoritmos iterativos *shrinkage-thresholding* (ISTA) constituem uma família de métodos de otimização convexa que utilizam o operador $S_a(x)$ de mesmo nome, definido como:

$$S_a(x) = \begin{cases} 0 & , a \geq |x| \\ x - a \cdot \text{sign}(x) & , a < |x| \end{cases} . \quad (42)$$

¹Esta é a designação usual na Engenharia. Na Matemática, o nome usual é *busca linear*.

Na Equação (42) o operando x é um escalar. Nos algoritmos, no entanto, usamos uma notação vetorial para indicar que o *shrinkage-thresholding* opera sobre uma variável multidimensional. Para um vetor $\mathbf{x} \in \mathbb{R}^w$, por exemplo, escrevemos $S_a(\mathbf{x})$ para indicar que o operador é aplicado individualmente a cada um de seus elementos, resultando em $S_a(\mathbf{x}) = [S_a(x_1) S_a(x_2) \dots S_a(x_w)]^T$.

Muitos métodos foram propostos com base nesse operador, os quais têm a vantagem de serem simples, e para os quais a convergência tem sido bastante estudada e garantida, como por Daubechies et al. (2004).

Entretanto, os algoritmos ISTA são reconhecidos por serem lentos e, por isso, várias propostas de aceleração vêm sendo apresentadas nos últimos anos. Nesterov (1983, 2005) propôs um método de aceleração que fez as taxas de convergência de métodos baseados no gradiente passarem da ordem $O(1/k)$ para $O(1/k^2)$, sendo k o número da iteração. Mais tarde, Beck e Teboulle (2009a) propuseram outra forma de aceleração do ISTA que também garante taxa de convergência da ordem $O(1/k^2)$, mas difere da proposta de Nesterov. Esses pesquisadores batizaram seu algoritmo de FISTA (*Fast Iterative Shrinkage-Thresholding Algorithm*) (ZIBULEVSKY; ELAD, 2010).

O método FISTA e suas variantes, apesar de relativamente recentes, têm mostrado ser muito eficientes na minimização $\ell_2\text{-}\ell_1$, especialmente em problemas de grandes dimensões. Todos eles baseiam-se nos mesmos passos básicos: o cálculo do vetor de resíduos $\mathbf{g} - \mathbf{H}\mathbf{f}$ no espaço dos dados e sua projeção de volta ao espaço da imagem pela multiplicação por \mathbf{H}^T , seguida da aplicação do operador de *shrinkage-thresholding* (ELAD, 2010).

As iterações prosseguem até que algum critério de parada seja alcançado. Tanto para o FISTA, quanto para os demais métodos avaliados neste estudo, os critérios de parada adotados estão descritos nos experimentos dos Capítulos 4 e 5.

Como observamos no Quadro 1, o algoritmo FISTA estima um novo \mathbf{f}_k por meio do operador de *shrinkage-thresholding* (passo 6) a cada iteração. O vetor passado como argumento ao operador, é obtido pela soma de um ponto especial \mathbf{y}_k ao negativo do gradiente nesse ponto, escalonado pelo parâmetro c .

O ponto \mathbf{y}_k é calculado pela combinação de duas soluções anteriores, e representa a modificação chave que acelera o FISTA em relação aos primeiros métodos ISTA. Enquanto as iterações evoluem, o vetor \mathbf{y}_k modifica o ponto da solução sobre o qual o operador de *shrinkage-thresholding* é aplicado, aumentando gradualmente o *momento (momentum)* do algoritmo, proporcionando boa velocidade de convergência. Embora rápido, o FISTA requer algumas

Quadro 1: Algoritmo FISTA

Entradas: $\lambda, c \geq \|\mathbf{H}^T \mathbf{H}\|_2$

- 1: $\mathbf{f}_0 = \mathbf{0}$
- 2: $\mathbf{y}_1 = \mathbf{f}_0$
- 3: $t_1 = 1$
- 4: $k = 1$
- 5: **enquanto** Critério de parada não atingido **faça**
- 6: $\mathbf{f}_k = S_{\lambda/c} \left[\frac{1}{c} \mathbf{H}^T (\mathbf{g} - \mathbf{H} \mathbf{y}_k) + \mathbf{y}_k \right]$
- 7: $k = k + 1$
- 8: $t_k = \left(1 + \sqrt{1 + 4t_{k-1}^2} \right) / 2$
- 9: $\mathbf{y}_k = \mathbf{f}_{k-1} + \frac{t_{k-1} - 1}{t_k} (\mathbf{f}_{k-1} - \mathbf{f}_{k-2})$
- 10: **fim enquanto**

Fonte: (BECK; TEBoulLE, 2009a)

iterações para ganhar momento; isso significa que suas primeiras iterações são semelhantes às dos algoritmos ISTA não acelerados.

A convergência do FISTA é garantida para valores do parâmetro c maiores que a constante de Lipschitz do gradiente do termo diferenciável da função custo, isto é, o termo quadrático. Essa constante está relacionada ao maior valor singular da matriz $\mathbf{H}^T \mathbf{H}$ e uma escolha típica é selecionar $c \geq \|\mathbf{H}^T \mathbf{H}\|_2$ (BECK; TEBoulLE, 2009a).

Apesar da vantagem de ser mais rápido que outros métodos de sua família, o FISTA não é um algoritmo *monotônico* porque não garante que uma nova solução \mathbf{f}_k tenha custo inferior à solução anterior. Por isso, os mesmos pesquisadores propuseram o *Monotone Fast Iterative Shrinkage-Thresholding Algorithm* (MFISTA) algum tempo depois, em (BECK; TEBoulLE, 2009b).

O algoritmo MFISTA é mostrado no Quadro 2. A monotonicidade é obtida no passo 7 por meio de um novo ponto intermediário \mathbf{z}_k , que também é usado em uma nova expressão para o cálculo do ponto \mathbf{y}_k .

Uma terceira variante, que chamamos de OMFISTA (*Over-Relaxation of Monotone Fast Iterative Shrinkage-Thresholding Algorithm*), foi proposta por Yamagishi e Yamada (2011).

O OMFISTA é mostrado no Quadro 3, e calcula a nova solução estimada \mathbf{f}_k usando um passo variável α_k . Embora use o mesmo ponto intermediário \mathbf{z}_k do MFISTA, a nova solução

Quadro 2: Algoritmo MFISTA

Entradas: $\lambda, c \geq \|\mathbf{H}^T \mathbf{H}\|_2$

1: $\mathbf{f}_0 = \mathbf{0}$

2: $\mathbf{y}_1 = \mathbf{f}_0$

3: $t_1 = 1$

4: $k = 1$

5: **enquanto** Critério de parada não atingido **faça**

6: $\mathbf{z}_k = S_{\lambda/c} \left[\frac{1}{c} \mathbf{H}^T (\mathbf{g} - \mathbf{H}\mathbf{y}_k) + \mathbf{y}_k \right]$

7: $\mathbf{f}_k = \arg \min_{\mathbf{f}} \Psi(\mathbf{f}) \mid \mathbf{f} \in \{\mathbf{z}_k, \mathbf{f}_{k-1}\}$

8: $k = k + 1$

9: $t_k = \left(1 + \sqrt{1 + 4t_{k-1}^2} \right) / 2$

10: $\mathbf{y}_k = \mathbf{f}_{k-1} + \frac{t_{k-1} - 1}{t_k} (\mathbf{f}_{k-1} - \mathbf{f}_{k-2}) + \frac{t_{k-1}}{t_k} (\mathbf{z}_{k-1} - \mathbf{f}_{k-1})$

11: **fim enquanto**

Fonte: (BECK; TEBoulLE, 2009b)

Quadro 3: Algoritmo OMFISTA

Entradas: $\alpha_1, \eta_1, \lambda, c \geq \|\mathbf{H}^T \mathbf{H}\|_2$

1: $\mathbf{f}_0 = \mathbf{0}$

2: $\mathbf{y}_1 = \mathbf{f}_0$

3: $t_1 = \alpha_1$

4: $k = 1$

5: **enquanto** Critério de parada não atingido **faça**

6: $\mathbf{z}_k = S_{\lambda/c} \left[\frac{1}{c} \mathbf{H}^T (\mathbf{g} - \mathbf{H}\mathbf{y}_k) + \mathbf{y}_k \right]$

7: Calcular α_k (via *Busca em Linha* se OLS)

8: $\mathbf{f}_k = \arg \min_{\mathbf{f}} \Psi(\mathbf{f}) \mid \mathbf{f} \in \{\mathbf{f}_{k-1} + \alpha_k (\mathbf{z}_k - \mathbf{f}_{k-1}), \mathbf{f}_{k-1}\}$

9: $k = k + 1$

10: $t_k = \left(\alpha_1 \alpha_{k-1} + \sqrt{\alpha_1^2 \alpha_{k-1}^2 + 4t_{k-1}^2} \right) / 2$

11: $\mathbf{y}_k = \mathbf{f}_{k-1} + \frac{t_{k-1} - \alpha_1}{t_k} (\mathbf{f}_{k-1} - \mathbf{f}_{k-2}) + \frac{t_{k-1}}{t_k} (\mathbf{z}_{k-1} - \mathbf{f}_{k-1})$
 $+ \frac{t_{k-1}}{t_k} (1 - \eta_{k-1}) (\mathbf{y}_{k-1} - \mathbf{z}_{k-1})$

12: **fim enquanto**

Fonte: (YAMAGISHI; YAMADA, 2011)

\mathbf{f}_k do OMFISTA é obtida por meio da comparação do custo da solução anterior com o custo atingido ao avançar com o passo dado em uma direção especial.

Essa formulação permite que o algoritmo seja ajustado com um passo apropriado, que pode ser fixo ou variável com o passar das iterações. Ao custo de ter esse parâmetro para ser ajustado, o OMFISTA pode apresentar um desempenho superior ao do MFISTA, segundo apontam os propositores.

Assim como t_k , o ponto \mathbf{y}_k também é calculado de modo diferente. O novo parâmetro η_k depende do c escolhido e do passo α_k , entre outros fatores (YAMAGISHI; YAMADA, 2011). Nos experimentos realizados nesse estudo, consideramos α_1 e α_k como um único passo fixo α , e η_k como um parâmetro η fixo.

O quarto algoritmo da família ISTA avaliado neste estudo é designado de OMFISTA-OLS (*Over-Relaxation of Monotone Fast Iterative Shrinkage-Thresholding Algorithm with Optimal Line Search*). Trata-se de uma variante do OMFISTA proposta por Zibetti et al. (2015a) para Tomografia Computadorizada. Nessa variante, o passo α_k é calculado pelo operador de busca em linha proposto por Zibetti et al. (2015b), diferentemente do passo fixo adotado no OMFISTA.

3.3 GRADIENTE CONJUGADO NÃO LINEAR

O método do Gradiente Conjugado (*Conjugate Gradient*, CG) é um dos principais algoritmos para a resolução de sistemas de equações *lineares* de grandes dimensões. Ele é particularmente interessante para matrizes mal-condicionadas, porque converge mais rápido que os métodos baseados apenas no gradiente e tem convergência garantida em um número finito de iterações (SHEWCHUK, 1994; LUENBERGER; YE, 2008).

Derivado do CG, o Método do Gradiente Conjugado Não Linear (*Nonlinear Conjugate Gradient*, NLCG) fornece uma solução aproximada para sistemas não lineares, tais como o da Equação (40). A solução obtida será mais próxima de exata para pequenos valores do parâmetro $\varepsilon \in \mathbb{R}_+^*$, utilizado para aproximar o subgradiente da norma ℓ_1 com uma função substituta estritamente convexa.

Embora similares, o NLCG e o CG diferem em alguns aspectos; um deles é relativo ao parâmetro β , elemento chave na definição da direção de busca em uma dada iteração. Considerando um sistema com uma matriz de coeficientes \mathbf{A} , esse parâmetro garante que a próxima direção de busca será sempre \mathbf{A} -ortogonal a todas as direções anteriores no caso linear, garantindo convergência após um número finito de iterações. Entretanto, de acordo com Shewchuk (1994), isso não é mais verdade para o caso não linear e muitos pesquisadores têm proposto diferentes escolhas para β . Um levantamento interessante sobre muitas dessas

Quadro 4: Algoritmo NLCG

Entradas: ε, λ

- 1: $\mathbf{f}_1 = \mathbf{0}$
 - 2: $\mathbf{r}_1 = \mathbf{g} - \mathbf{H}\mathbf{f}_1$
 - 3: $\mathbf{W}_1 = \text{diag}[1/(|\mathbf{f}_1| + \varepsilon)]$
 - 4: $\nabla\Psi_1 = -\mathbf{H}^T(\mathbf{g} - \mathbf{H}\mathbf{f}_1) + \lambda\mathbf{W}_1\mathbf{f}_1$
 - 5: $\mathbf{d}_1 = -\nabla\Psi_1$
 - 6: $k = 1$
 - 7: **enquanto** Critério de parada não atingido **faça**
 - 8: $\alpha_k = \frac{-\nabla\Psi_k^T \mathbf{d}_k}{[(\mathbf{H}\mathbf{d}_k)^T(\mathbf{H}\mathbf{d}_k) + \lambda(\mathbf{d}_k^T \mathbf{W}_k \mathbf{d}_k)]}$ (via *Busca em Linha* se OLS)
 - 9: $\mathbf{f}_{k+1} = \mathbf{f}_k + \alpha_k \mathbf{d}_k$
 - 10: $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{H}\mathbf{d}_k$
 - 11: $\mathbf{W}_{k+1} = \text{diag}[1/(|\mathbf{f}_{k+1}| + \varepsilon)]$
 - 12: $\nabla\Psi_{k+1} = -\mathbf{H}^T \mathbf{r}_{k+1} + \lambda \mathbf{W}_{k+1} \mathbf{f}_{k+1}$
 - 13: $\mathbf{y}_k = \nabla\Psi_{k+1} - \nabla\Psi_k$
 - 14: $\beta_k = (\nabla\Psi_{k+1}^T \mathbf{y}_k) / (\mathbf{d}_k^T \mathbf{y}_k)$
 - 15: $\mathbf{d}_{k+1} = -\nabla\Psi_{k+1} + \beta_k \mathbf{d}_k$
 - 16: $k = k + 1$
 - 17: **fim enquanto**
-

Fonte: (SHEWCHUK, 1994)

propostas é apresentado por Hager e Zhang (2006).

O algoritmo NLCG é mostrado no Quadro 4, em que \mathbf{r} representa o resíduo, \mathbf{W} é uma matriz diagonal de pesos cujos elementos são calculados com o referido parâmetro de aproximação; neste estudo adotamos $\varepsilon = 10^{-4}$. O parâmetro α é o passo na direção de busca \mathbf{d} e \mathbf{y} é uma variável auxiliar. A escolha para β expressa no passo 14 foi proposta por Hestenes e Stiefel (1952); outras alternativas são brevemente citadas na Seção 4.1.

Neste estudo também avaliamos uma variante do NLCG que chamamos de NLCG-OLS (*Nonlinear Conjugate Gradient with Optimal Line Search*). Ela consiste na substituição do passo 8 do NLCG, de modo que obtemos o passo ótimo α_k a cada iteração, por meio do operador de busca em linha $\ell_2\text{-}\ell_1$ proposto por Zibetti et al. (2015b).

3.4 QUADRADOS MÍNIMOS ITERATIVAMENTE REPONDERADOS

O Método dos Quadrados Mínimos Iterativamente Reponderados (*Iteratively Reweighted Least Squares*, IRLS) é bastante conhecido na literatura, e representa uma opção simples e atrativa para a solução de sistemas não lineares como os da Equação (40) (BJORCK,

Quadro 5: Algoritmo IRLS

Entradas: ε, λ

- 1: $\mathbf{f}_1 = \mathbf{0}$
 - 2: $k = 1$
 - 3: **enquanto** Critério de parada não atingido **faça**
 - 4: $\mathbf{r}_k = \mathbf{g} - \mathbf{H}\mathbf{f}_k$
 - 5: $\nabla \mathbf{f}_k = -\mathbf{H}^T \mathbf{r}_k$
 - 6: $\mathbf{W}_k = \text{diag}[1/(|\mathbf{f}_k| + \varepsilon)]$
 - 7: $\mathbf{d}_k = (\mathbf{H}^T \mathbf{H} + \lambda \mathbf{W}_k)^{-1} (-\nabla \mathbf{f}_k - \lambda \mathbf{W}_k \mathbf{f}_k)$
 - 8: $\mathbf{f}_{k+1} = \mathbf{f}_k + \mathbf{d}_k$
 - 9: $k = k + 1$
 - 10: **fim enquanto**
-

Fonte: (BJORCK, 1996)

1996).

Ele consiste na obtenção de uma aproximação quadrática para a função custo, reponderada a cada iteração, seguida do cálculo da sua solução de quadrados mínimos. Como no método NLCG, o IRLS também utiliza um parâmetro $\varepsilon \in \mathbb{R}_+^*$ para aproximar o subgradiente da norma ℓ_1 com uma função substituta estritamente convexa; neste estudo, adotamos $\varepsilon = 10^{-4}$. O algoritmo IRLS é mostrado no Quadro 5.

Enquanto o IRLS claramente usa um passo de busca unitário (ver o passo 8 do algoritmo), uma variante que designamos IRLS-OLS (*Iteratively Re-weighted Least Squares with Optimal Line Search*) é obtida ao calcularmos explicitamente um passo ótimo α_k por meio do procedimento de busca em linha de Zibetti et al. (2015b). Nesse caso, a nova solução \mathbf{f}_{k+1} torna-se $\mathbf{f}_k + \alpha_k \mathbf{d}_k$ ao invés de $\mathbf{f}_k + \mathbf{d}_k$, sendo \mathbf{d}_k a direção de busca da k -ésima iteração. O cálculo do passo ótimo a cada iteração promove uma convergência mais rápida que a do IRLS, como está mostrado nos experimentos do Capítulo 4.

Outra variante avaliada neste estudo vem do emprego do método CG para obter a solução \mathbf{d}_k para o sistema de equações $(\mathbf{H}^T \mathbf{H} + \lambda \mathbf{W}_k) \mathbf{d}_k = (-\nabla \mathbf{f}_k - \lambda \mathbf{W}_k \mathbf{f}_k)$, ao invés de realizar o cálculo explícito da matriz inversa $(\mathbf{H}^T \mathbf{H} + \lambda \mathbf{W}_k)^{-1}$, como indica o passo 7 do IRLS. Chamamos essa variante de algoritmo IRLS-CG (*Iteratively Re-weighted Least Squares with Conjugate Gradient*). A sua velocidade de convergência relativa ao IRLS depende das dimensões do problema, uma vez que elas podem impor um tempo maior ou menor para a inversão da matriz em relação ao tempo de cálculo via método CG.

A quarta variante do IRLS que avaliamos é designada de IRLS-CG-OLS (*Iteratively*

Re-weighted Least Squares with Conjugate Gradient and Optimal Line Search). Ela combina o uso do método CG para o cálculo da direção de busca \mathbf{d}_k a cada iteração, com o emprego do procedimento de busca em linha de Zibetti et al. (2015b) para obter o passo ótimo α_k nessa direção.

3.5 MULTIPLICADORES DE LAGRANGE COM DIREÇÕES ALTERNADAS

O algoritmo *Alternating Direction Method of Multipliers* (ADMM) foi inicialmente proposto por Gabay e Mercier (1976) e por Glowinski e Marroco (1975), enquanto Boyd (2010) mostra sua aplicação na otimização convexa, particularmente para problemas de grandes dimensões. Esse algoritmo propõe a solução de um problema convexo com restrição de igualdade por meio do método dos Multiplicadores de Lagrange, com um operador de Lagrange *aumentado*.

Dada uma função $q : \mathbb{R}^m \rightarrow \mathbb{R}$, consideramos o problema de minimização

$$\begin{aligned} \min \quad & q(\mathbf{x}), \\ \text{sujeito a:} \quad & \mathbf{Ax} = \mathbf{b}, \end{aligned} \tag{43}$$

que tem o lagrangiano dado por

$$L(\mathbf{x}, \mathbf{y}) = q(\mathbf{x}) + \mathbf{y}^T (\mathbf{Ax} - \mathbf{b}). \tag{44}$$

A chamada *função dual* é definida como

$$r(\mathbf{y}) = \inf_{\mathbf{x}} L(\mathbf{x}, \mathbf{y}) = -q^*(-\mathbf{A}^T \mathbf{y} - \mathbf{b}^T \mathbf{y}), \tag{45}$$

em que \mathbf{y} é a *variável dual* (ou multiplicador de Lagrange) e q^* é a função conjugada² convexa de q . O problema dual consiste em maximizar $r(\mathbf{y})$ com respeito a \mathbf{y} (BOYD; VANDENBERGHE, 2004).

A convergência dos métodos baseados no lagrangiano depende, entretanto, de algumas premissas especiais sobre a função q , tais como finitude e convexidade estrita. A convergência passa a não depender mais dessas premissas quando a Equação (44) recebe um termo adicional,

²Seja $f : \mathbb{R}^m \rightarrow \mathbb{R}$. A função $f^* : \mathbb{R}^m \rightarrow \mathbb{R}$, definida como $f^*(z) = \sup_{\mathbf{x} \in \text{dom} f} (\mathbf{z}^T \mathbf{x} - f(\mathbf{x}))$ é chamada de *conjugada* da função f . O domínio da função conjugada consiste nos pontos $\mathbf{z} \in \mathbb{R}^m$ para os quais o supremo é finito, isto é, para os quais a diferença $\mathbf{z}^T \mathbf{x} - f(\mathbf{x})$ é limitada superiormente no domínio de f (BOYD; VANDENBERGHE, 2004). A definição de função *convexa* é dada na Seção 5.1.1.

Quadro 6: Algoritmo ADMM

Entradas: ρ, λ

- 1: $\mathbf{f}_0 = \mathbf{0}$
 - 2: $\mathbf{y}_0 = \mathbf{0}$
 - 3: $k = 1$
 - 4: **enquanto** Critério de parada não atingido **faça**
 - 5: $\mathbf{x}_k = (\mathbf{H}^T \mathbf{H} + \rho \mathbf{I})^{-1} (\mathbf{H}^T \mathbf{g} + \rho \mathbf{f}_{k-1} - \mathbf{y}_{k-1})$
 - 6: $\mathbf{f}_k = S_{\lambda/\rho} \left(\mathbf{x}_k + \frac{\mathbf{y}_{k-1}}{\rho} \right)$
 - 7: $\mathbf{y}_k = \mathbf{y}_{k-1} + \rho (\mathbf{x}_k - \mathbf{f}_k)$
 - 8: $k = k + 1$
 - 9: **fim enquanto**
-

Fonte: (BOYD, 2010)

dando origem ao *lagrangiano aumentado*

$$L_\rho(\mathbf{x}, \mathbf{y}) = q(\mathbf{x}) + \mathbf{y}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) + \frac{\rho}{2} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2, \quad (46)$$

em que $\rho > 0$ é chamado de *parâmetro de penalidade*. Nesse caso, a função dual torna-se

$$r_\rho(\mathbf{y}) = \inf_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{y}). \quad (47)$$

O algoritmo que soluciona o problema (43) é chamado de *método dos Multiplicadores de Lagrange*, que consiste nos passos

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{y}_k) \quad (48)$$

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \rho (\mathbf{A}\mathbf{x}_{k+1} - \mathbf{b}), \quad (49)$$

com índices indicando os números das iterações. Esses dois passos representam, respectivamente, a minimização em \mathbf{x} e a atualização da variável dual por meio da avaliação do resíduo da restrição de igualdade (BOYD, 2010).

No caso da minimização ℓ_2 - ℓ_1 da Equação (40), o uso do ADMM consiste em reescrever o problema com *duas* variáveis (*variable split*), de modo que a função custo passa a ser escrita como a soma de duas funções de variáveis distintas, \mathbf{x} e \mathbf{f} , relacionadas por uma restrição de igualdade:

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda \|\mathbf{f}\|_1, \\ \text{sujeito a:} \quad & \mathbf{x} - \mathbf{f} = \mathbf{0}. \end{aligned} \quad (50)$$

A solução pode ser iterativamente obtida em três passos, como mostrado no laço do Quadro 6. No passo 5, $\mathbf{I} \in \mathbb{R}^{m \times m}$ é uma matriz identidade e \mathbf{x}_k representa a solução quadrática da k -ésima iteração. O passo 6 usa o mesmo operador de *shrinkage-thresholding* definido na Equação (42) para os algoritmos ISTA, para determinar a solução filtrada \mathbf{f}_k da k -ésima iteração. O passo 7 atualiza a variável dual \mathbf{y} , utilizando o valor obtido na iteração anterior e a diferença entre a solução filtrada e a quadrática da iteração corrente, ponderada pelo parâmetro de penalidade.

4 ESTUDO DE CASO EM ULTRASSONOGRAFIA ESPARSA

Este capítulo apresenta um estudo de caso em ultrassonografia esparsa, que visa avaliar o modelamento proposto no Capítulo 2 por meio dos algoritmos descritos no Capítulo 3. Neste estudo, um total de onze variantes são avaliadas, a saber: IRLS, IRLS-OLS, IRLS-CG, IRLS-CG-OLS, NLCG, NLCG-OLS, FISTA, MFISTA, OMFISTA, OMFISTA-OLS e ADMM.

