

**Métodos híbridos via energias quadráticas
em grafos direcionados para segmentação
de imagens**

Caio Lopes Demario

DISSERTAÇÃO APRESENTADA AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO TÍTULO DE
MESTRE EM CIÊNCIAS

Programa: Ciência da Computação
Orientador: Prof. Dr. Paulo A. V. de Miranda

São Paulo
3 de Maio de 2021

Métodos híbridos via energias quadráticas em grafos direcionados para segmentação de imagens

Caio Lopes Demario

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 3 de Maio de 2021.

Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Paulo A. V. de Miranda (orientador) – IME-USP
- Prof. Dr. Marcos de Sales Guerra Tsuzuki – EP-USP
- Prof. Dr. Letícia Rittner – Unicamp

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Agradecimentos

Prepara-se o cavalo para o dia da batalha, mas o Senhor é que dá a vitória

— Provérbios 21:31

Agradeço primeiramente a Deus, pela graça, suporte e força. Agradeço aos meus pais e irmão por todo o incentivo, por acreditarem em mim e pelas inúmeras caronas para o terminal de ônibus às 5 da manhã.

Ao ao meu orientador Paulo Miranda por todas as aulas, reuniões, dicas e ensinamentos e paciência com os meus horários complicados e minha disponibilidade.

Aos professores Marcos Tsuzuki, Letícia Ritnner e Roberto Hirata, que estiveram nas minhas bancas de qualificação e defesa e me deram muitas sugestões, correções e dicas que tornaram este trabalho muito melhor.

Aos meus amigos de longa data Mylena, Marcia, Gislaine e Calvin pela compreensão nas minhas ausências, pelos encorajamentos, orações e pelos necessários momentos de descontração.

À minha amiga Marcela, que me convenceu a me inscrever para o mestrado, me passou todas as informações e datas e me incentivou a sempre seguir em frente e não desistir. E aos outros amigos da *panelinha* que mesmo sem entender continuaram me motivando.

Aos meus amigos e colegas de trabalho que me ajudaram e incentivaram durante os últimos 4 anos: Felipe e Maiti, Ricardo, Gabriel (Jeff) e Jerry. A ajuda e a confiança de vocês em mim foi essencial para o término deste trabalho.

Ao meu amigo Gustavo, que praticou a defesa comigo muitas vezes e me deu muitas dicas, e ao meu amigo Iago, que passou vários sábados jogando *Fall Guys* comigo e me dando dicas sobre a escrita da dissertação e sobre o momento da defesa.

Obrigado a todos que tornaram esre trabalho possível.

Resumo

Caio Lopes Demario. **Métodos híbridos via energias quadráticas em grafos direcionados para segmentação de imagens**. Dissertação (Mestrado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2021.

Neste trabalho são investigados métodos híbridos para segmentação de imagens, tomando como base o algoritmo dos Passeios Aleatórios (Random Walks - RW), porém adotando grafos direcionados, de modo a explorar a polaridade de borda dos objetos. Como resultado temos um novo método híbrido baseado em sementes, chamado OIFT Relaxada (Relaxed OIFT - ROIFT), que estende o método proposto por Malmberg et al. para grafos direcionados, devidamente incorporando a restrição de polaridade de borda. Os resultados de segmentação via ROIFT se encontram entre os obtidos pela Transformada Imagem-Floresta Orientada (OIFT) e pela extensão do RW para grafos direcionados, tal como proposto por Singaraju et al., com estes refletindo casos particulares extremos do novo método. A ROIFT é avaliada em imagens médicas de Ressonância Magnética e Tomografia Computadorizada, produzindo resultados de segmentação superiores e mais intuitivamente corretos que os obtidos por Singaraju et al. e OIFT, além de ser facilmente extensível para imagens multidimensionais. Também é proposto um método híbrido chamado *Relaxed Deep Extreme Cut*, que estende o método *Deep Extreme Cut*, de Maninis et al., atuando em algumas de suas deficiências. Esse método é avaliado em imagens bidimensionais do conjunto de dados do GrabCut, apresentando resultados superiores aos obtidos pelo método de Maninis et al.

Palavras-chave: Segmentação de imagens, Passeios Aleatórios, Transformada Imagem-Floresta.

Abstract

Caio Lopes Demario. **Hybrid methods via quadratic energies in directed graphs for image segmentation**. Thesis (Masters). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2021.

In this work, we investigate hybrid methods for seeded image segmentation, based on the Random Walks algorithm (RW), but adopting directed graphs, in order to explore the boundary polarity of the objects. As results, we have a new seeded hybrid method, named Relaxed OIFT, which extends a method by Malmberg et al. to directed graphs, to properly incorporate the boundary polarity constraint. Relaxed OIFT lies between the pure Oriented Image Foresting Transform (OIFT) at one end and the extension of Random Walks (RW) to directed graphs as proposed by Singaraju et al., being OIFT and RW particular cases of the proposed method. Relaxed OIFT is evaluated in MR and CT medical images, producing more intuitively correct segmentation results than both OIFT and RW, and being easy to be extended to multi-dimensional images. We also propose a hybrid method, named Relaxed Deep Extreme Cut, which extends the method Deep Extreme Cut by Maninis et al., acting on some of its deficiencies. This method is evaluated in bidimensional images of the GrabCut Dataset, producing results superior to those obtained by the method of Maninis et al.

Keywords: Image segmentation, Random Walks, Image Foresting Transform.

Lista de Abreviaturas

CT	Tomografia Computadorizada (<i>Computed tomography</i>)
MR	Ressonância Magnética (<i>Magnetic resonance</i>)
RW	Passeios Aleatórios (<i>Random Walks</i>)
RFC	<i>Relative Fuzzy Connectedness</i>
GC	Corte em Grafo (<i>Graph Cut</i>)
GGC	Corte em Grafo Generalizado (<i>Generalized Graph Cut</i>)
IFT	Transformada Imagem-Floresta (<i>Image Foresting Transform</i>)
OIFT	Transformada Imagem-Floresta Orientada (<i>Oriented Image Foresting Transform</i>)
MI	Monotonicamente incremental (<i>Monotonically Incremental</i>)
DEXTR	<i>Deep Extreme Cut</i>
CNN	Rede Neural Convolutacional (<i>Convolutional Neural Network</i>)
ROIFT	Relaxed OIFT
RGB	<i>Red-Green-Blue</i>

Lista de Símbolos

\hat{I}	Imagem digital
\mathcal{V}	Conjunto de vértices de um grafo
\mathcal{A}	Conjunto de Arcos de um grafo
\mathcal{N}	Mapeamento de pixel para pixels adjacentes
\mathcal{S}_1	Conjunto dos pixels semente do objeto
\mathcal{S}_0	Conjunto dos pixels semente do fundo
\mathcal{L}	Mapeamento de vértices para rótulos
π_t	Caminho com término em um vértice t
$\pi_{s \rightsquigarrow t}$	Caminho com início em um vértice s término em um vértice t
$\Pi(G, t)$	Conjunto de caminhos no grafo G com término em um vértice t
$\omega : \mathcal{A} \rightarrow \mathbb{R}$	Função de peso dos arcos
ω_{st}	Abreviação de $\omega(\langle s, t \rangle)$

