

# DISSERTAÇÃO DE MESTRADO

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
CIÊNCIA DA COMPUTAÇÃO

**“Detector de pontos salientes 3D baseado na DT-  
CWT com aplicação no posicionamento de  
malhas deformáveis em imagens de ressonância  
magnética do cérebro”**

**ALUNO: Breno da Silveira Souza**  
**ORIENTADOR: Prof. Dr. Ricardo José Ferrari**

São Carlos  
Agosto/2019

CAIXA POSTAL 676  
FONE/FAX: (16) 3351-8233  
13565-905 - SÃO CARLOS - SP  
BRASIL

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**DETECTOR DE PONTOS SALIENTES 3D  
BASEADO NA DT-CWT COM APLICAÇÃO NO  
POSICIONAMENTO DE MALHAS  
DEFORMÁVEIS EM IMAGENS DE  
RESSONÂNCIA MAGNÉTICA DO CÉREBRO**

**BRENO DA SILVEIRA SOUZA**

**ORIENTADOR: PROF. DR. RICARDO JOSÉ FERRARI**

São Carlos – SP

Agosto/2019

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**DETECTOR DE PONTOS SALIENTES 3D  
BASEADO NA DT-CWT COM APLICAÇÃO NO  
POSICIONAMENTO DE MALHAS  
DEFORMÁVEIS EM IMAGENS DE  
RESSONÂNCIA MAGNÉTICA DO CÉREBRO**

**BRENO DA SILVEIRA SOUZA**

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

Agosto/2019



# 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 Breno da Silveira Souza, realizada em 30/08/2019:

---

Prof. Dr. Ricardo José Ferrari  
UFSCar

---

Prof. Dr. Marcelo Suetake  
UFSCar

---

Prof. Dr. Marcelo Zancheta do Nascimento  
UFU

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Marcelo Zancheta do Nascimento e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ao) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.

---

Prof. Dr. Ricardo José Ferrari

## **AGRADECIMENTOS**

Agradeço a todos que me ajudaram durante o período do mestrado, de forma direta ou indireta. Em especial, agradeço à minha família, aos amigos e principalmente ao meu orientador, Prof. Dr. Ricardo José Ferrari, cuja dedicação me inspira a ser um profissional melhor.

Agradeço também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) - número de processo: 2017/24391-0, pelo apoio financeiro durante a pesquisa.

## RESUMO

O imageamento por Ressonância Magnética (RM) é um método não invasivo que possibilita alta diferenciação de tecidos moles, proporcionando um bom nível de detalhamento anatômico. Por esse motivo, a imagem de RM é frequentemente usada no auxílio ao diagnóstico e acompanhamento de doenças neurodegenerativas, dentre as quais podemos citar a doença de Alzheimer (DA). A análise da alteração na forma e da redução gradual do volume de estruturas de interesse no cérebro auxilia no estagiamento da DA. Visto que a marcação manual de tais estruturas (para posterior cômputo do volume e avaliação de forma) na imagem de RM é altamente susceptível a erros, diversos métodos automáticos foram propostos na literatura para essa tarefa. Alguns dos mais populares são os chamados modelos geométricos deformáveis. Eles são capazes de capturar a variabilidade geométrica de estruturas anatômicas pois carregam em si informações *a priori* de formato e localização. Entretanto, apesar dessas vantagens, modelos geométricos deformáveis precisam ser posicionados próximos à estrutura que se deseja segmentar para se obter bons resultados. Uma maneira de minimizar esse problema é usar pontos salientes obtidos da imagem para estimar uma transformação deformável local, a qual será aplicada aos vértices da malha de uma determinada estrutura para realizar seu posicionamento inicial. Este projeto propõe uma nova abordagem para a detecção de pontos salientes 3D em imagens de RM baseada na transformada *wavelet* complexa de árvore dupla (*Dual-Tree Complex Wavelet Transform* - DT-CWT). O detector proposto combina os mapas de resposta dos filtros da DT-CWT calculados para diferentes escalas e orientações para criar um mapa acumulado de energia. A abordagem proposta foi testada para o posicionamento de malhas representando os hipocampos em imagens 3D de RM e os resultados obtidos foram comparados com anotações manuais realizadas por especialistas. Valores médios de Dice (DC) e da métrica *Hausdorff Average Distance* (HAD) obtidos pelo método proposto mostraram-se melhores (DC = 0,58/0,58 e HAD = 0,73/0,75 para o hipocampo esquerdo e direito, respectivamente) quando comparados com valores obtidos de um posicionamento guiado por uma transformação Afim (DC = 0,49/0,50, HAD = 1,04/1,11) e uma transformação estimada por pontos salientes detectados usando uma técnica baseada na informação de congruência de fase (DC = 0,55/0,56, HAD = 0,84/0,83).

**Palavras-chave:** Detector de pontos salientes 3D, transformada *wavelet* complexa de árvore dupla, modelos geométricos deformáveis, imagem de ressonância magnética.

## ABSTRACT

The magnetic resonance (MR) imaging is a noninvasive method that allows high soft tissue differentiation by providing a good anatomical detailing level. For this reason, the MR image is often used to help in diagnosis and supervision of neurodegenerative diseases, in which we can mention the Alzheimer's disease (AD). The analysis of shape changes and gradual volume reduction of brain structures of interest helps to identify the AD staging. Considering that the manual annotation of those structures in the MR image (to posterior volume and shape evaluation) is highly susceptible to measurement error, several automatic methods have been proposed in the literature for this task. Some of the most popular among the proposed methods are so-called geometric deformable models. They are able to catch the geometric variability of anatomical structures because they offer a priori knowledge of shape and localization. However, despite these advantages, geometric deformable models need to be positioned close to the structure to be segmented to achieve success. One way to minimize this problem is to use salient points to estimate a local deformable transformation, which will be applied to the mesh vertices of a given structure to perform its initial positioning. In this project, we propose a new approach for the detection of 3-D salient points in MR images which is based on the dual-tree complex wavelet transform (DT-CWT). Our method combines filter response maps calculated for different scales and orientations of the DT-CWT to create an accumulated map of energy responses. We tested our approach for the positioning of hippocampi meshes in 3-D brain MR images and compared the results with manual annotations made by a medical specialist. Mean values of Dice (DC) and Hausdorff Average Distance (HAD) measurements of our proposed method showed better results (DC = 0.58/0.58 and HAD = 0.73/0.75 for the left and right hippocampus, respectively) when compared to both an Affine transform guided initialization (DC = 0.49/0.50, HAD = 1.04/1.11) and a transform guided initialization using salient points detected by a phase congruency technique (DC = 0.55/0.56, HAD = 0.84/0.83).

**Keywords:** salient point 3D detector, dual-tree complex wavelet transform, geometric deformable models, magnetic resonance image.

## LISTA DE FIGURAS

2.1	Vistas axial (a), coronal (b) e sagital (c) de uma representação em 3D da localização dos hipocampos, representados na cor vermelha, no encéfalo humano. . . . .	23
2.2	Ilustração do movimento <i>spin</i> dos átomos. Extraído de (HANSON, 2009). . . . .	25
2.3	Movimento de precessão. Extraído de (HANSON, 2009). . . . .	25
2.4	Pulsos de RF. O pulso de $90^\circ$ é chamado de pulso de excitação, o de $180^\circ$ de pulso de inversão e o pulso $\alpha$ pode assumir qualquer valor. Extraído de (MAZZOLA, 2009). . . . .	26
2.5	Retorno do vetor de magnetização ao equilíbrio após a aplicação de um pulso de radiofrequência de $90^\circ$ . Extraído de (MAZZOLA, 2009). . . . .	27
2.6	Aplicação do gradiente no campo magnético. Extraído de (MAZZOLA, 2009). . . . .	28
4.1	Etapas gerais da metodologia de posicionamento inicial dos modelos deformáveis em imagens de RM do cérebro. . . . .	42
4.2	Projeções 2D de uma estrutura do tipo “ponta” - exemplos em suas versões não degradada (a), degradada por <i>bias field</i> ( $BF = 40\%$ ) (b), degradada por ruído Gaussiano ( $\sigma_{ruído} = 3\%$ ) (c) e degradada por ruído ( $\sigma_{ruído} = 3\%$ ) e <i>bias field</i> ( $BF = 40\%$ ) (d). . . . .	45
4.3	Visualização 3D de três imagens da base de dados que representa estruturas do tipo “ponta”. . . . .	46
4.4	Projeção 2D da imagem da base de dados NAC e da máscara dos hipocampos representando a marcação realizada por especialista (região em vermelho na imagem) nas vistas axial (a), coronal (b) e sagital (c). Na vista sagital é exibida a marcação do hipocampo esquerdo. Nas outras projeções, os dois hipocampos estão visíveis. . . . .	46
4.5	<i>Pipeline</i> do pré-processamento. . . . .	48

4.6	Análise multiescala da transformada <i>wavelet</i> para um sinal 1D (a). Translação da função <i>wavelet</i> ao longo do sinal para garantir a interpretação do sinal como um todo (b). . . . .	55
4.7	Na vizinhança de uma borda, a DWT real produz coeficientes com valores altos e baixos. Em contraste, a DT-CWT aproximadamente analítica produz coeficientes cuja magnitude representa melhor a sua proximidade a uma borda. A Figura (a) representa a resposta da DWT aplicada a uma função degrau, sendo que a transição ocorre em $n_0$ . A Figura mostra a resposta da DT-CWT aplicada ao mesmo sinal. Adaptado de (SELESNICK; BARANIUK; KINGSBURY, 2005).	56
4.8	Esquema de decomposição da DT-CWT 1D em três níveis usando representação por árvore dupla para um sinal de entrada $t$ (Adaptação de (KINGSBURY, 1998; SELESNICK; LI, 2003)). Os filtros reais $h_{*0}$ e $h_{*0}$ representam o primeiro par dos filtros em quadratura e os filtros reais $g_{*0}$ e $g_{*0}$ representam o segundo par. Os filtros com índice de valor final igual a 0 são filtros passa-baixa, enquanto que filtros com índices terminados em 1 representam filtros passa-alta (alguns deles são aplicados na saída de um filtro passa-baixa do nível anterior). . . . .	58
4.9	Espectro de Fourier idealizado para a <i>wavelet</i> complexa $\psi(x,y) = \psi(x) \psi(y)$ (a). Espectro de Fourier idealizado para a parte real (b). Adaptado de (SELESNICK; BARANIUK; KINGSBURY, 2005). . . . .	60
4.10	(a) <i>Wavelets</i> no domínio do espaço. (b) Suporte das <i>wavelets</i> do espectro de Fourier. Extraído de (SELESNICK; BARANIUK; KINGSBURY, 2005). . . . .	60
4.11	Espectro de Fourier idealizado para a <i>wavelet</i> discreta $\psi(x,y) = \psi(x) \psi(y)$ . Extraído de (SELESNICK; BARANIUK; KINGSBURY, 2005). . . . .	61
4.12	Isosuperfícies de 7 <i>wavelets</i> 3D para uma DWT 3D típica (a). Isosuperfícies de 7 das 28 <i>wavelets</i> 3D para uma típica DT-CWT 3D (b). Extraído de (SELESNICK; LI, 2003). . . . .	62
4.13	Pipeline de execução do detector proposto em uma imagem de RM. A palavra Up representa a operação de interpolação <i>upsampling</i> . . . . .	64
4.14	Representação gráfica da repartição do espaço log-esférico ao redor de um ponto saliente para o cômputo do descritor <i>shape context</i> . Extraído de (TOMBARI; SALTI; STEFANO, 2010). . . . .	65

4.15	Representação em forma de árvore da decomposição <i>wavelet</i> 2D. . . . .	69
4.16	Malha do hipocampo esquerdo em suas representações triangular (a) e <i>simplex</i> (b). . . . .	70
4.17	Exemplos de pontos salientes anatômicos no sistema ventricular do cérebro. Adaptação de (WÖRZ; ROHR, 2005). . . . .	72
5.1	Resultado visual da redução de ruído, nos planos axial, coronal e sagital. . . .	78
5.2	Resultado da correção do efeito <i>bias field</i> , nos planos axial, coronal e sagital. . .	79
5.3	Comparação entre o “tabuleiro de xadrez” formado por uma das imagens da base de dados e a imagem <i>template</i> antes e depois do corregristo, nos planos axial, coronal e sagital. . . . .	80
5.4	Representação da região de aplicação do algoritmo de detecção de pontos salientes (em vermelho) sobreposta à imagem da base NAC, nas vistas axial (a), coronal (b) e sagital (c). Nas vistas axial e coronal, as máscaras da região do hipocampo direito e esquerdo encontram-se representadas. Na vista sagital apenas a máscara do hipocampo esquerdo encontra-se representada. . . . .	81
5.5	Sobreposição da imagem original com a máscara do encéfalo, nos planos axial, coronal e sagital. . . . .	81
5.6	Detecção de pontos salientes na versão original (a), rotacionada (b), escalonada (c) e transladada (d) do cubo. . . . .	82
5.7	Pontos detectados em três imagens sintéticas sem degradação, do conjunto de imagens descrito na subseção 4.2.1. O ponto dentro de um círculo na estrutura (a) corresponde a um dos FPs detectados. . . . .	83
5.8	Valores de medida F1 obtidos para o descritor proposto na avaliação de casamentos entre a imagem <i>template</i> NAC e cada uma das 50 imagens moderadamente deformadas para os diferentes valores de $R_{max}$ - 16 (a), 32 (b), 64 (c) e 128 (d). . . . .	86
5.9	Valores de Dice para o posicionamento das malhas dos hipocampus esquerdo (E) e direito (D) usando três abordagens distintas: Afim, PC, e DT-CWT. . . .	87
5.10	Valores de HAD para o posicionamento das malhas dos hipocampus esquerdo (E) e direito (D) usando três abordagens distintas: Afim, PC, e DT-CWT. . . .	88

5.11	Valores de Dice para o posicionamento dos hipocampos esquerdo e direito estratificados por idade usando três abordagens distintas: Afim, PC, e DT-CWT. Em parêntesis está indicado o número de imagens em cada uma das classes. . .	90
5.12	Valores de Dice para o posicionamento dos hipocampos esquerdo e direito estratificados por diagnóstico clínico usando três abordagens distintas: Afim, PC, e DT-CWT. Em parêntesis está indicado o número de imagens em cada uma das classes. . . . .	91
5.13	Mistura de funções de distribuição Rayleigh, representando a magnitude dos coeficientes <i>wavelets</i> dentro de uma sub-banda, estimada para a escala mais fina na primeira (a) e na décima primeira (b) orientações de decomposição. . .	94
5.14	Valores de Dice médio obtidos para o posicionamento das malhas dos hipocampos esquerdo (na cor vermelha) e direito (na cor azul) na base EADC-ADNI após a seleção dos pontos salientes detectados considerando um número mínimo de orientações no qual os coeficientes equivalentes deveriam apresentar persistência em todas as escalas de decomposição. . . . .	96

## LISTA DE TABELAS

3.1	Resumo comparativo das técnicas apresentadas nessa revisão bibliográfica. O marcador (s) indica a que as imagens ou malhas foram geradas sinteticamente. DC representa a sigla para a métrica de Dice ( <i>Dice Coefficient</i> ), e MD para distância média ( <i>Mean Distance</i> ) que se encontra em milímetros (mm). . . . .	38
4.1	Quantidade de imagens na base de dados EADC-ADNI para cada uma das condições diagnósticas. . . . .	47
4.2	Resumo das bases de dados usadas nesse projeto de pesquisa, com destaque à quantidade de imagens, resolução e uma breve descrição de cada uma delas. . .	47
5.1	Resultados selecionados da combinação dos parâmetros $\alpha$ , $\beta$ e limiar $T_E$ . . . .	83
5.2	Valor U e <i>p-value</i> da comparação de cada uma das versões degradadas com a versão sem degradação das estruturas do tipo “ponta”. . . . .	84
5.3	Valor médio de FPs detectados no conjunto de imagens degradado e quantidade de FNs que ocorreram no conjunto de imagens pela quantidade de <i>bias field</i> e ruído especificadas. . . . .	85
5.4	Número de pontos mantidos na região do hipocampo esquerdo (a) e direito (b), usando o modelo HMT, considerando um número mínimo de orientações. . . .	95
5.5	Valores médios de Dice na base de dados EADC-ADNI após o ajuste usando a técnica <i>Simplex Mesh</i> para os modelos dos hipocampus direito e esquerdo posicionados usando uma transformação Afim, por meio de uma transformação guiada pelos pontos da DT-CWT, com a imagem original e com a imagem de energia, destacando os valores alcançados pelo posicionamento e pelo ajuste realizado a partir do posicionamento destacado. . . . .	97

## LISTA DE ABREVIATURAS E SIGLAS

ADNI	Alzheimer's Disease Neuroimaging Initiative
CBCT	Cone Beam Computed Tomography
CNN	Convolutional Neural Network
CT	Computed Tomography
DA	Doença de Alzheimer
DC	Dice Coefficient
DP	Desvio Padrão
DT-CWT	Dual-Tree Complex Wavelet Transform
DWT	Discrete Wavelet Transform
EADC	European Alzheimer's Disease Consortium
EM	Expectation Maximization
FLAIR	Fluid Attenuated Inversion Recovery
FN	Falso Negativo
FP	Falso Positivo
GT	Ground-Truth
HAD	Hausdorff Average Distance
HMC	Hidden Markov Chain
HMT	Hidden Markov Tree
ITK	Insight Toolkit

MCI	Mild Cognitive Impairment
MD	Mean Distance
MDP	Markov Decision Process
N3	Nonparametric Nonuniform Intensity Normalization
NAC	Neuroimage Analysis Center
NLM	Non-Local Means
PC	Phase Congruency
PD-w	Proton Density-weighted
RM	Ressonância Magnética
ROI	Region of Interest
SIFT	Scale Invariant Feature Transform
SIL	Sinal de Indução Livre
SNR	Signal-to-Noise Ratio
T1-w	T1-weighted
T2-w	T2-weighted
VP	Verdadeiro Positivo
WHO	World Health Organization

## LISTA DE SÍMBOLOS

Símbolo	Descrição
$\mathbb{C}$	Conjunto ou números complexos.
$T$	Período de uma função periódica.
$t$	Discretização de um sinal no tempo - equações 4.2, 4.3, 4.4, 4.5,4.6, 4.7, 4.8, 4.9.
$a_k$	Coefficiente que mede a amplitude da frequência $k$ da formulação da Série de Fourier representada na equação 4.2.
$j$	Unidade imaginária nas equações 4.2, 4.3, 4.7, 4.10, 4.13.
$e^{jt}$	Função moduladora exponencial nas equações 4.2, 4.3, 4.7.
$\omega_0$	Frequência fundamental na equação 4.2.
$s$	Fator de escala nas equações 4.4, 4.5,4.6, 4.8, 4.9, 4.17.
$n$	Fator de deslocamento no tempo nas equações 4.4, 4.5,4.6.
$\psi(t)$	Função <i>wavelet</i> passa alta no tempo $t$ .
$\phi(t)$	Função <i>wavelet</i> passa baixa no tempo $t$ .
$\mathcal{H}$	Símbolo da transformada de Hilbert.
$\alpha$	Parâmetro para controle do peso relativo das escalas na equação 4.17.
$\beta$	Parâmetro para diminuir o intervalo entre o menor e o maior resultado produzido dentro de uma escala na equação 4.17.
$T_E$	Parâmetro usado como limiar para definir pontos salientes no mapa de energia acumulado da equação 4.18.
$w$	Magnitude do coeficiente <i>wavelet</i> na equação 4.26.
$\sigma$	Parâmetro de escala da distribuição de probabilidade Rayleigh nas equações 4.26, 4.27.
$Z$	Variável aleatória na equação 4.27.
$\gamma$	Parâmetro de amortecimento na equação 4.28.
$\alpha_{sm}$	Parâmetro de controle da força interna na equação 4.28.
$\beta_{sm}$	Parâmetro de controle da força externa na equação 4.28.

<b>Símbolo</b>	<b>Descrição</b>
$\eta_r$	Número de distâncias radiais do espaço log-esférico para a construção do descritor <i>shape context</i> nas equações 4.19, 4.22.
$\eta_a$	Número de ângulos azimute do espaço log-esférico para a construção do descritor <i>shape context</i> nas equações 4.20, 4.22.
$\eta_e$	Número de ângulos de elevação do espaço log-esférico para a construção do descritor <i>shape context</i> na equação 4.21.
$R_{max}$	Tamanho do raio do espaço log-esférico para a construção do descritor <i>shape context</i> na equação 4.19.
U	Informação estatística do teste U de Mann-Whitney.
<i>p-value</i>	Informação estatística do teste U de Mann-Whitney.

# SUMÁRIO

<b>CAPÍTULO 1 – INTRODUÇÃO</b>	<b>18</b>
1.1 Motivação . . . . .	18
1.2 Objetivos . . . . .	20
1.2.1 Objetivo geral . . . . .	20
1.2.2 Objetivos específicos . . . . .	20
1.3 Organização do texto . . . . .	21
<b>CAPÍTULO 2 – INFORMAÇÕES COMPLEMENTARES</b>	<b>22</b>
2.1 Doença de Alzheimer . . . . .	22
2.2 Ressonância Magnética . . . . .	23
2.2.1 Princípios Físicos . . . . .	24
2.2.2 Imageamento . . . . .	26
2.2.3 Formação da imagem . . . . .	27
2.2.4 Espaço K . . . . .	28
2.2.5 Tipos de ponderação . . . . .	28
2.2.6 Principais problemas: ruído e <i>bias field</i> . . . . .	29
2.3 Considerações finais . . . . .	29
<b>CAPÍTULO 3 – REVISÃO BIBLIOGRÁFICA</b>	<b>30</b>
3.1 Visão Geral . . . . .	30
3.2 Trabalhos Correlatos . . . . .	31

3.2.1	Detectores de pontos salientes . . . . .	31
3.2.2	Detectores de pontos anatômicos específicos . . . . .	34
3.3	Discussão . . . . .	37
3.4	Considerações finais . . . . .	39
 <b>CAPÍTULO 4 – MATERIAIS E MÉTODOS</b>		<b>41</b>
4.1	Visão Geral . . . . .	41
4.2	Base de Dados . . . . .	43
4.2.1	Imagens sintéticas do tipo “ponta” . . . . .	44
4.2.2	NAC <i>Brain Atlas</i> . . . . .	44
4.2.3	EADC-ADNI . . . . .	46
4.3	Pré-processamento . . . . .	48
4.3.1	Redução de ruído . . . . .	49
4.3.2	Correção do efeito <i>bias field</i> . . . . .	50
4.3.3	Padronização de histogramas . . . . .	50
4.3.4	Alinhamento espacial (corregistro afim) . . . . .	50
4.3.5	Máscara da região de interesse . . . . .	51
4.3.6	Extração do encéfalo . . . . .	51
4.4	Detecção de pontos salientes . . . . .	52
4.4.1	Transformada de Fourier . . . . .	52
4.4.2	Transformada <i>wavelet</i> . . . . .	54
4.4.3	Transformada <i>wavelet</i> complexa de árvore dupla . . . . .	56
4.4.4	DT-CWT 2D . . . . .	58
4.4.5	DT-CWT 3D . . . . .	60
4.4.6	Detector de Fauqueur estendido para 3D . . . . .	62
4.5	Descrição de pontos salientes . . . . .	63
4.5.1	Técnica <i>Shape Context</i> . . . . .	64

4.6	Casamento de os pontos salientes . . . . .	66
4.7	Modelo probabilístico <i>Hidden Markov Tree</i> . . . . .	66
4.8	Modelo deformável <i>Simplex Mesh</i> . . . . .	68
4.9	Métodos de avaliação . . . . .	70
4.9.1	Métricas . . . . .	70
4.9.2	Avaliação do detector . . . . .	72
4.9.3	Avaliação do descritor . . . . .	73
4.9.4	Avaliação do posicionamento das malhas dos hipocampos . . . . .	74
4.9.5	Avaliação da aplicabilidade do modelo HMT e do algoritmo de deformação . . . . .	74
4.10	Considerações Finais . . . . .	75
<b>CAPÍTULO 5 – RESULTADOS E DISCUSSÕES</b>		<b>77</b>
5.1	Pré-processamento . . . . .	77
5.1.1	Redução de ruído . . . . .	77
5.1.2	Correção do efeito <i>bias field</i> . . . . .	78
5.1.3	Alinhamento espacial (corregistro) . . . . .	78
5.1.4	Máscara da região para a detecção . . . . .	78
5.1.5	Máscara da região para treinamento do modelo HMT . . . . .	79
5.2	Detecção de pontos salientes . . . . .	80
5.2.1	Análise de invariância à rotação, escala e translação . . . . .	80
5.2.2	Análise das estruturas do tipo “ponta” . . . . .	81
5.3	Descrição dos pontos salientes . . . . .	85
5.4	Posicionamento dos modelos deformáveis . . . . .	85
5.4.1	Otimização dos parâmetros . . . . .	85
5.4.2	Comparação com outros métodos de posicionamento . . . . .	87
5.5	Avaliação do modelo probabilístico HMT . . . . .	92

5.5.1	Treinamento . . . . .	92
5.5.2	Seleção de pontos para o posicionamento . . . . .	94
5.6	Avaliação do ajuste local usando a técnica <i>Simplex Mesh</i> . . . . .	97
5.7	Considerações finais . . . . .	98
<b>CAPÍTULO 6 – CONCLUSÕES</b>		<b>99</b>
6.1	Conclusões . . . . .	99
6.2	Contribuições . . . . .	100
6.3	Sugestões para trabalhos futuros . . . . .	101
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>		<b>102</b>

# Capítulo 1

## INTRODUÇÃO

---

---

*Esse capítulo apresenta a motivação para a realização desta pesquisa, os objetivos definidos no projeto e a descrição da organização do texto.*

### 1.1 Motivação

O imageamento por Ressonância Magnética (RM) é um método de imageamento não invasivo que possibilita alta diferenciação de tecidos moles, proporcionando um bom nível de detalhamento anatômico (MAZZOLA, 2009). Uma das aplicações da imagem de RM se dá no auxílio ao diagnóstico e acompanhamento de doenças neurodegenerativas, comumente realizados via análise do volume e da forma de estruturas atingidas por essas doenças. Entre as doenças em que a imagem de RM é utilizada para auxílio ao diagnóstico e ao acompanhamento, destaca-se a doença de Alzheimer (DA).

A DA é uma desordem neurodegenerativa que causa perda de memória e comportamentos anormais nos indivíduos acometidos (LUO; TANG, 2017). É uma doença associada à idade e com o inevitável envelhecimento populacional a probabilidade de ocorrência da DA tem aumentado. Por esse motivo, existe uma crescente demanda pela quantificação de mudanças anatômicas em estruturas de interesse do cérebro, como os hipocampos (direito e esquerdo), já que eles são atingidos pelas alterações iniciais da doença (CHAN, 2017). Visto que a marcação manual realizada por um especialista (para posterior cômputo do volume e avaliação do formato) na imagem de RM é altamente susceptível a erros, devido a variações subjetivas intra e interobservadores, considerando o nível de experiência do especialista e problemas decorrentes do cansaço e da fadiga visual, vários pesquisadores têm voltado seus esforços para o desenvolvimento de técnicas computacionais para auxiliar a segmentação de estruturas cerebrais em imagens de RM (BALAFAR et al., 2010; HIRALAL; MENON, 2016).

Existem diversos métodos na literatura para realizar a tarefa de segmentação automática e semiautomática (BALAFAR et al., 2010) de estruturas cerebrais, entretanto, a maioria deles têm sua performance comprometida devido a presença de ruído, pobre informação de contraste e *bias field* (JUNTU et al., 2005), características usuais em imagens médicas. Técnicas de segmentação baseadas em modelos geométricos deformáveis (ou malhas) (WANG et al., 2005; HEGADI; KOP; HANGARGE, 2010; FERRARI et al., 2011; VILLA-PINTO; FERRARI, 2016; POLONI et al., 2018) normalmente apresentam bons resultados em imagens médicas, pois oferecem conhecimento *a priori* da informação anatômica, intrinsecamente contida nas malhas, durante o processo de segmentação.

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 deseja 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) e com grande variação anatômica, caso típico da DA (FJELL et al., 2014).

Trabalhos anteriores voltados para o posicionamento inicial de modelos deformáveis utilizam-se da aplicação de uma transformação nos vértices do modelo (VILLA-PINTO; FERRARI, 2016; POLONI et al., 2018). A transformação é estimada a partir do casamento entre pontos automaticamente detectados na imagem de interesse e de pontos automaticamente detectados na imagem à qual o modelo se encontra alinhado. Esses pontos são chamados de pontos salientes e podem ser definidos como pontos localizados em regiões específicas da imagem que podem facilmente ser identificadas pelo sistema visual humano, como bordas, cantos e locais de alta curvatura (SOLOMON; BRECKON, 2011). No caso de imagens médicas, pontos salientes ainda podem corresponder a pontos de referência anatômicos (ou *landmarks*). Dessa maneira, para o posicionamento do modelo, primeiramente é necessário detectar (ou seleccionar) pontos que representem as mesmas regiões em diferentes imagens, descrevê-los e garantir que a correspondência entre esses pontos seja correta para estimar a transformação.

Dentre os diversos trabalhos propostos na literatura para a detecção de pontos salientes em imagens volumétricas, destacam-se aqueles que utilizam informações de intensidade da imagem (RISTER; HOROWITZ; RUBIN, 2017), informações do domínio da frequência da imagem (VILLA-PINTO; FERRARI, 2016; FERRARI et al., 2011) e, mais recentemente, trabalhos focados em detecção de *landmarks* por meio de técnicas de *deep learning* (ZHENG et al., 2015; GHESU et al., 2019). A transformada *wavelet* complexa de árvore dupla (*Dual-Tree*

*Complex Wavelet Transform* - DT-CWT) é uma ferramenta usada para a análise de sinais que permite a obtenção de informação de frequência com certa localização no tempo/espço. Ela foi proposta por Kingsbury (KINGSBURY, 1998) e representa um aprimoramento da transformada *wavelet* discreta (*Discrete Wavelet Transform* - DWT), na medida em que permite invariância aproximada ao deslocamento do sinal, menor efeito de *aliasing* e maior seletividade direcional. A DT-CWT constitui uma ferramenta poderosa, visto que permite a decomposição da imagem em diferentes escalas e orientações. Embora a DT-CWT já tenha sido aplicada para a detecção de pontos salientes em imagens 2D em (FAUQUEUR; KINGSBURY; ANDERSON, 2006), ela foi muito pouco explorada para a detecção de pontos salientes 3D (KOVACS; FERRARI, 2017).

## 1.2 Objetivos

Os objetivos geral e específicos desse projeto de pesquisa são descritos a seguir.

### 1.2.1 Objetivo geral

Desenvolver uma técnica para a detecção de pontos salientes 3D usando a DT-CWT com aplicação no posicionamento inicial de modelos deformáveis em imagens de RM do cérebro visando o auxílio ao diagnóstico e acompanhamento de doenças como a DA.

### 1.2.2 Objetivos específicos

Para tornar possível o desenvolvimento desta pesquisa, foram definidos os seguintes objetivos específicos:

- Estender a técnica de detecção de pontos salientes 2D proposta por Fauqueur, Kingsbury e Anderson (2006) para uso com imagens 3D de RM;
- Desenvolver um descritor para representar de maneira precisa e inequívoca cada ponto saliente detectado;
- Investigar o uso do modelo probabilístico *Hidden Markov Tree* (HMT) (CROUSE; NOWAK; BARANIUK, 1998; ROMBERG et al., 2001) para a seleção de pontos salientes nas imagens;
- Analisar características de repetibilidade e acurácia do detector desenvolvido usando imagens 3D sintéticas na presença de ruído e *bias field*;

- Comparar os resultados obtidos utilizando o detector desenvolvido com os resultados obtidos por posicionamento guiado por uma transformação Afim e com os resultados obtidos usando o detector de pontos salientes baseado na técnica de congruência de fase (*Phase Congruency* - PC) 3D (FERRARI et al., 2011; VILLA-PINTO; FERRARI, 2016).

### **1.3 Organização do texto**

O presente texto foi organizado em seis capítulos, incluindo esse capítulo introdutório. No capítulo 2 são expostas as informações complementares para a compreensão da pesquisa. No capítulo 3 é apresentada a revisão do estado da arte no tópico de pesquisa principal desse trabalho. O capítulo 4 apresenta a teoria por trás das principais técnicas utilizadas, desde o pré-processamento até o objetivo final. Os resultados e as discussões são apresentados no capítulo 5. Por fim, no capítulo 6 são apresentadas as conclusões.

# Capítulo 2

## INFORMAÇÕES COMPLEMENTARES

---

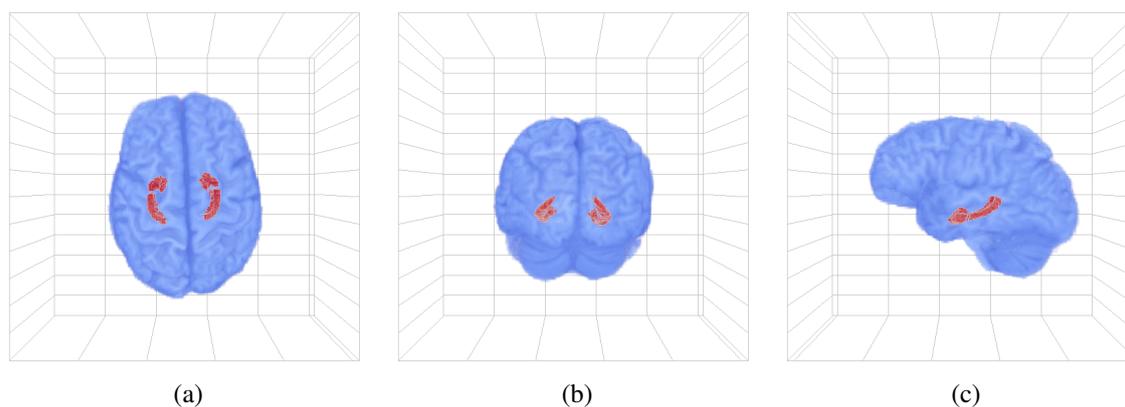
---

*Este capítulo apresenta informações complementares relativas à doença de Alzheimer e ao processo de formação da imagem de ressonância magnética.*

### 2.1 Doença de Alzheimer

Devido à melhoria no cuidado com a saúde no século passado, a população mundial como um todo tem apresentado uma maior expectativa de vida, fato que ocasiona um progressivo envelhecimento populacional e, conseqüentemente, um aumento no número de pessoas com doenças associadas à idade, grupo no qual a demência está incluída. Segundo um estudo realizado pelo *World Health Organization* (WHO) (CHAN, 2017), o número estimado de pessoas em 2012 vivendo com algum tipo de demência foi de cerca de 36 milhões. Esse número dobrará até 2030 e triplicará até 2050, aumentando consideravelmente gastos privados e públicos com esse segmento da população.