Dado que alguns desses algoritmos possuem parâmetros a serem ajustados, iniciamos com a descrição dos critérios adotados para escolhê-los. O objetivo é ajustar os parâmetros de modo a permitir que cada variante alcance seu melhor desempenho no estudo comparativo, em termos de tempo de convergência.

Nesse estudo, a avaliação do modelo \mathbf{H} consiste na análise das imagens originadas por dois conjuntos de dados distintos. Primeiramente, geramos dados a partir de uma imagem sintética, em que \mathbf{f} é conhecido e imita objetos imageados (*targets*) em um dispositivo de teste (*phantom*¹) de ultrassom, de uso profissional. Os dados gerados são então contaminados com ruído gaussiano, com uma relação sinal-ruído (*Signal-to-Noise Ratio*, SNR) conhecida, estimada a partir de dados reais. A avaliação desse caso “sintético” é necessária porque possibilita a comparação das imagens reconstruídas com um resultado previamente conhecido, por meio de uma métrica adequada ao imageamento por ultrassom.

Em seguida, esse mesmo modelo discreto \mathbf{H} é avaliado com um conjunto de dados *reais* coletados por um equipamento de ultrassonografia, obtidos ao imagear o *phantom* representado pela imagem sintética. Os resultados são então comparados com a imagem *beamforming* correspondente.

Finalmente, confirmamos a validade do modelamento proposto com um segundo conjunto de dados reais, utilizando outra matriz \mathbf{H} construída de modo similar; os resultados são igualmente comparados com a imagem *beamforming* correspondente.

¹No contexto da ultrassonografia, *phantom* é um objeto que imita as propriedades de um tecido (ou material) em relação à sua resposta ao estímulo acústico. É usado, em geral, para avaliar a qualidade e o comportamento de um sistema de imageamento por ultrassom, por ter propriedades físicas e geométricas bem conhecidas (HEDRICK et al., 1995).

4.1 AVALIAÇÃO DOS PARÂMETROS DOS ALGORITMOS

4.1.1 PARÂMETRO λ – AJUSTE DA ESPARSIDADE

O parâmetro λ da Equação (40) tem um papel importante no equilíbrio entre o termo quadrático $\|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2$ e o termo de regularização $\|\mathbf{f}\|_1$ que promove a esparsidade da solução. Para $\lambda > 0$, quanto mais o incrementamos, maior é o distanciamento entre a solução obtida e a solução de quadrados mínimos, que em geral é ruidosa. Isso pode dar origem a uma solução regularizada mais próxima da desejada, dada a premissa prévia de que a imagem procurada é esparsa.

Um dos métodos para estimar λ é a Curva L, apresentada na Seção 2.3.3. Como o termo de regularização da Equação (40) utiliza a norma ℓ_1 , o método é adaptado para mostrar a curva da norma ℓ_1 da solução estimada $\hat{\mathbf{f}}_\lambda$ em função da norma ℓ_2 do resíduo correspondente, para diversos valores de λ . Cabe ressaltar, entretanto, que a Curva L deve ser usada como ferramenta para obter apenas uma *estimativa inicial* de λ , porque em aplicações tais como a ultrassonografia o usuário final *deve* ser considerado no momento de escolher entre uma imagem mais ou menos ruidosa. Em plataformas reais que usem a abordagem IPB para o imageamento de ultrassom, seria interessante possibilitar o ajuste do parâmetro λ , tal como é feito para outros parâmetros em máquinas convencionais de *beamforming*.

A Figura 8 mostra a forma esperada da Curva L, agora com o eixo das ordenadas utilizando a norma ℓ_1 . Para $\lambda \rightarrow 0$, $\hat{\mathbf{f}}_\lambda$ tende à solução de quadrados mínimos e a imagem tende a ser ruidosa. No outro extremo, o efeito de filtragem torna-se proeminente com o aumento de λ , forçando a esparsidade até o limite em que $\|\hat{\mathbf{f}}_\lambda\|_1 = 0$. Um bom compromisso pode ser encontrado com o valor de λ que corresponde à região central de maior curvatura, que nem sempre é claramente identificável.

Em teoria, o parâmetro λ poderia tender ao infinito, mas acima de um certo limite obtemos $\|\hat{\mathbf{f}}_\lambda\|_1 = 0$. Esse valor máximo é dado por

$$\lambda_{max} = \max (\|\mathbf{H}^T \mathbf{g}\|_i) , \text{ para } 1 \leq i \leq m . \quad (51)$$

Para observar porque esse é um limite superior para λ , deve-se considerar o passo 6 do algoritmo FISTA no Quadro 1. Se a primeira iteração utilizar $\mathbf{y}_k = \mathbf{0}$, o operador resultará num vetor nulo. Nesse estudo, o valor de λ é determinado como uma fração desse limite superior:

$$\lambda = \kappa \lambda_{max} , \text{ com } 0 < \kappa < 1 . \quad (52)$$

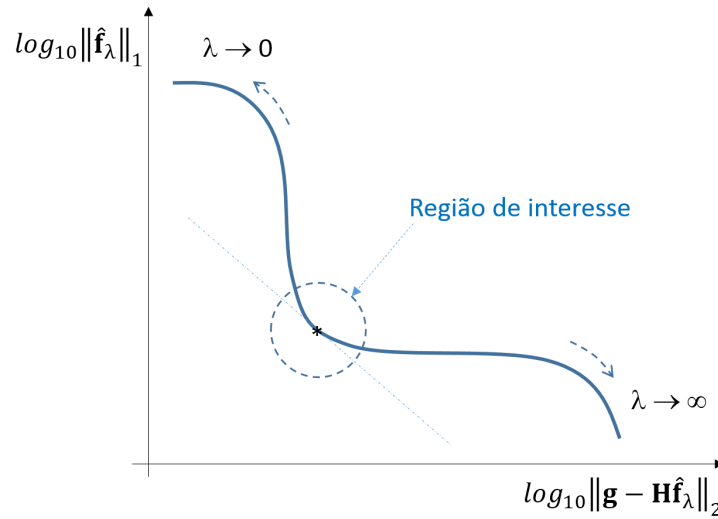


Figura 8: Forma geral da Curva L com as normas ℓ_1 e ℓ_2

Fonte: Adaptada de Hansen (1998)

Nos experimentos apresentados neste capítulo, o modelo de aquisição \mathbf{H} é inicialmente avaliado com dois conjuntos de dados e as Curvas L correspondentes são traçadas para ambos. A Figura 9 mostra a Curva L obtida para a solução da Equação (40) com o conjunto de dados \mathbf{g} sintético, enquanto a Figura 10 mostra a Curva L para o conjunto de dados real. Os pontos em ambas as curvas correspondem às soluções calculadas pelo algoritmo IRLS, com o parâmetro κ variando entre 10^{-6} e 1.

No lugar da escala \log_{10} proposta por Reginska (1996), adotamos escalas distintas para os termos ℓ_1 e ℓ_2 , de modo a melhor evidenciar as curvaturas na região de interesse. Na Figura 9, a região de curvatura é aproximadamente delimitada por $5 \times 10^{-4} \leq \kappa \leq 5 \times 10^{-2}$, enquanto na Figura 10 essa região é delimitada por $5 \times 10^{-4} \leq \kappa \leq 3 \times 10^{-2}$.

Embora as regiões de curvatura sejam claras em ambos os casos, na Seção 4.3 apresentamos uma métrica específica para possibilitar uma escolha mais precisa de κ .

4.1.2 PARÂMETRO β – NLCG

Conforme apontamos na Seção 3.3, há opções distintas para o parâmetro β do algoritmo NLCG. Dado que não temos um critério prévio para determinar a melhor opção para o caso de reconstrução em ultrassonografia, todas as sete opções a seguir foram avaliadas.

1. β_{HZ} proposto por Hager e Zhang (2006);
2. β_{FR} proposto por Fletcher e Reeves (1964);

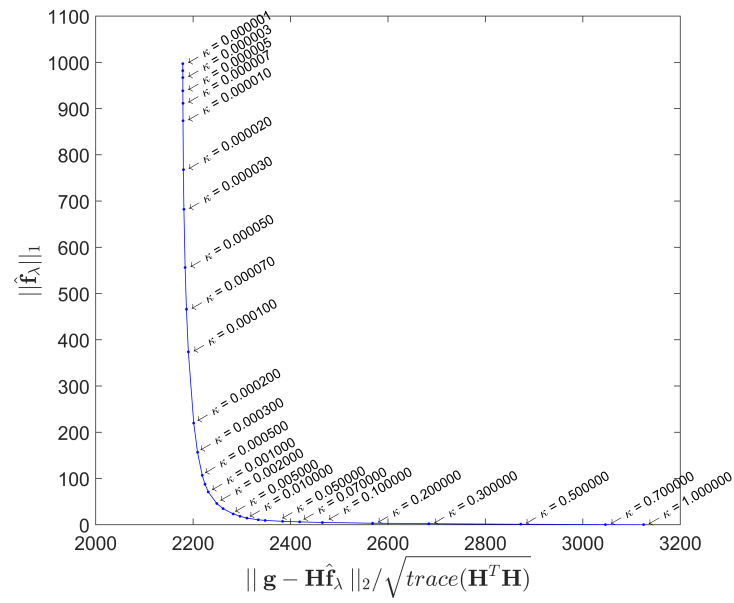


Figura 9: Curva L – Conjunto de dados sintético

Fonte: O autor (2017)

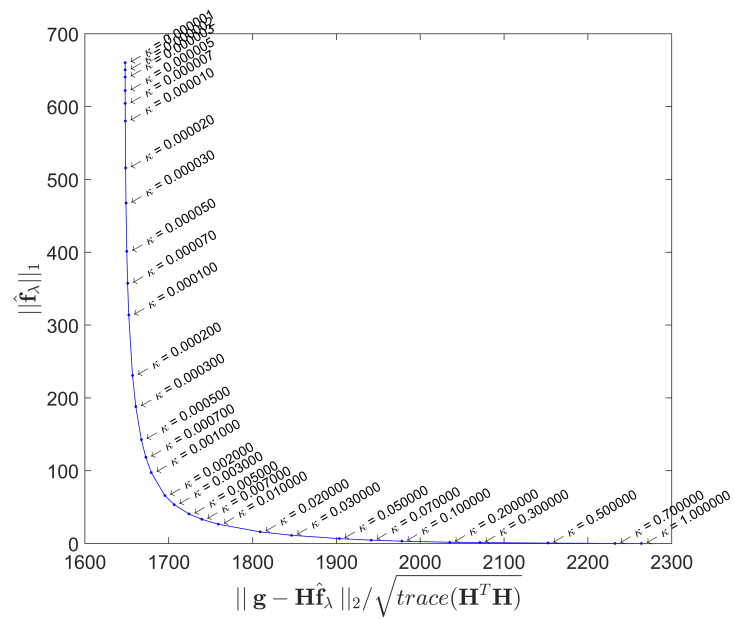


Figura 10: Curva L – Conjunto de dados real

Fonte: O autor (2017)

3. β_{DY} proposto por Dai e Yuan (1999);
4. β_{CD} proposto por Fletcher (1987);
5. β_{PRP} proposto por Polak e Ribière (1969) e Polyak (1969);

6. β_{HS} proposto por Hestenes e Stiefel (1952); e
7. β_{LS} proposto por Liu e Storey (1991).

Os detalhes sobre cada formulação são apresentados por Hager e Zhang (2006). Nas avaliações realizadas pelo autor, o tempo de convergência variou cerca de 3,5 vezes entre as diferentes escolhas para β . Os melhores resultados foram obtidos com β_{HS} e essa formulação foi então adotada neste estudo, conforme apontado no Quadro 4 da Seção 3.3.

4.1.3 PARÂMETROS α E η – OMFISTA

O parâmetro α do algoritmo OMFISTA foi inicialmente avaliado para as reconstruções deste capítulo, com valores $1,0 \leq \alpha \leq 2,0$, com incrementos de 0,1, combinados com valores para o parâmetro η no intervalo $1,0 \leq \eta \leq 2,0$, também com incrementos de 0,1. Os melhores resultados em termos de tempo de convergência foram obtidos para valores próximos aos limites dos intervalos inicialmente estabelecidos, isto é, $\alpha \approx 1,0$ e $\eta \approx 2,0$.

Novos intervalos foram então estabelecidos, com $0,5 \leq \alpha \leq 1,5$, com incrementos de 0,1, combinados com valores de η no intervalo $1,5 \leq \eta \leq 2,5$, também com incrementos de 0,1. A Figura 11 mostra as curvas da função custo da Equação (40) para todas as combinações do segundo experimento; a avaliação cuidadosa dessas curvas mostra que o melhor valor para o primeiro parâmetro é $\alpha = 1,0$. A Figura 12 mostra o detalhe das curvas para $\alpha = 1,0$, mostrando a evolução do custo para as primeiras e para as últimas iterações do experimento. O valor mais adequado para o segundo parâmetro é $\eta = 2,0$.

Os resultados $\alpha = 1,0$ e $\eta = 2,0$ corroboram os valores obtidos para parâmetros similares, observados por Kim e Fessler (2015).

No caso da variante OMFISTA-OLS, o parâmetro α_k é calculado a cada iteração pelo operador de busca em linha proposto por Zibetti et al. (2015b). Outro conjunto de experimentos com $1,0 \leq \eta \leq 2,5$ também revelou o melhor valor deste parâmetro como sendo $\eta = 2,0$.

4.1.4 PARÂMETRO ρ – ADMM

Se observarmos os Quadros 1 e 6, veremos que as escalas do operador de *shrinkage-thresholding* aplicado ao FISTA e ao ADMM são dadas, respectivamente, por $a = \lambda/c$ e $a = \lambda/\rho$. Em função dessa similaridade, os experimentos mostrados neste capítulo utilizam um parâmetro ρ proporcional a $c = \|\mathbf{H}^T \mathbf{H}\|_2$. O valor $\rho = 0,25 c$ é adotado para o ADMM, e todos os algoritmos da família ISTA avaliados usam esse mesmo valor para c .

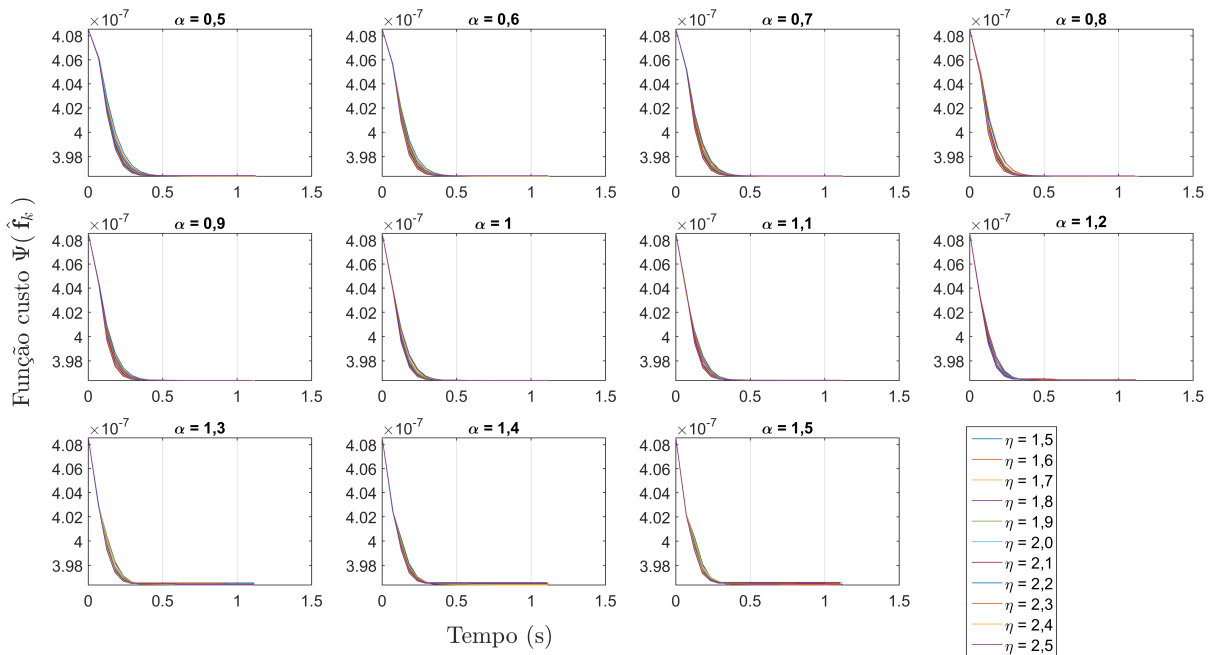


Figura 11: Avaliação dos parâmetros α e η do OMFISTA – geral

Fonte: O autor (2017)

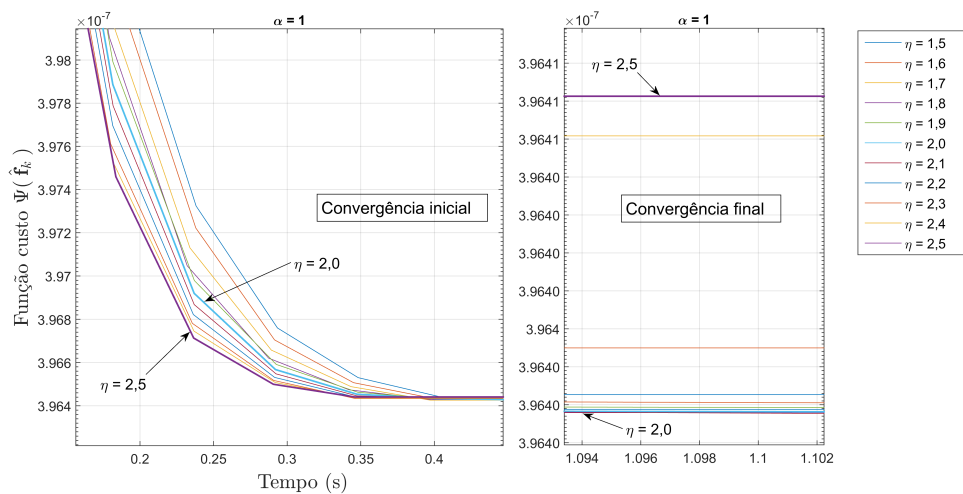


Figura 12: Avaliação dos parâmetros α e η do OMFISTA – detalhe

Fonte: O autor (2017)

Ressaltamos, no entanto, que essa escolha para o parâmetro ρ não otimiza o desempenho do ADMM. O Capítulo 5 propõe a aceleração do método ADMM, e os dois conjuntos de dados reais usados neste capítulo figuram entre os utilizados para validar a nova proposição. No contexto do Capítulo 5, no entanto, a proposta é mostrar que a aceleração de fato leva o ADMM a um desempenho que não pode ser alcançado com o algoritmo em sua forma original. Por isso, uma avaliação mais criteriosa para o parâmetro ρ é realizada e a relação ótima

é estabelecida como $\rho = 0,054 c$. Apesar disso, a escolha aproximada $\rho = 0,25 c$, adotada nos experimentos deste capítulo, possibilitou ao ADMM ser um dos dois métodos mais rápidos.

4.2 DESCRIÇÃO DOS EXPERIMENTOS – MATERIAIS E MÉTODOS

A Tabela 1 resume os principais parâmetros de operação do sistema de ultrassom usado nos experimentos. O equipamento que coleta os dados reais é desenvolvido pela Verasonics², modelo Vantage 128, operando com um transdutor L11-4v (VERASONICS, 2017). Ambos estão disponíveis no laboratório de ensaios de ultrassom LUS/CPGEI/UTFPR. Apesar de o transdutor possuir 128 elementos, apenas os 64 centrais são utilizados neste estudo, de modo a reduzir a quantidade de dados a ser processada.

Tabela 1: Parâmetros principais do sistema de imageamento por ultrassom

Parâmetro	Especificação
Frequência de ultrassom (f_c)	6,25 MHz
Frequência de amostragem (f_s)	25 MHz
Velocidade do som (v)	1.540 m/s
Taxa de atenuação (parênquima hepático)	0,5 dB/cm/MHz
Comprimento de onda (λ)	0,2464 mm
Número de elementos do transdutor (<i>array</i>)	128
Largura de banda do transdutor (B)	3,84 MHz
Largura de banda fracionária (BW)	61,44%

Fonte: (VERASONICS, 2017) e (FLUKE-BIOMEDICAL, 2005)

O dispositivo (*phantom*) de referência é designado *Multipurpose Tissue/Cyst Ultrasound Phantom*, modelo 84-317, desenvolvido pela Fluke Biomedical³, também disponível no laboratório LUS. Esse *phantom* contém diversos grupos de alvos (*targets*) com espaçamentos precisos, imersos em um meio que exibe respostas ultrassônicas similares às encontradas no parênquima hepático humano, incluindo a mesma atenuação, velocidade de propagação e características de espalhamento (*scattering*) (FLUKE-BIOMEDICAL, 2005).

A Figura 13 mostra o equipamento de ultrassom em (a) e o dispositivo *phantom* em (b), com a identificação da porção imageada no dispositivo real e na sua representação esquemática.

²Verasonics Inc. Kirkland, EUA. Detalhes em <http://verasonics.com/vantage-systems/>

³Fluke Corporation, Cleveland, EUA. Detalhes em <http://www.flukebiomedical.com/rms/>

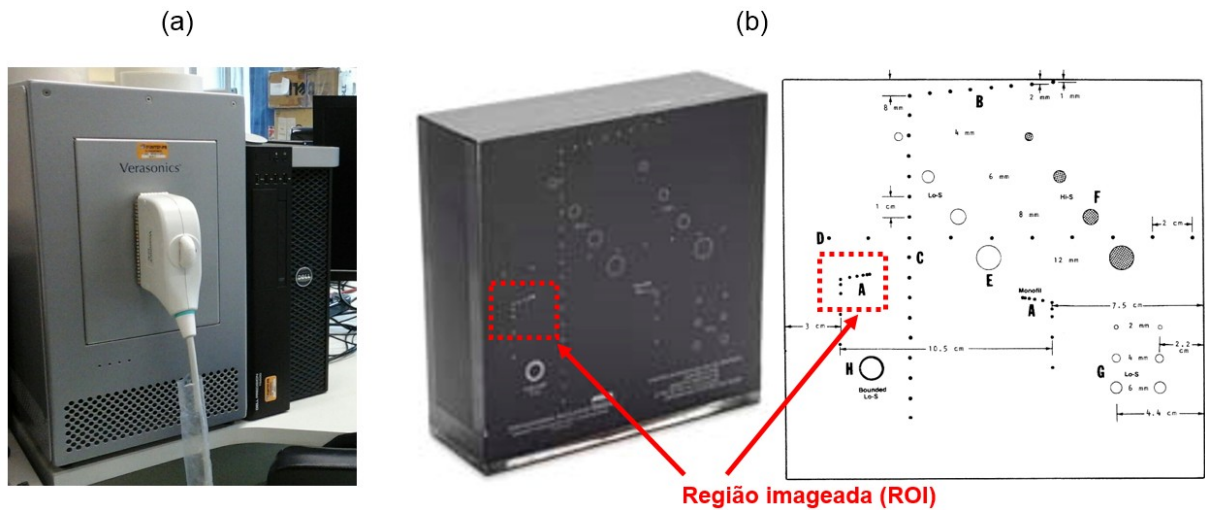


Figura 13: Equipamentos de ultrassom. (a) Verasonics Vantage 128 e transdutor L11-4v. (b) Phantom Fluke 84-317.

Fonte: (VERASONICS, 2017) e (FLUKE-BIOMEDICAL, 2005)

4.2.1 CONJUNTO DE DADOS SINTÉTICO

No caso que designamos de “sintético”, uma imagem bidimensional de referência é construída de modo a reproduzir com precisão os tamanhos e posições de oito pequenos *targets* dentro do *phantom* utilizado.

A Figura 14 mostra uma vista esquemática dos *targets* escolhidos, enumerados de 1 a 8. Todas as dimensões, tamanhos e posições relativas são exatamente as descritas em Fluke-Biomedical (2005). A ROI imageada é uma região quadrada de $19,96 \text{ mm} \times 19,96 \text{ mm}$, com $81 \times 81 \text{ pixels}$, e uma resolução espacial de $0,2464 \text{ mm/pixel}$ nos eixos x e z .

Dado que os *targets* reais são finas hastes de *nylon* com $0,24 \text{ mm}$ de diâmetro, e dadas as posições indicadas, nenhum deles encontra-se inserido completamente na região de um único *pixel*. Uma visão ampliada dos oito *targets* em relação à grade é mostrada na Figura 15. As porcentagens mostradas representam as frações das áreas de cada *target* dentro dos quadrantes indicados.

A imagem de referência é construída com a atribuição dos valores dessas porcentagens, no intervalo $[0, 1]$, aos *pixels* correspondentes. O *Target* 1, por exemplo, é representado por 4 *pixels*, com os valores $0,1686$ ($16,86\%$) e $0,1392$ ($13,92\%$) na linha 16, e com $0,3738$ ($37,38\%$) e $0,3183$ ($31,83\%$) na linha 17, conforme ilustra o detalhe apresentado na Figura 16. Após atribuir os valores dos *pixels* para os oito *targets* correspondentes, o vetor \mathbf{f}_{sint} é obtido pela simples concatenação das colunas da imagem bidimensional.