Lista de Figuras

1.1	Exemplos de segmentação por <i>Deep Extreme Cut</i>	3
1.2	IFT e IFT após procedimento de relaxação.	4
1.3	Ilustração do problema de ambiguidade devido à falta de informação de polaridade de borda.	5
2.1	Exemplo de imagem digital.	7
2.2	Exemplo de imagem tridimensional.	8
2.3	Exemplo de imagem e grafo gerado a partir da imagem.	9
2.4	Exemplos de adjacências circulares em grafos de imagens.	10
2.5	Exemplo de imagem tridimensional com adjacência circular de $\rho = 1$	11
2.6	Simulação do Algoritmo 1 com função de custo da Equação 2.3. (parte 1)	14
2.6	Simulação do Algoritmo 1 com função de custo da Equação 2.3. (parte 2)	15
2.6	Simulação do Algoritmo 1 com função de custo da Equação 2.3. (parte 3)	16
2.7	Resultado de segmentação utilizando restrição de polaridade de borda do objeto.	17
2.8	Imagem sintética com regiões de brilhos diferentes.	19
2.9	Simulação do algoritmo da OIFT com função de custo da Equação 2.8. (parte 1)	20
2.9	Simulação do algoritmo da OIFT com função de custo da Equação 2.8. (parte 2)	21
2.10	Exemplo de passeios aleatórios em um grafo unidimensional.	22
2.11	Circuito com N resistores em série e diferença de potencial $1V$ entre os nós 0 e N	23
3.1	Configuração do diodo resistor entre um par de nós do grafo.	28
3.2	Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.4 (Parte 1).	32
3.2	Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.4 (Parte 2).	33
3.3	Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.5 (Parte 1).	34

3.3	Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.5 (Parte 2).	35
3.4	Exemplo da segmentação de uma imagem natural por Deep Extreme Cut.	38
4.1	ROIFT - Resultados experimentais de segmentação do osso tálus.	43
4.2	ROIFT - Resultados experimentais de segmentação do osso calcâneo	44
4.3	ROIFT - Resultados experimentais de segmentação do fígado.	45
4.4	Segmentação do osso tálus.	46
4.5	Segmentação do osso calcâneo.	47
4.6	Segmentação do fígado em CT.	48
4.7	Exemplo #1: Segmentação 3D do cérebro em imagem de RM.	50
4.8	Exemplo #2: Segmentação 3D do cérebro em imagem de RM.	51
4.9	Exemplo #3: Segmentação 3D do cérebro em imagem de RM.	52
4.10	Exemplo #4: Segmentação 3D do cérebro em imagem de RM.	53
4.11	Exemplo #5: Segmentação 3D do cérebro em imagem de RM.	54
4.12	Exemplo #6: Segmentação 3D do cérebro em imagem de RM.	55
4.13	Exemplo #7: Segmentação 3D do cérebro em imagem de RM.	56
4.14	Exemplo de segmentação da imagem 227092.jpg da base do grabcut.	59
4.15	Exemplo de segmentação da imagem teddy.jpg da base do grabcut.	60
4.16	Exemplo de segmentação da imagem bool.jpg da base do grabcut.	61

Lista de Tabelas

4.1	Matriz de confusão para segmentação e gabarito da segmentação.	40
4.2	Resultados experimentais na base do Grabcut.	58

Sumário

1	Introdução	1
1.1	Objetivos	6
1.2	Organização do Trabalho	6
2	Conceitos e Métodos Relacionados	7
2.1	Imagens Digitais	7
2.1.1	Imagem multidimensional	8
2.2	Grafos	8
2.2.1	Relações de Adjacência	8
2.3	Grafos de Imagens	8
2.3.1	Grafos de Imagens tridimensionais	10
2.4	Transformada Imagem-Floresta	10
2.5	Transformada Imagem-Floresta Orientada	17
2.5.1	Polaridade de borda	17
2.5.2	Descrição do Algoritmo	18
2.5.3	Controlando o parâmetro α	19
2.5.4	Simulação do algoritmo da OIFT	20
2.6	Passeios Aleatórios	22
2.6.1	Motivação	22
2.6.2	Analogia com circuitos	22
2.6.3	Minimização de energia	24
2.7	Método das Relaxações	25
2.8	Deep Extreme Cut	25
3	Métodos híbridos via energias quadráticas em grafos direcionados	27
3.1	Método das relaxações com pesos assimétricos	27
3.2	OIFT Relaxada	30
3.3	Relaxed Deep Extreme Cut	36

4	Resultados Experimentais	39
4.1	Medidas de acurácia	39
4.1.1	Coeficiente de Dice	39
4.1.2	Coeficiente de Jaccard	40
4.1.3	Métrica Balanceada Geral (GB)	40
4.1.4	Quociente Isoperimétrico	41
4.2	OIFT Relaxada	41
4.2.1	Implementação em 3D	49
4.3	Relaxed Deep Extreme Cut	57
5	Conclusão	63
5.1	Trabalhos futuros	64

Apêndices

Anexos

Referências	65
--------------------	-----------

Capítulo 1

Introdução

A segmentação de imagens para extrair um objeto de um fundo é um tópico muito estudado em processamento de imagens e visão computacional, sendo útil para muitas aplicações, como na edição de vídeos/fotos e na análise de imagens médicas e biológicas (Bai e Sapiro, 2007; Maninis *et al.*, 2018). A segmentação precisa de objetos de interesse geralmente é difícil devido à presença de fundos heterogêneos, objetos com bordas mal definidas, inomogeneidade de campo, ruído, artefatos, efeitos de volume parcial e seus efeitos conjuntos (Madabhushi e J.K. Udupa, 2005), bem como devido à potencial complexidade dos objetos de interesse, com elevado padrão de variabilidade. Portanto, a incorporação de conhecimento prévio no algoritmo de segmentação é útil para obter resultados mais precisos.

Uma classe importante de métodos de segmentação abrange os métodos baseados em grafos sujeitos a restrições rígidas, nos quais uma rotulagem parcial, fornecida pela seleção de pixels sementes, é propagada a todos os pixels não rotulados, para a partição ótima de um grafo de imagem. A segmentação obtida pode ser então editada posteriormente pela adição de novas sementes de objetos e fundo a fim de melhorar os resultados. Exemplos incluem o método das Linhas Divisoras de Água (*Watersheds*) (Cousty *et al.*, 2010; Audigier e Lotufo, 2005), Passeios Aleatórios (*Random Walks - RW*) (Grady, 2006), Conexidade Fuzzy Relativa (*Relative Fuzzy Connectedness - RFC*) (Ciesielski, J.K. Udupa *et al.*, 2012), Corte em grafo (*Graph Cuts - GC*) (Y.Y. Boykov e Jolly, 2001), Distância da Barreira (*Barrier Distance*) (Strand *et al.*, 2013) e Corte de Crescimento (*Grow Cut*) (Vezhnevets e Konouchine, 2005). Com base em vários estudos teóricos de suas relações (Sinop e L.Grady, 2007; P. Miranda e Falcão, 2009; Ciesielski, J.K. Udupa *et al.*, 2012; C.Couprrie *et al.*, 2010; Allène *et al.*, 2007; P. Miranda e Falcão, 2011), métodos baseados em conectividade (Bai e Sapiro, 2007; Strand *et al.*, 2013; Protiere e Sapiro, 2007; Vezhnevets e Konouchine, 2005) e alguns métodos baseados em cortes em grafos puderam ser descritos de uma maneira unificada de acordo com os arcabouços (*frameworks*) da Transformada Imagem-Floresta (*Image Foresting Transform - IFT*) (Falcão, Stolfi e Lotufo, 2004; Ciesielski, Falcão e P. Miranda, 2018) e Corte em Grafo Generalizado (*Generalized Graph Cut - GGC*) (Ciesielski, J.K. Udupa *et al.*, 2012; C.Couprrie *et al.*, 2010), respectivamente.