A DA é um tipo de demência neurodegenerativa progressiva do sistema nervoso central e que possui uma alta prevalência na população idosa (MCKHANN et al., 1984), caracterizando-se por mudanças de personalidade, deterioração de função cognitiva, perda de neurônios e presença de placas senis em regiões específicas do cérebro (WEST et al., 1994). No diagnóstico clínico da doença, usualmente, os médicos entrevistam familiares e amigos para buscar informações relacionados às condições gerais de saúde do paciente, dieta, habilidade de realização de tarefas rotineiras, e em relação à mudanças comportamentais, usando testes cognitivos e de memória aplicados ao paciente (NIH, 2017). Além desses testes clínicos, é comum a solicitação de exames médicos padrão, como testes de sangue e urina, para identificar outras causas possíveis do problema; e exames de imageamento do cérebro para verificar alterações, como atrofia da região hipocampal e córtex, e dilatação dos ventrículos (FOX et al., 1996).



**Figura 2.1:** Vistas axial (a), coronal (b) e sagital (c) de uma representação em 3D da localização dos hipocampos, representados na cor vermelha, no encéfalo humano.

Os hipocampos, direito e esquerdo, são estruturas localizadas nos lobos temporais do cérebro humano, como podem ser observados na Figura 2.1. Envolvida principalmente em aspectos relacionados à memória, a formação hipocampal é uma das primeiras a desenvolver sinais patológicos da DA e, em casos avançados, é a região mais profundamente afetada (WEST et al., 1994). Isso acontece devido ao fato de que os marcadores anatômicos da doença, emaranhados neurofibrilares e placas senis (PERL, 2010), são encontrados em grande quantidade nos hipocampos (FOX et al., 1996). Macroscopicamente, essas alterações são acompanhadas pela atrofia bruta dos hipocampos e região do córtex e de um alargamento concomitante dos espaços de líquido cefalorraquiano.

A DA pode ser diagnosticada com precisão somente no pós-morte, juntando-se medidas clínicas com análises histopatológicas do tecido cerebral na autópsia (NIH, 2017). Portanto, existe uma grande necessidade de refinamento nos critérios de diagnóstico, considerando que boa parcela dos casos com diagnóstico clínico são descobertos, após a morte, como possuidores de outras condições, e não DA (MCKHANN et al., 1984).

## 2.2 Ressonância Magnética

A ressonância magnética é uma técnica de imageamento não invasiva usada principalmente em aplicações médicas para produzir imagens de alta qualidade do interior do corpo humano. A RM possui alta capacidade para diferenciar tecidos moles e para capturar características anatômicas e funcionais. A imagem por RM é o resultado da interação do forte campo magnético produzido pelo equipamento de imageamento com os prótons de hidrogênio do tecido humano, criando uma condição para que um pulso de radiofrequência possa ser enviado para excitar

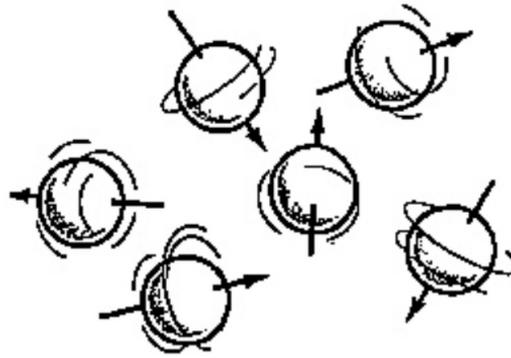
uma determinada fatia do corpo humano (MAZZOLA, 2009). Descrito inicialmente por Bloch, Hanson e Packard (1946) e Purcell, Torrey e Pound (1946) em 1946, o fenômeno da ressonância magnética nuclear foi utilizado para a aquisição de imagens do interior do corpo humano apenas em 1973, pela necessidade de avanços tecnológicos (GOSSUIN et al., 2010). Nas subseções a seguir são apresentados, de forma resumida, os princípios básicos para a aquisição de imagens de RM usando como base os trabalhos de Mazzola (2009), Hanson (2009) e Gossuin et al. (2010).

### 2.2.1 Princípios Físicos

O imageamento por RM é baseado na absorção e emissão de energia pelos átomos de hidrogênio do corpo humano. Os sistemas de imageamento usam a reação desses átomos quando eles se interagem com um campo magnético para formar uma imagem médica detalhada. O átomo de hidrogênio é o mais abundante do corpo humano, devido a cerca de 70% da massa corporal ser composta por moléculas de água. O fato do seu núcleo conter apenas um próton resulta em grande momento magnético e, portanto, numa maior sensibilidade ao fenômeno de ressonância magnética. Além disso, as características da ressonância magnética nuclear entre o hidrogênio presente no tecido normal e no tecido patológico se diferem bastante (MAZZOLA, 2009). Por esses motivos básicos, o átomo de hidrogênio é o escolhido para a utilização no imageamento por RM.

Os núcleos dos átomos possuem uma propriedade chamada de *spin*, ou momento angular, que faz com que eles se comportem como se fossem pequenos ímãs que giram em torno de si mesmos e que são espacialmente orientados de maneira aleatória. Esse movimento faz com que os átomos sejam magnéticos. Sob condições normais e considerando apenas a ação do campo magnético terrestre, os momentos magnéticos dos átomos do corpo humano não possuem uma orientação espacial definida, fazendo com que os núcleos se movimentem em diferentes direções ocasionando magnetização resultante do volume do tecido igual a zero. Na Figura 2.2 esse movimento dos átomos é ilustrado.

Os sistemas de imageamento, através de uma bobina ou antena receptora, detectam e processam os sinais gerados quando os átomos de hidrogênio são colocados em um forte campo magnético e excitados por um pulso de rádio frequência (RF). Quando estão sob a ação de um forte campo magnético, os prótons de hidrogênio tendem a se orientar de acordo com a direção do campo aplicado, tanto de forma paralela, quanto de forma antiparalela ao campo. Na tentativa de alinhamento ao campo, os átomos passam a se comportar analogamente a um pião, num movimento denominado precessão. A frequência de precessão dos prótons de hidrogênio

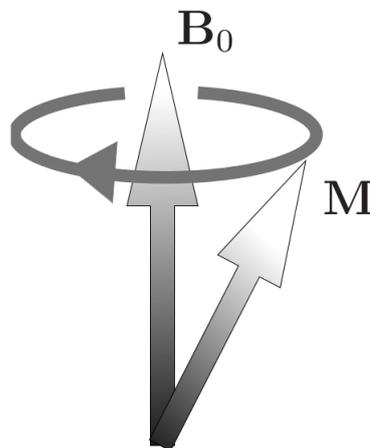


**Figura 2.2:** Ilustração do movimento *spin* dos átomos. Extraído de (HANSON, 2009).

é determinada pela equação de Larmor:

$$\omega = \kappa \mathbf{B}_0, \quad (2.1)$$

em que  $\kappa$  representa a razão giromagnética e  $\mathbf{B}_0$  representa o valor do campo magnético externo aplicado. Como pode ser observado nesta equação, a frequência de precessão do próton de hidrogênio é diretamente proporcional a intensidade do campo magnético. A Figura 2.3 ilustra o movimento de precessão de um núcleo magnetizado girando ao redor do campo magnético  $\mathbf{B}_0$ .



**Figura 2.3:** Movimento de precessão. Extraído de (HANSON, 2009).

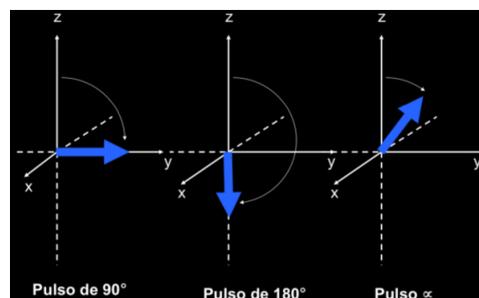
Fazendo-se uso de uma analogia com uma bússola (HANSON, 2009), se ela se encontrar próxima a um campo magnético poderoso, sua agulha irá se alinhar na direção do campo magnético. Normalmente, a agulha da bússola se encontra mergulhada em um líquido para amortecer suas oscilações. Sem o líquido, a agulha irá vibrar na direção norte por um período antes de atingir um estado de equilíbrio. A frequência das oscilações depende do campo magnético. Ondas de radiofrequência são campos magnéticos que mudam de acordo com o tempo

(oscilam), e enquanto a agulha vibra, elas serão emitidas na mesma frequência que a da agulha.

Considerando uma situação na qual a bússola é colocada em um campo magnético criado por um ou mais magnetos poderosos, depois de um certo período de tempo, a agulha passará a apontar na direção do campo magnético. Se um pequeno empurrão é dado à agulha, de forma perpendicular ao campo magnético, ela passará a vibrar pelo norte, mas gradualmente irá se acalmar novamente. As oscilações irão ocorrer em uma frequência chamada de frequência de ressonância. Enquanto a agulha estiver oscilando, ondas de radiofrequência com a mesma frequência de oscilação serão emitidas. Essas ondas podem ser medidas por uma bobina ou antena.

### 2.2.2 Imageamento

Partindo dos princípios apresentados anteriormente, no imageamento, primeiramente é aplicado um forte campo magnético  $\mathbf{B}_0$  normalmente no eixo longitudinal (eixo z). Para que uma corrente elétrica seja induzida em uma bobina posicionada de forma perpendicular ao plano transversal (plano xy), é necessário que o vetor de magnetização como um todo, ou parte dele, esteja no plano transversal e possua coerência de fase. Desse modo, para reorientar o vetor de magnetização, um pulso de radiofrequência (RF) de curta duração perpendicular a  $\mathbf{B}_0$  é aplicado. Esse pulso de radiofrequência é chamado de campo  $\mathbf{B}_1$  e deve estar em fase com a frequência de precessão. Quando esse pulso é aplicado, o vetor de magnetização que se encontrava alinhado a  $\mathbf{B}_0$  é desviado em  $90^\circ$  (esse é o ângulo de desvio mais utilizado e é chamado de pulso de excitação, pulsos de  $180^\circ$  também são utilizados e são chamados de pulsos de inversão, mas o ângulo pode assumir qualquer valor), transferindo assim o vetor para o plano transversal. A Figura 2.4 ilustra a aplicação do pulso de radiofrequência.



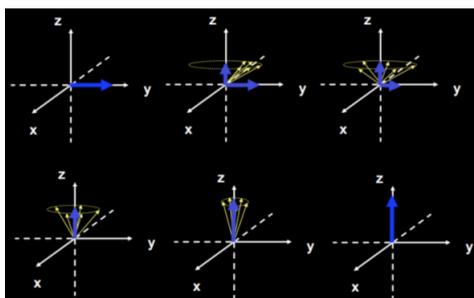
**Figura 2.4: Pulsos de RF. O pulso de  $90^\circ$  é chamado de pulso de excitação, o de  $180^\circ$  de pulso de inversão e o pulso  $\alpha$  pode assumir qualquer valor. Extraído de (MAZZOLA, 2009).**

Quando se encerra a aplicação do pulso de RF, ocorre o fenômeno de relaxação, no qual o vetor de magnetização retorna para sua posição de repouso, alinhando-se novamente a  $\mathbf{B}_0$ . Durante a relaxação, os prótons de hidrogênio induzem uma tensão elétrica que é captada pela

bobina. O formato do sinal induzido (ou sinal de indução livre - SIL) é uma onda senoidal amortecida. Duas constantes de tempo foram criadas para caracterizar a relaxação:

- **T1**: está relacionada com o tempo de retorno da magnetização para o eixo longitudinal. A constante **T1** representa o tempo necessário para que o vetor de magnetização no eixo longitudinal atinja 63% do seu valor inicial.
- **T2**: está relacionada com a redução da magnetização no plano transversal. A constante **T2** representa o tempo necessário para que o vetor de magnetização no plano transversal atinja 37% do seu valor inicial.

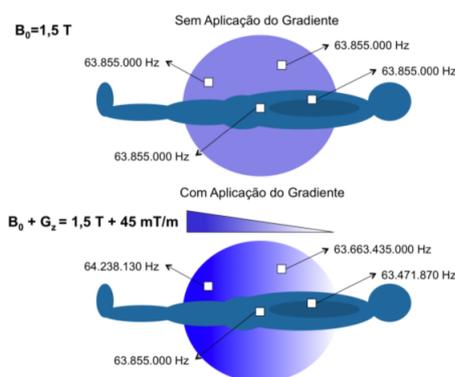
É importante conhecer os tempos de relaxação dos tecidos já que a intensidade dos sinais - e consequentemente o contraste dos tecidos - depende desses fatores. Na Figura 2.5 é ilustrado o retorno do vetor de magnetização ao equilíbrio após a aplicação de um pulso de radiofrequência de 90°.



**Figura 2.5: Retorno do vetor de magnetização ao equilíbrio após a aplicação de um pulso de radiofrequência de 90°. Extraído de (MAZZOLA, 2009).**

### 2.2.3 Formação da imagem

A formação da imagem de RM depende da codificação espacial do sinal recebido na bobina via uso de gradientes do campo magnético. Quando o volume está sob um campo magnético de valor único e uniforme, e um sinal de RF for enviado com valor de frequência exatamente igual à frequência de precessão dos prótons de hidrogênio, então, todo volume será excitado. O sinal que é enviado para a bobina irá conter informação de todo o volume, mas não é possível saber de qual parte vem. Para se ter localização espacial são adicionados gradientes de campo magnético, o que permite variar linearmente em uma dada direção a intensidade do campo magnético, modificando a frequência de precessão dos prótons de hidrogênio ao longo do campo e permitindo localizar diferentes pontos através de informações de seleção de corte, codificação de fase (a fase dos *spins* também é alterada com o acionamento do gradiente) e codificação de frequência. A Figura 2.6 demonstra o acionamento do gradiente.



**Figura 2.6:** Aplicação do gradiente no campo magnético. Extraído de (MAZZOLA, 2009).

Por exemplo, se o gradiente de seleção de corte for acionado na direção  $z$ , cada posição nesse eixo terá uma frequência de precessão distinta. Enviando um pulso de RF na mesma frequência de precessão da região que se deseja excitar, apenas aquela região enviará sinal para a bobina captar. O gradiente de codificação de fase permite diferenciar as fases dos prótons de hidrogênio, diferenciando cada um dos eixos do corte. O gradiente de codificação de frequência é acionado na direção restante no momento de leitura do sinal.

## 2.2.4 Espaço K

O sinal coletado está codificado em informações de fase e frequência. Ele é mapeado num conceito abstrato chamado de espaço K, que pode ser visto como uma matriz na qual cada elemento descreve a informação contida nos sinais amostrados, que são representados por números complexos, indicando as informações de magnitude e fase do sinal. A informação contida no espaço K consiste basicamente na transformada de Fourier da distribuição espacial do sinal, que, neste caso, corresponde à transformada de Fourier da imagem do corte. Então, para se obter a imagem a partir do espaço K, utiliza-se a transformada inversa de Fourier sobre o espaço K.

## 2.2.5 Tipos de ponderação

O contraste entre os tecidos em uma imagem de RM é controlado pela escolha de um método de ponderação. Por exemplo, uma imagem é chamada de *T2-weighted* (T2-w) se os parâmetros de aquisição são escolhidos de modo que o contraste na imagem reflita principalmente variações no tempo T2 (HANSON, 2009). A ponderação *T1-weighted* (T1-w), por sua vez, resulta em imagens com a distribuição dos níveis de cinza de maneira proporcional ao valor do contraste T1 de cada tecido, e dessa maneira, propicia às imagens um melhor delineamento da

anatomia do paciente (VILLA-PINTO, 2016). Existem outros tipos de ponderação como a ponderação *Proton Density-weighted* (PD-w) e a ponderação *Fluid Attenuated Inversion Recovery* (FLAIR), que são obtidas pelo ajuste dos parâmetros de aquisição da imagem.

### 2.2.6 Principais problemas: ruído e *bias field*

Um dos fatores de degradação da imagem de RM é o ruído (HANSON, 2009). Ele pode ser causado por efeitos fisiológicos (pulso, respiração, movimento) mas, mesmo se o paciente estiver completamente parado, ainda existe um limite superior para a qualidade alcançável da imagem em um dado período. Na ausência de ruído causado por efeitos fisiológicos, a relação sinal-ruído (*Signal-to-Noise ratio* - SNR) é proporcional ao tamanho do voxel e à raiz quadrada do tempo de aquisição da imagem. Tal ruído tem natureza descrita pela distribuição de Rice (GUDBJARTSSON; PATZ, 1995), mas para um SNR elevado, tal distribuição pode ser aproximada para uma distribuição Gaussiana. Outra degradação comum das imagens de RM é aquela causada pelo efeito *bias field*. O *bias field* é um sinal suave de baixa frequência que corrompe as imagens de RM, especialmente imagens produzidas por máquinas de RM antigas (JUNTU et al., 2005). Esse efeito faz com que a intensidade de um sinal medido a partir de um tecido homogêneo varie suavemente ao longo da imagem (SLED; ZIJDENBOS; EVANS, 1998), dificultando o desempenho de algoritmos automáticos que assumem a uniformidade de intensidade nos tecidos.

## 2.3 Considerações finais

Este capítulo apresentou uma breve descrição a respeito da DA, com destaque às estruturas dos hipocampos, que são inicialmente acometidas pela doença; e uma revisão básica da teoria por trás dos princípios físicos da formação da imagem de RM, que representa o método de imageamento comumente utilizado no caso de doenças neurodegenerativas. Além disso, foi exposto que a imagem de RM pode ter diferentes tipos de contraste, definidos pela ponderação escolhida. Em concordância com a maioria dos trabalhos publicados na literatura que utilizam detectores de pontos salientes em imagens de RM, neste estudo será usada apenas a ponderação T1-w, visto que esta fornece um melhor contraste nas estruturas anatômicas imageadas. Por fim, nesse capítulo foi apresentado alguns dos problemas intrínsecos da imagem de RM, como ruído e efeito *bias field*, os quais serão tratados pela etapa de pré-processamento das imagens.

# Capítulo 3

## REVISÃO BIBLIOGRÁFICA

---

---

*Este capítulo apresenta uma revisão dos principais trabalhos publicados na literatura relacionados à detecção de pontos salientes em imagens volumétricas, tópico desta pesquisa.*

### 3.1 Visão Geral

O objetivo principal desse projeto de pesquisa foi o desenvolvimento de uma técnica automática para a detecção e descrição de pontos salientes em imagens volumétricas de RM do cérebro com aplicação no posicionamento inicial de modelos geométricos deformáveis. A detecção de tais pontos é de grande interesse em diversas aplicações médicas, pois eles formam um conjunto de pontos bem identificados e bem localizados que podem servir como âncoras para a determinação de uma transformação (TUYTELAARS; MIKOLAJCZYK, 2008) para uso no corregristo entre imagens clínicas (KLEIN et al., 2009) ou entre imagens clínicas e um atlas (BAILLEUL; RUAN; BLOYET, 2003; IDE et al., 2008; POLONI et al., 2018), caso esse comumente usado para o posicionamento automático de modelos deformáveis (ALLAIRE et al., 2008a, 2008b; VILLA-PINTO; FERRARI, 2016; POLONI et al., 2018).

Dessa maneira, a revisão bibliográfica se ateu às técnicas de detecção de pontos salientes com foco na aplicação em imagens médicas volumétricas. Conforme será apresentado, mais recentemente foram propostas diversas abordagens para a seleção de pontos em imagens médicas, que vão desde o uso de modelos de intensidade paramétricos, até a utilização de técnicas de redes profundas (*deep learning*) para aprender as características de um ponto saliente anatômico. A seção a seguir foi dividida em duas subseções: a primeira (subseção 3.2.1) apresenta os trabalhos que propõem a detecção de pontos salientes gerais, sejam eles representantes de estruturas anatômicas ou não. Já a segunda (subseção 3.2.2) apresenta detectores de pontos salientes anatômicos específicos em imagens médicas volumétricas.

## 3.2 Trabalhos Correlatos

### 3.2.1 Detectores de pontos salientes

**Ferrari et al. (2011)** Neste trabalho, os autores propuseram uma técnica para a detecção de pontos salientes em imagens volumétricas a partir da extensão do modelo de congruência de fase - PC - proposto por Kovese (2003) para 3D. Ao invés de assumir que um ponto saliente é representado por um ponto de intensidade de gradiente maximal, o modelo de congruência de fase postula que características são percebidas em pontos na imagem onde os componentes de Fourier (obtidos a partir da decomposição do sinal via transformada de Fourier) se encontram maximalmente em fase. Por não ser diretamente dependente de informações de intensidade, em contraste às técnicas baseadas em informações de gradientes, a medida de PC é invariante à variações de intensidade, resolução e contraste contidas no sinal. Segundo os autores, ela oferece uma boa indicação da significância de uma característica visualmente distintiva presente em uma imagem.

Para obter a medida de congruência de fase, um banco de filtros log-Gabor 3D é projetado diretamente no domínio da frequência. A seguir, a imagem original é filtrada no domínio da frequência, gerando um conjunto de mapas de energia em diferentes orientações (representadas por ângulos de azimute e elevação), que são combinados para obter a medida de congruência de fase. Para realizar a detecção dos pontos salientes, os autores utilizam-se da teoria de análise de momentos discretos da imagem para definir uma matriz de momentos de segunda ordem e, a partir da análise dos autovalores dessa matriz, computada para todos os pontos da imagem, determinar se um ponto é saliente ou não.

A avaliação do detector se deu pela comparação dos pontos detectados pela técnica com pontos manualmente anotados por um especialista em imagens de CT 3D do pescoço e da cabeça. O método proposto foi capaz de detectar pontos salientes que são anatomicamente importantes. Nesse trabalho, a técnica foi utilizada para propagar contornos de estruturas de interesse entre imagens de diferentes modalidades (CT e *Cone Beam CT* - CBCT) via estimação de uma transformação que mapeia pontos anatômicos de referência para as posições na imagem alvo onde existe uma grande similaridade na vizinhança dos pontos automaticamente detectados. Foram gerados padrões ouro (*Ground-Truths* - GT) das estruturas do tronco encefálico e da medula espinhal por meio da marcação manual por especialista. Os resultados da aplicação da técnica foram avaliados por meio da distância média (*Mean Distance* - MD) do contorno inicializado em relação ao GT, com valores de 2,66mm para a medula espinhal e 3,23mm para o tronco encefálico.

**Villa-Pinto e Ferrari (2016)** Dando continuidade ao trabalho de Ferrari et al. (2011), Villa-Pinto e Ferrari (2016) aplicaram a técnica de congruência de fase 3D descrita anteriormente para a inicialização de modelos geométricos deformáveis representando as estruturas dos hipocampos e do corpo caloso em imagens volumétricas de RM do cérebro. A metodologia de inicialização proposta pelos autores faz a combinação da detecção de pontos salientes pela técnica de congruência de fase, descritores baseados na técnica *shape context* (BELONGIE; MALIK; PUZICHA, 2002; MORTENSEN; DENG; SHAPIRO, 2005), comparação de descritores baseada na técnica *nearest neighbor* e estimação de uma transformação *B-spline* multinível não rígida.

A avaliação da precisão do detector foi realizada utilizando um conjunto de imagens 3D sintéticas representando estruturas de alta curvatura geradas a partir do modelo paramétrico proposto por Wörz e Rohr (2005). Esse conjunto de imagens foi corrompido por ruído e *bias field* e foi realizada uma comparação da distância dos pontos detectados em relação ao ponto desejado na presença de diferentes níveis de degradação. Os resultados demonstraram que a técnica apresenta boa robustez à ruído e ao efeito *bias field*, e isso se deve principalmente ao fato de que a detecção leva em consideração informações de fase da imagem, e não informações de magnitude dos níveis de cinza. A menor distância obtida na detecção foi de 0 voxels e a maior foi de 5 voxels. A avaliação do casamento entre descritores e da inicialização dos modelos deformáveis foi realizada usando um conjunto de imagens de RM e de malhas das estruturas dos hipocampos e do corpo caloso gerado pela aplicação de uma série de transformações aleatórias em uma imagem e nas malhas alinhadas à essa imagem.

Na avaliação do descritor, os descritores calculados para cada uma das imagens sintéticas foram comparados com os descritores da imagem original para verificar a correspondência entre pontos das diferentes imagens, e a acurácia média foi de 84,04% de casamentos corretos. Na avaliação da inicialização foi verificada a sobreposição entre os modelos sintéticos após a aplicação da metodologia proposta e os modelos originais utilizando métricas de Dice (*Dice Coefficient* - DC) - e Jaccard. Os valores tiveram um aumento significativo para ambas as métricas com a aplicação da metodologia proposta, resultando em valor médio de sobreposição na métrica Dice para a estrutura do corpo caloso subindo de 0,41 para 0,88, de 0,34 para 0,84 para o hipocampo esquerdo e de 0,31 para 0,84 para o hipocampo direito. Apesar dos resultados animadores, o trabalho não apresentou nenhuma avaliação em imagens clínicas.

**Poloni et al. (2018)** No trabalho de Poloni et al. (2018), os autores propuseram a construção de um atlas probabilístico de pontos salientes 3D automaticamente detectados em um conjunto de imagens de RM do cérebro de pacientes saudáveis (obtidas da base de dados *Information*

*eXtraction from Images - IXI*<sup>1</sup>) utilizando a técnica de detecção apresentada nos trabalhos anteriores (FERRARI et al., 2011; VILLA-PINTO; FERRARI, 2016).

O atlas probabilístico tem como principal função informar quantitativamente sobre a estabilidade dos pontos salientes detectados, considerando uma população de imagens e, com essa informação, ponderar os pontos salientes considerando a ocorrência espacial ao longo desse conjunto de imagens de pacientes saudáveis. Dessa maneira, é possível aumentar a robustez do casamento com pontos detectados em uma imagem clínica. O principal objetivo da utilização do atlas probabilístico é obter uma transformação mais eficiente e menos ruidosa para o posicionamento inicial de malhas representando as estruturas dos hipocampos, direito e esquerdo.

Como métrica de avaliação quantitativa, os autores utilizaram a DC normalizado e a distância média de Hausdorff (*Hausdorff Average Distance - HAD*). O conjunto de malhas rotuladas dos hipocampos da base de dados EADC-ADNI<sup>2</sup>, que consiste em malhas obtidas a partir da marcação realizada por especialistas, foi utilizado como GT na comparação final desse trabalho. Os resultados das métricas mostraram que, em média, a abordagem baseada em atlas probabilístico foi significativamente melhor (DC = 0,74/0,70, HAD = 0,70/0,73, para os hipocampos esquerdo e direito, respectivamente) quando comparada à abordagem de inicialização de pontos salientes baseada em *template* (DC = 0,65/0,61, HAD = 0,88/0,91) e à abordagem de inicialização utilizando apenas informação de transformação afim (DC = 0,58/0,53, HAD = 1,10/1,22).

**Rister, Horowitz e Rubin (2017)** No trabalho de Rister, Horowitz e Rubin (2017), os autores utilizam pontos salientes 3D detectados automaticamente invariantes à escala e à rotação, para realizar o corregristo entre imagens de RM do cérebro. Nesse trabalho, a técnica *Scale Invariant Feature Transform* (SIFT) foi estendida para dimensões arbitrárias por meio de modificações chave na etapa de atribuição da orientação e de elaboração dos histogramas de gradiente que compõem o descritor.

Esse trabalho oferece três contribuições principais. Primeiramente, os autores apresentam uma generalização do algoritmo SIFT para  $\mathbb{R}^n$ , com atenção particular para  $\mathbb{R}^3$ . As modificações realizadas na orientação e na geometria do histograma de gradiente permitem aos pontos salientes resultantes apresentarem as mesmas invariâncias à escala e à rotação em  $\mathbb{R}^n$  que o algoritmo original possui em  $\mathbb{R}^2$ . A estimativa da orientação é baseada na decomposição dos autovalores de uma estrutura tensor (*structure tensor*), obtida a partir da correlação entre componentes do gradiente da imagem. A correção do histograma de gradientes foi realizada pela

<sup>1</sup><https://brain-development.org/ixi-dataset/>

<sup>2</sup><http://www.hippocampal-protocol.net>

substituição da abordagem na definição dos *bins*, que anteriormente utilizava-se de coordenadas esféricas uniformemente espaçadas. O novo histograma é formado usando valores de gradientes obtidos para um icosaedro regular, interpolando as contribuições entre os *bins* do histograma pela coordenada baricêntrica das faces triangulares.

A segunda contribuição refere-se a uma nova análise para a escolha da vizinhança na detecção do máximo local. Nesse caso, os autores demonstram experimentalmente que a vizinhança  $\ell^1$  tem resultados melhores que a vizinhança  $\ell^\infty$  para o correto registro de imagens médicas. Por fim, os autores desenvolveram um sistema de correto registro de imagens totalmente automático baseado no casamento de pontos salientes. Em imagens de RM de cérebro, utilizando a nova técnica SIFT, os autores obtiveram 92% de sobreposição média (métrica DC) entre um atlas rotulado e as imagens de RM. Em imagens de CT de abdômen, a coluna vertebral é registrada com um erro médio de 4,82mm.

### 3.2.2 Detectores de pontos anatômicos específicos

**Wörz e Rohr (2005)** Neste trabalho os autores apresentam uma técnica para a localização de pontos salientes anatômicos 3D em imagens volumétricas que aplica modelos de intensidade paramétricos ajustados aos níveis de cinza da imagem. Os modelos paramétricos considerados representam três estruturas principais: estrutura de “ponta” (*tip-like*), estrutura de sela (*saddle-like*), e estrutura do tipo esfera (*sphere-like*), cada um deles refletindo as variações de intensidade das estruturas em imagens tomográficas 3D.

A técnica proposta descreve as intensidades de pontos salientes anatômicos na imagem em uma região como uma função analítica, com um número determinado de parâmetros. Dessa maneira, é necessário definir uma região de interesse (*Region of Interest* - ROI) grande o suficiente ao redor de cada ponto saliente para capturar uma quantidade de informações satisfatória da imagem e assim garantir a localização precisa do ponto. No entanto, é necessário ter cuidados na definição da ROI, pois uma região muito grande pode conter estruturas vizinhas que influenciam negativamente na acurácia da localização. Sendo assim, os autores desenvolveram também uma nova abordagem para a seleção do tamanho ótimo da ROI 3D. Além disso, também propuseram um algoritmo para inicializar automaticamente os parâmetros do modelo, usando-se de propriedades diferenciais e de outras propriedades da estrutura na imagem.

A avaliação dos resultados foi realizada em imagens sintéticas 3D, representando estruturas do tipo “ponta”, “sela” e “esfera” com diferentes variações na orientação, tamanho e nível de ruído gaussiano, e em imagens de RM e de tomografia computadorizada (*Computed Tomography* - CT) do cérebro. Nas imagens sintéticas, a localização do ponto saliente é definida

pelo modelo usado para a geração dessas imagens. Os autores conseguiram alta acurácia na localização dos pontos com a técnica proposta, sendo que o erro médio da localização estimada foi de 0,04mm nas estruturas do tipo “ponta”, 0,03mm nas estruturas do tipo sela e 0,01mm nas estruturas do tipo “esfera”. Nas imagens clínicas os autores usaram pontos anatômicos manualmente marcados por especialistas como *ground truth*. O erro médio da localização dos pontos nessas imagens foi de 1,63mm utilizando a inicialização automática das regiões de interesse.

**Zheng et al. (2015)** Mais recentemente, dada a popularização da aplicação de técnicas de *deep learning* para a resolução de diversos tipos de problemas computacionais (NAJAFABADI et al., 2015), alguns trabalhos foram propostos para a detecção de pontos salientes anatômicos utilizando o aprendizado oferecido por essas técnicas. Tais trabalhos, embora focados em resolver problemas 2D, demonstraram que é possível aprender características importantes de um conjunto de treinamento.

O desafio de se utilizar técnicas de *deep learning* em um conjunto de imagens 3D reside no fato de que o vetor de entrada, nesse caso, é muito maior que o vetor de entrada do caso 2D. Isso aumenta drasticamente o tempo de computação e as chances de ocorrer *overfitting*<sup>3</sup>, especialmente quando combinado com um número de amostras de treinamento limitado, o que é típico em aplicações de imagens médicas (ZHENG et al., 2015).

Para tratar esse problema, Zheng et al. (2015) propuseram um algoritmo de *deep learning*, dividido em dois passos principais, que é capaz de realizar a detecção de pontos salientes anatômicos em dados volumétricos. Primeiramente, uma rede com apenas uma camada escondida é usada para testar todos os voxels da imagem. O objetivo dessa etapa inicial é obter um conjunto menor de voxels que são mais promissores a representarem pontos salientes anatômicos de interesse. A seguir, é aplicada uma classificação mais acurada, utilizando uma rede profunda composta por 3 camadas. Essa rede recebe como entrada os voxels selecionados pela rede rasa (de uma camada), de modo que a entrada passa a ter tamanho 2000, diminuindo consideravelmente a complexidade do algoritmo. Para mitigar o problema de *overfitting*, os autores extraem pequenos *patches* 3D de uma pirâmide multirresolução. As características da imagem aprendidas pela rede profunda são combinadas com características da transformada *wavelet* de Haar para aumentar a acurácia da detecção. O método foi avaliado quantitativamente para a detecção da bifurcação da artéria carótida em um conjunto de imagens de CT 3D da região da cabeça-pescoço composto por 455 pacientes. Segundo os autores, a técnica proposta consegue realizar a detecção desse ponto saliente anatômico em menos de um segundo e com uma distância média

---

<sup>3</sup>*Overfitting*, ou sobreajuste, ocorre quando um modelo se ajusta muito bem ao conjunto de treinamento, mas se mostra ineficaz para novos dados de entrada.

de 2,64mm.

**Ghesu et al. (2016)** Em (GHESU et al., 2016), os autores também utilizam técnicas de *deep learning* para a detecção de pontos salientes anatômicos em imagens médicas 2D e 3D. Nesse trabalho, os autores combinam uma rede neural convolucional (*Convolutional Neural Network* - CNN) com técnicas de aprendizado de máquina para modelar pontos anatômicos nas imagens. Os autores propuseram um novo paradigma de aprendizado que contrasta com o paradigma normalmente adotado na literatura, em que a maioria das soluções se restringem em aprender um modelo de aparência e escanear exaustivamente um espaço de parâmetros para detectar uma característica anatômica específica. Nesse caso, todos os aspectos relacionados ao entendimento do problema e a garantia da generalidade do algoritmo são responsabilidade do engenheiro, enquanto a máquina desacoplada desse alto nível de compreensão executa cegamente a solução.

