

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**CONSTRUÇÃO E APLICAÇÃO DE ATLAS DE
PONTOS SALIENTES 3D NA INICIALIZAÇÃO
DE MODELOS GEOMÉTRICOS
DEFORMÁVEIS EM IMAGENS DE
RESSONÂNCIA MAGNÉTICA**

CARLOS HENRIQUE VILLA PINTO

ORIENTADOR: PROF. DR. RICARDO JOSÉ FERRARI

São Carlos – SP

Março/2016

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**CONSTRUÇÃO E APLICAÇÃO DE ATLAS DE
PONTOS SALIENTES 3D NA INICIALIZAÇÃO
DE MODELOS GEOMÉTRICOS
DEFORMÁVEIS EM IMAGENS DE
RESSONÂNCIA MAGNÉTICA**

CARLOS HENRIQUE VILLA PINTO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação, área de concentração: Processamento de Imagens e Sinais

Orientador: Prof. Dr. Ricardo José Ferrari

São Carlos – SP

Março/2016

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária UFSCar
Processamento Técnico
com os dados fornecidos pelo(a) autor(a)

P659c Pinto, Carlos Henrique Villa
Construção e aplicação de atlas de pontos salientes
3D na inicialização de modelos geométricos
deformáveis em imagens de ressonância magnética /
Carlos Henrique Villa Pinto. -- São Carlos : UFSCar,
2016.
120 p.

Dissertação (Mestrado) -- Universidade Federal de
São Carlos, 2016.

1. Ressonância magnética. 2. Modelos geométricos
deformáveis. 3. Construção de atlas. 4. Detecção de
pontos salientes 3D. 5. Segmentação de imagens
médicas. I. Título.



UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a defesa de dissertação de mestrado do candidato Carlos Henrique Villa Pinto, realizada em 10/03/2016.

Prof. Dr. Ricardo José Ferrari
(UFSCar)

Prof. Dr. Marcelo Andrade da Costa Vieira
(USP)

Prof. Dr. Hélio Pedrini
(UNICAMP)

Certifico que a sessão de defesa foi realizada com a participação à distância do membro Prof. Dr. Prof. Dr. Hélio Pedrini e, depois das arguições e deliberações realizadas, o participante à distância está de acordo com o conteúdo do parecer da comissão examinadora redigido no relatório de defesa do aluno Carlos Henrique Villa Pinto.

Prof. Dr. Ricardo José Ferrari
Presidente da Comissão Examinadora
(UFSCar)

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Ricardo José Ferrari, por todos os seus valiosos ensinamentos e conselhos, pela paciência, por estar sempre presente e disposto a ajudar, e também por inspirar seus alunos com sua infindável dedicação e gosto pela pesquisa.

A todos os colegas do BIP-Group e do laboratório GAPIS, por todos os momentos de descontração e por toda a ajuda dada durante a elaboração deste trabalho.

Às agências CAPES e FAPESP (processos nº 2014/11988-0 e 2015/02232-1), pelo apoio financeiro que viabilizou minha participação neste projeto de pesquisa.

E aos meus pais, José Antônio e Maria Rita, por todo o suporte que somente um pai e uma mãe podem prover.

RESUMO

O imageamento por ressonância magnética (RM) tornou-se uma ferramenta indispensável no diagnóstico e estudo de diversas doenças e síndromes do sistema nervoso central, tais como a doença de Alzheimer (DA). Para que se possa realizar o diagnóstico preciso de uma doença, bem como o acompanhamento evolutivo de um determinado tratamento, o médico neurorradiologista frequentemente precisa medir e avaliar alterações de volume e forma em determinadas estruturas do cérebro ao longo de uma série de imagens de RM. Para isso, a delimitação prévia das estruturas de interesse nas imagens é necessária. Em geral, essa tarefa é realizada manualmente, com ajuda limitada de um computador, e portanto possui diversos problemas. Por esse motivo, vários pesquisadores têm voltado seus esforços para o desenvolvimento de técnicas automáticas de segmentação de estruturas cerebrais em imagens de RM. Dentre as várias abordagens propostas na literatura, técnicas baseadas em modelos deformáveis e atlas anatômicos estão entre as que apresentam os melhores resultados. No entanto, uma das principais dificuldades na aplicação de modelos geométricos deformáveis é o posicionamento inicial do modelo. Assim, esta pesquisa teve por objetivo desenvolver um atlas de pontos salientes 3D (automaticamente detectados em um conjunto de imagens de RM) e investigar a aplicabilidade de tal atlas em guiar o posicionamento inicial de modelos geométricos deformáveis representando estruturas cerebrais, com o propósito de auxiliar a segmentação automática de tais estruturas em imagens de RM. O arcabouço de processamento incluiu o uso de um detector de pontos salientes 3D baseado na medida de congruência de fase, uma adaptação da técnica *shape contexts* para a criação de descritores de pontos e a estimação de uma transformação *B-spline* para mapear pares de pontos correspondentes. Os resultados, avaliados com as métricas Jaccard e Dice antes e após a inicialização dos modelos, mostraram um ganho significativo em testes envolvendo imagens sinteticamente deformadas de pacientes normais, mas em imagens de pacientes clínicos com DA o ganho foi marginal e ainda pode ser melhorado em pesquisas futuras. Algumas maneiras de se realizar tais melhorias são discutidas neste trabalho.

Palavras-chave: ressonância magnética, modelos geométricos deformáveis, construção de atlas, detecção de pontos salientes 3D, segmentação de imagens médicas.

ABSTRACT

The magnetic resonance (MR) imaging has become an indispensable tool for the diagnosis and study of various diseases and syndromes of the central nervous system, such as Alzheimer's disease (AD). In order to perform the precise diagnosis of a disease, as well as the evolutionary monitoring of a certain treatment, the neuroradiologist doctor often needs to measure and assess volume and shape changes in certain brain structures along a series of MR images. For that, the previous delineation of the structures of interest is necessary. In general, such task is manually done, with limited help from a computer, and therefore it has several problems. For this reason, many researchers have turned their efforts towards the development of automatic techniques for segmentation of brain structures in MR images. Among the various approaches proposed in the literature, techniques based on deformable models and anatomical atlases are among those which present the best results. However, one of the main difficulties in applying geometric deformable models is the initial positioning of the model. Thus, this research aimed to develop an atlas of 3D salient points (automatically detected from a set of MR images) and to investigate the applicability of such atlas in guiding the initial positioning of geometric deformable models representing brain structures, with the purpose of helping the automatic segmentation of such structures in MR images. The processing pipeline included the use of a 3D salient point detector based on the phase congruency measure, an adaptation of the shape contexts technique to create point descriptors and the estimation of a B-spline transform to map pairs of matching points. The results, evaluated using the Jaccard and Dice metrics before and after the model initializations, showed a significant gain in the tests involving synthetically deformed images of normal patients, but for images of clinical patients with AD the gain was marginal and can still be improved in future researches. Some ways to do such improvements are discussed in this work.

Keywords: magnetic resonance, deformable geometric models, atlas construction, 3D salient point detection, medical image segmentation.

LISTA DE FIGURAS

- 2.1 Representação gráfica da distribuição dos momentos magnéticos dos prótons de hidrogênio em uma amostra qualquer de tecido (a) em condições usuais; e (b) sob ação do campo magnético \vec{B}_0 . Fonte: adaptada de Bushberg et al. (2011). 22
- 2.2 Produção de um SIL a partir da aplicação de um pulso de RF de 90°. O SIL apresenta variações positivas e negativas oscilando na frequência de Larmor, e decai com o tempo devido à posterior perda gradual de coerência de fase entre os momentos magnéticos no plano transversal. Fonte: adaptada de Bushberg et al. (2011). 22
- 2.3 Retorno gradual do componente de magnetização no eixo longitudinal (μ_Z) após a aplicação de um pulso de RF de 90°. O tempo $t = T1$ é o momento em que se tem $\mu_Z = 0,63\mu_Z^{(0)}$. Fonte: adaptada de Bushberg et al. (2011). 23
- 2.4 Decaimento gradual do componente de magnetização no plano transversal (μ_{XY}) após a aplicação de um pulso de RF de 90°. O tempo $t = T2$ é o momento em que se tem $\mu_{XY} = 0,37\mu_{XY}^{(0)}$. Fonte: adaptada de Bushberg et al. (2011). 24
- 2.5 Relação entre (a) o espaço K e (b) a imagem de RM formada, em termos da TF. Os valores nas regiões centrais do espaço K correspondem às densidades de baixas frequências (intensidade constante ou pouco variante na imagem de RM), enquanto os valores mais distantes do centro representam as frequências mais altas (detalhes e bordas na imagem de RM). Fonte: adaptada de Mazzola (2009). 26
- 2.6 Comparação entre as ponderações (a) T1-w, (b) T2-w, e (c) PD-w na imagem de um mesmo paciente. 27
- 2.7 Sistemas de coordenadas considerados no contexto do imageamento por RM: coordenadas (a) do mundo; (b) anatômicas; (c) da imagem. Fonte: SlicerWiki (2014). 30

4.1	Etapas para aplicação do atlas no auxílio à segmentação de estruturas cerebrais por modelos geométricos (malhas) deformáveis.	60
4.2	Distribuição etária dos pacientes nas imagens em T1-w da base de dados IXI.	62
4.3	Distribuição etária dos pacientes nas imagens da base de dados ADNI.	63
4.4	Influência das heterogeneidades locais de intensidade, causadas pelo <i>bias field</i> , na detecção de bordas em imagens. (a) Exemplo de <i>bias field</i> . (b) Magnitudes dos gradientes de imagem na ausência e (c) na presença de <i>bias field</i> . (d) Respostas de PC na ausência e (e) na presença de <i>bias field</i>	72
4.5	Exemplos de banco de filtros 3D de log-Gabor amostrados (a) uniformemente; (b) não-uniformemente no domínio da frequência.	74
4.6	Distribuição espacial das classes do histograma de descritor, para $\eta_r = 5$, $\eta_a = 12$ e $\eta_e = 1$ (caso 2D). A distância χ^2 entre os descritores (a) e (b) é de 0,05. Entre os descritores (a) e (c), a distância χ^2 vale 1,04.	79
4.7	Etapas para a construção dos atlas de pontos salientes 3D.	80
4.8	Exemplo de imagem sintética 3D usada para testar a técnica de detecção de pontos salientes. Projeções 2D de imagens representando a mesma estrutura 3D (a) sem <i>bias field</i> ou ruído; e (b) com <i>bias field</i> ($BF = 40\%$) e ruído Gaussiano ($\sigma_{ruído} = 5\%$). (c) Renderização 3D da estrutura, destacando a localização esperada para um ponto saliente no local em que a curvatura 3D da estrutura é máxima.	86
5.1	Resultado visual da redução de ruído em uma imagem de treinamento, nas vistas axial, sagital e coronal.	90
5.2	Resultado da correção do <i>bias field</i> em uma imagem de treinamento, nas vistas axial, sagital e coronal.	91
5.3	Resultado do alinhamento espacial (corregistro) entre uma imagem de treinamento e a imagem de referência, ilustrado pelo padrão “tabuleiro de xadrez”, nas vistas axial, sagital e coronal.	92
5.4	Resultado da extração da máscara do encéfalo em uma imagem de treinamento, utilizando a ferramenta ROBEX. Sobreposição da máscara obtida (em tons de cinza mais claros) sobre a imagem na qual a extração foi feita, nas vistas axial, sagital e coronal.	93

5.5	Resultados obtidos com a técnica de detecção de pontos salientes nos testes com imagens sintéticas de estruturas 3D de alta curvatura. Cada uma das 18 “caixas” representa um grupo de 105 imagens sintéticas degradadas por uma determinada quantidade de ruído e <i>bias field</i> . Os círculos indicam a média e as linhas horizontais mais espessas indicam a mediana. Os pontos cinzas isolados são os valores atípicos (<i>outliers</i>).	94
5.6	Valores individuais e valor médio de Medida-F1 obtidos com a técnica de descrição e casamento de pontos salientes baseada em <i>shape contexts</i> , envolvendo a imagem de RM de referência e as imagens de RM sinteticamente deformadas.	97
5.7	Sobreposição da densidade de distribuição de localização dos pontos salientes, estimada no atlas probabilístico, com a imagem de referência.	99
5.8	Renderização 3D da densidade de distribuição de localização dos pontos salientes, estimada no atlas probabilístico.	100
5.9	Sobreposição dos pontos de alta estabilidade, estimados no atlas de médias e frequências, com a imagem de referência.	101
5.10	Valores das métricas JC e DC obtidos para cada uma das três estruturas cerebrais testadas (CC, LH e RH) antes e depois da inicialização dos modelos com a abordagem desenvolvida neste trabalho, para o teste envolvendo imagens sinteticamente deformadas como imagens de entrada. Os círculos indicam a média e as linhas mais espessas indicam a mediana.	102
5.11	Valores das métricas JC e DC obtidos para cada uma das duas estruturas cerebrais testadas (LH e RH) antes e depois da inicialização dos modelos com a abordagem desenvolvida neste trabalho, para o teste envolvendo imagens clínicas da base EADC-ADNI como imagens de entrada. Os círculos indicam a média e as linhas horizontais mais espessas indicam a mediana.	103

LISTA DE TABELAS

2.1	Tempos de repetição (TR) e de eco (TE) utilizados para se obter as ponderações T1-w, T2-w e PD-w na sequência <i>spin-echo</i>	28
4.1	Características das imagens de RM da base de dados “IXI”.	61
5.1	Média e desvio padrão da contagem de falsos positivos (pontos salientes excessivos) detectados nas imagens sintéticas de estruturas 3D de alta curvatura. Cada uma das 18 células da tabela representa um grupo de 105 imagens sintéticas degradadas por uma determinada quantidade de ruído e <i>bias field</i>	95
5.2	Contagem das quantidades de pontos salientes detectados em cada uma das 18 imagens obtidas da base de dados <i>BrainWeb</i> , com seus respectivos níveis de ruído e <i>bias field</i>	95

LISTA DE ABREVIATURAS E SIGLAS

ADNI	Alzheimer's Disease Neuroimaging Initiative
ASM	Active Shape Model
CBCT	Cone Beam Computerized Tomography
CC	Corpo caloso
CCL	Comprometimento Cognitivo Leve
CDR	Clinical Dementia Rating
CT	Computerized Tomography
DA	Doença de Alzheimer
DC	Dice Coefficient
DGHT	Discriminative Generalized Hough Transform
DMC	Discriminative Model Combination
DoG	Difference of Gaussians
DoH	Determinant of Hessian
EADC	European Alzheimer's Disease Consortium
EM	Expectation-Maximization
FCCF	Feature Cross Correlation Function
FDG	F-fluorodeoxiglicose
FLAIR	Fluid Attenuated Inversion Recovery
FOV	Field of View

GH	Guy's Hospital
GHT	Generalized Hough Transform
GVF	Gradient Vector Flow
HD	Hausdorff Distance
HF	Hough Forests
HH	Hammersmith Hospital
IBSR	Internet Brain Segmentation Repository
IoP	Institute of Psychiatry
ITK	Insight Toolkit
IXI	Information eXtraction from Images
JC	Jaccard Coefficient
JS	Divergência de Jensen-Shannon
k-NN	k-Nearest Neighbors
KFM	Key Features Model
LH	Left hippocampus (hipocampo esquerdo)
MAP	Máxima a Posteriori
MBMRIA	Model Based Magnetic Resonance Image Alignment
MIUE	Model-based Image Understanding Environment
MPW	Manifold Parzen Windows
N3	Nonparametric Nonuniform Intensity Normalization
NAC	Neuroimage Analysis Center
NAC	Neuroimage Analysis Center
nD-SIFT	n-Dimensional Scale Invariant Feature Transform
NLM	Non-Local Means

PATS	Probabilistic Atlas in Template Space
PATSAS	Probabilistic Atlas in Template Space with Atlas Selection
PAUSAS	Probabilistic Atlas in Unseen data Space with Atlas Selection
PC	Phase congruency
PCA	Principal Component Analysis
PD-w	Proton Density-weighted
PDM	Point Distribution Model
PET	Positron Emission Tomography
PVC	Partial Volume Classifier
PW	Parzen Windows
RF	Radiofrequência
RH	Right hippocampus (hipocampo direito)
RM	Ressonância Magnética
SIFT	Scale Invariant Feature Transform
SIL	Sinal de Indução Livre
SPECT	Single-Photon Emission Computerized Tomography
SSM	Statistical Shape Model
T1-w	T1-weighted
T2-w	T2-weighted
TE	Tempo de Eco
TF	Transformada de Fourier
TPS	Thin-Plate Spline
TR	Tempo de Repetição

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO	16
1.1 Motivação	16
1.2 Objetivos	18
1.2.1 Objetivos específicos	18
1.3 Organização do trabalho	19
CAPÍTULO 2 – INFORMAÇÕES COMPLEMENTARES	20
2.1 Imageamento por ressonância magnética	20
2.1.1 Ressonância magnética nuclear	21
2.1.2 Formação da imagem	24
2.1.3 Características de imagem	26
2.1.3.1 Modalidade de ponderação	27
2.1.3.2 Resolução e dimensões	27
2.1.3.3 Sistemas de coordenadas	29
2.1.3.4 <i>Bias field</i>	30
2.1.3.5 Ruído	31
2.2 Doença de Alzheimer	31
2.2.1 Sintomas	32
2.2.2 Causas	33
2.2.3 Diagnóstico	33

2.3	Considerações finais	34
CAPÍTULO 3 – REVISÃO DO ESTADO DA ARTE		36
3.1	Visão geral	36
3.2	Detecção de pontos salientes 3D	36
3.2.1	Trabalhos correlatos	38
3.2.2	Discussão	43
3.3	Construção de atlas do cérebro	44
3.3.1	Trabalhos correlatos	44
3.3.2	Discussão	51
3.4	Inicialização de modelos deformáveis	52
3.4.1	Trabalhos correlatos	52
3.4.2	Discussão	57
3.5	Considerações finais	57
CAPÍTULO 4 – MATERIAIS E MÉTODOS		58
4.1	Visão geral	58
4.2	Bases de dados	60
4.2.1	NAC Brain Atlas	60
4.2.2	IXI	61
4.2.3	ADNI e EADC-ADNI	62
4.2.4	BrainWeb	63
4.2.5	Imagens 3D sintéticas	63
4.3	Pré-processamento	64
4.3.1	Redução de ruído	64
4.3.2	Correção de <i>bias field</i>	65
4.3.3	Padronização de histogramas	67

4.3.4	Alinhamento espacial (corregistro)	68
4.3.5	Extração de máscaras do encéfalo	69
4.4	Detecção de pontos salientes 3D	69
4.4.1	Congruência de fase 3D	70
4.4.2	Banco de filtros de log-Gabor	72
4.4.3	Momentos de segunda ordem	73
4.5	Descrição e casamento de pontos salientes 3D	75
4.5.1	Método 1	75
4.5.2	Método 2	76
4.5.3	Método 3	77
4.6	Atlas de pontos salientes 3D	80
4.6.1	Atlas probabilístico	80
4.6.2	Atlas de médias e frequências	82
4.7	Posicionamento e deformação dos modelos	83
4.8	Análise dos resultados	85
4.9	Considerações finais	88

CAPÍTULO 5 – RESULTADOS E DISCUSSÃO **89**

5.1	Pré-processamento	89
5.1.1	Redução de ruído	89
5.1.2	Correção de <i>bias field</i>	90
5.1.3	Alinhamento espacial (corregistro)	90
5.1.4	Extração de máscaras do encéfalo	90
5.2	Detecção de pontos salientes 3D	92
5.3	Descrição e casamento de pontos salientes 3D	95
5.3.1	Métodos 1 e 2	95
5.3.2	Método 3 (<i>shape contexts</i>)	96

5.4	Atlas de pontos salientes 3D	97
5.4.1	Atlas probabilístico	97
5.4.2	Atlas de médias e frequências	98
5.5	Posicionamento e ajuste dos modelos	98
5.5.1	Testes com imagens sinteticamente deformadas	98
5.5.2	Testes com imagens clínicas	100
5.5.3	Discussão	102
5.6	Considerações finais	105
CAPÍTULO 6 – CONCLUSÕES		106
6.1	Conclusões do trabalho	106
6.2	Contribuições	107
6.3	Ideias para trabalhos futuros	108
6.4	Produção científica relacionada a esta pesquisa	108
REFERÊNCIAS		110

Capítulo 1

INTRODUÇÃO

Este capítulo apresenta a motivação por trás deste trabalho de pesquisa, descrevendo o problema que foi abordado e os objetivos esperados. Serve como introdução ao conteúdo apresentado nos capítulos seguintes.

1.1 Motivação

Devido ao seu alto grau de detalhamento anatômico, o imageamento por ressonância magnética (RM) tornou-se uma ferramenta indispensável no diagnóstico e estudo de diversas doenças e síndromes do sistema nervoso central, como, por exemplo, a doença de Alzheimer (DA). O diagnóstico preciso de uma doença, bem como o acompanhamento evolutivo de um determinado tratamento, frequentemente exige que o médico neurorradiologista meça e avalie alterações de volume e forma de determinadas estruturas do cérebro ao longo de uma série de imagens de RM. Como exemplo, pode-se mencionar as alterações volumétricas e formais que ocorrem no hipocampo e no corpo caloso de pacientes portadores da DA (TEIPEL et al., 2002; SHEN et al., 2012).

O método convencional de medição volumétrica de estruturas do cérebro em imagens de RM ainda é o delineamento manual (ou semiautomático) das regiões ou estruturas de interesse (KELLER; ROBERTS, 2009). Esse processo é realizado por um especialista com ajuda limitada de um computador. Entretanto, tal procedimento, além de consumir muito tempo, é custoso e propenso à grande variabilidade intra e inter observadores. Além disso, com o crescente uso do imageamento por RM, o número gerado de imagens que precisa ser analisado também aumenta rapidamente, criando uma alta demanda por um número maior de neurorradiologistas e também por ferramentas computacionais automáticas que os auxiliem. Essas ferramentas devem, primeiramente, realizar a segmentação das estruturas de interesse na imagem para que

as medições e análises posteriores possam ser efetuadas.

Na última década, vários pesquisadores têm voltado seus esforços para o desenvolvimento de técnicas automáticas para a segmentação de tecidos e estruturas cerebrais em imagens de RM (KLAUSCHEN et al., 2009; BALAFAR et al., 2010; LIU et al., 2014). Tais ferramentas utilizam técnicas de processamento de imagens e reconhecimento de padrões para delinear as bordas ou regiões dos principais tecidos e subestruturas cerebrais, que são posteriormente analisadas isoladamente para o diagnóstico ou acompanhamento de uma determinada patologia. Dentre as várias abordagens propostas na literatura, técnicas de segmentação baseadas em modelos deformáveis e atlas anatômicos probabilísticos e topológicos (LIJN et al., 2008; CHUPIN et al., 2007; CARMICHAEL et al., 2005; ANBECK et al., 2004) estão entre algumas das que apresentam os melhores resultados, pois possibilitam a utilização da informação anatômica intrinsecamente contida nos modelos durante o processo de segmentação.

Modelos geométricos deformáveis têm sido utilizados com grande sucesso em várias aplicações de processamento de imagens nas últimas décadas, principalmente na área médica (LAI, 1994; MCINERNEY; TERZOPOULOS, 1996, 1999; MONTAGNAT; DELINGETTE; AYACHE, 2001; WANG et al., 2005a, 2005b; LEAVENS et al., 2008; KAUS; BROCK, 2009; HEGADI; KOP; HANGARGE, 2010; FERRARI et al., 2011). O principal atrativo desse tipo de abordagem em segmentação de imagens médicas é o fato de tais modelos trazerem em si a topologia do órgão anatômico que representam. Dessa forma, essa abordagem torna-se menos suscetível a ruído e pode ser aplicada inclusive em situações nas quais o nível de contraste da anatomia em estudo na imagem é demasiadamente baixo (PEKAR; MCNUTT; KAUS, 2004; KAUS; BROCK, 2009). Técnicas mais sofisticadas embutem conhecimento a priori sobre as regiões vizinhas da estrutura analisada, possibilitando considerar restrições espaciais adaptativas na deformação do modelo (LIU; IGLESIAS; TU, 2013), ou ainda melhorar a adaptação do modelo utilizando atlas obtidos de uma população rotulada de imagens (LINGURARU et al., 2007; GHOSE et al., 2012).

Uma das principais dificuldades na aplicação de modelos geométricos deformáveis para segmentação automática de imagens médicas é o posicionamento inicial adequado do modelo. Para o sucesso da segmentação, o modelo deve ser posicionado em uma região muito próxima da estrutura que se quer segmentar na imagem, algo que pode ser problemático, especialmente para imagens contendo um grande número de estruturas próximas umas das outras (BOGOVIC; PRINCE; BAZIN, 2013). Nesse contexto, alguns trabalhos presentes na literatura propõem o uso de pontos salientes¹ para servirem como âncoras durante o posicionamento dos modelos (FER-

¹O termo “ponto saliente” refere-se a um ponto da imagem que se destaca localmente por apresentar alguma característica distintiva específica (por exemplo, alto grau de curvatura), sendo muitas vezes um ponto óbvio para a visão humana. Ele não necessariamente representa um ponto de referência anatômica que seria marcado por um

RARI et al., 2011; QAZI et al., 2011; ZHANG et al., 2014). No entanto, em geral esse procedimento é realizado de maneira semiautomática usando marcações manuais na imagem, o que pode ser uma tarefa proibitiva se considerarmos o elevado número de imagens de RM que precisam ser processadas.

1.2 **Objetivos**

Com base no que foi exposto, o objetivo inicial deste trabalho de pesquisa foi criar um atlas probabilístico de pontos salientes 3D automaticamente detectados em um conjunto de imagens 3D de RM, e então investigar a aplicabilidade de tal atlas no processo de posicionamento inicial de modelos geométricos deformáveis para fins de segmentação automática de estruturas cerebrais. Até o momento da escrita deste texto, não se teve conhecimento de outro trabalho proposto na literatura científica que utilizasse a mesma abordagem para esse fim.

É importante frisar que o ponto de foco deste projeto restringiu-se à inicialização do modelo, e não em realizar a segmentação final, que envolveria a elaboração de funcionais de energia e estratégias para a deformação do modelo. Porém, é obvio que quanto mais esta puder se aproximar do resultado final da segmentação, melhor.

1.2.1 **Objetivos específicos**

Adicionalmente, os seguintes objetivos específicos foram estabelecidos para possibilitar o desenvolvimento deste trabalho:

- Organização e pré-processamento de um conjunto de imagens de RM de pacientes normais (sem patologia diagnosticada). Estabeleceu-se que as imagens seriam obtidas do projeto *Brain Development*² e o pré-processamento foi composto pela aplicação de técnicas de redução de ruído, correção da heterogeneidade de intensidade (*bias field*), padronização de histograma, correção das imagens com uma imagem de referência (*template*) e obtenção de máscaras binárias do encéfalo.
- Obtenção e pré-processamento de modelos geométricos (aqui representados por malhas triangulares 3D) das estruturas anatômicas de interesse. Os modelos disponíveis na base de dados *Neuroimage Analysis Center* (NAC) (HALLE et al., 2013) foram utilizados para realização desta tarefa.

especialista - embora isso muitas vezes aconteça.

²<http://www.brain-development.org>

- Desenvolvimento de um método para a detecção automática de pontos salientes 3D nas imagens de RM. O método se baseou no modelo de congruência de fase (KOVESI, 2000) e foi uma extensão da técnica proposta por Ferrari et al. (2011), com a inclusão de uma assinatura ou descritor para a representação, por um vetor de características, de cada ponto saliente detectado.
- Construção de dois tipos de atlas de pontos salientes para auxiliar o processo de inicialização dos modelos geométricos que representam as estruturas em estudo, utilizando para isso as imagens pré-processadas descritas no primeiro item.
- Construção de uma base de imagens 3D sintéticas que representam pequenas estruturas 3D de alta curvatura sob diferentes condições de degradação da imagem, para fins de avaliação da técnica de detecção de pontos salientes 3D.
- Avaliação da efetividade do atlas desenvolvido quando aplicado para guiar o posicionamento inicial e adaptação dos modelos das estruturas de interesse em um conjunto de imagens sintéticas (sem patologia diagnosticada) e clínicas (pacientes com DA).

1.3 Organização do trabalho

O texto deste trabalho está organizado da seguinte maneira: o Capítulo 2 apresenta uma série de informações complementares a respeito do tema da pesquisa, incluindo características do imageamento por RM e da DA. O Capítulo 3 contém uma revisão do estado da arte em termos das técnicas de detecção automática de pontos salientes, construção de atlas do cérebro e inicialização automática de modelos deformáveis para fins de segmentação de imagens 3D. O Capítulo 4 apresenta os materiais e métodos utilizados e desenvolvidos na pesquisa. O Capítulo 5 apresenta e discute todos os resultados obtidos. Por fim, o Capítulo 6 conclui o texto da dissertação, recapitulando as contribuições do trabalho e apresentando ideias para trabalhos futuros e a produção científica realizada no período de desenvolvimento do projeto.

Capítulo 2

INFORMAÇÕES COMPLEMENTARES

Este capítulo apresenta com maior riqueza de detalhes algumas informações complementares relevantes para o entendimento desta pesquisa. É feita uma explicação resumida do processo de formação e aquisição da imagem de ressonância magnética e das principais características desse tipo de imagem. Também são descritos os aspectos fundamentais da doença de Alzheimer, um dos principais exemplos de doenças neurodegenerativas que, acredita-se, poderiam ter seus diagnósticos auxiliados pela técnica computacional proposta neste trabalho.

2.1 Imageamento por ressonância magnética

O imageamento por RM é um método não invasivo e não ionizante de diagnóstico por imagem que possibilita a exploração de aspectos anatômicos e funcionais em diferentes partes do corpo humano, sendo amplamente utilizado em hospitais de todo o mundo nos dias atuais. Seu desenvolvimento tornou-se possível graças à observação do fenômeno da ressonância magnética nuclear na década de 1940 (BLOCH; HANSON; PACKARD, 1946; PURCELL; TORREY; POUND, 1946), muito embora a técnica e os equipamentos de imageamento que permitiram que as primeiras imagens médicas pudessem ser finalmente adquiridas só tenham ocorrido décadas depois. Desde então, o método original passou por grandes evoluções, contando atualmente com um tempo de aquisição cerca de cem vezes menor e alcançando resolução espacial em escala inferior a milímetros, além de ter inaugurado todo um novo campo de pesquisas com o desenvolvimento da RM funcional, que possibilita explorar a atividade cerebral envolvida nas funções de memória, linguagem, controle da motricidade, entre outras (MAZZOLA, 2009).

O estudo da cadeia de processos por trás da formação e aquisição de uma imagem de RM compreende diversas áreas de estudo, incluindo mecânica quântica, eletromagnetismo, trans-

formações matemáticas e processamento de sinais. Os parágrafos a seguir visam explicar tal processo de maneira simples e resumida a partir dos trabalhos publicados por Gossuin et al. (2010), Mazzola (2009) e Bushberg et al. (2011).

2.1.1 Ressonância magnética nuclear

Segundo a mecânica quântica, os núcleos dos átomos possuem uma propriedade de *spin*, \vec{I} , com a qual se relacionam uma medida de momento angular \vec{J} diretamente proporcional a \vec{I} e uma medida de momento magnético $\vec{\mu}$ diretamente proporcional a \vec{J} , que serão não nulas para núcleos que apresentam determinadas características de composição (número ímpar de prótons, por exemplo). Fazendo uma analogia simplificada, podemos entender tais propriedades imaginando os núcleos como se fossem pequenos ímãs que giram em torno de si mesmos e que são espacialmente orientados em uma determinada direção.

No contexto da RM, especificamente, o que mais nos interessa são os fenômenos quânticos relacionados ao núcleo do átomo de hidrogênio (composto por apenas 1 próton) visto que tal elemento é o mais abundante no corpo humano e possui alta sensibilidade ao fenômeno de RM nuclear. Numa amostra qualquer de tecido humano, sob ação do fraco campo magnético terrestre e em condições normais de temperatura, os momentos magnéticos dos prótons de hidrogênio distribuem-se de maneira indefinida pelas infinitas orientações espaciais, de maneira que a magnetização resultante em tal amostra é praticamente nula (Figura 2.1a). O primeiro passo de um aparelho de aquisição (scanner) de RM, então, é submeter o paciente à ação de um forte campo magnético externo \vec{B}_0 (comumente 1,5T ou 3T) para que os momentos magnéticos dos prótons de hidrogênio presentes em seus tecidos passem a se alinhar, paralelamente ou antiparalelamente, em orientação a tal campo. O eixo que indica a direção na qual o campo magnético externo é aplicado é chamado de eixo longitudinal, e o plano perpendicular a essa direção é denominado plano transversal. Pela combinação do *spin* com o alinhamento forçado, cada momento magnético acaba adquirindo um movimento de precessão em torno do eixo longitudinal, com frequência angular dada por $\omega_0 = \gamma B_0$, sendo γ a constante giromagnética do hidrogênio. Essa frequência de precessão denomina-se frequência de Larmor (Figura 2.1b).

Como a princípio não existe coerência de fase entre os momentos magnéticos dos prótons, não existe um componente de magnetização no plano transversal. Logo, não há como uma bobina posicionada perpendicularmente a esse plano detectar qualquer sinal, dado que não existe variação no fluxo magnético para induzir uma corrente elétrica. Por esse motivo, aplica-se no scanner um segundo campo magnético, \vec{B}_1 , que é perpendicular a \vec{B}_0 , possui curta duração (pulso) e está em fase com a frequência de Larmor. \vec{B}_1 é denominado um pulso de radiofrequên-

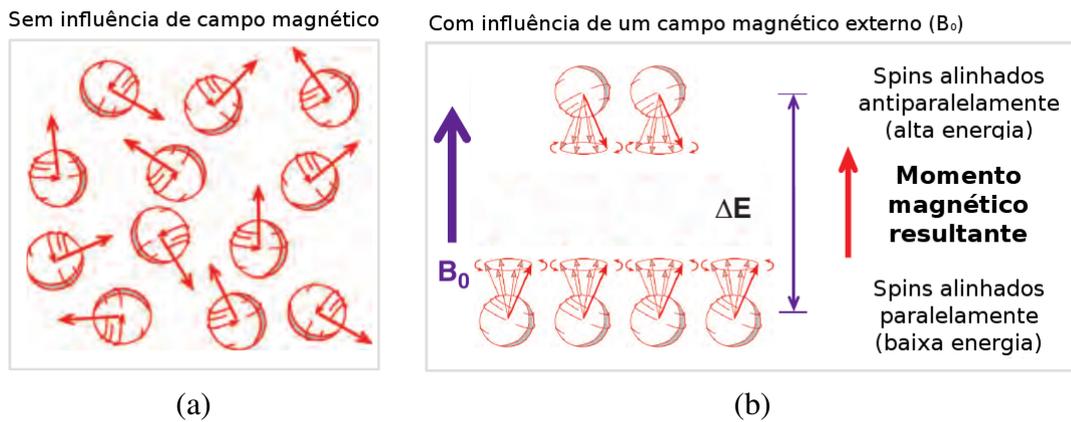


Figura 2.1: Representação gráfica da distribuição dos momentos magnéticos dos prótons de hidrogênio em uma amostra qualquer de tecido (a) em condições usuais; e (b) sob ação do campo magnético \vec{B}_0 . Fonte: adaptada de Bushberg et al. (2011).

cia (RF) e tem como consequência imediata de sua aplicação o fenômeno de ressonância que causa a reorientação dos vetores magnetização dos prótons de hidrogênio, afastando-os do alinhamento com \vec{B}_0 por um ângulo de desvio α (comumente igual a 90° ou 180°). Esse desvio faz com que os núcleos realizem um movimento de precessão momentaneamente em fase no plano transversal, induzindo assim uma tensão elétrica na bobina do scanner e fazendo surgir nela um sinal de indução livre (SIL). Cessado o pulso de RF, ocorre o processo de relaxação, no qual os vetores começam a se realinhar com \vec{B}_0 e o SIL gradualmente decai (Figura 2.2).

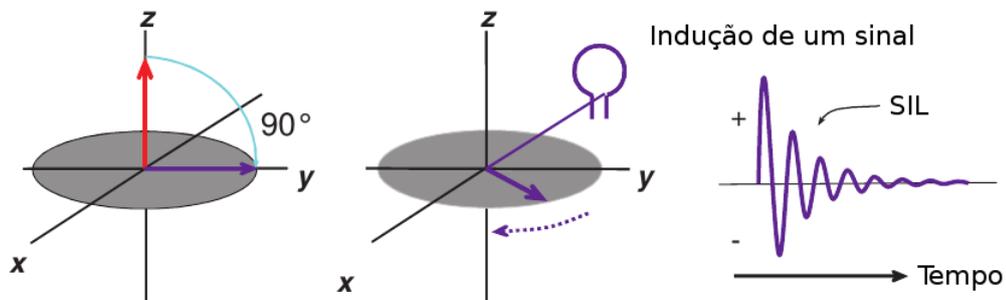


Figura 2.2: Produção de um SIL a partir da aplicação de um pulso de RF de 90° . O SIL apresenta variações positivas e negativas oscilando na frequência de Larmor, e decai com o tempo devido à posterior perda gradual de coerência de fase entre os momentos magnéticos no plano transversal. Fonte: adaptada de Bushberg et al. (2011).

Duas constantes de tempo de grande importância para a RM atuam durante a relaxação:

T1 Está relacionada ao tempo de retorno da magnetização para o eixo longitudinal (Z) após um pulso de RF de 90° . Esse retorno é descrito pela equação

$$\mu_Z(t) = \mu_Z^{(0)} \left(1 - e^{-t/T1} \right), \quad (2.1)$$

em que $\mu_Z(t)$ é a magnetização no eixo longitudinal no tempo t e $\mu_Z^{(0)}$ é a magnetização

inicial nesse eixo. Na prática, a constante $T1$ representa o tempo necessário para que $\mu_z(t)$ atinja 63% do seu valor inicial (Figura 2.3).

T2 Está relacionada ao tempo de decaimento da magnetização no plano transversal (XY) após um pulso de RF de 90° . Esse decaimento é descrito pela equação

$$\mu_{XY}(t) = \mu_{XY}^{(0)} e^{-t/T2}, \quad (2.2)$$

onde $\mu_{XY}(t)$ é a magnetização no plano transversal no tempo t e $\mu_{XY}^{(0)}$ é a magnetização inicial nesse plano. Na prática, a constante $T2$ representa o tempo necessário para que $\mu_{XY}(t)$ atinja 37% do seu valor inicial (Figura 2.4).

Os valores de $T1$ e $T2$ não são necessariamente iguais, já que alguns processos que participam da relaxação transversal não estão presentes na longitudinal. Em geral, $T1 \geq T2$, e para a maioria dos tecidos biológicos, $T1 \gg T2$. Conhecer os tempos de relaxação dos tecidos é algo crucial no imageamento por RM, visto que a intensidade dos sinais - e consequentemente o contraste das imagens adquiridas - depende desses fatores. Tipos diferentes de tecidos podem ser visualmente identificados numa imagem de RM justamente por causa das diferenças que os seus elementos constituintes apresentam nesses tempos.

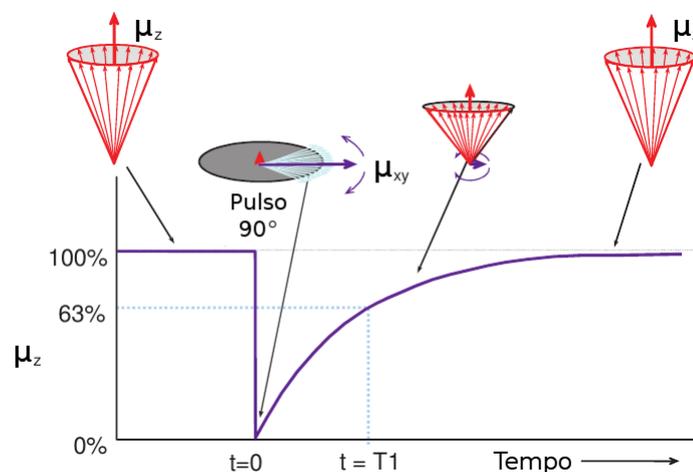


Figura 2.3: Retorno gradual do componente de magnetização no eixo longitudinal (μ_z) após a aplicação de um pulso de RF de 90° . O tempo $t = T1$ é o momento em que se tem $\mu_z = 0,63\mu_z^{(0)}$. Fonte: adaptada de Bushberg et al. (2011).