Em trabalhos mais recentes, a extensão para grafos direcionados de vários métodos, incluindo RW (Singaraju, Grady e Vidal, 2008), IFT Orientada (OIFT) (P. Miranda e Mansilla,

2014), RFC Orientada (ORFC) (Bejar e P. Miranda, 2015) e GC (Y. Boykov e Funka-Lea, 2006), possibilitaram o suporte ao processamento de propriedades globais de objetos, como restrições de forma (Gulshan *et al.*, 2010; Braz *et al.*, 2019; Mansilla e P. Miranda, 2013a), polaridade de borda (Singaraju, Grady e Vidal, 2008; P. Miranda e Mansilla, 2014) e restrições hierárquicas (DeLong e Y. Boykov, 2009; Leon e P. Miranda, 2017), permitindo sua customização para um dado objeto.

A segmentação de imagem também pode ser interpretada como um problema de classificação no nível de pixel. Não surpreendentemente, métodos baseados em aprendizagem de máquina estão entre as soluções mais proeminentes para o problema da segmentação, especialmente após o advento de técnicas de aprendizagem profunda (He *et al.*, 2016). Nesse contexto, um método avançado do estado-da-arte na literatura no estilo *grabcut* (Rother, Kolmogorov e Blake, 2004) para segmentação interativa com mínimo envolvimento do usuário é o *Deep Extreme Cut* (Maninis *et al.*, 2018). Ele considera pontos extremos (pixels mais à esquerda, mais à direita, superior e inferior nos limites do objeto) como 4 cliques para cada objeto e a Rede Neural Convolutiva (*Convolutional Neural Network* - CNN) produz as máscaras segmentadas (Figura 1.1). Esse tipo de entrada do usuário geralmente é uma boa opção para algoritmos de execução única (*single-shot algorithms*), pois dos 4 pontos indicados podemos inferir a caixa delimitadora do objeto e, conseqüentemente, todos os pontos fora da caixa pertencem ao plano de fundo. À primeira vista, isso parece ser melhor do que outros tipos de entrada do usuário, pois pontos arbitrários não nos permitem tirar conclusões adicionais. No entanto, em geral, não é possível obter um resultado de segmentação com qualidade de gabarito usando apenas uma seleção de 4 pontos, pois erros residuais consideráveis ao redor das bordas são bastante comuns (Figuras 1.1c e 1.1f). Por outro lado, as técnicas de aprendizado profundo exigem muitos dados anotados para o treinamento da rede e, para muitas aplicações, a disponibilidade do padrão ouro é bastante limitada.

Já que não há um método único que funciona melhor em todas as aplicações e critérios de avaliação, uma abordagem comum é combiná-los em métodos híbridos, a fim de atingir uma boa combinação entre sua robustez, aderência de borda, regularidade de forma e eficiência (Ciesielski, P. Miranda *et al.*, 2013). Por exemplo, Price *et al.* (Price, Morse e Cohen, 2010) propõem a combinação de uma abordagem de segmentação geodésica com o método de corte em grafo (GC), de modo a aumentar a acurácia. No entanto, os resultados de segmentação por esses métodos apresentam grande sensibilidade em relação à disposição das sementes, dado que ambas abordagens utilizadas (GC e Dijkstra) são sensíveis às sementes. Ciesielski *et al.* (Ciesielski, P. Miranda *et al.*, 2013) propõem uma associação entre os algoritmos de segmentação GC e RFC, preservando a robustez do RFC, e ao mesmo tempo mantendo o controle do GC sobre segmentos de borda mal definidos. Em outro método híbrido, uma IFT relaxada é proposta por Malmberg *et al.* (Malmberg *et al.*, 2010), a qual combina RW e IFT (Figura 1.2). Sua robustez é demonstrada pela redução da variação intra-usuário para segmentar músculos do tronco em imagens de Ressonância Magnética (RM) da região lombar (Malmberg *et al.*, 2010). Posteriormente, esse método híbrido foi também avaliado por Moya no contexto da IFT diferencial (Moya, 2015). Algoritmos baseados puramente na imagem podem também ser combinados com abordagens baseadas em modelos de aparência/forma, em uma tentativa de superar algumas deficiências particulares das abordagens, tal como nos métodos híbridos propostos por Udupa *et al.* (JK Udupa