No novo paradigma, a modelagem da aparência do modelo e a estratégia de busca de parâmetros são realizadas simultaneamente como uma tarefa comportamental unificada para um agente artificial, resultando em uma técnica que combina as vantagens de aprendizado comportamental (*behavior learning*), alcançados por meio do aprendizado por esforço (*reinforcement learning*), com extração hierárquica de características por meio de uma técnica de *deep learning*. A modelagem do aprendizado do agente (treinamento) é realizada por meio do modelo *Markov Decision Process* (MDP), em que os estados descrevem uma região de interesse com seu centro representando a posição atual do agente. As ações são modeladas denotando os movimentos do agente no espaço paramétrico e os ganhos são modelados codificando o *feedback* supervisionado recebido pelo agente, representado pela mudança da distância relativa até a localização do ponto saliente anatômico depois de executar o movimento. Dada a definição do modelo, o objetivo do agente é selecionar ações por meio da interação com o ambiente para maximizar o ganho futuro acumulado, e isso é feito utilizando uma CNN profunda para aproximar a função de ação-valor ótima para determinar as ações.

A avaliação da metodologia proposta foi realizada em conjuntos de imagens médicas 2D (de RM e ultrassom) e em um conjunto de imagens 3D (CT). Os conjuntos de dados consistem em 891 imagens de RM 2D cardíacas de 338 pacientes, 1186 imagens de ultrassom cardíaco de 361 pacientes e 455 imagens distintas de CT da região pescoço-cabeça. Nas diferentes imagens foram determinados diferentes pontos anatômicos, como o *LV-center* nas imagens de RM 2D e a bifurcação da artéria carótida em imagens de CT 3D. A métrica de avaliação foi a distância média do ponto detectado em relação ao GT demarcado nessas bases de dados, e os resultados mostraram que a técnica proposta pelos autores consegue detectar a uma distância menor em relação ao estado da arte de técnicas usadas para a detecção de pontos anatômicos,

como distância média de 1,8mm para o ponto anatômico *LV-center* em imagens de RM 2D usando a técnica proposta em comparação a 6,2mm do estado da arte. Para a bifurcação da artéria carótida em CT 3D, os autores conseguiram resultados com distância média de 1,8mm, valor superior aos 2,6mm do estado da arte.

**Ghesu et al. (2019)** No trabalho seguinte, Ghesu et al. (2019) estenderam a solução original proposta em (GHESU et al., 2016) para conjuntos de dados mais volumosos com intuito de realizar uma busca por objeto em uma representação multiescala de um conjunto de imagens de CT 3D de tamanhos arbitrários. Para isso, modelos de busca independentes em cada nível do espaço-escala são usados, garantindo que essa busca encontre as características mais discriminativas visíveis naquele nível.

Nesse trabalho, os autores avaliaram o desempenho do método proposto na detecção de 8 pontos salientes anatômicos representando diferentes estruturas, usando um conjunto de 1487 imagens de CT 3D de 532 pacientes que inclui imagens de CT cardíacas, do torax, da região do pescoço-cabeça e da região abdominal. Para avaliar a taxa de sucesso, os autores determinaram distâncias máximas em relação ao GT marcado por especialista de 30mm para o ponto anatômico representando o centro de cada rim e de 10mm para os pontos anatômicos restantes (cantos do osso do quadril, bifurcações de artérias e veias específicas, bifurcação bronquial) sendo que qualquer detecção acima desse limiar é considerada um *outlier*. Por meio de experimentos, o método proposto mostrou-se robusto em relação a *outliers* e alcançou uma acurácia média cerca de 20 a 30% melhor que outras soluções da literatura, com valores de distância média variando de 3,07mm até 6,89mm para os diferentes pontos anatômicos procurados. Ao mesmo tempo, a velocidade do algoritmo é 2-3 ordens de magnitude mais rápida que as soluções comparadas.

A Tabela 3.1 apresenta um resumo comparativo dos trabalhos descritos nessa revisão bibliográfica, destacando os principais pontos de cada um deles.

### 3.3 Discussão

Com base na descrição dos trabalhos apresentada neste capítulo, observa-se que a detecção de pontos salientes em imagens médicas 3D ainda constitui um tópico de pesquisa aberto na literatura. Alguns dos trabalhos selecionados nessa revisão propuseram técnicas para a detecção de pontos salientes em geral (cantos, regiões de alta curvatura), sejam eles anatômicos ou não (FERRARI et al., 2011; VILLA-PINTO; FERRARI, 2016; RISTER; HOROWITZ; RUBIN, 2017), enquanto outros se limitaram ao aprendizado de um modelo que descrevesse pontos

<b>Autores</b>	<b>Detector</b>	<b>Descritor</b>	<b>Base de imagens</b>	<b>Avaliação</b>	<b>Principais resultados</b>
(WÓRZ et al., 2005)	Modelos paramétricos	—	Imagens (s), RM e CT (3D)	MD	$tip=0,04$ , $saddle=0,03$ e $sphere-like=0,01$ ; $RM/CT=1,63$
(FERRARI et al., 2011)	PC 3D	Vizinhança local	CT e CBCT (3D)	MD	Medula espinhal= $2,66$ /Tronco encefálico= $3,23$
(ZHENG et al., 2015)	<i>Deep Learning</i>	—	CT (3D)	MD	Bifurcação da artéria carótida= $2,64$
(VILLA-PINTO et al., 2016)	PC 3D	<i>Shape Context</i>	Imagens e malhas (s)	DC/MD	Hipocampus esq. e dir.= $0,84$ , Corpo Caloso= $0,88$ - DC
(GHESU et al., 2016)	<i>Deep Learning</i>	—	RM e ultrassom (2D), CT (3D)	MD	CT 3D - Bifurcação da artéria carótida= $1,8$
(RISTER et al., 2017)	SIFT 3D	Descritor SIFT	RM e CT (3D)	DC/MD	$RM=92\%$ - DC, $CT=4,82$ - MD
(POLONI et al., 2018)	PC 3D	<i>Shape Context</i>	RM (3D)	DC/HAD	Hip. dir./esq. $0,74/0,70$ - DC — $0,70/0,73$ - HAD
(GHESU et al., 2019)	<i>Deep Learning</i>	—	CT (3D)	MD	MD de $3,07$ a $6,89$

**Tabela 3.1: Resumo comparativo das técnicas apresentadas nessa revisão bibliográfica. O marcador (s) indica a que as imagens ou malhas foram geradas sinteticamente. DC representa a sigla para a métrica de Dice (*Dice Coefficient*), e MD para distância média (*Mean Distance*) que se encontra em milímetros (mm).**

anatômicos (*landmarks*) referentes à estruturas específicas do cérebro humano (WÖRZ; ROHR, 2005; ZHENG et al., 2015; GHESU et al., 2016, 2019). Para o posicionamento das malhas dos hipocampus, é mais interessante que uma grande quantidade de pontos altamente distinguíveis sejam detectados, sejam eles representantes de estruturas anatômicas ou não.

Nessa revisão bibliográfica não foi encontrado nenhum trabalho que utilizasse informações provenientes da DT-CWT 3D para a detecção de pontos salientes em imagens 3D. Para imagens bidimensionais, no entanto, Fauqueur, Kingsbury e Anderson (2006) propuseram um detector simples utilizando a DT-CWT 2D, o qual é capaz de selecionar pontos salientes por meio da aplicação de um limiar a um mapa de energia acumulada gerado a partir dos valores de magnitude dos coeficientes. Esse detector foi estendido para imagens 3D em (KOVACS; FERRARI, 2017), contudo, esse trabalho não apresenta uma avaliação detalhada do detector e nem uma verificação da aplicabilidade em imagens médicas para a solução de problemas.

Além disso, apenas dois dos trabalhos apresentados oferecem uma solução para a descrição dos pontos salientes, que são: a utilização da técnica *shape context* em (VILLA-PINTO; FERRARI, 2016) e o descritor adaptado para 3D da técnica SIFT em (RISTER; HOROWITZ; RUBIN, 2017). Em (FERRARI et al., 2011) também é realizada a descrição dos pontos salientes baseando-se em informações da vizinhança de tais pontos, mas o artigo possui um foco maior na técnica de detecção.

### 3.4 Considerações finais

Nesse capítulo foram apresentados alguns dos trabalhos mais recentes voltados à detecção de pontos salientes em imagens médicas 3D. Alguns desses trabalhos, como citado anteriormente, propõem técnicas capazes de detectar pontos salientes mais gerais, enquanto que outros buscam detectar apenas pontos anatômicos específicos. Os trabalhos que se propõem detectar esses pontos anatômicos específicos não costumam apresentar uma aplicação para os pontos que foram detectados, sendo o objetivo final apenas encontrar esses pontos. Trabalhos que buscam encontrar pontos mais gerais oferecem a possibilidade de serem utilizados para fins variados.

As métricas de avaliação variam entre as aplicações desses detectores, sendo que em detectores de pontos anatômicos é comum utilizar a distância média do ponto detectado até o ponto demarcado na imagem, e nos detectores de pontos salientes no geral, é comum utilizar alguma métrica de sobreposição para verificar a aplicabilidade do detector em uma tarefa específica (posicionamento de malhas, correção do cérebro). Nesses casos, é necessário ter um descritor associado a cada ponto. Por fim, observou-se que a DT-CWT 3D, que constitui uma ferramenta

eficiente e robusta para análise de sinais, tem sido pouco utilizada para a detecção de pontos salientes.

# Capítulo 4

## MATERIAIS E MÉTODOS

---

---

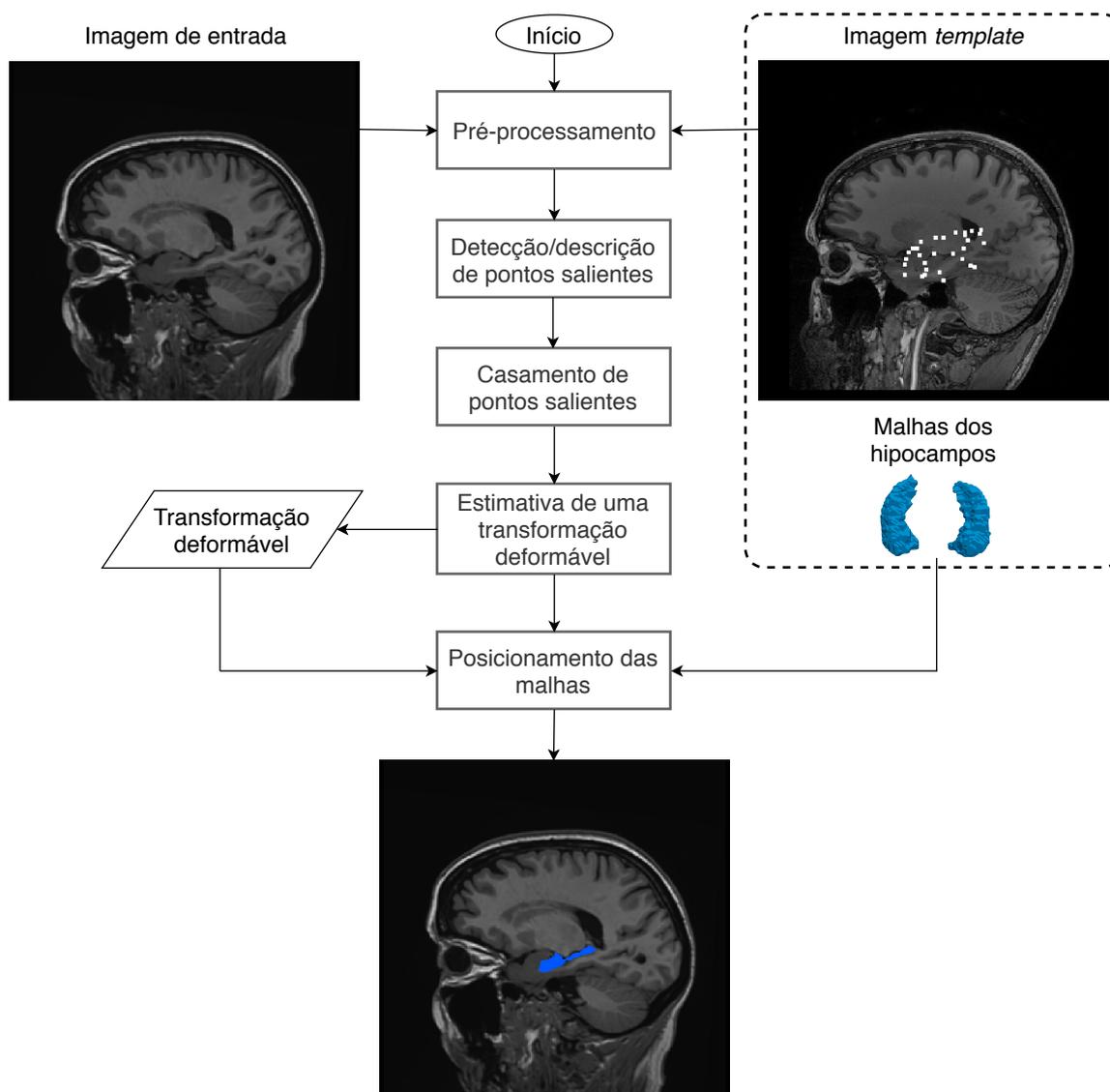
*Esse capítulo apresenta os materiais e métodos que foram usados nas etapas de detecção, descrição, seleção e casamento de pontos salientes em imagens de RM para o posicionamento de modelos geométricos deformáveis representativos de estruturas de interesse do cérebro. Primeiramente, é apresentada uma visão geral das principais etapas da metodologia de posicionamento proposta e, a seguir, são descritas as bases de dados e as técnicas utilizadas em cada etapa.*

### 4.1 Visão Geral

Nesse trabalho foi proposta uma abordagem para a detecção de pontos salientes em imagens volumétricas de RM usando a DT-CWT 3D e a aplicação dos pontos detectados em um *pipeline* de posicionamento de modelos geométricos deformáveis. O diagrama da Figura 4.1 ilustra as principais etapas e a sequência de execução da metodologia proposta.

No *pipeline* de execução, os dados de entrada são uma imagem clínica, uma imagem *template* e malhas que representam as estruturas dos hipocampos direito e esquerdo. As malhas encontram-se alinhadas às estruturas equivalentes na imagem *template*. Para o posicionamento inicial dos modelos, é necessário encontrar uma transformação deformável que mapeia o espaço de coordenadas da imagem *template* para o espaço de uma imagem clínica. Essa transformação é estimada a partir da informação de equivalência dos pontos salientes detectados em ambas imagens. Nesse projeto, a imagem e as malhas descritas na subseção 4.2.2 representam, respectivamente, a imagem *template* e as malhas nas quais a transformação deformável é aplicada. As imagens clínicas são obtidas da base de dados descrita na subseção 4.2.3.

A detecção dos pontos salientes utilizando a DT-CWT 3D segue os seguintes passos:



**Figura 4.1:** Etapas gerais da metodologia de posicionamento inicial dos modelos deformáveis em imagens de RM do cérebro.

1. A DT-CWT 3D é aplicada à imagem de entrada.
2. A informação de magnitude dos coeficientes das diferentes escalas e orientações é combinada por meio da multiplicação ao longo das orientações e do somatório desses valores interpolados ao longo das escalas (da escala mais grosseira até a escala mais fina), formando um mapa de energia acumulada.
3. Uma normalização é aplicada aos valores do mapa de energia, considerando apenas a região de interesse; por exemplo, a região do hipocampo esquerdo.
4. O mapa de energia acumulada representa informações salientes significantes de cada região da imagem, considerando uma decomposição multi-escala e multi-orientada. Dessa

maneira, pontos máximos de uma vizinhança, que possuem valor acima de um determinado limiar, são selecionados como pontos salientes.

A descrição dos pontos salientes segue a técnica *shape context* (subseção 4.5.1), com o mapa final de energia acumulada sendo utilizado para o cálculo do descritor. Essas etapas de detecção e descrição são realizadas para cada imagem clínica e uma única vez para a imagem *template* (da base de dados NAC), que possui os modelos alinhados nos quais a transformação, representada por um campo vetorial de deformação, é aplicada.

A correspondência entre os pontos salientes da imagem clínica e da imagem *template* é determinada por meio da distância  $\chi^2$  dos descritores dos pontos, como descrita na subseção 4.6. Encontrados pontos equivalentes em ambas imagens (na região do entorno do hipocampo), estes são utilizados para estimar uma transformação deformável *B-spline* (LEE; WOLBERG; SHIN, 1997), sendo a imagem clínica definida como imagem fixa e a imagem *template* como a imagem móvel, permitindo o mapeamento dos pontos do espaço de coordenadas da imagem *template* para o espaço de coordenadas da imagem clínica.

As métricas obtidas da comparação dos modelos posicionados a partir dos pontos obtidos pelo detector proposto com a marcação por especialista nas imagens clínicas foram comparados com as métricas obtidas por um posicionamento Afim e com as métricas obtidas pelo mesmo *pipeline* de inicialização utilizando um detector baseado na técnica de congruência de fase.

Além das principais etapas descritas no *pipeline* geral, nesse projeto foi verificada a aplicabilidade do modelo probabilístico HMT (seção 4.7) utilizando o algoritmo de recuperação de sequência de estados proposto por Durand, Goncalvès e Guédon (2004) para descartar pontos salientes que não fossem persistentes ao longo das escalas e orientações, antes de prosseguir para a descrição e equivalência dos pontos. Por fim, uma técnica de ajuste de modelos geométricos deformáveis (*Simplex Mesh* (DELINGETTE, 1994)) foi aplicada às malhas posicionadas.

## 4.2 Base de Dados

Nas subseções a seguir, são descritas as bases de dados utilizadas para a verificação da eficácia da metodologia proposta. As bases de dados consistem em imagens sintéticas representando pontos anatômicos, imagens clínicas de RM e malhas representando as estruturas dos hipocampos. Ao final dessa seção, as principais informações a respeito das bases de dados são organizadas em uma tabela.

### 4.2.1 Imagens sintéticas do tipo “ponta”

Para a verificação da eficácia do detector foi gerado um conjunto com 648 imagens sintéticas elaboradas por meio do modelo matemático paramétrico de estruturas do tipo “ponta” (*tip-like structures*) descrito no trabalho de Wörz e Rohr (2005). Essas imagens simulam estruturas de alta curvatura que se parecem com pontos anatômicos 3D comumente presentes nas imagens de RM do cérebro. As imagens são criadas usando versões de uma estrutura em formato elipsoide com diferentes níveis de afinação, curvatura e inclinação, e diferentes ângulos de rotação.

Dessas 648 imagens, existem 54 estruturas distintas com 12 versões cada uma, sendo que cada versão apresenta um nível de degradação causado por ruído e *bias field*. As imagens possuem dimensões  $64 \times 64 \times 64$  voxels e a localização do ponto saliente de cada imagem é definida pelo modelo usado para gerá-las. Para exemplificação da base de dados, as Figuras 4.2(a), (b), (c) e (d) mostram projeções 2D de uma das estruturas da base de dados com diferentes níveis de degradação. A Figura 4.3 mostra diferentes estruturas dessa base de dados sem degradação em uma visualização 3D.

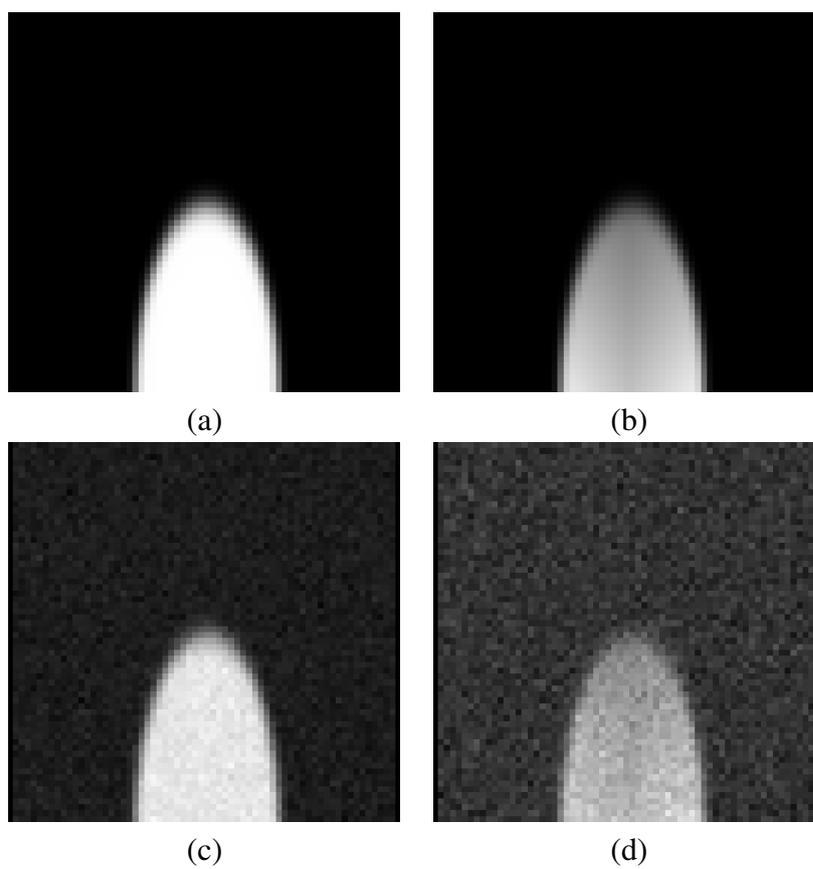
### 4.2.2 NAC *Brain Atlas*

Essa base de dados consiste em uma imagem de RM nas ponderações T1-w e T2-w obtida de um paciente saudável do sexo masculino de 42 anos e um conjunto de 149 malhas triangulares representando estruturas distintas do cérebro alinhadas à essa imagem. As imagens possuem dimensões  $256 \times 256 \times 256$  voxels, com resolução isotrópica de 1mm. Foi elaborada pelo *Neuroimage Analysis Center* (NAC)<sup>1</sup> afiliado ao *Surgical Planning Laboratory* e a Universidade de Harvard.

Essa imagem, em sua versão com ponderação T1-w, foi escolhida para representar a imagem *template* desse projeto, desse modo, a aplicação da transformação deformável para realizar o posicionamento inicial do modelo em uma imagem clínica é realizada nas malhas dos hipocampos associadas à essa imagem. A Figura 4.4 mostra a imagem da base de dados, em conjunto com a marcação das estruturas dos hipocampos realizada por especialistas.

---

<sup>1</sup><https://nac.spl.harvard.edu/>



**Figura 4.2:** Projeções 2D de uma estrutura do tipo “ponta” - exemplos em suas versões não degradada (a), degradada por *bias field* ( $BF = 40\%$ ) (b), degradada por ruído Gaussiano ( $\sigma_{\text{ruído}} = 3\%$ ) (c) e degradada por ruído ( $\sigma_{\text{ruído}} = 3\%$ ) e *bias field* ( $BF = 40\%$ ) (d).

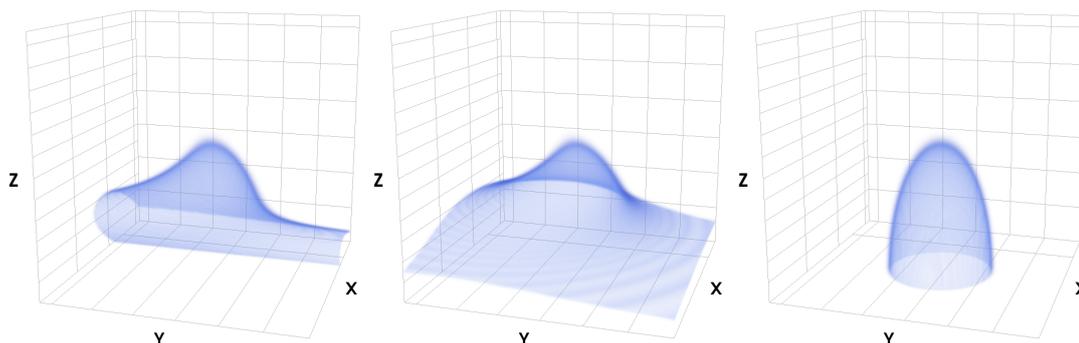


Figura 4.3: Visualização 3D de três imagens da base de dados que representa estruturas do tipo “ponta”.

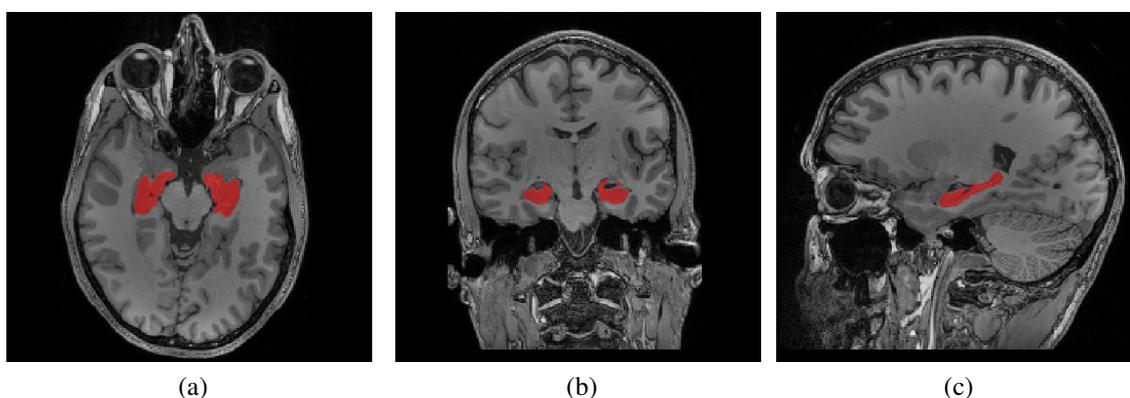


Figura 4.4: Projeção 2D da imagem da base de dados NAC e da máscara dos hipocampus representando a marcação realizada por especialista (região em vermelho na imagem) nas vistas axial (a), coronal (b) e sagital (c). Na vista sagital é exibida a marcação do hipocampo esquerdo. Nas outras projeções, os dois hipocampus estão visíveis.

### 4.2.3 EADC-ADNI

O projeto *Alzheimer’s Disease Neuroimaging Initiative* (ADNI)<sup>2</sup> foi elaborado em 2004 por um grupo de instituições de pesquisa, do setor público e privado, sob a liderança do Dr. Michael W. Weiner. O objetivo principal da iniciativa é desenvolver biomarcadores clínicos, por imagem, genéticos e bioquímicos para auxiliar na detecção precoce e no acompanhamento da doença de Alzheimer. A base de dados do projeto ADNI contém cerca de 2000 imagens de RM rotuladas em pacientes normais, pacientes com comprometimento cognitivo leve (*Mild Cognitive Impairment* - MCI), e pacientes com DA.

A base de imagens EADC-ADNI foi elaborada pela iniciativa *European Alzheimer’s Disease Consortium* (EADC) com o propósito de padronizar o protocolo para a segmentação manual dos hipocampus em imagens de RM. Ela possui 135 imagens com ponderação T1-w que foram

<sup>2</sup><http://adni.loni.usc.edu>

obtidas da base de dados ADNI, em conjunto com máscaras binárias indicando a marcação manual dos hipocampos (direito e esquerdo) em cada uma das imagens. A ponderação T1-w foi escolhida pois fornece melhor contraste nas estruturas anatômicas imageadas. As 135 imagens são de pacientes de ambos os sexos entre 60 e 90 anos com diferentes condições de saúde. A quantidade de imagens de cada classe é apresentada na Tabela 4.1.

	Normal	MCI	DA
Pacientes	44	46	45

**Tabela 4.1: Quantidade de imagens na base de dados EADC-ADNI para cada uma das condições diagnósticas.**

No subgrupo de pacientes normais, a faixa etária com maior número de imagens está presente no intervalo de idade 70 – 74, com 16 imagens no total. As faixas etárias de prevalência de ocorrência de diagnóstico clínico MCI são representadas nos intervalos de idade 75 – 79 e 80 – 85. Assim como para pacientes normais, para o subgrupo da DA, a faixa etária com prevalência de ocorrência é representada pelo intervalo de idade 70 – 75, com 11 imagens. A base de dados EADC-ADNI foi selecionada por possuir as rotulações das estruturas dos hipocampos, servindo como padrão ouro - GT - na validação da inicialização das malhas em imagens clínicas.

Na Tabela 4.2 são apresentadas as principais informações a respeito das três bases de dados utilizadas nesse projeto, especificando a quantidade de imagens, a resolução e uma descrição resumida de cada base. Algumas bases, além de imagens também possuem malhas associadas. Essa informação é destacada na descrição das bases de dados.

Base	Quantidade	Resolução	Descrição
Imagens sintéticas	648	$64 \times 64 \times 64$	Imagens simulando estruturas sintéticas do tipo ponta.
EADC-ADNI	135	variada	Imagens clínicas de RM do cérebro com rotulações das estruturas dos hipocampos associadas a cada uma delas.
NAC	1	$256 \times 256 \times 256$	Imagem de RM de um paciente saudável associada a 149 malhas triangulares representando estruturas cerebrais.

**Tabela 4.2: Resumo das bases de dados usadas nesse projeto de pesquisa, com destaque à quantidade de imagens, resolução e uma breve descrição de cada uma delas.**

### 4.3 Pré-processamento

O *pipeline* de pré-processamento aplicado às imagens da base de dados EADC-ADNI é ilustrado pelo diagrama na Figura 4.5. Primeiramente, elimina-se o ruído (que assume-se ser uma componente aditiva na equação de formação da imagem de RM (SLED; ZIJDENBOS; EVANS, 1998)) para depois tratar a componente multiplicativa causada pelo efeito *bias field*. Tendo um conjunto de imagens livres de ruído e *bias field*, as etapas seguintes incluem a padronização de histogramas, usando a imagem *template* da base NAC, o corregristo afim entre as imagens da base de dados e a imagem *template*, para que todas as imagens estejam no mesmo sistema de coordenadas, e a definição de máscaras binárias representando as regiões de interesse na imagem, já que a imagem como um todo inclui toda caixa craniana. A fundamentação teórica das técnicas utilizadas nessas etapas é apresentada nas subseções a seguir.

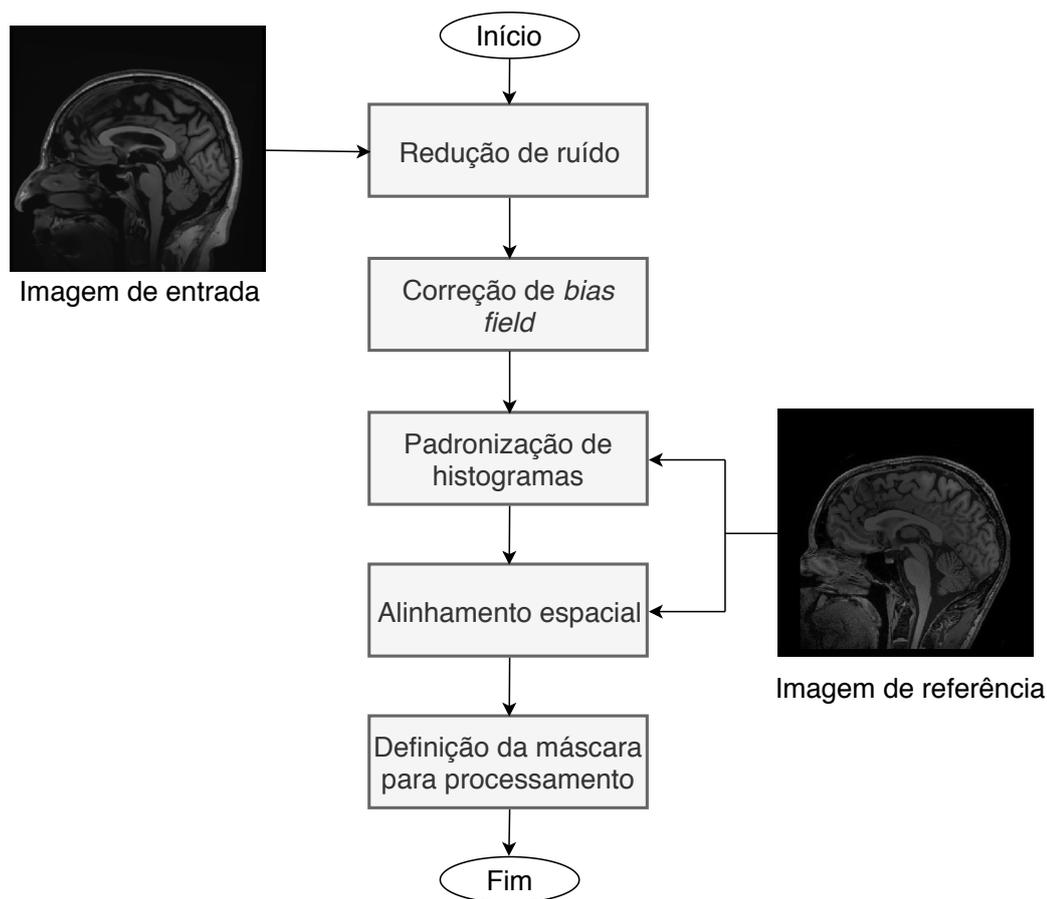


Figura 4.5: Pipeline do pré-processamento.

### 4.3.1 Redução de ruído

A imagem de RM é formada a partir do sinal coletado provindo da iteração de um forte campo magnético produzido pelo equipamento de imageamento e os átomos de hidrogênio do corpo humano (MAZZOLA, 2009). Devido a diversos fatores externos e internos, o sinal recebido é corrompido por ruído e pelo sinal de degradação *bias field*.

Nesse projeto, a etapa de redução de ruído foi realizada utilizando a técnica *Non-Local Means* (NLM) (COUPÉ et al., 2008; TRISTÁN-VEGA et al., 2012), com implementação disponível na ferramenta EZMINC<sup>3</sup>. Essa técnica baseia-se numa média ponderada dos valores dos voxels dentro de uma janela de busca relativamente grande (quando comparada às técnicas de vizinhança tradicionais), a qual corresponde a uma vizinhança não local. Para preservar as estruturas da imagem, os voxels são ponderados de acordo com uma medida de similaridade em relação ao voxel de interesse, de modo que voxels localizados na janela de busca mais similares ao voxel de interesse tenham maior peso para determinar o valor atribuído a ele (COUPÉ et al., 2008). A similaridade é estimada via diferença quadrada entre voxels correspondentes dentro de janelas locais centralizadas nas posições comparadas (TRISTÁN-VEGA et al., 2012).