Um outro aspecto muito importante em imageamento por RM, no que se refere à coleta do sinal que irá formar a imagem, é o fenômeno da formação de ecos. Basicamente significa que, se aplicarmos dois pulsos de RF separados por um intervalo de tempo Δt , surgirá na bobina não apenas o sinal decorrente do primeiro pulso (o SIL) como também um segundo sinal, que é um eco do primeiro, aparecendo após um tempo $2\Delta t$ com o alinhamento das fases temporárias

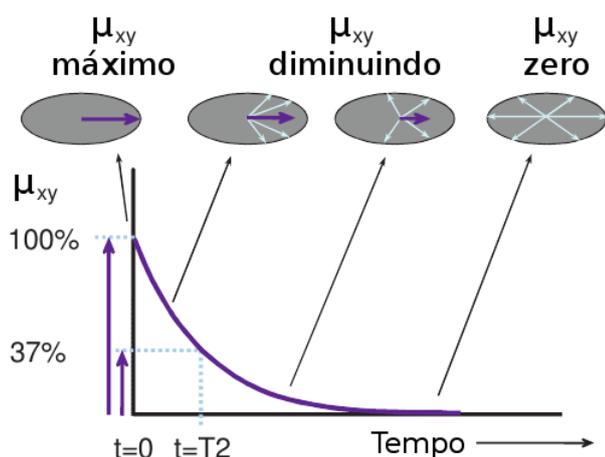


Figura 2.4: Decaimento gradual do componente de magnetização no plano transversal (μ_{XY}) após a aplicação de um pulso de RF de 90° . O tempo $t = T2$ é o momento em que se tem $\mu_{XY} = 0,37\mu_{XY}^{(0)}$. Fonte: adaptada de Bushberg et al. (2011).

dos momentos magnéticos, e que é coletado para formação da imagem. A partir disso, pode-se falar no conceito de famílias de sequências de pulso, as quais influenciam diretamente em certas características da imagem. Por exemplo, a família de sequências denominada *spin echo* é caracterizada pela aplicação de um pulso de RF de 90° seguido por um pulso de RF de 180° . A relação do tempo entre dois pulsos consecutivos de 90° (TR - tempo de repetição) com o tempo entre um pulso de 90° e o pico do eco (TE - tempo de eco) determina a ponderação de níveis de cinza que tal imagem terá¹.

Diversas outras sequências de pulso existem e possuem características e vantagens próprias em determinadas aplicações, mas tais detalhes vão além do escopo desse trabalho de Mestrado.

2.1.2 Formação da imagem

Para que o sinal obtido na bobina do scanner de RM tenha utilidade para a formação de uma imagem, é necessário realizar a sua codificação espacial, possibilitando tanto a seleção de cortes transversais (fatias) específicos do volume imageado como a localização espacial dos pontos presentes em uma mesma fatia. Isso permite associar a cada ponto um valor de intensidade de sinal que leva, na imagem posteriormente formada, à atribuição de um tom em uma escala de cinza. Na década de 1970, foi desenvolvido um método baseado em gradientes de campo magnético para esse fim. A ideia consistiu em fazer com que a magnitude do campo magnético variasse linearmente ao longo de um dado eixo, de maneira que certas propriedades dos momentos magnéticos em diferentes pontos do campo passassem a ser distintas e, com isso, se conseguisse localizar tais pontos em função da fatia selecionada, da frequência e da fase do

¹Vide Subseção 2.1.3.1.

sinal detectado.

Por exemplo, para fazer a seleção de uma fatia, varia-se o campo \vec{B}_0 ao longo da posição no eixo longitudinal (z) utilizando um gradiente G_z :

$$B_z(z) = B_0 + zG_z. \quad (2.3)$$

Fazendo isso, a frequência de Larmor dos prótons passa a depender linearmente de z , e dessa forma é possível excitar prótons de maneira seletiva pela aplicação de um pulso de RF com a exata frequência correspondente à frequência de Larmor dos prótons localizados na fatia desejada, aumentando ou diminuindo a frequência do pulso para selecionar outras fatias do mesmo volume.

Para a localização espacial dos prótons de uma mesma fatia, são feitas as codificações do sinal em termos de frequência e fase. A codificação de frequência é obtida aplicando-se um gradiente na direção do eixo X (G_X) durante a aquisição do sinal, também alterando a frequência de precessão dos prótons conforme suas posições ao longo desse eixo. A codificação de fase é feita pela aplicação de um pulso de gradiente ao longo do eixo Y (G_Y) antes da aquisição do sinal, fazendo assim com que os sinais obtidos dos prótons em diferentes posições ao longo desse eixo apresentem diferentes fases.

A partir dessas codificações, é definido o conceito de espaço K , que pode ser entendido como um mapa bidimensional $S(k_x, k_y)$ da distribuição das frequências contidas nos sinais fornecidos pelo volume amostrado (paciente). Como o scanner de RM realiza a detecção dos sinais em quadratura (isto é, com um par de bobinas sensíveis ao fluxo magnético em direções ortogonais), cada amostra de $S(k_x, k_y)$ é um número complexo. Os eixos de coordenadas desse espaço (k_X e k_Y) são estabelecidos, respectivamente, pelos gradientes de codificação de frequência e de codificação de fase. Observa-se, assim, que a informação contida no espaço K consiste na transformada de Fourier (TF) da distribuição espacial do sinal $I_C(x, y)$ que corresponde à imagem complexa de uma fatia do volume amostrado. Logo, utiliza-se a transformada 2D inversa de Fourier sobre o espaço K para obter $I_C(x, y)$:

$$\begin{aligned} I_C(x, y) &= \text{TF}^{-1} [S(x_k, y_k)] \\ &= \frac{1}{NM} \sum_{k_x=0}^{N-1} \sum_{k_y=0}^{M-1} S(k_x, k_y) \exp \left[j2\pi \left(\frac{k_x x}{N} + \frac{k_y y}{M} \right) \right], \end{aligned} \quad (2.4)$$

sendo N e M as quantidades de codificações de frequência e fase, respectivamente, e $j = \sqrt{-1}$. Por fim, a imagem final (imagem de magnitude) é obtida calculando-se o módulo dos coefi-

entes real e imaginário de $I_{\mathbb{C}}(x, y)$:

$$I(x, y) = \sqrt{\text{real}^2 [I_{\mathbb{C}}(x, y)] + \text{imag}^2 [I_{\mathbb{C}}(x, y)]}. \quad (2.5)$$

Na Figura 2.5, a relação entre os dados no espaço K e a imagem de RM correspondente, em termos da TF, é ilustrada.

Repetindo o mesmo procedimento para todos os cortes transversais (z) selecionados, adquire-se uma sequência de imagens 2D representando diferentes fatias do volume imageado. A partir do empilhamento dessas fatias, o resultado da aquisição pode ser considerado uma imagem 3D, isto é, uma matriz de voxels. Alternativamente, seria possível tratar cada fatia como uma imagem independente no momento do processamento (especialmente se existir um grande espaçamento entre as fatias impedindo que a imagem seja tratada como um volume contínuo), mas não é dessa maneira que as imagens de RM são interpretadas e processadas neste trabalho. Por isso, o termo “imagem 3D de RM” é empregado ao longo deste texto.

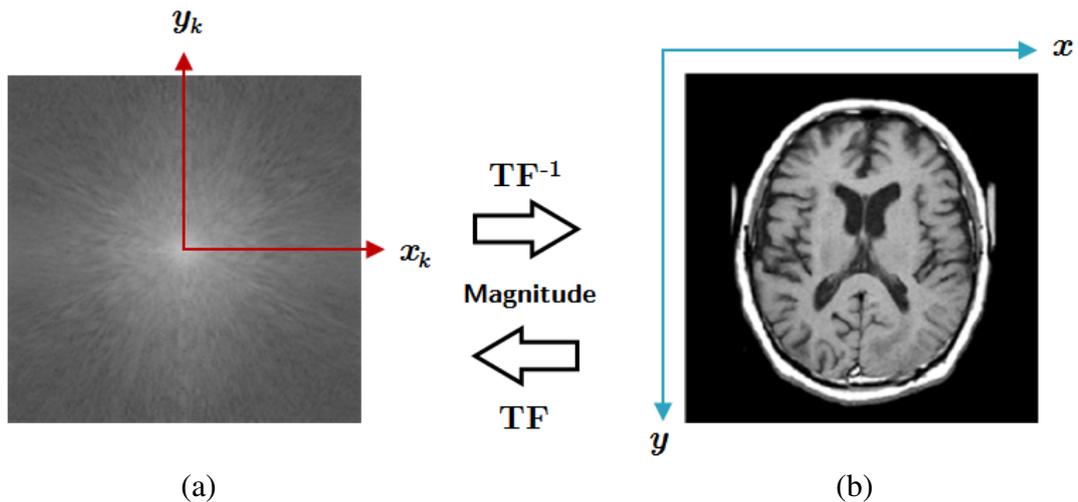


Figura 2.5: Relação entre (a) o espaço K e (b) a imagem de RM formada, em termos da TF. Os valores nas regiões centrais do espaço K correspondem às densidades de baixas frequências (intensidade constante ou pouco variante na imagem de RM), enquanto os valores mais distantes do centro representam as frequências mais altas (detalhes e bordas na imagem de RM). Fonte: adaptada de Mazzola (2009).

2.1.3 Características de imagem

A seguir são apresentadas algumas características fundamentais das imagens de RM, cujo entendimento se faz importante no contexto deste trabalho.

2.1.3.1 Modalidade de ponderação

A ponderação de uma imagem de RM diz respeito ao tipo de informação que determina a magnitude do sinal adquirido pelo scanner e, conseqüentemente, a intensidade de nível de cinza atribuída a cada ponto da imagem. Diversas ponderações existem atualmente, cada qual com características que as tornam mais ou menos vantajosas, dependendo da aplicação. Por exemplo, a ponderação T1-w (T1-weighted), que resulta em uma distribuição dos níveis de cinza de maneira proporcional ao valor da constante $T1$ de cada tecido, propicia às imagens um melhor delineamento da anatomia do paciente. Imagens em T2-w (distribuição dos níveis de cinza em função do valor da constante $T2$) possuem importante aplicação no diagnóstico da esclerose múltipla, uma vez que tal ponderação fornece maior contraste às lesões cerebrais na imagem. Outras ponderações comumente utilizadas em aplicações clínicas incluem *Proton Density-weighted* (PD-w) e *Fluid Attenuated Inversion Recovery* (FLAIR). Na Figura 2.6, são comparadas as diferenças visuais entre as ponderações T1-w, T2-w e PD-w na imagem de um mesmo paciente.

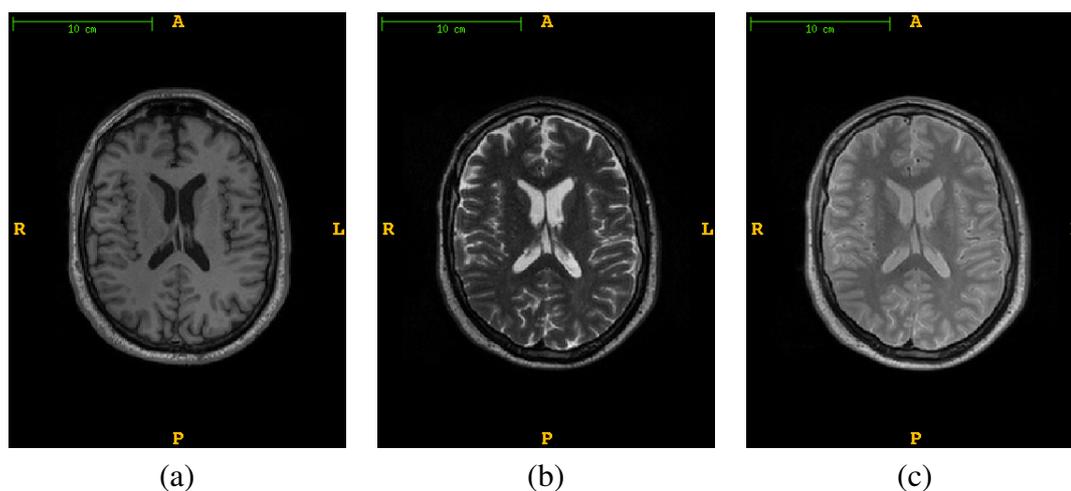


Figura 2.6: Comparação entre as ponderações (a) T1-w, (b) T2-w, e (c) PD-w na imagem de um mesmo paciente.

Conforme já mencionado, a ponderação de uma imagem de RM é controlada pelo tempo de repetição (TR) e pelo tempo de eco (TE) da sequência de pulso utilizada na aquisição. No caso da sequência *spin-echo*, temos a relação apresentada na Tabela 2.1.

2.1.3.2 Resolução e dimensões

A resolução espacial é uma característica diretamente relacionada ao grau de nitidez de uma imagem. Nas imagens 3D de RM, ela é determinada basicamente pelo tamanho dos voxels. Esse tamanho, por sua vez, depende de quatro outros fatores: das dimensões da matriz imagem, do

Tabela 2.1: Tempos de repetição (TR) e de eco (TE) utilizados para se obter as ponderações T1-w, T2-w e PD-w na sequência *spin-echo*.

Ponderação	TR	TE
T1-w	Curto (<500ms)	Curto (<30ms)
T2-w	Longo (>1500ms)	Longo (>90ms)
PD-w	Longo (>1500ms)	Curto (<30ms)

tamanho do campo de visão (*Field Of View* - FOV), da espessura de volume representado em cada fatia e do espaçamento entre fatias (BROWN; SEMELKA, 2003).

As dimensões da matriz imagem são dadas conforme o número de passos de codificação de frequência (eixo X) e de fase (eixo Y) e pelo número de fatias adquiridas (eixo Z). Assim, se todas as outras grandezas se mantiverem constantes, o aumento desses valores resulta em melhora na resolução da imagem. No entanto, aumentar tais valores também significa aumentar o tempo de aquisição da imagem, algo que pode não ser viável devido ao desconforto que seria causado ao paciente e ao atraso que isso acarretaria nos diagnósticos. Imagens de RM do cérebro tipicamente usadas em diagnósticos clínicos muitas vezes não ultrapassam o tamanho 256 em nenhum dos eixos, e também é bastante comum que tenham um tamanho reduzido em algum eixo específico apenas para fins de agilizar a aquisição.

O FOV é a medida da área do mundo real que está representada em cada fatia 2D da imagem 3D de RM. Dividindo o FOV pelo tamanho total da fatia, obtemos o tamanho dos voxels nos eixos X e Y . Logo, aumentar o FOV em qualquer direção (mantendo as demais grandezas constantes) aumenta também o tamanho dos voxels na imagem final, consequentemente causando piora na resolução espacial. Por outro lado, FOVs menores produzem imagens com melhor resolução, mas que só conseguem representar regiões mais localizadas da anatomia do paciente.

A espessura de volume representado em cada fatia vai depender do gradiente G_z usado para seleção dos cortes transversais. Quanto maior for a precisão que tal gradiente proporciona na seleção de um corte localizado, menor tende a ser a espessura das fatias e o tamanho dos voxels no eixo Z . Entretanto, os scanners também estabelecem um espaçamento entre fatias vizinhas - algo que acaba conduzindo ao aumento do tamanho dos voxels devido à limitação no número de fatias adquiridas. Teoricamente falando, poderíamos pensar que, quanto menor a espessura e o espaçamento entre fatias, melhor será a resolução; mas o que acontece na prática é que a redução excessiva desses valores acaba fazendo surgir interferências entre os sinais provenientes de diferentes fatias (fenômeno denominado *cross-talking*), consequentemente diminuindo a relação sinal-ruído da imagem adquirida.

Mesmo quando uma imagem de RM tem resolução razoável, a discretização espacial realizada no momento da aquisição pode acabar fazendo com que um único voxel contenha informações vindas de dois ou mais tecidos diferentes. Esse fenômeno é chamado de efeito de volume parcial. É considerado um problema especialmente para os algoritmos de segmentação de imagens baseados em classificação voxel a voxel.

Uma imagem de RM é denominada isotrópica quando o tamanho dos voxels, decorrente das características mencionadas no parágrafo anterior, é idêntico nas direções dos três eixos (por exemplo, $1 \times 1 \times 1 \text{ mm}$). Quando existem diferenças em pelo menos uma das direções, a imagem é dita anisotrópica.

2.1.3.3 Sistemas de coordenadas

De acordo com o trabalho de Kindlmann (KINDLMANN, 2010), três sistemas de coordenadas são considerados quando se trabalha com aquisição e processamento de imagens médicas: (a) coordenadas do mundo, (b) coordenadas anatômicas e (c) coordenadas da imagem. O sistema de coordenadas do mundo é basicamente um sistema cartesiano 3D único usado como referência para o posicionamento de um modelo (por exemplo, o paciente no scanner). O sistema de coordenadas anatômicas, por sua vez, é referente ao paciente e é composto por três planos ortogonais que descrevem seu posicionamento:

- O plano axial, que separa a parte superior do corpo (cabeça) da parte inferior (pés);
- O plano coronal, que separa a parte anterior (frente) da parte posterior (costas);
- O plano sagital, que separa o lado esquerdo do lado direito.

Dessa forma, uma base é definida a partir dos vetores representando eixos esquerda-direita, anterior-posterior e superior-inferior. A orientação de uma imagem 3D de RM pode então ser representada por um código de três letras denominado *RAI code*, pelo qual cada letra corresponde (na língua inglesa) a um sentido relativo a esses eixos. Esse código é utilizado por softwares médicos para fins de visualização na tela do computador. Por exemplo, o código *AIL* significa que o eixo *X* corresponde à direção anterior-posterior (*Anterior-posterior*), o eixo *Y* corresponde à direção inferior-superior (*Inferior-superior*) e o eixo *Z* corresponde à direção esquerda-direita (*Left-right*).

Por fim, o sistema de coordenadas da imagem descreve como a imagem foi adquirida em relação à anatomia do paciente. Pode ser visto como uma matriz 3D de pontos (voxels) na qual cada elemento em (i, j, k) representa uma amostra da anatomia do paciente em determinada

localização. Além das intensidades de cada voxel, o sistema de coordenadas da imagem compreende também as informações de origem (posição do primeiro voxel, $(0,0,0)$, no sistema de coordenadas anatômicas) e a medida de espaçamento entre voxels relativa a cada eixo.

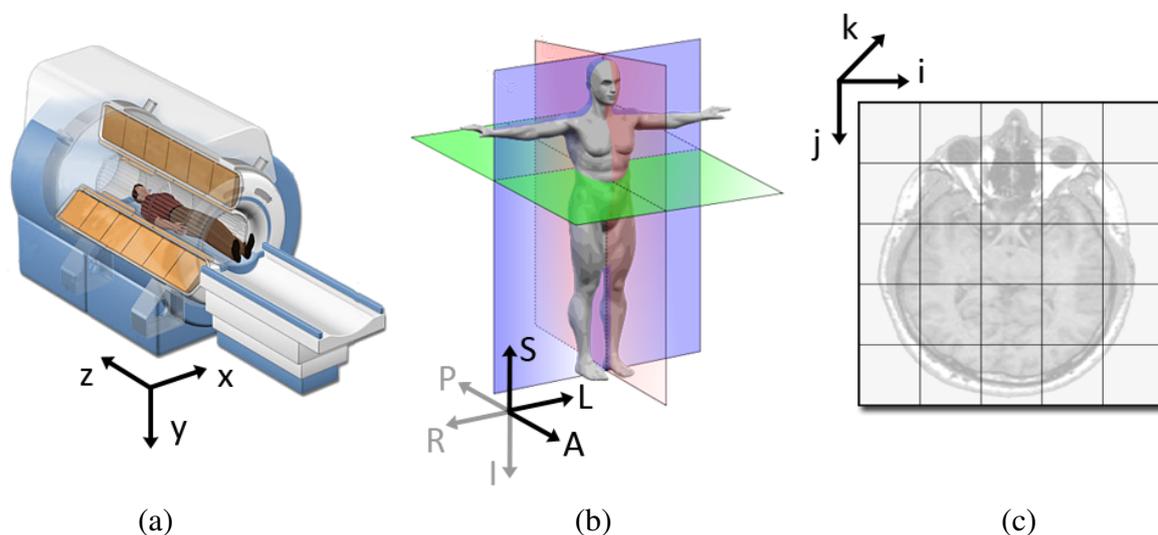


Figura 2.7: Sistemas de coordenadas considerados no contexto do imageamento por RM: coordenadas (a) do mundo; (b) anatômicas; (c) da imagem. Fonte: SlicerWiki (2014).

2.1.3.4 *Bias field*

Mesmo quando toda a sua informação é obtida de um tecido homogêneo, as heterogeneidades que inevitavelmente ocorrem no campo magnético fazem com que a intensidade de um sinal de RM raramente seja uniforme. Diversos são os fatores que acarretam tal fenômeno, como, por exemplo, a forma de resposta da bobina receptora e a própria anatomia do paciente no scanner. Como resultado, as imagens obtidas se mostram corrompidas por um sinal de baixa frequência que degrada bordas e altera as distribuições de intensidade dos tecidos conforme a localização espacial. Esse efeito tem pouco impacto em análises médicas visuais, porém técnicas automáticas que fazem uso dos valores de intensidade podem ter seus desempenhos prejudicados (GISPERT et al., 2004; JUNTU et al., 2005).

Esse campo de heterogeneidade (*bias field*) é matematicamente modelado como um sinal multiplicativo que suaviza o sinal original emitido pelo tecido do paciente numa determinada localidade. A partir dessa formulação, pode-se trabalhar em métodos computacionais para a correção de tal efeito².

²Vide Subseção 4.3.2.

2.1.3.5 Ruído

Os dados complexos representados no espaço K são corrompidos por ruído gerado, principalmente, por flutuações térmicas que ocorrem devido ao próprio corpo do paciente no scanner. Esse ruído pode ser estatisticamente modelado por uma distribuição Gaussiana que corrompe de maneira independente as partes real e imaginária de cada amostra no espaço K (MACOVSKI, 1996). Entretanto, com o cômputo da magnitude do sinal (Equação 2.5), o modelo de ruído deixa de ser Gaussiano e passa a ser descrito pela distribuição de Rice (GUDBJARTSSON; PATZ, 1995), expressa por $\Pr(b | a, \sigma)$, em que a é a intensidade de um voxel na ausência de ruído e b é a intensidade efetivamente obtida. Essa distribuição pode ser geometricamente interpretada como a distribuição das distâncias dos pontos $P = (x, y)$ até a origem de um plano, na condição de que os valores de x e y sejam escolhidos a partir de duas distribuições Gaussianas independentes: uma tendo média x_0 e a outra tendo média y_0 , ambas com desvio padrão σ .

Para valores de razão sinal-ruído não muito pequenos, a distribuição de Rice começa a se aproximar cada vez mais de uma distribuição Gaussiana. Por essa razão, a modelagem do ruído por esta distribuição é, em geral, uma aproximação suficientemente precisa em aplicações práticas envolvendo imagens de RM. Um caso especial da distribuição de Rice ocorre nas localidades da imagem em que há apenas ruído ($a = 0$): nesses casos, $\Pr(b | 0, \sigma)$ assume a forma da distribuição de Rayleigh.

2.2 Doença de Alzheimer

A doença de Alzheimer (DA) é uma doença crônica não transmissível, neurodegenerativa e progressiva que consiste na causa mais comum de demência em idosos (ALZHEIMER'S DISEASE EDUCATION AND REFERRAL CENTER, 2011). O termo demência refere-se, de maneira generalizada, ao processo de perda ou progressiva redução das capacidades cognitivas de um indivíduo (memória, raciocínio, comunicação, etc.) de maneira a interferir em seu funcionamento ocupacional e social e acarretar a perda de sua autonomia.

A DA é, em essência, uma doença que afeta idosos³. Sua prevalência aumenta rapidamente conforme a idade avança, dobrando a cada 5 anos a partir dos 65 anos de idade e abrangendo atualmente mais de 30% da população com idade na faixa dos 80 anos (APRAHAMIAN; MARTINELLI; YASSUDA, 2009). A característica de ter a idade como principal fator de risco torna a DA

³É conhecida a existência de casos de DA em pessoas mais jovens, com idade a partir dos 35 anos; entretanto, a ocorrência de tais casos é muito mais baixa em comparação aos índices da doença em pessoas acima dos 65 anos de idade (FLIER; SCHELTENS, 2005).

um preocupante problema de saúde pública, devido ao crescente envelhecimento populacional. No Brasil, a população com mais de 60 anos corresponde a cerca de 20 milhões de pessoas (dentre as quais estima-se por volta de 1,2 milhão de casos de DA) e as projeções indicam que, em 2030, os idosos irão compor aproximadamente 13,5% da população nacional (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2014). A nível global, estudos apontaram uma prevalência de cerca de 26,6 milhões de casos de DA no ano de 2006 e projetam que, até 2050, se não forem feitos avanços em termos de diagnóstico precoce e tratamento, tal número irá quase quadruplicar, ultrapassando a marca dos 106 milhões de casos. Mais do que isso, estima-se que 43% desses casos futuros serão graves o bastante para exigir a necessidade de cuidados caseiros contínuos (BROOKMEYER et al., 2007), causando um impacto social e econômico bastante significativo.

2.2.1 Sintomas

Os problemas de memória costumam ser o primeiro sinal de perda cognitiva observado nas pessoas que possuem a DA. À medida que a doença avança de estágio, o deficit adquirido em outros domínios cognitivos começa a ser cada vez mais evidente: a pessoa pode começar a se perder, a ter problemas em lidar com dinheiro e pagar suas contas, a repetir perguntas sem perceber, a necessitar de mais tempo do que o normal para desempenhar suas tarefas diárias básicas, e a sofrer mudanças de comportamento. É neste estágio que normalmente a família e amigos percebem que há algo de errado e acontece o diagnóstico clínico. Posteriormente, esse indivíduo pode passar a ter grande dificuldade em reconhecer seus próprios amigos e familiares e em desempenhar sozinho tarefas como vestir-se e cozinhar, além de ter seus problemas de memória e confusão mental consideravelmente piorados. No último estágio da doença, o paciente já não tem mais sequer capacidade de se comunicar e torna-se completamente dependente de outras pessoas, passando, em seus últimos dias, a maior parte do tempo acamado (ALZHEIMER'S DISEASE EDUCATION AND REFERRAL CENTER, 2011).

Existe também uma condição denominada Comprometimento Cognitivo Leve (CCL), que descreve pacientes que possuem um deficit cognitivo acima do normal para a idade, mas que ainda não é significativo o suficiente para interferir em suas atividades cotidianas, e portanto não é classificado como demência. Quando esse deficit se manifesta em termos de perda de memória, classifica-se tal condição como CCL amnésico, que pode ser considerado um estágio intermediário entre o envelhecimento natural e a DA. Segundo estudos, todos os anos um percentual de 10% a 15% dos pacientes com CCL amnésico acabam progredindo para um diagnóstico de DA (BLENNOW; LEON; ZETTERBERG, 2006).

2.2.2 Causas

A ciência ainda não tem uma resposta definitiva para quais são as causas da DA. Sabe-se que ela evolui a partir de uma série de eventos que ocorrem no cérebro ao longo de muitos anos, mas esses eventos ainda não são perfeitamente compreendidos. Acredita-se que seja uma combinação de fatores genéticos e comportamentais.

Atualmente, os cientistas procuram estudar como as típicas mudanças na estrutura cerebral que ocorrem devido à idade (atrofias, inflamações, produção de radicais livres, etc.) danificam os neurônios e contribuem para o aparecimento e agravamento da doença. Já se estima, por exemplo, que os danos ao cérebro comecem décadas antes de os sintomas da DA aparecerem e o problema se tornar evidente (BLENNOW; LEON; ZETTERBERG, 2006; ALZHEIMER'S DISEASE EDUCATION AND REFERRAL CENTER, 2011). Esses danos incluem a formação de placas senis e novos neurofibrilares em diferentes partes do cérebro. Placas senis são aglomerações de fragmentos proteicos beta-amiloide que, ao se formarem, têm efeito tóxico sobre os neurônios e podem levar ao bloqueio da comunicação sináptica. Novos neurofibrilares estão relacionados à mutação da proteína tau, responsável pela estabilização dos microtúbulos neuronais que, em condições normais, possibilitam o transporte de nutrientes e outros elementos essenciais entre essas células. O aparecimento generalizado das placas e novos acaba levando os neurônios a perderem gradativamente a capacidade de interagir uns com os outros e, conseqüentemente, a morrerem. Com a diminuição da quantidade de neurônios ativos, o resultado é a atrofia cerebral. Essa atrofia é primeiramente observada, e especialmente severa, nas regiões do córtex entorrinal e do hipocampo, as quais têm importância no desempenho de funções relacionadas ao raciocínio e à memorização. Em contraste, os ventrículos cerebrais, que produzem e são preenchidos pelo líquido cefalorraquidiano, aumentam de tamanho (MORISHIMA-KAWASHIMA; IHARA, 2002).

Além da genética e da idade avançada, outro fator de risco (ou proteção) para a DA é o estilo de vida do indivíduo. Estudos indicam que possuir um maior grau de escolaridade, alimentar-se bem, praticar atividade física e manter-se socialmente ativo são fatores que efetivamente reduzem o declínio cognitivo na velhice (APRAHAMIAN; MARTINELLI; YASSUDA, 2009; ALZHEIMER'S DISEASE EDUCATION AND REFERRAL CENTER, 2011).

2.2.3 Diagnóstico

O diagnóstico da DA é uma tarefa complexa, realizada a partir de um conjunto de critérios clínicos preestabelecidos e da exclusão de outras possíveis causas para a demência apresentada pelo paciente. Essa exclusão é feita através de exames clínicos, laboratoriais e de imagea-

mento cerebral. Os exames clínicos incluem uma avaliação cuidadosa do histórico do paciente (doenças preexistentes, traumas, uso de álcool e medicações, exposição ambientais a tóxicos, etc.) e também a condução de testes neuropsicológicos de memória e outras funções cognitivas. Os procedimentos laboratoriais podem incluir exames de sangue e urina para verificar, por exemplo, se ocorre deficiência de vitamina B_{12} , hipotireoidismo ou outras patologias que se sabe também terem relação com as demências em idosos (APRAHAMIAN; MARTINELLI; YASSUDA, 2009).

O uso do imageamento cerebral por RM ou tomografia computadorizada (*Computerized Tomography - CT*) é recomendado pela Academia Brasileira de Neurologia (e pelas análogas estrangeiras) porque, primeiramente, a análise dessas imagens possibilita excluir da lista de possíveis causas de demência os tumores e as lesões estruturais no cérebro. Mas, além disso, o imageamento por RM e CT também auxilia na identificação das atrofia que ocorrem, desde os estágios iniciais da DA, no lobo temporal medial, onde se incluem os já mencionados hipocampo e córtex entorrinal. Embora os padrões de atrofia ainda permaneçam incertos e difíceis de se correlacionar com a fisiopatologia subjacente, é mostrado em estudos que a avaliação da atrofia temporal medial permite discriminar com suficiente efetividade os pacientes com diagnóstico clínico de DA daqueles cognitivamente normais, estando por isso entre os mais aceitos biomarcadores para essa doença (DUBOIS et al., 2010; APRAHAMIAN; MARTINELLI; YASSUDA, 2009). A observação da atrofia hipocampal em pacientes com CCL também ajuda a estimar as chances de que o CCL evolua futuramente para a DA (APOSTOLOVA et al., 2006).

Outras técnicas de imageamento cerebral muito úteis ao diagnóstico da DA são a tomografia por emissão de pósitrons (*Positron Emission Tomography - PET*) e a tomografia computadorizada por emissão de fóton único (*Single-Photon Emission Computerized Tomography - SPECT*), que, por realizarem imageamento funcional, permitem avaliar o fluxo sanguíneo e o metabolismo da ^{18}F -fluorodeoxiglicose (FDG) nas regiões dos lobos temporais, identificando as reduções nesses dois aspectos que ocorrem no cérebro das pessoas portadoras da DA. Esse método consegue identificar com boa precisão os pacientes com a doença e, inclusive, diferenciá-los daqueles cuja demência se dá por outras causas (FLIER; SCHELTENS, 2005; APRAHAMIAN; MARTINELLI; YASSUDA, 2009).

2.3 Considerações finais

Este capítulo apresentou um apanhado de informações complementares necessárias para o pleno entendimento da presente proposta de pesquisa. O processo de formação da imagem de

RM, bem como as características tipicamente existentes em tais imagens em decorrência desse processo, foram abordados de maneira resumida e sem grande preocupação com rigor matemático ou físico. Também foram descritas as informações mais básicas relativas à incidência, aos sintomas, às causas e ao diagnóstico da DA, que é um exemplo de doença neurológica cujo diagnóstico espera-se conseguir facilitar utilizando as técnicas computacionais desenvolvidas neste trabalho.

Capítulo 3

REVISÃO DO ESTADO DA ARTE

Este capítulo apresenta uma revisão dos principais trabalhos encontrados na literatura científica relacionados aos temas desta pesquisa. Visa sintetizar as recentes técnicas, metodologias e resultados obtidos por diferentes pesquisadores na busca pela solução de problemas análogos ou outros relacionados, a fim de contextualizar a metodologia desenvolvida neste trabalho ante o estado da arte.

3.1 Visão geral

Como este trabalho envolve três grandes tipos de técnicas - detecção de pontos salientes 3D, construção de atlas e inicialização de modelos geométricos deformáveis, esta revisão da literatura é organizada de maneira a apresentar resumidamente as principais pesquisas e resultados obtidos nessas três áreas correlatas. A produção científica em cada uma delas é muito ampla e diversa e, portanto, buscou-se limitar o escopo dos trabalhos citados nesta revisão àqueles que possuem aplicações em imagens médicas 3D.

3.2 Detecção de pontos salientes 3D

Um ponto saliente 3D pode ser entendido como uma localização, no domínio de uma imagem 3D, que apresenta alguma característica distintiva em relação à sua vizinhança imediata. Normalmente essa característica está associada a uma mudança em uma ou mais propriedades da imagem, dentre as quais se pode considerar intensidade, cor, textura, ou medidas de mais alto nível, como grau de curvatura das bordas. Normalmente, um ponto saliente é um ponto óbvio para a visão humana e, no caso de imagens médicas, ainda podem corresponder - embora não necessariamente - a pontos de referência anatômicos que seriam marcados por um

especialista; por exemplo, ao esplênio do corpo caloso no cérebro. A utilidade de pontos salientes de imagens em aplicações computacionais é diversa, podendo servir, por exemplo, para o reconhecimento e rastreamento de objetos, para orientar certas transformações, ou mesmo como uma forma de representação de uma imagem para fins de indexação em bases de dados (TUYTELAARS; MIKOLAJCZYK, 2008).

Conforme será mostrado, as técnicas existentes para detecção de pontos salientes em imagens 3D utilizam abordagens diversas, indo desde o uso de operadores diferenciais de primeira ordem para se identificar “cantos” até o ajuste de modelos paramétricos e a aplicação de métodos baseados em aprendizado. E, embora não pareça existir muito consenso na literatura quanto à melhor forma de se classificar essas técnicas, os principais critérios que definem um bom detector são bem estabelecidos. Idealmente, um detector de pontos salientes qualquer (seja 2D ou 3D) deve possuir as seguintes características (TUYTELAARS; MIKOLAJCZYK, 2008):

1. Repetibilidade: dadas duas imagens de um mesmo objeto ou cena, adquiridas sob diferentes condições de visualização, uma alta porcentagem dos pontos salientes presentes na região comum a ambas as imagens deve efetivamente ser detectada em ambas as imagens. Essa, que é considerada a característica mais importante de um detector, pode ser alcançada se a técnica for construída de maneira a ter estas duas propriedades:
 - (a) Invariância: a detecção de um ponto saliente não deve ser afetada de maneira relevante por transformações matemáticas realizadas na imagem (por exemplo, a rotação da cena).
 - (b) Robustez: o detector deve ser capaz de lidar com deformações relativamente pequenas (algum grau de ruído, por exemplo) sem que isso afete drasticamente seu desempenho.
2. Informatividade: o local onde um ponto saliente é detectado precisa apresentar uma variação relativamente alta de alguma propriedade da imagem em relação à sua vizinhança, de maneira que tal ponto seja considerado distintivo naquela região e, conseqüentemente, informativo.
3. Acurácia: a detecção de um ponto saliente deve ser a mais precisa e localizada possível no domínio da imagem, e também em outros domínios nos quais isso seja relevante (por exemplo, em termos de escala para as técnicas que utilizam multirresolução).
4. Eficiência: idealmente, um detector de pontos salientes deveria poder ser utilizado em sistemas nos quais o tempo é crítico. Entretanto, esse é um critério cujo grau de importância depende da aplicação.

Em diversos trabalhos na literatura, pontos salientes também são chamados de pontos de interesse ou marcadores (*landmarks*). Nos casos em que estiver explícito que o ponto também possui correspondência a alguma referência anatômica, será usado, aqui, o termo “marcadores anatômicos”.

3.2.1 Trabalhos correlatos

Algumas das primeiras abordagens para a detecção e localização de pontos salientes em imagens 3D foram desenvolvidas ainda na década de 1990 por Rohr (1997). No referido trabalho, quatro operadores diferenciais usados para a detecção de “cantos” em imagens 2D foram estendidos para o caso 3D. O primeiro deles (Op1) é baseado no trabalho de Kitchen e Rosenfeld (1982) e utiliza derivadas parciais de primeira e segunda ordem (matriz Hessiana) para definir uma métrica de curvatura média das superfícies isointensas, a qual é multiplicada pela magnitude do gradiente local. O segundo (Op2) é baseado no trabalho de Blom (1992) e tem uma abordagem muito similar, modificando apenas a ordem do termo de magnitude do gradiente, fazendo com que este tenha mais peso. O terceiro operador (Op3) utiliza apenas derivadas parciais de primeira ordem e é baseado na métrica de Förstner e Gülch (1987), definida em termos da razão entre o determinante e o traço da matriz de momentos de segunda ordem da imagem. Por fim, o quarto operador (Op4), formulado a partir de um trabalho prévio do próprio Rohr (1992), é análogo ao anterior, mas utiliza como métrica apenas o determinante da matriz de momentos. Em todos esses casos, a localização dos pontos salientes da imagem é definida como os pontos maximais das referidas funções de métrica. Outras abordagens diferenciais adaptadas para imagens 3D, descritas em outros trabalhos (HARTKENS; ROHR; STIEHL, 2002), incluem as métricas baseadas em curvatura média e curvatura Gaussiana (FLORACK et al., 1994) e o operador de Beaudet (1978), definido como o determinante da matriz Hessiana (*Determinant of Hessian* - DoH) e que é equivalente ao termo numerador na fórmula da curvatura Gaussiana.

Hartkens, Rohr e Stiehl (2002) realizaram uma avaliação criteriosa de todas as técnicas mencionados no parágrafo anterior. Eles conduziram testes em imagens 3D sintéticas e imagens clínicas do cérebro humano. As 22 imagens sintéticas utilizadas representavam superfícies de tetraedros, elipsoides e paraboloides hiperbólicos modelados com diversos valores de parâmetros e suavizados com filtro Gaussiano. O conjunto de imagens médicas, por sua vez, foi constituído por 4 imagens 3D de RM em T1-w (tamanho de voxel entre 0,85mm e 1mm; espessura de fatia entre 1mm e 1,2mm) e 1 imagem de CT (tamanho de voxel 0,65×0,65×1mm). Em cada uma destas imagens, 10 marcadores anatômicos foram manualmente detectados por especialistas, constituindo o padrão ouro dos experimentos. As técnicas foram avaliadas em ter-

mos de: (a) invariância no número de pontos salientes detectados em relação ao nível de ruído Gaussiano e a diferentes transformações locais; (b) grau estatístico de confiança das detecções, medido pelas frações de pontos detectados, detecções corretas, detecções múltiplas e média de detecções por ponto do padrão ouro; e (c) separabilidade nas distribuições das respostas dos operadores, para fins de eliminação de falsos positivos mediante limiarização. Em resumo, os resultados apontaram que: (a-b) na maioria dos casos, tanto nas imagens sintéticas quanto nas imagens médicas, os operadores Op3 e Op4 das abordagens de Förstner e Gülch (1987), baseados apenas em diferenciais de primeira ordem, se mostram mais estáveis em relação ao ruído e às transformações locais, detectando portanto um menor número de pontos, mas uma maior fração de detecções corretas e uma menor fração de detecções falsas; (c) os mesmos operadores também mostraram os melhores resultados em termos da separabilidade entre as detecções corretas e os falsos positivos nas distribuições de suas respostas.