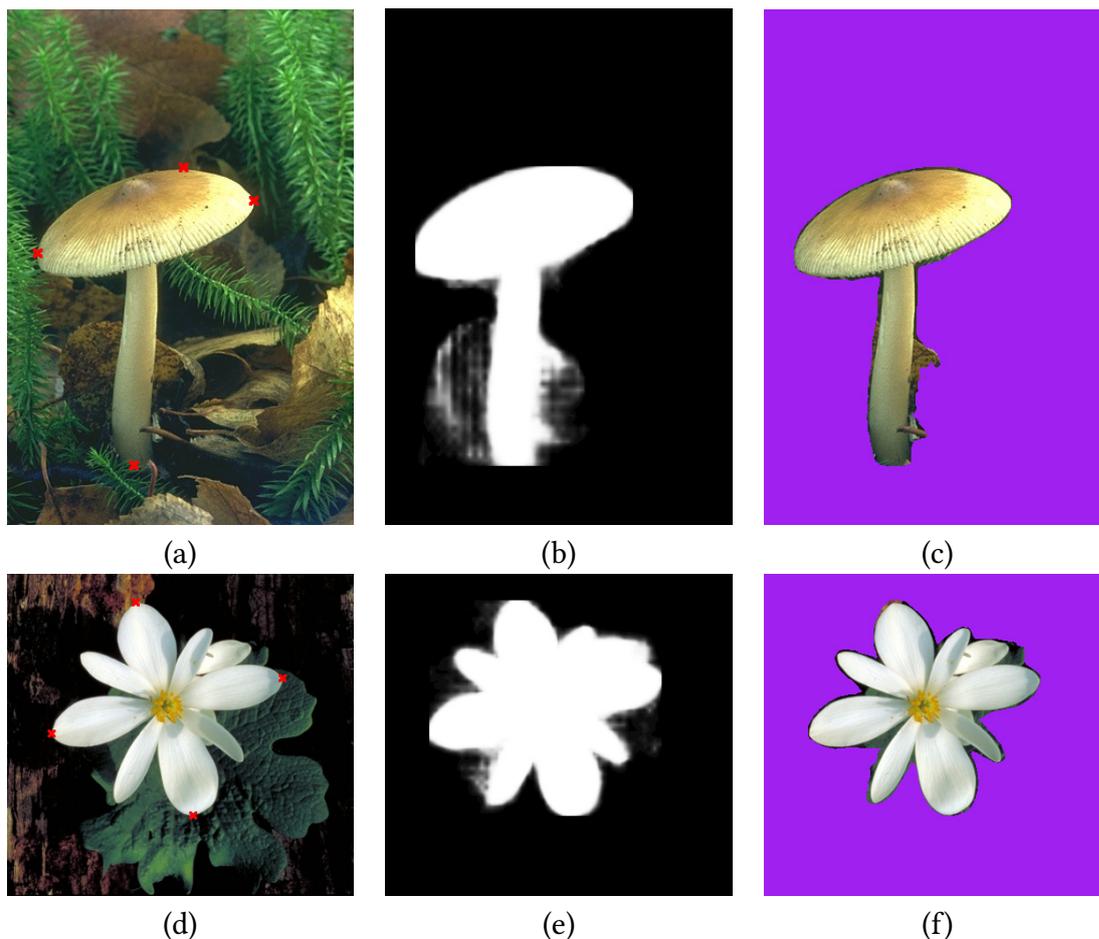


Figura 1.1: Exemplos de segmentação por Deep Extreme Cut. (a) e (d): Imagens de entrada com a seleção dos quatro pontos extremos para segmentar o cogumelo e a flor, respectivamente. (b) e (e): Mapas de probabilidade calculados pela Rede Neural Convolutiva. (c) e (f): As máscaras segmentadas são obtidas por limiarização dos mapas de probabilidade.

et al., 2014; J.Liu, 2006). Outro exemplo consiste em explorar as forças complementares de abordagens baseadas em restrições por região (marcadores/sementes) com abordagens de perseguição de borda, como na abordagem interativa do *Live Markers* (Spina, de Miranda e Xavier Falcão, 2014), em que segmentos ótimos de borda entre pontos de ancoragem são transformados em marcadores internos e externos para o delineamento por região.

Mais tarde, alguns métodos híbridos, como (Ciesielski, P. Miranda *et al.*, 2013), também foram estendidos para grafos direcionados (Bejar e P. Miranda, 2015). Novos métodos híbridos também foram propostos para digrafos, tais como o trabalho de Bejar et al. (Bejar, F. Cappabianco e P. Miranda, 2017), que combina *Live Wire* (LW) com OIFT, dados os seus pontos fortes complementares. Contudo, sua extensão para imagens tridimensionais ainda não é suportada.

Neste trabalho, visamos o desenvolvimento de métodos híbridos via energias quadráticas em grafos direcionados para segmentação de imagens, a fim de explorar a polaridade de borda dos objetos. A Figura 1.3 ilustra o problema teórico de ambiguidade devido à falta da informação da polaridade de borda. A imagem sintética da Figura 1.3a apresenta

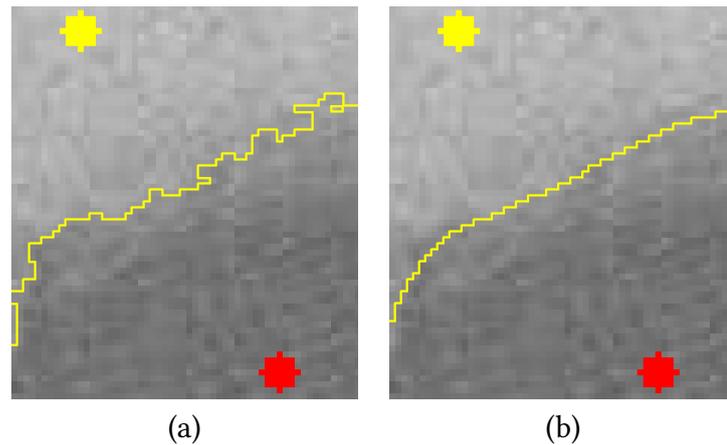


Figura 1.2: (a) Segmentação obtida da IFT em uma imagem de CT do fígado com dois pontos de semente, mostrados aqui como círculos. (b) Aplicando 100 iterações do procedimento de relaxação proposto por Malmberg et al. (Malmberg et al., 2010), é obtido um limite de segmentação mais suave e intuitivamente correto.

duas bordas similares candidatas, com transições dos tipos escuro/claro e claro/escuro da esquerda para a direita. Na Figura 1.3b é fornecida uma segmentação inicial hipotética com bordas irregulares. Observe que nesse caso o refinamento da segmentação via relaxações sucessivas tal como proposto por Malmberg et al. (Malmberg et al., 2010) não teria como distinguir apropriadamente entre as duas soluções apresentadas nas Figuras 1.3c e 1.3d, devido à simetria do problema. Portanto, a incorporação de conhecimento prévio dos objetos a serem segmentados é necessária para a resolução do conflito, através de algoritmos em grafos direcionados que possam distinguir as direções opostas das transições de borda.

Como resultado, estendemos Malmberg et al. (Malmberg et al., 2010) para digrafos e propomos a OIFT Relaxada (*Relaxed OIFT*) (C. L. Demario e P. A. V. Miranda, 2019). A OIFT tem como vantagens sua eficiência, flexibilidade (suportando diversas restrições de alto nível) e um bom equilíbrio entre robustez e aderência de borda (Tavares, Bejar e P. Miranda, 2017), mas possui deficiências com relação à regularidade de forma. Nosso método híbrido reside entre a OIFT pura em uma extremidade e a extensão do RW para grafos direcionados de Singaraju et al. (Singaraju, Grady e Vidal, 2008), produzindo resultados de segmentação mais intuitivamente corretos que os dos métodos individuais (OIFT e o RW).

Também é proposto o método híbrido entre o *Deep Extreme Cut* e a extensão do RW para grafos direcionados de Singaraju et al. (Singaraju, Grady e Vidal, 2008), tirando vantagem do tempo de entrada do usuário utilizando os quatro pontos extremos do *Deep Extreme Cut* e atenuando sua deficiência na aderência de borda através do componente RW.

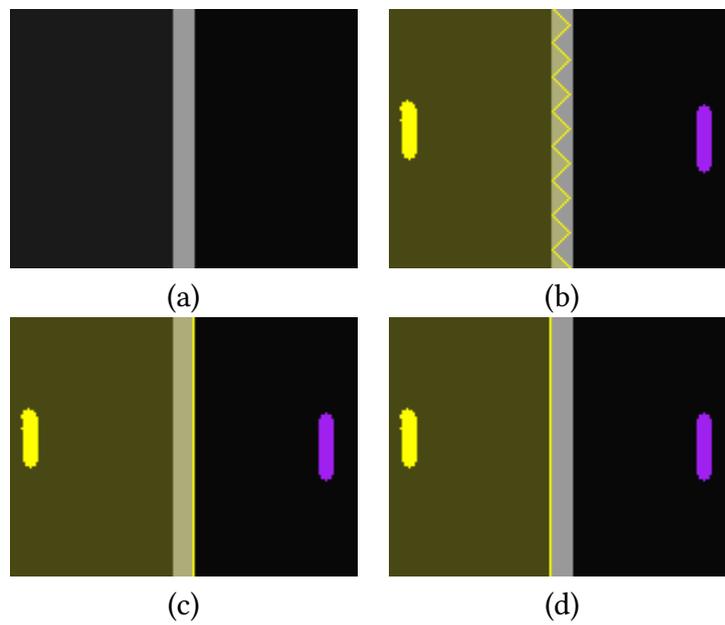


Figura 1.3: Ilustração do problema teórico de ambiguidade devido a falta da informação da polaridade de borda. (a) Imagem sintética com transições de escuro para claro e claro para escuro. (b) Uma segmentação irregular inicial hipotética. (c-d) Possíveis resultados candidatos por relaxações sucessivas.