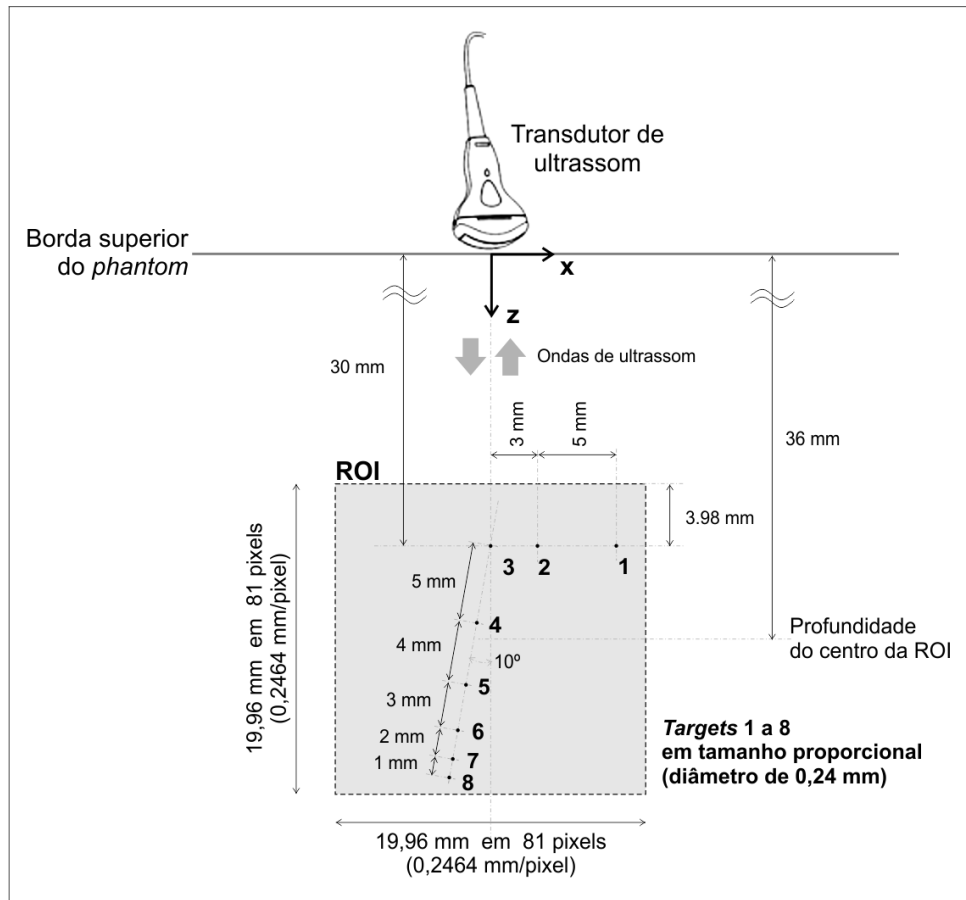


Figura 14: Vista esquemática dos *targets* dentro do *phantom*

Fonte: O autor (2017)

Uma vez obtido o vetor \mathbf{f}_{sint} , o conjunto de dados sintético é calculado por $\mathbf{g}_{sint} = \mathbf{H}\mathbf{f}_{sint} + \mathbf{e}$. O vetor de ruído gaussiano \mathbf{e} é calculado de modo a produzir uma SNR similar àquela encontrada no conjunto de dados real, que é estimada por

$$SNR = 10 \log_{10} \left(\frac{\|\mathbf{H}\mathbf{f}_{sint}\|_2^2}{\|\mathbf{g}_{real}\|_2^2 - \|\mathbf{H}\mathbf{f}_{sint}\|_2^2} \right), \quad (53)$$

em que \mathbf{g}_{real} é o conjunto de dados real coletado pelo equipamento de ultrassom. A SNR estimada por essa abordagem é de -0.94 dB. O conjunto \mathbf{g}_{sint} completo é mostrado na Figura 17, em que cada uma das 64 linhas verticais representa os dados de um elemento do transdutor, com o eixo do tempo disposto verticalmente, e com os instantes t_i sendo incrementados de cima para baixo.

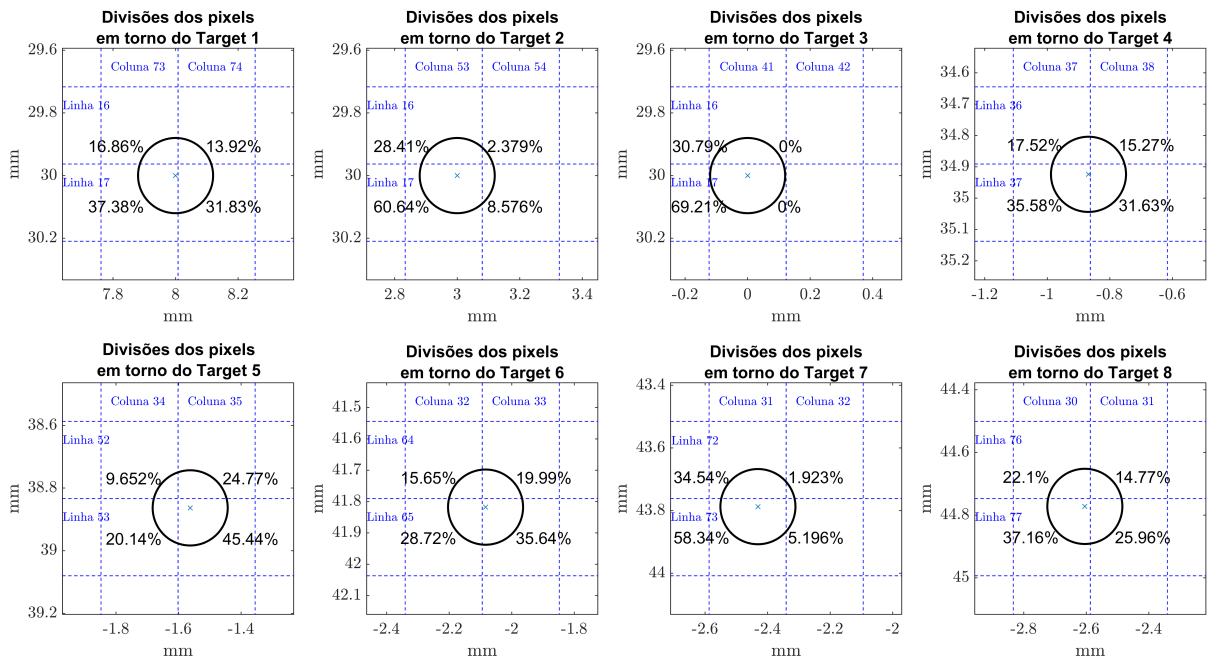


Figura 15: Vista ampliada dos *targets* na grade de *pixels* da imagem de referência

Fonte: O autor (2017)

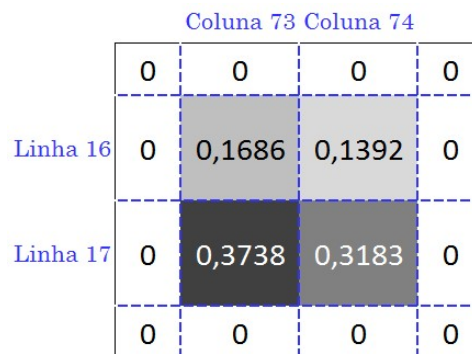


Figura 16: Atribuição de valores aos *pixels* da imagem de referência para o *Target 1*

Fonte: O autor (2017)

4.2.2 CONJUNTO DE DADOS REAL

Designado de \mathbf{g}_{real} , o conjunto de dados real é coletado pelo equipamento de ultrassom Verasonics ao imagear o dispositivo phantom Fluke na mesma ROI indicada na Figura 14, tomando a média de 10 medições subsequentes. O conjunto \mathbf{g}_{real} completo é mostrado na Figura 18, em que cada uma das 64 linhas verticais representa os dados de um elemento do transdutor, com o eixo do tempo disposto verticalmente, e com os instantes t_i sendo incrementados de cima para baixo.

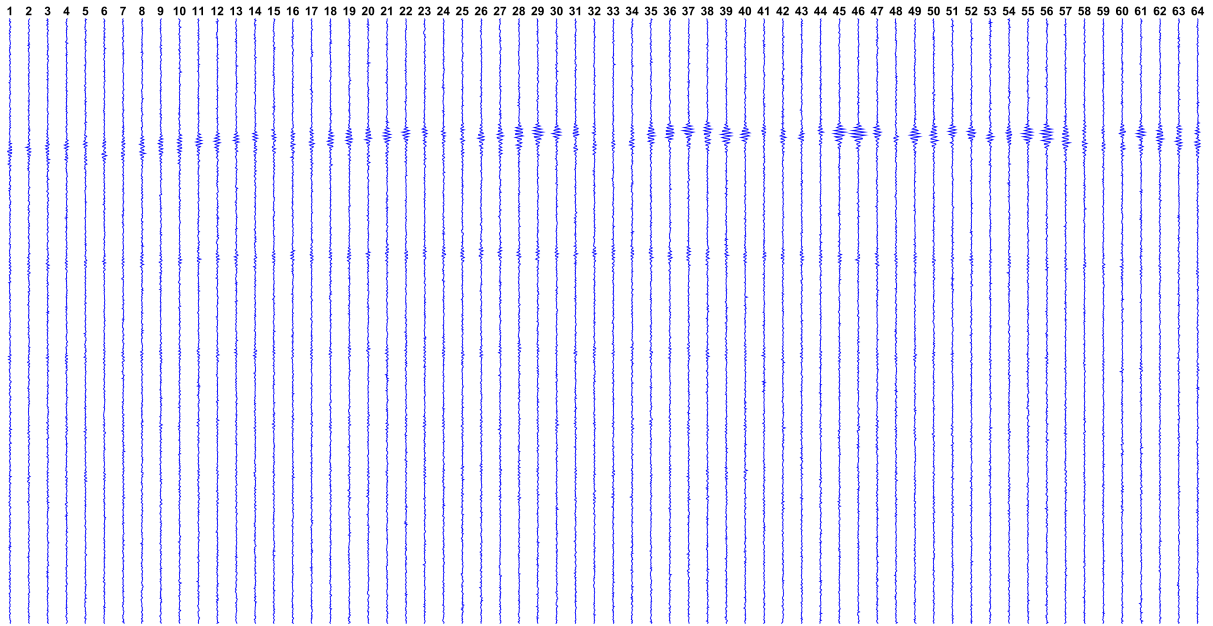


Figura 17: Conjunto de dados sintético $g_{sint} = Hf_{sint} + e$
Fonte: O autor (2017)

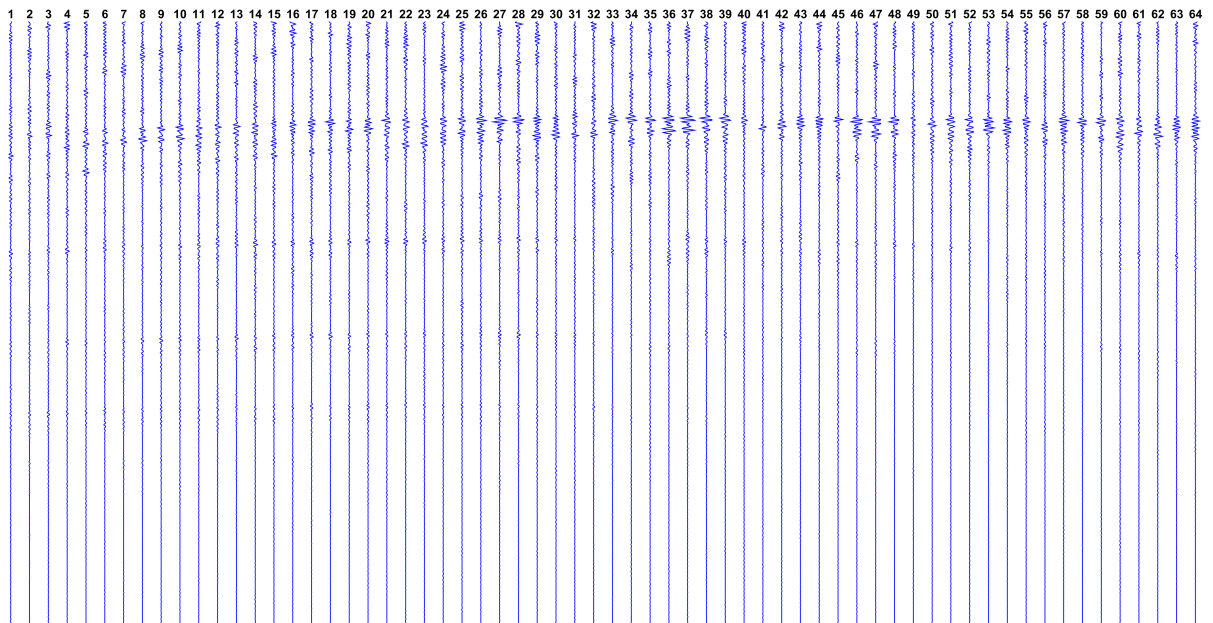


Figura 18: Conjunto de dados real g_{real}
Fonte: O autor (2017)

4.3 RESULTADOS DOS EXPERIMENTOS

Os resultados para o caso sintético e para o caso real são apresentados em quatro visões distintas. O mesmo panorama é apresentado para um segundo conjunto de dados real, com

vistas a confirmar a validade do modelamento.

Na primeira visão, as *imagens reconstruídas* são apresentadas para os onze algoritmos considerados. Uma métrica chamada *Array Performance Indicator* (API) proposta por Holmes et al. (2005) é usada para comparar as imagens. A API é uma métrica simples, que mede o desempenho de um método de reconstrução em termos de sua habilidade em imagear refletores pontuais, como os usados neste estudo. A API é adimensional e mede o tamanho da *Função de Espalhamento de Ponto* (*Point Spread Function*, PSF). Ela é definida como a razão entre a área (A_{-6dB}) dentro da qual a PSF está entre o valor de pico e -6 dB (meia amplitude) e o quadrado do comprimento de onda Λ do ultrassom,

$$API = \frac{A_{-6dB}}{\Lambda^2}. \quad (54)$$

Valores menores da API indicam que o método tem maior habilidade de reconstruir com precisão imagens que representam *targets* pontuais. Valores maiores da API significam que a PSF ocupa uma área maior na imagem, indicando menor habilidade de imagear objetos pontuais.

A segunda visão mostra as curvas de evolução dos *erros de estimação*, calculados como a distância de norma ℓ_2 entre a solução estimada $\hat{\mathbf{f}}_k$ na iteração k e uma *solução de referência*. Para o caso sintético, a referência é a imagem \mathbf{f}_{sint} construída com os oito *targets*, conforme descrito na Seção 4.2. Para o caso real, dado que a imagem final é desconhecida, adotamos como solução de referência \mathbf{f}^* aquela obtida pelo método IRLS após 500 iterações.

A terceira visão apresenta as curvas da *função custo* $\Psi(\mathbf{f})$, traçadas em função do *número da iteração* k . Essas curvas nos permitem observar a eficiência de minimização de cada algoritmo, deixando de lado o tempo requerido. Na última visão, traçamos as curvas de $\Psi(\mathbf{f})$ em função do *tempo*, para observar o custo computacional de cada método.

Antes de prosseguir, entretanto, devemos especificar como o valor de κ foi escolhido, uma vez que ele determina o parâmetro de regularização λ na Equação (52). Dado que o conjunto de dados sintético é obtido a partir da imagem de referência \mathbf{f}_{sint} que representa com precisão os *targets* do *phantom* real, podemos calcular a API para essa imagem, API_{ref} , e compará-la com os valores API_κ das imagens obtidas com valores de κ na região de curvatura da Figura 9.

A razão API_κ / API_{ref} é traçada em função de κ na Figura 19, usando a faixa $5 \times 10^{-4} \leq \kappa \leq 5 \times 10^{-2}$ apontada na Figura 9; as soluções para cada κ foram calculadas pelo método IRLS com 30 iterações. Está claro que além de certo ponto ($\kappa = 0,01$), a razão cai rapidamente,

indicando que a reconstrução forçou demais a esparsidade da imagem. Razões muito pequenas indicam a supressão de alguns *pixels* importantes que carregam informação sobre os *targets*, devida a valores excessivamente elevados para o parâmetro de regularização λ . Essa estimativa foi igualmente confirmada com o conjunto de dados real, mostrando que $\kappa = 0,01$ é o melhor compromisso em ambos os casos. Portanto, todos os resultados apresentados a seguir usam $\lambda = 0,01 \lambda_{max}$.

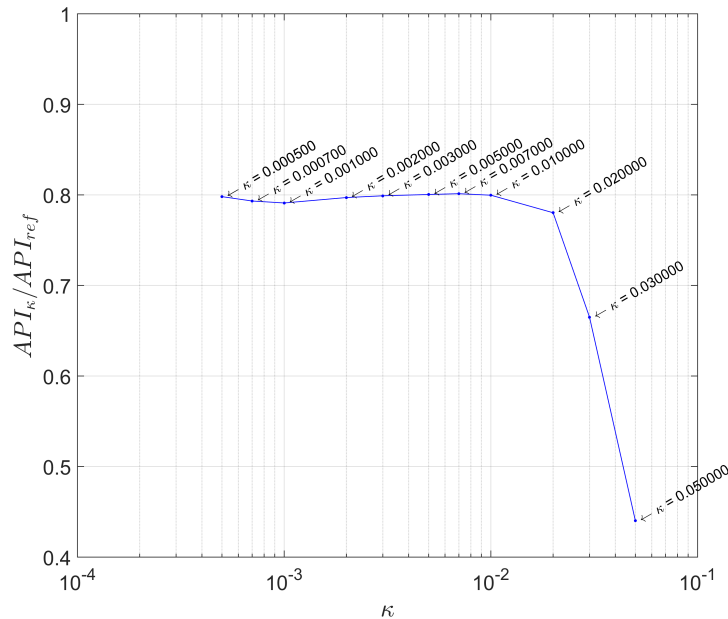


Figura 19: Curva de razões da métrica API (Array Performance Indicator)

Fonte: O autor (2017)

4.3.1 RESULTADOS PARA O CONJUNTO DE DADOS SINTÉTICO

As imagens reconstruídas a partir do conjunto de dados sintético são mostradas na Figura 20 e os valores API correspondentes estão na Tabela 2.

Visualmente, todas as imagens são muito similares, mas os valores e razões da métrica API apontam algumas diferenças. Todos os métodos da família IRLS mostram resultados similares sob a ótica da API, assim como os métodos NLCG, que possuem valores API ligeiramente maiores porque suas imagens são levemente mais borradas que as demais. Dentro do grupo ISTA, vemos uma distinção clara: enquanto FISTA e MFISTA mostram resultados API idênticos, OMFISTA e OMFISTA-OLS resultaram em valores mais distantes da referência. Esses dois métodos possuem um parâmetro adicional a ser ajustado (η_k), que pode adotar uma regra empírica de atualização ao longo das iterações. Neste estudo esse parâmetro é fixo, e

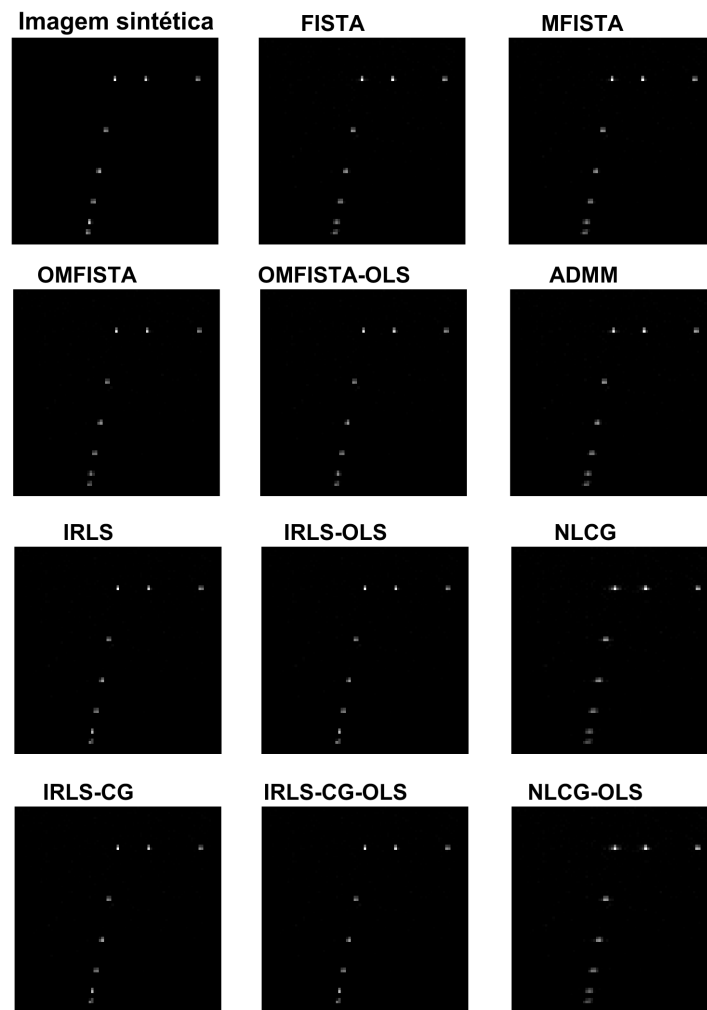


Figura 20: Imagens reconstruídas a partir de conjunto de dados sintético

Fonte: O autor (2017)

uma investigação posterior poderia revelar uma abordagem mais vantajosa. Apesar disso, o OMFISTA-OLS mostrou um resultado API mais perto da referência, em parte porque o seu passo não é fixo, mas calculado de modo ótimo a cada iteração. Juntamente com FISTA e MFISTA, o método ADMM foi um dos que resultaram na melhor solução pela métrica API, dado que esses três algoritmos possuem as razões API/API_{ref} mais próximas da unidade.

As curvas de evolução dos erros de estimação são traçadas na Figura 21 para as primeiras 30 iterações. Note-se que as curvas das variantes IRLS e IRLS-OLS estão visualmente sobrepostas às curvas das variantes IRLS-CG e IRLS-CG-OLS, respectivamente, por terem apresentado a mesma evolução; ocorre o mesmo entre os métodos FISTA e MFISTA. As curvas mostram que todos os métodos reduzem gradualmente os erros de estimação, com clara vantagem para as variantes do método IRLS em termos de velocidade. Embora não mostrado na figura, a extensão das curvas para 100 iterações mostrou que todos os métodos

Tabela 2: Valores da API para imagens reconstruídas com 30 iterações – Caso sintético

Imagem de referência	API _{ref}	
Imagem sintética do <i>phantom</i>	5,4813	
Método de reconstrução	API	API/API _{ref}
IRLS	4,4000	0,80
IRLS-OLS	4,4029	0,80
IRLS-CG	4,4000	0,80
IRLS-CG-OLS	4,4029	0,80
NLCG	6,1278	1,12
NLCG-OLS	6,1292	1,12
FISTA	5,7841	1,06
MFISTA	5,7841	1,06
OMFISTA	3,4132	0,62
OMFISTA-OLS	4,0858	0,75
ADMM	5,8548	1,07

Fonte: O autor (2017)

convergem aproximadamente para o mesmo erro; entretanto, nenhum dos erros igualou-se a zero. Isso significa que todos os algoritmos atingem a mesma *região* de mínimo no problema de minimização, mas não exatamente a imagem de referência. Isso provavelmente se deve ao alto nível de ruído introduzido no conjunto de dados sintético e, possivelmente, à convexidade não estrita do problema.

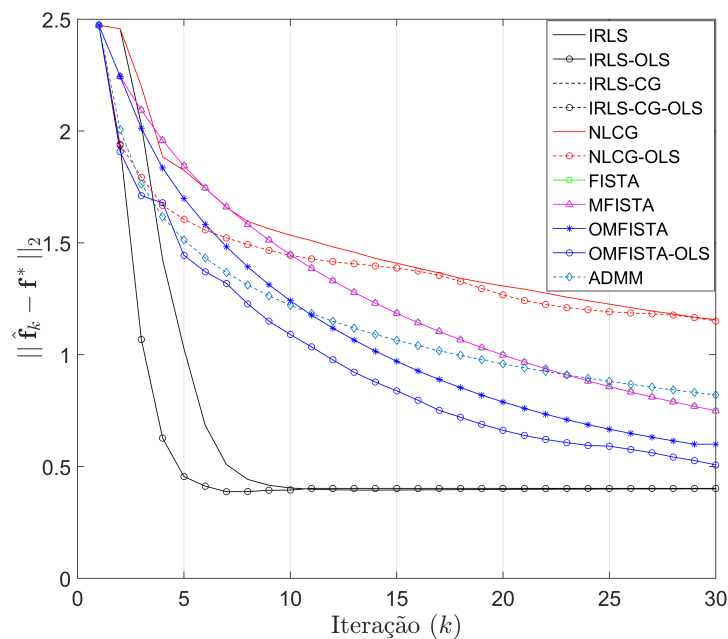


Figura 21: Erros de estimação $\|\hat{\mathbf{f}}_k - \mathbf{f}^*\|_2 \times$ iterações – Caso sintético

Fonte: O autor (2017)

As curvas de custo são traçadas em função do número da iteração k e em função do tempo nas Figuras 22 e 23, respectivamente. A primeira indica que todos os métodos tendem a convergir para o mesmo nível de custo, mas necessitam de um número diferente de iterações para alcançar o mínimo. A segunda é apresentada com uma visão ampliada dos primeiros instantes de cálculo. Todas as medidas de tempo referem-se a uma execução de cada algoritmo, rodando em um processador Intel® Core™ i7-4790⁴ a 3,60 GHz, com 32 GB DDR3 RAM, em Microsoft Windows 7 Professional[©] de 64 bits, rodando em ambiente MATLAB^{®5}.