Considerando-se uma imagem não processada, onde  $u(x_i)$  é o nível de cinza na posição  $x_i = (x, y, z)$ , a saída filtrada do sinal  $\hat{u}(x_i)$  é definida como

$$\hat{u}(x_i) = \sum_{x_j \in \Omega_i} w(x_i, x_j) u(x_j), \quad (4.1)$$

onde  $\Omega_i$  é a janela de busca centralizada no pixel  $x_i$  e  $w(x_i, x_j)$  é o peso atribuído ao voxel  $x_j$  em relação ao voxel  $x_i$  de acordo com os *patches* (de tamanho fixo) centralizados nesses dois pontos.

A técnica NLM gera resultados ótimos para ruído aditivo do tipo Gaussiano. Apesar da natureza dos distúrbios em algumas modalidades de imageamento se diferir da Gaussiana, como é o caso da imagem de RM, que possui ruído com distribuição do tipo Riciano, o NLM foi adaptado com resultados promissores para o contexto do imageamento por RM (TRISTÁN-VEGA et al., 2012). Em imagens de RM que possuem SNR alto, a distribuição representativa do ruído pode ser aproximada a uma distribuição Gaussiana, caso das imagens usadas nesta pesquisa. Um outro ponto importante em relação ao ruído das imagens de RM é a sua cor (AJA-FERNÁNDEZ; VEGAS-SÁNCHEZ-FERRERO, 2016), que neste caso foi assumida como sendo branca, ou seja, que o ruído afeta igualmente todas as frequências.

<sup>3</sup><https://github.com/BIC-MNI/EZminc>

### 4.3.2 Correção do efeito *bias field*

A correção do efeito *bias field* foi realizada usando o algoritmo N4ITK (TUSTISON et al., 2010), uma variante do algoritmo *Nonparametric Nonuniform Intensity Normalization* (N3), com implementação disponível na biblioteca *Insight Toolkit of the National Institutes of Health* (ITK). O método N3 é iterativo e procura o campo multiplicativo que maximiza o conteúdo de alta frequência da distribuição de intensidade do tecido (TUSTISON et al., 2010). Já o N4ITK representa um aperfeiçoamento do algoritmo N3 proposto em (SLED; ZIJDENBOS; EVANS, 1998) pela substituição da estratégia de suavização original por uma alternativa mais vantajosa. Além disso, o algoritmo N4ITK possui um esquema de otimização que melhora o desempenho de convergência.

### 4.3.3 Padronização de histogramas

Um dos principais problemas do processamento de imagens de RM é a falta de padronização na distribuição de intensidade das imagens, o que dificulta a aplicação de algoritmos de segmentação e quantificação (NYUL; UDUPA; ZHANG, 2000) baseados puramente na intensidade dos voxels. Neste projeto, a padronização dos histogramas das imagens da base de dados foi realizada utilizando uma variação do algoritmo proposto por Nyul, Udupa e Zhang (2000), com implementação disponível pela biblioteca ITK. Para o funcionamento do algoritmo é necessário utilizar uma imagem de referência a fim de permitir que as intensidades dos voxels nas imagens a serem corrigidas possam ter uma correspondência consistente entre os tecidos cerebrais em imagens distintas. Esse processo é realizado por meio de uma técnica similar a técnica convencional de casamento de histogramas.

A imagem de referência utilizada nessa etapa é a imagem da base de dados NAC (subseção 4.2.2), visto que ela contém um conjunto de malhas associadas que será utilizado posteriormente para a avaliação do método de posicionamento das malhas. Sendo assim, todas as imagens clínicas têm seus níveis de cinza padronizados de acordo com essa imagem.

### 4.3.4 Alinhamento espacial (corregistro afim)

Corregistro é o processo de sobrepor duas ou mais imagens representando um mesmo objeto, ou objetos similares, que foram obtidas em diferentes intervalos de tempo, de diferentes pontos de vista, e/ou por diferentes sensores (ZITOVA; FLUSSER, 2003). Para o posicionamento correto dos modelos deformáveis nas imagens de RM, é necessário que previamente à aplicação da transformação deformável as imagens estejam padronizadas em um mesmo sis-

tema de coordenadas. Desse modo, é importante realizar um corregristo afim, que envolve apenas transformações globais (rotação, translação, escala e cisalhamento), entre as imagens da base de dados e uma imagem de referência. Nesse projeto, a imagem da base de dados NAC é a imagem de referência.

O corregristo foi realizado por meio da ferramenta Nifty Reg<sup>4</sup>, que possui a implementação da técnica utilizada para esse fim, proposta por Ourselin et al. (2000). Essa técnica é baseada em uma abordagem de casamento de blocos, que oferece um conjunto de pontos correspondentes entre a imagem de referência e a imagem móvel. Com essas correspondências, encontra-se a melhor transformação afim entre os blocos. Esses passos são repetidos por meio de um processo iterativo até que a melhor transformação seja encontrada.

#### 4.3.5 Máscara da região de interesse

A imagem de RM do cérebro inclui informação de toda caixa craniana e de todas estruturas cerebrais. Para o posicionamento dos modelos dos hipocampos, é interessante que os parâmetros do detector sejam ajustados para permitir a detecção de variados pontos salientes nos entornos da região dessas estruturas. Assim é comum eliminar informação das outras regiões e estruturas imageadas antes do processamento.

Para realizar tal tarefa é comum utilizar máscaras que indiquem as regiões de interesse no processamento, nesse caso, a região dos hipocampos e das suas vizinhanças. A máscara da região de interesse para a detecção foi obtida como resultado da aplicação de uma operação morfológica de dilatação às máscaras da imagem de referência da base de dados NAC que representam as marcações manuais por especialista dos hipocampos, garantindo que tanto a região do hipocampo, quanto regiões vizinhas, sejam consideradas no momento do processamento. O elemento estruturante utilizado nessa operação tem formato esférico e raio de tamanho 12 voxels. O raio foi definido de modo que acomodasse a região do hipocampo em todas as imagens da base de dados.

#### 4.3.6 Extração do encéfalo

Além de incluir toda a caixa craniana, a imagem de RM do cérebro possui uma grande região de *background* de valor constante. No treinamento do modelo probabilístico HMT, essa região de *background* acarreta em uma distribuição dos valores de magnitude com um pico concentrado de coeficientes de valores iguais ou próximos a 0, dificultando o ajuste (*fitting*) de uma

<sup>4</sup><http://sourceforge.net/projects/niftyreg/>

mistura de funções Rayleigh representando coeficientes com magnitude alta e magnitude baixa. Considerando que o intuito da utilização do modelo HMT é a modelagem dos estados *small* (magnitude baixa) e *large* (magnitude alta) dos coeficientes *wavelet*, foi necessário utilizar uma máscara para delimitar a região de processamento contendo as informações de interesse para definir o comportamento dos coeficientes nesses estados. Para esse caso, a região de treinamento do modelo HMT foi definida como a região que engloba todo encéfalo, para que fosse possível ter uma quantidade considerável de coeficientes representando estruturas cerebrais imageadas em todas as escalas de decomposição.

Para extrair a região do encéfalo foi utilizada uma implementação da técnica *Robust Brain Extraction* (ROBEX) (IGLESIAS et al., 2011), que combina um modelo discriminativo e um modelo geral para realizar tal tarefa. O modelo discriminativo é um classificador *Random Forest* treinado, usando 92 imagens de uma base de dados proprietária, para detectar as bordas do encéfalo. O modelo geral é o modelo utilizado quando uma nova imagem é apresentada ao ROBEX, ele atua procurando o contorno com maior probabilidade de ser o contorno do encéfalo de acordo com o modelo discriminativo treinado.

## 4.4 Detecção de pontos salientes

Ao final dessa seção será apresentada a abordagem utilizada para detecção de pontos salientes nesse trabalho. Para uma melhor compreensão da fundamentação teórica e das vantagens do detector, as subseções a seguir apresentam os conceitos básicos relacionados à transformada de Fourier, a DWT e à DT-CWT. O conteúdo das subseções a seguir baseia-se em definições que podem ser encontradas em (GOMES; VELHO; GOLDENSTEIN, 1997; KINGSBURY, 1998; SELESNICK; LI, 2003; SELESNICK; BARANIUK; KINGSBURY, 2005; CHEN; KINGSBURY, 2012).

### 4.4.1 Transformada de Fourier

Um sinal pode ser definido como um conjunto de informações representado por uma função de uma ou mais variáveis. Uma imagem 2D, por exemplo, pode ser interpretada como uma função em duas variáveis  $I(x,y)$  que representa os tons de cinza para todas as combinações de  $(x,y)$  do seu domínio, sendo assim, pode ser vista também como um sinal 2D.

Diversas aplicações necessitam obter informações contidas em um sinal para alcançar sucesso. Os sinais se caracterizam pela sua frequência, dessa maneira, um bom começo para

analisar uma função é iniciar pelo estudo de sua frequência (GOMES; VELHO; GOLDENSTEIN, 1997), que representa o número de oscilações da função (ciclos completos) por unidade de tempo. Partindo desse princípio, o matemático francês Jean Baptiste Joseph Fourier propôs uma formulação matemática para análise de um sinal pelo estudo de sua frequência. Nos seus estudos, Fourier observou que toda função periódica pode ser decomposta em um somatório infinito de funções periódicas mais simples, a nomear, senos e cossenos.

Mais precisamente, se uma função é periódica de período  $T > 0$ , isto é,  $f(t + T) = f(t)$ , a teoria das séries de Fourier diz que ela pode ser reescrita da forma:

$$f(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_0 t}, \quad (4.2)$$

em que  $a_k \in \mathbb{R}$  e  $\omega_0$  é uma constante, e  $j = \sqrt{-1}$  é a unidade imaginária. Pela fórmula de Euler, temos que  $e^{jt} = \cos(t) + j \sin(t)$ .

Essa decomposição do sinal em um somatório infinito de senos e cossenos é chamada de série de Fourier e torna a análise das frequências de uma função trivial; existe uma frequência fundamental  $\omega_0 = \frac{2\pi}{T}$  e todas as outras frequências  $\omega_k = k\omega_0$  são múltiplos inteiros dessa frequência fundamental. A cada componente de frequência  $\omega_k$  é associado um coeficiente  $a_k$  que mede sua amplitude no sinal.

Quando a função não é periódica, não é possível realizar sua decomposição em frequências bem definidas pela série de Fourier. A transformada de Fourier é uma generalização das séries de Fourier para aplicação em sinais não periódicos, realizando a análise do sinal considerando que a periodicidade ocorre no infinito. Dessa maneira, a formulação se torna

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt, \quad (4.3)$$

sendo que  $F(\omega)$  mede a densidade da frequência  $\omega$  no sinal, e o componente  $e^{-j\omega t}$  é uma função exponencial moduladora. Quando a função  $f$  possui oscilações de frequência  $\omega$ , ou próximas de  $\omega$ , essas frequências entram em ressonância com a frequência da função moduladora e  $F(\omega)$  assume valores não nulos, sendo assim,  $F(\omega)$  mede a densidade da frequência  $\omega$  na função  $f$ .

Apesar de oferecer uma ferramenta robusta para a análise das frequências de um sinal, a transformada de Fourier não permite determinar a localização dessas frequências. Se uma função apresenta mudanças bruscas, as altas frequências relativas a essas mudanças são detectadas pela transformada, no entanto, ela é computada ao longo de todo domínio. Isso ocorre pois a função moduladora (exponencial), utilizada para medir a densidade de frequência, não possui

suporte compacto, visto que a integral estende-se ao infinito (GOMES; VELHO; GOLDENSTEIN, 1997). Desse modo, a transformada de Fourier permite apenas determinar se uma dada componente de frequência está presente, ou não, no sinal, mas é impossível localizá-la, o que a torna ineficiente para análise de sinais não estacionários.

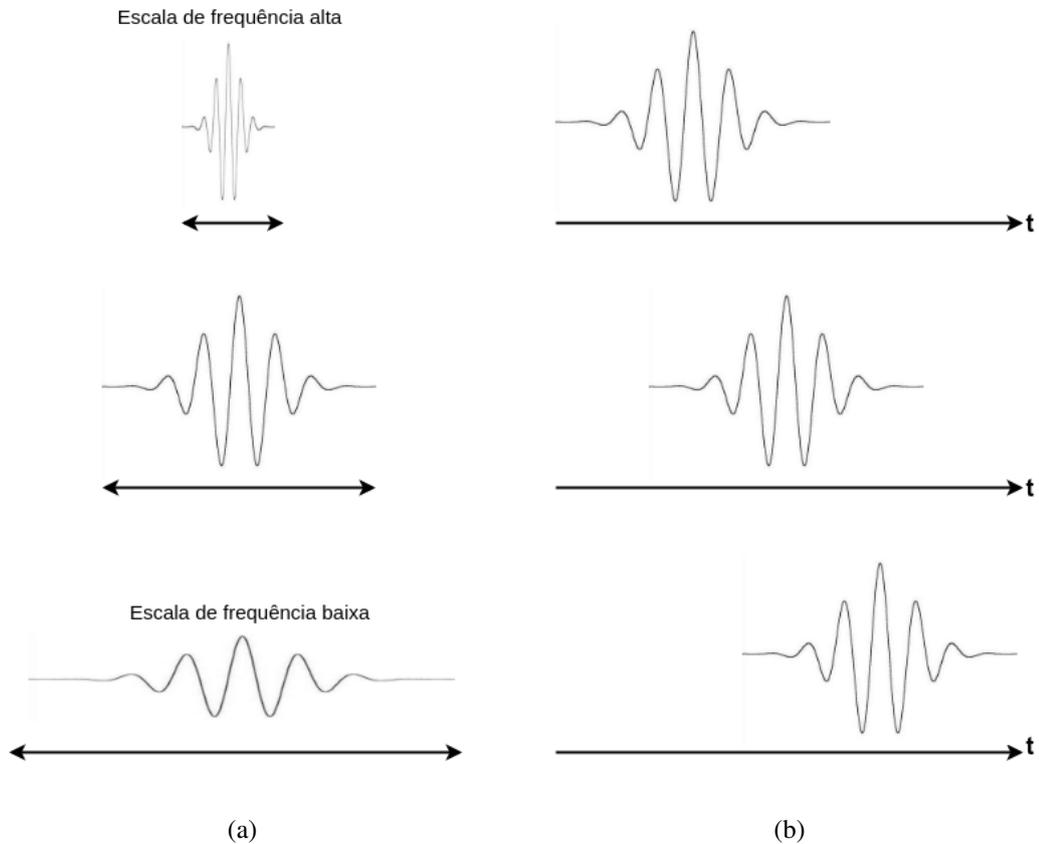
#### 4.4.2 Transformada *wavelet*

A transformada *wavelet* é uma ferramenta que permite a decomposição de uma função em funções mais simples (GOMES; VELHO; GOLDENSTEIN, 1997), de modo que o sinal original possa ser reescrito pela combinação de um conjunto de coeficientes *wavelet* com funções base *wavelet*.

A transformada *wavelet* é uma opção para análise localizada de frequências, pois substitui as funções base sinusoidais infinitas da transformada de Fourier por um conjunto de funções base de oscilação local, chamadas *wavelets* (SELESNICK; BARANIUK; KINGSBURY, 2005). *Wavelets* são versões esticadas e deslocadas de uma função moduladora passa banda de valor real  $\psi(t)$ , chamada de função *wavelet* mãe. Essa função deve atender algumas condições de admissibilidade, a nomear, deve possuir suporte compacto, ou seja, o sinal da função *wavelet* não pode durar infinitamente mas sim oscilar em um curto espaço de tempo e decair rapidamente para zero, e possuir uma área sob a curva da função igual a zero, garantindo que a energia seja distribuída igualmente nas direções positiva e negativa e que a resposta para um sinal constante seja nula.

A Figura 4.6 (a) ilustra a ideia por trás da definição das funções *wavelets* no caso de sinais 1D. As versões escalonadas da função *wavelet* mãe permitem a análise localizada no tempo de diferentes níveis de frequência no sinal, garantindo uma interpretação em diferentes níveis de detalhamento, isto é, escala, variando-a para evitar compromisso com uma escala específica. No caso 1D, a escala é definida pela largura da função de modulação. Uma versão mais esticada permite a análise de informação de baixas frequências, enquanto que uma versão mais comprimida permite a análise da informação de altas frequências. Essas funções são deslocadas avançando no tempo ao longo do sinal para garantir a interpretação do sinal como um todo, como representado pela Figura 4.6 (b).

Formalmente, considerando um sinal  $f(t)$  1D, versões deslocadas de uma função passa baixa de escala  $\phi$ , a transformada *wavelet* discreta - DWT - de  $f(t)$  pode ser definida da seguinte



**Figura 4.6:** Análise multiescala da transformada *wavelet* para um sinal 1D (a). Translação da função *wavelet* ao longo do sinal para garantir a interpretação do sinal como um todo (b).

maneira:

$$f(t) = \sum_{n=-\infty}^{\infty} c(n)\phi(t-n) + \sum_{s=0}^{\infty} \sum_{n=-\infty}^{\infty} d(s,n)2^{s/2}\psi(2^s t - n), \quad (4.4)$$

em que os coeficientes de escala  $c(n)$  e os coeficientes *wavelet* são computados pelos produtos internos:

$$c(n) = \int_{-\infty}^{\infty} f(t)\phi(t-n) dt, \quad (4.5)$$

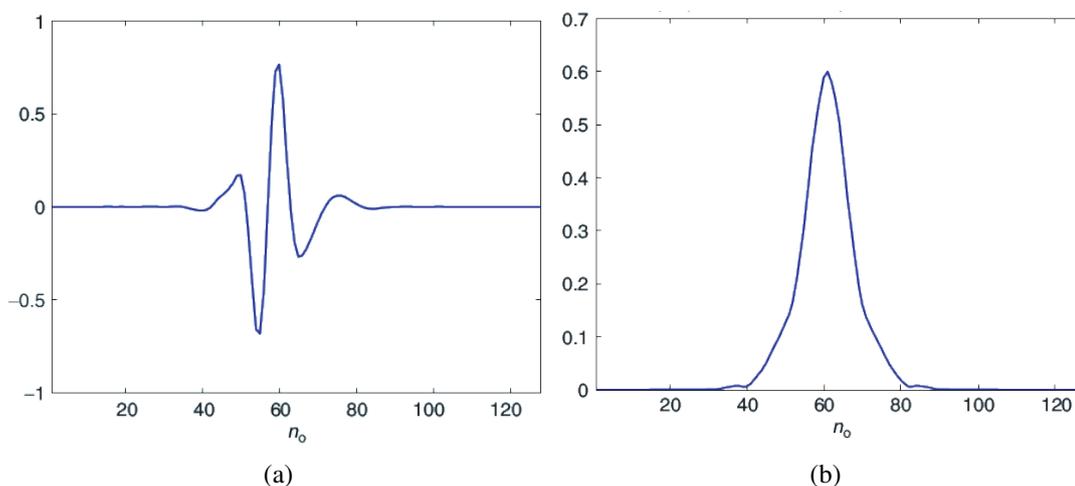
$$d(s,n) = 2^{s/2} \int_{-\infty}^{\infty} f(t)\psi(2^s t - n) dt. \quad (4.6)$$

Eles oferecem uma análise de frequência localizada no tempo medindo o conteúdo de frequência (controlado pelo fator de escala  $s$ ) em tempos distintos (controlado pelo deslocamento  $n$ ).

Apesar da representação eficiente para sinais não estacionários, a DWT sofre de quatro deficiências fundamentais: oscilação no valor dos coeficientes em pontos próximos a singularidades, variância ao deslocamento, *aliasing*<sup>5</sup> e direcionalidade pobre. Tendo em vista que esses problemas da transformada *wavelet* discreta não são observados na Transformada de Fourier, que decompõe o sinal pela combinação de funções base sinusoidais complexas que formam um par de Hilbert ( $\mathcal{H}$ ), Kingsbury (KINGSBURY, 1998) propôs a utilização funções bases *wavelets* complexas para resolver essas deficiências.

### 4.4.3 Transformada *wavelet* complexa de árvore dupla

A DT-CWT foi proposta por Kingsbury (KINGSBURY, 1998) e representa um aprimoramento da transformada DWT, na medida em que fornece informação de magnitude dos coeficientes não oscilante em pontos próximos a singularidades (Figuras 4.7 (a) e (b)), invariância aproximada ao deslocamento do sinal, menor efeito de *aliasing* e maior seletividade direcional. Além disso, a DT-CWT permite que o sinal seja perfeitamente reconstruído a partir de sua representação *wavelet*, possui redundância limitada ( $1 : 2^m$  onde  $m$  é a dimensão do sinal) e ainda possui fase linear.



**Figura 4.7:** Na vizinhança de uma borda, a DWT real produz coeficientes com valores altos e baixos. Em contraste, a DT-CWT aproximadamente analítica produz coeficientes cuja magnitude representa melhor a sua proximidade a uma borda. A Figura (a) representa a resposta da DWT aplicada a uma função degrau, sendo que a transição ocorre em  $n_0$ . A Figura mostra a resposta da DT-CWT aplicada ao mesmo sinal. Adaptado de (SELESNICK; BARANIUK; KINGSBURY, 2005).

Similarmente às funções base sinusoidais de Fourier, uma função *wavelet* complexa é defi-

<sup>5</sup>*Aliasing* ocorre quando não se pode reconstruir um sinal de forma satisfatória, devido a perda de alta frequência, gerando assim artefatos.

nida da seguinte maneira:

$$\psi_c(t) = \psi_h(t) + j\psi_g(t), \quad (4.7)$$

em que  $\psi_h(t)$  representa a parte real (par),  $j\psi_g(t)$  representa a parte imaginária (ímpar) e  $j = \sqrt{-1}$ . É importante definir as funções representando a parte real e imaginária como um par de Hilbert ( $\mathcal{H}$ ) ( $\psi_h(t)$  e  $\psi_g(t)$  devem possuir um deslocamento de fase de  $90^\circ$  entre si) para que  $\psi_c(t)$  seja um sinal analítico e apresente as características da transformada de Fourier.

O coeficiente complexo  $d_c(s, t) = d_h(s, t) + jd_g(s, t)$  tem magnitude definida como

$$|d_c(s, t)| = \sqrt{[d_h(s, t)]^2 + [d_g(s, t)]^2}, \quad (4.8)$$

e fase definida como

$$\angle d_c(s, t) = \arctan\left(\frac{d_h(s, t)}{d_g(s, t)}\right). \quad (4.9)$$

Um coeficiente com valor de magnitude alto representa uma região no sinal com algum tipo de singularidade (SELESNICK; BARANIUK; KINGSBURY, 2005) como, por exemplo, uma transição abrupta dos valores medidos.

A implementação da DT-CWT utiliza uma representação em árvore dupla, com duas transformadas do tipo DWT. A primeira DWT resulta na parte real da transformada complexa e a segunda resulta na parte imaginária. As funções *wavelet* mãe ( $\psi_h(t)$  e  $\psi_g(t)$ ) de cada uma das DWT são projetadas conjuntamente para garantir que uma seja aproximadamente a transformada de Hilbert da outra, ou seja  $\psi_h(t) \approx \mathcal{H}\{\psi_g(t)\}$ , e para garantir que o efeito de *aliasing* em um ramo da primeira árvore (árvore *h*) seja aproximadamente cancelado com o correspondente ramo na segunda árvore (árvore *g*).

A Figura 4.8 ilustra o esquema de decomposição de um sinal unidimensional  $t$  usando a estrutura de árvore dupla da DT-CWT. Embora o esquema corresponda a um sinal unidimensional, ele pode ser generalizado para dimensões maiores.

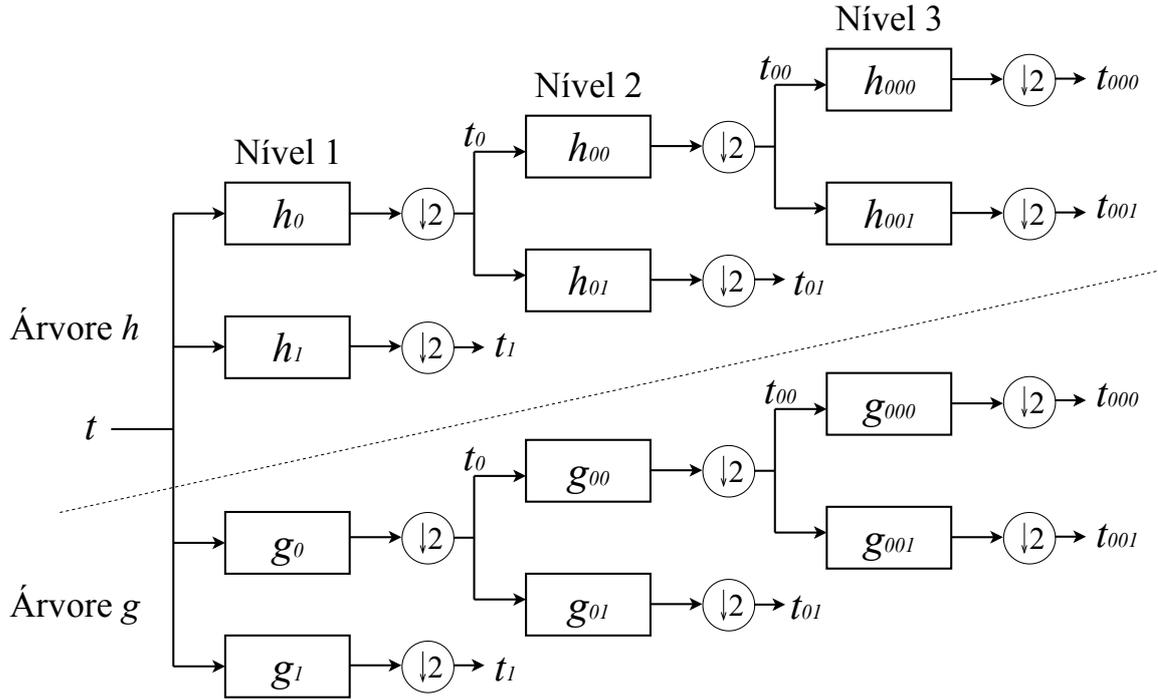


Figura 4.8: Esquema de decomposição da DT-CWT 1D em três níveis usando representação por árvore dupla para um sinal de entrada  $t$  (Adaptação de (KINGSBURY, 1998; SELESNICK; LI, 2003)). Os filtros reais  $h_{*0}$  e  $h_{*1}$  representam o primeiro par dos filtros em quadratura e os filtros reais  $g_{*0}$  e  $g_{*1}$  representam o segundo par. Os filtros com índice de valor final igual a 0 são filtros passa-baixa, enquanto que filtros com índices terminados em 1 representam filtros passa-alta (alguns deles são aplicados na saída de um filtro passa-baixa do nível anterior).

#### 4.4.4 DT-CWT 2D

Na extensão para sinais bidimensionais, a filtragem do sinal  $f(x,y)$  é realizada separadamente ao longo das linhas e das colunas, tendo como base *wavelets* complexas 1D  $\phi, \psi$ :

$$\begin{aligned}\phi(x) &= \phi_h(x) + j\phi_g(x), \\ \psi(x) &= \psi_h(x) + j\psi_g(x).\end{aligned}\quad (4.10)$$

As seis sub-bandas da transformada decompõem o sinal nas direções  $\pm 15^\circ$ ,  $\pm 45^\circ$  e  $\pm 75^\circ$ . Pela combinação das sub-bandas de duas DWTs separáveis (SELESNICK; LI, 2003), define-se:

$$\begin{aligned}\psi_{1,1}(x,y) &= \phi_h(x) \psi_h(y), & \psi_{2,1}(x,y) &= \phi_g(x) \psi_g(y), \\ \psi_{1,2}(x,y) &= \psi_h(x) \phi_h(y), & \psi_{2,2}(x,y) &= \psi_g(x) \phi_g(y), \\ \psi_{1,3}(x,y) &= \psi_h(x) \psi_h(y), & \psi_{2,3}(x,y) &= \psi_g(x) \phi_g(y),\end{aligned}\quad (4.11)$$

em que  $x$  e  $y$  correspondem as linhas e colunas da imagem. Partindo disso, as 6 *wavelets* orientadas são definidas da seguinte maneira:

$$\begin{aligned}\psi_i(x, y) &= \frac{1}{\sqrt{2}} (\psi_{1,i}(x, y) + \psi_{2,i}(x, y)), \\ \psi_{i+3}(x, y) &= \frac{1}{\sqrt{2}} (\psi_{1,i}(x, y) - \psi_{2,i}(x, y)),\end{aligned}\quad (4.12)$$

para  $1 \leq i \leq 3$ .

Considerando uma *wavelet* 2D,  $\psi(x, y) = \psi(x) \psi(y)$ , correspondente à orientação  $-45^\circ$  e associada à implementação linha-coluna da transformada *wavelet*, em que  $\psi(x)$  é uma *wavelet* complexa e aproximadamente analítica, dada por  $\psi(x) = \psi_h(x) + j\psi_g(x)$ , então obtém-se a seguinte expressão:

$$\begin{aligned}\psi(x, y) &= [\psi_h(x) + j\psi_g(x)] [\psi_h(y) + j\psi_g(y)], \\ \psi(x, y) &= \underbrace{\psi_h(x) \psi_h(y) - \psi_g(x) \psi_g(y)}_{\text{Parte real}} + \underbrace{j[\psi_g(x) \psi_h(y) + \psi_h(x) \psi_g(y)]}_{\text{Parte imaginaria}}.\end{aligned}\quad (4.13)$$

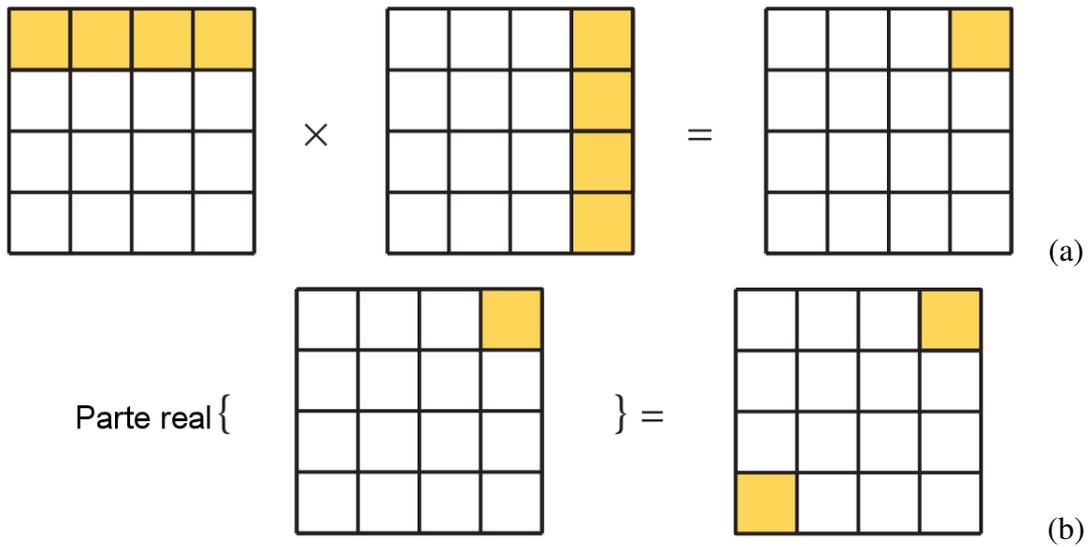
O espectro de Fourier dessa *wavelet* complexa é ilustrado pelo diagrama apresentado na Figura 4.9(a). Já que o espectro da *wavelet* 1D aproximadamente analítica é suportado em apenas um lado do eixo de frequência (SELESNICK; BARANIUK; KINGSBURY, 2005), o espectro da *wavelet* complexa 2D é suportado em apenas um quadrante do plano de frequência 2D. Considerando a parte real, obtém-se a soma de duas *wavelets* separáveis:

$$\text{Parte real}\{\psi(x, y)\} = \psi_h(x) \psi_h(y) - \psi_g(x) \psi_g(y).\quad (4.14)$$

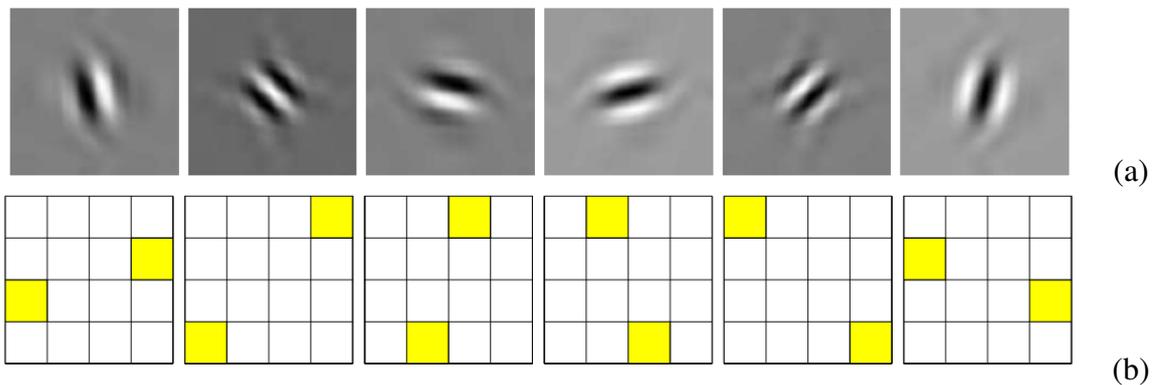
No caso da parte real, o espectro deve ser simétrico em relação à origem. Dessa maneira, o espectro dessa *wavelet* real é suportado em dois quadrantes do plano de frequência 2D, como ilustrado pelo diagrama da Figura 4.9(b).

As seis *wavelets* utilizadas no caso 2D no domínio do espaço são representadas na Figura 4.10(a), enquanto que na Figura 4.10(b) está representado o suporte do espectro de Fourier idealizado de cada uma das *wavelets* no plano de frequência 2D.

Quando comparada à DWT 2D, a DT-CWT 2D apresenta diversas vantagens, pois mantém as propriedades atrativas da DT-CWT 1D além de adicionar uma melhor seletividade direcional. A DWT 2D, por sua vez, permite a análise em apenas 3 orientações pela combinação de funções *wavelet* passa baixa e passa alta (LH) realçando características na direção horizontal, funções *wavelet* passa alta e passa baixa (HL) realçando características na direção vertical e funções *wavelet* passa alta e passa alta (HH) realçando características na diagonal. Dessas orientações, apenas LH e HL são bem definidas. A *wavelet* HH da DWT 2D mistura as orientações  $45^\circ$  e  $-45^\circ$ , não sendo capaz de separá-las, resultando em uma aparência de tabuleiro de



**Figura 4.9:** Espectro de Fourier idealizado para a *wavelet* complexa  $\psi(x,y) = \psi(x)\psi(y)$  (a). Espectro de Fourier idealizado para a parte real (b). Adaptado de (SELESNICK; BARANIUK; KINGSBURY, 2005).