1.1 Objetivos

- Estender Malmberg et al. (Malmberg *et al.*, 2010) para digrafos, a fim de explorar a polaridade de borda dos objetos, combinando o algoritmo da OIFT (P. Miranda e Mansilla, 2014) com a extensão do *Random Walks* para grafos direcionados (Singaraju, Grady e Vidal, 2008), via minimização de energias quadráticas, em um novo método híbrido, de modo a tirar proveito de suas forças complementares.
 - Pretende-se avaliar o método descrito acima em imagens médicas 2D e 3D, nas quais os objetos de interesse apresentem transições nas bordas com polaridade conhecida (isto é, transições do tipo claro para escuro ou escuro para claro), a fim de reduzir as ambiguidades na segmentação.
 - Espera-se obter bordas de segmentação mais suaves em comparação com as bordas irregulares geralmente produzidas pela OIFT (P. Miranda e Mansilla, 2014) e com melhor acurácia em comparação com Malmberg et al. (Malmberg *et al.*, 2010) e Moya (Moya, 2015).
- Avaliar o emprego de relaxação seguindo a minimização de energias quadráticas no desenvolvimento de outro método híbrido, tomando como base o algoritmo de *machine learning* do *Deep Extreme Cut* (Maninis *et al.*, 2018).
 - Pretende-se avaliar o método acima na base pública de imagens naturais do Grabcut (Rother, Kolmogorov e Blake, 2004).
 - Espera-se melhorar a aderência à borda (acurácia) da segmentação resultante.

1.2 Organização do Trabalho

O Capítulo 2 apresenta conceitos e métodos relacionados necessários para o desenvolvimento dos métodos propostos. Os métodos híbridos propostos são apresentados no Capítulo 3. Os experimentos e resultados obtidos são apresentados no Capítulo 4. Finalmente, no Capítulo 5, apresentamos nossas conclusões finais e sugestões de trabalhos futuros.

Capítulo 2

Conceitos e Métodos Relacionados

2.1 Imagens Digitais

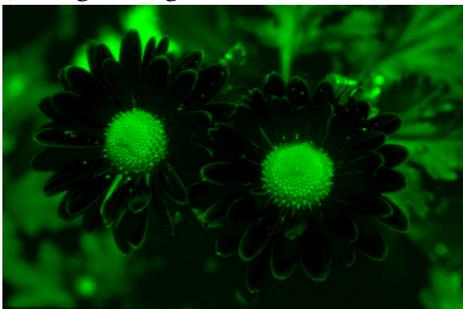
Uma imagem digital é definida como $\hat{I} = \langle D_I, \vec{I} \rangle$, com $D_I \subset \mathbb{Z}^n$ e n representando a dimensão da imagem. Os pontos em D_I são denominados *spels* (*spacial elements*), ou pixels (*picture elements*), no caso bidimensional. \vec{I} é um mapeamento que associa a cada *spel* um vetor $\vec{I}(p) = \{I_1(p), I_2(p), \dots, I_k(p)\} \in \mathbb{Z}^k$ de escalares que representam propriedades físicas da imagem, sendo k o número de bandas (por exemplo, cores no espaço RGB).



(a) imagem digital bidimensional RGB



(b) banda R (vermelho)



(c) banda G (verde)



(d) banda B (azul)

Figura 2.1: Exemplo de imagem digital bidimensional e com $k = 3$, onde \vec{I} associa um pixel a uma cor no espaço RGB.

2.1.1 Imagem multidimensional

Uma imagem $\hat{I} = \langle D_I, \vec{I} \rangle$ é dita multidimensional quando $D_I \subset \mathbb{Z}^n$, $n > 2$. Como exemplos de imagens multidimensionais temos as imagens médicas tridimensionais, formadas a partir de seqüências de fatias tomográficas ou de ressonância magnética (Figura 2.2a). No caso de $n = 3$, os *spels* são chamados de *voxels*, que são representados pelas coordenadas x , y e z e essa imagem é denominada volume (Figura 2.2b).

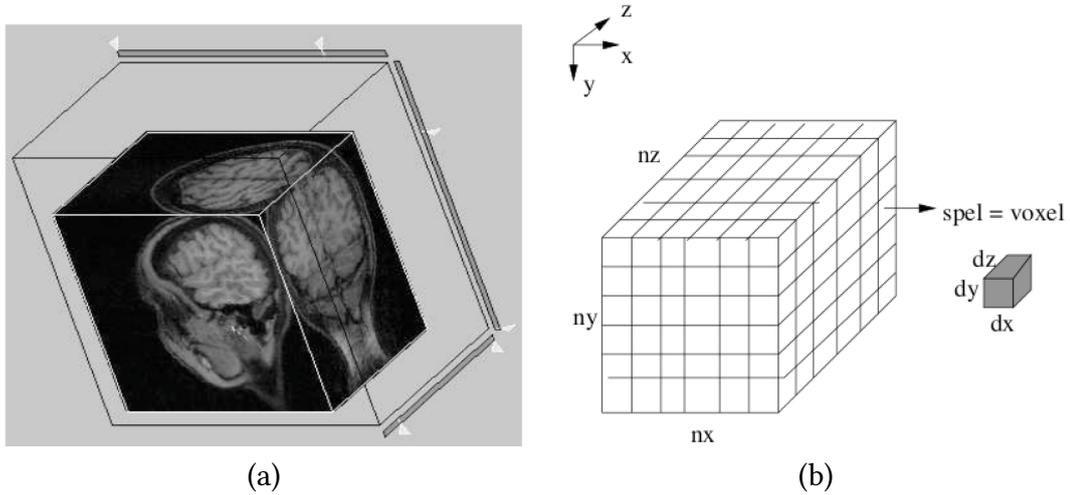


Figura 2.2: (a) Exemplo de imagem tridimensional gerada a partir de seqüências de fatias de ressonância magnética e (b) sua representação de voxels em um volume. Adaptado de [Paulo A. V. de Miranda, 2020](#)

2.2 Grafos

Um grafo é definido como um par $G = \langle \mathcal{V}, \mathcal{A} \rangle$, onde \mathcal{V} é um conjunto finito e não vazio que representa os vértices e $\mathcal{A} \subset \mathcal{V} \times \mathcal{V}$ é um conjunto que representa os arcos.

2.2.1 Relações de Adjacência

Dados dois nós de um grafo p e q , dizemos que p é adjacente a q se $\langle p, q \rangle \in \mathcal{A}$. No caso de grafos simétricos, $\langle p, q \rangle \in \mathcal{A} \rightarrow \langle q, p \rangle \in \mathcal{A}$.

2.3 Grafos de Imagens

Uma imagem pode ser interpretada como um digrafo ponderado $G = \langle \mathcal{V}, \mathcal{A}, \omega \rangle$, cujos vértices \mathcal{V} são os pixels (ou superpixels) da imagem no domínio da imagem e cujos arcos são os pares ordenados $\langle s, t \rangle \in \mathcal{A}$ de pixels vizinhos, e.g., vizinhança-4 (Figura 2.3) ou vizinhança-8 em caso de imagens 2D.