Wörz e Rohr (2005) desenvolveram um método para a localização de pontos salientes 3D que é baseado no ajuste da informação de intensidade da imagem a modelos paramétricos representativos de três classes distintas de estruturas 3D: “ponta”, “cela” e “esfera” (*tip-like*, *saddle-like* e *sphere-like*, respectivamente). Para melhorar a precisão dos ajustes, os parâmetros de cada modelo descrevem não apenas a forma básica da estrutura, como também algumas possíveis transformações globais e locais que esta pode sofrer. A posição de um ponto saliente corresponde aos pontos de máxima curvatura nas estruturas do tipo “ponta” e “cela”, e ao centro das estruturas do tipo “esfera”. O ajuste de cada modelo às intensidades da imagem é feito dentro de uma região de interesse seguindo um critério de minimização do erro quadrático; para isso, é utilizado o método de otimização iterativa de Levenberg-Marquardt (PRESS et al., 1993). O tamanho das regiões de interesse pode ser inicializado automaticamente ou semi-automaticamente. A posição inicial dos pontos salientes e regiões de interesse é definida conforme os pontos detectados pelos operadores Op3 e Op4 avaliados no trabalho de Hartkens, Rohr e Stiehl (2002). Por fim, ainda é utilizada uma abordagem multirresolução para aumentar a robustez e precisão do algoritmo. Os resultados foram avaliados em: (a) imagens sintéticas gerados pelos modelos paramétricos (com variações nos parâmetros e níveis de ruído Gaussiano); e (b) três imagens clínicas 3D (duas de RM e uma de CT) com dimensões e tamanho de voxel variados. A métrica de avaliação foi, em cada caso: (a) o erro de localização (em voxels) dos pontos em relação à localização definida no modelo; e (b) o erro de localização (em milímetros) dos pontos em relação a 30 marcadores anatômicos manualmente detectados, em 12 diferentes estruturas do cérebro, por especialistas. No caso (a), os resultados apontaram erros médios de 0,04mm nas estruturas do tipo “ponta”, 1,80mm nas do tipo “sela” e 0,16mm nas do tipo “esfera”. No caso (b), o erro médio global, envolvendo todos os tipos de marcadores, foi de 1,63mm com a inicializa-

ção automática das regiões de interesse, e 1,25mm com inicialização semiautomática; ambos desempenhos superiores aos operadores Op3 e Op4 usados isoladamente (erros de 2,73mm e 2,33mm, respectivamente).

Lowe (2004) propôs a técnica *Scale Invariant Feature Transform* (SIFT), que alcançou grande popularidade devido à sua capacidade em extrair pontos salientes estáveis e representá-los por descritores que são invariantes em relação à escala, translação, orientação, transformações afins e - em menor grau - mudanças de iluminação. A técnica pode ser resumida em quatro grandes passos: (a) Detecção de extremos no espaço-escala, utilizando a função de diferença de Gaussianas (*Difference of Gaussians* - DoG) e fazendo uma busca pelos máximos e mínimos das respostas na escala atual e nas adjacentes; (b) refinamento da localização dos pontos, pelo ajuste de uma função quadrática em torno de cada posição candidata a fim de eliminar pontos instáveis (com baixo contraste), e também pelo cômputo da razão entre os autovalores da matriz Hessiana na região de cada ponto para eliminar as respostas correspondentes a arestas (com razão dos autovalores maior que um determinado limiar); (c) atribuição de uma ou mais orientações a cada ponto detectado, de acordo com a magnitude e orientação dos gradientes na região, a fim de se poder representar cada ponto de maneira relativa à sua orientação, escala e localização (logo, alcançando invariância em relação a tais características); e (d) criação de um descritor para cada ponto, com base no cálculo da orientação e magnitude (ponderada por uma Gaussiana) dos gradientes de todos os pixels dentro de uma determinada região ao redor do ponto, cujo tamanho depende da escala em que este foi detectado. Essas informações são acumuladas em uma matriz de 16×16 histogramas e, posteriormente, em uma outra matriz de 4×4 histogramas, com 8 classes de orientações cada, representando a distribuição dos gradientes em cada uma dessas sub-regiões. O vetor que representa o descritor é normalizado para ter norma unitária e, assim, tornar o descritor menos sensível às condições de iluminação na imagem.

A técnica SIFT foi originalmente proposta para imagens 2D. No trabalho de Cheung e Hamarneh (2007), ela é generalizada para dados de n dimensões (n D-SIFT), com cada ponto descritor passando a representar as informações de uma região hipercúbica de tamanho 16^n ao redor da localização do ponto, a qual é dividida em 4^n sub-regiões com 8^{n-1} orientações em cada histograma. Os autores testaram a técnica para o caso 3D utilizando três imagens de RM obtidas da base de dados do projeto *BrainWeb*, com dimensões $90 \times 108 \times 90$ e voxels isotrópicos de 1mm^3 , sendo uma imagem ponderada em T1-w, outra em T2-w e a terceira em PD-w. Os resultados foram avaliados por meio da comparação entre os pontos descritores detectados nas imagens originais e em versões transformadas das mesmas (por rotação de 10° , redimensionamento em 80% do tamanho original, ou ambas juntas); ou seja, testou-se a invariância da técnica

a essas operações. Verificou-se que, para os menores valores de erro de localização tolerados para uma correspondência entre pontos (distância < 1,5), a fração de pontos correspondidos foi de 0,7 para a operação de redimensionamento, 0,83 para a rotação e 0,64 para ambas juntas. Para os maiores valores de erro de localização (distância < 7,5), as frações foram de 0,99 para o redimensionamento, 0,94 para a rotação e 0,89 para ambas juntas.

Liu et al. (2008) desenvolveram um método baseado em modelos pré-construídos e segmentação semi-global. Dado o atlas 3D de Tailarach-Tournoux (LIU; NOWINSKI, 2006; TALAIRACH; TOURNOUX, 1988) e um conjunto de imagens com marcadores manualmente detectados, a estrutura de interesse é segmentada usando a ferramenta *Model-based Image Understanding Environment* (MIUE) (LIU et al., 2007), obtendo-se uma superfície (malha 3D) que a representa em uma vizinhança local ao redor do marcador. As superfícies de todas as imagens são alinhadas ao atlas utilizando os marcadores correspondentes como âncoras. Com essas superfícies alinhadas, é estimado um tamanho ótimo para a região de interesse circular em torno de cada marcador. O modelo construído para cada ponto saliente, assim, envolve sua localização, a superfície local ao seu redor (dentro da região de interesse circular estimada) e a porção correspondente de volume no atlas. Quando uma imagem clínica é analisada, os seguintes passos são realizados: (a) corre-gistro do atlas e do modelo para o espaço da imagem clínica; (b) uso da técnica *Gradient Vector Flow* (GVF) (XU; PRINCE, 1998) para refinar a posição da superfície na imagem; (c) segmentação, na imagem clínica, da estrutura na vizinhança do ponto, usando uma técnica de superfícies ativas; (d) cálculo de posições candidatas dos marcadores, com base em propriedades da superfície segmentada; (e) seleção da melhor posição para o marcador, dentre as candidatas, através da minimização do erro do corre-gistro entre as superfícies obtidas em (a) e em (c) usando cada ponto candidato como âncora. Para avaliar a técnica, os autores verificaram o erro de localização na detecção de 10 marcadores anatômicos do sistema ventricular cerebral, em relação à detecção manual feita por dois especialistas. Os experimentos utilizaram ao todo 66 imagens de RM em T1-w, sendo 18 delas obtidas do projeto *BrainWeb* (dimensões $181 \times 217 \times 181$ e voxels isotrópicos de 1mm^3 , simulados com níveis variados de ruído e heterogeneidade do sinal), 18 da base *Internet Brain Segmentation Repository*¹ (IBSR) (dimensões $256 \times 256 \times 128$ e voxels de tamanho $1 \times 1 \times 1.5\text{mm}^3$), e 30 de imagens clínicas diversas com pacientes normais e patológicos (dimensões e voxels anisotrópicos de tamanhos variados). Os resultados apontaram erros médios de localização de 0,5mm nas imagens simuladas do *BrainWeb*, de 3,3mm nas imagens da base IBSR, e de 2,4mm nas imagens clínicas. Os erros máximos nas referidas bases de dados foram de 2mm, 21,3mm e 23mm, respectivamente. A técnica se mostrou robusta às influências do ruído, heterogeneidade de sinal e variações anatômicas.

¹<http://www.cma.mgh.harvard.edu/ibsr/data.html>

Ferrari et al. (2011) utilizaram uma abordagem baseada na medida de congruência de fase proposta por Kovesei (KOVESI, 2000), estendendo-a para o caso 3D utilizando o banco de filtros de log-Gabor 3D proposto por Dosil, Pardo e Fdez-Vidal (2005). A medida de congruência de fase tem importante aplicação na detecção de bordas e cantos em imagens (KOVESI, 2003), produzindo respostas de pico nas regiões em que os componentes de Fourier encontram-se maximalmente em fase e, assim, tendo a característica de invariância às mudanças nos níveis de contraste e iluminação da imagem. A partir do mapa de congruência de fase, uma matriz 3×3 de momentos de segunda ordem é computada. É definida, então, uma métrica de saliência baseada no produto normalizado dos autovalores dessa matriz. A localização dos pontos salientes é dada pelos voxels onde tal métrica produz valores maximais. A técnica foi testada no contexto dos métodos de propagação de contornos entre imagens de diferentes modalidades em terapia por radiação - no caso, de CT para *Cone Beam CT* (CBCT). Mais precisamente, os pontos salientes correspondentes que foram detectados nas imagens de CT e CBCT são utilizados como âncoras para a interpolação dos contornos de uma modalidade de imagem para outra, estimando-se uma transformação *Thin-Plate Spline* (TPS) para tal fim. A correspondência entre pontos é baseada na similaridade local entre a informação de congruência de fase ao redor de cada ponto. Para fins de validação, duas estruturas (tronco encefálico e medula espinhal) tiveram seus contornos manualmente delineados por um especialista em 6 imagens de cada tipo (padrão ouro). Os resultados foram então avaliados em termos do erro de localização de cada vértice interpolado em relação ao padrão ouro. Esse erro foi em média de 2,66mm para a medula espinhal e 3,23mm para o tronco encefálico - resultados superiores do que os obtidos pela simples transferência do contorno original de uma imagem para a outra (o que produziu erros de 4mm e 3,57mm, respectivamente).

Mais recentemente, Han et al. (2014) propuseram uma abordagem baseada em aprendizado, na qual a detecção dos marcadores anatômicos é formulada como um problema de regressão não-linear multivariada. Dado um conjunto de imagens de treinamento alinhadas e com marcadores previamente definidos, um algoritmo de floresta de regressão (combinação de múltiplas árvores de decisão) (BREIMAN, 2001) é utilizado para aprender o melhor conjunto de características que representa cada marcador, bem como os mapeamentos não-lineares que são utilizados para estimar a localização dos marcadores correspondentes em novas imagens clínicas. O algoritmo é treinado a partir de conjuntos de voxels \mathbf{p}_j que são amostrados da vizinhança de cada marcador \mathbf{v}_j e codificados, cada um, por um vetor de características que inclui medidas derivadas de padrões intensidade locais (CRIMINISI et al., 2013) e o desvio $\mathbf{d}_j = \mathbf{v}_j - \mathbf{p}_j$. Quando uma nova imagem clínica é analisada, cada voxel \mathbf{p}_i dessa imagem é codificado e processado pelas árvores de decisão e elas produzem, como saída, valores $\bar{\mathbf{d}}$ correspondentes aos desvios

médios estimados até um marcador \mathbf{v}_j . Sendo $\hat{\mathbf{d}}$ a média dos valores de $\bar{\mathbf{d}}$ em todas as árvores, a estimativa para a localização do marcador mais próximo é então dada por $\mathbf{p}_i + \hat{\mathbf{d}}$. Esses “saltos” são repetidos um certo número de vezes para cada voxel, e então a localização final dos marcadores de uma imagem é estabelecida como a posição final dos pontos \mathbf{p}_i para os quais os valores de $\|\hat{\mathbf{d}}\|$ são mínimos. Ainda é utilizada uma estratégia de multirresolução para aumentar a robustez do método. Os autores avaliaram a técnica de duas maneiras: analisando a precisão de localização dos marcadores automaticamente detectados em relação aos de imagens manualmente marcadas, e analisando os mapeamentos estimados no contexto do corregristo de imagens. Foram selecionadas 62 imagens 3D de RM da base de dados ADNI², com dimensões $256 \times 256 \times 256$ e voxels isotrópicos de 1mm^3 . 44 dessas imagens foram usadas para treinamento e 18 para testes. Na primeira avaliação, os resultados mostraram que os erros médios de localização de um marcador são de $1,48 \pm 0,77\text{mm}$ no melhor caso e de $2,77 \pm 1,24\text{mm}$ no pior caso. Na segunda avaliação, foram comparados os coeficientes de similaridade de Dice envolvendo versões segmentadas das imagens corregristadas, com a aplicação do corregristo acontecendo em três casos distintos: (a) somente usando os mapeamentos entre marcadores; (b) somente usando o algoritmo HAMMER (SHEN; DAVATZIKOS, 2002); e (c) usando ambas as técnicas combinadas. Os resultados mostraram que o caso (c) produziu os valores mais favoráveis, indicando a utilidade dos marcadores em melhorar o desempenho de outras técnicas de corregristo.

3.2.2 Discussão

Nesta seção, foram apresentadas algumas abordagens recentes propostas na literatura para a solução do problema de detecção de pontos salientes (ou marcadores anatômicos) em imagens 3D. Embora muitas delas tenham tido um desempenho considerado satisfatório nos experimentos realizados pelos trabalhos aqui mencionados, é importante notar que elas também possuem as suas limitações. Por exemplo, no próprio trabalho de Hartkens, Rohr e Stiehl (2002), constata-se que os operadores diferenciais Op1 e Op2, o DoH e a curvatura Gaussiana apresentam grande sensibilidade ao ruído e podem produzir respostas espúrias e mal localizadas espacialmente. Abordagens baseadas em aprendizado, como a de Han et al. (2014), têm a desvantagem de necessitar que o algoritmo seja previamente treinado com imagens marcadas por um especialista. A técnica baseada em modelos de Liu et al. (2008) é dependente da segmentação realizada por uma ferramenta externa, além de também depender de um conjunto de imagens manualmente marcadas.

²Vide Subseção 4.2.3.

Além disso, há também de se considerar que as outras abordagens baseadas em diferenciais e vetores gradiente - como os operadores Op3 e Op4 de Rohr (1997), o método de Wörz e Rohr (2005) e a técnica nD-SIFT (CHEUNG; HAMARNEH, 2007) - podem ser afetadas pela influência da heterogeneidade de intensidade (*bias field*) que frequentemente existe nas imagens de RM. Essa heterogeneidade causa mudanças não-monotônicas nos níveis de intensidade e contraste da imagem, consequentemente afetando a magnitude dos vetores gradiente. A técnica baseada em congruência de fase de Ferrari et al. (2011) não sofre tanto com esse problema, pelo fato de definir saliência a partir da informação de fase; porém, ela exige um ajuste preciso do termo de supressão de ruído no cálculo do mapa de congruência de fase (KOVESI, 2000), o que, em imagens 3D de RM, não se mostra algo tão trivial. Além disso, sua métrica de saliência tende a realçar não apenas cantos, mas também algumas respostas de bordas e linhas (embora em imagens de CT isso não seja tão óbvio devido ao maior contraste em relação a imagens de RM).

3.3 Construção de atlas do cérebro

No contexto do processamento e análise de imagens médicas, os atlas anatômicos são um recurso normalmente utilizado para embutir, em um determinado modelo matemático, um conhecimento *a priori* da informação anatômica de um determinado órgão, com o objetivo de melhorar o desempenho de algoritmos que utilizem tal modelo - como, por exemplo, métodos de segmentação (KALINIC, 2008). Um atlas anatômico pode descrever diferentes aspectos de um órgão, incluindo distribuição espacial de tecidos e formato e localização típica de estruturas. Além disso, pode ser construído tanto a partir de uma única amostra (imagem) quanto de uma população (várias imagens). Costuma-se fazer a diferenciação entre atlas topológicos (determinísticos) e atlas probabilísticos (que refletem graus de incerteza conforme o perfil de variabilidade anatômica da população na qual foram estimados) (CABEZAS et al., 2011).

Nos últimos anos, o desenvolvimento de técnicas eficientes de construção de atlas de diferentes características do cérebro para diferentes fins tem se tornado um tema relevante de pesquisa (CABEZAS et al., 2011). Assim, a subseção a seguir descreve algumas dessas técnicas.

3.3.1 Trabalhos correlatos

Uma das abordagens mais antigas e clássicas para construção de atlas probabilísticos foi proposta ainda na década de 1990 por Evans et al. (1993). Nesse trabalho, um mapa probabilístico de intensidades médias do cérebro é produzido a partir de um conjunto de imagens 3D de RM espacialmente alinhadas por uma técnica de corregristo usando transformações afins, que

realiza a otimização de 9 parâmetros (envolvendo translação, rotação e escala). O algoritmo também utiliza uma abordagem multirresolução para reduzir a influência de mínimos locais e o tempo de processamento. O método de otimização utilizado no correregistro é o algoritmo SIMPLEX (MURTY, 1983) e a métrica de similaridade é a função de correlação cruzada (*Feature Cross Correlation Function - FCCF*) (COLLINS et al., 1994) envolvendo características de intensidade e magnitude dos gradientes da imagem. O alinhamento é realizado em dois passos: (a) um sistema de coordenadas de referência é determinado a partir de um conjunto de marcadores anatômicos manualmente identificados em cada imagem do conjunto e no atlas topológico de Talairach-Tournoux (TALAIRACH; TOURNOUX, 1988), de maneira a se definir a transformação que mapeia cada imagem para este espaço. As imagens transformadas são reamostradas (utilizando interpolação tricúbica) para dimensões $256 \times 256 \times 150$ e tamanho de voxel $0,67 \times 0,86 \times 0,75$ mm, e a partir disso é computada uma imagem média de todo o conjunto (denominada MNI-305) de acordo com a distribuição dos níveis de intensidade em cada voxel. (b) Cada imagem original do conjunto é novamente transformada, desta vez para o espaço da imagem média. O objetivo disso é reduzir o viés introduzido pela escolha dos marcadores no primeiro estágio, além de resultar em um mapa de intensidades visualmente menos borrado (isto é, com menor variância estatística).

O método de Evans et al. (1993) foi testado em um conjunto de 305 imagens 3D de RM ponderadas em T1-w e obtidas no *Montreal Neurological Institute* (MNI), retratando pacientes com idades de $23 \pm 4,1$ anos, dentre os quais 239 eram homens e 66 eram mulheres. Para a aquisição, foi utilizado um scanner Philips 1,5T. Os resultados, em termos do atlas produzido, foram avaliados qualitativamente, constatando-se uma maior variância nas regiões de bordas salientes (como nos limites dos ventrículos) e no córtex, e menor variância em regiões de matéria branca, proporcionando assim uma melhor referenciação anatômica destas últimas.

Wang et al. (2008) desenvolveram uma técnica que realiza simultaneamente o correregistro de múltiplos conjuntos de pontos (que podem representar estruturas anatômicas) e a construção de um atlas probabilístico da distribuição desses pontos após o alinhamento. A ideia é modelar cada conjunto de pontos por uma função de densidade de probabilidade e definir uma métrica para quantificar a divergência entre essas densidades, de maneira que, pela otimização (minimização) de tal medida, se encontrem as transformações não-rígidas que melhor alinham os conjuntos de pontos em um mesmo espaço de coordenadas de maneira livre de viés. No caso, cada função de densidade de probabilidade é representada por um modelo de mistura de Gaussianas cujos componentes são determinados conforme os centroides dos agrupamentos (*clusters*) identificados em cada conjunto de pontos. Para se encontrar tais centroides, é utilizada a técnica de agrupamento *deterministic annealing* (YUILLE; STOLORZ; UTANS, 1994). A divergência

de Jensen-Shannon (JS) (LIN, 1991) - um valor computado a partir da medida de entropia - foi escolhida como métrica de dissimilaridade conjunta entre as densidades de probabilidade. Assim, pela formulação matemática desenvolvida no trabalho, o corregristo e a construção do atlas reduzem-se a um problema de otimização que busca encontrar os parâmetros das transformações não-rígidas que minimizam a métrica JS. Para restringir o grau de deformação dos pontos, utiliza-se ainda um termo de regularização/suavização das transformações (CHUI; RANGARAJAN, 2003). Como resultado, tem-se um atlas definido pela densidade do conjunto médio de pontos para o qual todos os demais conjuntos foram mapeados durante o corregristo.

Os resultados da técnica de Wang et al. (2008) foram avaliados para conjuntos de pontos 2D e 3D. No caso 2D, extraíram-se pontos ao longo dos contornos externos do corpo caloso em 7 imagens 2D de RM (vista sagital) de cérebros de pacientes normais. 3 variações de atlas foram construídos pelo uso de diferentes valores para o parâmetro de regularização. Em uma análise qualitativa, observou-se que os atlas produzidos permaneceram relativamente estáveis mesmo com tal variação. Além disso, também se analisou o efeito da adição de um oitavo conjunto de pontos, e novamente o resultado foi um atlas ainda estável em relação aos anteriores. Para o teste no caso 3D, foram extraídos conjuntos de pontos representando os hipocampus em 10 imagens 3D de RM de cérebros de pacientes com epilepsia. Neste caso, a avaliação foi feita pela comparação dos conjuntos de pontos sobrepostos antes e após o corregristo, mostrando que este último caso tem uma distribuição mais compacta. Quantitativamente, avaliou-se a capacidade do método em realizar o corregristo entre pontos na presença de ruído, *outliers* e transformações de rotação e translação. A métrica de avaliação foi a raiz do erro médio quadrático, para a qual se obteve valores de até 0,04 para a rotação e até 6,0 para a translação.

O trabalho de Shattuck et al. (2008) descreve o desenvolvimento de um atlas probabilístico de estruturas corticais do cérebro. As imagens de RM usadas para treinamento passam pelo seguinte pré-processamento: (i) alinhamento afim com o espaço de referência MNI-305 (EVANS et al., 1993); (ii) correção do campo de heterogeneidade usando o algoritmo N3 (SLED; ZIJDENBOS; EVANS, 1998)³; e (iii) extração automática do encéfalo com a ferramenta *Brain Extraction Tool* (BET) (SMITH, 2002). Após isso, um grupo de especialistas realizou nessas imagens a segmentação manual de 50 estruturas corticais, 4 áreas subcorticais, tronco cerebral e cerebelo. Três variações de atlas foram então produzidas de acordo com a imagem de referência escolhida e com a técnica de corregristo não-linear utilizada para o alinhamento com a referência: (a) algoritmo AIR 5.2.5 (WOODS et al., 1998), usando como referência o atlas ICBM-452 Warp5 (REX; MA; TOGAA, 2003); (b) algoritmo FLIRT (SMITH et al., 2004), usando como referência o atlas ICBM-152 T1; e (c) algoritmo SPM5 (ASHBURNER; FRISTON, 2005), também

³Vide Subseção 4.3.2.

tendo o ICBM-152 T1 como referência. Antes da realização dos corregristos, cada imagem foi automaticamente segmentada em 3 classes de tecidos - matéria branca, matéria cinzenta e líquido cefalorraquiano - usando a ferramenta *Partial Volume Classifier* (PVC) (SHATTUCK et al., 2001), a qual produz mapas com as estimativas da fração de tecido em cada voxel. As imagens de treinamento e os mapas de tecidos foram transformados para os espaços de seus respectivos atlas de referência conforme (a-c). O mesmo foi feito nas imagens com estruturas corticais segmentadas; porém, neste caso, o procedimento foi feito de maneira separada para cada estrutura de cada imagem. Pela média dos valores nas imagens de cada estrutura segmentada e interpolada para o espaço do atlas, ao longo das imagens de todos os pacientes, foi produzido um conjunto de mapas estruturais (um para cada estrutura segmentada). Multiplicando-se esses mapas pelo mapa de densidade de matéria cinzenta de cada paciente, e calculando-se as médias, foi obtido um novo conjunto de mapas com as densidades de cada estrutura na matéria cinzenta. O atlas desejado, por fim, é computado atribuindo-se a cada voxel o rótulo identificador da estrutura com o maior valor, naquela posição, ao longo dos mapas de densidades estruturais na matéria cinzenta.

Para o treinamento, foi utilizado um conjunto de 40 imagens obtidas no *North Shore-Long Island Jewish Health System* com um scanner GE 1,5T. As imagens foram adquiridas de pacientes saudáveis, com idades entre 16 e 40 anos, proporção igual de homens e mulheres, níveis educacionais variados (em média, nível superior incompleto) e etnicamente diversos (a maioria caucasianos). A segmentação manual das estruturas corticais foi feita por um grupo de 15 especialistas seguindo critérios rigorosos para assegurar a confiabilidade dos resultados e reduzir ao máximo a variabilidade inter observador. Qualitativamente, observou-se que o atlas construído em (a) possui uma melhor definição visual em relação às outras duas variações, pelo fato de a imagem de referência utilizada possuir maior resolução. Quantitativamente, avaliou-se o volume médio de cada estrutura em relação aos correspondentes volumes estimados em cada uma das 3 variações do atlas. Nas variações (b) e (c), os resultados apontaram uma expansão volumétrica média de 37,7% e 37,8%, respectivamente. Em (a), essa expansão foi de apenas 10,3%. Os autores explicam essa diferença no fato de que o cérebro usado na imagem de referência em (a) (ICBM-452 Warp5) possui um tamanho mais próximo dos cérebros da população analisada. Também se verificou a existência de assimetrias entre algumas estruturas, tais como o hipocampo (que, para 80% dos pacientes testados, é maior no lado direito do cérebro).

Leemput (2009) argumenta que os atlas probabilísticos construídos a partir da computação de valores médios (ou frequências relativas) ao longo de um conjunto de imagens de treinamento não necessariamente expressam da melhor maneira as estatísticas da população repre-

sentada nesse conjunto. Dois problemas com tal abordagem seriam o *overfitting*⁴ causado pelo número limitado de imagens de treinamento, em relação à grande variabilidade anatômica de estruturas corticais entre diferentes indivíduos, e o fato de que tais atlas não expressam os mapeamentos não-lineares que alinham as imagens em um espaço anatômico comum e descrevem variações anatômicas locais⁵. Diante desse problema, o autor propôs um modelo mais generalista para a construção de atlas probabilísticos, no qual os atlas são representados por malhas deformáveis e a inferência Bayesiana é utilizada para encontrar os parâmetros do modelo (configuração de malha) que melhor representa e generaliza os dados das imagens de treinamento pré-segmentadas. Esses parâmetros descrevem tanto as deformações locais da malha (posições dos vértices) como também as frequências relativas de cada rótulo (nas imagens segmentadas) na vizinhança espacial de cada vértice da malha. Os atlas tradicionalmente computados via média ou frequência relativa seriam, então, um caso particular dessa abordagem - isto é, uma escolha particular de modelo de atlas. O autor ainda faz um paralelo de sua abordagem com a teoria da informação, mostrando que computar o logaritmo negativo da verossimilhança dos dados de treinamento em relação a um dado modelo de atlas equivale matematicamente a medir o comprimento da mensagem binária mais curta que codifica os dados de treinamento quando esse mesmo modelo de atlas é utilizado. Isso também pode ser entendido como um processo de compressão dos dados.

Assim, dado o conjunto de imagens de treinamento e suas respectivas rotulações (estruturas segmentadas), o processo de inferência Bayesiana de Leemput (2009) consiste basicamente em otimizar as posições espaciais dos vértices da malha (o que se traduz no corregristo das imagens de treinamento) e, simultaneamente, otimizar também as probabilidades de cada rótulo em cada vértice da malha conforme as rotulações nas imagens segmentadas. Essa otimização é feita utilizando o algoritmo *Expectation-Maximization* (EM), inicializado com probabilidades iguais para todos os rótulos em todos os vértices⁶. Também é feita uma estimação por Máxima a Posteriori (MAP) para o parâmetro que define o grau de flexibilidade da deformação das malhas que melhor se compatibiliza às imagens de treinamento. Por fim, um terceiro processo de otimização computa a resolução ótima para a malha deformada, a fim de se reduzir o *overfitting* e produzir um atlas com o grau adequado de generalização de acordo com as imagens

⁴*Overfitting* é o termo usado para denominar o ajuste excessivo de um modelo estatístico ao seu respectivo conjunto de amostras, de maneira que esse modelo acaba descrevendo o ruído presente nos dados (erros de medição, casos anômalos, etc.) e, conseqüentemente, tem uma representação matemática demasiado complexa (com muitos parâmetros). Isso é um problema porque tais modelos são pobres em termos de generalização e capacidade de predição; logo, não constituem uma descrição adequada da população (MITCHELL, 1997).

⁵Essas observações afetam diretamente trabalhos como os de Evans et al. (1993) e Shattuck et al. (2008), descritos anteriormente.

⁶O atlas computado via média seria o caso particular em que a malha é regular, não admite deformações e há um vértice para cada voxel da imagem.

de treinamento. A malha final é convertida para uma imagem de atlas pela interpolação das probabilidades associadas a cada vértice aos pixels/voxels no domínio da imagem.

O método de Leemput (2009) é utilizável tanto em dados 2D (para os quais se tem malhas triangulares) quanto 3D (malhas tetraédricas). Para a avaliação do atlas, foram utilizados dois conjuntos de treinamento distintos, com imagens de RM do cérebro obtidas da base IBSR. O primeiro conjunto consistiu em 18 imagens de cortes coronais, segmentadas conforme as regiões de matéria branca, córtex, ventrículo lateral, caudado, putâmen e núcleo accumbens, em ambos os hemisférios cerebrais. O segundo conjunto foi composto por cortes axiais nas imagens obtidas dos mesmos 18 pacientes, desta vez com a segmentação feita em 4 classes: matéria branca, matéria cinzenta, fluido cerebrospinal e fundo. Quatro modelos diferentes de geração de atlas foram testados: (a) usando malha regular não-deformável e com máxima resolução (cada vértice corresponde a um pixel); (b) usando malha regular não-deformável e com resolução ótima estimada automaticamente; (c) usando malha não-regular não-deformável e automaticamente adaptável ao conteúdo (com fusão de vértices e simplificação das malhas onde for conveniente); e (d) usando malha não-regular otimamente deformável e adaptável ao conteúdo. Em termos quantitativos, a avaliação dos atlas foi realizada pela análise do comprimento da mensagem binária mais curta que codifica os dados em cada caso, em relação ao que seria uma mensagem sem compressão na qual se assume uma distribuição completamente aleatória para os rótulos. Em (a), foi observada uma compressão de aproximadamente 64% em relação à mensagem com rótulos aleatórios; entretanto, a maior parte dessa mensagem comprimida é usada para codificar os parâmetros do modelo, o que é um sinal de *overfitting*. Em (b), a malha produzida tem cerca de 30 vezes menos vértices do que em (a) e o comprimento da mensagem é cerca de 42% do comprimento obtido em (a), o que se traduz em uma imagem de atlas mais borrada. Em (c), o comprimento da mensagem binária é reduzido em mais 10%, e a irregularidade das malhas faz com que o grau de borramento seja distribuído de maneira não uniforme. Finalmente, o caso (d) produziu as mensagens de menor comprimento, ainda que a quantidade de bits necessária para se codificar as posições dos vértices na mensagem tenda a crescer conforme o grau de flexibilidade admitida para a deformação das malhas.

Lotjonen et al. (2011) propõem três abordagens distintas para geração de atlas probabilísticos visando uso em segmentação de imagens de RM do cérebro, dado um conjunto de atlas topológicos pré-segmentados e uma imagem de referência (*template*). Sendo definidas as transformações não-rígidas que fazem o mapeamento entre a imagem de referência e cada atlas topológico, a primeira abordagem (*Probabilistic Atlas in Template Space - PATS*) estabelece que o atlas probabilístico seja computado no espaço da imagem de referência, realizando-se as devidas transformações inversas e então calculando-se a probabilidade de cada voxel per-

tencer a uma determinada estrutura conforme a simples frequência relativa de sua classificação em cada estrutura feita nas segmentações dos atlas topológicos. Em seguida, uma transformação não-rígida da imagem de referência para a imagem clínica analisada transfere esse atlas probabilístico para o espaço da imagem clínica. A segunda abordagem (*PATS with Atlas Selection* - PATSAS) é bastante similar, com a diferença de empregar a métrica de informação mútua para selecionar apenas os atlas mais próximos da imagem clínica. A terceira abordagem (*Probabilistic Atlas in Unseen data Space with Atlas Selection* - PAUSAS), por fim, trabalha com mapeamentos entre a imagem clínica e os atlas topológicos, gerando o atlas probabilístico diretamente no espaço da imagem clínica sem a intermediação da imagem de referência.

As três abordagens de Lotjonen et al. (2011) foram testadas no contexto da segmentação multiatlas do hipocampo em imagens de RM. Foram utilizadas, nos testes, 340 imagens obtidas da base de dados ADNI. Essas imagens incluem pacientes cognitivamente normais (CN), com comprometimento cognitivo leve (CCL) e pacientes diagnosticados com a doença de Alzheimer (DA). Para o conjunto de atlas, foram usadas 30 imagens, mais 1 imagem escolhida para ser a referência (*template*). Como métricas de avaliação, foram usadas o coeficiente de similaridade de Dice (DC) e os coeficientes de correlação entre os volumes calculados para o hipocampo nas segmentações automáticas e também em segmentações manuais. Os melhores resultados foram obtidos com a abordagem PAUSAS, com $DC = 0,87 \pm 0,04$ e correlação 0,96 (equiparável às melhores segmentações manuais). No entanto, ela também apresenta o maior custo computacional, visto que não permite o pré-cômputo das transformações entre referência e atlas, como nas abordagens PATS e PATSAS.

Mais recentemente, Mondal et al. (2014) propuseram uma metodologia para a geração de um atlas de pontos descritores produzidos pela técnica 3D-SIFT. Dado um conjunto de imagens 3D de RM retratando cérebros normais, o objetivo é encontrar os pontos descritores anatomicamente invariantes, isto é, aqueles que são comuns a um grande número de imagens. Pelo uso de uma técnica de correção (*Model Based MRI Alignment* - MBMRIA (MONDAL et al., 2012)), as imagens de RM são primeiramente alinhadas a um sistema de coordenadas comum, e então, com a aplicação da 3D-SIFT, cada cérebro passa a ser representado por um mapa de pontos descritores com 512 dimensões (características) cada. Para se avaliar a existência de uma correspondência entre pontos descritores de diferentes imagens dentro de uma determinada região de busca, usa-se como métrica a razão entre as distâncias Euclidianas aos dois descritores mais próximos. O atlas, por fim, é computado pela mesclagem incremental dos mapas de descritores, na qual cada par de pontos identificados como correspondentes é mesclado em um terceiro ponto descritor cujos valores consistem na média dos anteriores. Para cada ponto descritor, existe um contador que se acumula a cada mesclagem, de maneira que, ao final do processa-

mento de todos os mapas, somente os pontos cuja contagem estiver acima de um determinado limiar serão considerados invariantes e serão incluídos no atlas.

É digno de nota que, ainda no mesmo trabalho, os autores tenham utilizado o atlas para diagnosticar automaticamente os estágios iniciais da DA. Isso é realizado através da análise da consistência geométrica (localização anatômica) entre os pontos descritores do atlas com os de uma dada imagem clínica, e da posterior contagem da fração de pontos consistentes em comparação a um determinado limiar. Os experimentos foram feitos utilizando as imagens 3D de RM em ponderação T1-w da base de dados *Open Access Series of Imaging Studies*⁷ (OASIS), com 30 imagens sendo usadas para geração do atlas (15 pacientes homens e 15 mulheres) e outras 165 para os testes de diagnóstico. Foram estabelecidas três classes para classificação da DA: *normal*, *very mild* e *mild*, de acordo com os valores da métrica de diagnóstico *Clinical Dementia Rating* (CDR) (MORRIS, 1993) em cada imagem da base. Os resultados foram analisados a partir da matriz de confusão que correlaciona os resultados da técnica com o padrão ouro. Eles apontaram uma precisão de 70% na classificação das imagens da classe *normal*, 66,7% das imagens da classe *very mild* e 73,3% das imagens da classe *mild*. Em um segundo teste, no qual se trabalhou apenas com as classes *normal* e *mild* e foi feita uma análise de sensibilidade e especificidade, os resultados para essas métricas foram de 73% e 76%, respectivamente.

3.3.2 Discussão

Nesta seção, foram apresentadas algumas das diferentes abordagens na literatura para construção de atlas do cérebro. Essas abordagens incluem métodos baseados na computação de médias de intensidade (EVANS et al., 1993), frequências relativas de rótulos em imagens pré-segmentadas (SHATTUCK et al., 2008; LOTJONEN et al., 2011), minimização da divergência entre densidades de conjuntos de pontos (WANG et al., 2008) e modelos geradores baseados em malhas (LEEMPUT, 2009). Apesar das diferenças, observa-se que, em todas os casos, o corregristo preciso das imagens de treinamento é um passo imprescindível para a geração dos atlas. Esse corregristo deve ser preferencialmente livre de viés, razão pela qual se procura alinhar todas as imagens a um espaço médio comum ao invés de se estabelecer o espaço de uma imagem específica como referência.

O trabalho de Mondal et al. (2014) foi o único encontrado na literatura que realiza a construção de um atlas de pontos salientes. Entretanto, observa-se que ele se não se trata de um atlas probabilístico, e por essa razão é limitado quanto à capacidade de representação da variabilidade anatômica da população na qual fora construído. Além disso, sua utilização é específica

⁷<http://www.oasis-brains.org/>

para o diagnóstico da doença de Alzheimer via similaridade com as informações do atlas.

3.4 Inicialização de modelos deformáveis

Nas últimas duas décadas, os métodos de segmentação baseados em modelos deformáveis têm se estabelecido como uma das mais bem-sucedidas ferramentas para análise de imagens, especialmente na área médica (LAI, 1994; MCINERNEY; TERZOPOULOS, 1999; MONTAGNAT; DELINGETTE; AYACHE, 2001; WANG et al., 2005a, 2005b; LEAVENS et al., 2008; KAUS; BROCK, 2009; HEIMANN; MEINZER, 2009). Nessas abordagens, um modelo - geralmente representado por uma curva paramétrica, superfície paramétrica, malha ou nuvem de pontos - é iterativamente ajustado à imagem de entrada de acordo com dois funcionais de energia que controlam sua deformação: um funcional de energia interna, que controla sua elasticidade e rigidez, preservando a continuidade e suavidade do modelo, e um funcional de energia externa, que induz o modelo a convergir e se moldar às bordas do objeto que se deseja segmentar (HEIMANN; MEINZER, 2009). A definição desses funcionais, bem como os pesos dados a cada um de seus componentes, variam conforme a aplicação. Um dos principais atrativos desse tipo de abordagem em segmentação de imagens médicas é o fato de tais modelos já poderem trazer em si a topologia do órgão anatômico que representam; dessa forma, essa abordagem torna-se menos suscetível a ruídos das imagens e pode ser aplicada inclusive em situações nas quais o nível de contraste da anatomia em estudo na imagem é demasiadamente baixo (PEKAR; MCNUTT; KAUS, 2004; KAUS; BROCK, 2009).

Uma das principais dificuldades da aplicação de modelos deformáveis para a segmentação de imagens médicas é o posicionamento inicial adequado do modelo. Para o sucesso da segmentação e redução do custo computacional, o modelo deve ser preferencialmente posicionado em uma região muito próxima da estrutura que se quer segmentar na imagem, algo problemático especialmente para imagens contendo um grande número de estruturas próximas umas das outras (BOGOVIC; PRINCE; BAZIN, 2013). A subseção a seguir faz uma revisão de algumas das abordagens propostas na literatura para inicialização desses modelos.

3.4.1 Trabalhos correlatos

Existem diversos trabalhos focados na segmentação automática de estruturas anatômicas específicas em imagens 3D usando modelos deformáveis. No entanto, os métodos que eles propõem muitas vezes utilizam técnicas de inicialização que também são muito específicas para essas estruturas de interesse, fazendo uso de heurísticas e de um conhecimento a priori da ana-

tomia que não se aplicariam em outros casos. Dentre tais tipos de trabalhos, pode-se citar, por exemplo, alguns que utilizam técnicas para localização e delineamento do coração (ECABERT et al., 2008; SAALBACH et al., 2011; ZAMBAL et al., 2008), da próstata (CHANDRA et al., 2012), do fígado (SOLER et al., 2001), das cartilagens do joelho (FRIPP et al., 2005), dos parênquimas cerebrais (WU; PAULSEN; JR., 2005), da espinha dorsal (DELEENER; KADOURY; COHEN-ADAD, 2014; KOREZ et al., 2015), entre outras. Devido à falta de generalidade desse tipo de técnica, esta revisão se limitará a relatar com mais detalhes as abordagens mais generalistas (não tão restritas a uma estrutura anatômica específica) propostas nos últimos anos.