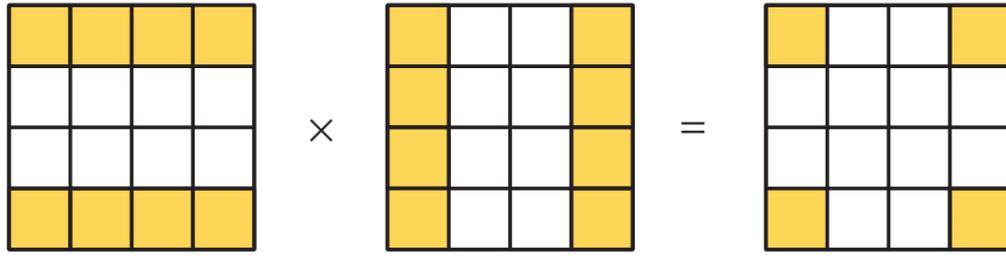


**Figura 4.10:** (a) *Wavelets* no domínio do espaço. (b) Suporte das *wavelets* do espectro de Fourier. Extraído de (SELESNICK; BARANIUK; KINGSBURY, 2005).

xadrez (*checkerboard*). Isso pode ser observado no domínio da frequência. Considerando que  $\psi(x)$  representa uma *wavelet* real e a *wavelet* 2D separável é dada por  $\psi(x,y) = \psi(x)\psi(y)$ , o espectro de Fourier dessa *wavelet* é ilustrado pelo diagrama da Figura 4.11. Como pode ser observado, essa *wavelet* possui passa-bandas nos quatro cantos do espectro, o que explica sua inabilidade em isolar as orientações  $45^\circ$  e  $-45^\circ$ . A DT-CWT 2D não sofre desse problema, pois possui *wavelets* bem definidas e isoladas em 6 orientações.

#### 4.4.5 DT-CWT 3D

De maneira similar ao caso 2D, as *wavelets* complexas da transformada DT-CWT para o caso 3D são definidas como combinações das *wavelets* 1D nos três eixos  $(x,y,z)$  (SELESNICK;



**Figura 4.11:** Espectro de Fourier idealizado para a *wavelet* discreta  $\psi(x, y) = \psi(x) \psi(y)$ . Extraído de (SELESNICK; BARANIUK; KINGSBURY, 2005).

LI, 2003) como

$$\psi(x, y, z) = \psi(x) \psi(y) \psi(z) = [\psi_h(x) + j\psi_g(x)][\psi_h(y) + j\psi_g(y)][\psi_h(z) + j\psi_g(z)]. \quad (4.15)$$

Portanto, para a parte real de  $\psi(x, y, z)$  tem-se:

$$\text{Re}\{\psi(x, y, z)\} = \psi_1(x, y, z) - \psi_2(x, y, z) - \psi_3(x, y, z) - \psi_4(x, y, z), \text{ sendo que}$$

$$\psi_1(x, y, z) = \psi_h(x) \psi_h(y) \psi_h(z) \quad (4.16)$$

$$\psi_2(x, y, z) = \psi_g(x) \psi_g(y) \psi_h(z)$$

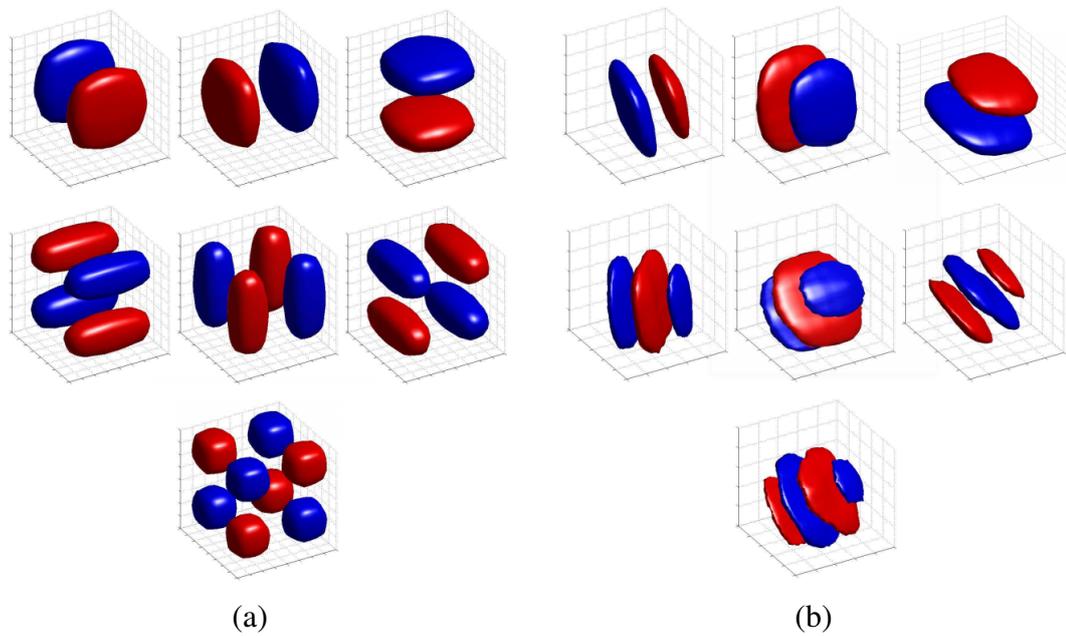
$$\psi_3(x, y, z) = \psi_g(x) \psi_h(y) \psi_g(z)$$

$$\psi_4(x, y, z) = \psi_h(x) \psi_g(y) \psi_g(z).$$

Essas 4 transformadas *wavelet* separáveis,  $\psi_1$ ,  $\psi_2$ ,  $\psi_3$  e  $\psi_4$  são combinadas linearmente gerando uma das árvores DT-CWT contendo uma sub-banda passa-baixa - L - e sete sub-bandas passa-altas - H - (LLL, HLL, LHL, HHL, LLH, HLH, LHH and HHH) (SELESNICK; LI, 2003). Para se obter as sub-bandas restantes, utiliza-se adicionalmente as partes reais de  $\psi(x) \psi(y) \overline{\psi(z)}$ ,  $\psi(x) \overline{\psi(y)} \psi(z)$  e  $\psi(x) \overline{\psi(y)} \overline{\psi(z)}$ , em que  $\overline{\psi(y)}$  representa o conjugado complexo de  $\psi(y)$ . Assim como a DT-CWT 2D, DT-CWT 3D também resolve o problema da mistura de orientações. As Figuras 4.12(a) e (b) ilustram a diferença entre as superfícies representando as funções *wavelet* da DT-CWT 3D e da DWT 3D. A *wavelet* associada à DWT 3D possui o mesmo fenômeno *checkerboard* presente na *wavelet* DWT 2D, consequência da mistura de orientações. Já a *wavelet* associada à DT-CWT 3D é livre desse efeito. Além disso, DT-CWT 3D possui alta seletividade direcional, sendo capaz de decompor os dados de entrada em 28 orientações, que correspondem às 7 orientações passa-alta obtidas em cada uma das DT-CWTs 3D  $\psi(x) \psi(y) \psi(z)$ ,  $\psi(x) \psi(y) \overline{\psi(z)}$ ,  $\psi(x) \overline{\psi(y)} \psi(z)$  e  $\psi(x) \overline{\psi(y)} \overline{\psi(z)}$ . A DWT trabalha com 7 orientações.

A DT-CWT utilizada nesse projeto encontra-se disponível pela biblioteca dtcwt<sup>6</sup> na lin-

<sup>6</sup><https://dctwt.readthedocs.io/en/0.12.0/>



**Figura 4.12:** Isosuperfícies de 7 *wavelets* 3D para uma DWT 3D típica (a). Isosuperfícies de 7 das 28 *wavelets* 3D para uma típica DT-CWT 3D (b). Extraído de (SELESNICK; LI, 2003).

guagem de programação Python. Essa biblioteca fornece as versões 1D, 2D e 3D da DT-CWT com opções de implementação via CPU por meio do uso da biblioteca Numpy e via GPU pela biblioteca OpenCL.

#### 4.4.6 Detector de Fauqueur estendido para 3D

A detecção dos pontos salientes nesse trabalho é realizada a partir da adaptação da técnica proposta no contexto de sinais bidimensionais por Fauqueur et al. (2006) para sinais tridimensionais em (KOVACS; FERRARI, 2017). O detector de Fauqueur faz uso do princípio de que coeficientes *wavelet* representando regiões suaves do sinal possuem valores de magnitude baixos, enquanto que coeficientes que representam singularidades no sinal possuem valores de magnitude altos (ROMBERG et al., 2001). Considerando essa propriedade, a ideia principal do detector é assumir que um ponto saliente corresponde a uma singularidade em muitas orientações de decomposição ao longo das escalas e, desse modo, multiplicando-se (voxel a voxel) todas as sub-bandas orientadas da DT-CWT dentro da mesma escala e acumulando os resultados ao longo das escalas obtém-se um mapa de energia acumulada com uma boa medida de saliência.

A aplicação da DT-CWT em uma imagem com tamanho  $N_x \times N_y \times N_z$  resulta em um conjunto de coeficientes complexos divididos em  $s$  escalas de decomposição. Cada escala  $s$  possui tamanho  $\frac{N_x}{2^s} \times \frac{N_y}{2^s} \times \frac{N_z}{2^s}$  e é subdividida em 28 sub-bandas com seletividade direcional, com  $s$  as-

sumindo valores no intervalo a partir de 1, para a escala mais fina, até o número total de escalas, para a escala mais grosseira.

Para cada localização espacial  $x = (x, y, z)$  em uma dada escala  $s$ , a DT-CWT gera um conjunto  $C$  de 28 coeficientes complexos,  $C = \{\rho_1(x)e^{j\theta_1}, \dots, \rho_{28}(x)e^{j\theta_{28}}\}$ , que correspondem às respostas dos filtros representando as 28 orientações. Então, para essa escala, um mapa de energia é construído a partir do produto, voxel a voxel, da magnitude dos coeficientes complexos da DT-CWT 3D nas 28 orientações da seguinte maneira:

$$E_s(x) = \alpha^s \left( \prod_{k=1}^{28} \rho_k(x) \right)^\beta, \quad (4.17)$$

no qual o parâmetro  $\alpha$  controla o peso relativo das escalas e o parâmetro  $\beta$  diminui o intervalo entre o menor e o maior resultado produzido dentro de uma escala. Definindo valores baixos para esses parâmetros, enfatiza-se escalas mais finas e permite-se que o resultado do produtório não tenha uma variação muito grande entre regiões com coeficientes de baixo valor de magnitude e regiões com coeficientes de alto valor de magnitude. O resultado do produtório em uma escala mais grosseira é combinado com o resultado do produtório em uma escala mais fina depois de uma operação de *upsampling* com fator 2, gerando o mapa de energia acumulada  $A$ :

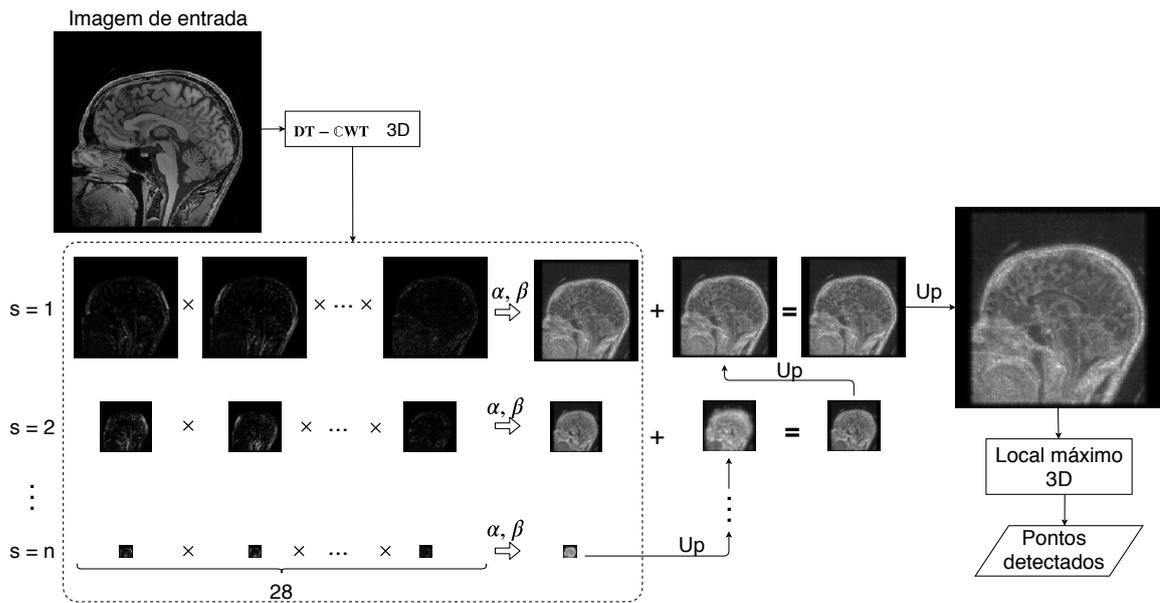
$$A = \sum_{s=1}^m f(E_{s-m+1}), \quad (4.18)$$

em que  $m$  é o número total de escalas e  $f$  representa a função de interpolação para a operação de *upsampling*. A detecção dos pontos salientes é realizada pela procura de máximos locais em  $A$  em uma vizinhança de tamanho  $n \times n \times n$  voxels e, se os valores forem maiores que um limiar  $T_E$ , então a localização do ponto é definida como um ponto saliente. Na Figura 4.13, encontra-se representado de forma gráfica a *pipeline* de execução do detector, dada a DT-CWT de uma imagem.

Considerando que a implementação da DT-CWT 3D encontra-se disponível na linguagem Python, o detector também foi implementado nessa linguagem de programação.

## 4.5 Descrição de pontos salientes

Após a detecção dos pontos salientes, é necessário descrevê-los usando informações extraídas de sua região de localização na imagem, as quais são armazenadas em um vetor que recebe a denominação de descritor. Esse descritor é utilizado para computar a similaridade entre pontos de diferentes imagens, permitindo estabelecer uma relação entre eles. Na etapa de



**Figura 4.13:** Pipeline de execução do detector proposto em uma imagem de RM. A palavra Up representa a operação de interpolação *upsampling*.

correspondência, cada ponto em uma imagem  $A$  é correspondido ao ponto mais similar em uma imagem  $B$ , dado um critério de similaridade (por exemplo, distância Euclidiana ou  $\chi^2$ ) baseado em seus descritores. É de extrema importância que o descritor de um ponto seja altamente discriminativo e sucinto, de forma que pontos correspondentes tenham descrições próximas e pontos não correspondentes possuam descrições suficientemente distintas. Nessa seção será apresentada a fundamentação teórica relativa à técnica *shape context* (BELONGIE; MALIK; PUZICHA, 2002; MORTENSEN; DENG; SHAPIRO, 2005), que é utilizada para descrever os pontos detectados nesse projeto.

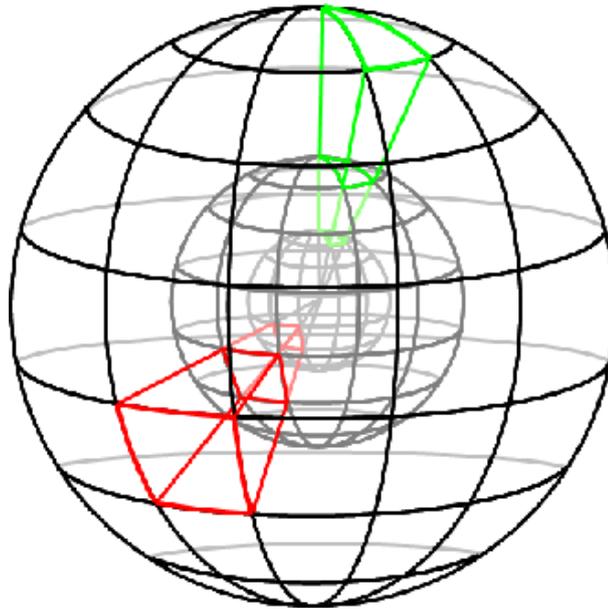
### 4.5.1 Técnica *Shape Context*

O método utilizado nesse trabalho para a descrição dos pontos salientes é baseado na técnica *shape context* (BELONGIE; MALIK; PUZICHA, 2002; MORTENSEN; DENG; SHAPIRO, 2005), cuja abordagem original para imagens 2D considera que a forma de um objeto pode ser capturada por um conjunto finito de pontos amostrados de seus contornos internos ou externos. Para um ponto qualquer  $x$  extraído da forma, computa-se um histograma grosseiro  $h$ , definido como o *shape context* de  $x$ , cujos *bins* armazenam o número de pontos de borda amostrados em um espaço log-polar, permitindo que o descritor seja mais sensível a pontos amostrados em regiões mais próximas do que daqueles mais distantes.

O descritor *shape context* foi adaptado para imagens 3D de RM tendo como base o trabalho de Villa-Pinto e Ferrari (VILLA-PINTO; FERRARI, 2016), que utiliza pontos provindos de um

mapa de bordas gerado a partir de uma relação dos maiores autovalores da matriz de momentos de segunda ordem do domínio da congruência de fase computada para cada ponto da imagem. No ajuste para o âmbito desse projeto de pesquisa, o mapa de bordas utilizado é obtido a partir do mapa de energia acumulada gerado pelo produtório dos coeficientes *wavelet*. Esse mapa de energia, assim como o mapa de autovalores da proposta original, consegue expressar simultaneamente informação de forma e significância das características, pois ele representa a magnitude acumulada em diferentes escalas e orientações de decomposição.

O espaço ao redor do ponto saliente é amostrado para gerar um histograma log-esférico em termos de  $\eta_r$  distâncias radiais,  $\eta_a$  ângulos azimutes relativos e  $\eta_e$  ângulos de elevação relativos. Na Figura 4.14, encontra-se uma representação gráfica da repartição do espaço log-esférico em setores que representam os *bins* do histograma. Assim como no trabalho de Villa-Pinto e Ferrari (2016), os parâmetros escolhidos para a criação do histograma para cada ponto saliente são  $\eta_r = 5$ ,  $\eta_a = 12$  e  $\eta_e = 6$ , gerando um descritor com 360 posições, em que cada posição representa a informação de borda contida no respectivo *bin*.



**Figura 4.14:** Representação gráfica da repartição do espaço log-esférico ao redor de um ponto saliente para o cômputo do descritor *shape context*. Extraído de (TOMBARI; SALTI; STEFANO, 2010).

As equações para determinar o *bin* (0 a 359) do histograma ao qual um ponto de borda  $x = (x, y, z)$  localizado no espaço log-esférico da vizinhança de um ponto saliente  $x_0 = (x_0, y_0, z_0)$  são as seguintes:

$$b_r = \max \left( 0, \eta_r + \left\lceil \log_2 \left( \frac{\|x - x_0\|}{R_{max}} \right) \right\rceil \right), \quad (4.19)$$

$$b_a = \text{mod} \left( \left\lceil \frac{\eta_a(\pi + \phi)}{2\pi} \right\rceil, \eta_a \right), \quad (4.20)$$

$$b_e = \begin{cases} \text{mod} \left( \left[ \frac{\eta_e(\frac{\pi}{2} + \theta)}{\pi} \right], \eta_e \right) & \text{se } -\frac{\pi}{2} \leq \theta < \frac{\pi}{2}, \\ \eta_e - 1 & \text{se } \theta = \frac{\pi}{2} \end{cases}, \quad (4.21)$$

$$b = b_r + \eta_r(b_a + \eta_a b_e), \quad (4.22)$$

em que

$$\phi = \arctan \left( \frac{y - y_0}{x - x_0} \right), \quad (4.23)$$

e

$$\theta = \arctan \left( \frac{z - z_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \right), \quad (4.24)$$

são, respectivamente, os ângulos relativos de azimute e elevação entre o ponto  $x$  e o ponto  $x_0$ , e  $R_{max}$  é a distância radial máxima que pode ser definida para reduzir o número de pontos de borda processados. O descritor final é normalizado para ter norma unitária.

O descritor shape context proposto encontra-se implementado na linguagem de programação C++.

## 4.6 Casamento de os pontos salientes

A correspondência entre os pontos salientes em duas imagens é realizada usando a distância  $\chi^2$  como medida de dissimilaridade. Assim, se  $h_1$  e  $h_2$  são dois descritores computados pela técnica *shape context* descrita anteriormente, a distância  $\chi^2$  entre eles é definida como

$$\chi^2(h_1, h_2) = \frac{1}{2} \sum_{b=0}^B \frac{\left( h_1^{(b)} - h_2^{(b)} \right)^2}{h_1^{(b)} + h_2^{(b)}}, \quad (4.25)$$

onde  $h_1^{(b)}$  e  $h_2^{(b)}$  são, respectivamente, o  $b$ -ésimo componente normalizado de  $h_1$  e de  $h_2$ , e  $B$  representa o número total de *bins* (nesse caso, 360). Essa correspondência é mantida apenas se existir reciprocidade entre os casamentos dos pontos da imagem *template* e da imagem clínica, ou seja, somente se o descritor mais próximo de  $h_1$  (na imagem *template*) for  $h_2$  (na imagem clínica) e, ao mesmo tempo, o descritor mais próximo de  $h_2$  for  $h_1$ .

## 4.7 Modelo probabilístico *Hidden Markov Tree*

A transformada *wavelet* de um sinal resulta em um conjunto de coeficientes com duas propriedades chaves: regiões suaves no sinal são representadas por coeficientes *wavelet* de

magnitude pequena, enquanto que singularidades são representadas por coeficientes *wavelet* de magnitude grande; e a intensidade dos valores dos coeficientes tende a persistir ao longo das escalas de decomposição do sinal (coeficientes filhos tendem a ter mesma intensidade de valor de coeficiente do coeficiente pai).

Esses coeficientes podem ser representados por uma estrutura em árvore, em que os coeficientes da escala mais grosseira podem ser vistos como coeficientes raiz e os coeficientes da escala mais fina podem ser vistos como coeficientes folhas. Para cada coeficiente em uma escala mais grosseira (pai), existem  $2^d$  coeficientes (filhos) equivalentes na escala mais fina adjacente, sendo  $d$  a dimensão do sinal. A Figura 4.15 demonstra a estrutura em árvore resultante da transformada *wavelet* de uma imagem 2D, sendo que os coeficientes equivalentes são representados pelas mesmas cores nas diferentes escalas. No caso da transformada *wavelet* 3D, cada coeficiente em uma escala mais grosseira possui 8 coeficientes equivalentes na escala mais fina adjacente.

Essas características da transformada *wavelet* são de grande interesse para modelagem estatística. Romberg et al. (2001) propuseram uma adaptação para *wavelets* complexas do modelo *Hidden Markov Tree* (HMT) proposto para *wavelets* discretas por Crouse, Nowak e Baraniuk (1998). Esse modelo busca capturar a estrutura estatística da decomposição *wavelet* modelando as duas propriedades chaves citadas, para determinar o estado (variável aleatória escondida) que a magnitude do coeficiente *wavelet* irá assumir. A primeira propriedade é modelada por uma mistura de densidades de probabilidade Rayleigh<sup>7</sup>, representando os estados *small* (coeficientes representando regiões suaves do sinal) ou *large* (coeficientes representando singularidades), parametrizada por  $\sigma$ :

$$pdf(w, \sigma) = \frac{w}{\sigma^2} e^{-w^2/(2\sigma^2)}, w \geq 0, \quad (4.26)$$

em que  $w$  representa a magnitude do coeficiente *wavelet*. A segunda propriedade é modelada pela propriedade de Markov<sup>8</sup> aplicada à estrutura de árvore resultante da transformada *wavelet*, por meio da definição de probabilidades de transição do estado do coeficiente pai para o estado dos coeficientes filhos  $A_{m,n} = P(Z_l = n | Z_{pai(l)} = m)$ , em que  $Z$  é a variável aleatória representando o estado,  $m$  representa o estado do coeficiente pai,  $n$  representa o estado do coeficiente filho,  $l$  representa o coeficiente atual e  $pai(l)$  determina o coeficiente pai do coeficiente atual.

<sup>7</sup>Quando se considera a DT-CWT, a magnitude dos coeficientes complexos possui a propriedade de seguir uma mistura de densidade de probabilidade Rayleigh. No caso da DWT, os coeficientes se distribuem seguindo uma mistura densidade de probabilidade Gaussiana.

<sup>8</sup>A propriedade de Markov diz que a probabilidade condicional de estados futuros de um processo estocástico (condicional tanto em estados passados, como presentes) depende apenas do estado presente, e não da sequência de estados que o precedeu.

Como não existe transição para a escala raiz da árvore, determina-se uma distribuição inicial  $\pi(i) = p(Z_0 = i)$ , onde  $i$  (estado)  $\in \{small, large\}$ . Dessa maneira, a probabilidade conjunta da magnitude dos coeficientes ( $w$ ) e dos estados ( $Z$ ) pode ser descrita da seguinte maneira:

$$p(w_0, \dots, w_L, Z_0, \dots, Z_L) = \pi(Z_0) pdf(w_0, \sigma_{Z_0}) \prod_{l=1}^L P(Z_l | Z_{pai(l)}) pdf(w_l, \sigma_{Z_l}), \quad (4.27)$$

em que  $L$  representa o número de elementos na sequência.

Sendo assim, o modelo HMT é totalmente parametrizado pela probabilidade inicial, pelas probabilidades de transição e pelos parâmetros  $\sigma$  representando as distribuições Rayleigh dos estados em cada sub-banda de decomposição. O algoritmo utilizado para encontrar esses parâmetros é uma adaptação do algoritmo *Backward-Forward* utilizado nos modelos escondidos de Markov para sua aplicação em árvores, denominado *Upward-Downward* e descrito em (DURAND; Goncalvès; Guédon, 2004). Esse algoritmo é iterativo e segue uma estrutura baseada no algoritmo *Expectation Maximization* (EM) para determinar os parâmetros.

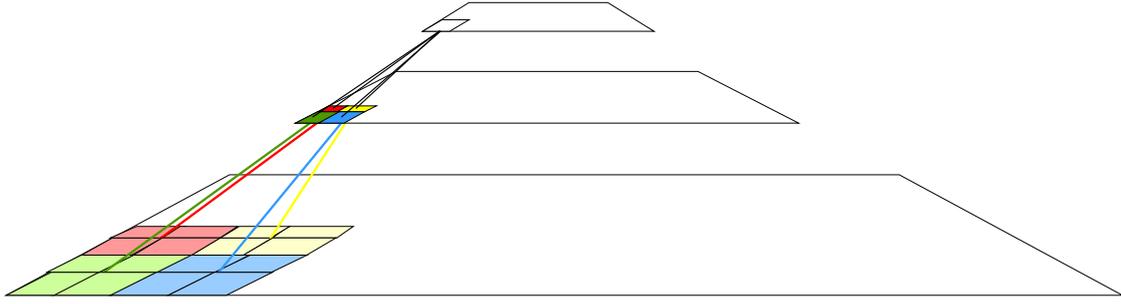
A recuperação da sequência de estados segue também o algoritmo proposto em (DURAND; Goncalvès; Guédon, 2004). Esse algoritmo emula o algoritmo de Viterbi no contexto de cadeias escondidas de Markov (*Hidden Markov Chain* - HMC) para a restauração da árvore de estados escondidos no contexto de HMT. O principal objetivo do algoritmo é encontrar os estados que compõem a árvore escondida que maximizam a probabilidade conjunta de todas as variáveis visíveis (coeficientes *wavelet*) e escondidas (estados).

Diferentemente do algoritmo de treinamento do modelo, que necessita de várias iterações percorrendo todos coeficientes organizados em formato de árvore, das folhas a raiz e da raiz as folhas, até que seja atingida a convergência, o algoritmo de recuperação de estados necessita apenas de uma iteração *Upward-Downward*, tornando a classificação dos estados mais simples e rápida.

A informação de persistência oferecida pelo modelo probabilístico HMT foi utilizada para verificar se a remoção de pontos não persistentes em muitas orientações poderia ajudar no posicionamento inicial dos modelos deformáveis.

## 4.8 Modelo deformável *Simplex Mesh*

Modelos deformáveis podem ser descritos como materiais deformáveis que possuem equações associadas modelando o comportamento de curvas não rígidas, superfícies e sólidos em função do tempo (TERZOPOULOS et al., 1987). São modelos ativos que respondem de ma-



**Figura 4.15: Representação em forma de árvore da decomposição *wavelet* 2D.**

neira natural as forças aplicadas e as restrições associadas. A maioria das estratégias utilizadas para o ajuste desse tipo de modelo passa pela otimização de funções objetivo (equações associadas) procurando encontrar um compromisso entre um termo de energia baseado na imagem e outro termo relacionado com a energia interna do modelo (MCINERNEY; TERZOPOULOS, 1996).

Nesse projeto foi aplicada a técnica de ajuste de modelos deformáveis *Simplex Mesh* (DELINGETTE, 1994) aos modelos posicionados. Formalmente, um modelo *k-simplex mesh* é definido como um par  $(V(\mathcal{M}), N(\mathcal{M}))$ , onde  $V(\mathcal{M})$  é um conjunto de vértices  $p_i$  e  $N(\mathcal{M})$  é uma função específica de conectividade dos vértices.

Cada vértice de uma *k-simplex mesh* é conectado a exatamente  $k+1$  vértices vizinhos. Seguindo essa definição, uma *1-simplex mesh* é um contorno deformável representado por uma linha poligonal e uma *2-simplex mesh* é uma superfície deformável na qual cada vértice encontra-se conectado a exatamente três vértices vizinhos. Os modelos posicionados, originalmente na forma de malhas triangulares, são convertidos para a representação *2-simplex mesh*, representação dual à representação triangular, antes de realizar o ajuste. Na Figura 4.16 é demonstrada uma malha do hipocampo esquerdo em suas representações triangular e *simplex*.

A equação de deformação de uma malha *simplex* é modelada por funções de estabilização locais que controlam a curvatura média da malha. A posição de um vértice  $p_i$ , dentro de uma iteração no tempo  $t$ , é definida pela seguinte equação:

$$p_i^{t+1} = p_i^t + (1 - \gamma)(p_i^t - p_i^{t-1}) + \alpha_{sm} \mathbf{f}_{\text{int}}(p_i^t) + \beta_{sm} \mathbf{f}_{\text{ext}}(p_i^t), \quad (4.28)$$

em que  $\mathbf{f}_{\text{int}}$  e  $\mathbf{f}_{\text{ext}}$  são funções que representam as ações de uma força interna e uma força externa ao modelo. Os parâmetros  $\alpha_{sm}$  e  $\beta_{sm}$  controlam o peso da aplicação da força interna e externa ao modelo, respectivamente, e o parâmetro  $\gamma$  de amortecimento influencia na regularidade da malha. A força interna é computada pela minimização de um critério local baseado em métricas

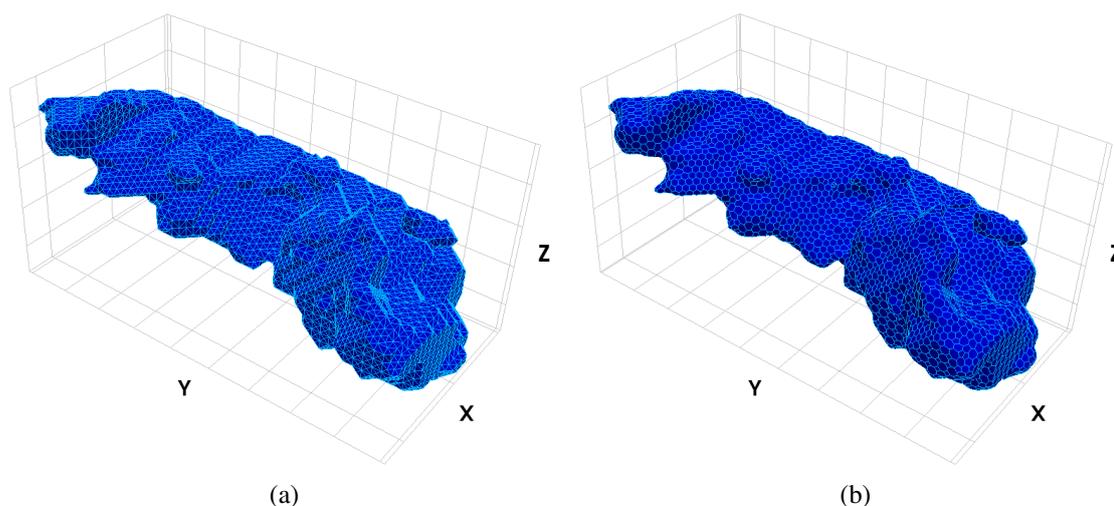


Figura 4.16: Malha do hipocampo esquerdo em suas representações triangular (a) e *simplex* (b).

extraídas da relação da posição entre o vértice sendo analisado e seus vizinhos. Já a força externa restringe a proximidade do encaixe do modelo aos dados tridimensionais.

O ajuste do modelo utilizando a técnica *Simplex Mesh* ocorre por meio de um processo iterativo que atualiza a posição de cada vértice da malha com base na equação 4.28. Para isso, é necessário determinar o valor dos parâmetros da equação e o número total de iterações. O cômputo da força externa é realizado com base na informação de gradiente obtida da imagem na qual se deseja ajustar o modelo deformável. Devido ao acúmulo de informações de intensidade multiescala, o mapa de energia gerado pela combinação dos coeficientes da DT-CWT descrito na subseção 4.4.6 é utilizado como fonte para o cômputo do gradiente utilizado na energia externa. O algoritmo da técnica de deformação *Simplex Mesh* utilizado nesse projeto encontra-se disponível por meio da biblioteca ITK.

## 4.9 Métodos de avaliação

Nessa seção são descritas as métricas e os experimentos para a avaliação do detector, do descritor, e da metodologia de posicionamento propostos.