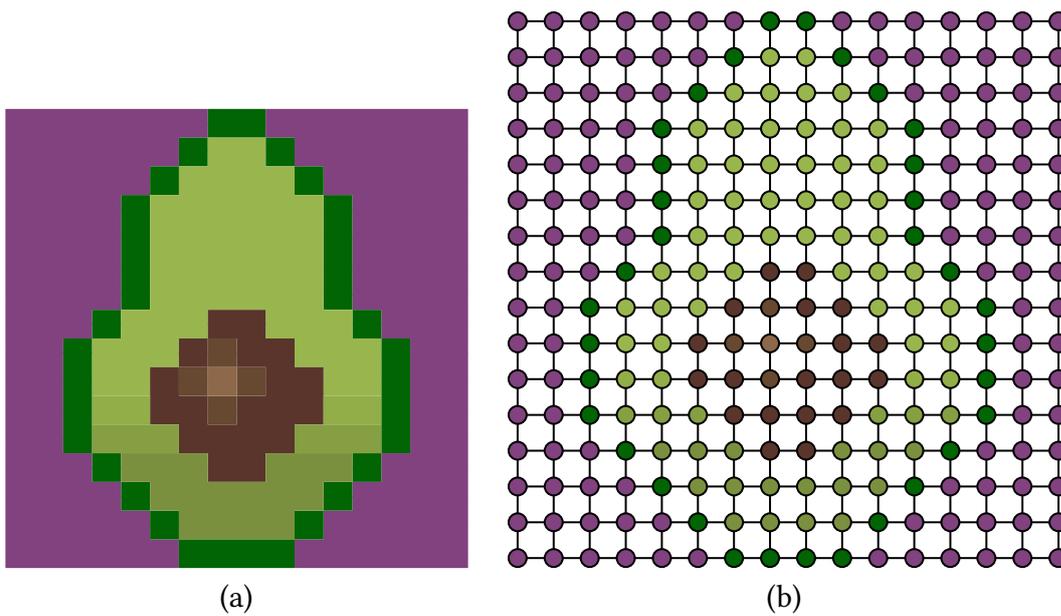


Figura 2.3: Exemplo de imagem e grafo gerado a partir da imagem com cada pixel sendo representado por um nó e utilizando vizinhança-4 como adjacência.

O conjunto de nós adjacentes a s é denotado por $\mathcal{N}(s) = \{t \in \mathcal{V} : \langle s, t \rangle \in \mathcal{A}\}$ e é denominado vizinhança de s . Alguns exemplos de relações de adjacência para imagens bidimensionais são:

- **Circular:** $\langle p, q \rangle \in \mathcal{A}$ se $d(p, q) \leq \rho$ sendo d a função distância euclidiana e ρ um escalar. Na Figura 2.4 são apresentados alguns exemplos de adjacências circulares. A Figura 2.4a representa o que é chamado vizinhança-4 e a Figura 2.4b representa o que é chamado vizinhança-8. Outro exemplo de adjacência circular com maior raio ρ é apresentado na Figura 2.4c
- **Retangular:** $\langle p, q \rangle \in \mathcal{A}$ se $|x_q - x_p| \leq \frac{a}{2}$ e $|y_q - y_p| \leq \frac{b}{2}$, em que a e b são os comprimentos dos lados de um retângulo com centro em $\langle x_p, y_p \rangle$.
- **Baseada em conjunto:** $\langle p, q \rangle \in \mathcal{A}$ se $q - p \in \{(-1, -1), (1, -1)\}$.
- **Baseada em propriedades locais:** $\langle p, q \rangle \in \mathcal{A}$ se $d(p, q) \leq \rho_i$ e $d(\vec{I}(p), \vec{I}(q)) \leq \rho_a$, em que ρ_i e ρ_a são os raios nos espaços da imagem e de atributos.

Utilizamos $\omega_{st} \geq 0$ para abreviar $\omega(\langle s, t \rangle)$, em que $\omega : \mathcal{A} \rightarrow \mathbb{R}$ é a função de peso dos arcos. O peso dos arcos pode ser calculado de várias formas baseado em propriedades da imagem e dos objetos de interesse. O digrafo G é simétrico se para cada um de seus arcos $\langle s, t \rangle \in \mathcal{A}$, o par $\langle t, s \rangle$ também é um arco de G , mas podemos ter $\omega_{st} \neq \omega_{ts}$.

Para um dado digrafo de imagem G , um *caminho* é uma sequência de vértices $\pi = \langle t_1, t_2, \dots, t_k \rangle$, na qual $\langle t_i, t_{i+1} \rangle \in \mathcal{A}$, para $1 \leq i \leq k - 1$, e $t_i \neq t_j$ para $i \neq j$. Um caminho é dito ser *trivial* quando $k = 1$. Usamos π_t para indicar um caminho com término em um vértice t . Para indicar explicitamente a origem do caminho, usamos a notação $\pi_{s \rightsquigarrow t} = \langle s = t_1, t_2, \dots, t_k = t \rangle$, em que s indica a origem e t o término do caminho. Usamos $\pi_{r \rightsquigarrow s} \cdot \langle s, t \rangle = \langle r, \dots, s, t \rangle$ para denotar a extensão de um caminho $\pi_{r \rightsquigarrow s} = \langle r, \dots, s \rangle$ de r até s por um arco

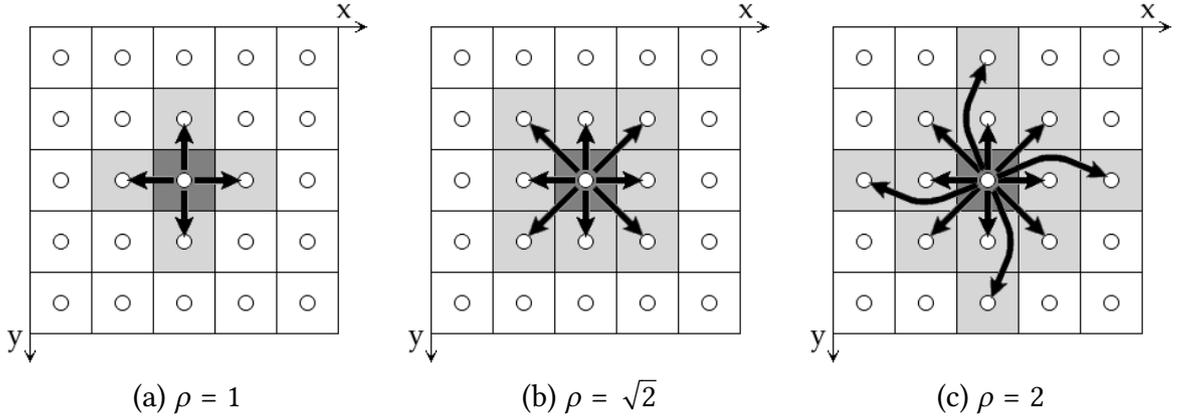


Figura 2.4: Exemplos de adjacências circulares em grafos de imagens bidimensionais. Adaptado de Paulo A. V. de Miranda, 2020

$\langle s, t \rangle$. Denotamos por $\Pi(G, t)$ o conjunto de todos caminhos no digrafo G com término em t , enquanto que $\Pi(G) = \bigcup_{t \in \mathcal{V}} \Pi(G, t)$ denota todos possíveis caminhos existentes no digrafo G .