Vemos claramente que alguns métodos são muito mais eficientes que outros em termos do tempo requerido. A aceleração promovida pelo procedimento de busca em linha para os algoritmos OLS torna-se evidente quando comparados com seus correspondentes não-OLS. A eficiência em tempo também é evidenciada quando os métodos são agrupados em suas famílias. As variantes do NLCG e do IRLS são as menos eficientes. Os métodos ISTA são adequados a este tipo de problema, com uma clara vantagem para o OMFISTA-OLS, que apresenta quase a mesma velocidade de convergência do método ADMM.

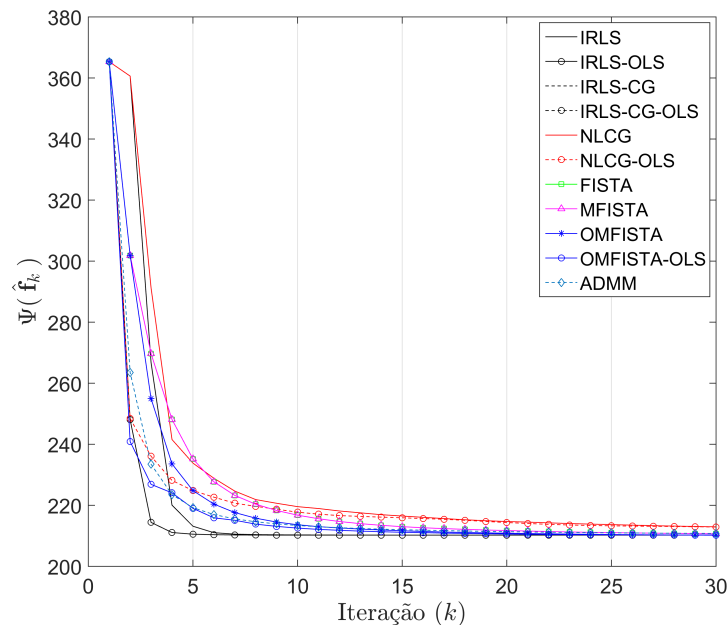


Figura 22: Função custo $\Psi(\hat{f}_k) \times$ iterações – Caso sintético

Fonte: O autor (2017)

⁴Intel Corporation, Santa Clara, EUA.

⁵The MathWorks Inc., Natick, EUA.

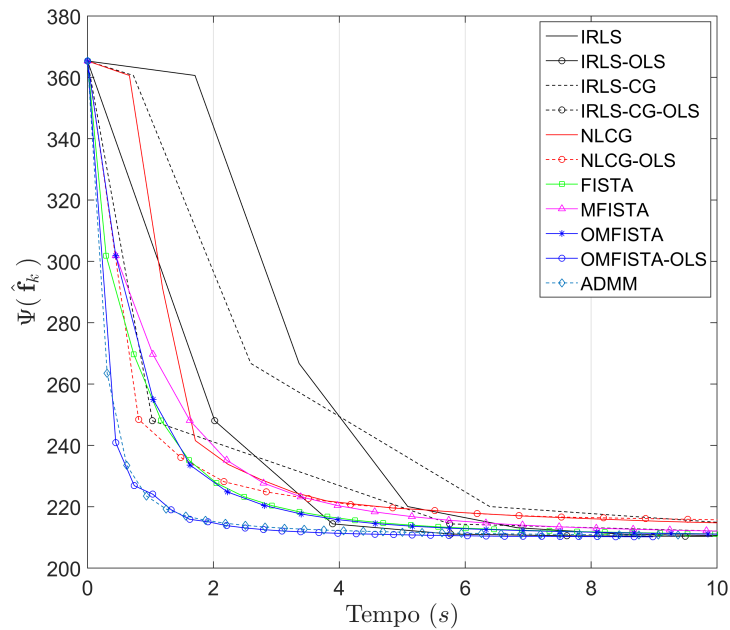


Figura 23: Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ tempo – Caso sintético
Fonte: O autor (2017)

4.3.2 RESULTADOS PARA O CONJUNTO DE DADOS REAL

Enquanto o conjunto de dados sintético é bem conhecido e controlado, precisamos destacar algumas particularidades do conjunto de dados real. Além dos erros de amostragem e de outros sinais de ruído inerentes aos dados amostrados de transdutores reais, destacamos duas outras fontes de variação entre os casos sintético e real.

Em primeiro lugar, todo dispositivo real apresenta dimensões e propriedades dos materiais dentro de uma certa faixa de variação. Por exemplo, a velocidade real do som pode variar até ± 6 m/s no material que preenche internamente o *phantom* utilizado, e as posições e diâmetros declarados para os *targets* podem variar em até $\pm 0,10$ mm e $\pm 5\%$, respectivamente (FLUKE-BIOMEDICAL, 2005).

Em segundo lugar, a distância, a inclinação e a translação do transdutor de ultrassom em relação aos *targets* podem ser ligeiramente diferentes daquelas indicadas na Figura 14, devido ao uso de gel acoplador, e devido à inabilidade e/ou imprecisão do operador durante a coleta de dados com o transdutor.

As imagens reconstruídas obtidas a partir do conjunto de dados real são mostradas na Figura 24, e os valores API correspondentes estão na Tabela 3. Novamente, todas as imagens são visualmente muito similares, e os valores e razões da métrica API são usados para destacar as diferenças.

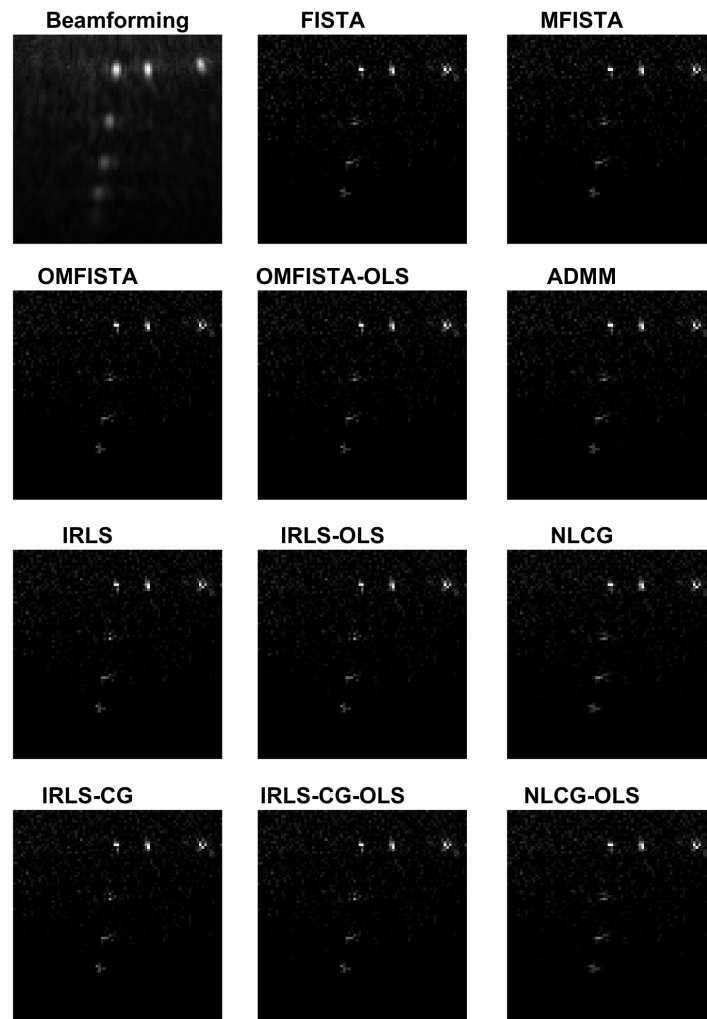


Figura 24: Imagens reconstruídas a partir do conjunto de dados real

Fonte: O autor (2017)

Sob a ótica da métrica API, as conclusões para as famílias IRLS, NLCG e ISTA são em geral as mesmas descritas para o caso sintético, exceto pelo fato de que as razões API/API_{ref} têm todas, aqui, resultados próximos da unidade. A exceção clara refere-se ao API da imagem do *beamforming*, que evidencia a capacidade limitada desse método tradicional de representar *targets* com diâmetros reduzidos em proporção real na imagem. Esse resultado suporta nossa afirmação de que métodos iterativos podem entregar resultados mais precisos se o devido cuidado for tomado na construção de um modelo adequado de aquisição.

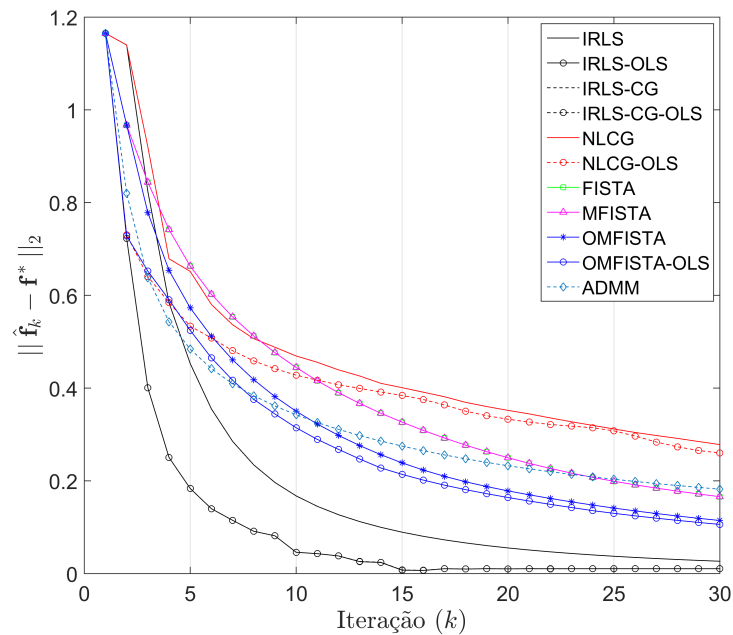
As curvas de evolução dos erros de estimação são mostradas na Figura 25 para as primeiras 30 iterações. As curvas indicam que todos os algoritmos tendem a convergir para o mesmo nível de erro, embora com velocidades distintas. A observação das curvas estendidas mostrou que todos os erros tendem a zero em aproximadamente 100 iterações, indicando que

Tabela 3: Valores da API para imagens reconstruídas com 30 iterações – Caso real

Imagem de referência	API _{ref}	
IRLS após 500 iterações	7,2347	
Método de reconstrução	API	API/API _{ref}
Beamforming	20,8626	2,88
IRLS	7,2194	1,00
IRLS-OLS	7,2307	1,00
IRLS-CG	7,2194	1,00
IRLS-CG-OLS	7,2303	1,00
NLCG	6,9467	0,96
NLCG-OLS	6,9748	0,96
FISTA	7,4016	1,02
MFISTA	7,4016	1,02
OMFISTA	6,5906	0,91
OMFISTA-OLS	6,5703	0,91
ADMM	6,9921	0,97

Fonte: O autor (2017)

os métodos avaliados alcançam a solução de referência \mathbf{f}^* .



As curvas de custo são traçadas em função das iterações e em função do tempo nas Figuras 26 e 27, respectivamente. Como para o caso sintético, a primeira indica que todos os métodos convergem para o mesmo nível de custo, mas com diferentes números de iterações,

enquanto a segunda destaca distintos custos computacionais; os desempenhos de cada grupo de algoritmos e de suas variantes são muito semelhantes àqueles obtidos para o caso sintético, com clara vantagem para OMFISTA-OLS e ADMM. Todas as medidas de tempo foram realizadas no mesmo ambiente computacional descrito anteriormente.

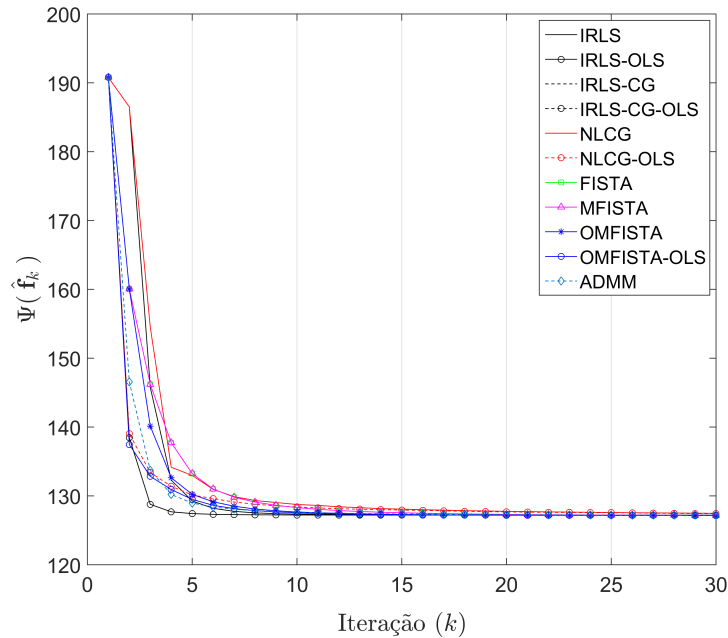


Figura 26: Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ iterações – Caso real

Fonte: O autor (2017)

4.3.3 RESULTADOS PARA OUTRO CONJUNTO DE DADOS REAL

Com vistas a verificar a validade do modelamento proposto, os algoritmos são agora usados para reconstruir as imagens a partir de um novo conjunto de dados real. Os dados são obtidos pelo mesmo equipamento de ultrassom, mas neste caso temos 4 *targets* em uma região distinta do mesmo *phantom* previamente utilizado. A matriz \mathbf{H} é construída para a ROI apropriada, que engloba os novos *targets*, e consiste em uma região retangular de 17,99 mm \times 21,93 mm, com 73 \times 89 *pixels*, e resolução espacial de 0,2464 mm/*pixel* em ambos os eixos x e z . Nenhuma informação sobre a imagem “real” está disponível, exceto a natureza dos *targets*, que são finas hastes de *nylon* com 0,24 mm de diâmetro, como no experimento anterior.

Em termos de parâmetros, os mesmos valores para λ , β , α e η são usados. O parâmetro $c = \|\mathbf{H}^T \mathbf{H}\|_2$ é calculado para o novo modelo \mathbf{H} , e $\rho = 0,25c$, como antes.

As imagens reconstruídas por todos os métodos são mostradas na Figura 28 e os valores API correspondentes estão na Tabela 4. Novamente, todas as imagens são muito similares, e os

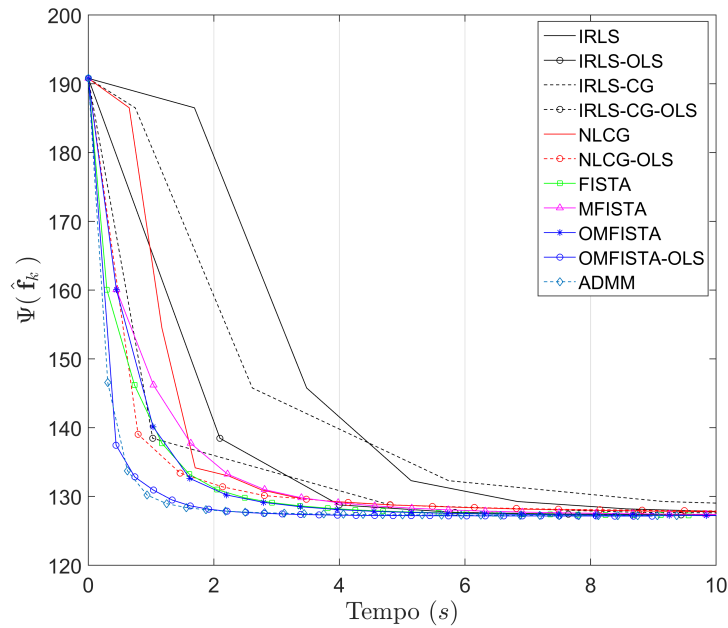


Figura 27: Função custo $\Psi(\hat{f}_k) \times \text{tempo}$ – Caso real

Fonte: O autor (2017)

valores e razões da métrica API destacam as diferenças entre elas.

Do ponto de vista da métrica API, notamos que todos os algoritmos do grupo IRLS atingem resultados muito similares. A família ISTA também é consistente, novamente com o OMFISTA-OLS atingindo um valor API inferior. Os métodos do grupo NLCG e o ADMM resultaram em imagens ligeiramente menos nítidas, o que explica as razões API/API_{ref} maiores. Novamente, observamos que o método *beamforming* mostrou o resultado API mais pobre.

As curvas dos erros de estimação são mostradas na Figura 29. Observamos que as variantes do NLCG, em particular, convergem mais lentamente para a solução de referência do IRLS. De modo geral, o comportamento relativo em cada família de métodos é o mesmo do experimento anterior.

Finalmente, as curvas de custo são mostradas nas Figuras 30 e 31, traçadas em função das iterações e em função do tempo, como antes. Como nos casos anteriores, a primeira indica que todos os métodos convergem para o mesmo nível de custo, mas agora as diferenças nas velocidades de convergência são mais evidentes. A segunda figura destaca os diferentes custos computacionais, com medições de tempo realizadas como nos casos anteriores.

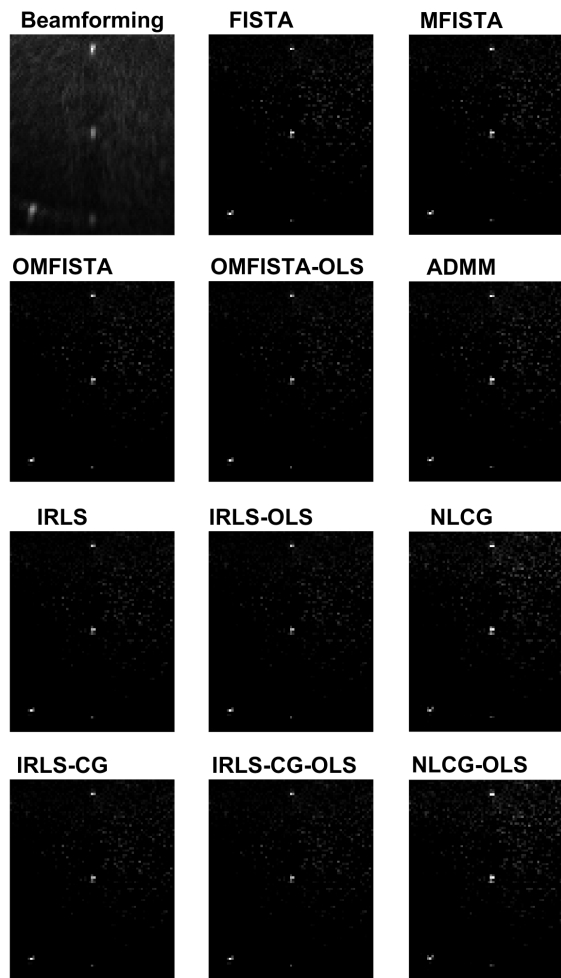


Figura 28: Imagens reconstruídas a partir de outro conjunto de dados real, com outra ROI

Fonte: O autor (2017)

4.3.4 SOBRE O CRITÉRIO DE PARADA

Devemos destacar a razão pela qual escolhemos um número fixo de $k = 30$ iterações como critério de parada para os algoritmos iterativos.

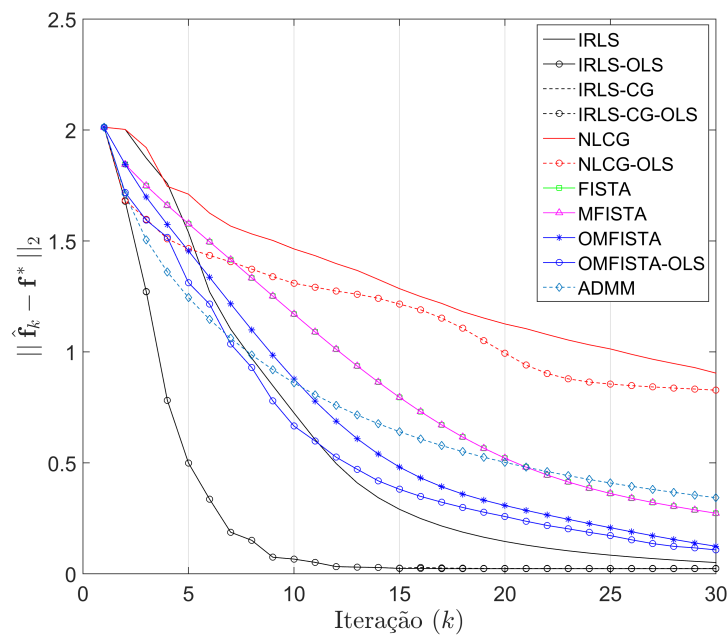
Ao observarmos as curvas do erro de estimação $\|\hat{\mathbf{f}}_k - \mathbf{f}^*\|_2$ para o IRLS nas Figuras 25 e 29, notamos que seus valores finais representam aproximadamente 2,3% e 2,5% dos erros iniciais, respectivamente. Essas baixas porcentagens explicam o por quê de as imagens obtidas pelo IRLS após 30 iterações serem visualmente tão próximas das imagens de referência \mathbf{f}^* alcançadas por esse mesmo método após 500 iterações, como mostram as Figuras 32 e 33.

Na prática, entretanto, os algoritmos iterativos são usualmente utilizados com um critério de parada definido em termos da *variação do custo*, uma vez que em reconstrução de imagens raramente dispomos de uma imagem de referência com a qual possamos calcular o

Tabela 4: Valores da API para imagens reconstruídas com 30 iterações – Outro caso real

Imagem de referência	API _{ref}	
IRLS após 500 iterações	3,7327	
Método de reconstrução	API	API/API _{ref}
Beamforming	5,3886	1,44
IRLS	3,7378	1,00
IRLS-OLS	3,7312	1,00
IRLS-CG	3,7378	1,00
IRLS-CG-OLS	3,7298	1,00
NLCG	4,8393	1,30
NLCG-OLS	4,6477	1,25
FISTA	4,3360	1,16
MFISTA	4,3360	1,16
OMFISTA	3,8639	1,04
OMFISTA-OLS	3,1716	0,85
ADMM	4,6643	1,25

Fonte: O autor (2017)

Figura 29: Erros de estimação $\|\hat{\mathbf{f}}_k - \mathbf{f}^*\|_2 \times \text{iterações}$ – Outro caso real

Fonte: O autor (2017)

erro de estimação. O algoritmo em geral é encerrado quando a função custo não decresce além de certa quantia entre duas iterações consecutivas. Nas Figuras 26 e 30 vemos essa tendência para $k > 25$, e portanto escolhemos $k = 30$ como critério de parada para todos os métodos neste estudo.

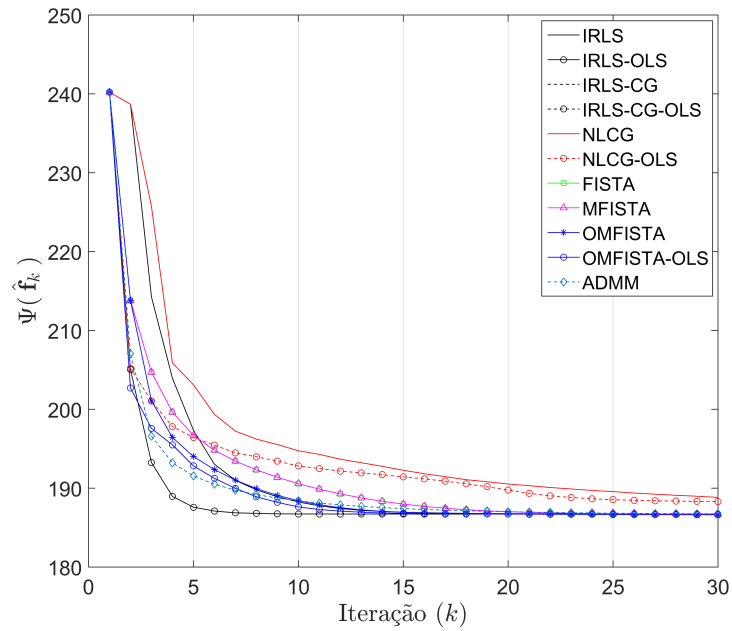


Figura 30: Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ iterações – Outro caso real
Fonte: O autor (2017)

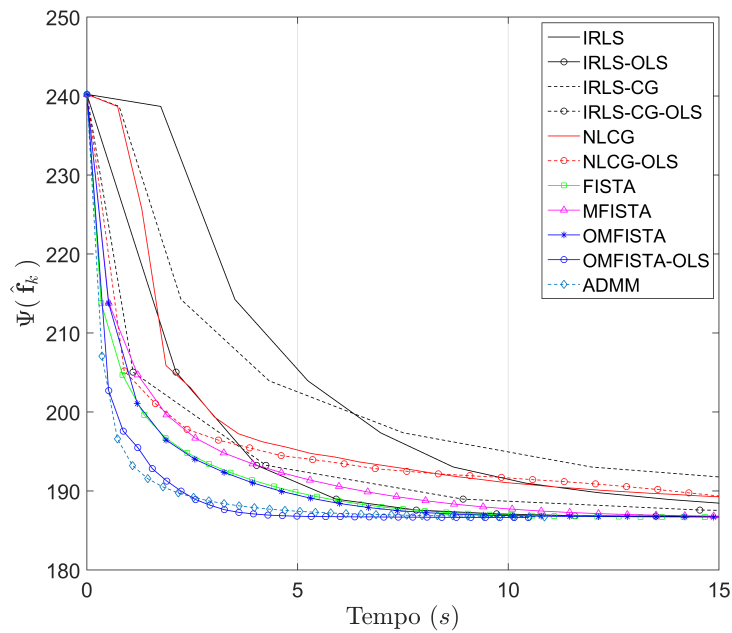


Figura 31: Função custo $\Psi(\hat{\mathbf{f}}_k) \times$ tempo – Outro caso real
Fonte: O autor (2017)

4.4 CONCLUSÕES SOBRE OS EXPERIMENTOS

Embora limitado, o modelo de aquisição proposto mostrou resultados satisfatórios, tanto visualmente quanto pela métrica API. A estratégia de iniciar com a elaboração de um

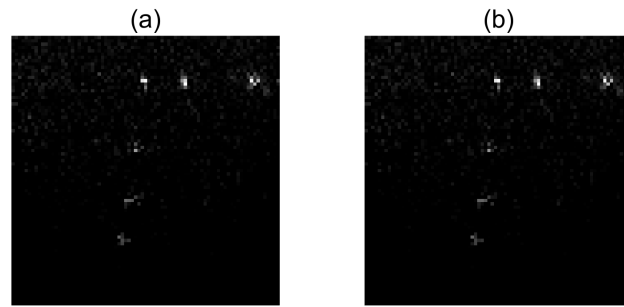


Figura 32: Conjunto de dados real. (a) Imagem de referência f^* , IRLS 500 iterações. (b) Imagem \hat{f}_k , IRLS 30 iterações. API na Tabela 3.