### 4.9.1 Métricas

**Teste estatístico U de Mann-Whitney** Na presença de poucas amostras é difícil determinar a eficácia dos resultados de algum estudo ou técnica baseando-se apenas nas amostras disponíveis. Não existe garantia de que o comportamento obtido nesse pequeno conjunto de amostras

se repetirá na população que essas amostras representam se o conjunto não for grande o suficiente e determinado de acordo com uma distribuição normal. Entretanto, mesmo que poucas amostras possam ser metodologicamente questionáveis, pois a generalização é difícil, elas podem ser úteis para inferir conclusões em uma população se o teste estatístico adequado for aplicado (NACHAR, 2008).

O teste U de Mann-Whitney (MANN; WHITNEY, 1947) é um teste não paramétrico (independente da estrutura de distribuição dos dados) e pode ser usado para responder questões relativas à diferenças entre grupos de amostras, mesmo que esses grupos possuam poucos exemplos. Nesse teste, assume-se a hipótese nula ( $H_0$ ) que estipula que os dois grupos vêm da mesma população, ou seja, ela diz que grupos independentes são homogêneos e são determinados pela mesma distribuição (NACHAR, 2008).

O teste fornece a informação estatística U, que é calculada para ambos os grupos, e seu valor mínimo mostra o quanto existe de sobreposição nos valores dos dois grupos. Dessa maneira, é possível determinar o grau de diferença entre os valores dos grupos. Quanto maior o valor da estatística U, menor a diferença entre os grupos. Quanto menor esse valor, maior a diferença entre os valores dos dois grupos. Seguido do cálculo da estatística U, e dado um nível de significância (que pode assumir valores 0,01 e 0,05), a hipótese nula ( $H_0$ ) pode ser rejeitada ou não, pela análise da tabela de Mann-Whitney, se o *p-value* correspondente ao valor mínimo de U for menor que o nível de significância.

**Métrica de Dice** A métrica de Dice é utilizada para verificar a sobreposição obtida na comparação de dois conjuntos de dados, no caso desse projeto, na comparação entre duas imagens binárias, uma representando a malha posicionada e a outra representando a demarcação por especialista. O resultado da medida DC varia entre 0 e 1, sendo que, quanto mais próximo do valor 1, melhor a sobreposição dos modelos. A métrica DC é definida da seguinte maneira:

$$DC(A, B) = \frac{2|A \cap B|}{|A| + |B|}, \quad (4.29)$$

em que A e B representam os conjuntos comparados.

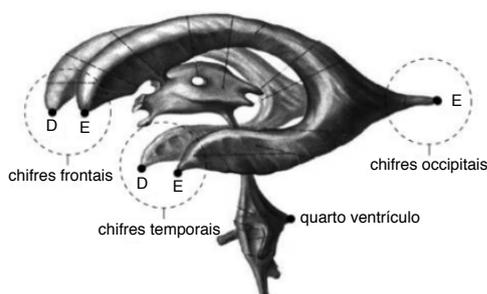
**Distância média de Hausdorff** Para computar a distância média de Hausdorff (HAD), primeiramente computa-se a distância entre o conjunto de voxels não nulos de duas imagens usando a seguinte fórmula:

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|, \quad (4.30)$$

em que  $A$  e  $B$  representam os conjuntos de voxels não nulos das imagens de entrada, que normalmente são imagens binárias. A métrica HAD, então, é definida como a média de todas as distâncias mínimas encontradas nesse comparação. A interpretação da métrica HAD segue a seguinte premissa: quanto mais próximo de 0, mais próximos os conjuntos de dados comparados se encontram.

## 4.9.2 Avaliação do detector

Idealmente, um detector de pontos salientes deve ser invariante a um conjunto de transformações que podem estar presentes em diferentes imagens de um mesmo objeto, sendo as mais comuns as transformações de escala, rotação e translação. Além dessa propriedade, no contexto de imagens de RM do cérebro, é importante que dentre os pontos detectados, o detector seja capaz de localizar pontos salientes anatômicos 3D (WÖRZ; ROHR, 2005), tais como os pontos exemplificados na Figura 4.17, que ilustram o sistema ventricular do cérebro humano.



**Figura 4.17:** Exemplos de pontos salientes anatômicos no sistema ventricular do cérebro. Adaptação de (WÖRZ; ROHR, 2005).

Considerando as propriedades desejáveis mencionadas, para a análise da eficiência do detector foram realizadas duas avaliações principais. A verificação da invariância do detector às transformações citadas foi realizada utilizando um conjunto de imagens geradas a partir da imagem de um cubo. Esse conjunto inclui uma versão rotacionada, transladada e escalonada da imagem original.

A análise da eficiência do detector para a localização de pontos salientes anatômicos 3D foi realizada utilizando a base de dados de estruturas sintéticas descrita na subseção 4.2.1. O teste estatístico U de Mann-Whitney (MANN; WHITNEY, 1947) foi utilizado para comparar a distância da localização entre o ponto detectado e a localização desejada para cada um dos subgrupos degradados (com ruído, *bias field* ou ambos) em relação ao subgrupo sem degradação (estritamente sem ruído e sem *bias field*). Além disso, para os diferentes subgrupos, foi

realizada uma contagem do número de imagens do subgrupo nas quais o detector não foi capaz de detectar ponto saliente em uma dada distância máxima (Falso Negativo - FN) e o número de pontos salientes que foram detectados de forma errada (Falso Positivo - FP) dentro desse subgrupo. Para essa avaliação, o ponto detectado mais próximo da posição desejada a uma distância Euclidiana máxima de 10 voxels é considerado como a detecção correta (Verdadeiro Positivo - VP). Se a detecção ocorre fora da região definida por essa distância, o ponto detectado é considerado um FP. Se mais de um ponto é detectado dentro dessa região, esses pontos também são considerados FPs. Se nenhum ponto é detectado dentro dessa região, têm-se um FN.

### 4.9.3 Avaliação do descritor

Para a avaliação do descritor proposto foi gerado um conjunto de 50 imagens moderadamente deformadas pela aplicação de transformações deformáveis aleatórias na imagem *template* da base de dados NAC. Nessa imagem, foram detectados pontos salientes usando parâmetros que permitiram um bom número de pontos na região do hipocampo esquerdo (usando a máscara definida para essa região). Os mesmos mapas de deformação foram aplicados às coordenadas dos pontos detectados para definir suas localizações equivalentes em cada uma das imagens deformadas. Com essa informação foi possível definir um conjunto de casamentos esperados para todos os pontos salientes nas imagens deformadas em relação aos pontos da imagem *template* NAC. Para cada ponto em cada imagem o descritor equivalente foi computado e, depois, casamentos foram procurados entre os pontos de cada imagem deformada e os pontos da imagem *template* NAC.

Na análise do descritor foi utilizada a Medida-F1 (POWERS, 2011), que pode ser definida da seguinte maneira:

$$F1 = \frac{2VP}{2VP + FP + FN}, \quad (4.31)$$

em que  $VP$  representa os verdadeiros positivos (casamentos esperados que ocorreram de forma correta),  $FP$  representa os falsos positivos (casamentos entre pares de pontos não equivalentes) e  $FN$  representa os falsos negativos (quantidade de casamentos que não ocorreram). O valor do raio do espaço log-esférico  $R_{max}$  para a construção do descritor foi variado para determinar seu melhor valor no contexto de posicionamento de modelos deformáveis.

É importante salientar que os pontos para cada imagem deformada são definidos por meio de uma transformação aplicada aos pontos detectados na imagem *template* NAC, e, dessa maneira, todas as imagens possuem a mesma quantidade de pontos.

#### 4.9.4 Avaliação do posicionamento das malhas dos hipocampus

A metodologia de posicionamento foi avaliada utilizando as métricas de Dice e HAD medindo, respectivamente, a sobreposição e a distância média entre o modelo posicionado e os GTs demarcados por especialista das imagens clínicas da base EADC-ADNI. Para isso, os modelos posicionados são primeiramente convertidos para uma imagem binária e então comparados com as imagens binárias representando os GTs.

Com essas métricas foram realizadas comparações entre os resultados obtidos pelo posicionamento usando técnica baseada na DT-CWT e os resultados obtidos por duas outras técnicas, posicionamento Afim e posicionamento guiado pela técnica PC 3D (FERRARI et al., 2011; VILLA-PINTO; FERRARI, 2016; POLONI et al., 2018). As comparações foram realizadas considerando o comportamento na base de imagens clínicas EADC-ADNI como um todo e também considerando o comportamento em diferentes estratificações da base (idade, doença). O teste estatístico U de Mann-Whitney também é utilizado para avaliar as diferenças na comparação dos grupos com os valores das métricas Dice e HAD definidos pelas diferentes técnicas.

Uma das técnicas comparadas, PC 3D, utiliza a informação de fase computada a partir das informações obtidas pela filtragem da imagem original por um conjunto de filtros log-Gabor. Pontos salientes são detectados em regiões nas quais os componentes de Fourier do domínio da frequência da imagem encontram-se maximamente em fase. Para uma comparação equivalente, a mesma metodologia de posicionamento proposta usando a técnica DT-CWT foi aplicada utilizando o detector PC 3D, com o descritor assumindo os mesmos parâmetros e sendo computado a partir da imagem representando os maiores autovalores da análise da matriz de segunda ordem gerada para detectar os pontos salientes (VILLA-PINTO; FERRARI, 2016).

#### 4.9.5 Avaliação da aplicabilidade do modelo HMT e do algoritmo de deformação

O modelo probabilístico HMT foi empregado para a seleção, dentre os pontos detectados pela técnica proposta que detecta erroneamente alguns pontos não interessantes, de pontos salientes persistentes em todas as escalas de decomposição considerando um determinado número de orientações. Para isso, a imagem *template* da base NAC foi utilizada na etapa de treinamento para estimar os parâmetros que definem o modelo. Tendo os parâmetros do modelo definido, a sequência de estados de todas as imagens clínicas e da imagem *template* foram recuperadas usando o algoritmo proposto por Durand, Goncalvès e Guédon (2004).

Considerando um número  $s$  de escalas de decomposição (a mesma quantidade de escalas

usada para a detecção), o modelo HMT foi usado para verificar quais dos pontos detectados pelo detector proposto possuíam persistência no estado *large* nos coeficientes equivalentes em todas as escalas de decomposição dado um número mínimo de orientações em que essa característica deveria ocorrer. Por exemplo, se  $s$  assumir valor 3 e for determinado que a persistência em todas as escalas deve ocorrer em pelo menos 5 orientações, foram mantidos apenas os pontos que tiveram os coeficientes equivalentes classificados no estado *large* nas escalas 1, 2 e 3 em 5 ou mais orientações. Considerando um ponto detectado com coordenadas (125, 135, 140), em uma imagem de tamanho  $256 \times 256 \times 256$ , a escala de decomposição DT-CWT mais fina ( $s = 1$ ) terá tamanho  $128 \times 128 \times 128$  e o coeficiente equivalente a esse ponto estará localizado na posição (62, 67, 70), cujas coordenadas são encontradas considerando a parte inteira da divisão por 2 das coordenadas na escala adjacente. O mesmo processo é repetido para encontrar as posições equivalentes nas outras escalas e verificar o estado associado com o coeficiente localizado nessa posição.

A DT-CWT decompõe a imagem em 28 orientações e para a avaliação do HMT como seletor de pontos, o número de orientações foi variado e foi verificado se a sobreposição final pela métrica de Dice dos modelos representando os hipocampos para a base de dados EADC-ADNI aumentava com a exclusão de pontos.

A avaliação da técnica de ajuste *Simplex Mesh* também foi realizada utilizando a métrica de Dice para verificar a média de sobreposição obtida para a base de dados EADC-ADNI após a aplicação do ajuste nos modelos posicionados. O mapa de energia acumulada foi usado como fonte para o cálculo do gradiente usado na definição da força externa dessa técnica.

## 4.10 Considerações Finais

Este capítulo apresentou a fundamentação teórica das técnicas utilizadas nessa pesquisa, desde o pré-processamento até o posicionamento dos modelos dos hipocampos em imagens de RM. A DT-CWT é uma ferramenta que permite a análise localizada de frequências em um sinal e representa um aprimoramento da DWT, já que fornece informação de magnitude e fase dos coeficientes, invariância aproximada ao deslocamento do sinal, menor efeito de *aliasing* e melhor seletividade direcional. Essas características se devem ao projeto dos filtros da DT-CWT que busca emular as propriedades analíticas da transformada de Fourier.

A técnica de detecção proposta, baseada na DT-CWT da imagem, é uma técnica relativamente simples, já que depende apenas de operações de produto, soma e interpolação. A técnica proposta para a descrição dos pontos salientes já foi utilizada com sucesso em outra abordagem

de inicialização de modelos deformáveis (VILLA-PINTO; FERRARI, 2016) e por isso foi escolhida para ser aplicada nesse projeto. Entretanto, diferente do trabalho de Villa-Pinto e Ferrari (2016), que utiliza a informação de congruência de fase, nesta pesquisa a técnica *shape context* é calculada a partir dos coeficientes da DT-CWT.

Nesse capítulo também foram apresentados o modelo probabilístico HMT e a técnica de ajuste *Simplex Mesh*. O modelo HMT oferece uma solução para a modelagem estatística dos coeficientes *wavelet* por uma relação de dependência ao longo das escalas e essa informação foi usada para verificar se a seleção de pontos persistentes poderia ser útil para o posicionamento dos modelos. A técnica *Simplex Mesh* foi aplicada aos modelos posicionados usando o mapa de energia acumulada dos coeficientes da DT-CWT.

Por fim, os métodos de avaliação para cada uma das principais etapas desse projeto foram detalhados, destacando as métricas utilizadas (Dice e HAD) e o teste estatístico empregado para inferir resultados para uma população. Os resultados das principais etapas da metodologia descrita nesse capítulo são apresentados no capítulo de resultados.

# Capítulo 5

## RESULTADOS E DISCUSSÕES

---

---

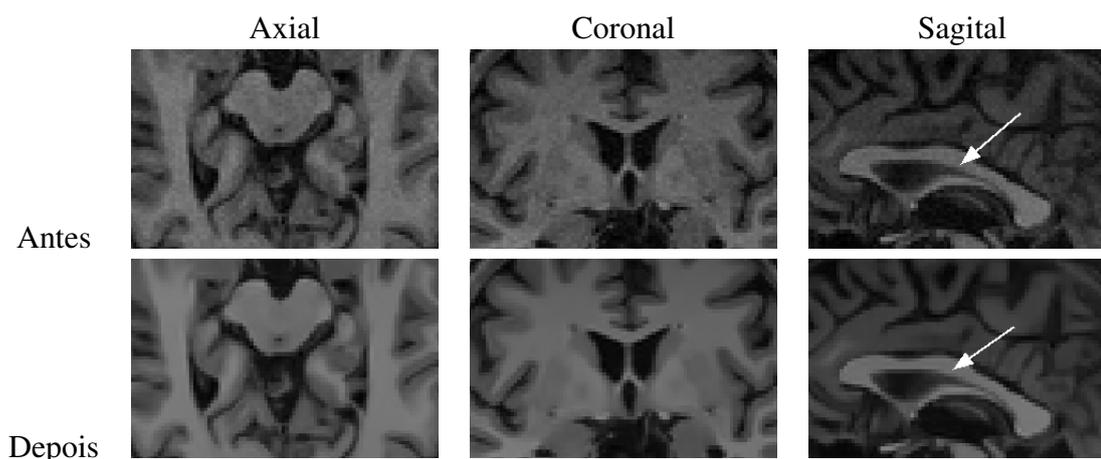
*Esse capítulo apresenta os resultados obtidos e discussões acerca das principais etapas do projeto, a nomear, as etapas de pré-processamento, detecção e descrição de pontos salientes, posicionamento das malhas, uso do modelo HMT e da técnica Simplex Mesh.*

### 5.1 Pré-processamento

As técnicas de pré-processamento são aplicadas a todas as imagens da base de dados EADC-ADNI descrita no Capítulo 4. Nas etapas de redução de ruído e correção de *bias field*, a avaliação dos resultados é realizada via comparação visual de uma imagem exemplo da base de dados anteriormente e posteriormente ao processamento. Na etapa de correção, o resultado é qualitativamente avaliado por meio da análise do padrão “tabuleiro de xadrez”. Por fim, os resultados de definição de máscaras das regiões de interesse são demonstrados sobrepondo a máscara obtida com a imagem original.

#### 5.1.1 Redução de ruído

A Figura 5.1 demonstra o resultado do filtro NLM (subseção 4.3.1) aplicado em uma das imagens de RM da base de dados EADC-ADNI (subseção 4.2.3), destacando uma região na qual pode ser observada a degradação da imagem e sua correção em cada um dos planos de visualização. É possível observar que o ruído presente na imagem é amenizado sem causar perdas de informação de borda. Na vista sagital, a estrutura do corpo caloso é destacada por uma seta para indicar a melhoria que essa etapa trouxe.



**Figura 5.1:** Resultado visual da redução de ruído, nos planos axial, coronal e sagital.

### 5.1.2 Correção do efeito *bias field*

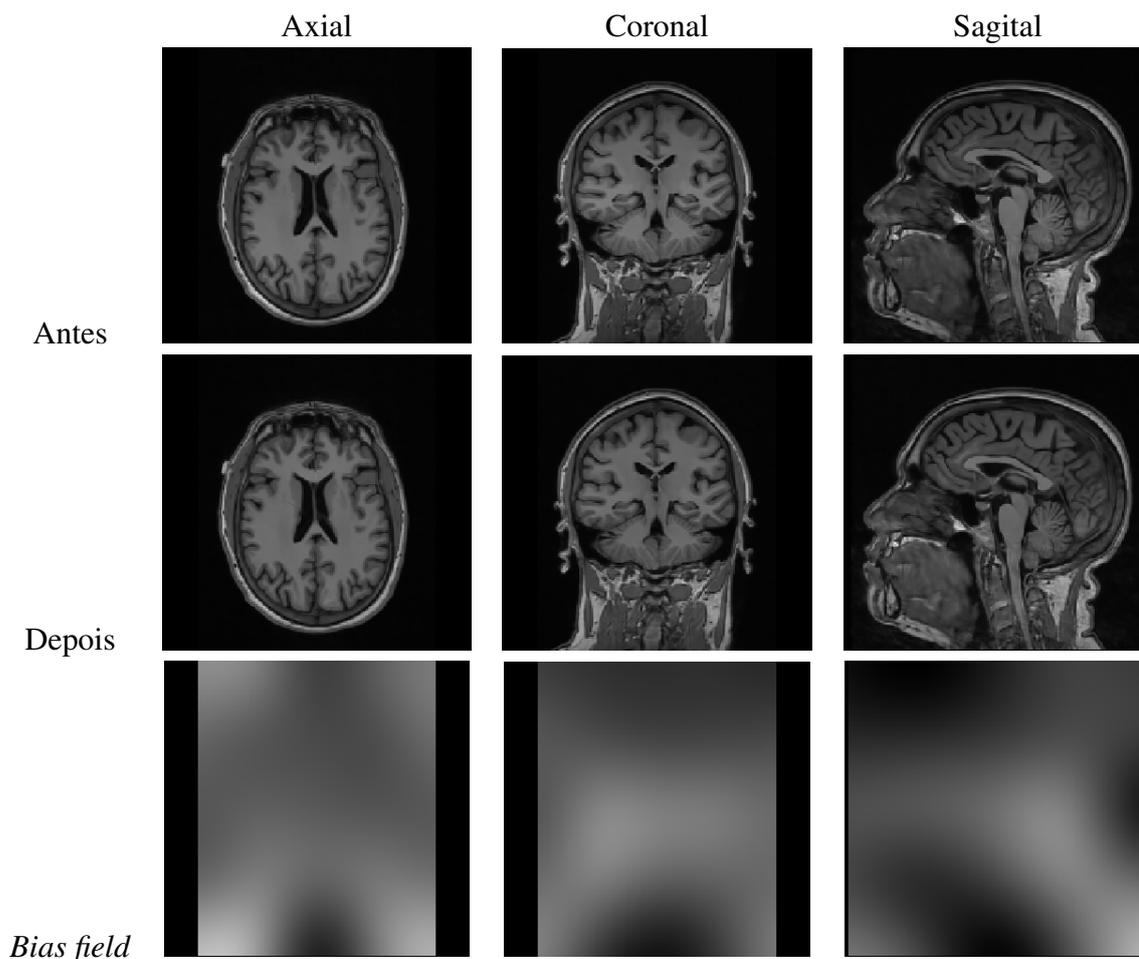
A Figura 5.2 demonstra o resultado da correção do efeito *bias field* em uma das imagens da base de dados EADC-ADNI (subseção 4.2.3). Apesar de não ser fortemente perceptível visualmente, existe um leve nível de degradação nas imagens antes do processamento, que pode ser observado pelo campo de inomogeneidade (*bias field*) removido da imagem. Esse campo causa variação na tonalidade dos níveis de cinza em regiões que deveriam apresentar os mesmos valores. A correção deixa os níveis de cinza mais uniformes nas mesmas estruturas.

### 5.1.3 Alinhamento espacial (corregistro)

A Figura 5.3 apresenta um resultado obtido da etapa de alinhamento espacial das imagens. Neste caso, uma imagem da base de dados e a imagem *template* de referência da base NAC (subseção 4.2.2) foram sobrepostas utilizando o padrão “tabuleiro de xadrez”, que secciona ambas as imagens e monta uma representação final pela combinação dos blocos subsequentes vindos alternadamente das duas imagens. Pode-se observar que o corregistro Afim conseguiu realizar uma boa padronização, garantindo que estruturas semelhantes (como por exemplo, o corpo caloso, indicado por uma seta na vista sagital) estejam localizadas em regiões próximas nas diferentes imagens da base de dados.

### 5.1.4 Máscara da região para a detecção

As máscaras das regiões de interesse para a detecção dos pontos salientes estão representadas nas imagens da Figura 5.4 sobrepostas à imagem *template* em suas projeções 2D para as vistas axial (a), coronal (b) e sagital (c). Elas foram obtidas pela aplicação da operação mor-



**Figura 5.2:** Resultado da correção do efeito *bias field*, nos planos axial, coronal e sagital.

fológica de dilatação, usando um elemento estruturante esférico de raio igual a 12 voxels, às máscaras representando as marcações manuais dos hipocampus na imagem de referência da base de dados NAC. Todas as imagens clínicas da base EADC-ADNI encontram-se alinhadas à imagem NAC (vide 5.1.3), e como pode ser visto pelas imagens da Figura 5.4, as regiões dos hipocampus e suas vizinhanças encontram-se dentro da área demarcada pelas máscaras.

### 5.1.5 Máscara da região para treinamento do modelo HMT

O resultado da extração da máscara do encéfalo pelo algoritmo ROBEX (IGLESIAS et al., 2011) pode ser observado na Figura 5.5, que exibe a sobreposição da máscara obtida com a imagem original. A região do encéfalo é a região de interesse no treinamento do modelo HMT, dessa maneira, é importante ter uma máscara para realizar o processamento apenas nos voxels presentes na região que ela delimita. Essa máscara foi definida para a imagem da base de dados NAC, pois o treinamento para encontrar os parâmetros do modelo foi realizado com essa imagem.

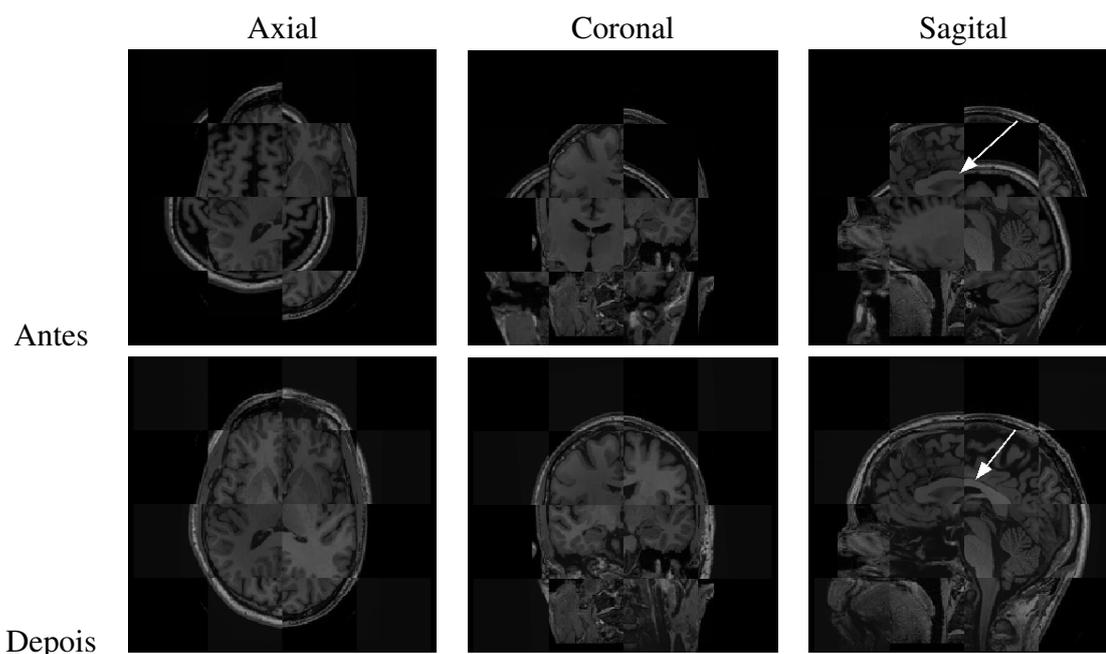


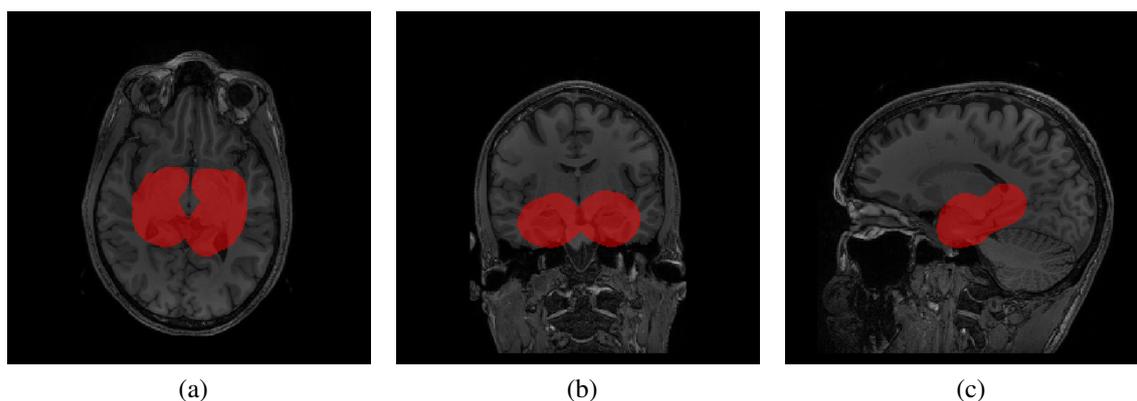
Figura 5.3: Comparação entre o “tabuleiro de xadrez” formado por uma das imagens da base de dados e a imagem *template* antes e depois do corregristo, nos planos axial, coronal e sagital.

## 5.2 Detecção de pontos salientes

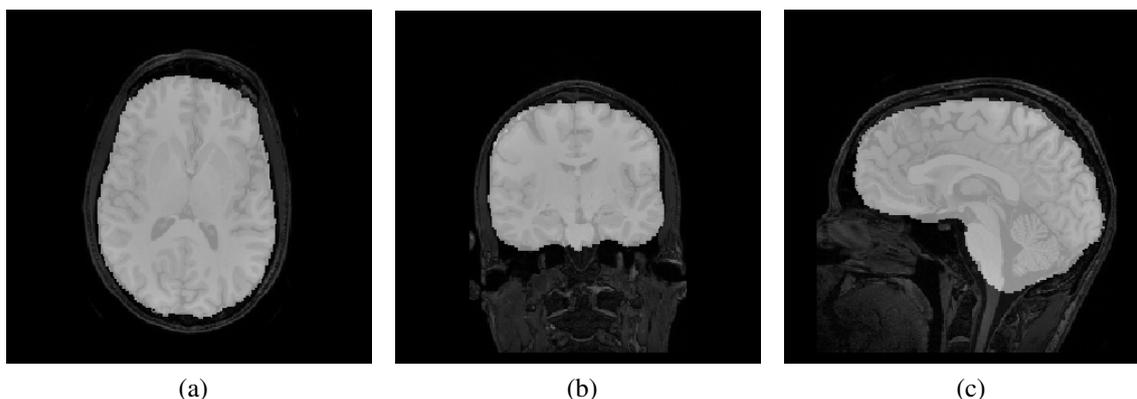
### 5.2.1 Análise de invariância à rotação, escala e translação

A Figura 5.6 ilustra o resultado da detecção dos pontos salientes, sobrepostos à imagem do cubo utilizada nesse experimento, nas suas versões original (a), rotacionada (b), escalonada (c) e transladada (d). A imagem do cubo foi escolhida para esse experimento uma vez que ela facilita a verificação visual dos resultados, já que os pontos que devem ser detectados são os oito cantos do cubo. A imagem possui dimensões  $128 \times 128 \times 128$  voxels e é uma imagem binária, sendo que o fundo possui valor 0 e a região do cubo possui valor 1.

Primeiramente, foram determinados os parâmetros da técnica (número de escalas,  $\alpha$ ,  $\beta$ , tamanho da vizinhança e limiar  $T_E$ ) capazes de detectar os pontos na imagem original do cubo. Depois, esses mesmos parâmetros foram utilizados na aplicação do algoritmo nas outras versões da imagem. Os parâmetros que permitiram ao detector encontrar os oito cantos nas quatro versões da imagem, dentro do conjunto de valores verificado nesse experimento, foram 3 escalas de decomposição,  $\alpha = 1$ ,  $\beta = 0,25$ , vizinhança de tamanho  $3 \times 3 \times 3$  e limiar  $T_E$  assumindo valor 0,05. Os resultados demonstram que, independente da transformação aplicada à imagem original, considerando apenas esse conjunto mais simples de transformações, o detector proposto nessa pesquisa para imagens volumétricas consegue detectar todos os pontos de interesse. É importante notar também que os pontos detectados pela técnica estão localizados no interior



**Figura 5.4:** Representação da região de aplicação do algoritmo de detecção de pontos salientes (em vermelho) sobreposta à imagem da base NAC, nas vistas axial (a), coronal (b) e sagital (c). Nas vistas axial e coronal, as máscaras da região do hipocampo direito e esquerdo encontram-se representadas. Na vista sagital apenas a máscara do hipocampo esquerdo encontra-se representada.

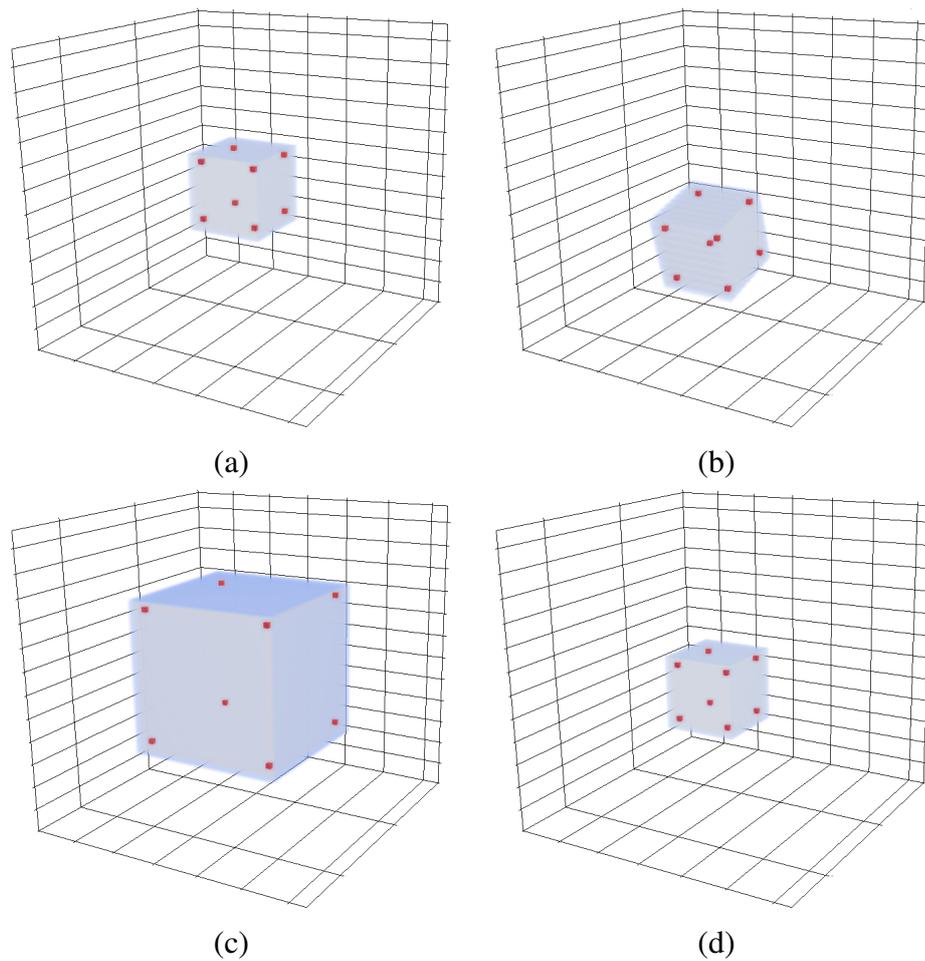


**Figura 5.5:** Sobreposição da imagem original com a máscara do encéfalo, nos planos axial, coronal e sagital.

do objeto, e não exatamente no canto.

### 5.2.2 Análise das estruturas do tipo “ponta”

**Otimização de parâmetros** Para a avaliação do detector de pontos salientes baseado na DT-CWT, no subgrupo de 54 imagens sintéticas sem degradação, primeiramente foi realizada uma busca exaustiva por um conjunto de parâmetros que fosse capaz de detectar apenas um ponto, a uma distância inferior a 10 voxels da localização do ponto de interesse em cada imagem. Nessa busca exaustiva, procurou-se minimizar a quantidade de falsos positivos (FP), ou seja, pontos excedentes detectados dentro e fora da região de interesse, e falsos negativos (FN), que correspondem aos casos da não detecção do ponto saliente dentro da região de interesse das imagens.