Um digrafo é fortemente conexo (*strongly connected*) se para qualquer par $\langle s, t \rangle$ de seus vértices existe um caminho de s a t e também um caminho de t para s .

Por simplicidade, vamos nos focar em segmentação de objeto/fundo (*object/background*), indicada por dois conjuntos de sementes $S_1 \subset \mathcal{V}$ e $S_0 \subset \mathcal{V}$, respectivamente. Ela pode ser um mapeamento rígido (*crisp*) $\mathcal{L} : \mathcal{V} \rightarrow \{0, 1\}$, o qual atribui para cada vértice um rótulo (1 para o objeto e 0 para o fundo), ou um mapeamento não-rígido (*fuzzy*) $\mathcal{L} : \mathcal{V} \rightarrow [0, 1]$, no qual $\mathcal{L}(s)$ indica o grau de pertinência de s ao objeto. Ambos os casos estão sujeitos a $\mathcal{L}(s) = 1$ para $s \in S_1$ e $\mathcal{L}(s) = 0$ para $s \in S_0$. Note que, uma segmentação \mathcal{L} rígida é a *função característica* χ_P do subconjunto $P = \{s \in \mathcal{V} : \mathcal{L}(s) = 1\}$ de \mathcal{V} .

2.3.1 Grafos de Imagens tridimensionais

Também podemos utilizar um grafo para representar uma imagem tridimensional de forma semelhante a imagens bidimensionais. No caso de imagens tridimensionais, cada nó $t \in \mathcal{V}$ representa um *voxel* do volume e definimos as relações de adjacência de forma análoga às adjacências em imagens bidimensionais. Na Figura 2.5 temos um exemplo de adjacência circular em um volume com $\rho = 1$, representando o que é chamado de vizinhança-6 de voxels.

A seguir são apresentados alguns métodos relacionados que servem de fundamento para os resultados obtidos nesta dissertação.

2.4 Transformada Imagem-Floresta

Uma *função de conectividade* é uma função $f : \Pi(G) \rightarrow \mathcal{C}$ que associa um valor escalar de custo de caminho $f(\pi)$, em um conjunto ordenado \mathcal{C} de valores de custo (e.g., $\mathcal{C} = \mathbb{R}$), a qualquer caminho π no grafo G , com base em propriedades da imagem ao longo deste

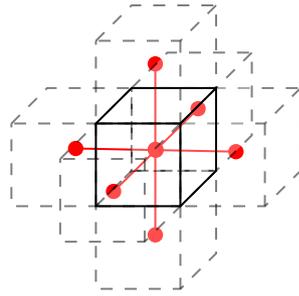


Figura 2.5: Exemplo de imagem tridimensional com adjacência circular de $\rho = 1$.

caminho. Geralmente, o custo do caminho depende dos pesos ω dos arcos ao longo do caminho.

A Transformada Imagem-Floresta (*Image Foresting Transform* - IFT) (Falcão, Stolfi e Lotufo, 2004) é uma modificação do algoritmo de caminho mais curto de Dijkstra (Dijkstra, 1959), permitindo múltiplas fontes e funções de conexidade mais gerais, por exemplo, incluindo a classe de funções monotonicamente incrementais (*Monotonically Incremental (MI) functions*). Uma função monotonicamente incremental é uma função definida pela seguinte equação:

$$\begin{aligned} f_{mi}(\langle t \rangle) &= H(t) \\ f_{mi}(\pi_s \cdot \langle s, t \rangle) &= f_{mi}(\pi_s) \odot \langle s, t \rangle \end{aligned} \quad (2.1)$$

em que $H(t)$ representa um valor inicial e $\odot : \mathcal{C} \times \mathcal{A} \rightarrow \mathcal{C}$ é uma operação binária entre o custo de um caminho e um arco. Essa operação binária satisfaz as seguintes condições:

1. $a \geq b \rightarrow a \odot \langle s, t \rangle \geq b \odot \langle s, t \rangle$
2. $a \odot \langle s, t \rangle \geq a$

em que $a, b \in \mathcal{C}$ e $\langle s, t \rangle \in \mathcal{A}$.

Um mapa de predecessores é uma função P que atribui para cada pixel t em \mathcal{V} , ou um outro pixel adjacente em \mathcal{V} , ou um marcador distintivo nil que não está em \mathcal{V} . Uma floresta de espalhamento é um mapa de predecessores que não contém ciclos. Logo, para qualquer pixel $t \in \mathcal{V}$, uma floresta de espalhamento P define um caminho π_t^P recursivamente como a seguir:

$$\pi_t^P = \begin{cases} \langle t \rangle & \text{se } P(t) = nil \quad (\text{nó raiz}) \\ \pi_s^P \cdot \langle s, t \rangle & \text{se } P(t) = s \neq nil \quad (\text{nó predecessor de } t). \end{cases} \quad (2.2)$$

Neste trabalho, utilizamos funções de conexidade f restritas a caminhos iniciados em um dado conjunto de pixels *semente* $S \subseteq \mathcal{V}$. Essa restrição é modelada definindo uma nova função de custo de caminho $f^S(\pi)$, que é igual a $f(\pi)$ quando a origem de π pertence a S ,

e é igual a $+\infty$ caso contrário.

Alguns exemplos de funções de custo de caminho que podem ser utilizadas na IFT são:

$$\begin{aligned} f_{max}(\langle t \rangle) &= H(t) \\ f_{max}(\pi_s \cdot \langle s, t \rangle) &= \max \{f_{max}(\pi_s), \omega(s, t)\} \end{aligned} \quad (2.3)$$

$$\begin{aligned} f_{sum}(\langle t \rangle) &= H(t) \\ f_{sum}(\pi_s \cdot \langle s, t \rangle) &= f_{sum}(\pi_s) + \omega(s, t) \end{aligned} \quad (2.4)$$

$$\begin{aligned} f_{euc}(\langle t \rangle) &= \begin{cases} 0 & \text{se } t \in \mathcal{S} \\ +\infty & \text{caso contrário} \end{cases} \\ f_{euc}(\pi_s \cdot \langle s, t \rangle) &= \|t - R(s)\|^2 \end{aligned} \quad (2.5)$$

em que $H(t)$ é geralmente definido como:

$$H(t) = \begin{cases} 0 & \text{se } t \in \mathcal{S} \\ +\infty & \text{caso contrário} \end{cases}$$

Um caminho π_t é *ótimo* se $f(\pi_t) \leq f(\tau_t)$ para qualquer outro caminho $\tau_t \in \Pi(G, t)$. O *valor do caminho ótimo* $V_{opt}(t)$ é obtido tomando para cada pixel t o custo de um caminho ótimo com término em t , o qual é definido por $V_{opt}(t) = \min_{\forall \pi_t \in \Pi(G, t)} \{f(\pi_t)\}$. O algoritmo da IFT (Algoritmo 1) resolve esse problema de otimização de custo de caminho. Ele recebe um grafo de imagem G , uma função de conectividade monotonicamente incremental f e um conjunto de sementes \mathcal{S} , e computa um mapa de rótulos \mathcal{L} , uma floresta de espalhamento $P : \mathcal{V} \rightarrow \mathcal{V} \cup \{nil\}$ de caminhos ótimos π_t^P associados a cada pixel $t \in \mathcal{V}$ e um mapa de conectividade V , tal que $V(t) = f(\pi_t^P)$. Por simplicidade, o Algoritmo 1 é apresentado aqui para segmentação binária.