Heimann et al. (2007) propuseram um método de segmentação de estruturas anatômicas baseado em modelos de formas estatísticas (*Statistical Shape Model - SSM*) inicializados por uma estratégia de busca global inspirada em técnicas de programação evolutiva. Dado um conjunto de imagens de treinamento manualmente marcadas com pontos que delineiam a superfície de uma determinada estrutura anatômica, o SSM tem seu modelo geométrico composto por um modelo de distribuição de pontos (*Point Distribution Model - PDM*) estimado via aplicação de análise de componentes principais (*Principal Component Analysis - PCA*) aos vetores de coordenadas dos pontos manualmente marcados. A partir do modelo médio obtido no treinamento, a inicialização dos pontos em imagens clínicas é guiada por uma técnica evolutiva na qual se define uma métrica de “aptidão” e se mantém uma população de soluções que vão evoluindo conforme a regra de “sobrevivência do mais apto”. No caso, cada possível solução é dada por um conjunto de coordenadas de pontos e as modificações nas soluções são feitas adicionando-se a elas vetores aleatórios seguindo uma distribuição Gaussiana multivariada (também estimada a partir das imagens de treinamento). A métrica de aptidão é baseada na soma das probabilidades a posteriori que expressam a evidência estatística da estrutura representada pelo modelo na imagem. Quanto maior a aptidão de uma solução, maior a probabilidade de esta ser selecionada para se reproduzir e permanecer na próxima iteração. As soluções sofrem ainda uma mutação de acordo com um desvio padrão σ_t que vai decrescendo em 5% a cada iteração. A solução final (inicialização do modelo geométrico) é dada pelo modelo que possui a maior aptidão quando o algoritmo atinge a convergência. Podem ser usadas, ainda, estratégias para simplificação do SSM obtido, que é representado por uma malha 3D.

A técnica de Heimann et al. (2007) foi testada no contexto da segmentação do fígado em imagens 3D de CT. A segmentação é obtida a partir das deformações locais da malha do modelo geométrico, de acordo com os funcionais de energia interna e externa que procuram adaptá-la da melhor maneira em relação às informações da imagem. As imagens utilizadas tinham tamanho 512×512 pixels por fatia, com 60 a 130 fatias cada, e voxel de aproximadamente $0,7 \times 0,7 \times 3 \text{mm}^3$. Elas foram obtidas de pacientes majoritariamente patológicos (com tumores)

utilizando protocolos de aquisição variados. No total, foram usadas 86 imagens, dentre as quais 32 foram selecionadas para treinamento do modelo, e as demais 54, para testes. O algoritmo evolutivo foi executado numa população de 1000 instâncias (pontos) ao longo de 40 iterações, e a métrica de aptidão foi avaliada num conjunto de 204 instâncias na imagem com resolução 4 vezes menor. O desvio padrão inicial foi definido como $\sigma_0 = 0,4$. As 54 segmentações produzidas foram avaliadas pela comparação com a delimitação manual do fígado feita por um grupo de especialistas, utilizando três métricas de erro. Em média, os resultados apresentados se mostraram melhores do que os das técnicas com as quais foram comparados (*Active Shape Model* - ASM, (LAMECKER; LANGE; SEEBASS, 2004) e (SOLER et al., 2001)). Não foi feita uma análise específica do ganho obtido pelo método de inicialização proposto nesse trabalho.

Qazi et al. (2011) desenvolveram um método de segmentação automática de estruturas anatômicas em imagens de CT (cabeça e pescoço) utilizando uma abordagem híbrida de atlas e modelos deformáveis. Nesse trabalho, os modelos deformáveis são representados por malhas triangulares 3D e têm suas posições inicializadas conforme o resultado de transformações que mapeiam um espaço de referência ao espaço da imagem do paciente. Mais especificamente, o método funciona da seguinte maneira: a partir de um conjunto de imagens de treinamento, dentre as quais uma é escolhida como referência, 14 marcadores anatômicos são manualmente detectados por um especialista. Na imagem de referência, o especialista faz a segmentação manual das estruturas anatômicas de interesse e são geradas malhas 3D para as referidas estruturas a partir dessas segmentações. A posição média e covariância entre os marcadores nas diferentes imagens é calculada, e os marcadores médios são então usados para alinhar, via transformações deformáveis, as imagens de treinamento. Com isso se tem um atlas de intensidades médias por voxel e de posições médias dos marcadores. Quando uma nova imagem clínica é analisada, esses marcadores médios são mapeados para o espaço da imagem clínica via correção afim do atlas, e depois outra transformação deformável é usada para mapear os marcadores da imagem de referência aos marcadores estimados da imagem clínica. Essa mesma transformação também mapeia os vértices da malha ao espaço da imagem clínica, assim inicializando o modelo deformável. Posteriormente à convergência do modelo, ainda é aplicado um refinamento da segmentação final, usando máscaras probabilísticas das estruturas (para representar graus de incerteza) e classificação de voxels pelo algoritmo *k-Nearest Neighbors* (k-NN), com $k = 100$, utilizando características de intensidade, localização, textura e formato locais.

A abordagem de Qazi et al. (2011) foi testada na segmentação de diversas estruturas anatômicas em 10 imagens de CT da cabeça e pescoço. 15 outras imagens foram usadas para a construção do atlas e o treinamento do algoritmo de refinamento. Todas as imagens foram obtidas do *Princess Margaret Hospital* (Toronto, Canadá) e possuíam resolução de $1 \times 1 \times 2 \text{mm}^3$.

Para validação, os resultados obtidos foram comparados às segmentações manuais realizadas pelo especialista nas 10 imagens de teste, utilizando para isso duas métricas de similaridade: coeficiente de Dice (DC) e distância de Hausdorff (HD). Obteve-se valores médios de fração de sobreposição (DC) de 0,93 para a mandíbula, 0,91 para o tronco cerebral, 0,83 para as glândulas parótidas e submandibular, e 0,74 para os linfonodos. Esses valores são considerados comparáveis (alguns melhores, outros piores) ao estado da arte na época da publicação do artigo.

Ruppertshofen et al. (2013) desenvolveram a Transformada de Hough Generalizada Discriminativa (*Discriminative Generalized Hough Transform* - DGHT) como técnica genérica para localização de objetos em imagens 3D. Ela também pode usada para o posicionamento inicial de modelos deformáveis. A DGHT é uma extensão da Transformada de Hough Generalizada (GHT) (BALLARD, 1981), que utiliza um sistema de “votações” para encontrar em uma imagem as ocorrências de um objeto de formato arbitrário representado por um modelo (nuvem de pontos, no caso). Informações das bordas da imagem (extraídas via detector de Canny) e dos pontos do modelo são usadas para mapear a imagem para um espaço de parâmetros (espaço de Hough), no qual cada “célula” de seu domínio representa uma região no domínio da imagem e seus valores maximais indicam as mais prováveis ocorrências do objeto. Juntamente à GHT, a DGHT emprega a técnica de aprendizado *Discriminative Model Combination* (DMC) para encontrar um conjunto de pesos a serem associados a cada célula do espaço de Hough, visando, assim, fazer com que ocorra pico maximal somente na célula correspondente à localização real do objeto. Especificamente, dado um conjunto de imagens de treinamento com as localizações do objeto de interesse marcadas, os pesos ótimos são obtidos com a minimização de uma métrica de erro que envolve as distâncias euclidianas de cada célula c_i em relação à célula alvo \tilde{c}_n e as densidades $\Pr(c_i | I_n)$ (probabilidade de que a célula c_i represente a verdadeira localização do objeto na imagem I_n). A criação do modelo do objeto de interesse é feita via amostragem das bordas extraídas das imagens de treinamento nas regiões ao redor desse objeto. Além disso, a DGHT também utiliza uma abordagem iterativa para o treinamento (testes e refinamento do modelo a cada iteração) e uma estratégia de multirresolução para reduzir o tempo de execução em imagens clínicas.

A DGHT de Ruppertshofen et al. (2013) foi testada em três bases de dados: (a) 626 imagens de radiografias da perna, todas com resolução de $1,1\text{mm}^3$, com o modelo sendo treinado para localização dos joelhos e 50 imagens tendo sido usadas para o treinamento. (b) 22 imagens de CT do tórax, com tamanhos de $256 \times 256 \times 198$ voxels e resolução isotrópica de 1mm^3 , com o modelo sendo treinado para localização das vértebras torácicas T3 a T11, com 10 imagens tendo sido usadas para treinamento. (c) 22 imagens 3D de RM do corpo inteiro, com $1,875\text{mm}$ de resolução interna e 6mm de espessura em cada fatia, sendo o modelo treinado para localização

da cabeça femoral e 10 imagens tendo sido usadas para treinamento. Em todos os casos, duas métricas de avaliação foram utilizadas: a taxa de localização (percentual de resultados com erro abaixo de 10mm) e o erro médio de todas as localizações bem-sucedidas. Para a base (a), os resultados da técnica foram comparados aos obtidos com treinamento do modelo via florestas de Hough (*Hough Forests* - HF) (GALL et al., 2011). Os testes foram feitos em imagens de pacientes normais, com próteses e com patologias ósseas. A técnica proposta produziu, em seus melhores resultados, valores de 97,6% para a taxa de localização (97,8% com HF) e erro médio de localização de 3mm (4,3mm com HF). Imagens com próteses ou patologias produziram valores considerados piores em ambas as métricas. Para a base (b), os modelos foram treinados (i) apenas para a vértebra T7, e (ii) com vértebras T3-T11. As taxa de localização estimadas foram de 58,3% no caso (i) e 95,5% em (ii), e os erros médios de localização foram de 2,6mm em (i) e 4,3mm em (ii). Na base (c), por fim, a técnica foi testada (i) sem e (ii) com a estratégia de multirresolução para o treinamento, produzindo como resultado taxas de localização de 100% em (i) e 91,6% em (ii), e erros médios de 6,7% em (i) e 3,8% em (ii).

Em um trabalho mais recente, Zhang et al. (2014) propuseram uma abordagem para a inicialização de modelos deformáveis baseada na extração e correspondência entre características locais invariantes. O modelo geométrico é representado por um PDM construído a partir da delimitação manual de marcadores em um conjunto de imagens de treinamento, do posterior alinhamento dessas imagens e conjuntos de pontos via transformações afins, e da obtenção de um perfil médio do formato da estrutura representada. As características locais invariantes das imagens são extraídas usando a técnica SIFT, e a seguir, as características com alto grau de repetibilidade entre as imagens de treinamento (de acordo com limiares preestabelecidos) são selecionadas para compor um modelo de características chave (*Key Features Model* - KFM). Este modelo é composto por K descritores de características SIFT, onde em cada conjunto se tem descritores de características localizados em posições correspondentes (suficientemente próximas). O relacionamento geométrico entre o PDM e o KFM é estabelecido por uma modelagem estatística que mapeia as características SIFT alinhadas aos pontos do PDM. A partir dessa relação, a inicialização de um modelo deformável em uma imagem clínica se dá pela extração das características SIFT em tal imagem, casamento entre as características extraídas com o KFM, e mapeamento para o PDM utilizando as transformações estimadas, obtendo assim, na imagem clínica, os marcadores que representam o posicionamento inicial do modelo deformável.

A técnica de Zhang et al. (2014) foi avaliada quantificando-se a precisão e o custo computacional da inicialização do PDM e da convergência do modelo deformável quando utilizado para correção de imagens. Os testes foram feitos em uma base de imagens de RM cardíaca, consistindo em 500 fatias obtidas de 16 pacientes (10 homens e 6 mulheres) com idades entre 45

e 64 anos. 60 dessas fatias, provenientes de 5 pacientes, foram usadas para treinamento. Para avaliar a precisão da inicialização e do resultado final, foi utilizada uma métrica de distância média entre os pontos das formas obtidas automaticamente pela técnica e os pontos da estrutura manualmente marcada. Os resultados foram comparados aos do método de Chen *et al.* (CHEN *et al.*, 2011), observando-se uma melhoria entre 6,2% e 8,3% na precisão de inicialização do modelo e uma redução de 50% no número de iterações necessárias para a convergência.

3.4.2 Discussão

Esta seção apresentou alguns dos diferentes métodos propostos na literatura nos últimos anos para inicialização de modelos deformáveis. As abordagens mais generalistas incluem o uso de técnicas de programação evolutiva (HEIMANN *et al.*, 2007), correção a partir de um atlas de malhas (QAZI *et al.*, 2011), transformada de Hough generalizada discriminativa (RUPPERTSHOFEN *et al.*, 2013) e modelo de características chave usando SIFT (ZHANG *et al.*, 2014). Observa-se que todos esses métodos exigem treinamento e interação humana em algum momento, inclusive o método de Zhang *et al.* (2014), que utiliza uma técnica de detecção de pontos de interesse na imagem e ainda requer um treinamento específico para cada estrutura de interesse. Dependendo do trabalho manual de um especialista é muitas vezes um fator limitante para aplicações práticas de uma técnica.

3.5 Considerações finais

Este capítulo apresentou a revisão do estado da arte dividida em três grandes tópicos: detecção de pontos salientes 3D, construção de atlas do cérebro e inicialização automática de modelos deformáveis, todos com aplicações em imagens médicas 3D. Foram descritas as técnicas, formas de avaliação, bases de dados utilizadas e resultados obtidos por cada trabalho selecionado. Ao final de cada seção, uma breve discussão apontou algumas das limitações existentes nessas técnicas. Com base nessa revisão, a abordagem desenvolvida neste trabalho pode ser apresentada a seguir.

Capítulo 4

MATERIAIS E MÉTODOS

Este capítulo apresenta em detalhes os materiais e métodos do presente trabalho, contextualizando-os em relação ao estado da arte. São descritas as bases de dados utilizadas e os procedimentos realizados para o pré-processamento das imagens, bem como a fundamentação teórica utilizada para desenvolver as técnicas de detecção, descrição e casamento de pontos salientes, construção dos atlas de pontos salientes e inicialização de modelos geométricos deformáveis. Também são descritos os métodos utilizados para análise dos resultados.

4.1 Visão geral

A proposta deste trabalho de pesquisa consistiu na construção de atlas de pontos salientes 3D, automaticamente detectados em imagens de RM do cérebro, para fins de representação da distribuição dos pontos salientes detectados entre os indivíduos de uma determinada população, considerando-se a variabilidade anatômica existente entre eles. A aplicação prática do atlas é demonstrada através do seu uso no contexto da segmentação automática de estruturas anatômicas relevantes ao estudo de doenças neurodegenerativas como a doença de Alzheimer (DA). Especificamente, a ideia é utilizar as informações do atlas como guias para estabelecer uma inicialização, em termos de posicionamento e ajustes de forma, de modelos geométricos deformáveis, buscando assim amenizar os típicos problemas de inicialização e convergência que os métodos de segmentação baseados nesse tipo de técnica comumente apresentam. Até o presente momento, não se tem conhecimento de outro trabalho proposto na literatura científica que utilize tal abordagem para esse fim.

A Figura 4.1 representa esquematicamente a abordagem utilizada para a aplicação de um atlas de pontos salientes na segmentação de estruturas anatômicas. Os elementos primordiais necessários para essa tarefa são: (a) uma imagem de RM de referência; (b) um banco de mode-

los geométricos (aqui representados como malhas 3D), alinhados à imagem de referência, que representem as estruturas anatômicas que se planeja segmentar; (c) um banco de imagens pre-processadas e espacialmente alinhadas conforme a necessidade da aplicação, para a construção do atlas; (d) uma técnica de detecção automática de pontos salientes (correspondentes a marcadores anatômicos ou não) em imagens 3D de RM; (e) uma técnica de descrição e casamento de pontos em imagens 3D de RM; e (f) uma técnica para mapeamento entre pontos casados (estimação de um campo de deformação) e consequente ajuste (deformações locais) dos modelos geométricos representados pelas malhas 3D. A partir da técnica de detecção de pontos salientes e do banco de imagens de RM alinhadas, o atlas de pontos salientes pode ser construído. E assim, para cada nova análise clínica, os seguintes passos são realizados:

1. É dada como entrada uma imagem clínica (imagem de um paciente) e o modelo (malha) da estrutura anatômica a ser segmentada em tal imagem, o qual inicialmente está alinhado ao atlas e à imagem de referência;
2. A imagem clínica é alinhada ao espaço de coordenadas do atlas (referência) via corregristo afim;
3. A técnica de detecção automática de pontos salientes é aplicada à imagem clínica alinhada, obtendo-se, além da localização, um descritor ou assinatura para cada ponto saliente detectado;
4. Com a técnica de casamento de pontos salientes, identifica-se um conjunto de pares de pontos similares (em termos de seus descritores) envolvendo o atlas e os pontos na imagem clínica alinhada;
5. É estimada uma transformação não-rígida (representada por um campo de deformação) que realiza o mapeamento entre os pontos salientes detectados na imagem clínica e aqueles presentes no atlas;
6. A transformação estimada é aplicada ao modelo (aos vértices da malha) da estrutura anatômica, ajustando-se, assim, o posicionamento e a forma deste em relação à imagem clínica, conferindo-lhe uma inicialização próxima da estrutura na imagem alinhada;
7. A inversa da transformação afim obtida pelo corregristo realizado na etapa 2 é aplicada ao modelo inicializado, para alinhá-lo à imagem clínica em seu espaço original.

Considerando que o corregristo das etapas 2 e 7 e o campo de deformação estimado na etapa 4 já representam, em conjunto, tanto uma inicialização em termos de posição como também de

ajustes locais de forma, não foi considerado necessário explorar técnicas adicionais para refinar o ajuste da malha. Entretanto, técnicas que aproveitassem as informações da própria imagem clínica, dos pontos salientes, do atlas e de outros funcionais de energia poderiam eventualmente ser usadas para refinar a adaptação da malha em relação à imagem clínica, de maneira a concluir a segmentação. Mas, neste trabalho, o foco está na inicialização do modelo.

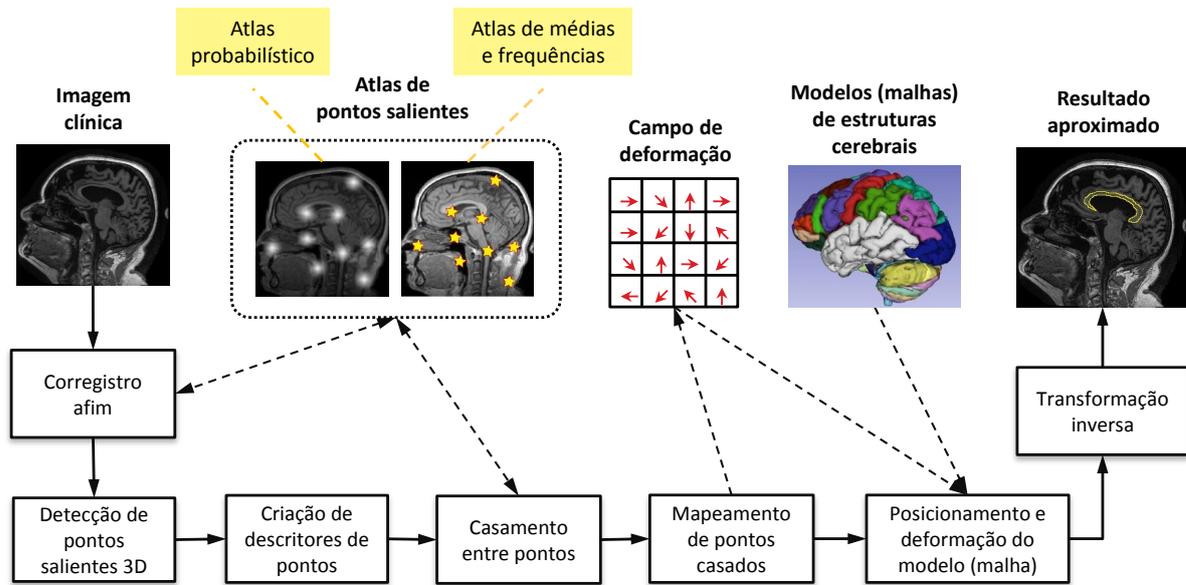


Figura 4.1: Etapas para aplicação do atlas no auxílio à segmentação de estruturas cerebrais por modelos geométricos (malhas) deformáveis.

As próximas seções descrevem com detalhes a metodologia que foi utilizada neste trabalho para atingir os objetivos esperados.

4.2 Bases de dados

As bases de dados utilizadas neste trabalho incluem imagens 3D de RM, malhas triangulares 3D de estruturas cerebrais e também imagens sintéticas geradas por modelos paramétricos. Todas elas são descritas a seguir.

4.2.1 NAC Brain Atlas

Esta base de dados foi desenvolvida pelo *Neuroimage Analysis Center* (NAC) afiliado ao *Surgical Planning Laboratory* e à Universidade de Harvard. Consiste em um conjunto de 149 malhas triangulares 3D que representam, cada uma, uma estrutura cerebral distinta. Tais malhas estão espacialmente alinhadas a duas imagens de RM de um mesmo paciente, obtidas nas modalidades T1-w e T2-w, ambas com resolução isotrópica de 1mm e de tamanho $256 \times 256 \times 256$

voxels. Juntamente às malhas, a base fornece essas imagens de referência e a rotulação das estruturas cerebrais feita por um especialista. Mais detalhes podem ser encontrados no endereço eletrônico do NAC¹ e no trabalho de Halle et al. (2013).

As malhas triangulares da base NAC foram escolhidas para representar os modelos geométricos deformáveis neste trabalho. E, por já estar espacialmente alinhada com tais malhas, a imagem de referência em T1-w dessa base foi escolhida como objeto de referência para todas as tarefas de pré-processamento e avaliação que envolvem padronização do sistema de coordenadas e alinhamento.

4.2.2 IXI

Desenvolvida pelo *Neuroimage Analysis Center* da *Imperial College London*, a base *Information eXtraction from Images* (IXI), que faz parte do projeto *Brain Development*, é constituída por imagens clínicas de RM do cérebro adquiridas nas modalidades T1-w, T2-w, PD-w, MRA e DTI. As imagens foram obtidas de pacientes saudáveis em três diferentes hospitais londrinos, a saber: *Hammersmith Hospital* (HH), *Guy's Hospital* (GH) e *Institute of Psychiatry* (IoP). Cada hospital utilizou seu próprio scanner de RM e protocolo de aquisição, cujas especificações podem ser encontradas em detalhes no endereço eletrônico do projeto². Para este trabalho, foram utilizadas apenas imagens ponderadas em T1-w, que possuem as características apresentadas na Tabela 4.1³. Embora a base tenha um total de 581 imagens, 15 delas foram descartadas por não haver informação sobre as idades dos seus respectivos pacientes. A faixa etária dos pacientes considerados varia entre 20 e 86 anos e está distribuída conforme expressa o gráfico da Figura 4.2.

Tabela 4.1: Características das imagens de RM da base de dados “IXI”.

Hospital	Scanner	Dimensões	Resolução (mm)	Orientação	Imagens
HH	Philips 3T	256×256×150	0,94×0,94×1,2	AIL	185
GH	Philips 1,5T	256×256×150	0,94×0,94×1,2	AIL	322
IoP	GE 1,5T	256×256×146	0,94×0,94×1,2	AIL	74

¹<http://nac.spl.harvard.edu>

²<http://www.brain-development.org>

³O código de três letras que representa a orientação já foi explicado na Seção 2.1.3.3. Entretanto, neste caso ele descreve uma orientação aproximada: na prática, todas as imagens dessa base têm orientação oblíqua.

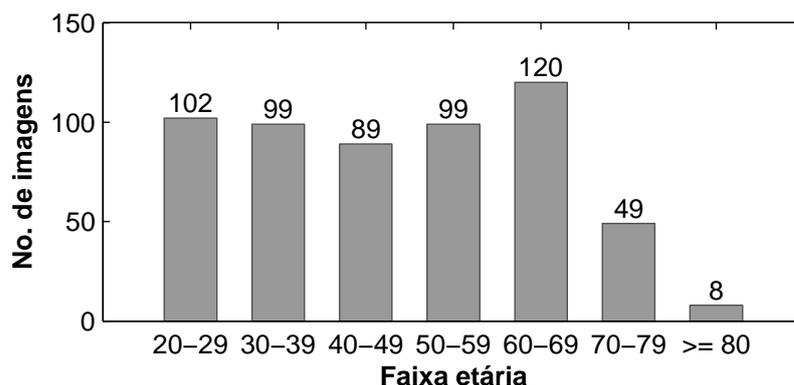


Figura 4.2: Distribuição etária dos pacientes nas imagens em T1-w da base de dados IXI.

4.2.3 ADNI e EADC-ADNI

A base de imagens do *Alzheimer's Disease Neuroimaging Initiative* (ADNI) faz parte de um projeto iniciado em 2003 por um grupo de instituições de pesquisa, empresas privadas do ramo farmacêutico e organizações não governamentais visando o estudo da progressão da DA. A iniciativa tem como objetivo coletar e disponibilizar à comunidade científica uma ampla base de dados de biomarcadores dessa doença para fins de posterior análise e exploração. No endereço eletrônico do projeto⁴ são disponibilizadas gratuitamente mais de duas mil imagens de RM pré-classificadas conforme a ocorrência dos sintomas da doença nos pacientes examinados. A base inclui imagens em T1-w já pré e pós-processadas, além de rotulações dos hipocampos para algumas imagens. O pré-processamento inclui correções de escala (devido a diferenças no tamanho do voxel entre uma aquisição e outra) (CLARKSON et al., 2009) e de heterogeneidade dos níveis de intensidade (devido ao *bias field*). As imagens foram adquiridas utilizando uma grande variedade de scanners de 1,5T e 3T das três principais fabricantes (Philips, General Electric e Siemens), portanto as especificações das imagens (tamanho da matriz, resolução, etc.) são diversas e serão aqui omitidas. Tais informações podem ser encontradas no endereço eletrônico previamente mencionado. E dado que o foco do projeto ADNI está na doença de Alzheimer, os pacientes imageados são majoritariamente idosos, conforme pode-se observar no gráfico de distribuição da faixa etária presente na Figura 4.3.

A base de imagens EADC-ADNI, por sua vez, faz parte de um projeto iniciado em 2012 pelo *European Alzheimer's Disease Consortium* (EADC) que procura estabelecer um protocolo harmonizado para segmentação manual do hipocampo em imagens clínicas de RM, a fim de possibilitar uma maior concordância entre as delimitações feitas por diferentes especialistas. A base disponibiliza 135 imagens de RM em T1-w extraídas diretamente da base ADNI (ou seja, imagens de pacientes idosos e com diferentes graus de evolução da DA) e também as respec-

⁴<http://adni.loni.usc.edu/>

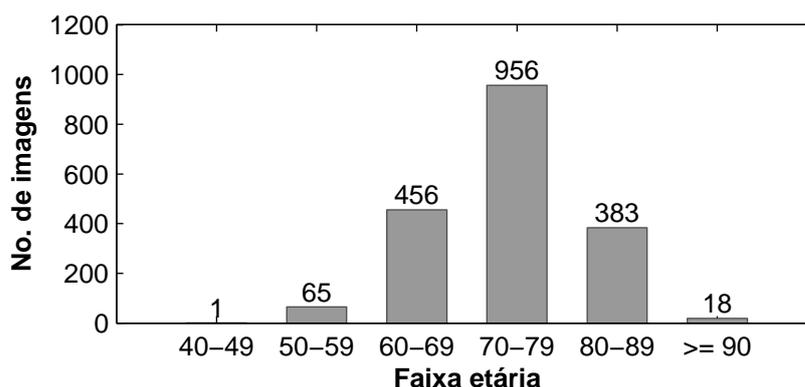


Figura 4.3: Distribuição etária dos pacientes nas imagens da base de dados ADNI.

tivas rotulações feitas para os hipocampos em cada uma dessas imagens seguindo o protocolo proposto. Informações a respeito desse protocolo e das rotulações obtidas por meio dele podem ser encontradas na página web do projeto⁵ e nas publicações de Boccardi et al. (2015) e Frisoni et al. (2015).

4.2.4 BrainWeb

O projeto *BrainWeb* (AUBERT-BROCHE et al., 1996) é uma iniciativa do *McConnel Brain Imaging Centre* da Universidade de McGill, Canadá, para auxiliar a validação de experimentos em pesquisas envolvendo imagens de RM do cérebro. Consiste em uma base de imagens sintéticas de RM de livre acesso, as quais são construídas por um simulador a partir de dois modelos anatômicos principais, um para cérebros normais e outro para cérebros com EM. Na página web do projeto⁶, é possível ajustar uma série de parâmetros que definirão as características das imagens que se deseja adquirir, incluindo: modalidade de ponderação (T1-w, T2-w ou PD-w), espessura da fatia (1mm, 3mm ou 5mm), quantidade de *bias field* (atenuação de até 0%, 20% ou 40% do valor máximo de intensidade da imagem) e quantidade de ruído Gaussiano (desvio padrão de 0%, 1%, 3%, 5%, 7% ou 9% do valor máximo de intensidade da imagem).

4.2.5 Imagens 3D sintéticas

Esta base de dados contém 1890 imagens 3D sintéticas que representam pequenas estruturas 3D de alta curvatura sob diferentes condições (simuladas) de degradação por *bias field* e ruído Gaussiano. Foi especialmente construída nesta pesquisa para fins de teste da técnica de detecção de pontos salientes (sendo, portanto, uma das contribuições deste trabalho). Em ima-

⁵<http://www.hippocampal-protocol.net>

⁶<http://brainweb.bic.mni.mcgill.ca/brainweb/>

gens de RM do cérebro, estruturas similares a essas podem ser encontradas, por exemplo, na cabeça do corpo caloso e próximo a certas estruturas ventriculares e corticais. Elas são sinteticamente geradas nas imagens utilizando-se o modelo paramétrico de estruturas do tipo “ponta” descrito no trabalho de Wörz e Rohr (2005). Esse modelo é baseado em uma função de erro Gaussiano que constrói uma forma 3D análoga à metade de um elipsoide e, adicionalmente, admite a aplicação de transformações de rotação e deformações locais para modelar variações dessa estrutura. Os parâmetros utilizados para simulação de *bias field* e ruído Gaussiano são exatamente os mesmos que a base *BrainWeb* utiliza. As imagens geradas possuem resolução isotrópica de 1mm e tamanho $33 \times 33 \times 33$ voxels.

4.3 Pré-processamento

O pré-processamento tem o objetivo de organizar e preparar os dados de entrada para que as posteriores etapas de processamento possam ter os melhores resultados possíveis. No contexto deste trabalho, ele consiste no uso de técnicas de processamento de imagens 3D para redução de ruído, correção de *bias field*, padronização de histogramas, alinhamento espacial com a imagem de referência e extração de máscaras da região do encéfalo em um conjunto de imagens de “treinamento”. Essas imagens pré-processadas são então usadas para construir o atlas de pontos salientes.

Todos os passos do pré-processamento são detalhados nas subseções a seguir. O conjunto de imagens de treinamento utilizado foi o banco de imagens em T1-w da base IXI (descrita na Seção 4.2.2), pelo fato de tais imagens apresentarem boa discriminação visual dos tecidos cerebrais e abrangerem pacientes em uma ampla variabilidade etária.

4.3.1 Redução de ruído

Neste trabalho, a redução do ruído nas imagens de treinamento é feita usando o algoritmo *Non Local Means* (NLM), proposto por Buades, Coll e Morel (2005). Segundo os autores, é um algoritmo que possui a desejável característica de eliminar o ruído minimizando a perda de informação estrutural da imagem original, ao contrário do que acontece com outras técnicas conhecidas na literatura, tais como o filtro Gaussiano (que causa perdas severas nas bordas) e o filtro de difusão anisotrópica (PERONA; MALIK, 1990) (que pode ter perdas maiores em bordas e texturas curvas).

Como já foi mencionado na Subseção 2.1.3.5, imagens de RM possuem ruído descrito pela

distribuição de Rice. Em contrapartida, o algoritmo NLM foi originalmente proposto para lidar com ruído aditivo Gaussiano. No entanto, isso não é um problema neste trabalho, porque as imagens de treinamento utilizadas não se encontram muito degradadas por ruído (isto é, uma aproximação por uma distribuição Gaussiana não seria, em tese, tão distante assim da realidade, levando em conta que o ruído depende do sinal) e a redução de ruído não se trata de uma etapa crítica para o desempenho das técnicas posteriores (além de ser empregada apenas nas imagens de treinamento para construção dos atlas). Assim, o uso de um algoritmo bem estabelecido como o NLM já é suficiente para o propósito que aqui se buscou.

A ideia básica do algoritmo NLM, no caso da sua aplicação em imagens 3D, é determinar a intensidade de cada elemento (voxel) $\mathbf{x} = (x, y, z)$ de uma imagem I por uma soma ponderada que leva em consideração a similaridade de \mathbf{x} com os demais voxels \mathbf{y} de toda a imagem. Essa similaridade é representada por uma função de ponderação $\omega(\mathbf{x}, \mathbf{y})$ que deve ser tal que $0 \leq \omega(\mathbf{x}, \mathbf{y}) \leq 1$ e $\sum_{\mathbf{y}} \omega(\mathbf{x}, \mathbf{y}) = 1$. No trabalho original, os autores estabelecem a similaridade entre dois elementos \mathbf{x} e \mathbf{y} de acordo com o quão semelhantes são as suas respectivas regiões de vizinhança ($\mathcal{N}_{\mathbf{x}}$ e $\mathcal{N}_{\mathbf{y}}$, respectivamente). Assim, $\omega(\mathbf{x}, \mathbf{y})$ é definida de maneira a variar inversamente em relação à distância Euclidiana entre os vetores de intensidades $I(\mathcal{N}_{\mathbf{x}})$ e $I(\mathcal{N}_{\mathbf{y}})$ ponderada pelo desvio padrão $\sigma > 0$ de um filtro Gaussiano.

Assim, o valor de cada elemento \mathbf{x}_i da imagem fica definido pelo algoritmo NLM como

$$NLM[I](\mathbf{x}) = \sum_{\mathbf{y} \in I} \omega(\mathbf{x}, \mathbf{y}) I(\mathbf{y}), \quad (4.1)$$

em que

$$\omega(\mathbf{x}, \mathbf{y}) = \frac{1}{Z(\mathbf{x})} \exp \left\{ - \frac{\|I(\mathcal{N}_{\mathbf{x}}) - I(\mathcal{N}_{\mathbf{y}})\|_{2, \sigma}^2}{h^2} \right\}, \quad (4.2)$$

sendo $Z(\mathbf{x})$ uma constante de normalização e h um parâmetro de filtragem que controla o decaimento das ponderações em função da distância Euclidiana.

A implementação do NLM utilizada neste trabalho está disponível na ferramenta de processamento de imagens médicas EZMINC⁷.

4.3.2 Correção de *bias field*

Para a correção do *bias field* nas imagens de treinamento, foi usado o algoritmo N4-ITK (TUSTISON et al., 2010), que é diretamente derivado do algoritmo *Nonparametric Nonuniform intensity Normalization* (N3) proposto por Sled, Zijdenbos e Evans (1998). Ambos os algorit-

⁷<https://github.com/BIC-MNI/EZminc>

mos assumem um modelo matemático de formação da imagem de RM no qual o *bias field* é representado por um sinal multiplicativo $BF(\mathbf{x})$ que suaviza o sinal $I_0(\mathbf{x})$ originalmente emitido pelo tecido do paciente numa localidade \mathbf{x} . Na ausência de ruído aditivo, esse modelo pode ser expresso pela equação

$$I(\mathbf{x}) = I_0(\mathbf{x})BF(\mathbf{x}). \quad (4.3)$$

Equivalentemente, adotando-se a notação $u(\mathbf{x}) = \log[I_0(\mathbf{x})]$, $f(\mathbf{x}) = \log[BF(\mathbf{x})]$ e $v(\mathbf{x}) = \log[I(\mathbf{x})]$, podemos escrever

$$v(\mathbf{x}) = u(\mathbf{x}) + f(\mathbf{x}). \quad (4.4)$$

O algoritmo N3 assume que u , f e v são variáveis aleatórias independentes e com distribuições de probabilidade U , F e V , respectivamente. Como V é a distribuição da soma das variáveis independentes u e f , temos que

$$V(v) = \int_{-\infty}^{\infty} F(v-f)U(f)df = F(v) * U(v), \quad (4.5)$$

em que $*$ denota convolução (DEVROYE, 1986). Assim, F pode ser entendida como um operador de borrimento que reduz os componentes de alta frequência de U , suavizando-a. Esse resultado tem relação direta com o Teorema da Convolução em processamento de imagens (GONZALEZ; WOODS, 2006). Assim, o algoritmo N3 busca resolver um problema de otimização que consiste em encontrar uma solução que maximize as frequências de U . Isso é realizado por meio do seguinte procedimento iterativo:

$$\begin{aligned} u^n &= v - f_e^n \\ &= v - BS\{v - E[u | u^{n-1}]\}, \end{aligned} \quad (4.6)$$

em que $u^0 = v$ e $BS\{\cdot\}$ é um aproximador *B-spline* que atua como termo de regularização para suavizar o componente $v - E[u | u^{n-1}]$. Essa suavização é necessária porque assume-se como hipótese que o *bias field* é uma função lentamente variante no espaço. Na prática, apenas v é conhecida a princípio, então inicia-se a densidade F como uma distribuição Gaussiana e as estimativas de $E[u^n | u^{n-1}]$ são feitas por meio de deconvolução no domínio da frequência. As iterações se repetem até que f_e e u converjam. Detalhes adicionais de implementação desta técnica incluem a estimação da densidade de v pela técnica não-paramétrica *Parzen Windows* (DUDA; HART; STORK, 2000).

O algoritmo N4-ITK é basicamente uma versão aprimorada do algoritmo N3. Ele utiliza um aproximador *B-spline* diferente para a regularização, sendo mais estável e mais robusto a ruído, e também modifica o processo de iteração empregando uma abordagem multirresolução

(aumentando gradativamente a quantidade de pontos de controle no *B-spline*) e estimando, a cada iteração de cada nível de resolução, o campo de heterogeneidade residual f_r em vez do campo total f_e . Dessa forma, as iterações do algoritmo N4 são descritas pela equação:

$$\begin{aligned} u^n &= u^{n-1} - f_r^n \\ &= u^{n-1} - BS^* \{u^{n-1} - E[u | u^{n-1}]\}, \end{aligned} \quad (4.7)$$

sendo que $f_e = \sum_{i=1}^n f_r^i$, $f_e^0 = \mathbf{0}$ e $BS^* \{.\}$ é o novo operador *B-spline* de suavização.

A implementação do algoritmo N4-ITK utilizada neste trabalho foi obtida da biblioteca *Insight ToolKit*⁸ (ITK).

4.3.3 Padronização de histogramas

A base de dados utilizada para a construção do atlas contém imagens provenientes de diferentes scanners, portanto existem diferenças tanto entre os histogramas das imagens de diferentes hospitais como também entre as imagens da base e a imagem de referência espacialmente alinhada aos modelos geométricos de estruturas cerebrais. Assim, foi feita a padronização da faixa de intensidades de cada uma dessas imagens em relação à imagem de referência.

Neste trabalho, a padronização de histogramas foi feita utilizando uma variação do algoritmo proposto por Nyul, Udupa e Zhang (2000). O algoritmo realiza o seguinte procedimento: dadas as imagens I_s e I , cujas faixas intensidades são $[s_1, s_2]$ e $[r_1, r_2]$ (respectivamente), estima-se um conjunto de l pontos de referência $\{\bar{q}_k | 1 \leq k \leq l\}$ no histograma de I_s , e também um conjunto de l pontos de referência $\{q_k | 1 \leq k \leq l\}$ no histograma de I . Após isso, é feita uma sucessão de mapeamentos lineares envolvendo os intervalos que separam pontos de referência correspondentes em $\{q_k\}$ e $\{\bar{q}_k\}$ - isto é, mapeamentos de $[q_k, q_{k+1}]$ a $[\bar{q}_k, \bar{q}_{k+1}]$, para todo $k = 1 \dots l - 1$, e também de $[r_1, q_1]$ a $[s_1, \bar{q}_1]$ e de $[q_l, r_2]$ a $[\bar{q}_l, s_2]$. Na prática, esses pontos de referência dos histogramas são definidos como os seus respectivos percentis, de acordo com a quantidade escolhida de pontos. Por exemplo, para $l = 3$, eles seriam os quartis dos histogramas.