**Figura 5.6: Detecção de pontos salientes na versão original (a), rotacionada (b), escalonada (c) e transladada (d) do cubo.**

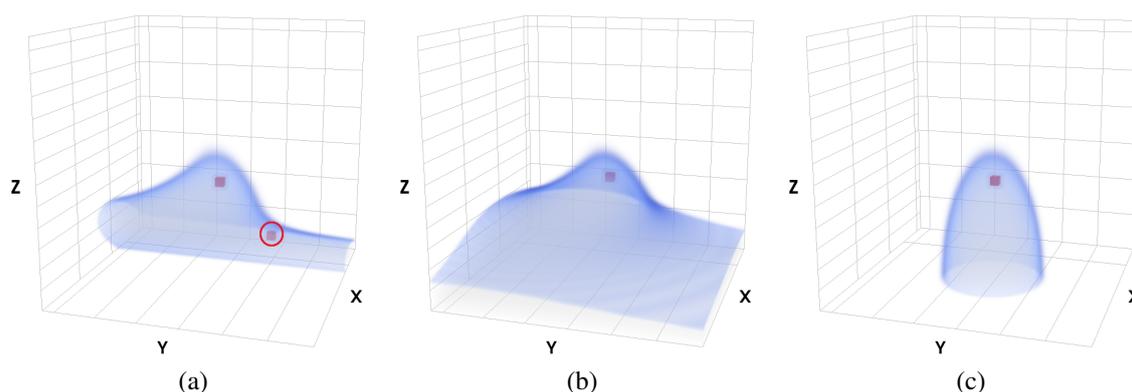
Considerando o tamanho das imagens sintéticas,  $64 \times 64 \times 64$  voxels, o número de escalas  $s$  foi definido com valor fixo 3 e o tamanho da vizinhança local foi definido com tamanho  $3 \times 3 \times 3$  voxels. A busca exaustiva foi realizada para os parâmetros nas seguintes variações:  $\alpha = \{0,25, 0,5, 0,75, 1, 2\}$ ,  $\beta = \{0,05, 0,1, 0,15, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 1\}$  e limiar  $T_E = \{0,1, 0,15, 0,20, 0,25, 0,5, 0,7, 0,8, 0,9\}$ .

Dentre os valores considerados na busca exaustiva, o melhor conjunto de parâmetros para o grupo de imagens sem degradação foi  $\alpha = 1$ ,  $\beta = 0,05$  e limiar  $T_E = 0,5$ . Com esses parâmetros, o detector proposto foi capaz de detectar um ponto saliente a uma distância máxima de 10 voxels do ponto de interesse em todas as imagens, com ocorrência de apenas 3 FPs. A distância média foi 5,67 voxels, com desvio padrão 0,75. Na tabela 5.1 é demonstrado outros resultados interessantes para algumas das combinações de parâmetros, elucidando a distância média do ponto mais próximo ao ponto de interesse detectado, o desvio padrão (DP) das distâncias, o número de FPs e o número de FNs.

$\alpha$	$\beta$	limiar	média	DP	FP	FN
0,25	0,1	0,15	6,00	0,97	4	0
0,5	0,15	0,1	6,21	1,14	4	0
0,75	0,1	0,2	6,03	0,98	4	0
2	0,05	0,5	5,70	0,84	3	0

**Tabela 5.1:** Resultados selecionados da combinação dos parâmetros  $\alpha$ ,  $\beta$  e limiar  $T_E$ .

Na Figura 5.7 é demonstrado o resultado do ponto detectado, sobreposto à imagem da estrutura do tipo “ponta” (representação tridimensional) para algumas das imagens da base de dados considerando os melhores parâmetros desse experimento. A estrutura representada na Figura (a) é uma das imagens na qual ocorre a detecção de um ponto FP (destacado por um círculo vermelho). Nesse caso, é observa-se que a detecção desse ponto ocorreu em uma região que apresenta uma certa curvatura, e considerando uma abordagem mais ampla, poderia ser um ponto interessante.



**Figura 5.7:** Pontos detectados em três imagens sintéticas sem degradação, do conjunto de imagens descrito na subseção 4.2.1. O ponto dentro de um círculo na estrutura (a) corresponde a um dos FPs detectados.

**Análise da robustez do detector na presença de ruído e *bias field*** Com o melhor conjunto de parâmetros da busca exaustiva no conjunto de imagens sem degradação ( $\alpha = 1$ ,  $\beta = 0,05$ , limiar  $T_E = 0,5$ ,  $s = 3$ , vizinhança =  $3 \times 3 \times 3$ ) foi verificado o comportamento do detector proposto nas mesmas imagens que simulam as estruturas anatômicas presentes no cérebro, descritas na subseção 4.2.1, em suas versões degradadas por 4 níveis de ruído e 3 níveis de *bias field* (o que resulta em 12 subgrupos, contando o subgrupo sem degradação em que o nível de ruído e *bias field* assume valor 0, ou seja, o número total de subgrupos com algum tipo de degradação é igual a 11).

Como descrito na subseção 4.9.2, nesse experimento foi analisada a reprodutibilidade da

distância do ponto mais próximo detectado a uma distância máxima de 10 voxels, ponto que é considerado o ponto saliente detectado, nos diferentes subgrupos. Usando o teste estatístico U de Mann-Whitney foram comparadas os valores das distâncias entre a localização do ponto detectado e a localização do ponto desejado, obtidas para os 11 subgrupos com degradação com os valores das distâncias obtidas para subgrupo estritamente sem degradação. Com essa informação é possível avaliar se as distâncias obtidas para cada subgrupo degradado podem ser consideradas oriundas da mesma população que define as distâncias do subgrupo sem degradação (hipótese nula do teste U de Mann-Whitney). Além dessa análise, também foi realizada uma análise do número de FPs e FNs que ocorrem na detecção quando a imagem é acometida por algum nível de degradação.

A Tabela 5.2 apresenta as medidas U e os  $p$  – *values* resultantes da aplicação do teste U de Mann-Whitney com nível de significância = 0,05 na comparação das distâncias dos pontos salientes detectados entre cada um dos grupos de imagens degradadas pela combinação de ruído e *bias field* e o grupo sem degradação. Para todas as comparações, a medida U apresentou valores altos, o que significa que existe uma grande sobreposição nas observações dos grupos comparados, ou seja, a distância dos pontos detectados possuem valores similares. Entretanto, esses resultados demonstraram que, na medida em que o *bias field* aumenta, a medida U diminui, pois os pontos passam a ser detectados mais distantes da posição desejada. A análise dos resultados do  $p$  – *value* demonstra que todos os conjuntos de imagens com degradação ainda apresentam uma distribuição similar ao conjunto de imagens sem degradação, já que todos possuem  $p$  – *value* > 0,05.

	<i>BF</i> = 0%		<i>BF</i> = 20%		<i>BF</i> = 40%	
	U	$p$ – <i>value</i>	U	$p$ – <i>value</i>	U	$p$ – <i>value</i>
$\sigma_r = 0\%$	---	---	655	0,83366	240	0,96012
$\sigma_r = 1\%$	1438,5	0,90448	660	0,88076	226,5	0,75656
$\sigma_r = 2\%$	1445	0,93624	724,5	0,9681	168,5	0,14706
$\sigma_r = 3\%$	1368	0,58232	662,5	0,89656	196,5	0,68910

**Tabela 5.2:** Valor U e  $p$ -*value* da comparação de cada uma das versões degradadas com a versão sem degradação das estruturas do tipo “ponta”.

Na Tabela 5.3 são apresentados, respectivamente, os valores médios de FPs detectados nos grupos de imagens e as quantidades de imagens sem detecção dentro da ROI (determinada pela distância máxima de 10 voxels da localização esperada) para cada uma das combinações de ruído e *bias field*. Os resultados demonstram que, na medida em que ruído e *bias field* aumentam, a quantidade de pontos indesejáveis detectados também aumenta, gerando maiores quantidades de FPs. Além disso, o aumento do *bias field* é o principal fator para a ocorrência de FNs.

	$BF = 0\%$		$BF = 20\%$		$BF = 40\%$	
	FP	FN	FP	FN	FP	FN
$\sigma_r = 0\%$	$0,056 \pm 0,23$	0	$0,056 \pm 0,23$	29	$0,018 \pm 0,13$	45
$\sigma_r = 1\%$	$0,074 \pm 0,26$	0	$0,074 \pm 0,26$	29	$0,111 \pm 0,31$	45
$\sigma_r = 2\%$	$0,074 \pm 0,26$	0	$0,074 \pm 0,26$	27	$0,278 \pm 0,65$	45
$\sigma_r = 3\%$	$0,074 \pm 0,26$	0	$0,074 \pm 0,26$	29	$1,241 \pm 4,49$	46

**Tabela 5.3:** Valor médio de FPs detectados no conjunto de imagens degradado e quantidade de FNs que ocorreram no conjunto de imagens pela quantidade de *bias field* e ruído especificadas.

Com base nesses resultados, nota-se que, apesar de ainda apresentar um comportamento similar na localização dos pontos detectados com diferentes níveis de degradação, o detector é bastante sensível à presença do efeito *bias field*, diminuindo consideravelmente a quantidade de imagens com detecção na medida em que a degradação por *bias field* aumenta. Entretanto, esse comportamento pode ser minimizado por uma correção de *bias field* nas imagens antes do processamento.

## 5.3 Descrição dos pontos salientes

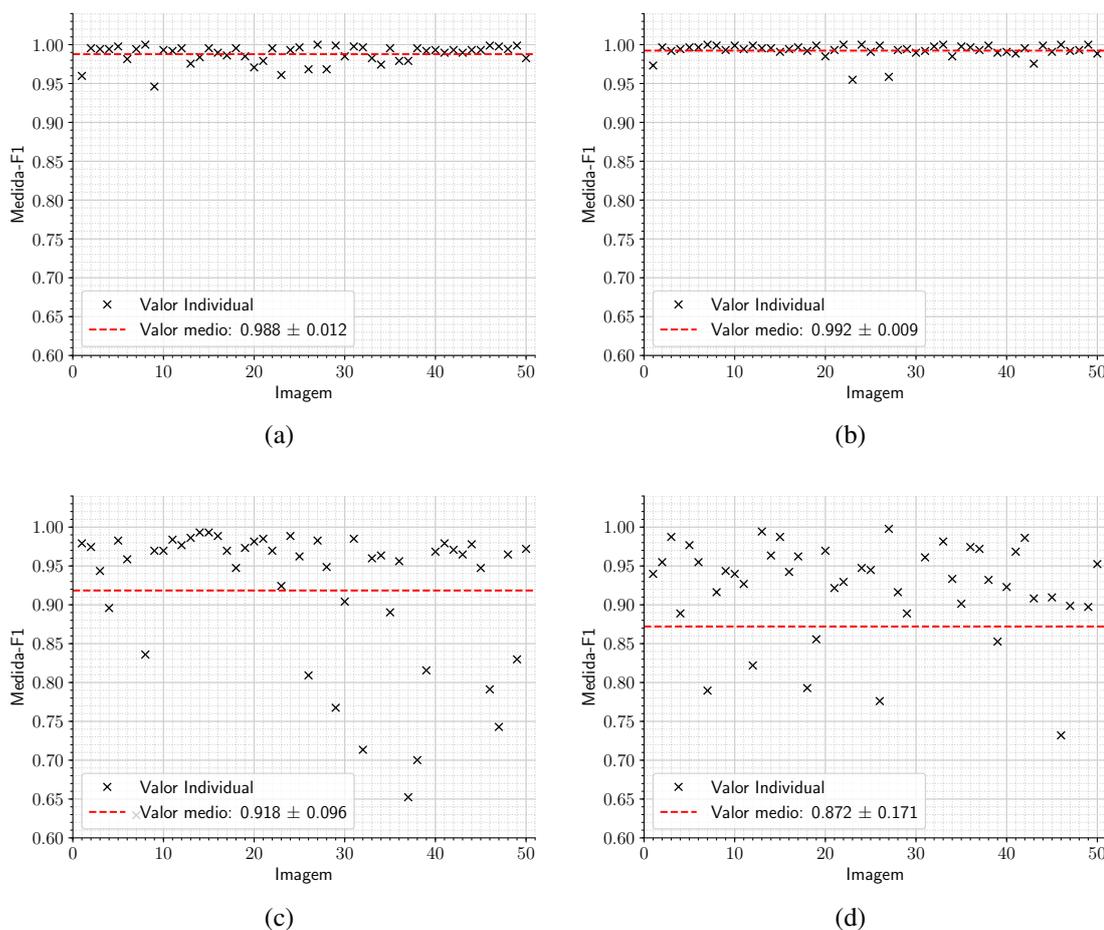
A avaliação do descritor foi realizada considerando a metodologia descrita na subseção 4.9.3. Os pontos salientes foram detectados uma única vez para a imagem *template* NAC na região do hipocampo esquerdo e suas respectivas localizações nas imagens deformadas foram estimadas.

A Figura 5.8 ilustra os valores obtidos da medida F1 para as imagens deformadas usando descritores com parâmetro  $R_{max}$  assumindo os valores 16 (a), 32 (b), 64 (c) e 128 (d). Como pode ser visto, o descritor *shape context* proposto apresentou uma boa performance nas imagens de teste, resultando em valores médios da medida F1 iguais a 0,988, 0,992, 0,918 e 0,872 para  $R_{max}$  com valores 16, 32, 64 e 128, respectivamente. O melhor resultado ocorreu quando  $R_{max}$  assumiu o valor 32, resultando no maior valor médio e menor dispersão da medida F1.

## 5.4 Posicionamento dos modelos deformáveis

### 5.4.1 Otimização dos parâmetros

Para avaliar o detector proposto na metodologia de posicionamento das malhas dos hipocampos, primeiramente foi encontrado o melhor conjunto de parâmetros usando uma abordagem similar à técnica *cross-validation*. Foi realizada uma busca exaustiva por parâmetros nos



**Figura 5.8:** Valores de medida F1 obtidos para o descritor proposto na avaliação de casamentos entre a imagem *template* NAC e cada uma das 50 imagens moderadamente deformadas para os diferentes valores de  $R_{max}$  - 16 (a), 32 (b), 64 (c) e 128 (d).

grupos de treinamento e os melhores parâmetros encontrados em cada grupo de treinamento foram aplicados ao grupo de teste correspondente. A métrica escolhida para determinar os melhores parâmetros foi o valor de Dice médio obtido nas diferentes combinações de parâmetros no conjunto de treinamento. Essa abordagem foi escolhida porque a busca por parâmetros dividindo as imagens em grupos de treinamento e teste permite saber como o detector se comportaria em um conjunto de dados não vistos.

A base de dados EADC-ADNI (composta por 135 imagens, como descrito na subseção 4.2.3) foi dividida em 5 *folds* resultando em 108 imagens para a etapa de treinamento e 27 para a etapa de teste. Na etapa de treinamento foram procurados os melhores valores para os parâmetros  $\alpha$ ,  $\beta$ ,  $T_E$  e tamanho da vizinhança usando os seguintes espaços de busca:  $\alpha = \{0,05, 0,5, 1\}$ ,  $\beta = \{0,005, 0,015, 0,03\}$ ,  $T_E = \{0,25, 0,5\}$  e tamanho da vizinhança =  $\{3 \times 3 \times 3, 5 \times 5 \times 5\}$ . O número de escalas  $s$  foi fixado em 3 para garantir que a escala mais grosseira ainda apresentasse detalhes suficientes. Finalmente, com os melhores parâmetros encontrados,

foi realizado o posicionamento do hipocampo esquerdo usando os pontos selecionados pelo detector baseado na DT-CWT. O tamanho do raio do descritor foi definido de acordo com o melhor resultado do experimento do descritor ( $R_{max} = 32$ ).

A combinação dos valores dos parâmetros  $\alpha = 0,05$ ,  $\beta = 0,015$ ,  $T_E = 0,25$ , e tamanho da vizinhança =  $3 \times 3 \times 3$  resultou em melhor valor de sobreposição para todos os  $5 \text{ folds}$ , com a métrica de Dice assumindo valores 0,59, 0,58, 0,57, 0,58 e 0,58 no grupo de treinamento, e valores correspondentes 0,55, 0,58, 0,61, 0,58 e 0,57 nos grupos de teste. Usando esses parâmetros, o valor médio de Dice, considerando a base toda, foi de 0,58, tanto para o hipocampo esquerdo, quando para o direito.

### 5.4.2 Comparação com outros métodos de posicionamento

Os *boxplots* nas Figuras 5.9 e 5.10 mostram uma comparação dos valores das métricas de Dice e HAD resultantes da sobreposição entre os GTs da EADC-ADNI (subseção 4.2.3) e os modelos dos hipocampos da NAC (subseção 4.2.2) posicionados usando uma transformação Afim, o detector PC 3D (FERRARI et al., 2011; VILLA-PINTO; FERRARI, 2016) e o detector DT-CWT proposto.

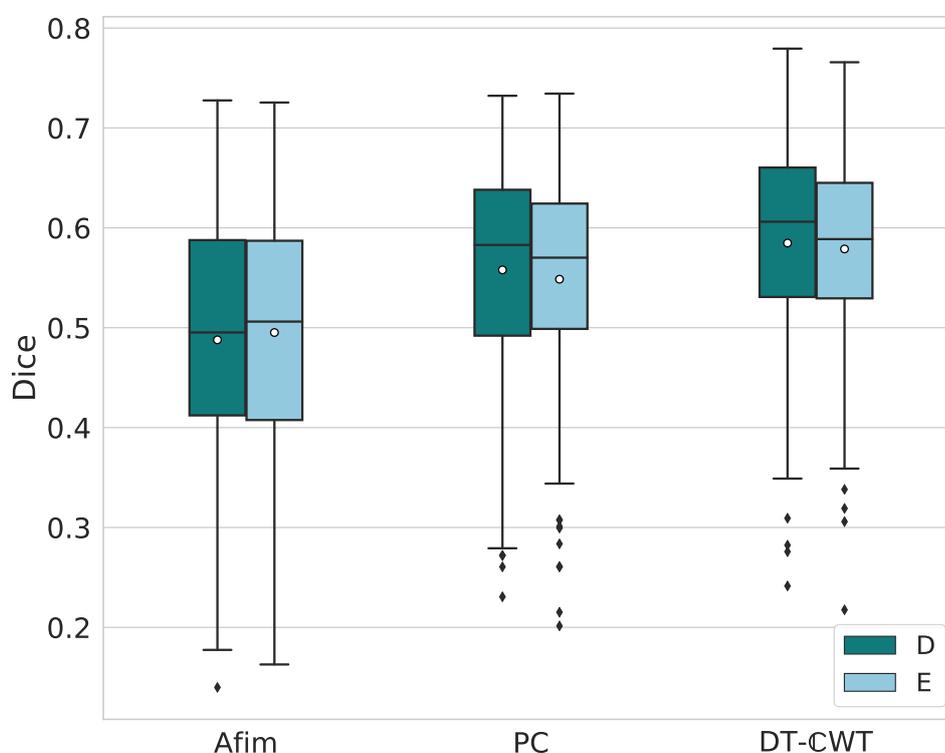
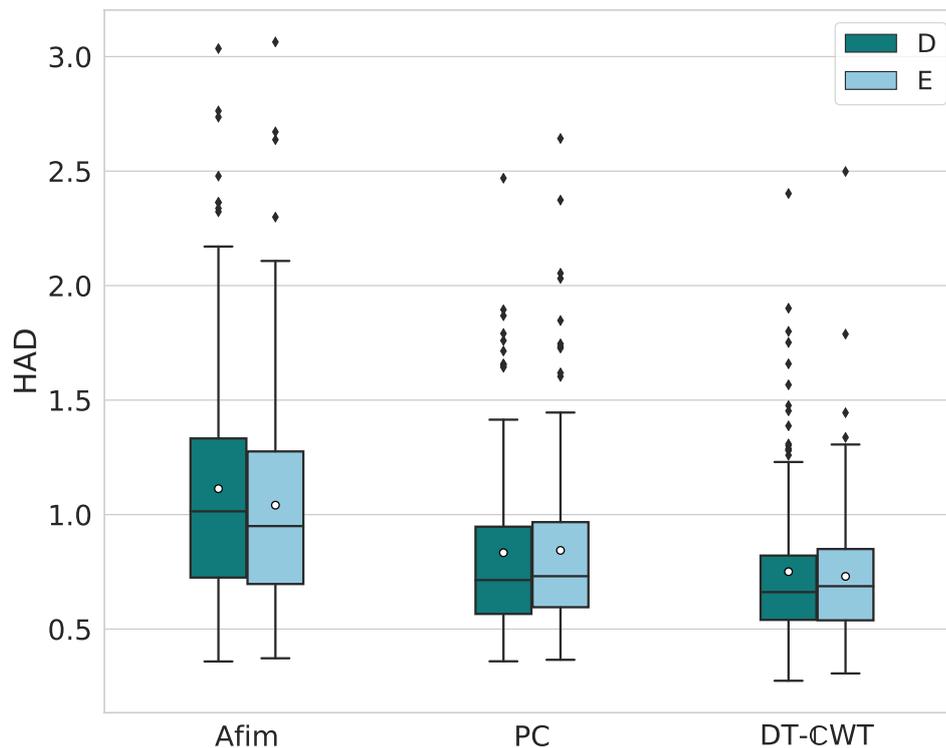


Figura 5.9: Valores de Dice para o posicionamento das malhas dos hipocampos esquerdo (E) e direito (D) usando três abordagens distintas: Afim, PC, e DT-CWT.



**Figura 5.10:** Valores de HAD para o posicionamento das malhas dos hipocampus esquerdo (E) e direito (D) usando três abordagens distintas: Afim, PC, e DT-CWT.

Como pode ser observado nos gráficos, o posicionamento guiado por uma transformação Afim possui os menores valores médios de Dice, 0,50 e 0,49 para os hipocampus esquerdo e direito, respectivamente, e a maior dispersão para essa métrica. Para a métrica HAD, o posicionamento Afim também apresenta os maiores valores médios e as maiores dispersões para ambos hipocampus, com o hipocampo esquerdo assumindo  $HAD = 1,04$  e o direito assumindo  $HAD = 1,11$ . Os posicionamentos usando tanto o detector PC 3D, quanto o detector baseado na DT-CWT apresentaram melhores valores médios de Dice e HAD, com menor dispersão, apesar de também apresentarem alguns *outliers*. Mais do que isso, o posicionamento usando o detector DT-CWT apresentou valores melhores do que o posicionamento usando o PC, já que foi atingido valor médio de 0,58 para ambos hipocampus, e valores médios de HAD de 0,73 para o hipocampo esquerdo e 0,75 para o direito, em contraste com a média de Dice/HAD assumindo valores 0,55/0,84 para o hipocampo esquerdo e 0,56/0,83 para o direito usando o detector PC 3D.

Além disso, a técnica de detecção baseada na DT-CWT 3D é computacionalmente mais eficiente quando comparado à técnica de detecção PC, que depende da construção de um conjunto de filtros log-Gabor proporcional ao tamanho da imagem.

O teste U de Mann-Whitney também foi usado para comparação dos resultados das duas

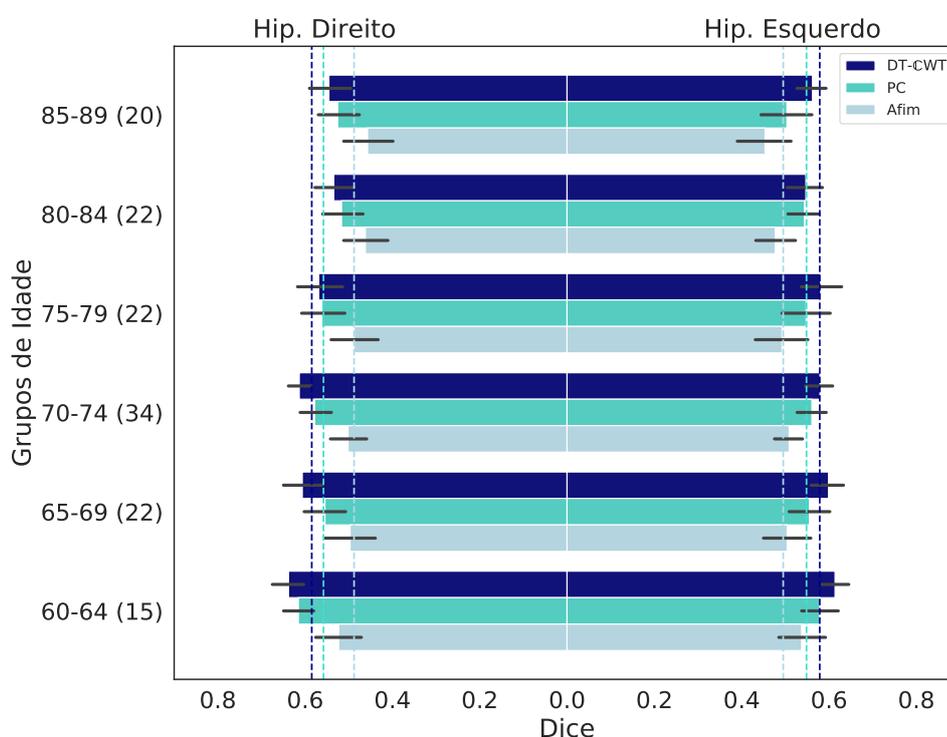
métricas (Dice, HAD) para a base de dados EADC-ADNI. Comparando o posicionamento usando a DT-CWT com o posicionamento Afim, o teste resultou em valores de *p-value* menores do que 0,00001 para ambos hipocampos usando as duas métricas. Com esses valores, usando nível de significância 0,05 ou 0,01 é possível descartar a hipótese nula do teste e concluir que os resultados são estatisticamente distintos e não vêm da mesma população. Quando comparado ao detector PC, para o hipocampo esquerdo, o *p-value* assume valores 0,063 e 0,043 para as métricas Dice e HAD, respectivamente, enquanto que os valores correspondentes para o hipocampo direito são 0,040 (Dice) e 0,041 (HAD). As conclusões variam considerando o nível de significância. Usando o nível de significância 0,01, não é possível descartar a hipótese nula, o que significa que os resultados, para os hipocampos esquerdo e direito, vêm da mesma população, e dessa maneira, os resultados da técnica DT-CWT não são estatisticamente distintos dos resultados da técnica PC. Por outro lado, com um nível de significância 0,05, apenas os resultados da métrica Dice para o hipocampo esquerdo da técnica PC não vem da mesma distribuição que os resultados Dice da DT-CWT, com os outros resultados descartando a hipótese nula (fato que ocorre por uma margem muito pequena).

**Análise estratificada** A população da base de dados EADC-ADNI foi dividida em grupos rotulados por idade e por diagnóstico clínico. A comparação estratificada foi realizada usando a métrica de Dice, no entanto, as conclusões para essa métrica podem ser estendidas para a métrica HAD, já que o comportamento dos resultados considerando a base toda foi similar em ambas métricas.

Os gráficos ilustrados nas Figuras 5.11 e 5.12 são similares à um gráfico de pirâmide populacional, e foram usados para ilustrar a estratificação por idade e diagnóstico. Cada barra horizontal representa o valor de Dice médio de um subgrupo com o desvio padrão correspondente indicado ao fim da barra por uma pequena linha horizontal. Em cada gráfico, a linha vertical tracejada representa o valor médio de Dice para cada técnica, considerando a população total da base de dados.

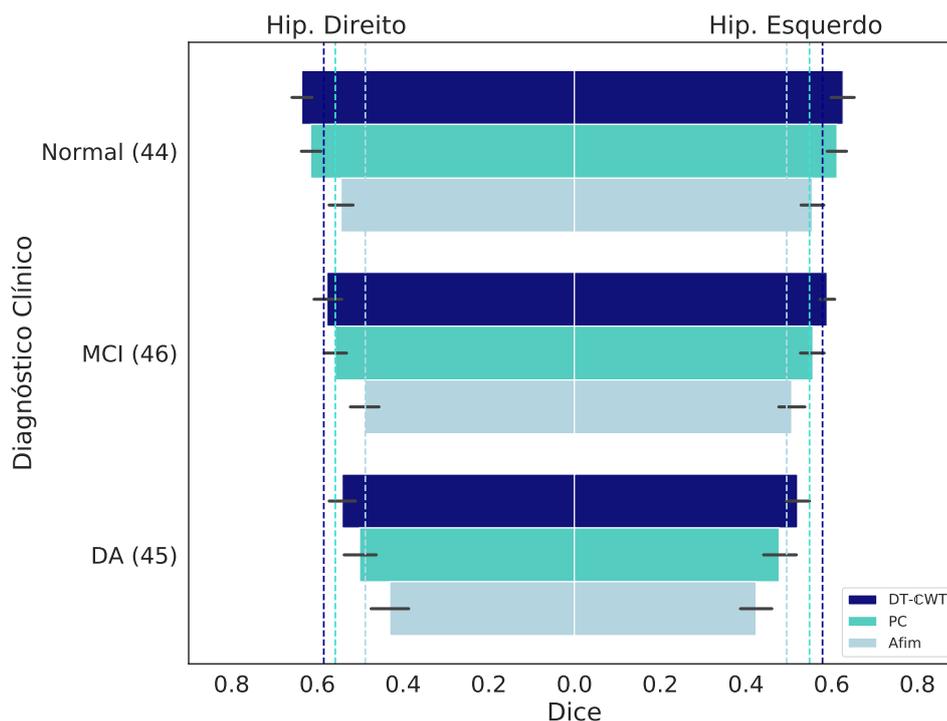
**Avaliação da base EADC-ADNI estratificada por idade** Nesse estudo, foi realizada uma avaliação comparativa dos resultados para o posicionamento usando a métrica de Dice das três técnicas, Afim, PC e DT-CWT, na população da base de dados EADC-ADNI estratificada por idade. Nessa análise, são considerados seis grupos definidos pelos seguintes intervalos de idade: [60 – 64, 65 – 69, 70 – 74, 75 – 79, 80 – 84, 85 – 89], com o número de pacientes em cada subgrupo igual a 15, 22, 34, 22, 22 e 20, respectivamente.

Como ilustrado na Figura 5.11, em todos os grupos estratificados os valores médios de Dice para o detector DT-CWT proposto supera os valores médios alcançados pelo posicionamento Afim e pelo detector PC. Mais do que isso, os melhores valores do posicionamento Afim (esquerdo = 0,54, direito = 0,52 para o intervalo de idade 60 – 64) são menores que os piores casos obtidos para a metodologia proposta usando o detector DT-CWT (esquerdo = 0,55, direito = 0,53 para o intervalo de idade 80 – 84). Pode-se notar também que, na medida em que o intervalo de idade aumenta, a acurácia no posicionamento tende a decrescer para todos os métodos, apesar de ser possível observar certa estabilidade nos intervalos 65 – 69, 70 – 74 e 75 – 79 para os detectores PC e DT-CWT.



**Figura 5.11:** Valores de Dice para o posicionamento dos hipocampos esquerdo e direito estratificados por idade usando três abordagens distintas: Afim, PC, e DT-CWT. Em parêntesis está indicado o número de imagens em cada uma das classes.

Mais precisamente, considerando o posicionamento usando o detector DT-CWT, que apresentou os melhores resultados nesse experimento, o decréscimo no valor médio de Dice entre o intervalo com indivíduos mais jovens e o intervalo com indivíduos mais velhos é da ordem de 8,2% para o hipocampo esquerdo e 14,1% para o direito. Para o posicionamento Afim, que representa os piores resultados, os valores correspondentes são de aproximadamente 16,7% e 11,5%. Esse comportamento pode ser explicado pelo fato de que existe uma diferença muito grande de forma e volume entre o modelo posicionado da base de dados NAC (de um paciente saudável de 42 anos) e os GTs dos indivíduos desses grupos de idade avançados.



**Figura 5.12:** Valores de Dice para o posicionamento dos hipocampos esquerdo e direito estratificados por diagnóstico clínico usando três abordagens distintas: Afim, PC, e DT-CWT. Em parêntesis está indicado o número de imagens em cada uma das classes.

**Avaliação da base EADC-ADNI estratificada por diagnóstico clínico** A base de dados EADC-ADNI também pode ser subdividida em indivíduos afetados por diagnósticos clínicos distintos, os quais se encontram detalhados na Tabela 4.1. Assim como na estratificação por idade, a comparação entre as três técnicas, considerando a condição de saúde da população, foi realizada usando a métrica de Dice.

O gráfico da Figura 5.12 mostra os valores médios de Dice do posicionamento dos hipocampos para os três subgrupos de indivíduos rotulados por diagnóstico clínico na base de dados. Como pode ser visto, a metodologia de posicionamento usando o detector DT-CWT apresenta médias superiores em todos os grupos estratificados para ambos hipocampos. Entretanto, é possível também observar que para todas as técnicas a acurácia do posicionamento decresce na medida em que a condição de saúde piora.

Para as três formas de posicionamento avaliadas nesse estudo, as médias de Dice tanto para o hipocampo esquerdo quanto para o direito para o grupo de indivíduos normais resultaram em valores maiores quando comparadas às médias correspondentes nos grupos MCI e DA. Esse comportamento era esperado, pois a atrofia nos hipocampos tende a ser mais severa em indivíduos acometidos por essas condições.

Os valores médios de Dice do posicionamento usando o detector DT-CWT para o grupo de

pacientes normais foram 0,63 e 0,64 para os hipocampos esquerdo e direito, respectivamente. Já para os grupos MCI e DA os valores médios foram 0,59 e 0,58, e 0,52 e 0,54, respectivamente.

Para o hipocampo esquerdo, quando se compara os resultados obtidos pela técnica DT-CWT em indivíduos normais e do grupo DA, é possível notar um decréscimo de aproximadamente 17,5% nos valores. Os valores de decréscimo correspondentes avaliando os resultados das técnicas PC e Afim são de 21,3% e 25%, respectivamente. Similarmente, esses os valores de decréscimo para o hipocampo direito são de 15,6%, 19,4% e 20,4%, para o detector DT-CWT, PC e o posicionamento Afim, respectivamente. Apesar de apresentar um comportamento semelhante, a metodologia usando o detector DT-CWT apresenta uma taxa de decréscimo menor quando comparada aos métodos PC e Afim.

É importante salientar que a base de dados EADC-ADNI é composta por imagens adquiridas de pacientes em faixas etárias avançadas e afetados por condições de diagnóstico distintas. O envelhecimento saudável já causa alterações na forma e redução no volume das estruturas dos hipocampos. Quando o paciente tem uma doença neurodegenerativa, esse cenário piora. A malha de referência da base de dados NAC, adquirida de um paciente saudável de 42 anos, é significativamente maior do que as malhas GT da base EADC-ADNI o que implica que, nem mesmo para o melhor cenário (malhas completamente sobrepostas), o valor de Dice poderia ser 1. A fim de verificar quais seriam os maiores valores médios possíveis que poderiam ser atingidos se o posicionamento fosse perfeito, foi realizado um experimento simulando a situação ideal. Nesse experimento, foi assumido total sobreposição entre a malha de referência e a malha GT, e conforme esperado, os valores foram menores do que 1. Para o hipocampo esquerdo, os valores foram 0,83, 0,75 e 0,70 para os grupos Normal, MCI e DA, respectivamente. Para o hipocampo direito, os valores correspondentes foram 0,89, 0,81 e 0,77. Além disso, esses resultados também comprovaram a tendência observada para as classes do diagnóstico clínico, com valores ordenados da seguinte maneira: Normal > MCI > DA.