Fonte: O autor (2017)

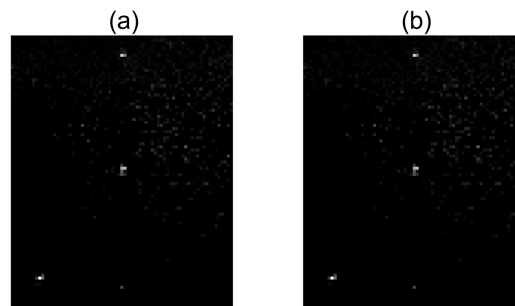


Figura 33: Outro caso real. (a) Imagem de referência f^* , IRLS 500 iterações. (b) Imagem \hat{f}_k , IRLS 30 iterações. API na Tabela 4.

Fonte: O autor (2017)

conjunto de dados sintético a partir de uma imagem de referência que imita o *phantom* real, nos permitiu ajustar apropriadamente alguns parâmetros em uma situação controlada, antes de avaliar o modelo com os dados reais coletados. Um conjunto diferente de dados reais, coletados de *targets* distintos em outra ROI, mostrou que o modelamento é aceitável, dado que todos os métodos alcançaram bons resultados.

Conforme havíamos suposto, embora um alto esforço computacional seja requerido pelo métodos iterativos, a reconstrução esparsa pela abordagem IPB pode conduzir a imagens mais representativas dos objetos imageados (*ground-truth*) do que o método tradicional de *beamforming*.

Entretanto, ainda que todos os algoritmos considerados tenham sido capazes de mostrar boas imagens reconstruídas, os desempenhos computacionais variaram consideravelmente, e a capacidade de processamento das plataformas atuais deve ainda evoluir antes que as taxas de quadros atualmente disponíveis nos equipamentos de ultrassom sejam atingidas.

Juntamente com o OMFISTA-OLS, o método ADMM apresentou o melhor desempenho na convergência da função custo em todas as reconstruções obtidas. O Capítulo 5 apresenta nossa contribuição para a aceleração do ADMM, tornando-o ainda mais rápido que os demais algoritmos aqui avaliados.

5 PROPOSTA DE ACELERAÇÃO DO MÉTODO ADMM

Neste capítulo apresentamos uma proposta de aceleração do método ADMM apresentado na Seção 3.5. Iniciamos a análise pela generalização do termo de regularização da função custo, mostrando a solução do problema de minimização por meio da versão do ADMM com o *operador proximal*. Em seguida, apresentamos e fundamentamos a modificação do ADMM, e validamos a proposta em quatro experimentos.

Os dois primeiros testes são os mesmos experimentos com dados reais mostrados no Capítulo 4. Eles são aqui repetidos com o propósito de mostrar que o ADMM modificado pode alcançar desempenho superior àquele obtido com o ajuste ótimo do parâmetro de penalidade ρ do algoritmo original.

Os dois últimos experimentos referem-se a reconstruções de imagens em Ressonância Magnética (MRI) com *Compressed Sensing* (CS), e visam mostrar que a modificação do ADMM pode ser usada em outras aplicações além da ultrassonografia esparsa.

Para fins de comparação, também apresentamos a versão proximal do algoritmo FISTA. Os resultados obtidos por esse método são mostrados para os três primeiros experimentos, enquanto o quarto utiliza a versão proximal da variante MFISTA.

5.1 GENERALIZAÇÃO DO TERMO DE REGULARIZAÇÃO

Generalizamos o termo de regularização da Equação (40), substituindo a norma ℓ_1 por um novo operador de regularização $R(\mathbf{f})$ que mantenha o problema bem-posto e que garanta a convexidade da nova função custo $\Psi_R(\mathbf{f})$. O problema de otimização é então modificado, de modo que a solução estimada é dada por

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi_R(\mathbf{f}) = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda R(\mathbf{f}) . \quad (55)$$

A aceleração proposta neste capítulo é válida para operadores $R(\mathbf{f})$ para os quais haja um operador proximal definido. Isso inclui a norma ℓ_1 e também muitos outros operadores, de

modo que passamos a ter mais possibilidades de introduzir informação *a priori* sobre a solução desejada. Quando procuramos uma solução esparsa, podemos fazer $R(\mathbf{f}) = \|\mathbf{f}\|_1$ como discutido no Capítulo 4. Nesse caso, o proximal é implementado pelo operador de *shrinkage-thresholding* S_a , definido na Seção 3.2. No caso do FISTA e do ADMM, esse caso particular é implementado pelo *shrinkage-thresholding* com escalas $a = \lambda/c$ e $a = \lambda/\rho$, respectivamente, como mostrado nos Quadros 1 e 6.

Sugerimos ao(à) leitor(a) consultar os trabalhos de Parikh e Boyd (2013) e de Combettes e Pesquet (2011) para verificar os proximais já definidos para uma grande variedade de operadores de regularização.

5.1.1 O OPERADOR PROXIMAL

Seja $f : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa própria fechada¹. O operador proximal $\text{prox}_{\gamma f} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ da função escalonada γf , em que $\gamma > 0$, é definido como

$$\text{prox}_{\gamma f}(\mathbf{v}) = \arg \min_{\mathbf{x}} \left[f(\mathbf{x}) + \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{v}\|_2^2 \right]. \quad (56)$$

Esse operador, frequentemente usado nos *algoritmos proximais*, está relacionado às subderivadas de funções $f(\mathbf{x})$ não suaves, necessárias em muitos métodos de otimização e em condições de optimalidade. Os algoritmos proximais são utilizados em diversas aplicações de processamento de sinais e também na reconstrução de imagens. Ver, por exemplo, os trabalhos de Combettes e Pesquet (2011) e de Parikh e Boyd (2013) e as referências citadas nessas publicações para maiores detalhes. As versões dos métodos ADMM e FISTA apresentadas a seguir compõem esse conjunto de algoritmos proximais.

5.1.2 VERSÃO PROXIMAL DO ADMM

Consideremos uma formulação para o método ADMM um pouco diferente daquela adotada na Seção 3.5. Seja um problema de minimização em que a função custo é escrita como a soma de duas funções de variáveis distintas, $\mathbf{x}, \mathbf{f} \in \mathbb{R}^m$, relacionadas por uma restrição de

¹Uma função $f : \mathbb{R}^m \rightarrow \mathbb{R}$ é dita *convexa* se $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^m, \forall t \in [0, 1] : f[t\mathbf{x}_1 + (1-t)\mathbf{x}_2] \leq tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2)$. Essa mesma função é dita *fechada* se, para cada $\alpha \in \mathbb{R}$, o conjunto $\{\mathbf{x} \in \text{dom}f | f(\mathbf{x}) < \alpha\}$ for fechado. Uma função *convexa própria* é uma função convexa, tal que $f(\mathbf{x}) < +\infty$ para pelo menos um ponto \mathbf{x} e $f(\mathbf{x}) > -\infty$ para todo \mathbf{x} (BOYD; VANDENBERGHE, 2004).

igualdade:

$$\begin{aligned} \min \quad & p(\mathbf{x}) + q(\mathbf{f}), \\ \text{sujeito a:} \quad & \mathbf{x} - \mathbf{f} = \mathbf{0}. \end{aligned} \quad (57)$$

O lagrangiano aumentado é

$$L_\rho(\mathbf{x}, \mathbf{f}, \mathbf{y}) = p(\mathbf{x}) + q(\mathbf{f}) + \mathbf{y}^T (\mathbf{x} - \mathbf{f}) + \frac{\rho}{2} \|\mathbf{x} - \mathbf{f}\|_2^2, \quad (58)$$

em que $\rho > 0$ é o parâmetro de penalidade e $\mathbf{y} \in \mathbb{R}^m$ é a variável dual associada à restrição. O ADMM soluciona o Problema (57) iterativamente, calculando:

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{f}_k, \mathbf{y}_k) \quad (59)$$

$$\mathbf{f}_{k+1} = \arg \min_{\mathbf{f}} L_\rho(\mathbf{x}_{k+1}, \mathbf{f}, \mathbf{y}_k) \quad (60)$$

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \rho(\mathbf{x}_{k+1} - \mathbf{f}_{k+1}), \quad (61)$$

com os índices subscritos indicando os números das iterações. Esses passos alternados representam, respectivamente, a minimização nas *variáveis primais* \mathbf{x} e \mathbf{f} , e a atualização da variável dual por meio da avaliação da restrição de igualdade escalonada (PARIKH; BOYD, 2013).

Movendo os termos lineares para os termos quadráticos das Equações (59) e (60), obtemos:

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x}} \left[p(\mathbf{x}) + \frac{\rho}{2} \left\| \mathbf{x} - \mathbf{f}_k + \frac{\mathbf{y}_k}{\rho} \right\|_2^2 \right] \quad (62)$$

$$\mathbf{f}_{k+1} = \arg \min_{\mathbf{f}} \left[q(\mathbf{f}) + \frac{\rho}{2} \left\| \mathbf{x}_{k+1} - \mathbf{f} + \frac{\mathbf{y}_k}{\rho} \right\|_2^2 \right]. \quad (63)$$

Podemos escrever essas equações usando os operadores proximais das funções escalonadas $(1/\rho)p$ e $(1/\rho)q$, o que resulta em:

$$\mathbf{x}_{k+1} = \text{prox}_{p/\rho} \left(\mathbf{f}_k - \frac{\mathbf{y}_k}{\rho} \right) \quad (64)$$

$$\mathbf{f}_{k+1} = \text{prox}_{q/\rho} \left(\mathbf{x}_{k+1} + \frac{\mathbf{y}_k}{\rho} \right). \quad (65)$$

Para calcular a solução estimada da Equação (55) com o ADMM, fazemos $p(\mathbf{x}) =$

$(1/2)\|\mathbf{g} - \mathbf{H}\mathbf{x}\|_2^2$ e $q(\mathbf{f}) = \lambda R(\mathbf{f})$, o que transforma o problema de otimização em

$$\begin{aligned} \min \quad & \frac{1}{2}\|\mathbf{g} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda R(\mathbf{f}), \\ \text{sujeito a:} \quad & \mathbf{x} - \mathbf{f} = \mathbf{0}. \end{aligned} \quad (66)$$

Nesse caso, o lagrangiano aumentado torna-se

$$L_\rho(\mathbf{x}, \mathbf{f}, \mathbf{y}) = \frac{1}{2}\|\mathbf{g} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda R(\mathbf{f}) + \frac{\rho}{2}\left\|\mathbf{x} - \mathbf{f} + \frac{\mathbf{y}}{\rho}\right\|_2^2. \quad (67)$$

A minimização de L_ρ em \mathbf{x} é obtida por

$$\begin{aligned} -\mathbf{H}^T(\mathbf{g} - \mathbf{H}\mathbf{x}) + \rho\left(\mathbf{x} - \mathbf{f} + \frac{\mathbf{y}}{\rho}\right) &= \mathbf{0} \\ \hat{\mathbf{x}} &= (\mathbf{H}^T\mathbf{H} + \rho\mathbf{I})^{-1}(\mathbf{H}^T\mathbf{g} + \rho\mathbf{f} - \mathbf{y}), \end{aligned} \quad (68)$$

em que $\mathbf{I} \in \mathbb{R}^{m \times m}$ é uma matriz identidade. A minimização em \mathbf{f} requer a avaliação do operador proximal da função de regularização escalonada λR , que dá origem à solução estimada

$$\hat{\mathbf{f}} = \text{prox}_{\lambda R/\rho}\left(\mathbf{x} + \frac{\mathbf{y}}{\rho}\right). \quad (69)$$

A solução final é então obtida iterativamente, com cada iteração do algoritmo sendo calculada em três passos, que incluem a atualização da variável dual:

$$\mathbf{x}_{k+1} = (\mathbf{H}^T\mathbf{H} + \rho\mathbf{I})^{-1}(\mathbf{H}^T\mathbf{g} + \rho\mathbf{f}_k - \mathbf{y}_k) \quad (70)$$

$$\mathbf{f}_{k+1} = \text{prox}_{\lambda R/\rho}\left(\mathbf{x}_{k+1} + \frac{\mathbf{y}_k}{\rho}\right) \quad (71)$$

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \rho(\mathbf{x}_{k+1} - \mathbf{f}_{k+1}). \quad (72)$$

5.1.3 PARÂMETRO DE PENALIDADE DO ADMM

O aumento de ρ torna mais significativo o termo de penalidade no lagrangiano aumentado da Equação (58), o que resulta em *resíduos primais* $(\mathbf{x}_k - \mathbf{f}_k)$ menores, mas em *resíduos duais* $\rho(\mathbf{f}_k - \mathbf{f}_{k-1})$ maiores (XU et al., 2017). Por outro lado, a diminuição de ρ produzirá resíduos duais menores, mas pode aumentar os primais. Dado que esses resíduos precisam ser pequenos na convergência, é razoável procurar por um valor fixo intermediário para ρ , ajustando-o de modo a manter ambos os resíduos com amplitudes similares (BOYD, 2010; XU et al., 2017). Essa é a estratégia adotada nos experimentos apresentados neste capítulo.

Entretanto, há outras abordagens na literatura. No cenário de ρ fixo, Ghadimi et al. (2015) propõem critérios para a escolha do parâmetro ótimo para problemas regularizados. Os autores adotam um esquema de aceleração, no qual um parâmetro de relaxação faz o balanço de uma combinação linear das soluções \mathbf{x}_k e \mathbf{f}_k obtidas em iterações anteriores.

Também há propostas para a determinação dinâmica dos valores de ρ . Uma delas é discutida no trabalho de Boyd (2010), no qual se propõe inflar e desinflar ρ ao longo das iterações, de modo a sempre manter os resíduos primais e duais com a mesma ordem de grandeza.

Um método de ajuste adaptativo do parâmetro de penalidade é apresentado por Xu et al. (2017). Ele alcança convergência rápida com relativa insensibilidade ao seu valor inicial e ao tamanho do problema de minimização.

Chan et al. (2017) aplicam o ADMM ao problema de filtragem (*denoising*) e propõem um esquema de continuação (*continuation scheme*), no qual o parâmetro de penalidade é incrementado pela regra $\rho_{k+1} = \mu_k \rho_k$, com $\mu_k \geq 1$. Os autores argumentam que o valor inicial ρ_0 deve ser ajustado com o problema, e fornecem faixas de valores típicos para ρ_0 e μ_k para o contexto daquele trabalho.

5.1.4 VERSÃO PROXIMAL DO FISTA

A versão proximal do algoritmo FISTA é apresentada por Beck e Teboulle (2009b). O propósito de utilizá-la neste capítulo é possibilitar o uso do FISTA com uma função de regularização não restrita à norma ℓ_1 , como nos experimentos do Capítulo 4. Com a modificação da função custo da Equação (40) para a da Equação (55), a única diferença em relação ao método apresentado na Seção 3.2 está no passo 6 do Quadro 1.

O algoritmo estima uma nova solução a cada iteração por meio do operador proximal da função de regularização $\lambda R(\mathbf{f})$. O operador é aplicado a um vetor resultante, obtido pela soma de um ponto especial \mathbf{y}_k ao negativo do gradiente nesse ponto, escalonado pelo parâmetro c . A nova solução \mathbf{f}_k é dada por:

$$\mathbf{f}_k = \text{prox}_{\lambda R/c} \left[\frac{1}{c} \mathbf{H}^T (\mathbf{g} - \mathbf{H} \mathbf{y}_k) + \mathbf{y}_k \right]. \quad (73)$$

5.2 ACELERAÇÃO DA PRIMEIRA ITERAÇÃO DO ADMM

Podemos iniciar as iterações do ADMM com valores nulos para os vetores \mathbf{f} e \mathbf{y} . Nesse caso, as Equações (70)-(72) para a *primeira iteração* tornam-se:

$$\mathbf{x}_1 = (\mathbf{H}^T \mathbf{H} + \rho \mathbf{I})^{-1} \mathbf{H}^T \mathbf{g} \quad (74)$$

$$\mathbf{f}_1 = \text{prox}_{\lambda R/\rho}(\mathbf{x}_1) \quad (75)$$

$$\mathbf{y}_1 = \rho(\mathbf{x}_1 - \mathbf{f}_1). \quad (76)$$

Para $n \geq m$ e $\text{posto}(\mathbf{H}) = m$, a solução \mathbf{x}_1 obtida pela Equação (74) é *única* e corresponde à solução de quadrados mínimos regularizada de Tikhonov, que minimiza $(1/2)\|\mathbf{g} - \mathbf{H}\mathbf{x}\|_2^2 + (\rho/2)\|\mathbf{x}\|_2^2$.

Uma vez que a solução quadrática é única nessa situação, propomos *desacoplar* o parâmetro de penalidade ρ no trio de equações do ADMM para a primeira iteração.

Na Equação (74) o parâmetro de penalidade ρ é substituído por um novo parâmetro ξ^2 , enquanto nas Equações (75) e (76) ρ é substituído pelo produto $\delta\rho$, sendo δ um parâmetro adicional a ser determinado. Em seguida, procuramos por uma combinação de valores para ξ^2 e δ que reduza o custo de \mathbf{f}_1 e que promova a atualização apropriada de \mathbf{y}_1 para a iteração seguinte.

Assim, a primeira iteração do ADMM é modificada para:

$$\mathbf{x}_1 = (\mathbf{H}^T \mathbf{H} + \xi^2 \mathbf{I})^{-1} \mathbf{H}^T \mathbf{g} \quad (77)$$

$$\mathbf{f}_1 = \text{prox}_{\lambda R/(\delta\rho)}(\mathbf{x}_1) \quad (78)$$

$$\mathbf{y}_1 = \delta\rho(\mathbf{x}_1 - \mathbf{f}_1), \quad (79)$$

e o algoritmo ADMM acelerado completo é mostrado no Quadro 7.

A ideia de separar o parâmetro de penalidade ρ nas equações do ADMM não é nova. Por exemplo, Sawatzky et al. (2014) propõem uma *matriz de penalidades* para problemas com múltiplas restrições, o que permite a definição de valores distintos de penalidade para cada uma das restrições.

Outra proposta é apresentada por Deng e Yin (2016), em que ρ é usado para calcular \mathbf{x} e \mathbf{f} , e é então escalonado como $\theta\rho$ apenas na atualização de \mathbf{y} . O novo parâmetro θ atua como um *passo* para a atualização da variável dual.

Essas e outras proposições consistem em modificações do parâmetro ρ que usualmente

Quadro 7: Algoritmo ADMM acelerado

Entradas: $\rho, \xi, \delta, \lambda$

- 1: $\mathbf{f}_0 = \mathbf{0}$
- 2: $\mathbf{y}_0 = \mathbf{0}$
- 3: $k = 1$
- 4: **enquanto** Critério de parada não atingido **faça**
- 5: **se** $k = 1$ **então**
- 6: $\mathbf{x}_1 = (\mathbf{H}^T \mathbf{H} + \xi^2 \mathbf{I})^{-1} \mathbf{H}^T \mathbf{g}$
- 7: $\mathbf{f}_1 = \text{prox}_{\lambda R / (\delta \rho)}(\mathbf{x}_1)$
- 8: $\mathbf{y}_1 = \delta \rho (\mathbf{x}_1 - \mathbf{f}_1)$
- 9: **senão**
- 10: $\mathbf{x}_k = (\mathbf{H}^T \mathbf{H} + \rho \mathbf{I})^{-1} (\mathbf{H}^T \mathbf{g} + \rho \mathbf{f}_{k-1} - \mathbf{y}_{k-1})$
- 11: $\mathbf{f}_k = \text{prox}_{\lambda R / \rho} \left(\mathbf{x}_k + \frac{\mathbf{y}_{k-1}}{\rho} \right)$
- 12: $\mathbf{y}_k = \mathbf{y}_{k-1} + \rho (\mathbf{x}_k - \mathbf{f}_k)$
- 13: **fim se**
- 14: $k = k + 1$
- 15: **fim enquanto**

Fonte: O autor (2017)

requerem provas de convergência. Além disso, elas podem introduzir instabilidade ao ADMM em certos casos. Um exemplo é encontrado na avaliação da proposta de Deng e Yin (2016) feita por Ghadimi et al. (2015).

A modificação aqui proposta, no entanto, não afeta a garantia de convergência do algoritmo. Isso se dá em virtude de (55) ser um problema convexo² e porque *apenas a primeira iteração é modificada*, dando origem a uma solução de custo inferior e, portanto, finita. Deste modo, nossa proposição pode ser usada para conferir ganho *adicional* a outras formas de aceleração do ADMM existentes na literatura, tais como as citadas acima e aquela proposta por Goldstein et al. (2014), entre outras.

Determinação dos Parâmetros ξ^2 e δ

A escolha de valores apropriados para os novos parâmetros ξ^2 e δ depende de fatores como a distribuição dos valores singulares da matriz \mathbf{H} , os erros do modelo e as propriedades da imagem (HANSEN, 1998).

Para determinar apenas ξ^2 , pode-se utilizar o *princípio de discrepância* proposto

²A convexidade da função custo da Equação (55) depende da escolha da função de regularização R . Nos experimentos apresentados neste estudo, todas as funções custo são convexas.

por Morozov (1966). Ele é baseado na suposição de que o vetor de dados sem ruído $\mathbf{g}_{\text{exato}}$ ou a norma $\|\mathbf{e}\|_2 = \|\mathbf{g} - \mathbf{g}_{\text{exato}}\|_2$ do erro é conhecida. Se essas informações não estiverem disponíveis, a melhor opção para a escolha de ξ^2 de acordo com Hansen (1998) vem, em geral, da avaliação de métodos distintos, tais como o GCV e o método da Curva L, que descrevemos na Seção 2.3.3.

Como discutido na Seção 2.3.1, o parâmetro ξ^2 atua conjuntamente com os valores singulares σ_i da matriz \mathbf{H} na determinação dos *fatores de filtro*

$$\phi_i = \frac{\sigma_i^2}{\sigma_i^2 + \xi^2} \quad (80)$$

na regularização de Tikhonov, os quais multiplicam cada inversão $1/\sigma_i$ que ocorre na inversão matricial da Equação (77).

A filtragem de Tikhonov tem a função de suavizar as inversões dos valores singulares muito pequenos, que caracterizam o mal condicionamento do problema.

Observamos esse efeito na Equação (13), ao notar que valores singulares $\sigma_i > \xi$ resultam em fatores $\phi_i \approx 1$ e as suas inversas filtradas tendem a $1/\sigma_i$. Por outro lado, valores singulares $\sigma_i < \xi$ resultam em fatores $\phi_i \approx \sigma_i^2/\xi^2$, que fazem as inversões tenderem a σ_i/ξ^2 e a terem, conseqüentemente, amplitudes significativamente reduzidas para valores singulares muito pequenos. Por isso, essa filtragem reduz a amplificação do ruído em \mathbf{g} ao calcular \mathbf{x}_1 na Equação (77) (HANSEN, 1998).

Seja σ_{\min} o menor valor singular não nulo de \mathbf{H} e σ_{\max} o seu maior valor singular. Considerando $\sigma_{\min} < \xi < \sigma_{\max}$, definimos $\xi = \sqrt{\alpha} \sigma_{\max} = \sqrt{\alpha} \|\mathbf{H}\|_2$ ou, de modo equivalente,

$$\xi^2 = \alpha \|\mathbf{H}^T \mathbf{H}\|_2. \quad (81)$$

Devemos, então, determinar o fator $\alpha < 1$ que resulte em uma solução de custo reduzido na primeira iteração modificada, calculada pelas Equações (77) e (78).