A implementação desse algoritmo utilizada neste trabalho foi obtida da biblioteca ITK. No caso, escolheu-se $l = 4$ e determinou-se I_s como sendo a imagem de referência em T1-w da base NAC (imagem alinhada às malhas).

⁸<http://www.itk.org/>

4.3.4 Alinhamento espacial (corregistro)

Para a correta construção de um atlas, as imagens de treinamento precisam estar alinhadas em um mesmo sistema de coordenadas (origem, orientação) e ter as mesmas dimensões e resolução. Além disso, também é necessário haver o alinhamento entre o atlas e a imagem de referência alinhada aos modelos geométricos. Assim sendo, todo o banco de imagens precisa ser alinhado à imagem de referência antes da construção dos atlas. Esse alinhamento é feito por técnicas de corregistro de imagens, que iterativamente estimam uma transformação T que mapeia uma imagem “fixa” I_1 para uma imagem “móvel” I_2 , ambas com dimensões idênticas, a fim de que a saída $I = T^{-1}(I_2)$ esteja alinhada a I_1 .

Neste trabalho, o alinhamento espacial foi feito usando uma técnica de corregistro afim, que envolve apenas transformações globais (rotação, translação, escala, cisalhamento). A técnica utilizada para isso está implementada na ferramenta Nifty Reg⁹, sendo baseada no método proposto em (OURSELIN et al., 2000; OURSELIN; STEFANESCU; PENNEC, 2002). O algoritmo basicamente realiza dois grandes passos:

1. Uso de uma estratégia de casamento de blocos (*block matching*) para encontrar localidades correspondentes entre I_1 e a atual I . Isso é feito dividindo-se ambas as imagens em blocos de tamanho uniforme e comparando-se cada bloco de I_1 com a sua correspondente vizinhança em I , selecionando como saída em cada caso o bloco de I para o qual a similaridade com o bloco em I_1 é máxima. Essa similaridade é medida via coeficiente de correlação (GONZALEZ; WOODS, 2006). O resultado é uma lista de pares $(\mathbf{x}_k, \mathbf{y}_k)$ de pontos (centros de blocos) correspondentes entre I_1 e a atual I .
2. Estimação de uma matriz de rotação \mathbf{R} e um vetor de translação \mathbf{t} a partir das correspondências locais encontradas no passo 1. Esse cálculo é feito por meio da seguinte minimização:

$$(\hat{\mathbf{R}}, \hat{\mathbf{t}}) = \arg \min_{\mathbf{R}, \mathbf{t}} \sum_k \rho(\|\mathbf{y}_k - \mathbf{R}\mathbf{x}_k - \mathbf{t}\|), \quad (4.8)$$

onde $\|\cdot\|$ denota a norma Euclidiana e $\rho(x)$ é uma função de valores não-negativos com único ponto de mínimo em zero. Neste caso, é utilizada a função $\rho(x) = |x|$.

Esses dois passos são alternados repetidas vezes com o objetivo de se seguir uma estratégia de refinamento progressivo do resultado final. Na implementação contida no Nifty Reg, os blocos do passo 1 são mantidos em tamanho constante ($4 \times 4 \times 4$ voxels) e as estimações são feitas em um esquema multirresolução, inicialmente em versões de baixa resolução das imagens ($G_\sigma * I_1$

⁹<http://sourceforge.net/projects/niftyreg/>

e $G_\sigma * I$, onde G_σ é um filtro Gaussiano com desvio padrão σ) e, por último, nas imagens originais.

4.3.5 Extração de máscaras do encéfalo

A informação anatômica fornecida por uma imagem de RM do cérebro inclui não somente o encéfalo em si como também a caixa craniana do paciente e outras estruturas secundárias. Entretanto, tal informação extra muitas vezes não é necessária e nem desejada para os métodos computacionais de processamento. Por essa razão, foi utilizada uma implementação da técnica *RObust Brain EXtraction*¹⁰ (ROBEX) (IGLESIAS et al., 2011) para a criação de máscaras binárias que isolam as regiões onde se deseja realizar o processamento (encéfalo). Essas máscaras são resultantes da limiarização da imagem do encéfalo extraído (produzida pela ROBEX) seguida por uma operação morfológica de fechamento para eliminar possíveis “buracos”.

A técnica ROBEX realiza a extração da região do encéfalo na imagem de RM por meio da combinação de dois modelos: um modelo discriminativo, que consiste em um classificador de floresta aleatória (*Random Forest Classifier*) para rotular os voxels da imagem de entrada como pertencentes ou não às bordas do encéfalo; e um modelo gerador, que é baseado em uma distribuição Gaussiana sobre um conjunto de pontos de referência que definem o formato do cérebro e permitem sua representação usando uma malha triangular. A ideia então é, para cada nova imagem a ser classificada, encontrar a instância do modelo gerador que maximiza a verossimilhança da superfície do encéfalo de acordo com o resultado dado pelo modelo discriminativo.

4.4 Detecção de pontos salientes 3D

Embora exista na literatura um grande número de técnicas aplicáveis em imagens 3D para detecção de pontos salientes, e boa parte delas afirme ser invariante a mudanças globais nos níveis de intensidade da imagem, pouca ou nenhuma discussão é feita sobre como os resultados dessas técnicas são afetados por heterogeneidades locais nos níveis de intensidade (XIA; ZHAO; LIU, 2013). Em imagens de RM, conforme mencionado na Seção 2.1.3.4, essas heterogeneidades acontecem devido ao inevitável *bias field*. Na Figura 4.4a, a representação visual de um *bias field* é mostrada, e nas Figuras 4.4b e 4.4c são exibidas as magnitudes dos gradientes de imagem computados em uma imagem de RM na ausência e na presença desse *bias field*, respectivamente. A comparação entre os dois casos mostra que o *bias field* pode afetar os valores de magnitude de gradiente de maneira não desprezável, portanto uma técnica que se baseie em tal

¹⁰<http://www.nitrc.org/projects/robex/>

tipo de informação (como a n D-SIFT¹¹ e diversas outras) poderia ter seus resultados distorcidos por isso.

4.4.1 Congruência de fase 3D

A técnica de detecção automática de pontos salientes 3D utilizada neste trabalho é baseada no modelo de congruência de fase (*phase congruency* - PC) (MORRONE et al., 1986). Em vez de assumir que as características salientes de uma imagem (bordas, cantos, etc.) estão localizadas nos pontos de mudanças maximais de intensidade (gradiente), esse modelo assume que tais características são percebidas nos pontos da imagem em que os componentes de Fourier encontram-se maximalmente em fase. Isso está de acordo com uma série de fenômenos psico-visuais que ocorrem na visão humana (WANG; SIMONCELLI, 2004) e provê uma medida que é robusta a variações de intensidade da imagem, uma vez que é baseada principalmente na informação de fase do sinal. Além disso, pode ser mostrado que a medida de PC é proporcional à energia do sinal, podendo ser calculada via convolução da imagem original com um banco de filtros em quadratura (VENKATESH; OWENS, 1990).

O modelo de PC pode ser estendido para uso em imagens 3D fazendo-se a combinação da PC 1D em um conjunto de orientações $\Theta = (\phi, \theta)$, em que ϕ é um ângulo de azimute e θ é um ângulo de elevação. Com base na formulação matemática proposta para imagens 2D por Kovési (2000), a medida de PC 3D é definida como:

$$PC(\mathbf{x}) = \frac{\sum_{\Theta} \sum_n W_{\Theta}(\mathbf{x}) \max(0, A_{\Theta n}(\mathbf{x}) \Delta \Phi_{\Theta n}(\mathbf{x}) - T)}{\sum_{\Theta} \sum_n A_{\Theta n}(\mathbf{x}) + \varepsilon}, \quad (4.9)$$

em que $\mathbf{x} = (i, j, k)$ denota uma localização espacial no domínio da imagem. O termo $A_{\Theta n}(\mathbf{x})$ indica a amplitude da energia da imagem, na localização \mathbf{x} , após a convolução com um par de filtros em quadratura com parâmetro de escala n ($1 \leq n \leq N_s$) e orientação Θ . A constante T é um limiar automaticamente estimado que representa a parcela da energia que se espera corresponder a ruído, e ε é uma constante de valor pequeno para evitar divisão por zero. O termo $W_{\Theta}(\mathbf{x})$, definido como:

$$W_{\Theta}(\mathbf{x}) = \frac{1}{1 + \exp[\gamma(c - I_{\Theta}(\mathbf{x}))]}, \quad (4.10)$$

é uma função sigmoide de ponderação que penaliza as distribuições de frequências que são particularmente estreitas (pois um ponto de congruência de fase só é significativo se ocorrer em uma ampla faixa de frequências). Os parâmetros γ e c nesta função são constantes representando

¹¹Vide Subseção 3.2.1.

um fator de ganho e um valor de corte, respectivamente, e

$$l_{\Theta}(\mathbf{x}) = \frac{1}{N_s} \left(\frac{\sum_n A_{\Theta n}(\mathbf{x})}{A_{\Theta \max}(\mathbf{x}) + \varepsilon} \right), \quad (4.11)$$

é uma medida do espalhamento da resposta do filtro calculada como a soma das amplitudes das respostas de energia ($A_{\Theta n}(\mathbf{x})$) dividida pela maior resposta ao longo das escalas ($A_{\Theta \max}(\mathbf{x})$). Isso é realizado para a obtenção de uma noção da largura da distribuição da resposta do filtro. Por fim, o termo $\Delta\Phi_{\Theta n}(\mathbf{x})$ é uma função de desvio de fase usada para melhorar a localização da resposta dos filtros, sendo definida como:

$$\begin{aligned} \Delta\Phi_{\Theta n}(\mathbf{x}) &= \cos(\phi_{\Theta n}(\mathbf{x}) - \bar{\phi}_{\Theta}(\mathbf{x})) - \\ &\quad |\sin(\phi_{\Theta n}(\mathbf{x}) - \bar{\phi}_{\Theta}(\mathbf{x}))| \\ &= e_{\Theta n}(\mathbf{x}) \left(\frac{\sum_n e_{\Theta n}(\mathbf{x})}{E_{\Theta}(\mathbf{x})} \right) + o_{\Theta n}(\mathbf{x}) \left(\frac{\sum_n o_{\Theta n}(\mathbf{x})}{E_{\Theta}(\mathbf{x})} \right) - \\ &\quad \left| e_{\Theta n}(\mathbf{x}) \left(\frac{\sum_n e_{\Theta n}(\mathbf{x})}{E_{\Theta}(\mathbf{x})} \right) - o_{\Theta n}(\mathbf{x}) \left(\frac{\sum_n o_{\Theta n}(\mathbf{x})}{E_{\Theta}(\mathbf{x})} \right) \right| \end{aligned} \quad (4.12)$$

em que

$$E_{\Theta}(\mathbf{x}) = \sqrt{\left(\sum_n e_{\Theta n}(\mathbf{x}) \right)^2 + \left(\sum_n o_{\Theta n}(\mathbf{x}) \right)^2},$$

corresponde à energia local da imagem ao longo de todas as escalas e $e_{\Theta n} = I * G_{\Theta n}^e$ e $o_{\Theta n} = I * G_{\Theta n}^o$ são os resultados da convolução da imagem de entrada I com o par de filtros em quadratura ($G_{\Theta n}^e, G_{\Theta n}^o$) para uma escala n e orientação Θ .

A medida de PC possui uma série de propriedades que a tornam interessante para a detecção de características salientes em imagens. Primeiramente, trata-se de uma medida normalizada (valores entre 0 e 1), invariante a mudanças nos níveis de intensidade e que permite prover uma noção mais apurada da significância de importantes características de uma imagem. Segundo, ela captura informações da imagem em várias escalas e orientações, sendo portanto invariante também em tais aspectos. Terceiro, o fato de ser computada a partir de um banco de filtros em quadratura a torna uma técnica bastante flexível, já que os parâmetros desses filtros podem ser controlados de maneira a permitir a captura de informações direcionais e em bandas de frequências específicas, se necessário. E, embora uma pesquisa mais aprofundada seja necessária, a medida de PC parece ser melhor adequada para imagens de RM do que técnicas baseadas em gradiente: nas Figuras 4.4d e 4.4e, são mostradas as respostas de PC na ausência e na presença (respectivamente) do mesmo *bias field* mostrado na Figura 4.4a. Embora seja notável que, nesse exemplo, a medida de PC não tenha detectado bordas com o mesmo grau de detalhamento do operador de gradiente, é visível que as respostas de PC praticamente não são afetadas pelas heterogeneidades locais de intensidade causadas pelo *bias field*.

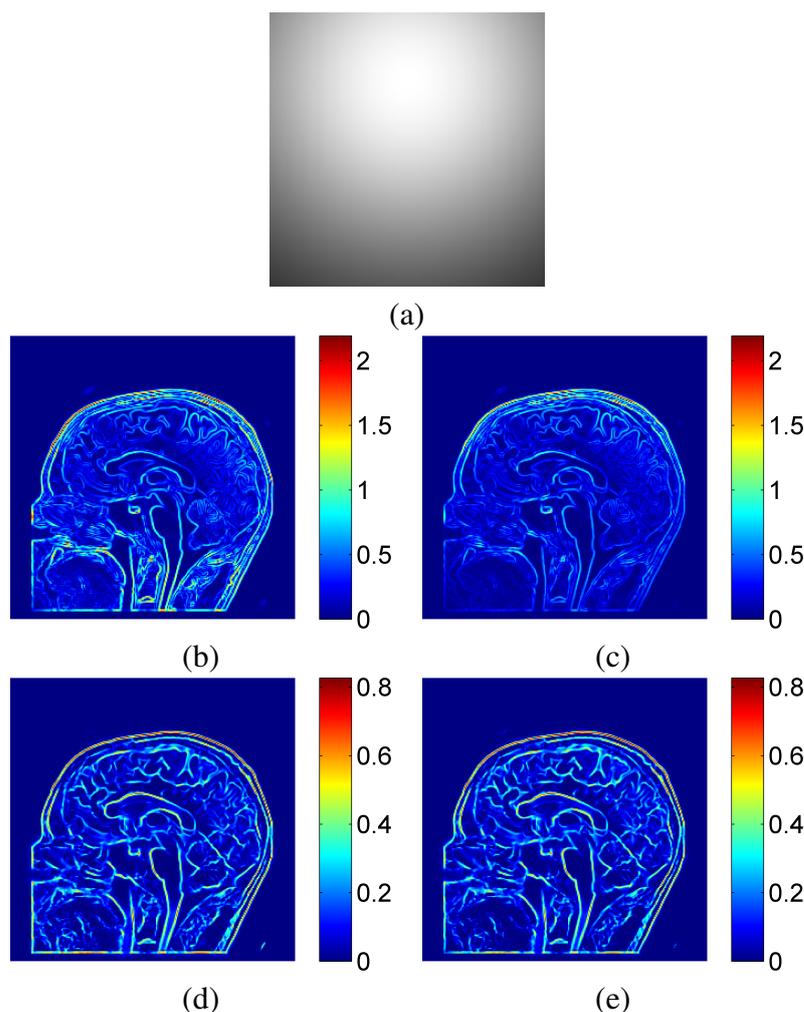


Figura 4.4: Influência das heterogeneidades locais de intensidade, causadas pelo *bias field*, na detecção de bordas em imagens. (a) Exemplo de *bias field*. (b) Magnitudes dos gradientes de imagem na ausência e (c) na presença de *bias field*. (d) Respostas de PC na ausência e (e) na presença de *bias field*.

4.4.2 Banco de filtros de log-Gabor

Os pares de filtros em quadratura usados para computar a PC 3D provêm de um banco de filtros 3D de log-Gabor construído no domínio da frequência, seguindo a formulação matemática de Dosil, Pardo e Fdez-Vidal (2005). Filtros de log-Gabor são uma escolha adequada para esse fim porque não produzem resposta em regiões com intensidade constante e se estendem até as altas frequências, possibilitando que se obtenha larguras de banda arbitrariamente amplas e uma cobertura relativamente uniforme do espectro de frequências.

Cada filtro 3D de log-Gabor é definido como o resultado da multiplicação de um termo

radial, $G_{rad}(\omega)_{\sigma_\omega, \omega_0}$, por um termo angular, $G_{ang}(\phi, \theta)_{\sigma_\alpha, \phi_0, \theta_0}$, sendo

$$G_{rad}(\omega)_{\sigma_\omega, \omega_0} = \exp\left[-\frac{\log^2(\omega/\omega_0)}{2\log^2(\sigma_\omega/\omega_0)}\right], \quad (4.13)$$

$$G_{ang}(\phi, \theta)_{\sigma_\alpha, \phi_0, \theta_0} = \exp\left[-\frac{\alpha^2(\phi, \theta)_{\phi_0, \theta_0}}{2\sigma_\alpha^2}\right], \quad (4.14)$$

e $\alpha(\phi, \theta)_{\phi_0, \theta_0} = \arccos\left(\frac{\mathbf{v} \cdot \mathbf{f}}{\|\mathbf{f}\|}\right)$. O vetor \mathbf{f} é um ponto no domínio da frequência expresso em coordenadas cartesianas, e $\mathbf{v} = (\cos \phi_0 \cos \theta_0, \cos \phi_0 \sin \theta_0, \sin \theta_0)$. Os parâmetros ω_0 , ϕ_0 e θ_0 representam a frequência central e os ângulos centrais de azimute e elevação, respectivamente. σ_ω controla a largura de banda e σ_α controla o espalhamento angular do filtro.

Neste trabalho, os valores dos parâmetros do banco de filtros de log-Gabor foram escolhidos de maneira a se conseguir uma cobertura uniforme do domínio da frequência e um bom equilíbrio entre uma boa precisão das respostas de PC e um tempo de processamento aceitável em aplicações clínicas¹². Os seguintes valores foram atribuídos: $N_s = 4$ níveis de escala, $N_a = 6$ ângulos de azimute e $N_e = 3$ ângulos de elevação. A frequência central para um filtro na escala n foi definida como $\omega_0 = \omega_{max}/(1,6)^{n-1}$, com $\omega_{max} = 1/3$ (o suficiente para capturar altas frequências). Uma largura de banda de aproximadamente 1 oitava foi atribuída para todos os filtros mantendo-se constante a razão $\sigma_\omega/\omega_0 = 0.75$. Finalmente, o parâmetro de espalhamento angular foi definido como $\sigma_\alpha \approx 25^\circ$. Adicionalmente, é feita uma amostragem não-uniforme dos filtros no domínio da frequência, reduzindo-se o número de filtros conforme o parâmetro de elevação aumenta. Isso tem o objetivo de manter a “densidade” de filtros constante e, assim, evitar uma sobreposição excessiva de filtros nas regiões do espectro de frequências que correspondem aos ângulos de elevação mais altos.

Nas Figuras 4.5a e 4.5b, são exibidos dois exemplos de bancos de filtros 3D de log-Gabor com parâmetros bastante similares aos utilizados neste trabalho. O primeiro deles foi construído via amostragem uniforme dos filtros, e enquanto o segundo emprega a amostragem não-uniforme.

4.4.3 Momentos de segunda ordem

Após a computação dos mapas de respostas de PC para cada orientação individual Θ (denotados por PC_Θ), a detecção dos pontos salientes 3D é feita usando uma abordagem similar àquela apresentada nos trabalhos de Ferrari et al. (2011) e Kovési (2003). Com base na teoria

¹²Embora muito ainda possa ser melhorado, em termos de desempenho, com o uso de processamento paralelo via GPU.

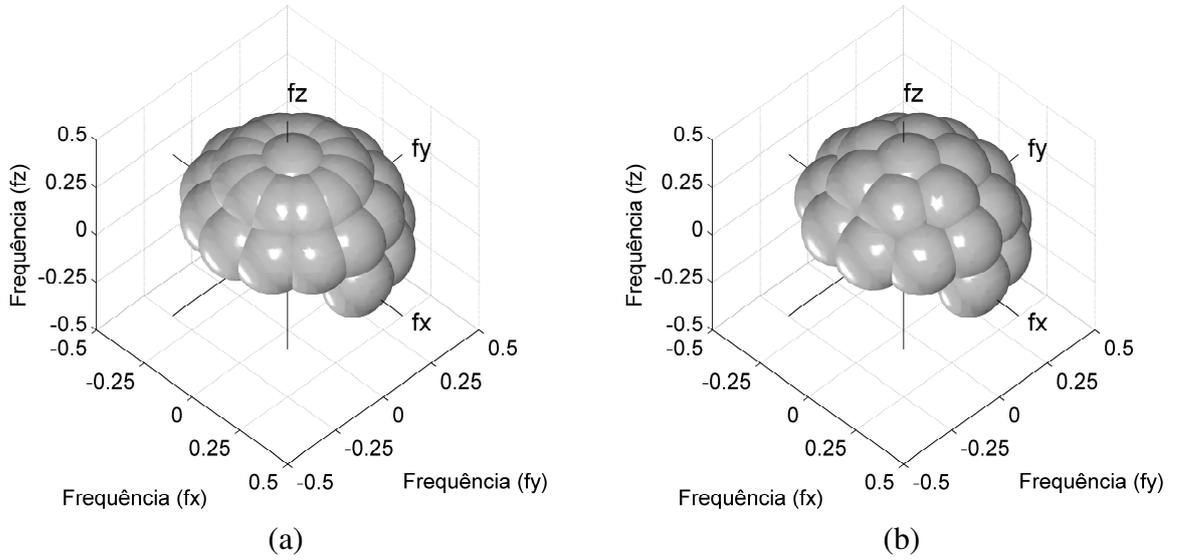


Figura 4.5: Exemplos de banco de filtros 3D de log-Gabor amostrados (a) uniformemente; (b) não-uniformemente no domínio da frequência.

de análise de momentos discretos de imagens (LO; DON, 1989), um mapa de momentos locais de ordem $p + q + r$ é calculado para cada voxel com a seguinte fórmula:

$$M_{p,q,r}(\mathbf{x}) = \sum_{\Theta} (x_{\Theta}^p y_{\Theta}^q z_{\Theta}^r), \quad (4.15)$$

em que

$$\begin{aligned} x_{\Theta} &= PC_{\Theta}(\mathbf{x}) \cos \theta \cos \phi \\ y_{\Theta} &= PC_{\Theta}(\mathbf{x}) \cos \theta \sin \phi \\ z_{\Theta} &= PC_{\Theta}(\mathbf{x}) \sin \theta, \end{aligned} \quad (4.16)$$

e então uma matriz simétrica de momentos de segunda ordem com dimensões 3×3 pode ser definida como

$$M_{PC}(\mathbf{x}) = \begin{bmatrix} M_{2,0,0}(\mathbf{x}) & M_{1,1,0}(\mathbf{x}) & M_{1,0,1}(\mathbf{x}) \\ M_{1,1,0}(\mathbf{x}) & M_{0,2,0}(\mathbf{x}) & M_{0,1,1}(\mathbf{x}) \\ M_{1,0,1}(\mathbf{x}) & M_{0,1,1}(\mathbf{x}) & M_{0,0,2}(\mathbf{x}) \end{bmatrix}. \quad (4.17)$$

Essa matriz confere uma noção da covariância entre os momentos de PC nas variáveis x_{Θ} , y_{Θ} e z_{Θ} . Havendo uma característica saliente de imagem (uma borda, por exemplo) em \mathbf{x} , a direção do eixo principal de dispersão indicará sua orientação dominante, e as dispersões nos eixos perpendiculares indicarão a significância dessa característica em três dimensões. Se os três valores forem altos, isso significa que a característica saliente em \mathbf{x} possui um forte componente 3D associado a ela. Portanto, \mathbf{x} é um candidato a ponto saliente a ser detectado. Normalmente isso ocorre em locais de “cantos” e pontos de alta curvatura na imagem 3D.

Os eixos de dispersão de $M_{PC}(\mathbf{x})$ podem ser aproximados pelos autovalores e autovetores dessa matriz. Os autovalores (representados como $\lambda_1 \geq \lambda_2 \geq \lambda_3$) são sempre não-negativos e indicam os comprimentos dos eixos de dispersão, enquanto os autovetores indicam as orientações associadas a tais eixos. No já referido trabalho de Ferrari *et al.*, os pontos salientes são detectados nas localizações em que se tem valores maximais para uma métrica de saliência baseada no produto normalizado dos três autovalores. Entretanto, uma métrica mais simples e precisa é simplesmente o valor normalizado do menor autovalor, λ_3 . Assim, para cada localização \mathbf{x} em que $M_{PC}(\mathbf{x})$ e seus autovalores são computados, um ponto saliente é detectado se o valor de

$$\text{Saliência}(\mathbf{x}) = \lambda_3 / \lambda_{3max}$$

estiver acima de um determinado limiar, T_λ , e for maximal em \mathbf{x} . Verificou-se experimentalmente que o valor $T_\lambda = 1/3$ produz resultados bastante aceitáveis para as imagens de RM usadas neste trabalho. Mas uma estimação automática do valor de tal limiar, conforme os dados da imagem de entrada, seria o ideal.

4.5 Descrição e casamento de pontos salientes 3D

Para que o casamento entre pontos salientes de diferentes imagens seja possível, é necessário associar a cada ponto saliente um descritor (assinatura) suficientemente discriminativo, bem como definir uma função de distância que meça a dissimilaridade entre pares de descritores. Após computar as distâncias de descritores entre todos os pares de pontos salientes das imagens de entrada, os pontos correspondentes podem ser encontrados e casados com base em uma busca pelo vizinho mais próximo nos sentidos “ida e volta”: \mathbf{h}_1 e \mathbf{h}_2 são casados se \mathbf{h}_2 é o vizinho mais próximo de \mathbf{h}_1 e \mathbf{h}_1 é também o vizinho mais próximo de \mathbf{h}_2 .

Neste trabalho, os descritores necessariamente precisavam ser baseados em informações invariantes a mudanças de intensidade e que não fossem afetadas pelo *bias field*, portanto, não poderiam ser usadas abordagens baseadas em gradiente. Ao todo, três métodos de construção de descritores de pontos salientes foram investigados e implementados durante esta pesquisa, os quais são descritos a seguir. Para todos os casos, considere $\mathbf{s} = (i_s, j_s, k_s)$ como sendo a localização de um ponto saliente detectado na imagem.

4.5.1 Método 1

Foi baseado no trabalho de Shojaeilangaria, Yaub e Teoha (2014) e utiliza uma abordagem relativamente similar ao que é implementado na técnica *nD-SIFT*(CHEUNG; HAMARNEH, 2007)

para a criação de pontos descritores. Mais especificamente, a ideia básica consiste em associar a cada voxel do mapa de PC uma informação direcional e agrupar essas informações, ao longo de uma determinada região de vizinhança \mathcal{N} centrada em \mathbf{s} e de tamanho $\mathcal{N}_x \times \mathcal{N}_y \times \mathcal{N}_z$, na forma de histogramas cujas classes são dadas por um conjunto pré-definido de orientações, O . Neste caso, O corresponde ao conjunto das orientações $\Theta = (\phi, \theta)$ utilizadas pelos filtros de log-Gabor. A construção do descritor de um ponto saliente em uma localização \mathbf{s} é feita, então, a partir da seguinte sequência de passos:

1. Para cada voxel \mathbf{x} da imagem dentro da região \mathcal{N} , obtém-se o valor de congruência de fase para a orientação dominante nesse voxel - isto é, o valor $PC_{\Theta}(\mathbf{x})$ tal que

$$\hat{\Theta} = \arg \max_{\Theta} [PC_{\Theta}(\mathbf{x})]. \quad (4.18)$$

2. Considerando cada orientação $\Theta = (\phi, \theta)$ como uma classe de um histograma, soma-se à contagem correspondente à classe $\hat{\Theta}$ o valor da congruência de fase nessa orientação, $PC_{\hat{\Theta}}(\mathbf{x})$, ponderado por uma janela Gaussiana 3D com $\boldsymbol{\mu} = \mathbf{s}$ e $\boldsymbol{\Sigma}$ sendo uma matriz 3×3 diagonal de elementos não-nulos $\frac{\lfloor \mathcal{N}_x/2 \rfloor}{\alpha}$, $\frac{\lfloor \mathcal{N}_y/2 \rfloor}{\alpha}$ e $\frac{\lfloor \mathcal{N}_z/2 \rfloor}{\alpha}$, dado $\alpha = 2, 1$. Com essa ponderação, voxels mais distantes de \mathbf{s} têm uma menor influência no descritor do que os voxels mais próximos.
3. O histograma de orientações é normalizado para ter soma 1, de maneira que passa a representar uma distribuição de probabilidade discreta $\mathbf{h} : O \rightarrow [0, 1]$. O vetor de atributos representando o descritor de cada ponto saliente em uma localização \mathbf{s} , assim, fica definido por essa distribuição.

Para medir as semelhanças (ou distâncias) entre pares de descritores, duas funções de dissimilaridade foram testadas: a distância Euclidiana e a divergência de Jensen-Shannon (LIN, 1991), que é definida como

$$JS(\mathbf{h}_1, \mathbf{h}_2) = H(\pi_1 \mathbf{h}_1 + \pi_2 \mathbf{h}_2) - \pi_1 H(\mathbf{h}_1) - \pi_2 H(\mathbf{h}_2),$$

em que \mathbf{h}_1 e \mathbf{h}_2 são descritores de pontos salientes distintos, $H(\cdot)$ denota entropia e $\pi_1 = \pi_2 = \frac{1}{2}$.

4.5.2 Método 2

Trata-se de uma abordagem bastante similar à do método anterior. A primeira diferença, neste caso, é que O corresponde ao conjunto de 26 orientações expressas pelos vetores 3D

$(-1, -1, -1), (0, -1, -1), (1, -1, -1), (-1, 0, -1), \dots, (1, 1, 1)$. Os seguintes passos são então realizados para a construção de um ponto saliente em uma localização \mathbf{s} :

1. Para cada voxel \mathbf{x} da imagem dentro da região \mathcal{N} , obtém-se o valor final da congruência de fase ao longo de todas as orientações ($PC(\mathbf{x})$) e a orientação do eixo principal do tensor $M_{PC}(\mathbf{x})$, a qual é representada pelo autovetor \mathbf{v}_1 que está associado ao autovalor λ_1 da matriz.
2. A distância angular entre \mathbf{v}_1 e cada vetor $\mathbf{o}_i \in O$ é calculada, usando a equação

$$d_{angular} = \cos^{-1} \left(\frac{\mathbf{v}_1 \cdot \mathbf{o}_i}{\|\mathbf{v}_1\| \|\mathbf{o}_i\|} \right). \quad (4.19)$$

Considerando cada vetor \mathbf{o}_i como sendo uma classe de um histograma de orientações, o autovetor \mathbf{v}_1 é associado ao \mathbf{o}_i que lhe for angularmente mais próximo, e então soma-se à contagem dessa classe o valor da congruência de fase, $PC(\mathbf{x})$, ponderado pela mesma janela Gaussiana 3D descrita no método anterior.

3. Assim como no método anterior, o histograma de orientações é normalizado para ter soma 1, e o vetor de atributos que representa o descritor de cada ponto saliente em uma localização \mathbf{s} fica definido pela distribuição de probabilidade discreta $\mathbf{h} : O \rightarrow [0, 1]$ assim obtida.

As mesmas funções de dissimilaridade empregadas no método anterior também foram utilizadas neste método.

4.5.3 Método 3

O último método utilizado para a descrição e casamento de pontos salientes foi baseado na técnica *shape contexts* (BELONGIE; MALIK; PUZICHA, 2002; MORTENSEN; DENG; SHAPIRO, 2005), originalmente proposta para quantificar semelhanças de forma e encontrar correspondências locais entre objetos cujos contornos são representados por conjuntos de pontos 2D. Essa técnica parte da ideia de que a distribuição espacial relativa dos pontos de contorno de um objeto constitui um descritor robusto, compacto e altamente discriminativo. A adaptação da *shape contexts* para uso em imagens 3D envolve a utilização de um mapa de bordas da imagem como fonte de informação de forma do objeto, bem como a definição de um histograma log-esférico que divide o espaço de imagem ao redor de cada ponto saliente em termos de η_r distâncias radiais (e logarítmicas) relativas, η_a ângulos de azimute relativos e η_e ângulos de

elevação relativos. Cada classe desse histograma acumula os valores que expressam a significância de borda dos voxels que nela podem ser posicionados. Adicionalmente, uma máxima distância radial, r_{max} , pode ser definida de maneira a restringir o número de voxels processados e reduzir o tempo de processamento (caso não sejam necessários muitos voxels para se construir descritores suficientemente discriminativos). O descritor final é formado pela concatenação das $\eta_r \eta_a \eta_e$ classes do histograma em um vetor \mathbf{h} , o qual é normalizado para ter norma unitária.

Neste trabalho, um mapa dos maiores autovalores (λ_1) da matriz $M_{PC}(\mathbf{x})$ foi usado como mapa de bordas da imagem, uma vez que λ_1 simultaneamente confere uma noção de localização e de significância de borda. O histograma log-esférico para cada ponto saliente é criado com parâmetros $\eta_r = 5$, $\eta_a = 12$ (valores sugeridos por Belongie, Malik e Puzicha (2002)) e $\eta_e = 6$. Logo, cada descritor de ponto saliente tem um total de 360 atributos.

As equações a seguir são utilizadas para calcular o índice b ($0 \leq b \leq 359$) da classe do histograma log-esférico em que algum voxel de borda em $\mathbf{x} = (i, j, k)$, ao redor de um ponto saliente em $\mathbf{s} = (i_s, j_s, k_s)$, se posiciona:

$$\begin{aligned}
 b_r &= \max \left(0, \eta_r + \left\lfloor \log_2 \left(\frac{\|\mathbf{x} - \mathbf{s}\|}{r_{max}} \right) \right\rfloor \right) \\
 b_a &= \text{mod} \left(\left\lfloor \frac{\eta_a (\pi + \phi)}{2\pi} \right\rfloor, \eta_a \right) \\
 b_e &= \begin{cases} \text{mod} \left(\left\lfloor \frac{\eta_e (\frac{\pi}{2} + \theta)}{\pi} \right\rfloor, \eta_e \right) & \text{se } -\frac{\pi}{2} \leq \theta < \frac{\pi}{2} \\ \eta_e - 1 & \text{se } \theta = \frac{\pi}{2} \end{cases} \\
 b &= b_r + \eta_r (b_a + \eta_a b_e),
 \end{aligned} \tag{4.20}$$

sendo que

$$\begin{aligned}
 \phi &= \arctan \left(\frac{j - j_s}{i - i_s} \right), \\
 \theta &= \arctan \left(\frac{k - k_s}{\sqrt{(i - i_s)^2 + (j - j_s)^2}} \right)
 \end{aligned} \tag{4.21}$$

são, respectivamente, os ângulos de azimute e elevação entre \mathbf{x} e \mathbf{s} . As classes do histograma têm suas distâncias radiais limitadas pelos valores $r_{max}/2^{\eta_r-1}$, $r_{max}/2^{\eta_r-2}$, ..., $r_{max}/2$ e englobam, em conjunto, ângulos nos intervalos $\phi \in [-\pi, \pi)$ e $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

Na Figura 4.6, são mostradas algumas representações gráficas para o histograma e os descritores computados no caso 2D ($\eta_e = 1$). O primeiro descritor (Figura 4.6a) foi computado a partir de uma imagem diferente da usada nos outros dois casos. Uma comparação entre as

Figuras 4.6a e 4.6b revela que duas regiões anatomicamente similares na imagem também produzem descritores similares. Em contraste, os descritores obtidos de duas regiões mais distintas (Figuras 4.6a e 4.6c) mostram-se notavelmente diferentes. Com o emprego de uma dimensão adicional, é razoável esperar ainda que a capacidade discriminativa dos descritores possa aumentar ainda mais.

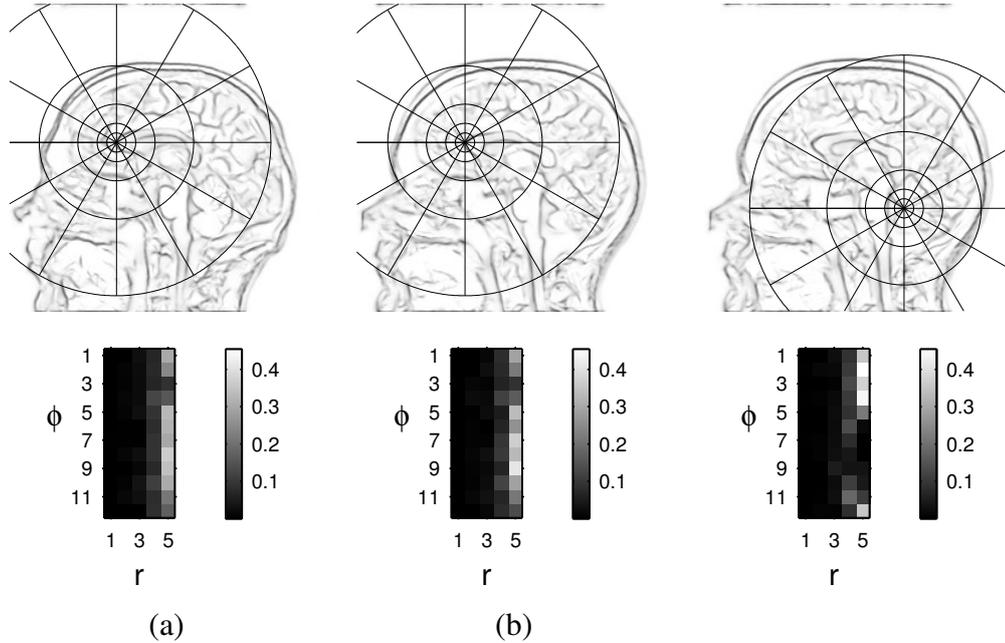


Figura 4.6: Distribuição espacial das classes do histograma de descritor, para $\eta_r = 5$, $\eta_a = 12$ e $\eta_e = 1$ (caso 2D). A distância χ^2 entre os descritores (a) e (b) é de 0,05. Entre os descritores (a) e (c), a distância χ^2 vale 1,04.

Para medir a semelhança entre pares de descritores, a distância χ^2 foi usada como função de dissimilaridade. Tal métrica é particularmente apropriada a este caso porque leva em consideração as diferenças entre os tamanhos das classes do histograma, de maneira que pequenas diferenças entre valores obtidos das classes de maior tamanho (que tipicamente acumulam valores muito maiores) resultam em uma distância menor do que uma pequena diferença entre valores das classes de menor tamanho (que acumulam valores pequenos) (BELONGIE; MALIK; PUZICHA, 2002). Assim, se \mathbf{h}_1 e \mathbf{h}_2 são dois descritores computados pela técnica de *shape contexts* descrita acima, a distância χ^2 entre eles é definida como

$$\chi^2(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{2} \sum_{b=0}^{\eta_r \eta_a \eta_e - 1} \frac{(h_1^{(b)} - h_2^{(b)})^2}{h_1^{(b)} + h_2^{(b)}}, \quad (4.22)$$

onde $h_1^{(b)}$ e $h_2^{(b)}$ são, respectivamente, o b -ésimo componente de \mathbf{h}_1 e de \mathbf{h}_2 .

4.6 Atlas de pontos salientes 3D

As subseções abaixo descrevem a construção dos dois atlas de pontos salientes desenvolvidos e testados neste trabalho: um atlas probabilístico e um atlas de médias e frequências. Ambos são construídos após o pré-processamento e a detecção de pontos salientes em todas as imagens de treinamento, não exigindo interação humana em nenhuma etapa do processo. A Figura 4.7 ilustra de maneira resumida toda a sequência de passos descrita nas seções anteriores, incluindo o pré-processamento e a detecção de pontos salientes, para a construção dos dois tipos de atlas.