## 5.5 Avaliação do modelo probabilístico HMT

### 5.5.1 Treinamento

Considerando o fato de que todas as imagens de RM do cérebro possuem características semelhantes, o treinamento para encontrar os parâmetros do modelo HMT foi realizado em apenas uma imagem, a imagem *template* da base de dados NAC, já que essa imagem foi utilizada como referência nas etapas do pré-processamento.

Como descrito na subsecção 4.7, no treinamento do modelo probabilístico HMT procura-se encontrar os seguintes parâmetros:

- A distribuição inicial para os coeficientes na raiz  $\pi(i) = p(Z_0 = i)$ , em que  $Z_0$  representa a variável de estado e  $i$  representa os estados que ela pode assumir, sendo que o valor 0 representa o estado *small* e o valor 1 representa o estado *large*.
- As probabilidades de transição entre as escalas, do coeficiente pai para os coeficientes filhos, representadas pela matriz  $A_{m,n} = P(Z_l = n | Z_{pai(l)} = m)$ , em que  $n$  representa o estado do nó atual  $l$ , e  $m$  representa o estado do nó pai de  $l$ ,  $pai(l)$ :

$$\begin{bmatrix} P(Z_l = 0 | Z_{pai(l)} = 0) & P(Z_l = 1 | Z_{pai(l)} = 0) \\ P(Z_l = 0 | Z_{pai(l)} = 1) & P(Z_l = 1 | Z_{pai(l)} = 1) \end{bmatrix}. \quad (5.1)$$

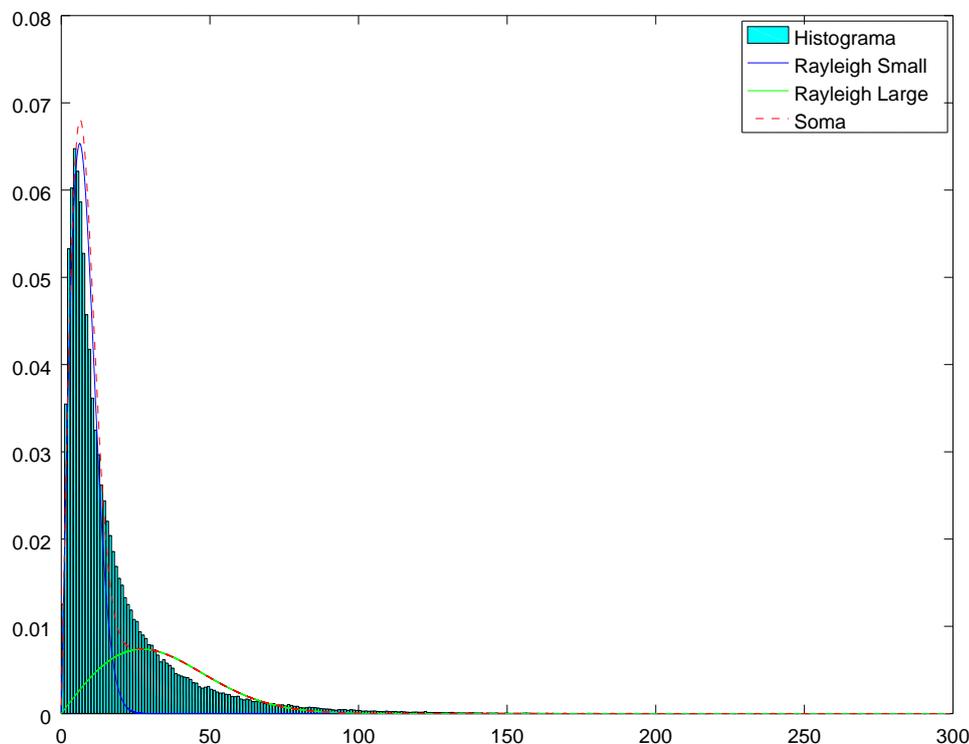
- Os valores de  $\sigma_{small}$  e  $\sigma_{large}$  que parametrizam a mistura de distribuições Rayleigh aproximando a distribuição das magnitudes dos coeficientes complexos *wavelet* em uma dada sub-banda.

A decomposição de uma imagem 3D pela DT-CWT resulta em 28 sub-bandas orientadas. No treinamento do modelo HMT, cada orientação representa uma árvore separada e, em concordância com a quantidade de escalas usada no detector, a DT-CWT foi aplicada na imagem *template* usando três escalas de decomposição. Os parâmetros descritos anteriormente foram estimados para cada escala, ou seja, cada escala de uma das 28 árvores possui uma matriz de transição e parâmetros modelando a mistura de distribuições Rayleigh. A escala mais grosseira, representando a raiz da árvore, possui valores de probabilidade inicial ao invés de uma matriz de transição.

A imagem de RM possui uma grande região de *background* (nível de cinza próximo de zero), o que faz com que muitos coeficientes assumam valor 0 ou próximos de 0 acarretando em uma concentração muito alta nessa região. Como essa informação não é de interesse para a modelagem, o treinamento do modelo HMT foi realizado apenas na região do encéfalo demarcada pela máscara definida em 5.1.5.

Para exemplificação dos resultados do treinamento, a Figura 5.13 mostra o resultado do ajuste (*fitting*) da mistura de funções de distribuição Rayleigh para uma das 28 orientações na escala mais fina de decomposição para a distribuição dos valores de magnitude dos coeficientes na sub-banda considerada. A linha delineada na cor azul representa a função estimada para a distribuição Rayleigh de variância *small* e a na cor verde representa a função estimada para a distribuição Rayleigh de variância *large*. O somatório da mistura de distribuições Rayleigh está

representado pela linha pontilhada vermelha. Como pode ser visto, a mistura de distribuições estimada teve um bom ajuste na sub-banda representada. Os parâmetros estimados para as outras sub-bandas também garantiram um bom ajuste à distribuição dos dados.



**Figura 5.13:** Mistura de funções de distribuição Rayleigh, representando a magnitude dos coeficientes *wavelets* dentro de uma sub-banda, estimada para a escala mais fina na primeira (a) e na décima primeira (b) orientações de decomposição.

### 5.5.2 Seleção de pontos para o posicionamento

O experimento para a seleção de pontos para o posicionamento consistiu em uma etapa extra no *pipeline* representado no diagrama da Figura 4.1. Essa etapa está localizada entre as etapas de detecção e casamento dos pontos salientes. Logo após a detecção (usando os melhores parâmetros encontrados no experimento do posicionamento), os pontos detectados são selecionados usando a informação de persistência oferecida pela classificação dos estados dos coeficientes ao longo das escalas em diferentes orientações. Somente aqueles pontos que apresentaram persistência em um dado número de orientações foram mantidos para a etapa seguinte. Essa seleção foi realizada para tentar aumentar a robustez dos casamentos dos pontos detectados, já que pontos não interessantes, erroneamente detectados e que podem atrapalhar os casamentos, não costumam apresentar persistência no estado *large* ao longo das escalas (ROMBERG et al., 2001).

Os parâmetros do modelo HMT estimados usando a imagem *template* NAC foram usados para classificar os coeficientes da decomposição DT-CWT de cada uma das imagens clínicas da base de dados EADC-ADNI, usando o algoritmo de recuperação de estados proposto por Durand, Goncalvès e Guédon (2004) (subseção 4.7). A decomposição *wavelet* é a mesma usada para a detecção de pontos, ou seja, composta por três escalas  $s = 3$  e 28 orientações, sendo que cada orientação representa uma HMT.

Em cada uma das imagens clínicas, para todos os pontos detectados usando o detector proposto, foram verificados os estados que os coeficientes equivalentes nas diferentes escalas de decomposição assumiam dentro de uma dada orientação. Se o ponto detectado apresentasse estados escondidos equivalentes, classificados como *large* nas três escalas de decomposição, essa orientação seria considerada como uma orientação na qual os coeficientes apresentaram persistência do estado *large*. Pontos salientes foram mantidos somente se apresentassem persistência do estado *large* em um determinado número mínimo de orientações das 28 oferecidas pela decomposição.

No experimento realizado para a avaliação do modelo HMT foi realizada uma variação na quantidade de orientações em que os pontos detectados deveriam apresentar persistência, no intervalo de 0 a 9 (considerando o número 9 de orientações, cerca de apenas 10% dos pontos detectados foram mantidos). A eficácia da seleção dos pontos salientes usando o modelo HMT foi verificada pela análise do valor de sobreposição da métrica Dice alcançado após eliminar os pontos salientes que não apresentassem persistência na quantidade de orientações determinada.

A Figura 5.14 ilustra um gráfico com os valores médios de sobreposição obtidos para os hipocampus esquerdo (em vermelho) e direito (na cor azul) após a seleção dos pontos para a quantidade mínima de orientações no intervalo de 0 a 9 (usar o número de orientações igual a 0 tem o mesmo resultado de não selecionar pontos). As Tabelas 5.4 (a) e (b) mostram a quantidade de pontos mantidos dado um número mínimo de orientações.

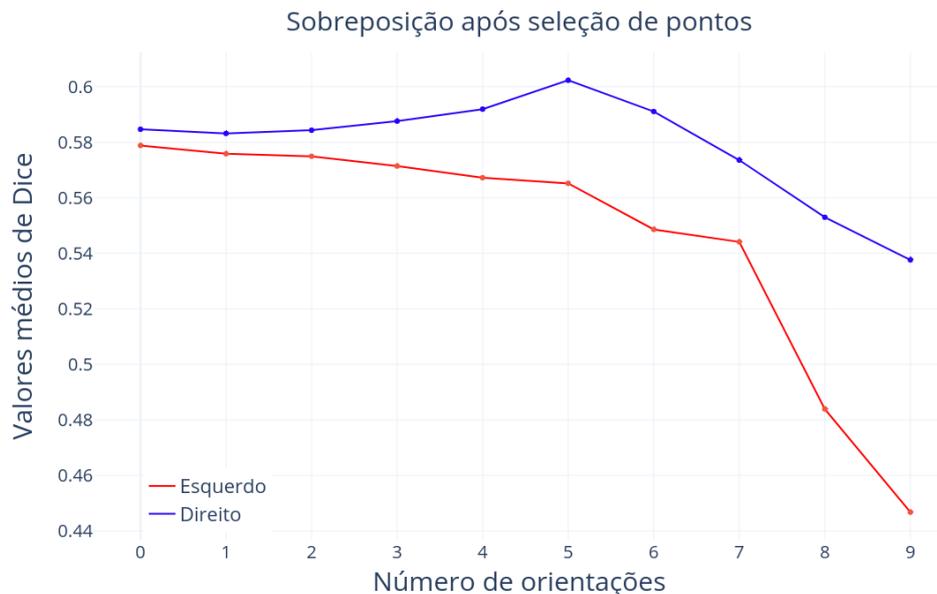
Orientações	0	1	2	3	4	5	6	7	8	9
Pontos	281,5	250,2	226,3	195,8	160,5	125,0	94,0	66,8	45,9	30,0

(a)

Orientações	0	1	2	3	4	5	6	7	8	9
Pontos	268,8	239,1	216,5	187,5	155,7	122,0	91,4	65,8	45,5	30,4

(b)

**Tabela 5.4:** Número de pontos mantidos na região do hipocampo esquerdo (a) e direito (b), usando o modelo HMT, considerando um número mínimo de orientações.



**Figura 5.14:** Valores de Dice médio obtidos para o posicionamento das malhas dos hipocampus esquerdo (na cor vermelha) e direito (na cor azul) na base EADC-ADNI após a seleção dos pontos salientes detectados considerando um número mínimo de orientações no qual os coeficientes equivalentes deveriam apresentar persistência em todas as escalas de decomposição.

Como pode ser observado, o uso do modelo HMT para a seleção de pontos salientes na região do hipocampo esquerdo não se mostrou benéfica para o posicionamento do modelo equivalente. Na medida em que a quantidade mínima de orientações com persistência para os pontos salientes detectados aumenta, o valor médio de Dice obtido para a base de dados EADC-ADNI diminui. No caso do hipocampo direito, os resultados de sobreposição aumentaram ligeiramente até o número mínimo de 5 orientações, sendo este caso o que resultou em melhor média (0,60) para seu posicionamento na base.

Apesar do aumento para o lado direito, o valor de sobreposição alcançado não é suficiente para a adição dessa etapa no *pipeline*, já que o mesmo não ocorre para a estrutura do hipocampo esquerdo. Dessa maneira, na metodologia de posicionamento foi escolhido não realizar a seleção dos pontos salientes por meio do modelo HMT. Apesar de não ser realizada uma seleção explícita de pontos salientes, na etapa de casamento de descritores da metodologia proposta, a verificação de reciprocidade entre os casamentos<sup>1</sup> consegue eliminar uma boa quantidade de pontos, o que pode ser visto também como uma seleção implícita dos pontos.

<sup>1</sup>A não reciprocidade entre os casamentos implica que um ponto da imagem clínica se casa com o ponto equivalente na imagem *template*, mas existe outro ponto com descritor mais semelhante que faz com que o ponto da imagem *template* se case com um ponto não equivalente na imagem clínica.

## 5.6 Avaliação do ajuste local usando a técnica *Simplex Mesh*

A técnica de deformação *Simplex Mesh* foi usada para o ajuste local dos modelos posicionados. Para comparação, tanto o mapa de energia acumulada da DT-CWT, gerado para a detecção dos pontos salientes, quanto a imagem original, foram usados como fonte de informação de bordas. Os parâmetros do modelo foram determinados de acordo com as recomendações da implementação da técnica pela biblioteca ITK, com parâmetro  $\alpha_{sm} = 0,3$ , o que garante uma maior rigidez da malha dando maior peso à  $\mathbf{f}_{int}$ ;  $\beta_{sm} = 0,05$ , diminuindo a influência da  $\mathbf{f}_{ext}$  (um valor mais alto poderia fazer com que a malha perdesse seu formato original), e  $\gamma = 0,001$ , valor que garante maior regularidade à malha.

A Tabela 5.5 apresenta os resultados médios de Dice para a sobreposição obtida na base de dados EADC-ADNI comparando as abordagens de posicionamento guiada por uma transformação Afim, técnica que resultou nos piores valores, e o posicionamento guiado pelos pontos detectados pela DT-CWT, técnica que resultou nos melhores valores.

Imagem				
	Esquerdo		Direito	
	Posicionamento	Ajuste	Posicionamento	Ajuste
Afim	0,50	0,50	0,49	0,49
DT-CWT	0,58	0,58	0,58	0,58
Mapa de Energia				
	Esquerdo		Direito	
	Posicionamento	Ajuste	Posicionamento	Ajuste
Afim	0,50	0,51	0,49	0,50
DT-CWT	0,58	0,60	0,58	0,60

**Tabela 5.5:** Valores médios de Dice na base de dados EADC-ADNI após o ajuste usando a técnica *Simplex Mesh* para os modelos dos hipocampos direito e esquerdo posicionados usando uma transformação Afim, por meio de uma transformação guiada pelos pontos da DT-CWT, com a imagem original e com a imagem de energia, destacando os valores alcançados pelo posicionamento e pelo ajuste realizado a partir do posicionamento destacado.

Como pode ser observado, o ajuste resulta em um ligeiro aumento de sobreposição média na base de dados, para os dois hipocampos, quando se usa o mapa de energia como fonte de borda, principalmente a partir do posicionamento DT-CWT. Usando a imagem original os resultados não apresentam um aumento significativo na média geral de sobreposição da base, sendo que os valores resultantes arredondados para duas casas decimais são os mesmos do posicionamento.

## 5.7 Considerações finais

Este capítulo apresentou os resultados obtidos para as principais etapas e para a metodologia proposta nesse projeto de pesquisa. No pré-processamento foi demonstrado que as técnicas aplicadas são capazes de padronizar todas as imagens da base de dados EADC-ADNI, eliminando problemas que poderiam prejudicar a performance das técnicas da metodologia de posicionamento. O detector proposto se mostrou eficiente para sua aplicação em imagens sintéticas representativas de estruturas de interesse, apesar de possuir algumas limitações como, a dependência na definição de alguns parâmetros, detecção de pontos um pouco distantes (média de 5,67 voxels de distância no melhor caso) da localização correta e a sensibilidade do detector à degradação causada pelo efeito *bias field*.

A análise do descritor baseado na técnica *shape context* no domínio *wavelet* apresentou resultados muito bons no conjunto de imagens moderadamente deformadas, com o raio  $R_{max}$  de tamanho 32 assumindo os melhores valores nessa avaliação (Medida F1 com valor médio de 0,992). Esse valor de raio foi mantido para os descritores no *pipeline* de posicionamento. Os resultados do *pipeline* de posicionamento usando o detector proposto nesse projeto (Dice = 0,58/0,58 e HAD = 0,73/0,75 para o hipocampo esquerdo e direito, respectivamente) se mostraram superiores aos resultados alcançados pelo posicionamento Afim (Dice = 0,49/0,50, HAD = 1,04/1,11) e pelo posicionamento baseado no detector PC (Dice = 0,55/0,56, HAD = 0,84/0,83), tanto para a base toda, quanto para a base estratificada por idade ou diagnóstico clínico. Entretanto, é importante frisar que, para a técnica PC, os parâmetros foram definidos de acordo com valores *default* e não foi procurado um conjunto de melhores parâmetros para essa base de dados.

Além desses resultados principais, nesse capítulo também foram apresentados resultados da análise do modelo HMT como seletor de pontos salientes e da técnica de deformação *Simplex Mesh*. No caso do modelo HMT, os resultados de seleção de pontos não se mostraram eficientes para o hipocampo esquerdo e se mostraram pouco eficientes para o direito. Dessa maneira, a seleção de pontos usando o modelo HMT não foi adicionada à metodologia de posicionamento proposta. O ajuste usando a técnica *Simplex Mesh* também não resultou em um aumento considerável na média de sobreposição da métrica Dice para as imagens da base EADC-ADNI.

# Capítulo 6

## CONCLUSÕES

---

---

*Esse capítulo apresenta as principais conclusões e observações desse projeto de pesquisa, além de uma listagem das contribuições e algumas sugestões para trabalhos futuros.*

### 6.1 Conclusões

Este texto apresentou a pesquisa realizada durante esse projeto de mestrado. O objetivo principal desse trabalho foi o desenvolvimento de uma técnica automática para a detecção e descrição de pontos salientes em imagens de RM 3D do cérebro baseada nos coeficientes da DT-CWT. O detector proposto foi aplicado para o posicionamento inicial de modelos deformáveis representando as estruturas dos hipocampos esquerdo e direito, que são acometidas nos estágios iniciais da DA. O posicionamento inicial das malhas, próximo às estruturas que elas representam em imagens 3D de RM, constitui uma etapa necessária e importante para algoritmos de segmentação baseados na deformação de modelos, já que uma boa inicialização aumenta a chance de sucesso desses algoritmos.

A DT-CWT fornece uma alternativa para a análise de frequências localizadas de um sinal, mantendo propriedades analíticas desejáveis da transformada de Fourier que a transformada *wavelet* discreta (DWT) não oferece. Conforme descrito na seção de revisão bibliográfica, apesar de suas excelentes propriedades para análise de sinais, a DT-CWT foi pouco explorada para a detecção de pontos salientes, sendo o principal trabalho voltado para imagens bidimensionais. Com base nisso, nesse projeto de pesquisa buscou-se verificar a aplicabilidade de um detector baseado em um mapa de energia acumulada obtido dos coeficientes da DT-CWT 3D para detectar pontos salientes em imagens volumétricas, e do uso desse mesmo mapa como fonte de informação na construção do descritor dos pontos.

A detecção e a descrição de pontos salientes usando a DT-CWT alcançaram resultados animadores. Para o posicionamento dos modelos dos hipocampos, a técnica proposta obteve resultados superiores (Dice = 0,58/0,58 e HAD = 0,73/0,75 para o hipocampo esquerdo e direito, respectivamente) quando comparada ao posicionamento baseado em outras técnicas (Dice = 0,49/0,50, HAD = 1,04/1,11, para o posicionamento Afim e Dice = 0,55/0,56, HAD = 0,84/0,83 para o detector PC). Além dos resultados de posicionamento superiores, a técnica DT-CWT demanda menor complexidade computacional quando comparada à técnica PC. O uso do modelo probabilístico HMT para selecionar pontos salientes não se mostrou eficiente quando aplicado ao posicionamento dos modelos. Mais precisamente, para o modelo do hipocampo direito, os resultados obtidos foram ligeiramente superiores quando se usou o modelo HMT, entretanto, para o esquerdo, a seleção de pontos fez com que os resultados sofressem um considerável decréscimo. Apesar de não ter funcionado bem nessa metodologia, observou-se que o modelo HMT pode ser utilizado de forma independente, pois pontos interessantes do cérebro apresentam persistência em uma quantidade considerável de orientações.

## 6.2 Contribuições

Com base no que foi exposto no texto, as seguintes contribuições podem ser listadas:

- Desenvolvimento de um detector de pontos salientes 3D baseado na DT-CWT com aplicação no posicionamento de malhas do hipocampo em imagens de RM do cérebro;
- Desenvolvimento de um descritor baseado no mapa de energia resultante da combinação dos coeficientes da DT-CWT;
- Avaliação do comportamento do detector na presença de *bias field* e ruído;
- Avaliação comparativa dos resultados do posicionamento das malhas em relação à demarcações por especialistas em imagens clínicas entre a técnica proposta e os resultados obtidos por duas outras técnicas, a nomear, posicionamento Afim e posicionamento guiado pela técnica PC 3D. Essa avaliação foi realizada utilizando as métricas de Dice e a distância média de Hausdorff (HAD);
- Avaliação do modelo probabilístico HMT para a seleção de pontos salientes por meio da métrica de Dice;
- Avaliação de um algoritmo de ajuste de modelos deformáveis às malhas posicionadas pela técnica baseada na DT-CWT usando a métrica de Dice.

## 6.3 Sugestões para trabalhos futuros

O modelo probabilístico HMT não se mostrou fortemente eficiente quando usado em conjunto com o detector proposto nesse projeto de pesquisa. Entretanto, pontos que apresentam persistência em muitas orientações parecem também representar pontos que apresentam informação de saliência. Uma sugestão de trabalho futuro é a utilização do modelo HMT, de forma independente, para a detecção de pontos salientes no cérebro.

Além da informação de magnitude, a DT-CWT também fornece informação de fase. Outra sugestão para trabalho futuro é o estudo da informação de fase da DT-CWT, e a verificação de sua aplicabilidade para a detecção e/ou descrição de pontos salientes.

Por fim, a aplicação da técnica de ajuste *Simplex Mesh* não resultou em um bom acréscimo de sobreposição usando a imagem original ou a imagem de energia como fontes de borda para a aplicação do algoritmo. Uma sugestão para trabalho futuro é a verificação da informação de textura como fonte de borda para o ajuste local dos modelos usando essa técnica.

## REFERÊNCIAS BIBLIOGRÁFICAS

AJA-FERNÁNDEZ, S.; VEGAS-SÁNCHEZ-FERRERO, G. *Statistical Analysis of Noise in MRI*. [S.l.]: Springer, 2016.

ALLAIRE, S.; KIM, J.; BREEN, S.; JAFFRAY, D.; PEKAR, V. Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Anchorage, AK, USA: IEEE, 2008. p. 1–8.

ALLAIRE, S.; PEKAR, V.; HOPE, A. J.; BREEN, S. L.; JAFFRAY, D. A. Automatic contour propagation in head and neck IGRT based on 3D salient interest points. In: *Proceedings of the 50th Annual ASTRO Meeting*. Boston, MA, USA: Elsevier, 2008. p. 87.

BAILLEUL, J.; RUAN, S.; BLOYET, D. Automatic atlas-based building of point distribution model for segmentation of anatomical structures from brain MRI. In: *Seventh International Symposium on Signal Processing and Its Applications Proceedings*. Paris, France: IEEE, 2003. v. 2, p. 629–630.

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

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.

BLOCH, F.; HANSON, W. W.; PACKARD, M. Nuclear induction. *Physical Review*, v. 70, n. 7, 1946.

BOGOVIC, J. A.; PRINCE, J. L.; BAZIN, P.-L. A multiple object geometric deformable model for image segmentation. *Computer Vision and Image Understanding*, v. 117, n. 2, p. 145–157, 2013.

CHAN, M. *World Health Organization and Alzheimer's Disease International - Dementia: a public health priority*. Geneva, Switzerland, 2017.

CHEN, H.; KINGSBURY, N. Efficient registration of nonrigid 3-D bodies. *IEEE Transactions on Image Processing*, v. 21, n. 1, p. 262–272, 2012.

COUPÉ, P.; YGER, P.; PRIMA, S.; HELLIER, P.; KERVRANN, C.; BARILLOT, C. An Optimized Blockwise Nonlocal Means Denoising Filter for 3-D Magnetic Resonance Images. *IEEE Transactions on Medical Imaging*, v. 27, 05 2008.

CROUSE, M. S.; NOWAK, R. D.; BARANIUK, R. G. Wavelet-based statistical signal processing using hidden markov models. *IEEE Transactions on Signal Processing*, v. 46, n. 4, april 1998.

DELINGETTE, H. Simplex meshes: a general representation for 3D shape reconstruction. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*. Seattle, WA, USA: IEEE, 1994. p. 856–860.

DURAND, J. B.; Goncalvès, P.; Guédon, Y. Computational methods for hidden markov tree models - an application to wavelet trees. *IEEE Transactions on Signal Processing*, v. 52, n. 9, september 2004.

FAUQUEUR, J.; KINGSBURY, N.; ANDERSON, R. Multiscale keypoint detection using the dual-tree complex wavelet transform. In: *IEEE International Conference on Image Processing*. Atlanta, GA, USA: IEEE, 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.

FJELL, A. M.; MCEVOY, L.; HOLLAND, D.; DALE, A. M.; WALHOVD, K. B. What is normal in normal aging? Effects of aging, amyloid and Alzheimer's disease on the cerebral cortex and the hippocampus. *Progress in Neurobiology*, v. 117, p. 20–40, 2014.

FOX, N. C.; WARRINGTON, E. K.; FREEBOROUGH, P. A.; HARTIKAINEN, P.; KENNEDY, A. M.; STEVENS, J. M.; ROSSOR, M. N. Presymptomatic hippocampal atrophy in Alzheimer's disease: A longitudinal MRI study. *Brain: A Journal of Neurology*, v. 119, p. 2001–2007, 1996.

GHESU, F.; GEORGESCU, B.; ZHENG, Y.; GRBIC, S.; MAIER, A.; HORNEGGER, J.; COMANICIU, D. Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 41, n. 1, p. 176–189, 2019.

GHESU, F. C.; GEORGESCU, B.; MANSI, T.; NEUMANN, D.; HORNEGGER, J.; COMANICIU, D. An artificial agent for anatomical landmark detection in medical images. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Athens, Greece: Springer International Publishing, 2016. p. 229–237.

GOMES, J.; VELHO, L.; GOLDENSTEIN, S. *Wavelets: teoria, software e aplicações*. Rio de Janeiro: IMPA - 21o Colóquio Brasileiro de Matemática, 1997.

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. 213001–1–15, 2010.

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

HANSON, L. P. G. *Introduction to Magnetic Resonance Imaging Techniques*. [S.l.], Aug 2009.

- 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)*, v. 2, p. 64–75, 2010.
- HIRALAL, R.; MENON, H. P. A survey of brain MRI image segmentation methods and the issues involved. In: *Intelligent Systems Technologies and Applications 2016*. Boston: Springer, 2016. v. 530, p. 245–259.
- IDE, J.; CHEN, R.; SHEN, D.; HERSKOVITS, E. Robust brain registration using adaptive probabilistic atlas. In: *Medical Image Computing and Computer-Assisted Intervention: MICCAI*. New York City, USA: Springer, 2008. v. 11, p. 1041–1049.
- 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.
- JUNTU, J.; SIJBERS, J.; DYCK, D. V.; GIELEN, J. Bias field correction for MRI images. In: *Computer Recognition Systems*. [S.l.]: Springer Berlin Heidelberg, 2005. p. 543–551.
- KINGSBURY, N. The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters. In: *IEEE Digital Signal Processing Workshop - DSP*. Bryce Canyon, USA: [s.n.], 1998. p. 2543–2560.
- KLEIN, A.; ANDERSSON, J.; ARDEKANI, B. A.; ASHBURNER, J.; AVANTS, B.; CHIANG, M. C.; CHRISTENSEN, G. E.; COLLINS, D. L.; GEE, J.; HELLIER, P.; AL. et. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, v. 46, n. 3, p. 786–802, 02 2009.
- KOVACS, N. Z.; FERRARI, R. J. Detection of 3D salient points in magnetic resonance images using the dual-tree complex wavelet transform. In: *Proceedings of the Conference on Graphics, Patterns and Images (SIBGRAPI) - Undergraduate Work*. Niterói, RJ, Brazil: [s.n.], 2017. p. 1–4.
- KOVESI, P. Phase congruency detects corners and edges. In: *The Australian Pattern Recognition Society Conference: DICTA 2003*. Sydney, Australia: CSIRO, 2003. p. 309–318.
- 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.
- LUO, Y.; TANG, X. Automated diagnosis of Alzheimer's disease with multi-atlas based whole brain segmentations. In: . Orlando, Florida, United States: SPIE MEDICAL IMAGING, 2017. v. 10137, p. 10137 – 10137 – 9.
- MANN, H. B.; WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, v. 18, n. 1, p. 50–60, 1947.
- 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. 840–850, 1996.

MCKHANN, G.; DRACHMAN, D.; FOLSTEIN, M.; KATZMAN, R.; PRICE, D.; STADLAN, E. M. Clinical diagnosis of Alzheimer's disease. *Neurology*, v. 34, n. 7, p. 939–939, 1984.

MORTENSEN, E. N.; DENG, H.; SHAPIRO, L. A SIFT descriptor with global context. In: *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Washington, DC, USA: IEEE Computer Society, 2005. (CVPR '05, v. 01), p. 184–190.

NACHAR, N. The mann-whitney u: A test for assessing whether two independent samples come from the same distribution. *Tutorials in Quantitative Methods for Psychology*, v. 4, 03 2008.

NAJAFABADI, M. M.; VILLANUSTRE, F.; KHOSHGOFTAAR, T. M.; SELIYA, N.; WALD, R.; MUHAREMAGIC, E. Deep learning applications and challenges in big data analytics. *Journal of Big Data*, n. 1, february 2015.

NATIONAL INSTITUTE ON AGING. *How Is Alzheimer's Disease Diagnosed?* [S.l.], 2017. Disponível em: <<https://www.nia.nih.gov/health/how-alzheimers-disease-diagnosed>>.

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, p. 25–31, 2000.

PERL, D. P. Neuropathology of Alzheimer's disease. *Mount Sinai Journal Of Medicine*, v. 77, p. 32–42, january 2010.

POLONI, K. M.; VILLA-PINTO, C. H.; SOUZA, B. S.; FERRARI, R. J. Construction and application of a probabilistic atlas of 3D landmark points for initialization of hippocampus mesh models in brain MR images. In: *Computational Science and Its Applications – ICCSA 2018*. Melbourne, Australia: Springer International Publishing, 2018. p. 310–322.

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.

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.

RISTER, B.; HOROWITZ, M. A.; RUBIN, D. L. Volumetric image registration from invariant keypoints. *IEEE Transactions on Image Processing*, v. 26, n. 10, p. 4900–4910, 2017.

ROMBERG, J. K.; CHOI, H.; BARANIUK, R. G.; KINGSBURY, N. A hidden markov tree model for the complex wavelet transform. In: *IEEE Transactions on Signal Processing*. USA: IEEE, 2001. p. 133–136.

SELESNICK, I. W.; BARANIUK, R. G.; KINGSBURY, N. C. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*, v. 22, n. 6, p. 123–151, 2005.

SELESNICK, I. W.; LI, K. Y. Video denoising using 2D and 3D dual-tree complex wavelet transforms. In: *Proceeding SPIE - Wavelet Applications Signal Image Processing X*. San Diego, CA, USA: Society of Photo Optical, 2003. v. 5207, p. 607–618.

SLED, J. G.; ZIJDENBOS, A. P.; EVANS, A. C. A non-parametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions in Medical Imaging*, v. 17, n. 1, p. 87–97, 1998.

SOLOMON, C.; BRECKON, T. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*. [S.l.]: Wiley, 2011. 1-328 p.

TERZOPOULOS, D.; PLATT, J. C.; BARR, A. H.; FLEISCHER, K. W. Elastically deformable models. In: *ACM SIGGRAPH Computer Graphics*. California, USA: ACM, 1987. v. 21, n. 4.

TOMBARI, F.; SALTI, S.; STEFANO, L. D. Unique shape context for 3d data description. In: *Proceedings of the ACM workshop on 3D object retrieval*. Firenze, Italy: ACM, 2010. p. 57–62.

TRISTÁN-VEGA, A.; GARCÍA-PÉREZ, V.; AJA-FERNÁNDEZ, S.; WESTIN, C.-F. Efficient and robust nonlocal means denoising of MR data based on salient features matching. *Computer Methods and Programs in Biomedicine*, v. 105, n. 2, p. 131–144, 2012.

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

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

VILLA-PINTO, C. H. *Construção e aplicação de atlas de pontos salientes 3D na inicialização de modelos geométricos deformáveis em imagens de ressonância magnética*. Dissertação (Mestrado) — Universidade Federal de São Carlos, March 2016. Disponível em: <<https://repositorio.ufscar.br/handle/ufscar/7861>>.

VILLA-PINTO, C. H.; FERRARI, R. J. Initialization of deformable models in 3D magnetic resonance images guided by automatically detected phase congruency point landmarks. *Pattern Recognition Letters*, v. 79, p. 1–7, 2016.

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.

WEST, M. J.; COLEMAN, P. D.; FLOOD, D. G.; TRONCOSO, J. C. Differences in the pattern of hippocampal neuronal loss in normal ageing and Alzheimer's disease. *The Lancet*, v. 344, p. 769–772, 1994.

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.

ZHENG, Y.; LIU, D.; GEORGESCU, B.; NGUYEN, H.; COMANICIU, D. 3D deep learning for efficient and robust landmark detection in volumetric data. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Munich, Germany: Springer International Publishing, 2015. p. 565–572.

ZITOVA, B.; FLUSSER, J. Image registration methods: a survey. *Image and Vision Computing*, v. 21, p. 977 – 1000, 2003.