A escolha de δ , entretanto, é uma tarefa distinta. Esse parâmetro afeta o efeito de filtragem (ou *denoising*) promovido pelo operador proximal. É intuitivo pensar que, se ξ^2 for pequeno na Equação (77), a solução \mathbf{x}_1 será ruidosa; então, δ deveria ser menor³ que 1 para compensar esse efeito com a suavização da solução final da primeira iteração, \mathbf{f}_1 . Entretanto, essa intuição se confirmou em apenas um dos quatro experimentos apresentados neste capítulo (ver o caso 4 na Seção 5.5). Isso indica que, em alguns casos, uma solução mais ruidosa na primeira iteração pode apresentar custo inferior àquele que se esperaria de uma solução mais

³Observe que o parâmetro δ pertence ao *denominador* na escala do proximal da Equação (78).

filtrada.

Para solucionar o problema de determinação dos dois parâmetros, utilizamos neste estudo uma abordagem exaustiva. Os melhores valores para os parâmetros foram escolhidos como

$$(\xi, \delta) = \arg \min_{\xi, \delta} \Psi(\mathbf{f}_1^{\xi, \delta}), \quad (82)$$

em que

$$\mathbf{f}_1^{\xi, \delta} = \text{prox}_{\lambda R/(\delta\rho)} [(\mathbf{H}^T \mathbf{H} + \xi^2 \mathbf{I})^{-1} \mathbf{H}^T \mathbf{g}]. \quad (83)$$

É importante notar que a busca prévia exaustiva por parâmetros não é um problema do ponto de vista prático. Em muitos sistemas de imageamento, todos os parâmetros de aquisição e de reconstrução, tais como a matriz \mathbf{H} do sistema, ρ , ξ^2 , δ , entre outros, são pré-definidos e estão armazenados na plataforma de reconstrução *antes* da coleta de dados (*scanning*). Isso significa que o uso de parâmetros ótimos, armazenados em memória para propiciar aceleração na reconstrução, é algo factível.

5.3 EXPERIMENTOS EM ULTRASSONOGRRAFIA ESPARSA

5.3.1 DESCRIÇÃO DOS EXPERIMENTOS – MATERIAIS E MÉTODOS

Nesta seção, avaliamos a proposta de aceleração do ADMM com um termo de regularização de norma ℓ_1 , fazendo $R(\mathbf{f}) = \|\mathbf{f}\|_1$. A função custo é a mesma da Equação (40), descrita novamente por conveniência:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi(\mathbf{f}) = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \|\mathbf{f}\|_1. \quad (84)$$

Apresentamos dois experimentos, que designamos de “caso 1” e “caso 2”. Tratam-se de reconstruções baseadas nos mesmos modelos \mathbf{H} e dados reais \mathbf{g} utilizados no Capítulo 4. Nesta seção, identificamos como “caso 1” as reconstruções baseadas no conjunto de dados definido na Seção 4.2.2, e como “caso 2” aquelas que usam o conjunto definido na Seção 4.3.3. Dado que o propósito aqui é observar as primeiras iterações para confirmar a aceleração proposta, os algoritmos são encerrados após 20 iterações e não mais 30, como na validação do modelamento.

Em termos de dimensões, temos $n = 50.944$ e $m = 6.561$ para o caso 1, e $n = 58.368$ e $m = 6.497$ para o caso 2. O parâmetro λ é o mesmo definido na Seção 4.3, obtido com $\kappa = 0,01$.

5.3.2 ESTIMATIVA DOS PARÂMETROS

Inicialmente, buscamos pelo valor ótimo do parâmetro de penalidade ρ que resulte na convergência mais rápida para o ADMM não modificado. Como feito nos experimentos do Capítulo 4, partimos da suposição inicial de que ρ é proporcional ao parâmetro $c = \|\mathbf{H}^T \mathbf{H}\|_2$:

$$\rho = \beta c, \text{ em que } 0 < \beta < 1, \quad (85)$$

e buscamos por um valor apropriado para o fator de multiplicação β .

A Figura 34 apresenta o custo variando em função de β para os casos 1 e 2, e indica que nossa suposição é plausível. Uma vez que buscamos pela convergência mais rápida nas primeiras iterações do ADMM, apenas cinco são calculadas antes de avaliarmos $\Psi(\mathbf{f}_\beta)$.

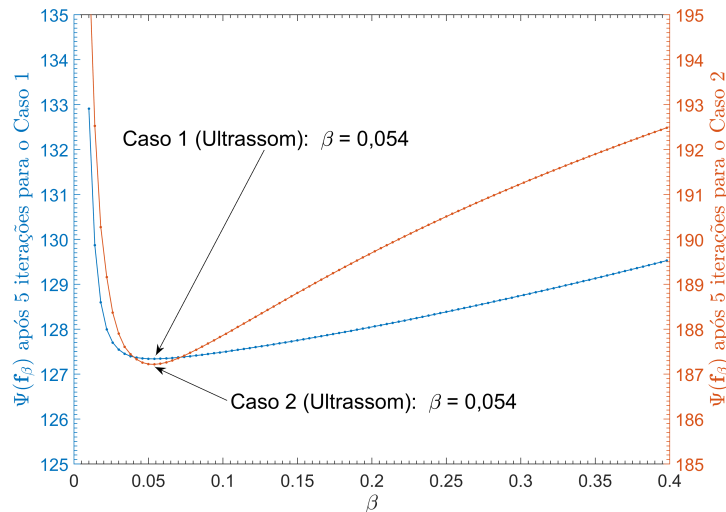


Figura 34: Estimativa do fator β para os casos 1 e 2 (Ultrassom).

Fonte: O autor (2017)

Os pontos de mínimo correspondem a $\beta = 0,054$ em ambos os casos, e fazemos $\rho = 0,054 c$ para executar o método ADMM original com o parâmetro ρ fixo, assim como todas as iterações do ADMM modificado a partir da segunda.

Uma vez determinado o parâmetro ρ ótimo, traçamos as curvas de custo para diferentes combinações de α e δ , em busca da melhor solução para a primeira iteração. A Figura 35 mostra as curvas obtidas. O ponto destacado **A** corresponde a $\alpha = \beta = 0,054$ e $\delta = 1$, que dão origem à solução com custo 143,6 atingida pela primeira iteração do ADMM padrão com o melhor parâmetro de penalidade, $\rho = 0,054 c$.

O ponto **B** corresponde a $\alpha = 0,007$ e $\delta = 1$, que promovem a solução com o menor

custo (134,8) que poderia ser atingido se apenas substituíssemos ρ por ξ^2 no termo quadrático, isto é, se trocássemos a Equação (74) pela Equação (77), sem nenhuma alteração nas Equações (75) e (76).

O ponto C indica o melhor caso, com $\alpha = 0,027$ e $\delta = 3,5$. Ele correspondente à aceleração proposta, em que as Equações (77) a (79) são utilizadas. O custo 130,2 é alcançado pela primeira iteração do ADMM modificado, com ρ substituído por $\xi^2 = 0,027 c$ no termo quadrático, e aumentado 3,5 vezes na escala do proximal e na atualização da variável dual.

Uma avaliação similar foi feita para o caso 2, e o resumo encontra-se na Tabela 5.

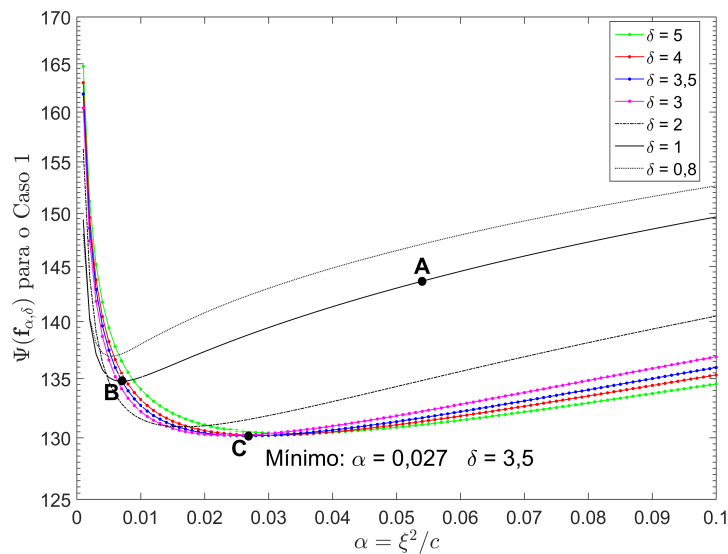


Figura 35: Estimativa de α e δ para o caso 1 (Ultrassom).

Fonte: O autor (2017)

Tabela 5: Parâmetros β , α e δ para os casos 1 e 2

Experimento	ADMM padrão	ADMM acelerado (apenas 1ª it. modificada)
Caso 1 - Ultrassom	$\beta = 0,054$	$\alpha = 0,027, \delta = 3,5$
Caso 2 - Ultrassom	$\beta = 0,054$	$\alpha = 0,043, \delta = 3,0$

Fonte: O autor (2017)

5.3.3 RESULTADOS DA ACELERAÇÃO EM ULTRASSONOGRRAFIA ESPARSA

A evolução da função custo $\Psi(\mathbf{f})$ da Equação (84) é mostrada nas Figuras 36 e 37 para os casos 1 e 2, respectivamente, variando em função do número da iteração k e em função do tempo.

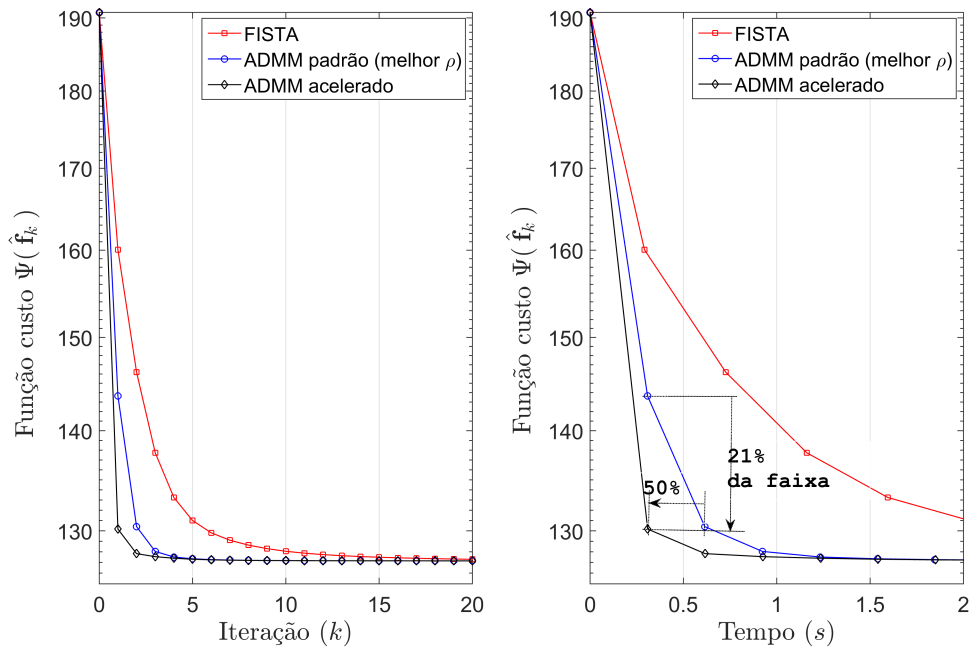


Figura 36: Caso 1 (Ultrassom) - Custo \times iterações (esquerda) e tempo (direita).

Fonte: O autor (2017)

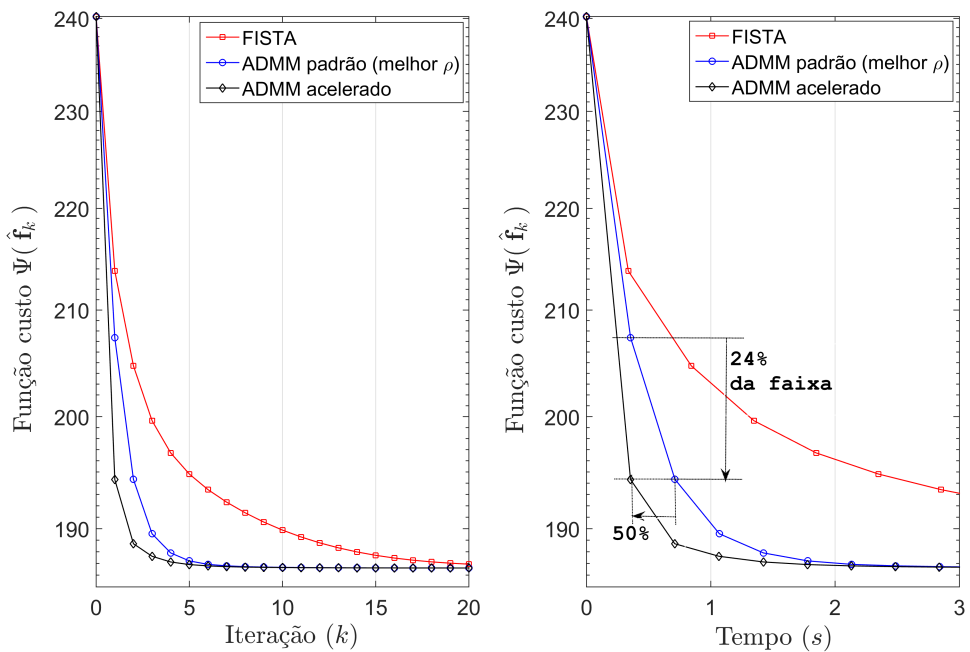


Figura 37: Caso 2 (Ultrassom) - Custo \times iterações (esquerda) e tempo (direita).

Fonte: O autor (2017)

Para avaliar a modificação proposta, o desempenho do ADMM é apresentado em duas configurações distintas. A primeira (linha azul) mostra o ADMM não modificado, sendo executado com o melhor parâmetro ρ determinado para cada experimento. A segunda curva

(linha preta) mostra o ADMM acelerado proposto. O algoritmo FISTA também é apresentado como referência para a comparação (linha vermelha).

À direita das Figuras 36 e 37 os níveis de custo para a primeira iteração estão indicados para o ADMM não modificado e para a modificação proposta. A aceleração alcançada representa mais de 21% da faixa para o caso 1, e mais de 24% da faixa para o caso 2. Em ambos os casos, o ADMM modificado atinge o mesmo nível de custo alcançado pelo ADMM padrão em aproximadamente metade do tempo.

Na Figura 38 vemos as evoluções dos erros de estimação $\|\hat{\mathbf{f}}_k - \mathbf{f}^*\|_2$ para os casos 1 e 2. O vetor \mathbf{f}^* é a solução de referência, obtida por um algoritmo distinto para evitar uma avaliação tendenciosa. Usamos o método IRLS descrito na Seção 3.4 e, em ambos os casos, a solução \mathbf{f}^* foi obtida após 500 iterações. As curvas mostram que a modificação proposta para o ADMM produz soluções $\hat{\mathbf{f}}_k$ com erros inferiores aos do ADMM padrão, tanto na primeira iteração quanto nas subsequentes.

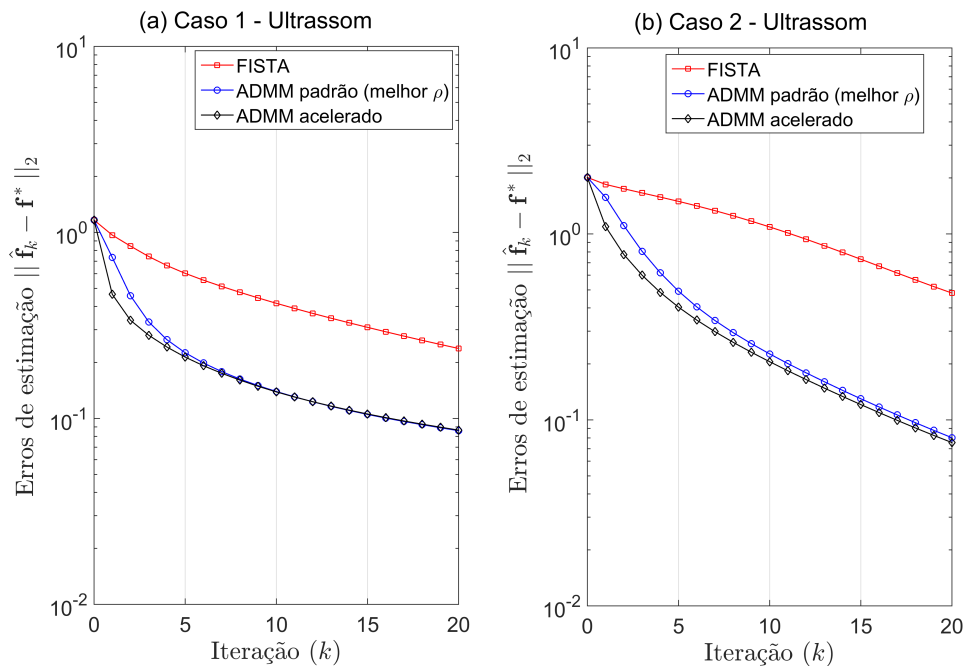


Figura 38: Casos 1 e 2 (Ultrassom) - Erros de estimação \times iterações.

Fonte: O autor (2017)

As imagens reconstruídas são mostradas nas Figuras 39 e 40, nas quais as reconstruções realizadas pelo equipamento de ultrassom são identificadas como *Beamforming*. A fileira inferior de ambas as figuras mostra os resultados do FISTA (esquerda), do ADMM padrão com o melhor ρ (centro) e para o ADMM acelerado proposto (direita), todos após 20 iterações. No canto superior direito de cada figura, vemos a imagem reconstruída pela primeira

iteração do ADMM modificado.

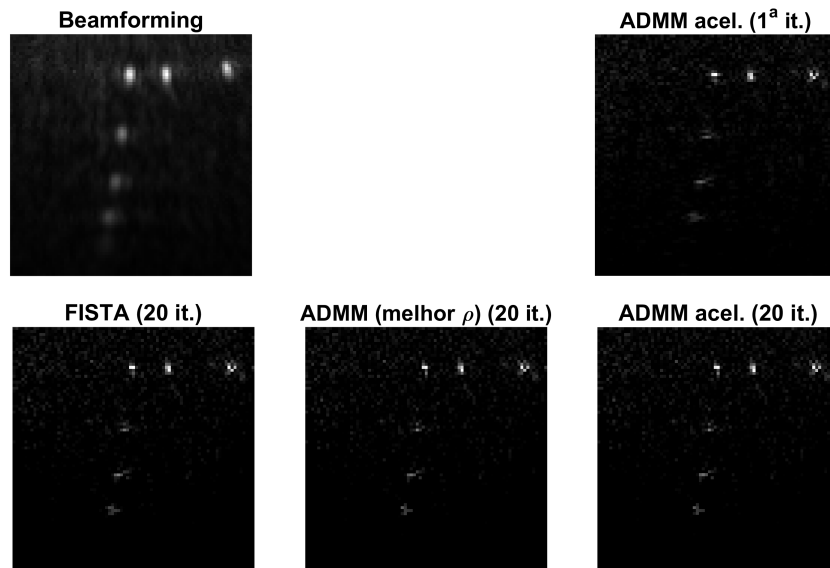


Figura 39: Caso 1 (Ultrassom) - Imagens reconstruídas.

Fonte: O autor (2017)

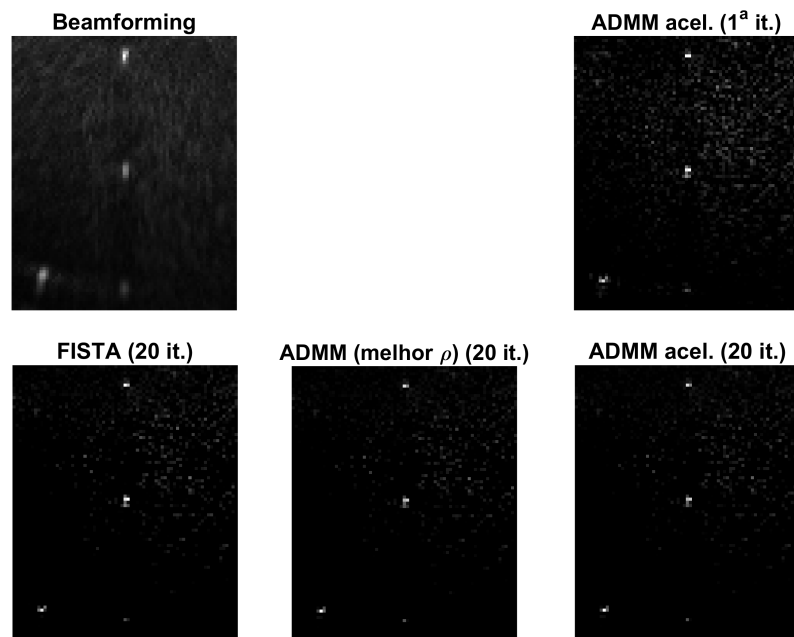


Figura 40: Caso 2 (Ultrassom) - Imagens reconstruídas.

Fonte: O autor (2017)

A métrica API definida na Seção 4.3 é usada para comparar as imagens, e os resultados são mostrados na Tabela 6.

Tabela 6: Valores API para as imagens de ultrassom reconstruídas.

Reconstrução	API para Caso 1	API para Caso 2
<i>Beamforming</i>	20,8626	5,3886
ADMM acelerado (primeira iteração)	6,7047	3,5369
FISTA	7,0334	4,7327
ADMM padrão com melhor ρ	6,6164	3,9155
ADMM acelerado	6,6143	3,9036

Fonte: O autor (2017)

Em ambos os casos o ADMM acelerado alcança melhores resultados que todos os demais métodos de reconstrução. Além disso, a imagem gerada pela primeira iteração do ADMM acelerado é melhor que a produzida pelo método tradicional de *beamforming*. Esse é um passo bastante importante na direção de reconstruções de alta qualidade em tempo real baseadas em IPB para sistemas ultrassônicos.

5.4 EXPERIMENTO EM RESSONÂNCIA MAGNÉTICA NO DOMÍNIO WAVELET

Designamos como “caso 3” o terceiro experimento formulado para avaliar a aceleração proposta do ADMM. Neste caso, utilizamos um termo de regularização com norma ℓ_1 no domínio Wavelet⁴ para reconstrução de imagens de Ressonância Magnética (MRI). A função custo é expressa com os coeficientes $\mathbf{d} = \mathbf{W}\mathbf{f}$:

$$\hat{\mathbf{d}} = \arg \min_{\mathbf{d}} \Psi(\mathbf{d}) = \arg \min_{\mathbf{d}} \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{W}^T \mathbf{d}\|_2^2 + \lambda \|\mathbf{d}\|_1, \quad (86)$$

em que \mathbf{W} representa uma transformada Wavelet de 4 níveis do tipo Daubechies, e a imagem estimada é calculada após a minimização por meio do produto $\hat{\mathbf{f}} = \mathbf{W}^T \hat{\mathbf{d}}$. O parâmetro λ é calculado com $\kappa = 0,001$.

5.4.1 DESCRIÇÃO DO EXPERIMENTO – MATERIAIS E MÉTODOS

A matriz de aquisição \mathbf{H} e sua transposta \mathbf{H}^T são implementadas como funções *in-line* no MATLAB[®], o que caracteriza um procedimento de programação usualmente chamado de *fake matrix*. Isso significa que não temos as matrizes \mathbf{H} e \mathbf{H}^T verdadeiras armazenadas em memória, mas sim funções que retornam o produto de \mathbf{H} ou de \mathbf{H}^T por um vetor passado como argumento. Portanto, a inversão matricial e a transposição utilizadas nas Equações (70) e (77)

⁴Detalhes sobre a definição da transformada Wavelet, suas aplicações e vantagens no processamento de imagens podem ser encontrados na Seção 5.3 de Barrett e Myers (2004) e nas referências ali citadas.

do ADMM não podem ser calculadas explicitamente, de modo que a solução quadrática deve ser obtida por algum algoritmo iterativo. Utilizamos o método do Gradiente Conjugado (CG), brevemente abordado na Seção 3.3, para esse propósito.

O *phantom* utilizado é baseado em um corte transaxial do crânio humano, e imita uma imagem médica verdadeira. Para maiores detalhes, veja as Seções 4.3 e 4.4 de (HERMAN, 2009). O *phantom* possui uma distribuição aleatória de não-homogeneidades e a inserção de pequenos “tumores”; os detalhes podem ser encontrados na Seção 5.2 da mesma referência. O *phantom* é uma digitalização com tamanho de 512×512 *pixels*, produzida pelo programa SNARK14 (DAVIDI et al., 2017; HERMAN, 2009).

Os dados simulados de MRI referem-se a uma aquisição paralela de *Compressed Sensing* (CS) com 6 bobinas (*coils*). A subamostragem utiliza um disco de Poisson com uma área central de espaço-k de 85×85 *pixels*. O fator de aceleração do CS é ajustado em 4. Os dados sintetizados são contaminados com ruído gaussiano, resultando em uma relação sinal-ruído (SNR) de 40 dB. Em relação às dimensões do problema de minimização, temos $n > m$.

5.4.2 ESTIMATIVA DOS PARÂMETROS

Como para os casos de ultrassom, buscamos inicialmente o valor ótimo para o parâmetro de penalidade ρ , supondo-o proporcional ao parâmetro c . A Figura 41 mostra que o custo mínimo após cinco iterações é obtido para $\beta = 0,062$. Portanto, utilizamos $\rho = 0,062 c$ para o ADMM padrão e para todas as iterações do ADMM modificado a partir da segunda.