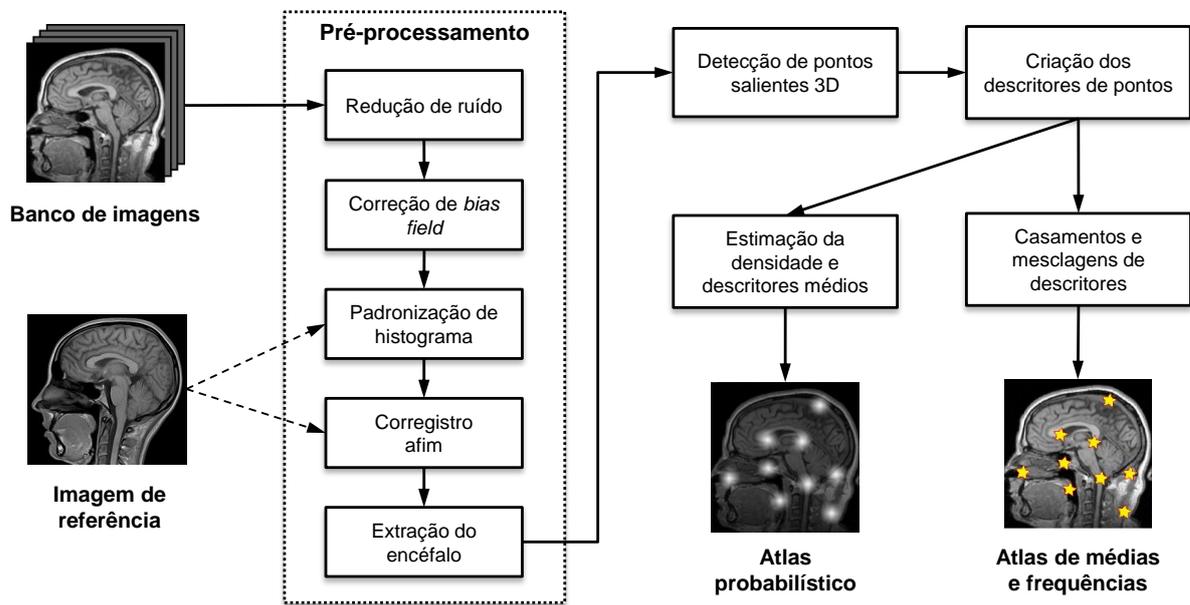


Figura 4.7: Etapas para a construção dos atlas de pontos salientes 3D.

4.6.1 Atlas probabilístico

A construção do atlas probabilístico de pontos salientes é feita em duas etapas: primeiramente, realiza-se a estimação de uma função de densidade de probabilidade $p_x : X \subset \mathbb{N}^3 \rightarrow [0, 1]$, que expressa a distribuição de localização dos pontos detectados nas imagens de treinamento. Depois, é feito o cálculo dos descritores médios \bar{h} ao redor de cada posição maximal de p_x . A densidade de localização dos pontos salientes é estimada com uso da técnica *Manifold Parzen Windows* (MPW) (VINCENT; BENGIO, 2003), que, assim como a técnica *Parzen Windows* (PW) tradicional (DUDA; HART; STORK, 2000), é um método não-paramétrico para estimação de uma função de densidade de probabilidade a partir de um conjunto de amostras, representando tal densidade na forma de uma mistura de Gaussianas $G_{\mu, \Sigma}$. A diferença básica entre essas técnicas está na maneira como essas Gaussianas são parametrizadas e escolhidas. Com a PW, toda

amostra é expressa por um componente Gaussiano fixo e normalmente isotrópico cuja matriz de covariância é expressa por $\Sigma = \sigma^2 \mathbf{I}_3$, sendo \mathbf{I}_3 a matriz identidade de ordem 3. Mas o problema dessa abordagem, segundo (VINCENT; BENGIO, 2003), é que as Gaussianas de formato fixo estão limitadas a fornecer apenas uma expressão da estrutura global dos dados, não capturando características localmente anisotrópicas da distribuição das amostras. Nos locais onde se tem uma maior concentração das amostras d -dimensionais em uma determinada variação topológica de menor dimensionalidade, uma Gaussiana isotrópica acabaria expressando uma probabilidade muito alta para regiões irrelevantes, e muito baixas ao longo da variedade topológica em que as amostras realmente se distribuem. Portanto, faz sentido se pensar em utilizar Gaussianas adaptáveis para as quais o formato de cada componente (expresso pela matriz de covariância) é dado em função das direções principais da distribuição local dos dados.

Com base nisso, a técnica MPW define a estimação de uma densidade $p(\mathbf{x})$ da seguinte maneira:

$$\begin{aligned} p(\mathbf{x}) &= \frac{1}{l} \sum_{i=1}^l G_{\mu=\mathbf{x}_i, \Sigma=\Sigma_i}(\mathbf{x}) \\ &= \frac{1}{l} \sum_{i=1}^l \frac{1}{\sqrt{(2\pi)^d \det(\Sigma_i)}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{x}_i)' \Sigma_i^{-1} (\mathbf{x} - \mathbf{x}_i) \right], \end{aligned} \quad (4.23)$$

sendo l o número de amostras, d a dimensionalidade das amostras, e Σ_i a matriz de covariância do i -ésimo componente Gaussiano. Conforme já mencionado, a ideia é que esses componentes sejam dados em função das direções principais das amostras que estão na vizinhança de \mathbf{x}_i . Assim, temos

$$\Sigma_{\mathcal{H}_i} = \frac{\sum_{j=1 \dots l, j \neq i} \mathcal{H}(\mathbf{x}_j, \mathbf{x}_i) (\mathbf{x}_j - \mathbf{x}_i)' (\mathbf{x}_j - \mathbf{x}_i)}{\sum_{j=1 \dots l, j \neq i} \mathcal{H}(\mathbf{x}_j, \mathbf{x}_i)}, \quad (4.24)$$

onde $\mathcal{H}(\mathbf{x}, \mathbf{x}_i)$ é uma função de vizinhança que associa uma ponderação a cada amostra \mathbf{x} na vizinhança de \mathbf{x}_i , podendo-se limitar essa vizinhança às k_{cov} amostras mais próximas.

Na prática, o cálculo da inversa da matriz de covariância feito na Equação (4.23) pode se tornar um problema mal condicionado. A fim de evitar isso, uma pequena Gaussiana isotrópica e de variância σ^2 é adicionada a $\Sigma_{\mathcal{H}_i}$, de maneira que $\Sigma_i = \Sigma_{\mathcal{H}_i} + \sigma^2 \mathbf{I}_3$. Além disso, a avaliação da densidade em cada ponto do espaço pode ser simplificada se restringirmos a combinação de Gaussianas a apenas os k_{eval} componentes correspondentes aos k_{eval} vizinhos mais próximos¹³.

¹³Na Seção 3.3, já havia sido comentado que, segundo Leemput (2009), técnicas baseadas em simples médias ou frequências relativas de voxels não constituem a abordagem mais adequada para se estimar as densidades em um atlas do cérebro. Entretanto, é importante observar que o uso da técnica MPW no contexto deste trabalho de Mestrado não necessariamente contraria tais afirmações, visto que a estimação por ela realizada incorpora informações direcionais a respeito da distribuição das amostras em função da vizinhança, logo, não faz simplesmente uma contagem totalmente desprovida de uma noção da forma da função de densidade (como seria o caso com

Para a aplicação da MPW neste trabalho, foi utilizada a implementação existente na biblioteca ITK. As amostras são constituídas pelas localizações de pontos salientes, $\{\mathbf{s}\}$. Os valores padrão foram mantidos para os parâmetros ajustáveis, isto é: $\sigma^2 = 1$, $k_{cov} = 50$, $k_{eval} = 5$, e $\mathcal{K}(\mathbf{s}, \mathbf{s}_i) = G_{\mu=\mathbf{s}_i, \Sigma=\sigma^2 \mathbf{I}_3}(\mathbf{s})$. O resultado desse processo é uma imagem volumétrica cujos valores de intensidade em cada voxel são proporcionais à maior ou menor presença de pontos salientes naquela localidade ao longo das imagens de treinamento.

Após a estimação da densidade, é feito o cálculo dos descritores médios $\bar{\mathbf{h}}$: para cada voxel correspondente a um ponto maximal de $p_{\mathbf{x}}$, são computadas as médias dos descritores de todos os pontos salientes que foram detectados, em qualquer imagem de treinamento, dentro de uma região de vizinhança ao redor do pico maximal. Um ponto saliente médio é então criado na mesma localização desse pico maximal, e a ele é atribuído o descritor $\bar{\mathbf{h}}$. Os pontos salientes criados dessa maneira representam, ao mesmo tempo, os locais de maior probabilidade de ocorrência de detecção de um ponto saliente e os descritores com os quais o casamento com tais pontos é, em tese, mais provável.

4.6.2 Atlas de médias e frequências

Na construção do atlas probabilístico, a média dos descritores é computada em função da proximidade espacial entre os pontos salientes detectados em diferentes imagens. Entretanto, uma abordagem que talvez funcionasse melhor seria computar a média em função da similaridade entre os próprios descritores. Pela própria maneira como os descritores são construídos, a proximidade espacial seria implicitamente capturada, e supostamente sem riscos de se incluir na média descritores de pontos que na verdade não são tão similares aos demais (*outliers*). Assim sendo, foi construído um segundo atlas baseado nessa abordagem.

O procedimento para construção do atlas de médias e frequências é análogo ao método utilizado no já referido trabalho de Mondal et al. (2014) para descritores SIFT¹⁴. Dado um par de imagens de treinamento e seus respectivos conjuntos de pontos salientes, realiza-se o casamento entre os pares de pontos com descritores similares (usando o método descrito na Seção 4.5.3) e a posterior mesclagem de cada par de pontos casados em um único ponto saliente médio (com valores médios tanto para a posição espacial como para cada componente do vetor que compõe o descritor). Um contador $c(\mathbf{s})$ associado a cada ponto saliente conta o número de vezes em

a PW tradicional). Além disso, há de se considerar também que utilizar o método de Leemput (2009) seria um trabalho excessivamente complexo para este trabalho de Mestrado, e que não se tem na literatura resultados da sua aplicação em atlas de pontos salientes para justificar tal esforço na presente etapa do projeto. Nada impede, porém, que isso possa ser testado em um trabalho futuro.

¹⁴Vide Seção 3.3

que cada ponto foi mesclado. Inicialmente, todo ponto saliente tem $c(\mathbf{s}) = 1$. Quando ocorre a mesclagem entre dois pontos, \mathbf{s}_1 e \mathbf{s}_2 , o ponto saliente médio resultante tem $c(\bar{\mathbf{s}}) = c(\mathbf{s}_1) + c(\mathbf{s}_2)$. Todo esse procedimento se repete sucessivamente até que todas as imagens de treinamento tenham sido processadas.

Ao final do processo, as contagens $c(\mathbf{s})$ expressam a frequência relativa de cada ponto saliente e dão uma noção do quão estável é cada ponto ao longo da população representada nas imagens de treinamento. Quanto mais estável, maior a probabilidade de que um casamento seja feito entre tal ponto e algum ponto saliente de uma nova imagem clínica. Adicionalmente, para evitar a ocorrência de um número excessivamente grande de pontos salientes no atlas resultante, uma limiarização é aplicada para eliminar do atlas todos os pontos para os quais a contagem $c(\mathbf{s})$ foi muito baixa (mesclaram-se com poucos ou talvez nenhum outro ponto saliente durante o treinamento). No caso, eliminaram-se todos os pontos com $c(\mathbf{s})$ inferior a 15% do número total de imagens de treinamento (valor escolhido empiricamente).

4.7 Posicionamento e deformação dos modelos

Dado um conjunto $\left\{ \left(\mathbf{s}_1^{(m)}, \mathbf{s}_2^{(m)} \right) \right\}$ de pares de pontos salientes casados entre uma imagem clínica e um atlas, a inicialização dos modelos geométricos deformáveis é finalmente feita por meio da estimação de uma transformação deformável *B-spline* que mapeia as localizações espaciais dos pontos $\left\{ \mathbf{s}_1^{(m)} \right\}$ para os pontos $\left\{ \mathbf{s}_2^{(m)} \right\}$ com erro mínimo, interpolando valores para o restante do domínio da imagem. Essa mesma transformação, que pode ser representada por um campo vetorial de deformação, pode então ser aplicada ao modelo (vértices da malha) de uma estrutura de interesse para, por fim, conferir-lhe uma configuração de posição e forma que seja próxima da estrutura de interesse na imagem alvo. Para realizar essa tarefa de mapeamento, foi utilizada uma generalização do método multinível de interpolação *B-spline* de Lee, Wolberg e Shin (1997), proposta por Tustison e Gee (2006), e cuja implementação encontra-se disponível na biblioteca ITK. A técnica é brevemente explicada nos parágrafos a seguir.

Em sentido simples, o termo *spline* é usado para se referir a qualquer função de suporte limitado que interpole um conjunto de pontos de maneira contínua em suas vizinhanças. Já o termo *B-spline* (*basis spline*) refere-se a um conjunto de $d + 1$ funções polinomiais de grau d usadas para especificar *splines*. Um objeto *B-spline* de dimensão n , que aproxima um conjunto de pontos em um domínio de coordenadas $\mathbf{x} = (x_1, \dots, x_n)$, é definido em termos das funções *B-spline* $B_{i,d}(\cdot)$ e de um reticulado de pontos de controle, $\mathbf{P}_{i_1, \dots, i_n}$. Matematicamente,

$$S(x_1, \dots, x_n) = \sum_{i_1=1}^{M_1} \dots \sum_{i_n=1}^{M_n} \mathbf{P}_{i_1, \dots, i_n} \prod_{j=1}^n B_{i_j, d_j}(u_j), \quad (4.25)$$

em que $S(\cdot)$ é o objeto *B-spline*, M_1, \dots, M_n são os números de pontos de controle em cada dimensão, e $u_j = x_j - i_j \delta_j$, com δ_j sendo o espaçamento entre pontos de controle na dimensão j (TUSTISON; GEE, 2006; LEE; WOLBERG; SHIN, 1997). O produtório nessa equação denota a formação de uma função *B-spline* nD a partir da combinação de n funções *B-spline* 1D com graus d_1, \dots, d_n . Na prática, o que a equação faz é calcular uma média ponderada que determina a contribuição de cada ponto de controle a $S(x_1, \dots, x_n)$ em função de sua distância a $\mathbf{x} = (x_1, \dots, x_n)$. A rapidez com que essa contribuição decai depende dos graus das funções $B_{i_j, d_j}(\cdot)$.

A partir de tal formulação, o problema original reduz-se a encontrar a configuração de pontos de controle $\mathbf{P}_{i_1, \dots, i_n}$ que melhor aproxima por $S(\cdot)$ um conjunto de pontos de entrada $\{\mathbf{x}\}$. Tustison e Gee (2006) estimam a solução utilizando um critério de mínima soma dos erros quadráticos entre cada ponto de entrada e todos os pontos de controle, incluindo ainda um peso ω_m a cada ponto de entrada para permitir que se indique, caso necessário, que certos pontos devem ser melhor aproximados do que outros. Ou seja, minimiza-se

$$e(\mathbf{P}_{i_1, \dots, i_n}) = \sum_{m=1}^M \omega_m \left[\mathbf{P}_{i_1, \dots, i_n} \prod_{j=1}^n B_{i_j, d_j}(u_j^{(m)}) - \mathbf{x}_m \prod_{j=1}^n B_{i_j, d_j}(u_j^{(m)}) \right]^2, \quad (4.26)$$

onde o sobrescrito (m) indica os parâmetros para o m -ésimo ponto de entrada e \mathbf{x}_m é o m -ésimo ponto de entrada em uma ordenação qualquer. A minimização de tal critério leva à seguinte solução:

$$\mathbf{P}_{i_1, \dots, i_n} = \frac{\sum_{m=1}^M \omega_m \prod_{j=1}^n B_{i_j, d_j}^2(u_j^{(m)})}{\sum_{m=1}^M \omega_m \prod_{j=1}^n B_{i_j, d_j}(u_j^{(m)})}. \quad (4.27)$$

Adicionalmente, a técnica emprega um método de refinamento hierárquico do reticulado de pontos de controle, computando os resultados em múltiplos níveis de resolução e melhorando-os a cada nível. Neste trabalho, foram utilizados $5 \times 5 \times 5$ pontos de controle no primeiro nível e, a cada nível posterior, o número de pontos de controle em cada dimensão é dobrado. No total, foram aplicados 3 níveis de refinamento. As funções *B-spline* foram definidas também com grau 3.

Para a aplicação pretendida neste trabalho, cada ponto de entrada a ser aproximado por $S(\cdot)$ é dado pelo vetor diferença $\mathbf{s}_2^{(m)} - \mathbf{s}_1^{(m)}$, em que $\mathbf{s}_1^{(m)}$ e $\mathbf{s}_2^{(m)}$ são as localizações dos pontos salientes no m -ésimo par de pontos previamente casados, com $\mathbf{s}_1^{(m)}$ correspondendo a um ponto “móvel” e $\mathbf{s}_2^{(m)}$ correspondendo a um ponto “fixo”. Fazendo dessa maneira, a estimação pelo

método aqui descrito (na implementação da biblioteca ITK) produz como saída uma imagem volumétrica 3D cujos voxels são vetores 3D que expressam o deslocamento que aproximadamente mapeia a localização dos pontos em $\{\mathbf{s}_1^{(m)}\}$ para os pontos $\{\mathbf{s}_2^{(m)}\}$, interpolando valores para o restante do domínio da imagem. Quando se faz uso de algum dos atlas de pontos salientes, o conjunto $\{\mathbf{s}_1^{(m)}\}$ passa a representar pontos salientes do atlas que foram casados aos pontos da imagem clínica, podendo, portanto, corresponder a pontos de pico da função de densidade de probabilidade da distribuição espacial dos pontos (no caso do atlas probabilístico) ou a pontos médios de alta estabilidade (no caso do atlas de médias e frequências). Além disso, os atlas são utilizados para determinar os valores dos pesos ω_m : para o atlas probabilístico, ω_m é o valor da função de densidade de probabilidade na posição espacial do m -ésimo ponto saliente do atlas. Já para o atlas de médias e frequências, ω_m é o valor da frequência $c(\mathbf{s})$ para o m -ésimo ponto saliente do atlas. Em ambos os casos, isso significa que pontos salientes para os quais há uma maior “confiança” têm uma maior tendência a serem mais precisamente mapeados. Essa é uma vantagem que não se conseguiria sem os atlas.

É importante mencionar ainda que, com essa técnica, já se realiza ao mesmo tempo o posicionamento e o ajuste do modelo da estrutura de interesse às informações da imagem.

4.8 Análise dos resultados

A técnica de detecção de pontos salientes 3D foi avaliada utilizando o conjunto de imagens sintéticas de pequenas estruturas 3D de alta curvatura descrito na Subseção 4.2.5. Como essas imagens foram geradas a partir de um modelo de equações paramétricas conhecidas, a exata localização onde se esperava detectar um ponto saliente em cada uma dessas estruturas, de acordo com o funcionamento esperado para a técnica, era previamente conhecida (Figura 4.8). Então as distâncias entre os pontos detectados e as localizações esperadas para esses pontos, além do número de falsos positivos (pontos excessivos), são usadas como métricas de avaliação.

Adicionalmente, também foi feita a detecção dos pontos salientes em 18 imagens da base de dados *BrainWeb*, as quais se diferenciam apenas no nível de degradação porruído e *bias field* que cada uma possui. Nesse caso não houve avaliação de precisão ou falsos positivos, apenas são contadas as quantidades de pontos salientes detectados em cada caso. Esse teste foi feito apenas para dar uma noção mais realista de como o número de pontos detectados em uma imagem de RM do cérebro varia conforme as condições de degradação da imagem.

Para as demais técnicas, uma das formas de avaliação envolveu a criação de um conjunto de imagens de RM deformadas e de modelos (malhas) deformados que simulam variações anatô-

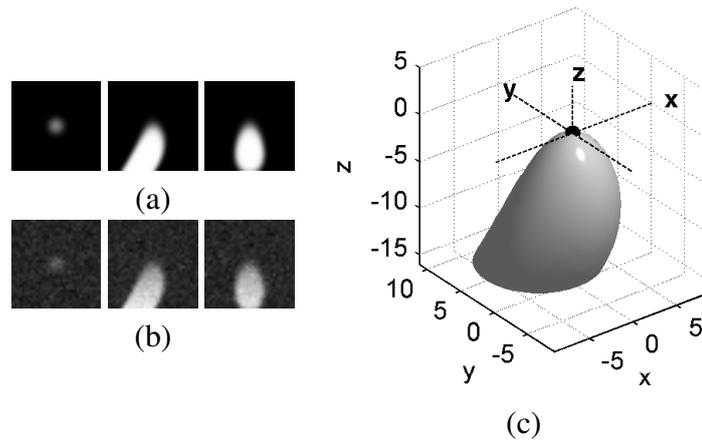


Figura 4.8: Exemplo de imagem sintética 3D usada para testar a técnica de detecção de pontos salientes. Projeções 2D de imagens representando a mesma estrutura 3D (a) sem *bias field* ou ruído; e (b) com *bias field* ($BF = 40\%$) e ruído Gaussiano ($\sigma_{ruído} = 5\%$). (c) Renderização 3D da estrutura, destacando a localização esperada para um ponto saliente no local em que a curvatura 3D da estrutura é máxima.

micadas do cérebro. No total, 105 imagens e 105×3 malhas foram geradas a partir da imagem de referência da base de dados NAC¹⁵ e das malhas do corpo caloso (CC - *corpus callosum*), hipocampo esquerdo (LH - *left hippocampus*) e hipocampo direito (RH - *right hippocampus*) presentes nessa mesma base. As deformações foram realizadas através da criação de transformações *B-spline* cúbicas aleatoriamente parametrizadas e da aplicação dessas transformações à imagem de RM de referência e às malhas das estruturas mencionadas. Adicionalmente, também essas transformações também foram aplicadas aos pontos salientes detectados na imagem de referência, resultando em uma lista de localizações esperadas para cada ponto saliente de uma imagem deformada - e, conseqüentemente, numa lista de casamentos esperados entre os pares de pontos dessas imagens.

Assim, a técnica de descrição e casamento de pontos salientes pôde ser quantitativamente avaliada por meio da comparação entre os casamentos de pontos salientes esperados e aqueles efetivamente encontrados pela técnica. Como métrica de avaliação, foi usada a Medida-F1 (POWERS, 2011):

$$F1 = \frac{2VP}{2VP + FP + FN}, \quad (4.28)$$

em que VP é a contagem de Verdadeiros Positivos (casamentos esperados que foram encontrados), FP é a contagem de Falsos Positivos (casamentos encontrados que não eram esperados), e FN é a contagem de Falsos Negativos (casamentos esperados que não foram encontrados). A Medida-F1 equivale à média ponderada entre a precisão (valor preditivo positivo) e a revocação (taxa de verdadeiros positivos) de um sistema de classificação binária; logo, é um valor numé-

¹⁵Vide Subseção 4.2.1.

rico único que resume o desempenho de tal sistema e produz um valor baixo se qualquer uma dessas medidas também tiver um valor baixo. Isso a torna uma medida eficaz para se avaliar os casamentos de pontos salientes simultaneamente em seus acertos e erros.

Por fim, a etapa de inicialização dos modelos geométricos deformáveis foi avaliada calculando-se a sobreposição espacial entre as malhas ajustadas pelo método aqui proposto e um padrão ouro que indica a configuração de posição e forma reais da estrutura de interesse na imagem clínica. Para realizar esse cálculo, as malhas foram convertidas para imagens binárias (usando um algoritmo de rasterização) e calcularam-se os coeficientes de similaridade de Jaccard (JC) e Dice (DC) entre elas (AGRAWAL; SHARMA, 2014). Tais medidas, cujos valores de saída variam entre 0 e 1, são definidas como

$$\begin{aligned} JC(A, B) &= \frac{|A \cap B|}{|A \cup B|}, \\ DC(A, B) &= \frac{2|A \cap B|}{|A| + |B|}, \end{aligned} \quad (4.29)$$

em que A e B são os conjuntos dos voxels de valor 1 nas imagens binárias.

Primeiramente, foi feito um teste preliminar da técnica de inicialização dos modelos sem uso de atlas (apenas os pontos salientes detectados na imagem de referência da base NAC são usados, todos com o mesmo peso), utilizando as imagens de RM sinteticamente deformadas como imagens de entrada. Neste caso, a inicialização tentou aproximar os modelos da imagem de referência original (da base NAC) aos modelos das imagens deformadas, que servem como padrão ouro para quantificação dos resultados. Esses experimentos envolveram modelos de três estruturas cerebrais: CC, LH e RH.

Em seguida, foi feito o teste principal deste trabalho, que utilizou as imagens clínicas da base EADC-ADNI¹⁶ como imagens de entrada. Neste caso, a inicialização tentou aproximar os modelos da imagem de referência original (da base NAC) às marcações manuais das estruturas de interesse (delineadas por especialistas), as quais servem como padrão ouro para quantificação dos resultados. Como as imagens dessa base só possuem marcações para o LH e o RH, apenas essas duas estruturas foram avaliadas nesses experimentos. Esta avaliação foi subdividida em três casos:

1. Inicialização sem uso de atlas;
2. Inicialização utilizando o atlas probabilístico;
3. Inicialização utilizando o atlas de médias e frequências.

¹⁶Vide Subseção 4.2.3.

4.9 Considerações finais

Este capítulo apresentou a metodologia desenvolvida neste trabalho para atingir os objetivos definidos na Seção 1.2. É importante observar que, embora tenha sido feita uma especificação precisa das técnicas que se pretende utilizar para a detecção dos pontos salientes, descrição e casamento desses pontos, construção dos atlas e inicialização dos modelos geométricos deformáveis, nada impede, em tese, que a mesma abordagem geral seja implantada com outras técnicas específicas em suas diferentes etapas. Por exemplo, usando métodos alternativos de detecção de pontos salientes ou outros descritores para descrever esses pontos. Nada impede, também, que se utilize a mesma abordagem em outros tipos de imagens - por exemplo, em imagens de CT, raio-X, ou mesmo imagens volumétricas não médicas, desde que se tenha todos os elementos necessários para sua aplicação.

Capítulo 5

RESULTADOS E DISCUSSÃO

Este capítulo apresenta e discute os resultados obtidos nesta pesquisa, no que diz respeito ao pré-processamento, à detecção de pontos salientes 3D, ao casamento de pontos salientes, e à inicialização (posicionamento e ajuste) dos modelos geométricos deformáveis.

5.1 Pré-processamento

São apresentados a seguir os resultados obtidos nas etapas do pré-processamento das imagens de treinamento para construção do atlas.

5.1.1 Redução de ruído

Na Figura 5.1, é exibida uma comparação visual entre regiões de uma das imagens de treinamento antes e após a redução de ruído pelo algoritmo NLM. A comparação mostra que o ruído é eliminado sem que ocorra degradação aparente nas bordas ou perdas significativas de informação estrutural na região do encéfalo. Portanto, pode-se dizer que o algoritmo teve resultados satisfatórios para o propósito deste trabalho.

Convém mencionar que, embora existam métodos quantitativos para melhor avaliação de um algoritmo de redução de ruído (como a razão sinal-ruído), não se fez uso destes aqui porque normalmente é requerida a existência de uma imagem modelo (sem ruído) para que tais medidas possam ser calculadas. Também não foi feita a computação da diferença entre a imagem ruidosa original e a imagem após redução de ruído, porque o ruído de Rice presente nas imagens de RM é dependente do sinal.

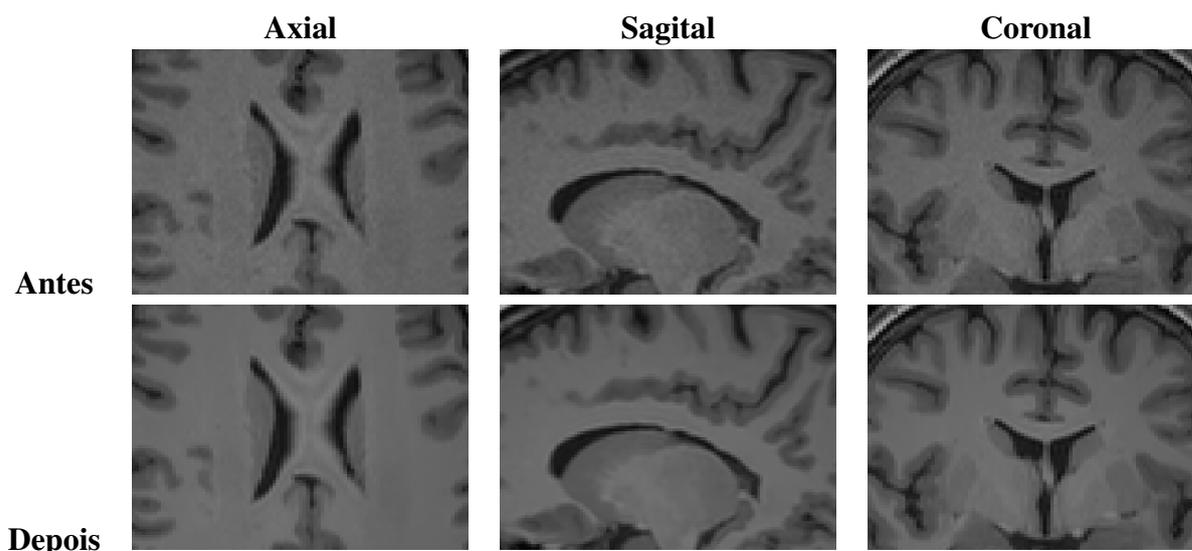


Figura 5.1: Resultado visual da redução de ruído em uma imagem de treinamento, nas vistas axial, sagital e coronal.

5.1.2 Correção de *bias field*

Dentre os três hospitais nos quais as imagens que compõem a base de dados IXI (imagens de treinamento) foram adquiridas, o *Institute of Psychiatry (IoP)* é aquele que contém as imagens mais fortemente afetadas por *bias field*. A Figura 5.2 mostra o resultado da correção de uma dessas imagens nas vistas axial, sagital e coronal, juntamente com o respectivo *bias field* estimado e corrigido pelo algoritmo N4. Comparando-se visualmente as imagens antes e depois do processamento, é possível perceber a correção.

5.1.3 Alinhamento espacial (corregistro)

Os resultados do alinhamento espacial (corregistro) podem ser qualitativamente analisados pela análise do padrão “tabuleiro de xadrez”, que compõe partes (“blocos”) da imagem de referência e partes da imagem de treinamento antes e depois do alinhamento entre elas. Esse padrão é mostrado na Figura 5.3. Percebe-se, pela figura, que o corregistro afim proporcionou uma grande melhoria no alinhamento entre as imagens, mesmo realizando transformações que são apenas globais. Esse alinhamento é essencial para que a construção dos atlas tenha sucesso.

5.1.4 Extração de máscaras do encéfalo

O resultado da extração de máscaras binárias do encéfalo com utilização da ferramenta ROBEX pode ser visto na Figura 5.4. Para fins de comparação, é feita uma sobreposição com

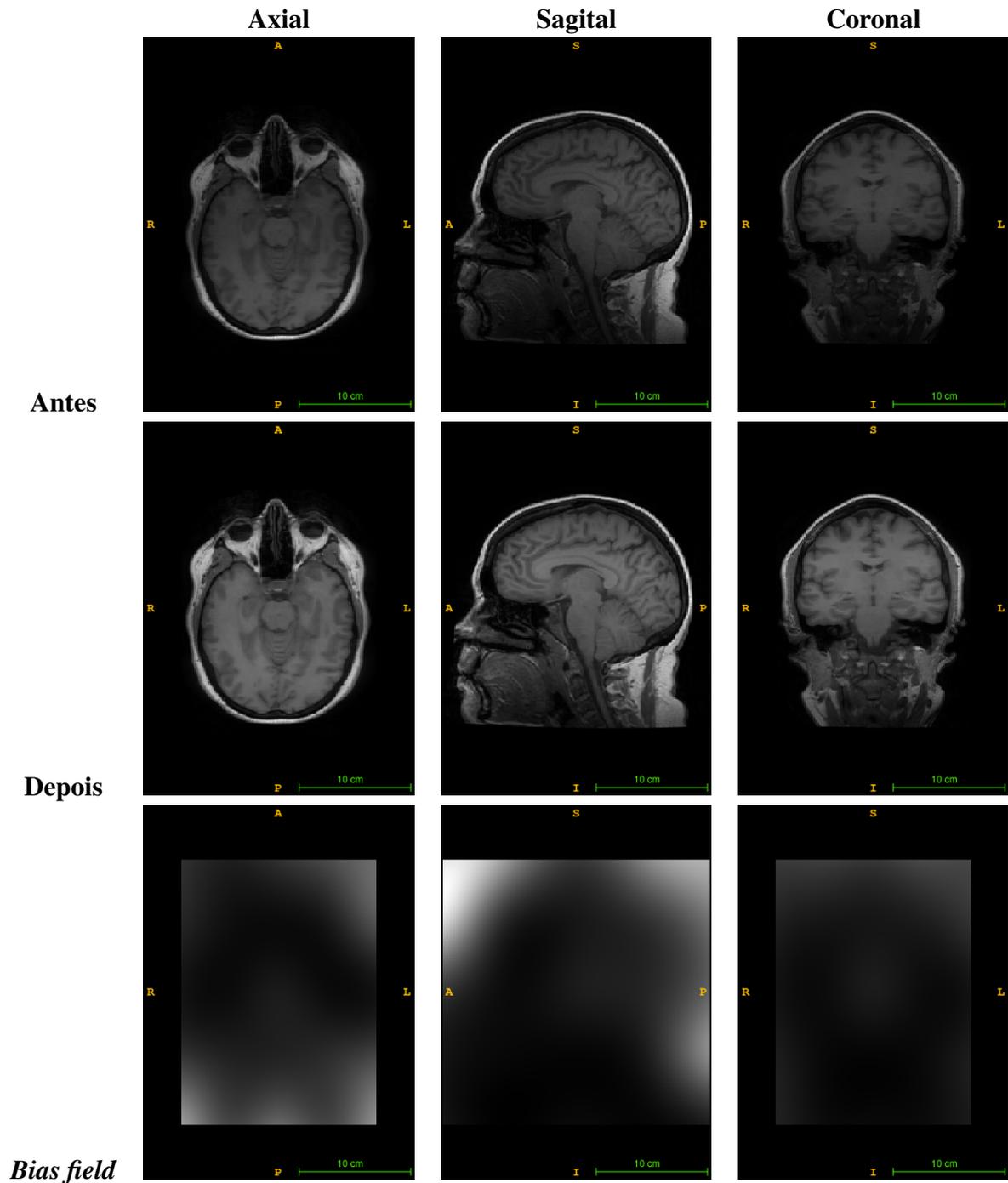


Figura 5.2: Resultado da correção do *bias field* em uma imagem de treinamento, nas vistas axial, sagital e coronal.

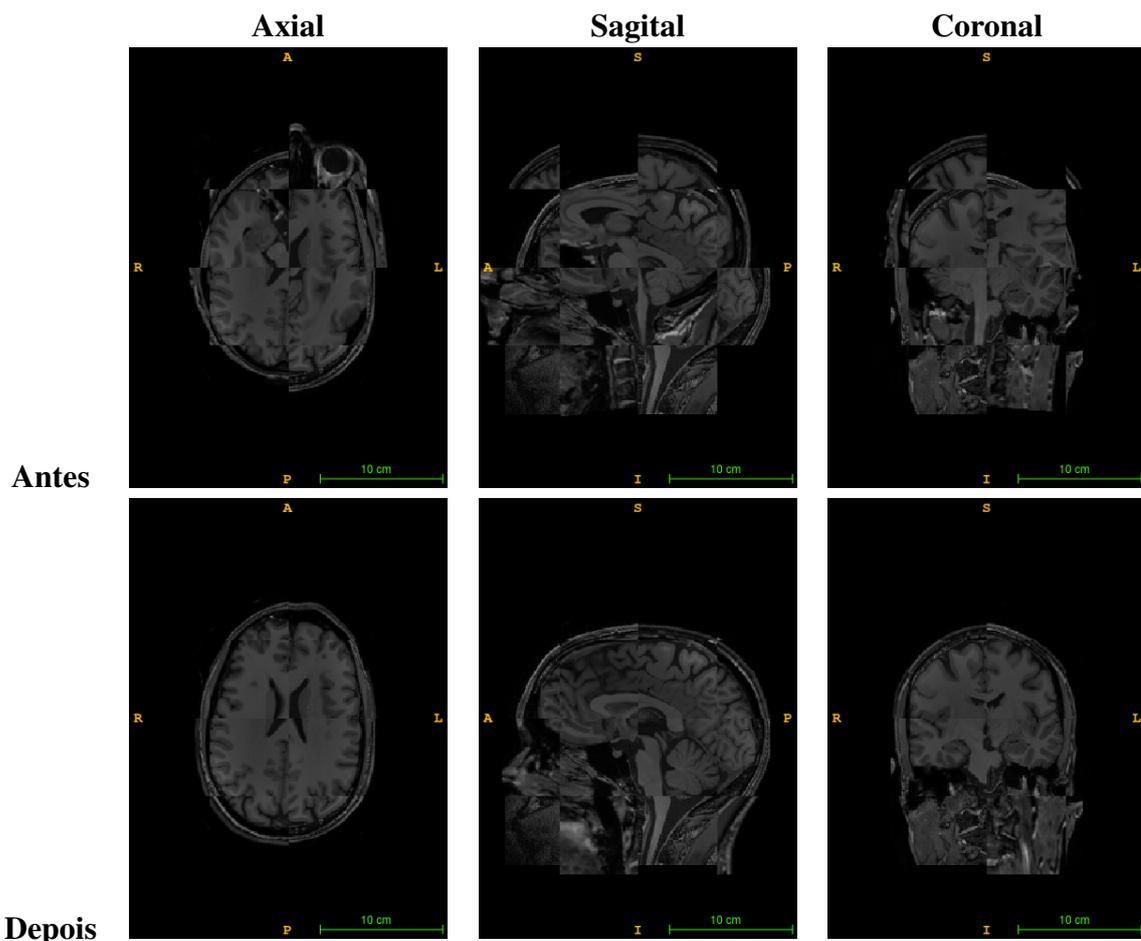


Figura 5.3: Resultado do alinhamento espacial (corregistro) entre uma imagem de treinamento e a imagem de referência, ilustrado pelo padrão “tabuleiro de xadrez”, nas vistas axial, sagital e coronal.

a imagem original na qual a extração foi realizada. Visualmente, os resultados se mostraram bastante satisfatórios. Todo o processamento relacionado à detecção de pontos salientes na imagem de um paciente pode então ser realizado somente nos voxels contidos dentro da região delimitada pela máscara binária do seu respectivo encéfalo.

5.2 Detecção de pontos salientes 3D

O gráfico da Figura 5.5 resume os resultados obtidos nos testes envolvendo as imagens sintéticas de estruturas 3D de alta curvatura. Esses resultados foram agrupados em 6 níveis de ruído e 3 níveis de *bias field*, totalizando 18 grupos de 105 imagens cada. Cada grupo no diagrama de caixas representa a distribuição das distâncias (erros) entre as localizações esperadas para a detecção dos pontos salientes e a localização do ponto saliente mais próximo efetivamente detectado pela técnica.

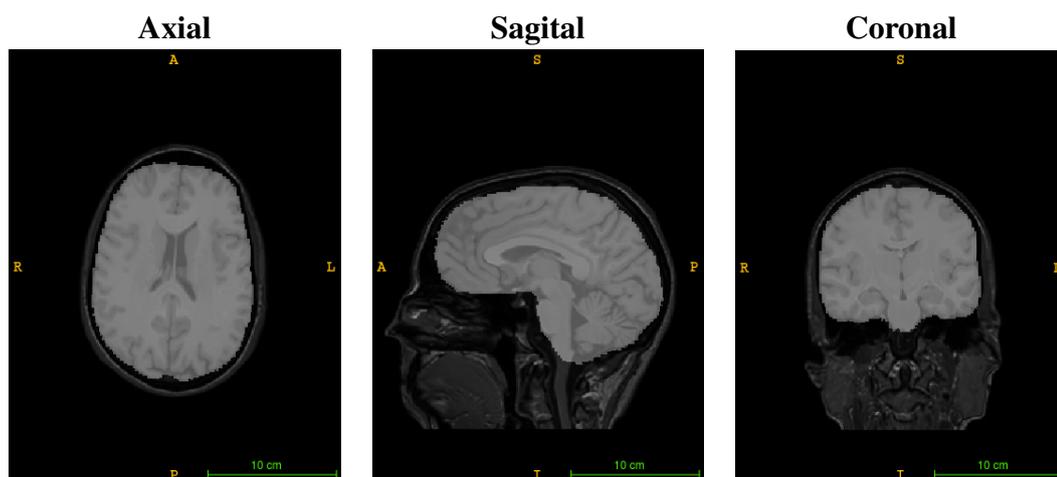


Figura 5.4: Resultado da extração da máscara do encéfalo em uma imagem de treinamento, utilizando a ferramenta ROBEX. Sobreposição da máscara obtida (em tons de cinza mais claros) sobre a imagem na qual a extração foi feita, nas vistas axial, sagital e coronal.