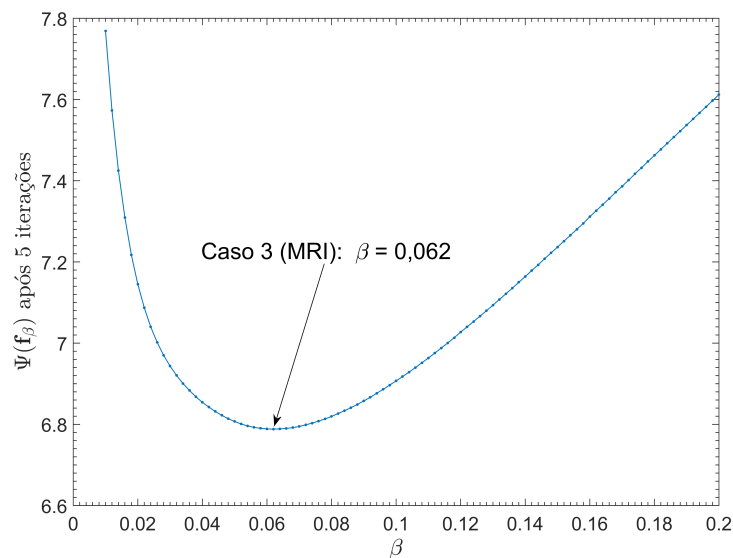


Figura 41: Estimativa do fator β para o caso 3 (MRI).

Fonte: O autor (2017)

Seguindo o mesmo procedimento adotado para os casos 1 e 2, avaliamos o custo da primeira iteração para diferentes combinações dos parâmetros α e δ . Os melhores valores são mostrados na Tabela 7.

Tabela 7: Parâmetros β , α e δ para o caso 3

Experimento	ADMM padrão	ADMM acelerado (apenas 1ª it. modificada)
Caso 3 - MRI	$\beta = 0,062$	$\alpha = 8 \times 10^{-4}$, $\delta = 3,0$

Fonte: O autor (2017)

5.4.3 RESULTADOS DA ACELERAÇÃO EM MRI

A evolução do custo é mostrada na Figura 42, variando em função do número da iteração k e em função do tempo.

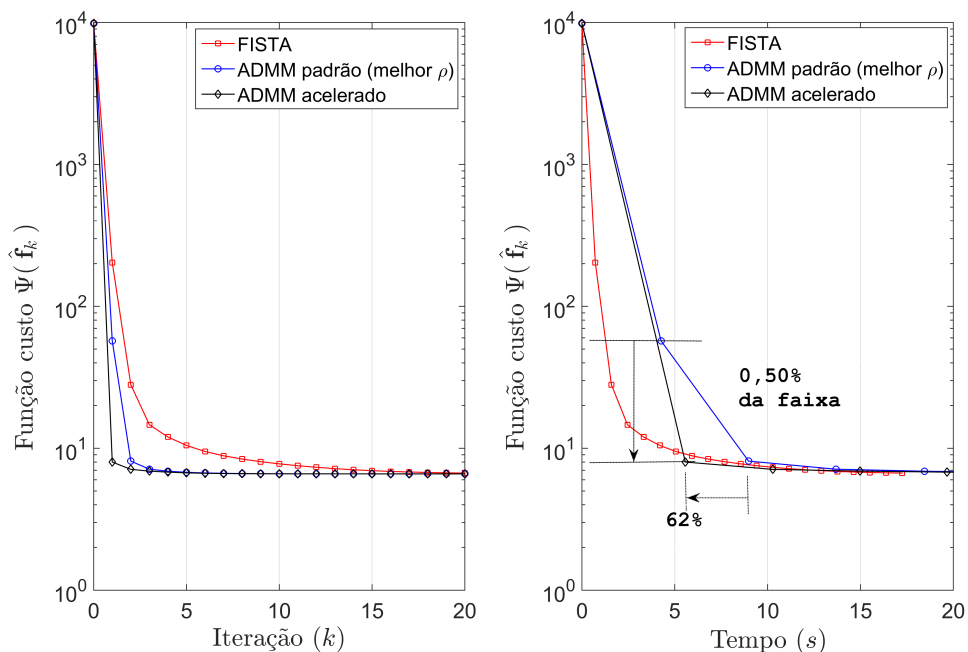


Figura 42: Caso 3 (MRI) - Custo \times iterações (esquerda) e tempo (direita).

Fonte: O autor (2017)

Como nos casos de ultrassom, observamos que a aceleração do ADMM reduz o custo da primeira iteração. O ganho representa apenas 0,50% da faixa, mas o custo atingido é inferior ao da 9ª iteração do FISTA; além disso, esse nível é atingido em cerca de 62% do tempo requerido pelo ADMM padrão em sua melhor configuração.

Na Figura 43 vemos a evolução dos erros de estimação $\|\hat{\mathbf{f}}_k - \mathbf{f}^*\|_2$, em que a referência \mathbf{f}^* é a imagem do *phantom*. Como nos casos de ultrassom, as soluções $\hat{\mathbf{f}}_k$ geradas pelo ADMM acelerado resultam em erros inferiores aos do ADMM não modificado, tanto na primeira iteração quanto nas demais.

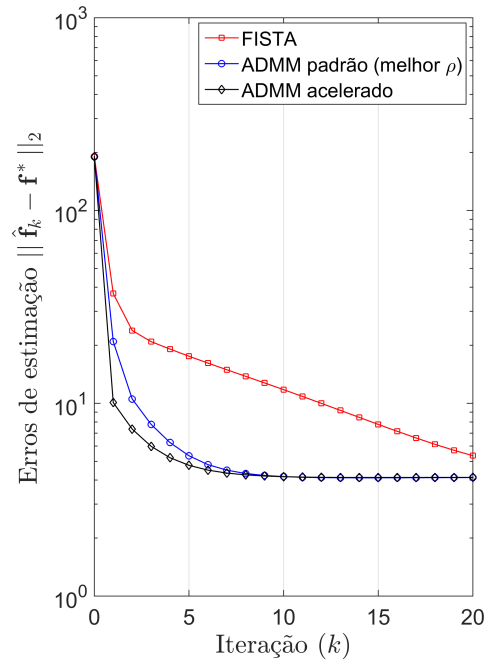


Figura 43: Caso 3 (MRI) - Erros de estimação \times iterações.

Fonte: O autor (2017)

As imagens reconstruídas são mostradas na Figura 44, com os resultados SNR correspondentes referidos ao *phantom*, na fileira superior à esquerda. Ao centro da fileira superior vemos a reconstrução $\hat{\mathbf{f}}_{CG} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{g}$ obtida pelo método CG, enquanto à direita vemos a reconstrução obtida na primeira iteração do ADMM acelerado proposto. Essa última apresenta uma SNR superior à da reconstrução via CG, graças à regularização e ao passo proximal (observe que a solução $\hat{\mathbf{f}}_{CG}$ corresponde à solução quadrática não regularizada).

Na fileira inferior, da esquerda para a direita, observamos as imagens reconstruídas após 20 iterações do FISTA, do ADMM padrão em sua melhor configuração e do ADMM modificado. As SNR resultantes para ambas as soluções do ADMM são bastante próximas, e melhores que a SNR da solução gerada pelo FISTA.

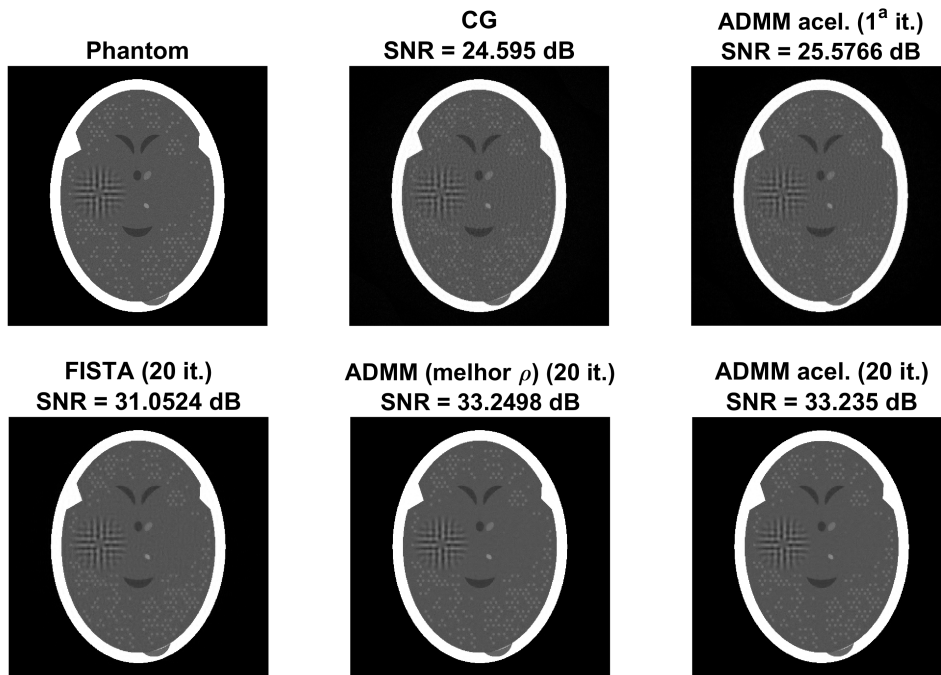


Figura 44: Caso 3 (MRI) - Imagens reconstruídas.

Fonte: O autor (2017)

5.5 EXPERIMENTO EM RESSONÂNCIA MAGNÉTICA COM *TOTAL VARIATION*

O quarto experimento, designado de “caso 4”, avalia a aceleração proposta para o ADMM com um termo de regularização do tipo *Total Variation*⁵ aplicado à reconstrução em MRI. A solução estimada $\hat{\mathbf{f}}$ é obtida por

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \Psi(\mathbf{f}) = \arg \min_{\mathbf{f}} \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \text{TV}(\mathbf{f}), \quad (87)$$

em que TV representa o *Total Variation* anisotrópico.

Nesse caso, o operador proximal que minimiza a função quadrática somada ao termo $\text{TV}(\mathbf{f})$ é o algoritmo *Fast Gradient Projection* (FGP). Tanto a penalidade TV quanto o método FGP estão definidos no trabalho de Beck e Teboulle (2009b). Nesse experimento seguimos a orientação dos autores, substituindo o FISTA por sua variante monotônica MFISTA, apresentada na Seção 3.2. De acordo com os proponentes do método FGP, a variante monotônica apresenta melhor estabilidade que o FISTA quando esses algoritmos são

⁵Na área de Processamento de Sinais, a eliminação de ruídos (*denoising*) por meio da regularização do tipo *Total Variation* baseia-se no princípio de que os sinais (neste estudo, imagens) com detalhes excessivos e possivelmente espúrios (ruídos) têm uma variação total elevada, ou seja, a integral do gradiente absoluto do sinal é alta. De acordo com este princípio, a redução da variação total do sinal sujeito a esses ruídos remove detalhes indesejados, preservando detalhes importantes, como as bordas.

implementados com o FGP. O parâmetro λ é calculado com $\kappa = 0,001$, como no caso 3.

5.5.1 DESCRIÇÃO DO EXPERIMENTO – MATERIAIS E MÉTODOS

Assim como no experimento em MRI no domínio Wavelet, a matriz de aquisição \mathbf{H} e sua transposta \mathbf{H}^T são implementadas como funções *in-line* no MATLAB, e o CG também é utilizado como método iterativo para calcular as inversões da Equações (70) e (77). O *phantom* é o mesmo do caso 3, os dados adquiridos são simulados do mesmo modo, e temos $n > m$.

5.5.2 ESTIMATIVA DOS PARÂMETROS

Como nos três casos anteriores, buscamos inicialmente o melhor parâmetro de penalidade ρ para o ADMM não modificado. De acordo com a Figura 45, devemos utilizar $\rho = 0,180c$ para o ADMM padrão, bem como para todas as iterações do ADMM modificado a partir de segunda.

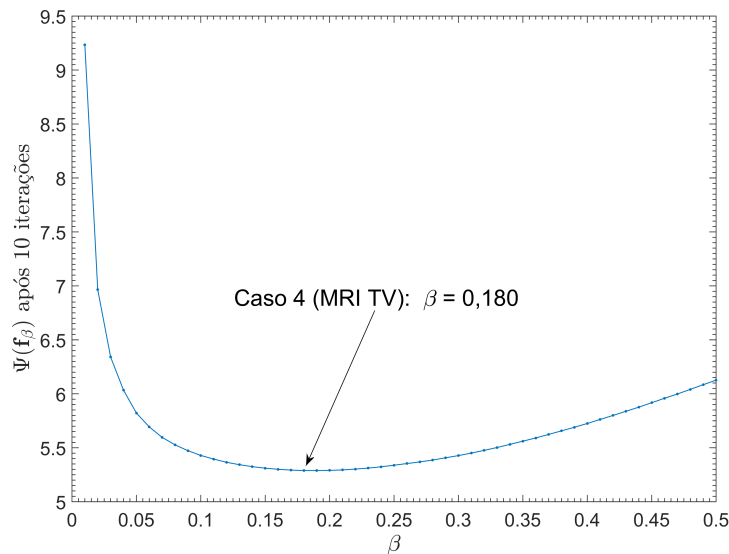


Figura 45: Estimativa do fator β para o caso 4 (MRI TV).

Fonte: O autor (2017)

Como antes, avaliamos em seguida o custo da *primeira iteração* para diferentes combinações de α e δ . Os melhores valores são mostrados na Tabela 8.

Tabela 8: Parâmetros β , α e δ para o caso 4

Experimento	ADMM padrão	ADMM acelerado (apenas 1ª it. modificada)
Caso 4 - MRI com TV	$\beta = 0,180$	$\alpha = 2 \times 10^{-4}$, $\delta = 0,4$

Fonte: O autor (2017)

5.5.3 RESULTADOS DA ACELERAÇÃO EM MRI TV

A evolução da função custo é mostrada na Figura 46, variando com o número da iteração k e com o tempo. Como nos casos 1, 2 e 3, vemos que a aceleração proposta também reduz o custo da primeira iteração. O ganho representa cerca de 2,69% da faixa, está aproximadamente no mesmo nível da 7ª iteração do MFISTA, e atinge esse nível em aproximadamente 56% do tempo requerido pelo ADMM padrão em sua melhor configuração.

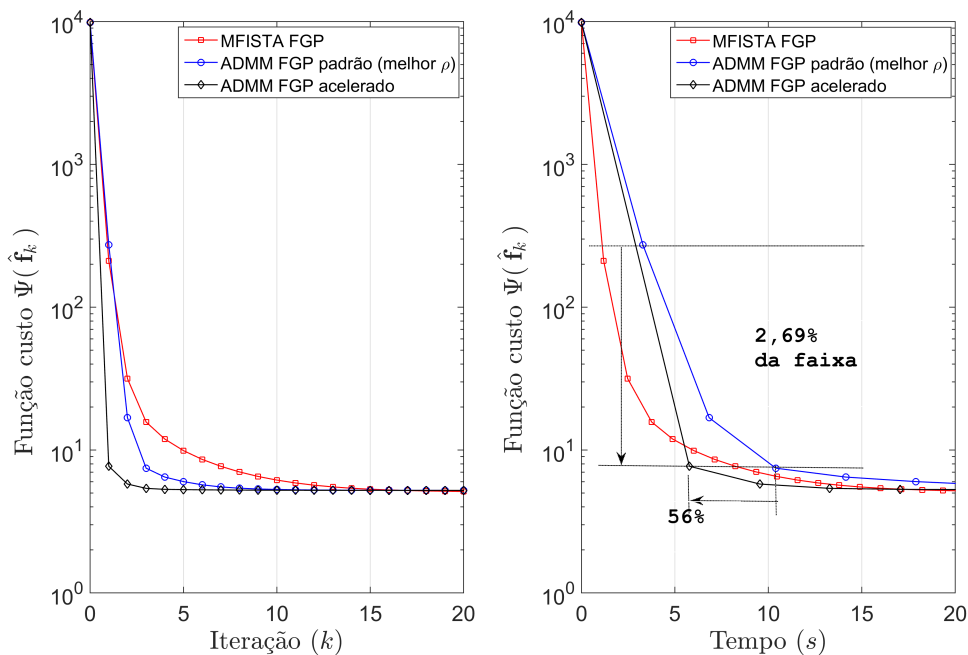


Figura 46: Caso 4 (MRI TV) - Custo \times iterações (esquerda) e tempo (direita).

Fonte: O autor (2017)

A Figura 47 apresenta a evolução dos erros de estimação $\|\hat{\mathbf{f}}_k - \mathbf{f}^*\|_2$, adotando novamente a imagem do *phantom* como a referência \mathbf{f}^* . Como antes, as soluções $\hat{\mathbf{f}}_k$ produzidas pelo ADMM acelerado apresentam erros inferiores na primeira iteração e nas subsequentes, quando comparadas às soluções do ADMM não modificado. Entretanto, nesse caso, ambas as variantes do ADMM tendem a soluções finais com erros ligeiramente maiores que o da solução obtida pelo MFISTA. A diferença entre os erros representa aproximadamente 0,4% da faixa em

ambos os casos, o que corrobora os resultados das SNR mostrados a seguir.

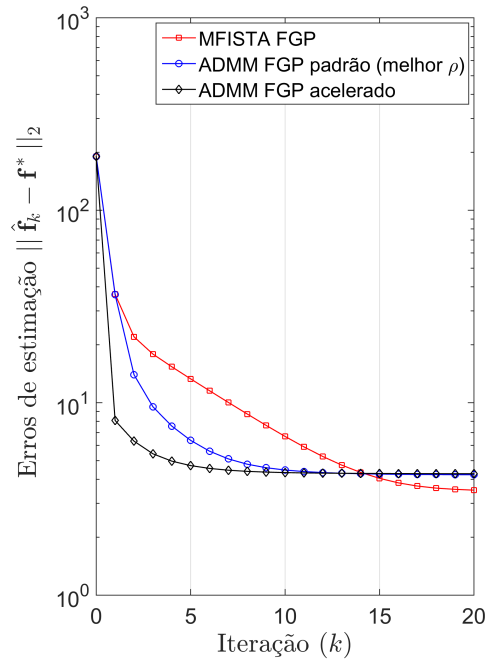


Figura 47: Caso 4 (MRI TV) - Erros de estimação \times iterações.

Fonte: O autor (2017)

As imagens reconstruídas são mostradas na Figura 48, novamente com as SNR referidas ao *phantom* à esquerda, na fileira superior. A distribuição das imagens reconstruídas é a mesma do caso 3. Notamos que a reconstrução da primeira iteração do ADMM acelerado tem novamente uma SNR superior à da reconstrução via CG.

Na fileira inferior, as imagens reconstruídas após 20 iterações para ambas as variantes do ADMM são muito próximas, mas apresentam SNR inferior à da solução do MFISTA devido aos maiores erros de estimação.

5.6 CONCLUSÕES SOBRE A ACELERAÇÃO DO ADMM

Neste capítulo propusemos um esquema de aceleração para o ADMM e avaliamos o seu desempenho para a solução de problemas de quadrados mínimos com penalizações. Esta proposta se enquadra no caso em que $n \geq m$ e $\text{posto}(\mathbf{H}) = m$, que são condições encontradas em alguns cenários de reconstrução de sinais e de imagens.

A questão central é encontrar valores apropriados para os parâmetros substitutos ξ^2 e $\delta\rho$ para uma dada aplicação. A Tabela 9 resume as modificações do parâmetro de penalidade realizadas nos quatro experimentos apresentados.

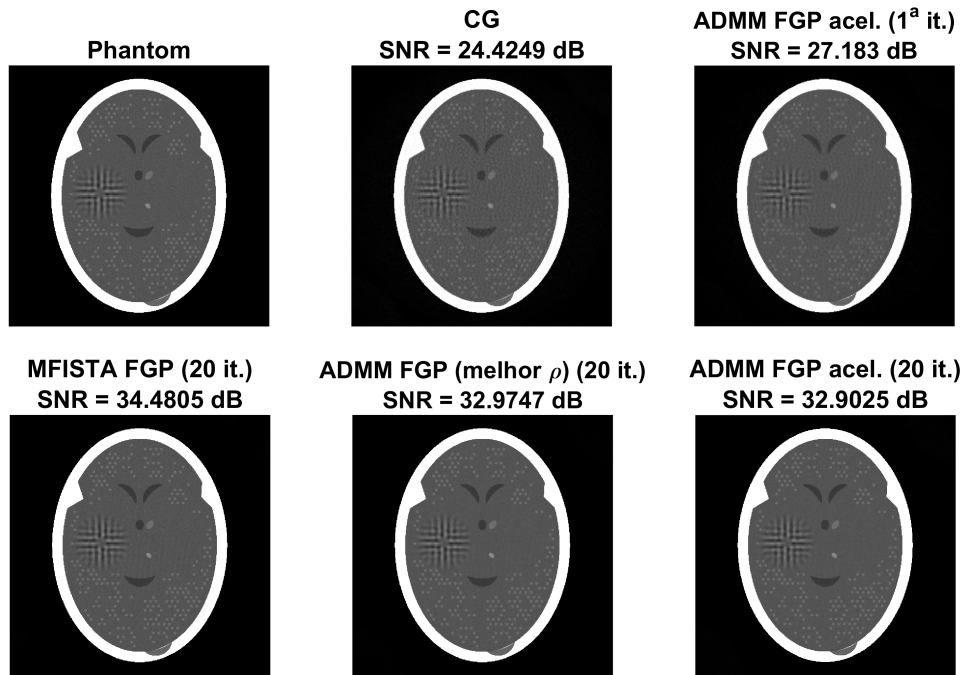


Figura 48: Caso 4 (MRI TV) - Imagens reconstruídas.

Fonte: O autor (2017)

Tabela 9: Modificação do parâmetro de penalidade na primeira iteração do ADMM

Experimento	Termo quadrático (ρ subst. por ξ^2) razão: $r_Q = \xi^2/\rho = \alpha/\beta$	Operador proximal (ρ subst. por $\delta\rho$) razão: $r_P = \delta\rho/\rho = \delta$
Caso 1 - Ultrassom	$r_Q = 0,5000$	$r_P = 3,5$
Caso 2 - Ultrassom	$r_Q = 0,7963$	$r_P = 3,0$
Caso 3 - MRI Wavelet	$r_Q = 0,0129$	$r_P = 3,0$
Caso 4 - MRI com TV	$r_Q = 0,0011$	$r_P = 0,4$

Fonte: O autor (2017)

Em todos os casos, o custo reduzido para a primeira iteração foi alcançado com uma regularização de Tikhonov *menos intensa* no cálculo da inversa para a solução quadrática. Note que todas as razões r_Q na coluna central da tabela são inferiores a 1. Nos casos 3 e 4, em particular, a regularização necessária é muito menor, provavelmente devido a um melhor condicionamento da matriz \mathbf{H} nos experimentos de MRI, em relação aos de ultrassom.

A modificação na escala do operador proximal é muito semelhante nos casos 1, 2 e 3, nos quais o operador de *shrinkage-thresholding* é utilizado. As razões $r_P > 1$ significam que o encolhimento (*shrinkage*) necessário para filtrar a solução quadrática \mathbf{x}_1 é *menos intenso*, e dá

origem a uma primeira solução \mathbf{f}_1 mais ruidosa. No caso 4, o proximal foi implementado pelo algoritmo FGP. Como $r_p < 1$, o valor original de ρ é reduzido, o que *intensifica* a regularização TV.

Nas reconstruções em ultrassom, o esquema de aceleração possibilitou uma **redução de custo de até 24% da faixa**. Em ambos os experimentos, isso equivaleu a uma **primeira iteração do ADMM com aproximadamente o mesmo custo da 6ª iteração do FISTA**.

Para o imageamento MR, embora os ganhos em termos de faixa de custo sejam relativamente pequenos, a aceleração levou a **primeira iteração do ADMM à vizinhança da 7ª e da 9ª iterações das variantes do FISTA avaliadas**.

Como observação final, dado que a primeira iteração pode gerar soluções aproximadas razoáveis em alguns casos, o esquema proposto pode ser útil em problemas de **estimação de sinais em tempo real**, em que o tempo de reconstrução é normalmente crítico. Se armazenarmos previamente a matriz $\mathbf{B} = (\mathbf{H}^T \mathbf{H} + \xi^2 \mathbf{I})^{-1} \mathbf{H}^T$ da Equação (77), podemos obter rapidamente soluções aproximadas para cada vetor de dados \mathbf{g} proveniente de um conjunto de sensores. Isso é feito por uma simples multiplicação matriz-vetor $\mathbf{x}_1 = \mathbf{B}\mathbf{g}$ e pela aplicação do operador proximal para obter $\mathbf{f}_1 = \text{prox}_{\lambda R/(\delta\rho)}(\mathbf{x}_1)$.

No caso $\ell_2\text{-}\ell_1$, em particular, as reconstruções podem ser obtidas em tempos **bastante inferiores** aos apresentados nos experimentos deste capítulo se implementadas em plataformas de processamento paralelo. Isso ocorre porque tanto o produto matriz-vetor quanto o operador de *shrinkage* são operações matemáticas **facilmente paralelizáveis**.

Essa abordagem é especialmente interessante face à exigência de reconstruções em tempo real da Ultrassonografia, e para visualizações rápidas durante procedimentos de varredura no imageamento paralelo de Ressonância Magnética com *Compressed Sensing*.