Primeiramente, os valores no gráfico mostram que, mesmo utilizando um detector baseado em PC, o *bias field* ainda pode conseguir ter alguma influência negativa na eficácia da detecção dos pontos salientes, especialmente quando ele se distribui da maneira que foi simulada nos testes: com a máxima atenuação da intensidade ao longo da imagem ocorrendo justamente na região em que um ponto saliente deveria ser detectado. No entanto, os valores também mostram que essa influência só começa a ser relevante quando o *bias field* atinge o nível mais alto: $BF = 40\%$ de atenuação máxima, o que não é mais uma ocorrência tão comum nos scanners de RM mais modernos. Para níveis menores de atenuação por *bias field*, a média e a mediana do erro de localização não chegaram sequer a 2 voxels, mesmo nas imagens com os maiores níveis de ruído ($\sigma_{ruído} = 9\%$).

Por sua vez, uma análise dos resultados obtidos para cada nível fixo de *bias field* também indica um crescimento da distância à medida que o nível de ruído aumenta. Entretanto, trata-se de um crescimento lento, pois mesmo as distâncias médias e medianas obtidas nas imagens mais ruidosas ainda não ultrapassam, ou mal ultrapassam, a marca dos 2 voxels. Podemos dizer que a conhecida sensibilidade a ruído do modelo de PC não atrapalhou significativamente a precisão de localização dos pontos salientes nos locais onde se esperava detectá-los.

Os valores atípicos (*outliers*) que se vê no gráfico correspondem, na maior parte dos casos, às imagens em que a estrutura 3D representada aparece rotacionada. Uma vez que todas as funções de *bias field* geradas para um nível fixo de atenuação são iguais, alguns ângulos de rotação particulares podem fazer com que uma estrutura tenha uma maior parte dos seus voxels localizada em regiões nas quais a atenuação é mais intensa, dificultando uma detecção precisa. Ao longo de todo o experimento, a maior distância registrada foi de aproximadamente 4,5

voxels e a menor foi 0 (que ocorreu na grande maioria dos casos em que há pouco ruído e pouco *bias field*).

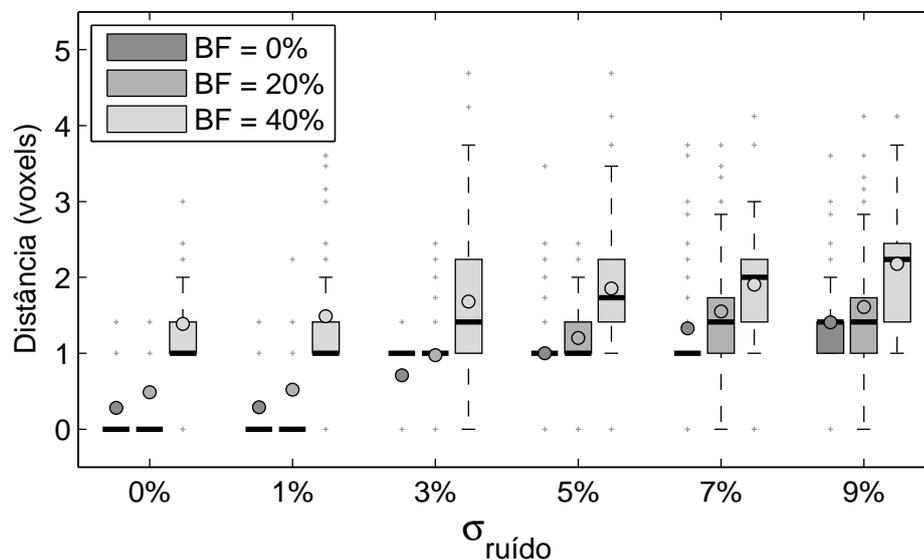


Figura 5.5: Resultados obtidos com a técnica de detecção de pontos salientes nos testes com imagens sintéticas de estruturas 3D de alta curvatura. Cada uma das 18 “caixas” representa um grupo de 105 imagens sintéticas degradadas por uma determinada quantidade de ruído e *bias field*. Os círculos indicam a média e as linhas horizontais mais espessas indicam a mediana. Os pontos cinzas isolados são os valores atípicos (*outliers*).

As estruturas 3D nas imagens de teste foram construídas de maneira que apenas 1 único ponto saliente fosse esperado em cada imagem. Entretanto, poderia ocorrer a detecção de falsos positivos, especialmente quando a imagem possuísse muito ruído. Por essa razão, foram computadas e analisadas as médias e desvios padrão obtidos na contagem desses falsos positivos ao longo das imagens testadas. Os resultados são mostrados na Tabela 5.1. Novamente, os valores encontram-se agrupados conforme o nível de ruído e *bias field*. Com base nos números obtidos, observações análogas àquelas já feitas quanto à influência do *bias field* e do ruído podem ser mantidas. As quantidades de falsos positivos encontrados nesse teste são consideradas pequenas, com valores de média e desvio padrão permanecendo abaixo de 1; porém, é importante notar que esses valores se devem, também, ao fato de a imagem ter tamanho pequeno e conter apenas uma única estrutura com ponto saliente detectável. O teste de detecção feito com imagens da base *BrainWeb*, cujos resultados são apresentados na Tabela 5.2, dá uma noção um pouco mais realista de como o número de pontos salientes detectados no total (incluindo os falsos positivos) poderia variar em uma imagem de RM (que contém várias estruturas detectáveis) de acordo com as condições de ruído e *bias field*. Neste caso não havia uma relação de pontos salientes esperados, então apenas se contou o número de pontos salientes detectados em cada caso para fins de comparação entre eles. Os valores novamente indicam que, para ruído pequeno, o *bias field* pouco influencia, mas quando a quantidade de ruído é alta, os resultados

podem ser bem mais significativamente afetados e o *bias field* começa a ter maior influência.

Tabela 5.1: Média e desvio padrão da contagem de falsos positivos (pontos salientes excessivos) detectados nas imagens sintéticas de estruturas 3D de alta curvatura. Cada uma das 18 células da tabela representa um grupo de 105 imagens sintéticas degradadas por uma determinada quantidade de ruído e *bias field*.

	$BF = 0\%$	$BF = 20\%$	$BF = 40\%$
$\sigma_{\text{ruído}} = 0\%$	0 ± 0	0 ± 0	$0,1 \pm 0,31$
$\sigma_{\text{ruído}} = 1\%$	0 ± 0	0 ± 0	$0,09 \pm 0,28$
$\sigma_{\text{ruído}} = 3\%$	0 ± 0	$0,02 \pm 0,14$	$0,16 \pm 0,37$
$\sigma_{\text{ruído}} = 5\%$	0 ± 0	$0,04 \pm 0,19$	$0,24 \pm 0,49$
$\sigma_{\text{ruído}} = 7\%$	$0,02 \pm 0,14$	$0,16 \pm 0,44$	$0,39 \pm 0,61$
$\sigma_{\text{ruído}} = 9\%$	$0,05 \pm 0,21$	$0,19 \pm 0,44$	$0,69 \pm 0,71$

Tabela 5.2: Contagem das quantidades de pontos salientes detectados em cada uma das 18 imagens obtidas da base de dados *BrainWeb*, com seus respectivos níveis de ruído e *bias field*.

	$BF = 0\%$	$BF = 20\%$	$BF = 40\%$
$\sigma_{\text{ruído}} = 0\%$	224	224	224
$\sigma_{\text{ruído}} = 1\%$	236	230	236
$\sigma_{\text{ruído}} = 3\%$	300	252	283
$\sigma_{\text{ruído}} = 5\%$	388	426	271
$\sigma_{\text{ruído}} = 7\%$	492	438	316
$\sigma_{\text{ruído}} = 9\%$	532	650	429

5.3 Descrição e casamento de pontos salientes 3D

As subseções a seguir apresentam e discutem resultados para os três métodos descritos na Seção 4.5 para criação e casamento de descritores de pontos salientes. Dos três, apenas o terceiro deles, baseado na técnica *shape contexts*, foi capaz de efetivamente prover uma representação discriminativa para os pontos salientes detectados nas imagens.

5.3.1 Métodos 1 e 2

Ambos falharam em conseguir representar e casar pontos salientes de regiões similares de imagem até mesmo nos casos de teste mais simples possíveis. A razão da falha parece ter sido a mesma para os dois métodos: uma fraca representabilidade direcional que impossibilitava a boa discriminação entre os descritores de pontos salientes.

No Método 1, as orientações em O corresponderam às orientações do banco de filtros 3D de log-Gabor. Mas a construção dos filtros no domínio da frequência sempre é feita de maneira a

se capturar informações de apenas metade do espectro, pois o domínio é simétrico em relação à origem (frequência zero). Conseqüentemente, as orientações do descritor também limitavam-se a representar apenas metade do espaço de coordenadas esféricas, com θ variando entre 0° e 90° e ϕ variando entre 0° e 180° (quando $\theta = 0^\circ$) ou entre 0° e 360° (quando $\theta > 0^\circ$). O problema é que, devido a essa limitação, regiões pouco semelhantes na imagem ainda poderiam ter perfis similares de distribuição das orientações de máxima PC. Assim, pontos salientes detectados em regiões cerebrais consideravelmente distintas eram, muitas vezes, classificados como pontos correspondentes.

No Método 2, a causa do problema foi similar: devido à própria maneira como cada $PC_{\mathbf{x}}$ é computada, também sendo baseada nas orientações determinadas pelos filtros de log-Gabor, os autovetores de $M_{PC}(\mathbf{x})$ não tinham uma representação direcional suficientemente precisa para uma boa discriminação entre regiões da imagem, especialmente quando estas continham estruturas cerebrais simétricas. Além disso, havia a presença de “instabilidades” na direção dos autovetores mais próximos dos pontos salientes, porque justamente na proximidade desses pontos os três autovalores começavam a se tornar mais próximos em valor e reduzia-se a predominância de uma orientação em relação às demais. O problema foi ainda mais intensificado pelo fato de que justamente os voxels mais próximos do ponto saliente eram os que possuíam maior peso na construção do descritor.

Por essas razões, os experimentos descritos na Seção 4.8, que proveriam resultados quantitativos para avaliação, nem chegaram a ser realizados para tais métodos. Isso foi feito apenas com o método 3, que foi o único efetivamente utilizado nos demais passos do processamento. Seus resultados são apresentados a seguir.

5.3.2 Método 3 (*shape contexts*)

O gráfico na Figura 5.6 ilustra os resultados obtidos nos casamentos entre os pontos salientes detectados na imagem de RM de referência da base IXI e os pontos salientes de cada uma das imagens de RM sinteticamente deformadas que foram geradas para este teste. De acordo com os valores obtidos, o uso do descritor baseado em *shape contexts* e da distância χ^2 como métrica de dissimilaridade produziu valores individuais de Medida-F1 entre 0,7 e 0,89. O valor médio, destacado pela linha tracejada em cor vermelha, foi de 0,797. Em geral, as falhas que ocorreram na identificação de algum par de pontos correspondentes aconteceram não apenas por questões de diferenças nos descritores, mas também porque as variações anatômicas (simuladas pelas deformações) podiam fazer com que não ocorresse a detecção, em uma imagem deformada, de algum ponto saliente equivalente a algum outro que houvesse sido detectado na

imagem de RM de referência.

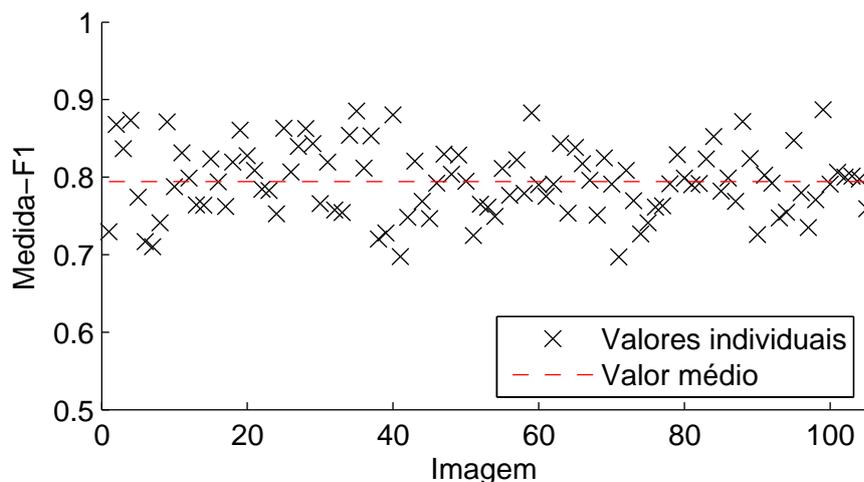


Figura 5.6: Valores individuais e valor médio de Medida-F1 obtidos com a técnica de descrição e casamento de pontos salientes baseada em *shape contexts*, envolvendo a imagem de RM de referência e as imagens de RM sinteticamente deformadas.

5.4 Atlas de pontos salientes 3D

As subseções a seguir apresentam os resultados obtidos na construção dos dois tipos de atlas desenvolvidos neste trabalho.

5.4.1 Atlas probabilístico

Na Figura 5.7, são mostradas algumas projeções do resultado obtido para a função de densidade expressa no atlas probabilístico. A imagem volumétrica que representa essa densidade foi sobreposta à imagem de referência da base NAC para fins de referenciação anatômica. Nessas projeções, é possível observar quais foram as regiões com maior incidência de pontos salientes com base nos níveis de intensidade de cada voxel, que mapeiam diretamente os valores da densidade em uma escala de cores. Analisando visualmente a distribuição, é possível observar fortes picos nas proximidades de certas estruturas cerebrais importantes, como na cabeça do corpo caloso, nas extremidades dos ventrículos, nos hipocampos, e em várias outras regiões do cérebro, especialmente as mais centrais. Em regiões onde a variabilidade anatômica é muito maior (como em basicamente todo o sistema cortical nas extremidades do encéfalo) ou onde o contraste é muito baixo e normalmente não há informação suficiente para a detecção de um ponto saliente (como na região do tálamo), os valores de densidade foram consideravelmente mais baixos. Mas as estruturas que mais interessavam no contexto deste trabalho aparecem bem representadas.

A Figura 5.8, por sua vez, contém uma renderização 3D da mesma função de densidade. Nessa figura é possível observar com mais clareza o que foi mencionado quanto à fraca representabilidade do sistema cortical e a predominância de picos de densidade em regiões mais centrais do encéfalo. Os descritores médios do atlas probabilístico, como mencionado na Subseção 4.6.1, são computados ao redor dos pontos maximais vistos nas figuras.

5.4.2 Atlas de médias e frequências

Na Figura 5.9, são mostradas algumas projeções do resultado obtido para o atlas de médias e frequências. Assim como feito na Figura 5.7, a imagem resultante, com as localizações e frequências de cada ponto médio, foi sobreposta à imagem de referência da base NAC para fins de referência anatômica. Nessas projeções, é possível observar quais foram as regiões com os pontos salientes mais estáveis a partir dos níveis de intensidade em cada voxel, que mapeiam diretamente os valores da contagem $c(\mathbf{s})$ na mesma escala de cores usada na Figura 5.7. Analisando visualmente, nota-se que os locais de pontos salientes estáveis acabaram correspondendo a muitos dos locais de pico encontrados no atlas probabilístico; isto é, próximos às mesmas estruturas cerebrais, salvo pequenas variações. E, analogamente, ocorre a ausência de pontos praticamente nos mesmos locais em que as densidades do atlas probabilístico são mais baixas. Ao final, os valores obtidos para $c(\mathbf{s})$ nesse atlas atingiram o valor máximo de $c(\mathbf{s}) = 231$.

5.5 Posicionamento e ajuste dos modelos

As subseções a seguir apresentam e discutem os resultados obtidos na inicialização (posicionamento e ajuste) dos modelos geométricos deformáveis em cada teste realizado.

5.5.1 Testes com imagens sinteticamente deformadas

Os resultados obtidos no teste preliminar usando imagens sinteticamente deformadas como imagens de entrada são apresentados nos gráficos da Figura 5.10. Para ambas as métricas de avaliação (JC e DC), os valores foram computados tanto antes da inicialização dos modelos (medindo-se a sobreposição entre as malhas de referência e as malhas deformadas antes de qualquer processamento) como após a inicialização, a fim de evidenciar o ganho que a inicialização trouxe nesses casos.

Nesse teste, com a abordagem desenvolvida neste trabalho, ocorreu um aumento significativo nos valores de ambas as métricas de avaliação, com a média dos valores de JC indo de

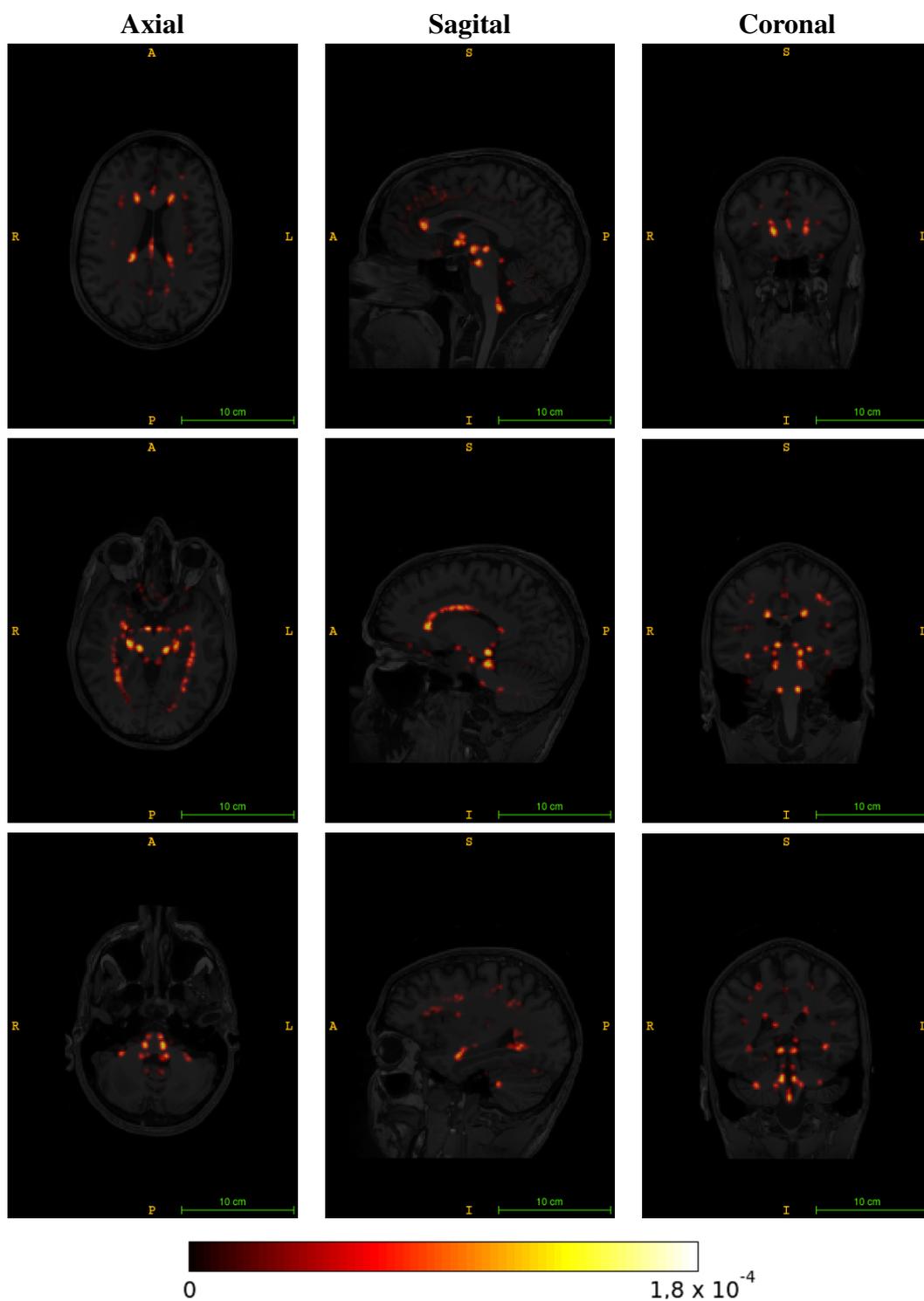


Figura 5.7: Sobreposição da densidade de distribuição de localização dos pontos salientes, estimada no atlas probabilístico, com a imagem de referência.

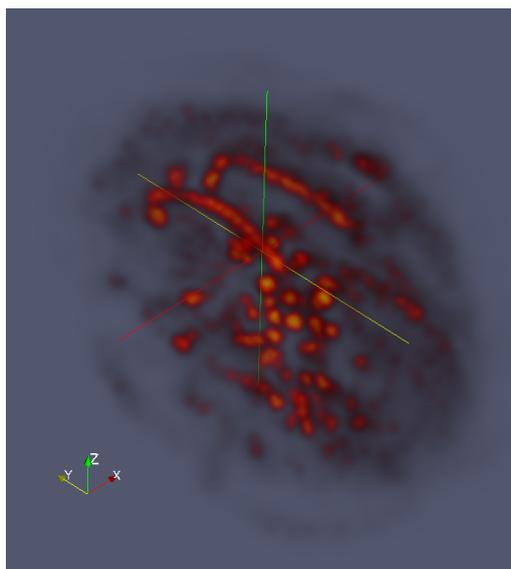


Figura 5.8: Renderização 3D da densidade de distribuição de localização dos pontos salientes, estimada no atlas probabilístico.

0,26 a 0,79 para a malha do CC, de 0,21 a 0,73 para a malha do LH, e de 0,19 a 0,73 para a malha do RH. Com a métrica DC, os valores médios aumentaram de 0,41 a 0,88 para a malha do CC, de 0,34 a 0,84 para a malha do LH, e de 0,31 a 0,84 para a malha do RH. Valores aparentemente promissores.

5.5.2 Testes com imagens clínicas

Os gráficos da Figura 5.11 resumem os resultados obtidos para os testes utilizando as imagens clínicas da base EADC-ADNI como imagens de entrada. Novamente, ambas as métricas de avaliação foram computadas tanto antes da inicialização dos modelos (logo após o corregristo afim que coloca a imagem e os modelos no mesmo espaço de coordenadas que o atlas) como após a inicialização. Desta vez, a avaliação foi feita tanto sem uso de atlas como com uso de cada um dos tipos de atlas, conforme descrito na Seção 4.8.

Novamente, a inicialização com a abordagem proposta neste trabalho trouxe um aumento em ambas as métricas de avaliação, mesmo no caso sem atlas. Porém, neste caso o aumento foi bem menos expressivo do que no teste anterior. Antes da inicialização dos modelos, os valores médios das métricas eram de $JC = 0,29$ e $DC = 0,43$ no LH, e de $JC = 0,28$ e $DC = 0,42$ no RH. Fazendo a inicialização sem uso de atlas, essas médias aumentaram para $JC = 0,35$ e $DC = 0,52$ no LH, e para $JC = 0,35$ e $DC = 0,51$ no RH. Com uso do atlas probabilístico, as médias aumentaram para $JC = 0,39$ e $DC = 0,55$ no LH, e para $JC = 0,40$ e $DC = 0,57$ no RH. Com uso do atlas de médias e frequências, por fim, as médias aumentaram para $JC = 0,37$

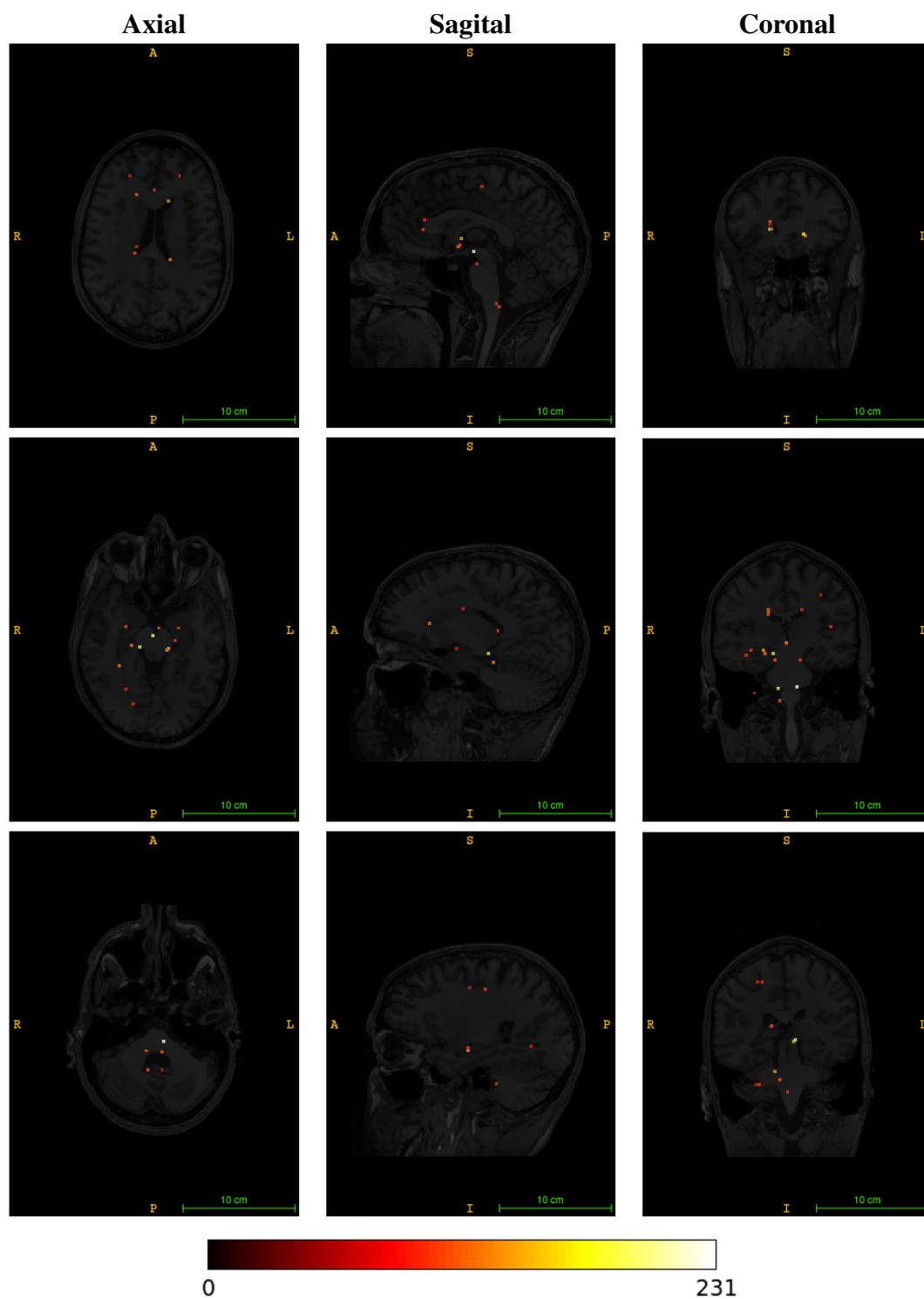


Figura 5.9: Sobreposição dos pontos de alta estabilidade, estimados no atlas de médias e frequências, com a imagem de referência.

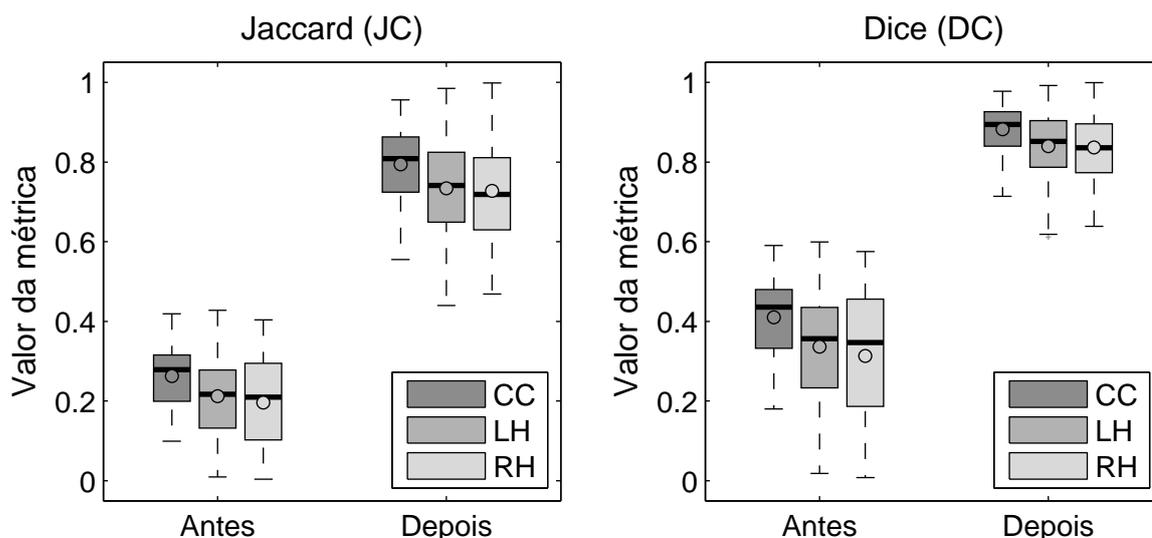


Figura 5.10: Valores das métricas JC e DC obtidos para cada uma das três estruturas cerebrais testadas (CC, LH e RH) antes e depois da inicialização dos modelos com a abordagem desenvolvida neste trabalho, para o teste envolvendo imagens sinteticamente deformadas como imagens de entrada. Os círculos indicam a média e as linhas mais espessas indicam a mediana.

e $DC = 0,54$ tanto no LH como no RH. Assim, as maiores médias, assim como os valores máximos ao longo de todo o teste, foram obtidos quando se usa o atlas probabilístico; mas a diferença em relação aos resultados obtidos sem uso de atlas ainda permaneceu muito pequena, de qualquer maneira.

5.5.3 Discussão

Duas questões importantes precisam ser discutidas: (i) a razão da discrepância entre os resultados obtidos com as imagens sinteticamente deformadas e aqueles obtidos com as imagens clínicas da base EADC-ADNI no caso sem atlas; e (ii) a razão pela qual o uso dos atlas não trouxe diferenças significativas nos resultados do teste com imagens clínicas.

Primeiramente, convém questionar se as imagens e modelos sinteticamente deformados estavam realmente deformados o suficiente para que a dificuldade em se inicializar precisamente um modelo, utilizando os pontos salientes, fosse comparável à dificuldade que se teria com uma imagem clínica de paciente com DA. Nesse sentido, um primeiro fator a se investigar seria a proximidade dos modelos sinteticamente deformados em relação às estruturas alvo nas imagens antes da inicialização. Afinal, se eles já estivessem muito próximos, então em tese seria mais fácil se chegar a um bom resultado a partir deles. No entanto, os valores obtidos para as métricas JC e DC antes da inicialização (Figuras 5.10 e 5.11) se mostraram ainda menores no teste com as imagens sinteticamente deformadas do que no teste com as imagens clínicas; ou seja, os modelos passaram por uma transformação maior no primeiro caso do que no último. Então

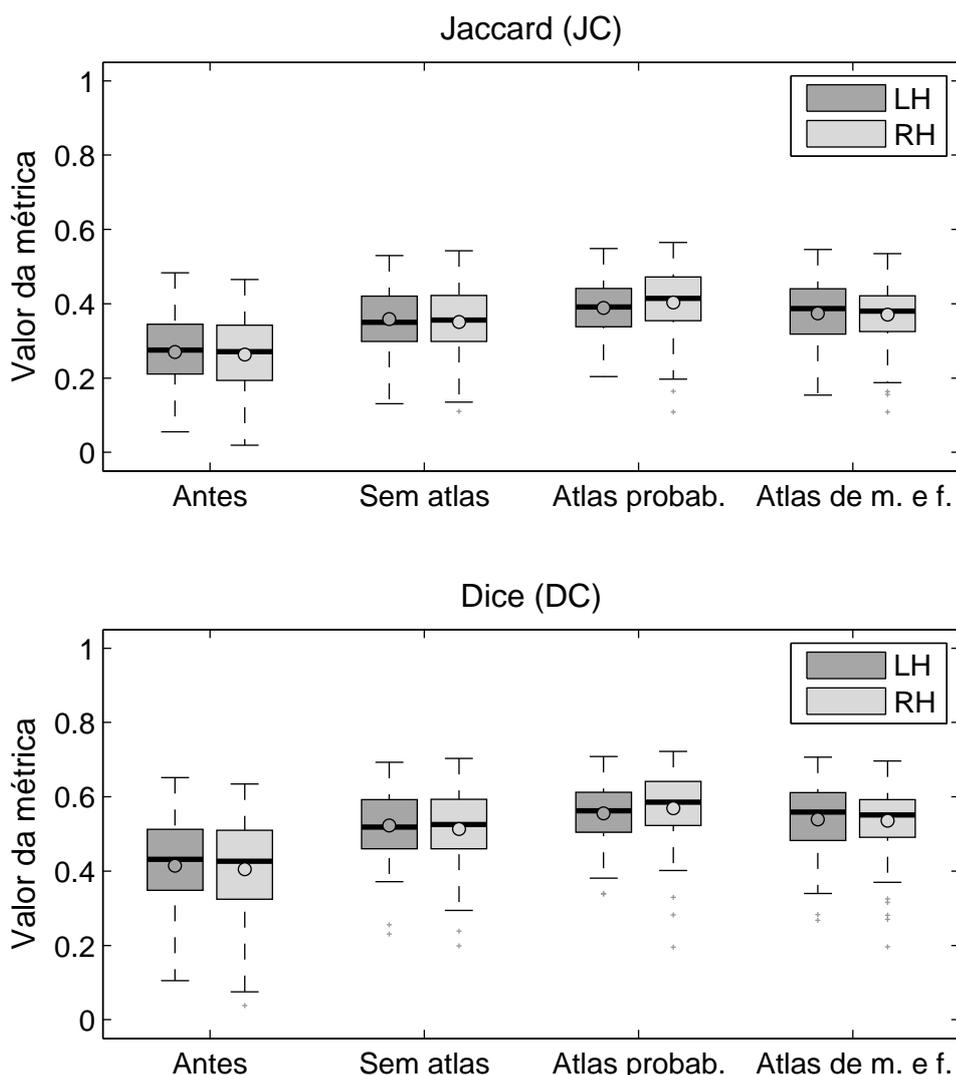


Figura 5.11: Valores das métricas JC e DC obtidos para cada uma das duas estruturas cerebrais testadas (LH e RH) antes e depois da inicialização dos modelos com a abordagem desenvolvida neste trabalho, para o teste envolvendo imagens clínicas da base EADC-ADNI como imagens de entrada. Os círculos indicam a média e as linhas horizontais mais espessas indicam a mediana.

o correto a se questionar não seria o quão distantes os modelos estavam das estruturas de interesse antes da inicialização, mas sim o quão finamente localizadas eram as variações anatômicas simuladas nas imagens sinteticamente deformadas em comparação às imagens clínicas de pacientes com DA. Não há como medir isso quantitativamente, mas pela própria maneira como as imagens sinteticamente deformadas foram geradas, parece claro que suas variações anatômicas locais são mais grosseiras se comparadas às de uma imagem clínica real de um paciente com atrofia cerebral. Logo, parece ser mais difícil se estimar uma transformação *B-spline* suficientemente precisa para captar bem todas as nuances anatômicas existentes entre uma imagem de referência e uma imagem clínica dessas - ainda mais quando os pacientes retratados em cada imagem são muito diferentes em faixa etária e patologia.

Esperava-se que o uso dos atlas pudesse amenizar o problema exposto no parágrafo anterior, já que eles capturam informações de uma população toda e não apenas de uma única imagem de referência. Porém, isso trouxe pouca diferença nos resultados. Embora seja possível observar que existe coerência entre os pontos salientes casados (isto é, que os pontos casados entre o atlas e a imagem clínica correspondem, na maioria dos casos, a pontos localizados em regiões muito similares anatomicamente), a transformação final, que ajusta o posicionamento e forma dos modelos, ainda não conseguiu aproximá-los o suficiente das estruturas de interesse nas imagens. A razão, mais uma vez, parece estar no fato de que as variações anatômicas locais nas imagens clínicas de pacientes com DA são mais finas do que a transformação estimada consegue mapear. Uma solução aparentemente óbvia para isso seria aumentar o número de pontos salientes detectados nas imagens (reduzindo o limiar T_λ da técnica de detecção), já que, em tese, quanto mais pontos salientes forem detectados e corretamente casados próximos a uma estrutura cerebral de interesse, mais chances o método tem de conseguir chegar a uma inicialização precisa para essa estrutura em particular. No entanto, fazer isso aumenta também a quantidade de falsos positivos na imagem toda (especialmente em imagens ruidosas), o que acaba prejudicando o desempenho do casamento e da estimação da transformação ao invés de trazer melhorias. Assim, uma solução que provavelmente funcionaria melhor seria embutir nos atlas algumas informações de referência anatômica dos pontos salientes detectados, a fim de que a transformação de inicialização possa dar mais peso aos pontos salientes que, além de serem mais estáveis ao longo da população, também tenham alta probabilidade de se localizarem próximos à estrutura de interesse na imagem alvo. As informações vindas dos atuais atlas expressam apenas a primeira dessas características, portanto não seria possível ter essa noção da anatomia a partir deles. Para se implementar essa nova estratégia, no entanto, seria necessário que rotulações de diversas estruturas do cérebro estivessem presentes em todo o conjunto de imagens de treinamento - algo que não se tinha durante esta pesquisa.

Adicionalmente, um último ponto a ser discutido é a questão da população de pacientes a partir da qual os atlas são computados. Idealmente, essa população precisa ser capaz de prover uma representação estatística tão precisa quanto possível das imagens clínicas que serão posteriormente testadas. Como o cérebro humano passa por mudanças estruturais devido ao envelhecimento e também em decorrência da DA, o ideal para este trabalho teria sido utilizar uma abordagem “multiatlas”, fazendo a estratificação dos atlas conforme as diferentes faixas etárias e graus de evolução da doença vistas na população. Dessa maneira, talvez o uso dos atlas também pudesse ter trazido melhorias mais significativas em comparação ao caso sem uso de atlas (que não teria como prover resultados muito melhores de qualquer maneira). O grande problema, no entanto, é que tal abordagem necessitaria um número muito maior de

imagens de treinamento (e a prévia estratificação delas), inclusive de pacientes com DA, e também precisaria de um mecanismo para seleção automática do atlas mais adequado a se utilizar de acordo com a similaridade em relação ao paciente retratado numa imagem clínica. Não havia como desenvolver tudo isso ainda no escopo deste trabalho de Mestrado, mas é uma possibilidade de trabalho futuro.

5.6 Considerações finais

Este capítulo apresentou e discutiu os resultados que foram obtidos utilizando a abordagem proposta e desenvolvida neste trabalho, no que diz respeito ao pré-processamento das imagens de RM, à detecção de pontos salientes 3D, à descrição e casamento desses pontos, à construção dos atlas de pontos salientes, e à inicialização de modelos geométricos deformáveis usando os pares de pontos casados como âncoras para guiar a transformação, com ou sem o auxílio do atlas nesse processo. Os experimentos acabaram mostrando que, embora a abordagem desenvolvida tenha melhorado marginalmente o posicionamento e forma dos modelos, os resultados ainda são preliminares e provavelmente não são suficientes para justificar o uso de tal abordagem em casos práticos (aplicações clínicas) envolvendo imagens de RM de pacientes com DA, muito embora tanto os testes das etapas anteriores (detecção, descrição, casamento) quanto os testes preliminares com imagens sinteticamente deformadas tenham tido bons resultados. Assim, mais pesquisas ainda precisam ser realizadas para se chegar a um método verdadeiramente eficaz em casos clínicos. Todavia, uma discussão sobre os possíveis direcionamentos de tais pesquisas foi feita e espera-se que ela possa servir para ajudar a orientar trabalhos futuros.

Capítulo 6

CONCLUSÕES

Este capítulo apresenta as conclusões finais do trabalho, enumerando suas contribuições, ideias para trabalhos futuros, e a produção científica no decorrer da pesquisa.

6.1 Conclusões do trabalho

Este trabalho propôs e desenvolveu uma abordagem para inicialização automática de modelos geométricos deformáveis com o objetivo de auxiliar a segmentação de estruturas cerebrais em imagens 3D de RM. A motivação por trás da pesquisa foi a crescente demanda por técnicas computacionais automáticas que proporcionem maior agilidade e confiabilidade à tarefa de análise de imagens de exames para o diagnóstico de doenças neurodegenerativas tais como a DA, na qual a atrofia de determinadas estruturas do cérebro é um importante biomarcador. Isso é necessário devido às limitações que existem na análise puramente visual e ao volume cada vez maior de imagens que necessitam ser analisadas pelos especialistas.