6 CONCLUSÃO E TRABALHOS FUTUROS

Avaliando os objetivos estabelecidos na Seção 1.1, estabelecemos a seguir algumas conclusões e propostas de trabalhos futuros.

O modelamento apresentado no Capítulo 2 para a reconstrução de imagens em ultrassonografia esparsa pela abordagem IPB, mostrou resultados satisfatórios para os casos de estudo apresentados. A validação do modelo discreto por meio de um conjunto de dados sintético de resultado conhecido permitiu ajustar os parâmetros de um conjunto de algoritmos iterativos. De um modo geral, todos os métodos selecionados obtiveram bom desempenho, dado que se mostraram capazes de realizar as reconstruções com os dados sintéticos e com os dados reais, coletados a partir de um equipamento de ultrassom. Deste modo, consideramos que o primeiro e o segundo objetivos específicos foram atingidos.

Como esperado, alguns algoritmos apresentaram melhor eficiência computacional. As avaliações realizadas no Capítulo 4 confirmaram a tendência observada na literatura, de considerar a família ISTA como uma das melhores opções para a minimização $\ell_2-\ell_1$. Os experimentos desse capítulo também confirmaram que o ADMM é um método rápido, mas bastante sensível à escolha do parâmetro de penalidade. Essas conclusões nos dão base para afirmar que o terceiro objetivo específico foi igualmente atingido.

O bom desempenho do método ADMM nos levou, então, a avaliar alternativas de torná-lo ainda mais rápido. Ainda que a proposta de aceleração não imponha modificações ao longo de todas as iterações, a substituição do parâmetro de penalidade na primeira iteração nos permitiu alcançar uma aceleração substancial na convergência inicial do algoritmo.

Os experimentos apresentados no Capítulo 5 confirmaram o bom desempenho da proposta de aceleração do ADMM, com ganhos expressivos para as reconstruções em ultrassom e em MRI. Além da reconstrução de imagens, o ADMM acelerado pode igualmente propiciar ganhos em outras aplicações de reconstrução de sinais que atendam as condições de $n \geq m$ e $\text{posto}(\mathbf{H}) = m$. Concluimos, portanto, que o último objetivo específico que estabelecemos foi atingido.

O esquema de aceleração proposto também pode ser usado para se obter reconstruções aproximadas rápidas em aplicações em tempo real, nas quais o tempo de reconstrução é crítico.

Como trabalho futuro, sugerimos o aperfeiçoamento do modelo discreto para representar pontos imageados fora da grade. Com a discretização homogênea da região imageada, há perda de informação no modelo relativa a pontos nos interstícios, dado que os objetos imageados são de natureza contínua.

Observamos, ainda, que o modelamento é sensível a variações dos parâmetros de ultrassom, tais como a frequência de pulso, as resoluções horizontal e axial, as taxas de atenuação etc. Com a abordagem apresentada, é necessária a construção de um modelo \mathbf{H} distinto para cada situação, o que limita a aplicação. Modelos representados por funções matemáticas ao invés de matrizes fixas podem ser uma linha de pesquisa interessante para reduzir essa limitação, uma vez que funções são facilmente parametrizáveis.

Os experimentos do Capítulo 5 validaram a proposta de aceleração do ADMM, mas critérios práticos para estimar os novos parâmetros ξ^2 e δ de modo não exaustivo precisam ser investigados.

REFERÊNCIAS

- ANDERSON, M. E.; TRAHEY, G. E. **A Seminar on k-Space Applied to Medical Ultrasound**. [S.l.], 2000.
- AZHARI, H. **Basics of Biomedical Ultrasound for Engineers**. Hoboken, NJ: John Wiley & Sons, Inc., 2010. ISBN 9780470465479.
- BARRETT, H. H.; MYERS, K. J. **Foundations of Image Science**. Hoboken, NJ: Wiley-Interscience, 2004. ISBN 0471153001.
- BECK, A.; TEOULLE, M. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. **SIAM Journal on Imaging Sciences**, v. 2, n. 1, p. 183–202, jan 2009. ISSN 1936-4954.
- BECK, A.; TEOULLE, M. Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems. **IEEE Transactions on Image Processing**, v. 18, n. 11, p. 2419–2434, 2009. ISSN 10577149.
- BJORCK, A. **Numerical Methods for Least Squares Problems**. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 1996. ISBN 0898713609.
- BOVIK, A. **Handbook of Image and Video Processing**. Austin, Texas: Academic Press, 2000. 1–974 p. ISSN 00942405. ISBN 0-471-23766-3.
- BOYD, S. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. **Foundations and Trends in Machine Learning**, v. 3, n. 1, p. 1–122, 2010. ISSN 1935-8237.
- BOYD, S. P.; VANDENBERGHE, L. **Convex Optimization**. Cambridge: Cambridge University Press, 2004. ISBN 978-0-521-83378-3.
- BYRD, R. H.; CHIN, G. M.; NOCEDAL, J.; OZTOPRAK, F. A Family of Second-Order Methods for Convex L1-Regularized Optimization. **Mathematical Programming**, Springer Berlin Heidelberg, v. 159, n. 1-2, p. 435–467, sep 2016. ISSN 0025-5610.
- BYRD, R. H.; NOCEDAL, J.; OZTOPRAK, F. An Inexact Successive Quadratic Approximation Method for L-1 Regularized Optimization. **Mathematical Programming**, p. 1–22, 2015.
- BYRD, R. H.; NOCEDAL, J.; SOLNTSEV, S. An Algorithm for Quadratic L1-Regularized Optimization with a Flexible Active-Set Strategy. **Optimization Methods and Software**, v. 30, n. 6, p. 1213–1237, 2015.
- BYRNE, C. L. **Iterative Optimization in Inverse Problems (Monographs and Research Notes in Mathematics)**. [S.l.]: CRC Press, 2014. ISBN 9781482223333.

CARCREFF, E.; BOURGUIGNON, S.; IDIER, J.; SIMON, L. A Linear Model Approach for Ultrasonic Inverse Problems With Attenuation and Dispersion. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 61, n. 7, p. 1191–1203, 2014.

CHAN, S. H.; WANG, X.; ELGENDY, O. A. Plug-and-Play ADMM for Image Restoration: Fixed Point Convergence and Applications. **IEEE Transactions on Computational Imaging**, v. 3, n. 1, p. 84–97, 2017. ISSN 2333-9403.

CHAUFFERT, N.; WEISS, P.; KAHN, J.; CIUCIU, P. A Projection Algorithm for Gradient Waveforms Design in Magnetic Resonance Imaging. **IEEE Transactions on Medical Imaging**, v. 35, n. 9, p. 2026–2039, sep 2016. ISSN 0278-0062.

CHRISTENSEN, D. **Ultrasonic Bioinstrumentation**. New York: John Wiley & Sons, Inc., 1988.

COMBETTES, P. L.; PESQUET, J.-C. Proximal Splitting Methods in Signal Processing. In: **Fixed-Point Algorithms for Inverse Problems in Science and Engineering**. New York: Springer, 2011. cap. 10, p. 185–212. ISBN 978-1-4419-9568-1.

DAI, Y. H.; YUAN, Y. A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property. **SIAM Journal on Optimization**, v. 10, n. 1, p. 177–182, jan 1999. ISSN 1052-6234.

DAUBECHIES, I.; DEFRISE, M.; MOL, C. D. E. An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint. **Communications on Pure and Applied Mathematics**, LVII, n. 2004, p. 1413–1457, 2004.

DAVIDI, R.; HERMAN, G. T.; LANGTHALER, O.; SARDANA, S.; YE, Z. **SNARK14: A Programming System for the Reconstruction of 2D Images from 1D Projections**. 2017. Disponível em: <<http://dig.cs.gc.cuny.edu/software/snark14/SNARK14.pdf>>.

DEMIRLI, R.; SANIIE, J. A High-Fidelity Time-Frequency Representation for Ultrasonic Signal Analysis. **IEEE Ultrasonics Symposium**, v. 2, p. 1376–1379, 2003.

DENG, L.-J.; GUO, W.; HUANG, T.-Z. Single-Image Super-Resolution via an Iterative Reproducing Kernel Hilbert Space Method. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 26, n. 11, p. 2001–2014, nov 2016. ISSN 1051-8215.

DENG, W.; YIN, W. On the Global and Linear Convergence of the Generalized Alternating Direction Method of Multipliers. **Journal of Scientific Computing**, v. 66, n. 3, p. 889–916, 2016. ISSN 0885-7474.

EFTEKHARI, A.; WAKIN, M. B. Greed Is Super: A New Iterative Method for Super-Resolution. In: **2013 IEEE Global Conference on Signal and Information Processing**. [S.l.]: IEEE, 2013. p. 631–631. ISBN 978-1-4799-0248-4.

ELAD, M. **Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing**. New York: Springer, 2010. ISBN 9781441970107.

ELLIS, M. A.; VIOLA, F.; WALKER, W. F. Super-Resolution Image Reconstruction Using Diffuse Source Models. **Ultrasound in Medicine and Biology**, v. 36, n. 6, p. 967–977, 2010. ISSN 03015629.

FLETCHER, R. **Practical Methods of Optimization vol. 1: Unconstrained Optimization**. New York: John Wiley & Sons, Inc., 1987.

FLETCHER, R.; REEVES, C. M. Function Minimization by Conjugate Gradients. **The Computer Journal**, v. 7, n. 2, p. 149–154, 1964.

FLUKE-BIOMEDICAL. Nuclear Associates: 84-317 and 84-317-700 Multipurpose Tissue/Cyst Ultrasound Phantoms - Users Manual. 2005. Disponível em: <<http://www.flukebiomedical.com/biomedical/user/diagnostic-imaging-qa/ultrasound-qa/84-317-multi-purpose-tissue-cyst-ultrasound-phantom.htm?PID=55394>>.

GABAY, D.; MERCIER, B. A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite Element Approximation. **Computers & Mathematics with Applications**, v. 2, n. 1, p. 17–40, 1976. ISSN 08981221.

GHADIMI, E.; TEIXEIRA, A.; SHAMES, I.; JOHANSSON, M. Optimal Parameter Selection for the Alternating Direction Method of Multipliers (ADMM): Quadratic Problems. **IEEE Transactions on Automatic Control**, v. 60, n. 3, p. 644–658, 2015.

GLOWINSKI, R.; MARROCO, A. Sur l'Approximation, par Eléments Finis d'Ordre Un, et la Résolution, par Pénalisation-Dualité d'une Classe de Problèmes de Dirichlet Non Linéaires. **Revue Française d'Automatique, Informatique, Recherche Opérationnelle - Analyse Numérique**, v. 9, n. 2, p. 41–76, 1975.

GOLDSTEIN, T.; O'DONOGHUE, B.; SETZER, S.; BARANIUK, R. Fast Alternating Direction Optimization Methods. **SIAM Journal on Imaging Sciences**, v. 7, n. 3, p. 1588–1623, jan 2014. ISSN 1936-4954.

GOLUB, G. H.; HEATH, M.; WAHBA, G. Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. **Technometrics**, v. 21, n. 2, p. 215–223, 1979.

HAGER, W. W.; ZHANG, H. A Survey of Nonlinear Conjugate Gradient Methods. **Pacific Journal of Optimization**, v. 2, n. 1, p. 35–58, 2006.

HANSEN, P. C. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. **SIAM Review**, v. 34, n. 4, p. 561–580, 1992.

HANSEN, P. C. **Rank-Deficient and Discrete Ill-Posed Problems - Numerical Aspects of Linear Inversion**. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 1998. ISBN 0898714036.

HEDRICK, W. R.; HYKES, D. L.; STARCHMAN, D. E. **Ultrasound Physics and Instrumentation**. 3rd. ed. St. Louis: Mosby, 1995.

HERMAN, G. T. **Fundamentals of Computerized Tomography: Image Reconstruction from Projections**. 2nd. ed. London: Springer London, 2009.

HESTENES, M. R.; STIEFEL, E. Methods of Conjugate Gradients for Solving Linear Systems. **Journal of Research of the Natural Bureau of Standards (NBS)**, v. 49, n. 6, p. 409–436, 1952.

HOLMES, C.; DRINKWATER, B. W.; WILCOX, P. D. Post-Processing of the Full Matrix of Ultrasonic Transmit-Receive Array Data for Non-Destructive Evaluation. **NDT & E International**, v. 38, n. 8, p. 701–711, dec 2005. ISSN 09638695.

HUANG, C. Aberration Correction for Transcranial Photoacoustic Tomography of Primates Employing Adjunct Image Data. **Journal of Biomedical Optics**, v. 17, n. 6, p. 066016, jun 2012. ISSN 1083-3668.

HUANG, C.; NIE, L.; SCHOONOVER, R. W.; WANG, L. V.; ANASTASIO, M. A. Photoacoustic Computed Tomography Correcting for Heterogeneity and Attenuation. **Journal of Biomedical Optics**, v. 17, n. 6, p. 061211, 2012. ISSN 10833668.

HUANG, C.; WANG, K.; NIE, L.; WANG, L. V.; ANASTASIO, M. A. Full-Wave Iterative Image Reconstruction in Photoacoustic Tomography With Acoustically Inhomogeneous Media. **IEEE Transactions on Medical Imaging**, v. 32, n. 6, p. 1097–1110, jun 2013. ISSN 0278-0062.

JENSEN, J. A. **Linear Description of Ultrasound Imaging Systems**. Lyngby, Denmark, 1999.

JENSEN, J. A.; GANDHI, D.; O'BRIEN, W. D. Ultrasound Fields in an Attenuating Medium. In: LEVY, M.; MCAVOY, B. R. (Ed.). **Proceedings of IEEE Ultrasonics Symposium**. Baltimore, Maryland, USA: IEEE Ultrasonics, Ferroelectric, and Frequency Control Society, 1993. v. 2, p. 943–946. ISBN 0-7803-1278-3.

JENSEN, J. A.; LYNGBY, D.; MEDICAL, P.; ENGINEERING, B.; TECHNOLOGY, I. Field: A Program for Simulating Ultrasound Systems. In: NOUSIAINEN, J.; ESKOLA, H.; MALMIVUO, J. (Ed.). **10th Nordic-Baltic Conference on Biomedical Engineering**. Tampere, Finland: Ragnar Granit Foundation, 1996. v. 34, p. 351–353.

JENSEN, J. A.; SVENDSEN, N. B. Calculation of Pressure Fields from Arbitrarily Shaped, Apodized, and Excited Ultrasound Transducers. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 39, n. 2, p. 262–267, 1992. ISSN 08853010.

JOLY, M. **Introdução à Análise da Imagem**. [S.l.]: Papyrus Editora, 2006.

KELLEY, C. T. **Iterative Methods for Optimization**. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 1999. 1–188 p. ISBN 0898714338.

KIM, D.; FESSLER, J. A. An Optimized First-Order Method for Image Restoration. In: **2015 IEEE International Conference on Image Processing (ICIP)**. [S.l.]: IEEE, 2015. p. 3675–3679. ISBN 978-1-4799-8339-1.

LAVARELLO, R.; KAMALABADI, F.; O'BRIEN, W. D. A Regularized Inverse Approach to Ultrasonic Pulse-Echo Imaging. **IEEE Transactions on Medical Imaging**, v. 25, n. 6, p. 712–722, 2006. ISSN 02780062.

LAWSON, C. L.; HANSON, R. J. **Solving Least Squares Problems**. Philadelphia: Prentice-Hall, 1974.

LEE, J.; SUN, Y.; SAUNDERS, M. **Proximal Newton-Type Methods for Convex Optimization**. Stanford, CA, 2012. Disponível em: <http://www.stanford.edu/class/cme334/docs/2012-12-08-lee_sun_paper.pdf>.

- LINGVALL, F.; OLOFSSON, T. On Time-Domain Model-Based Ultrasonic Array Imaging. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 54, n. 8, p. 1623–1633, 2007.
- LIU, Y.; STOREY, C. Efficient Generalized Conjugate Gradient Algorithms, Part 1: Theory. **Journal of Optimization Theory and Applications**, v. 69, n. 1, p. 129–137, 1991.
- LUENBERGER, D. G.; YE, Y. **Linear and Nonlinear Programming**. 3rd. ed. New York: Springer, 2008. 1–551 p. ISBN 978-0-387-74503-9.
- MANGUEL, A. **Lendo Imagens: Uma História de Amor e Ódio**. São Paulo: Companhia das Letras, 2001.
- MARKS, R. J. **Introduction to Shannon Sampling and Interpolation Theory**. New York, NY: Springer New York, 1991. (Springer Texts in Electrical Engineering). ISBN 978-1-4613-9710-6.
- MEYER, C. D. **Matrix Analysis and Applied Linear Algebra**. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 2000. ISBN 0898714540.
- MONTERO, R. J. L. **Pulse-Echo Image Formation Using Nonquadratic Regularization With Speckle-Based Images**. 73 p. Tese (Master thesis) — University of Illinois, Urbana, 2005.
- MOROZOV, V. On the Solution of Functional Equations by the Method of Regularization. **Soviet Math. Dokl.**, v. 7, p. 414–417, 1966.
- MUCKLEY, M. J.; NOLL, D. C.; FESSLER, J. A. Fast Parallel MR Image Reconstruction via B1-Based , Adaptive Restart , Iterative Soft Thresholding Algorithms (BARISTA). **IEEE Transactions on Medical Imaging**, v. 34, n. 2, p. 578–588, 2015.
- NATTERER, F.; WUBBELING, F. **Mathematical Methods in Image Reconstruction**. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 2001. ISBN 0898714729.
- NESTEROV, Y. A Method for Solving the Convex Programming Problem with Convergence Rate $O(1/k^2)$. **Dokl. Akad. Nauk SSSR**, v. 269, p. 543–547, 1983.
- NESTEROV, Y. Smooth Minimization of Non-Smooth Functions. **Mathematical Programming**, v. 103, n. 1, p. 127–152, 2005. ISSN 00255610.
- O'BRIEN, M. S.; SINCLAIR, A. N.; KRAMER, S. M. Recovery of a Sparse Spike Time Series by L1 Norm Deconvolution. **IEEE Transactions on Signal Processing**, v. 42, n. 12, p. 3353–3365, 1994. ISSN 1053587X.
- OSHER, S.; LI, Y. Coordinate Descent Optimization for L1 Minimization with Application to Compressed Sensing; a Greedy Algorithm. **Inverse Problems and Imaging**, v. 3, n. 3, p. 487–503, 2009. ISSN 1930-8337.
- PARIKH, N.; BOYD, S. Proximal Algorithms. **Foundations and Trends in Optimization**, v. 1, n. 3, p. 123–231, 2013. ISSN 2167-3888.
- PENG, Z.; XU, Z.; HUANG, J. RSPiRiT: Robust Self-Consistent Parallel Imaging Reconstruction Based on Generalized LASSO. In: **2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)**. [S.l.]: IEEE, 2016. p. 318–321. ISBN 978-1-4799-2349-6.

POLAK, E.; RIBIÈRE, G. Note Sur La Convergence de Méthodes de Directions Conjuguées. **Revue Française d'Informatique et de Recherche Opérationnelle**, v. 3, n. 1, p. 35–43, 1969.

POLYAK, B. The Conjugate Gradient Method in Extremal Problems. **USSR Computational Mathematics and Mathematical Physics**, v. 9, n. 4, p. 94–112, jan 1969. ISSN 00415553.

REGINSKA, T. A Regularization Parameter in Discrete Ill-Posed Problems. **SIAM Journal on Scientific Computing**, v. 17, n. 3, p. 740–749, 1996.

SAWATZKY, A.; XU, Q.; SCHIRRA, C. O.; ANASTASIO, M. A. Proximal ADMM for Multi-Channel Image Reconstruction in Spectral X-ray CT. **IEEE Transactions on Medical Imaging**, v. 33, n. 8, p. 1637–1668, 2014. ISSN 1558254X.

SHEN, J.; EBBINI, E. A New Coded-Excitation Ultrasound Imaging System. I. Basic Principles. **IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control**, v. 43, n. 1, p. 131–140, jan 1996. ISSN 0885-3010.

SHEWCHUK, J. R. An Introduction to the Conjugate Gradient Method Without the Agonizing Pain. 1994.

STEPANISHEN, P. R. An Approach to Computing Time-Dependent Interaction Forces and Mutual Radiation Impedances Between Pistons in a Rigid Planar Baffle. **The Journal of the Acoustical Society of America**, v. 49, n. 1B, p. 283–292, 1971. ISSN 00014966.

STEPANISHEN, P. R. The Time-Dependent Force and Radiation Impedance on a Piston in a Rigid Infinite Planar Baffle. **The Journal of the Acoustical Society of America**, v. 49, n. 3B, p. 841–849, 1971. ISSN 00014966.

STEPANISHEN, P. R. Transient Radiation from Pistons in an Infinite Planar Baffle. **The Journal of the Acoustical Society of America**, v. 49, n. 5B, p. 1629–1638, 1971. ISSN 00014966.

STEPANISHEN, P. R. Pulsed Transmit/Receive Response of Ultrasonic Piezoelectric Transducers. **The Journal of the Acoustical Society of America**, v. 69, n. 6, p. 1815–1827, 1981. ISSN 00014966.

VALENTE, S.; ZIBETTI, M.; PIPA, D.; MAIA, J.; SCHNEIDER, F. An Assessment of Iterative Reconstruction Methods for Sparse Ultrasound Imaging. **Sensors**, v. 17, n. 3, p. 533, 2017. ISSN 1424-8220.

VERASONICS. Vantage Family Specifications. 2017. Disponível em: <http://verasonics.com/wp-content/uploads/2017/03/Vantage_Family_Brochure_Mar2017.pdf>.

VIOLA, F.; ELLIS, M. A.; WALKER, W. F. Time-Domain Optimized Near-Field Estimator for Ultrasound Imaging: Initial Development and Results. **IEEE Transactions on Medical Imaging**, v. 27, n. 1, p. 99–110, 2008. ISSN 02780062.

VOGEL, C. R. **Computational Methods for Inverse Problems**. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 2002. 1–200 p. ISBN 0898715075.

WEBB, A. **Introduction to Biomedical Imaging**. [S.l.: s.n.], 2002. ISBN 0471237663.

WEN, Z.; YIN, W.; GOLDFARB, D.; ZHANG, Y. A Fast Algorithm for Sparse Reconstruction Based on Shrinkage, Subspace Optimization, and Continuation. **SIAM Journal on Scientific Computing**, v. 32, n. 4, p. 1832–1857, 2010. ISSN 1064-8275.

WU, H.; CHEN, J.; WU, S.; JIN, H.; YANG, K. A Model-Based Regularized Inverse Method for Ultrasonic B-Scan Image Reconstruction. **Measurement Science and Technology**, IOP Publishing, v. 26, n. 10, p. 105401, 2015. ISSN 0957-0233.

XU, Z.; FIGUEIREDO, M. A. T.; GOLDSTEIN, T. Adaptive ADMM with Spectral Penalty Parameter Selection. In: **Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)**. Fort Lauderdale, Florida, USA: JMLR: W&CP, 2017. v. 54, n. 1.

YAMAGISHI, M.; YAMADA, I. Over-Relaxation of the Fast Iterative Shrinkage-Thresholding Algorithm with Variable Stepsize. **Inverse Problems**, v. 27, n. 10, p. 105008, 2011. ISSN 0266-5611.

YORK, G.; KIM, Y. Ultrasound Processing and Computing: Review and Future Directions. **Annu. Rev. Biomed. Eng.**, v. 1, p. 559–588, 1999.

ZANIN, L. G. S.; ZIBETTI, M. V. W.; SCHNEIDER, F. K. Conjugate Gradient and Regularized Inverse Problem-Based Solutions Applied to Ultrasound Image Reconstruction. **IEEE International Ultrasonics Symposium**, IEEE, n. 1, p. 377–380, 2011.

ZHANG, S.; XIA, Y.; ZOU, C. Comparison of Sparse-View CT Image Reconstruction Algorithms. In: **2016 International Conference on Audio, Language and Image Processing (ICALIP)**. [S.l.]: IEEE, 2016. p. 385–390. ISBN 978-1-5090-0654-0.

ZIBETTI, M. V. W.; BAZÁN, F. S. V.; MAYER, J. Determining the Regularization Parameters for Super-Resolution Problems. **Signal Processing**, v. 88, n. 12, p. 2890–2901, dec 2008. ISSN 01651684.

ZIBETTI, M. V. W.; HELOU, E. S.; MIGUELES, E. X.; De Pierro, A. R. Accelerating the Over-Relaxed Iterative Shrinkage-Thresholding Algorithms with Fast and Exact Line Search for High Resolution Tomographic Image Reconstruction. In: **2015 IEEE International Conference on Image Processing (ICIP)**. Quebec: IEEE, 2015. p. 2305–2308. ISBN 978-1-4799-8339-1.

ZIBETTI, M. V. W.; PIPA, D. R.; De Pierro, A. R. Fast and Exact Unidimensional L2-L1 Optimization as an Accelerator for Iterative Reconstruction Algorithms. **Elsevier Digital Signal Processing**, p. 1–9, 2015.

ZIBULEVSKY, M.; ELAD, M. L1-L2 Optimization in Signal and Image Processing - Iterative shrinkage and beyond. **IEEE Signal Processing Magazine**, n. May, p. 76–88, 2010.