A abordagem proposta para alcançar esse objetivo envolveu a construção e uso de atlas de pontos salientes 3D. Os pontos salientes foram automaticamente detectados usando a medida de PC, por esta ser invariante às variações heterogêneas de intensidade causadas pelo *bias field* em imagens de RM. Descritores (assinaturas) baseados em *shape contexts* e na própria informação de PC foram utilizados para permitir que pontos salientes correspondentes pudessem ser identificados em imagens distintas. Dois tipos de atlas foram construídos a partir de uma mesma população de imagens: (i) um atlas probabilístico, obtido via estimação de uma função de densidade da localização dos pontos salientes na população, e (ii) um atlas de médias e frequências, obtido via casamentos sucessivos entre os pontos salientes da população e posterior eliminação dos pontos menos frequentes. Usando a informação dos descritores médios calculados nesses atlas, os pontos salientes detectados em uma nova imagem clínica poderiam encontrar pontos

correspondentes estatisticamente estáveis e, a partir dos casamentos com esses pontos, ser feita a estimação de uma transformação *B-spline* para mapear os modelos alinhados ao espaço do atlas para o espaço da imagem clínica, concluindo a inicialização tanto em posição quanto em ajuste de forma. Os modelos são representados por malhas 3D de estruturas cerebrais obtidas de uma base de uso livre.

Os resultados obtidos são preliminares e as técnicas desenvolvidas ainda precisam de uma melhor investigação para serem aplicáveis em casos clínicos reais. Embora as técnicas de detecção, descrição e casamento de pontos salientes tenham tido resultados quantitativos bastante interessantes nos testes realizados, e os atlas tenham obtido representações coerentes com o que se esperava, e bons resultados tenham sido obtidos até mesmo nos testes preliminares de inicialização de modelos (envolvendo imagens e modelos sinteticamente deformados), os testes finais, envolvendo imagens clínicas de pacientes idosos com DA, não produziram resultados suficientemente satisfatórios para tais casos - ainda que o foco do trabalho não estivesse na segmentação final das estruturas, e sim apenas em uma inicialização dos modelos para auxiliar tal segmentação. Todavia, foi feita uma discussão sobre os resultados atuais e o que se poderia fazer em pesquisas futuras para melhorá-los. Então acredita-se que o trabalho aqui desenvolvido sirva como um primeiro passo e como uma referência para orientar trabalhos futuros dentro do mesmo tema, para que enfim se consiga chegar a um método verdadeiramente aplicável em casos clínicos reais.

6.2 Contribuições

Com base no que foi realizado nesta pesquisa, pode-se dizer que este trabalho teve as seguintes contribuições principais:

- Desenvolvimento das seguintes inovações:
 - Metodologia para inicialização de modelos geométricos deformáveis em imagens 3D de RM usando atlas de pontos salientes 3D;
 - Método de descrição e casamento de pontos salientes 3D baseado na informação de PC, *shape contexts* e distância χ^2 ;
 - Dois métodos de criação de atlas de pontos salientes (sem necessidade de interação humana): um baseado em probabilidade de localização e outro baseado em casamentos sucessivos de pontos salientes similares;

- Refinamento de uma técnica invariante (e pouco afetada pelo *bias field*) para detecção de pontos salientes 3D, baseada em PC e momentos de segunda ordem;
- Criação de uma base de 1089 imagens 3D sintéticas que representam pequenas estruturas 3D de alta curvatura sob diferentes condições de degradação da imagem;
- Análises e discussões que podem ajudar a orientar trabalhos de pesquisa futuros para que estes melhorem os resultados obtidos até então.

6.3 Ideias para trabalhos futuros

Com base nas discussões que foram feitas sobre os resultados deste trabalho, as seguintes ideias são sugeridas para futuros trabalhos de pesquisa:

- Inclusão de informações de referenciação anatômica dos pontos salientes detectados nas imagens de treinamento, a fim de que também se possa, através dessa informação, dar mais relevância aos pontos que estiverem localizados próximos à estrutura de interesse no momento da inicialização dos modelos;
- Construção e uso de múltiplos atlas de pontos salientes, estratificados conforme a idade e a patologia dos pacientes, e desenvolvimento de um mecanismo de seleção automática do atlas que melhor se adapta às características do paciente em uma nova imagem clínica;
- Investigação de outras técnicas para a detecção dos pontos salientes, como por exemplo, o método multiescala baseado em *wavelets* de Fauqueur, Kingsbury e Anderson (2006) (que, diferentemente da PC, mostra-se robusto a ruído);
- Investigação de métodos iterativos de adaptação de modelos deformáveis que possam aproveitar as informações dos atlas (e da própria imagem alvo) na formulação de seus funcionais de energia, para assim realizar a segmentação final das estruturas de interesse após a inicialização dos modelos.

6.4 Produção científica relacionada a esta pesquisa

Até o momento de finalização desta dissertação, dois artigos científicos haviam sido escritos e submetidos:

- Um artigo já condicionalmente aceito na revista *Pattern Recognition Letters*, contendo resultados parciais obtidos na inicialização dos modelos do CC, LH e RH utilizando a abordagem proposta neste trabalho (ainda sem o uso de atlas);
- Um artigo ainda em revisão para a conferência *2016 IEEE International Conference on Image Processing (IEEE-ICIP)*, que apresenta um novo método para detecção automática do plano sagital médio do cérebro em imagens de RM utilizando PC, momentos de segunda ordem e uma medida de planaridade. Tal artigo está indiretamente relacionado a este trabalho porque surgiu da observação de que a medida de PC produzia respostas particularmente altas exatamente na localização do plano sagital médio.

REFERÊNCIAS

AGRAWAL, R.; SHARMA, M. Review of segmentation methods for brain tissue with magnetic resonance images. *International Journal of Computer Network and Information Security*, v. 6, n. 4, p. 55–62, 2014.

ALZHEIMER'S DISEASE EDUCATION AND REFERRAL CENTER. *Alzheimer's Disease Fact Sheet*. Julho 2011. Disponível em: <<http://www.nia.nih.gov/alzheimers/publication/alzheimers-disease-fact-sheet>>.

ANBECK, P.; VINCKEN, K. L.; OSCH, M. J. P. van; BISSCHOPS, R. H. C.; GROND, J. van der. Automatic segmentation of different-sized white matter lesions by voxel probability estimation. *Medical Image Analysis*, v. 8, n. 3, p. 205–215, 2004.

APOSTOLOVA, L. G.; DUTTON, R. A.; DINOV, I. D.; HAYASHI, K. M.; TOGA, A. W.; CUMMINGS, J. L.; THOMPSON, P. M. Conversion of mild cognitive impairment to Alzheimer disease predicted by hippocampal atrophy maps. *Archives of Neurology*, v. 63, n. 5, p. 693–699, 2006.

APRAHAMIAN, I.; MARTINELLI, J. E.; YASSUDA, M. S. Doença de Alzheimer: revisão da epidemiologia e diagnóstico. *Revista da Sociedade Brasileira de Clínica Médica*, v. 7, n. 1, p. 27–35, 2009.

ASHBURNER, J.; FRISTON, K. J. Unified segmentation. *Neuroimage*, v. 26, n. 3, p. 839–851, 2005.

AUBERT-BROCHE, B.; GRIFFIN, M.; PIKE, G.; EVANS, A.; COLLINS, D. Twenty new digital brain phantoms for creation of validation image databases. *IEEE Transactions on Medical Imaging*, v. 36, n. 3, p. 494–497, 1996.

BALAFAR, M. A.; RAMLI, A.; SARIPAN, M. I.; MASHOHOR, S. Review of brain MRI image segmentation methods. *Artificial Intelligence Review*, v. 33, n. 3, p. 261–274, 2010.

BALLARD, D. H. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, v. 13, n. 2, p. 111–122, 1981.

BEAUDET, P. Rotationally invariant image operators. In: *Proceedings of the 4th International Joint Conference of Pattern Recognition*. Kyoto, Japan: [s.n.], 1978. p. 579–583.

BELONGIE, S.; MALIK, J.; PUZICHA, J. Shape matching and object recognition using shape contexts. *Pattern Analysis and Machine*, v. 24, n. 4, p. 509–522, 2002.

- BLENNOW, K.; LEON, M. J. de; ZETTERBERG, H. Alzheimer's disease. *The Lancet*, v. 368, n. 9533, p. 387–403, 2006.
- BLOCH, F.; HANSON, W. W.; PACKARD, M. Nuclear induction. *Physical Review*, v. 69, p. 127, 1946.
- BLOM, J. *Topological and Geometrical Aspects of Image Structure*. Tese (Doutorado) — Utrecht University, Netherlands, 1992.
- BOCCARDI, M.; BOCCHETTA, M.; MORENCY, F. C.; COLLINS, D. L.; NISHIKAWA, M.; GANZOLA, R.; GROTHE, M. J.; WOLF, D.; REDOLFI, A.; PIEVANI, M.; ANTELM, L.; FELLGIEBEL, A.; MATSUDA, H.; TEIPEL, S.; DUCHESNE, S.; JR, C. R. J.; FRISONI, G. B. Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, v. 11, n. 2, p. 175–183, 2015.
- BOGOVIC, J. A.; PRINCE, J. L.; BAZIN, P.-L. A multiple object geometric deformable model for image segmentation. *Computer Vision and Image Understanding (CVIU)*, v. 117, n. 2, p. 145–157, 2013.
- BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001.
- BROOKMEYER, R.; JOHNSON, E.; ZIEGLER-GRAHAM, K.; ARRIGHI, H. M. Forecasting the global burden of Alzheimer's disease. *Alzheimer's & Dementia*, v. 3, n. 3, p. 186–191, 2007.
- BROWN, M. A.; SEMELKA, R. C. *MRI Basic Principles and Applications*. 3a. ed. [S.l.]: John Wiley & Sons, Inc., 2003.
- BUADES, A.; COLL, B.; MOREL, J. M. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation*, v. 4, n. 2, p. 490–530, 2005.
- BUSHBERG, J. T.; SEIBERT, J. A.; JR., E. M. L.; BOONE, J. M. *Essential Physics of Medical Imaging*. 3a. ed. [S.l.]: Lippincott Williams & Wilkins, 2011.
- CABEZAS, M.; OLIVER, A.; LLADÓ, X.; FREIXENET, J.; CUADRA, M. B. A review of atlas-based segmentation for magnetic resonance brain images. *Computer Methods and Programs in Biomedicine*, v. 104, n. 3, p. e158–e177, 2011.
- CARMICHAEL, O. T.; AIZENSTEIN, H. A.; DAVIS, S. W.; BECKER, J. T.; THOMPSON, P. M.; MELTZER, C. C.; LIU, Y. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *NeuroImage*, v. 27, n. 4, p. 979–990, 2005.
- CHANDRA, S. S.; DOWLING, J. A.; SHEN, K.-K.; RANIGA, P.; PLUIM, J. P. W.; GREER, P. B.; SALVADO, O.; FRIPP, J. Patient specific prostate segmentation in 3-D magnetic resonance images. *IEEE Transactions on Medical Imaging*, v. 31, n. 10, p. 1955–1964, 2012.
- CHEN, S. Y.; ZHANG, J.; GUAN, Q.; LIU, S. Detection and amendment of shape distortions based on moment invariants for active shape models. *Image Processing, IET*, v. 5, n. 3, p. 273–285, 2011.
- CHEUNG, W.; HAMARNEH, G. n-SIFT: n-dimensional scale invariant feature transform for matching medical images. In: *Proceedings of the Fourth IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2007)*. Metro Washington, DC, USA: [s.n.], 2007. p. 720–723.

- CHUI, H.; RANGARAJAN, A. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, Elsevier Science Inc., New York, NY, USA, v. 89, n. 2-3, p. 114–141, 2003.
- CHUPIN, M.; MUKUNA-BANTUMBAKULU, A. R.; HASBOUN, D.; BARDINET, E.; BAILLET, S.; KINKINGNÉHUN, S.; LEMIEUX, L.; DUBOIS, B.; GARNERO, L. Automatic constrained region deformation for the automated segmentation of the hippocampus and the amygdala: Method and validation on controls and patients with Alzheimer's disease. *NeuroImage*, v. 34, n. 3, p. 996–1019, 2007.
- CLARKSON, M. J.; OURSELIN, S.; NIELSEN, C.; LEUNG, K. K.; BARNES, J.; WHITWELL, J. L.; GUNTER, J. L.; HILL, D. L. G.; WEINER, M. W.; JR., C. R. J.; FOX, N. C.; The Alzheimer's Disease Neuroimaging Initiative. Comparison of phantom and registration scaling corrections using the ADNI cohort. *NeuroImage*, v. 47, n. 4, p. 1506–1513, 2009.
- COLLINS, D. L.; NEELIN, P.; PETERS, T. M.; EVANS, A. C. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, v. 18, n. 2, p. 192–205, 1994.
- CRIMINISI, A.; ROBERTSON, D.; KONUKOGLU, E.; SHOTTON, J.; PATHAK, S.; WHITE, S.; SIDDIQUI, K. Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical Image Analysis*, v. 17, n. 8, p. 1293–1303, 2013.
- DELEENER, B.; KADOURY, S.; COHEN-ADAD, J. Robust, accurate and fast automatic segmentation of the spinal cord. *Neuroimage*, v. 98, p. 528–536, 2014.
- DEVROYE, L. *Non-Uniform Random Variate Generation*. [S.l.]: Prentice Hall, 1986.
- DOSIL, R.; PARDO, X. M.; FDEZ-VIDAL, X. R. Decomposition of three-dimensional medical images into visual patterns. *IEEE Transactions on Biomedical Engineering*, v. 52, n. 12, p. 2115–2118, 2005.
- DUBOIS, B.; FELDMAN, H. H.; JACOVA, C.; CUMMINGS, J. L.; DEKOSKY, S. T.; BARBERGER-GATEAU, P.; DELACOURTE, A.; FRISONI, G.; FOX, N. C.; GALASKO, D.; GAUTHIER, S.; HAMPEL, H.; JICHA, G. A.; MEGURO, K.; O'BRIEN, J.; PASQUIER, F.; ROBERT, P.; ROSSOR, M.; SALLOWAY, S.; SARAZIN, M.; SOUZA, L. C. de; STERN, Y.; VISSER, P. J.; SCHELTENS, P. Revising the definition of Alzheimer's disease: a new lexicon. *The Lancet Neurology*, v. 9, n. 11, p. 1118 – 1127, 2010.
- DUDA, R. O.; HART, P. E.; STORK, D. G. Pattern classification. In: _____. 2. ed. New York, NY: Wiley-Interscience, 2000. cap. 4, p. 161–174.
- ECABERT, O.; PETERS, J.; SCHRAMM, H.; LORENZ, C.; BERG, J. von; WALKER, M. J.; VEMBAR, M.; OLSZEWSKI, M. E.; SUBRAMANYAN, K.; LAVI, G.; WEESE, J. Automatic model-based segmentation of the heart in CT images. *IEEE Transactions on Medical Imaging*, v. 27, n. 9, p. 1189–1201, 2008.
- EVANS, A. C.; COLLINS, D. L.; MILLS, S. R.; BROWN, E. D.; KELLY, R. L.; PETERS, T. M. 3d statistical neuroanatomical models from 305 mri volumes. In: *1993 IEEE Nuclear Science Symposium and Medical Imaging Conference*. [S.l.: s.n.], 1993. v. 3, p. 1813–1817.

- FAUQUEUR, J.; KINGSBURY, N.; ANDERSON, R. Multiscale keypoint detection using the dual-tree complex wavelet transform. In: *2006 IEEE International Conference on Image Processing*. [S.l.: s.n.], 2006. p. 1625–1628.
- FERRARI, R. J.; ALLAIRE, S.; HOPE, A.; KIM, J.; JAFFRAY, D.; PEKAR, V. Detection of point landmarks in 3D medical images via phase congruency model. *Journal of the Brazilian Computer Society*, v. 17, n. 2, p. 117–132, 2011.
- FLIER, W. M. van der; SCHELTENS, P. Epidemiology and risk factors of dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, v. 76, n. 5, p. v2–v7, 2005.
- FLORACK, L. M. J.; ROMENY, B. M. ter H.; KOENDERINK, J. J.; VIERGEVER, M. A. General intensity transformations and differential invariants. *Journal of Mathematical Imaging and Vision*, v. 4, n. 2, p. 171–187, 1994.
- FÖRSTNER, W.; GÜLCH, E. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: *Proceedings of the ISPRS Intercommission Workshop*. Interlaken, Switzerland: [s.n.], 1987. p. 149–155.
- FRIPP, J.; CROZIER, S.; WARFIELD, S.; OURSELIN, S. Automatic initialization of 3D deformable models for cartilage segmentation. In: *Proceedings of the International Conference on Digital Image Computing: Techniques and Applications*. Cairns, Australia: [s.n.], 2005. p. 74–74.
- FRISONI, G. B.; JR., C. R. J.; BOCCHETTA, M.; BAUER, C.; FREDERIKSEN, K. S.; LIU, Y.; PREBOSKE, G.; SWIHART, T.; BLAIR, M.; CAVEDO, E.; GROTHE, M. J.; LANFREDI, M.; MARTINEZ, O.; NISHIKAWA, M.; PORTEGIES, M.; STOUB, T.; WARD, C.; APOSTOLOVA, L. G.; GANZOLA, R.; WOLF, D.; BARKHOF, F.; BARTZOKIS, G.; DECARLI, C.; CSERNANSKY, J. G.; MORRELL, L. deToledo; GEERLINGS, M. I.; KAYE, J.; KILLIANY, R. J.; LEHÉRICY, S.; MATSUDA, H.; O'BRIEN, J.; SILBERT, L. C.; SCHELTENS, P.; SOININEN, H.; TEIPEL, S.; WALDEMAR, G.; FELLGIEBEL, A.; BARNES, J.; FIRBANK, M.; GERRITSEN, L.; HENNEMAN, W.; MALYKHIN, N.; PRUESSNER, J. C.; WANG, L.; WATSON, C.; WOLF, H.; DELEON, M.; PANTEL, J.; FERRARI, C.; BOSCO, P.; PASQUALETTI, P.; DUCHESNE, S.; DUVERNOY, H.; BOCCARDI, M. The EADC-ADNI harmonized protocol for manual hippocampal segmentation on magnetic resonance: Evidence of validity. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, v. 11, n. 2, p. 111–125, 2015.
- GALL, J.; YAO, A.; RAZAVI, N.; GOOL, L. V.; LEMPITSKY, V. Hough forests for object detection, tracking, and action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 33, n. 11, p. 2188–2202, 2011.
- GHOSE, S.; MITRA, J.; OLIVER, A.; MARTI, R.; LLADO, X.; FREIXENET, J.; VILANOVA, J. C.; SIDIBE, D.; MERIAUDEAU, F. A coupled schema of probabilistic atlas and statistical shape and appearance model for 3D prostate segmentation in MR images. In: *19th IEEE International Conference on Image Processing (ICIP)*. Orlando, FL: IEEE, 2012. p. 541–544.
- GISPERT, J. D.; REIG, S.; PASCAU, J.; VAQUERO, J. J.; GARCÍA-BARRENO, P.; DESCO, M. Method for bias field correction of brain T1-weighted magnetic resonance images minimizing segmentation error. *Human Brain Mapping*, v. 22, n. 2, p. 133–144, 2004.

GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 3a. ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.

GOSSUIN, Y.; HOCQ, A.; GILLIS, P.; VUONG, Q. L. Physics of magnetic resonance imaging: from spin to pixel. *Journal of Physics D: Applied Physics*, v. 43, n. 21, p. 213001, 2010.

GUDBJARTSSON, H.; PATZ, S. The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, v. 34, n. 6, p. 910–914, 1995.

HALLE, M.; TALOS, I.-F.; JAKAB, M.; MAKRIS, N.; MEIER, D.; WALD, L.; FISCHL, B.; KIKINIS, R. *Multi-modality MRI-based Atlas of the Brain*. 2013. Disponível em: <<http://nac.spl.harvard.edu/publications/item/view/2037>>.

HAN, D.; GAO, Y.; WU, G.; YAP, P. T.; SHEN, D. Robust anatomical landmark detection for MR brain image registration. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, v. 17, n. 1, p. 186–193, 2014.

HARTKENS, T.; ROHR, K.; STIEHL, H. S. Evaluation of 3D operators for the detection of anatomical point landmarks in MR and CT images. *Computer Vision and Image Understanding*, v. 86, n. 2, p. 118–136, 2002.

HEGADI, R.; KOP, A.; HANGARGE, M. A survey on deformable model and its applications to medical imaging. *IJCA Special Issue on 'Recent Trends in Image Processing and Pattern Recognition' (RTIPPR)*, n. 2, p. 64–75, 2010.

HEIMANN, T.; MEINZER, H.-P. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, v. 13, n. 4, p. 543–563, 2009.

HEIMANN, T.; MÜNZING, S.; MEINZER, H.-P.; WOLF, I. A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. In: KARSSEMEIJER, N.; LELIEVELDT, B. (Ed.). *Information Processing in Medical Imaging*. Berlin, Germany: Springer Berlin Heidelberg, 2007, (Lecture Notes in Computer Science, v. 4584). p. 1–12.

IGLESIAS, J. E.; LIU, C. Y.; THOMPSON, P.; TU, Z. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Transactions on Medical Imaging*, v. 30, n. 9, p. 1617–1634, 2011.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Projeção da população do Brasil e das Unidades da Federação*. Nov 2014. Disponível em: <<http://www.ibge.gov.br/apps/populacao/projecao/index.html>>.

JUNTU, J.; SIJBERS, J.; DYCK, D. V.; GIELEN, J. Bias field correction for MRI images. In: _____. Berlin, Germany: Springer Berlin Heidelberg, 2005. (Advances in Soft Computing, v. 30), cap. 4, p. 543–551.

KALINIC, H. *Atlas-based image segmentation: A Survey*. Zagreb, Croatia, 2008. Disponível em: <<https://bib.irb.hr/datoteka/435355.jnrl.pdf>>.

KAUS, M. R.; BROCK, K. K. Biomechanical System Technology: Computational Methods. In: _____. Los Angeles, USA: World Scientific Publishing, 2009. (Biomechanical Systems Technology, v. 1), cap. Deformable Image Registration for radiation therapy planning, p. 1–28.

- KELLER, S. S.; ROBERTS, N. Measurement of brain volume using MRI: software, techniques, choices and prerequisites. *Journal of Anthropological Sciences*, v. 87, p. 127–151, 2009.
- KINDLMANN, G. L. *An self-contained explanation of image orientation and the 'measurement frame', with connections to the NRRD format*. Mar 2010. Disponível em: <<http://people.cs.uchicago.edu/~glk/unlinked/nrrd-iomf.pdf>>.
- KITCHEN, L.; ROSENFELD, A. Gray-level corner detection. *Pattern Recognition Letters*, v. 1, n. 2, p. 95–102, 1982.
- KLAUSCHEN, F.; GOLDMAN, A.; BARRA, V.; MEYER-LINDENBERG, A.; LUNDERVOLD, A. Evaluation of automated brain MR image segmentation and volumetry methods. *Human Brain Mapping*, v. 30, n. 4, p. 1310–1327, 2009.
- KOREZ, R.; IBRAGIMOV, B. A.; LIKAR, B.; PERNUS, F.; VRTOVEC, T. A framework for automated spine and vertebrae interpolation-based detection and model-based segmentation. *IEEE Transactions on Medical Imaging*, v. 34, n. 8, p. 1649–1662, 2015.
- KOVESI, P. Phase congruency: A low-level image invariant. *Psychological Research*, v. 64, n. 2, p. 136–148, 2000.
- KOVESI, P. Phase congruency detects corners and edges. In: *The Australian Pattern Recognition Society Conference: DICTA 2003*. [S.l.: s.n.], 2003. p. 309–318.
- LAI, K. F. *Deformable contours: modeling, extraction, detection and classification*. Tese (Doutorado) — Department of Electrical Engineering, University of Wisconsin-Madison, Madison, WI, USA, 1994.
- LAMECKER, H.; LANGE, T.; SEEBASS, M. *Segmentation of the liver using a 3D statistical shape model*. [S.l.], 2004.
- LEAVENS, C.; VIK, T.; SCHULZ, H.; ALLAIRE, S.; KIM, J.; DAWSON, L.; O'SULLIVAN, B.; BREEN, S.; JAFFRAY, D.; PEKAR, V. Validation of automatic landmark identification for atlas-based segmentation for radiation treatment planning of the head-and-neck region. In: *Proceedings of the SPIE Conference on Medical Imaging*. San Diego, CA, USA: [s.n.], 2008. v. 6914, p. 3G1–3G7.
- LEE, S.; WOLBERG, G.; SHIN, S. Y. Scattered data interpolation with multilevel B-splines. *IEEE Transactions on Visualization and Computer Graphics*, v. 3, n. 3, p. 228–244, 1997.
- LEEMPUT, K. van. Encoding probabilistic brain atlases using bayesian inference. *IEEE Transactions on Medical Imaging*, v. 28, n. 6, p. 822–837, 2009.
- LIJN, F. van der; HEIJER, T. den; BRETILER, M. M. B.; NIESSEN, W. J. Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. *NeuroImage*, v. 43, n. 4, p. 708–720, 2008.
- LIN, J. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, v. 37, n. 1, p. 145–151, 1991.
- LINGURARU, M. G.; VERCAUTEREN, T.; REYES-AGUIRRE, M.; BALLESTER, M. A. G.; AYACHE, N. Segmentation propagation from deformable atlases for brain mapping and analysis. *Brain Research Journal*, v. 1, n. 4, p. 1935–2875, 2007.

- LIU, C.-Y.; IGLESIAS, J. E.; TU, Z. Deformable templates guided discriminative models for robust 3D brain MRI segmentation. *NeuroInformatics*, v. 11, n. 4, p. 447–468, 2013.
- LIU, J.; GAO, W.; HUANG, S.; NOWINSKI, W. L. A model-based, semi-global segmentation approach for automatic 3-D point landmark localization in neuroimages. *IEEE Transactions on Medical Imaging*, v. 27, n. 8, p. 1034–1044, 2008.
- LIU, J.; HUANG, S.; AZIZ, A.; NOWINSKI, W. L. Three dimensional digital atlas of the orbit constructed from multi-modal radiological images. *International Journal of Computer Assisted Radiology and Surgery*, Springer-Verlag, v. 1, n. 5, p. 275–283, 2007.
- LIU, J.; LI, M.; WANG, J.; WU, F.; LIU, T.; PAN, Y. A survey of MRI-based brain tumor segmentation methods. *Tsinghua Science and Technology*, v. 19, n. 6, p. 578–595, 2014.
- LIU, J.; NOWINSKI, W. L. A hybrid approach to shape-based interpolation of stereotactic atlases of the human brain. *NeuroInformatics*, Humana Press, v. 4, n. 2, p. 177–198, 2006.
- LO, C. H.; DON, H. S. 3D moment forms: Their construction and application to object identification and positioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 11, n. 10, p. 1053–1064, 1989.
- LOTJONEN, J.; WOLZ, R.; KOIKKALAINEN, J.; THURFJELL, L.; LUNDQVIST, R.; WALDEMAR, G.; SOININEN, H.; RUECKERT, D. Improved generation of probabilistic atlases for the expectation maximization classification. In: *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. Chicago, IL, USA: [s.n.], 2011. p. 1839–1842.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 60, n. 2, p. 91–110, 2004.
- MACOVSKI, A. Noise in MRI. *Magnetic Resonance in Medicine*, v. 36, n. 3, p. 494–497, 1996.
- MAZZOLA, A. A. Ressonância magnética: princípios de formação da imagem e aplicações em imagem funcional. *Revista Brasileira de Física Médica*, v. 3, n. 1, p. 117–129, 2009.
- MCINERNEY, T.; TERZOPOULOS, D. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, v. 1, n. 2, p. 91–108, 1996.
- MCINERNEY, T.; TERZOPOULOS, D. Topology adaptive deformable surfaces for medical image volume segmentation. *IEEE Transactions on Medical Imaging*, v. 18, n. 10, p. 840–850, 1999.
- MITCHELL, T. M. *Machine Learning*. 1a. ed. New York, NY, USA: McGraw-Hill Science/Engineering/Math, 1997. 432 p. (Mcgraw-Hill Series in Computer Science).
- MONDAL, P.; MUKHOPADHYAY, J.; SURAL, S.; BHATTACHARYYA, P. P. An efficient model-guided framework for alignment of brain MR image sequences. In: *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. Seoul, South Korea: [s.n.], 2012. p. 2201–2206.
- MONDAL, P.; MUKHOPADHYAY, J.; SURAL, S.; BHATTACHARYYA, P. P. 3D-SIFT feature based brain atlas generation: An application to early diagnosis of Alzheimer’s disease. In: *2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*. Greater Noida, India: [s.n.], 2014. p. 342–347.

- MONTAGNAT, J.; DELINGETTE, H.; AYACHE, N. A review of deformable surfaces: topology, geometry and deformation. *Image and Vision Computing*, v. 19, n. 14, p. 1023–1040, 2001.
- MORISHIMA-KAWASHIMA, M.; IHARA, Y. Alzheimer's disease: beta-amyloid protein and tau. *Journal of Neuroscience Research*, v. 70, n. 3, p. 392–401, 2002.
- MORRIS, J. C. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*, v. 43, n. 11, p. 2412–2414, 1993.
- MORRONE, M. C.; ROSS, J.; BURR, D. C.; OWENS, R. Mach bands are phase dependent. *Nature*, v. 324, n. 6094, p. 250–253, 1986.
- MORTENSEN, E. N.; DENG, H.; SHAPIRO, L. A SIFT descriptor with global context. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, CA, USA: [s.n.], 2005. p. 184–190.
- MURTY, K. G. *Linear Programming*. 1a. ed. [S.l.]: John Wiley & Sons Inc., 1983.
- NYUL, L. G.; UDUPA, J. K.; ZHANG, X. New variants of a method of MRI scale standardization. *IEEE Transactions on Medical Imaging*, v. 19, n. 2, p. 143–150, 2000.
- OURSELIN, S.; ROCHE, A.; SUBSOL, G.; PENNEC, X.; AYACHE, N. Reconstructing a 3D structure from serial histological sections. *Image and Vision Computing*, v. 19, n. 1-2, p. 25–31, 2000.
- OURSELIN, S.; STEFANESCU, R.; PENNEC, X. Robust registration of multi-modal images: Towards real-time clinical applications. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2002)*. [S.l.]: Springer Berlin Heidelberg, 2002. p. 140–147.
- PEKAR, V.; MCNUTT, T. R.; KAUS, M. R. Automated model-based organ delineation for radiotherapy planning in prostatic region. *International Journal of Radiation Oncology Biology Physics*, v. 60, n. 3, p. 973–980, 2004.
- PERONA, P.; MALIK, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 12, n. 7, p. 629–639, 1990.
- POWERS, D. M. W. Evaluation: from precision, recall and F-factor to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, v. 2, n. 1, p. 37–63, 2011.
- PRESS, W. H.; FLANNERY, B. P.; TEUKOLSKY, S. A.; VETTERLING, W. T. *Numerical Recipes in C*. Cambridge, UK: University Press, 1993.
- PURCELL, E. M.; TORREY, H. C.; POUND, R. V. Resonance absorption by nuclear magnetic moments in a solid. *Physical Review*, v. 69, p. 37–38, 1946.
- QAZI, A. A.; PEKAR, V.; KIM, J.; XIE, J.; BREEN, S. L.; JAFFRAY, D. A. Auto-segmentation of normal and target structures in head and neck CT images: A feature-driven model-based approach. *Medical Physics*, v. 38, n. 11, p. 6160–6170, 2011.
- REX, D. E.; MA, J. Q.; TOGAA, A. W. The LONI pipeline processing environment. *NeuroImage*, v. 19, n. 3, p. 1033–1048, 2003.

- ROHR, K. Modelling and identification of characteristic intensity variations. *Image and Vision Computing*, v. 10, n. 2, p. 66–76, 1992.
- ROHR, K. On 3D differential operators for detecting point landmarks. *Image and Vision Computing*, v. 15, n. 3, p. 219–233, 1997.
- RUPPERTSHOFEN, H.; LORENZ, C.; ROSE, G.; SCHRAMM, H. Discriminative generalized hough transform for object localization in medical images. *International Journal of Computer Assisted Radiology and Surgery*, Springer Berlin Heidelberg, v. 8, n. 4, p. 593–606, 2013.
- SAALBACH, A.; WÄCHTER-STEHLE, I.; KNESER, R.; MOLLUS, S.; PETERS, J.; WE-ESE, J. Optimizing GHT-based heart localization in an automatic segmentation chain. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, v. 14, n. 3, p. 463–470, 2011.
- SHATTUCK, D. W.; MIRZA, M.; ADISETIYO, V.; HOJATKASHANI, C.; SALAMON, G.; NARR, K. L.; POLDRACK, R. A.; BILDER, R. M.; TOGAA, A. W. Construction of a 3D probabilistic atlas of human cortical structures. *NeuroImage*, v. 39, n. 3, p. 1064–1080, 2008.
- SHATTUCK, D. W.; SANDOR-LEAHY, S. R.; SCHAPER, K. A.; ROTTENBERG, D. A.; LEAHY, R. M. Magnetic resonance image tissue classification using a partial volume model. *Neuroimage*, v. 13, n. 5, p. 856–876, 2001.
- SHEN, D.; DAVATZIKOS, C. HAMMER: hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging*, v. 21, n. 11, p. 1421–1439, 2002.
- SHEN, K. K.; FRIPP, J.; MÉRIAUDEAU, F.; CHÉTELAT, G.; SALVADO, O.; BOURGEAT, P. Detecting global and local hippocampal shape changes in Alzheimer’s disease using statistical shape models. *NeuroImage*, v. 59, n. 3, p. 2155–2166, 2012.
- SHOJAEILANGARIA, S.; YAUB, W.-Y.; TEOHA, E.-K. A novel phase congruency based descriptor for dynamic facial expression analysis. *Pattern Recognition Letters*, v. 49, p. 55–61, 2014.
- SLED, J. G.; ZIJDENBOS, A. P.; EVANS, A. C. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, v. 17, n. 1, p. 87 – 97, 1998.
- SLICERWIKI. *Coordinate systems*. Set 2014. Disponível em: <http://www.slicer.org/slicerWiki/index.php/Coordinate_systems>.
- SMITH, S. M. Fast robust automated brain extraction. *Human Brain Mapping*, Wiley Subscription Services, Inc., A Wiley Company, v. 17, n. 3, p. 143–155, 2002.
- SMITH, S. M.; JENKINSON, M.; WOOLRICH, M. W.; BECKMANN, C. F.; BEHRENS, T. E.; JOHANSEN-BERG, H.; BANNISTER, P. R.; LUCA, M. D.; DROBNJAK, I.; FLITNEY, D. E.; NIAZY, R. K.; SAUNDERS, J.; VICKERS, J.; ZHANG, Y.; STEFANO, N. D.; BRADY, J. M.; MATTHEWS, P. M. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, v. 23, Supplement 1, p. S208–S219, 2004.
- SOLER, L.; DELINGETTE, H.; MALANDAIN, G.; MONTAGNAT, J.; AYACHE, N.; KOEHL, C.; DOURTHE, O.; MALASSAGNE, B.; SMITH, M.; MUTTER, D.; MARESCAUX, J. Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surgery*, v. 6, n. 3, p. 131–142, 2001.

TALAIRACH, J.; TOURNOUX, P. *Co-planar stereotaxic Atlas of the Human Brain*. [S.l.]: Thieme Medical Publishers, 1988.

TEIPEL, S. J.; BAYER, W.; ALEXANDER, G. E.; BOKDE, A. L. W.; ZEBUHR, Y.; TEICHBURG, D.; MÜLLER-SPAHN, F.; SCHAPIRO, M. B.; MÖLLER, H. J.; RAPOPORT, S. I.; HAMPEL, H. Regional pattern of hippocampus and corpus callosum atrophy in Alzheimer's disease in relation to dementia severity: evidence for early neocortical degeneration. *Neurobiology of Aging*, v. 24, n. 1, p. 85–94, 2002.

TUSTISON, N. J.; AVANTS, B. B.; COOK, P. A.; ZHENG, Y.; EGAN, A.; YUSHKEVICH, P. A.; GEE, J. C. N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, v. 29, n. 6, p. 1310–1320, 2010.

TUSTISON, N. J.; GEE, J. C. Generalized n-D Ck b-spline scattered data approximation with confidence values. In: *Medical Imaging and Augmented Reality*. [S.l.]: Springer Berlin Heidelberg, 2006, (Lecture Notes in Computer Science, v. 4091). p. 76–83.

TUYTELAARS, T.; MIKOLAJCZYK, K. Local invariant feature detectors: A survey. *Foundations and Trends in Computer Graphics and Vision*, Now Publishers Inc., Hanover, MA, USA, v. 3, n. 3, p. 177–280, 2008.

VENKATESH, S.; OWENS, R. On the classification of image features. *Pattern Recognition Letters*, v. 11, n. 5, p. 339–349, 1990.

VINCENT, P.; BENGIO, Y. Manifold parzen windows. In: *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2003. v. 15.

WANG, F.; VEMURI, B. C.; RANGARAJAN, A.; EISENSCHENK, S. J. Simultaneous non-rigid registration of multiple point sets and atlas construction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 30, n. 11, p. 2011–2022, 2008.

WANG, H.; DONG, L.; LII, M. F.; LEE, A. L.; CREVOISIER, R.; MOHAN, R.; COX, J. D.; KUBAN, D. A.; CHEUNG, R. Implementation and validation of a three-dimensional deformable registration algorithm for target prostate cancer radiotherapy. *International Journal of Radiation Oncology Biology Physics*, v. 61, n. 3, p. 725–735, 2005.

WANG, H.; DONG, L.; O'DANIEL, J.; MOHAN, R.; GARDEN, A. S.; ANG, K. K.; KUBAN, D. A.; BONNEN, M.; CHANG, J. Y.; CHEUNG, R. Validation of an accelerated demons algorithm for deformable image registration in radiation therapy. *Physics in Medicine and Biology*, v. 50, n. 12, p. 2887–2905, 2005.

WANG, Z.; SIMONCELLI, E. P. Local phase coherence and the perception of blur. In: *Advances in Neural Information Processing Systems*. [S.l.]: MIT Press, 2004. v. 16, p. 1435–1442.

WOODS, R. P.; GRAFTON, S. T.; WATSON, J. D. G.; SICOTTE, N. L.; MAZZIOTTA, J. C. Automated image registration: Ii. intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, v. 22, n. 1, p. 153–165, 1998.

WÖRZ, S.; ROHR, K. Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. *Medical Image Analysis*, v. 10, n. 1, p. 41–58, 2005.

- WU, Z.; PAULSEN, K. D.; JR., J. M. S. Adaptive model initialization and deformation for automatic segmentation of T1-weighted brain MRI data. *IEEE Transactions on Biomedical Engineering*, v. 52, n. 6, p. 1128–1131, 2005.
- XIA, R.; ZHAO, J.; LIU, Y. A robust feature-based registration method of multimodal image using phase congruency and coherent point drift. In: *MIPPR 2013: Pattern Recognition and Computer Vision*. [S.l.: s.n.], 2013. v. 8919, p. 891903–891903–8.
- XU, C.; PRINCE, J. L. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, v. 7, n. 3, p. 359–369, 1998.
- YUILLE, A.; STOLORZ, P.; UTANS, J. Statistical physics, mixtures of distributions, and the em algorithm. *Neural Computation*, v. 6, n. 2, p. 334–340, 1994.
- ZAMBAL, S.; SCHÖLLHUBER, A.; BÜHLER, K.; HLADUVKA, J. Fast and robust localization of the heart in cardiac MRI series - a cascade of operations for automatically detecting the heart in cine MRI series. In: *Proceedings of the Third International Conference on Computer Vision Theory and Applications (VISAPP)*. Funchal, Madeira, Portugal: [s.n.], 2008. v. 1, p. 341–346.
- ZHANG, J.; CHEN, L.; WANG, X.; TENG, Z.; BROWN, A. J.; GILLARD, J. H.; GUAN, Q.; CHEN, S. Compounding local invariant features and global deformable geometry for medical image registration. *PLoS ONE*, Public Library of Science, v. 9, n. 8, p. e105815, 2014.