Algoritmo 1. – ALGORITMO DA IFT

- ENTRADA:** Grafo de imagem $G = \langle \mathcal{V}, \mathcal{A}, \omega \rangle$, conjuntos disjuntos não-vazios de sementes $S_0, S_1 \subset \mathcal{V}$, e a função de custo de caminho $f : \Pi(G) \rightarrow \mathcal{C}$.
- SAÍDA:** Floresta de caminho ótimo $P : \mathcal{V} \rightarrow \mathcal{V} \cup \{nil\}$, mapa de custo de caminho $V : \mathcal{V} \rightarrow \mathcal{C}$ e mapa de rótulos $\mathcal{L} : \mathcal{V} \rightarrow \{0, 1\}$.
- AUXILIARES:** Fila de prioridade Q , variável tmp , e uma função status $S : \mathcal{V} \rightarrow \{0, 1\}$, em que $S(t) = 1$ para nós processados e $S(t) = 0$ para nós não processados.
1. **Para cada** $t \in \mathcal{V}$, **Faça**
 2. $S(t) \leftarrow 0, P(t) \leftarrow nil$ e $V(t) \leftarrow f(\langle t \rangle)$;
 3. **Se** $t \in S_0$, **então**
 4. $\mathcal{L}(t) \leftarrow 0$ e insira t em Q ;
 5. **Se** $t \in S_1$ **então**
 6. $\mathcal{L}(t) \leftarrow 1$ e insira t em Q .
 7. **Enquanto** $Q \neq \emptyset$ **Faça**
 8. Remova s de Q tal que $V(s)$ é mínimo;
 9. $S(s) \leftarrow 1$;
 10. **Para cada** $\langle s, t \rangle \in \mathcal{A}$ tal que $S(t) = 0$ **Faça**
 11. Calcule $tmp \leftarrow f(\pi_s^P.\langle s, t \rangle)$
 12. **Se** $tmp < V(t)$ **então**
 13. $P(t) \leftarrow s, V(t) \leftarrow tmp$ e $\mathcal{L}(t) \leftarrow \mathcal{L}(s)$
 14. **Se** $t \notin Q$ **então** insira t em Q .

No caso das funções não monotonicamente incrementais, o algoritmo da IFT produz uma floresta de espalhamento P , mas os caminhos π_t^P podem não ser ótimos (i.e., podemos ter $f(\pi_t^P) = V(t) \neq V_{opt}(t)$). Entretanto, a floresta de espalhamento calculada P pode ser ótima de acordo com outros critérios de otimalidade, como provado para a função de custo da Transformada Imagem-Floresta Orientada (OIFT) (P. Miranda e Mansilla, 2014), apresentada na Seção 2.5.

Na Figura 2.6 é apresentada uma simulação do algoritmo da IFT utilizando a função f_{max} (Equação 2.3) como função de custo de caminho.

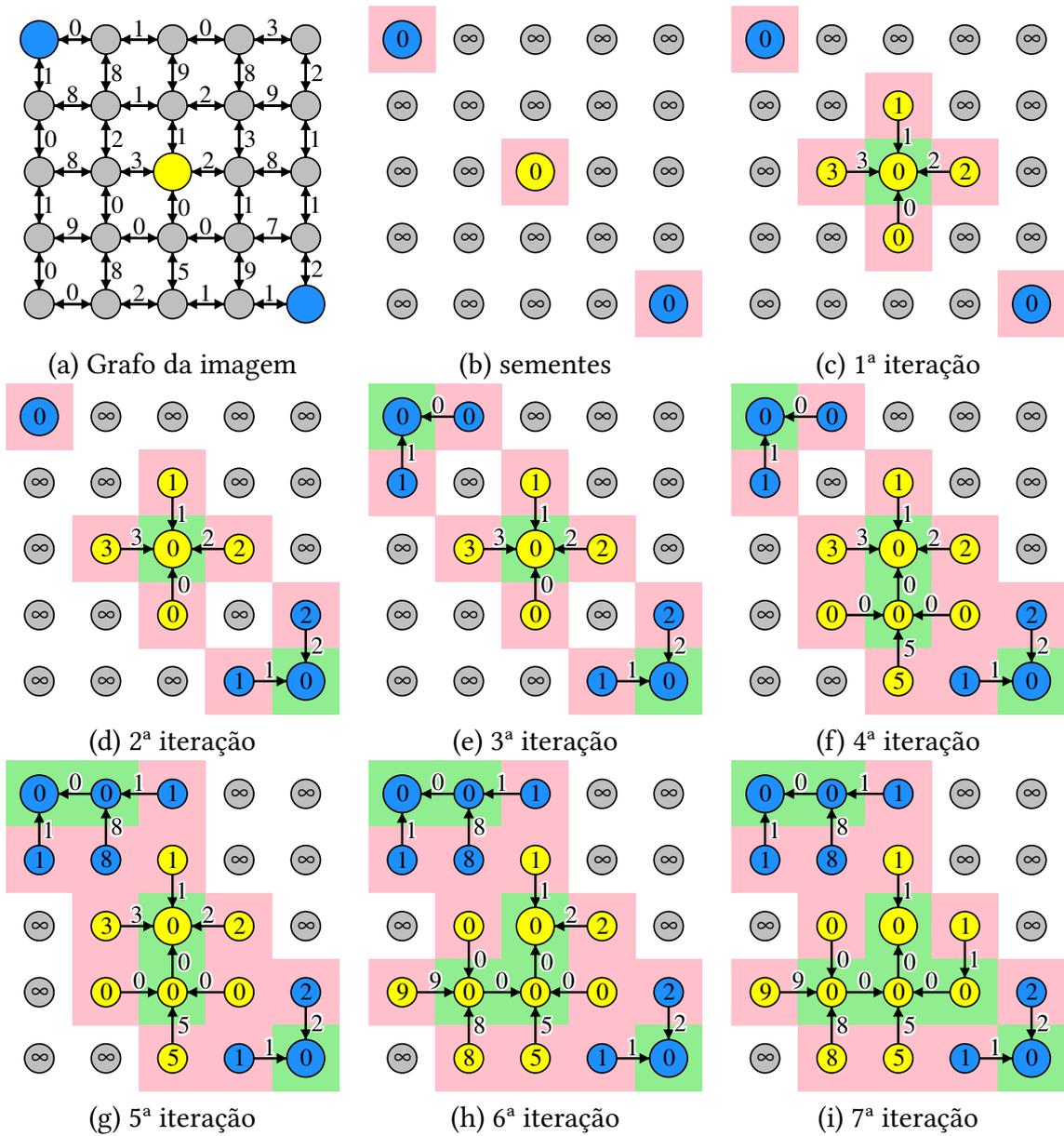


Figura 2.6: Simulação do Algoritmo 1 com função de custo da Equação 2.3. (parte 1) Adaptado de (Paulo A. V. de Miranda, 2020).

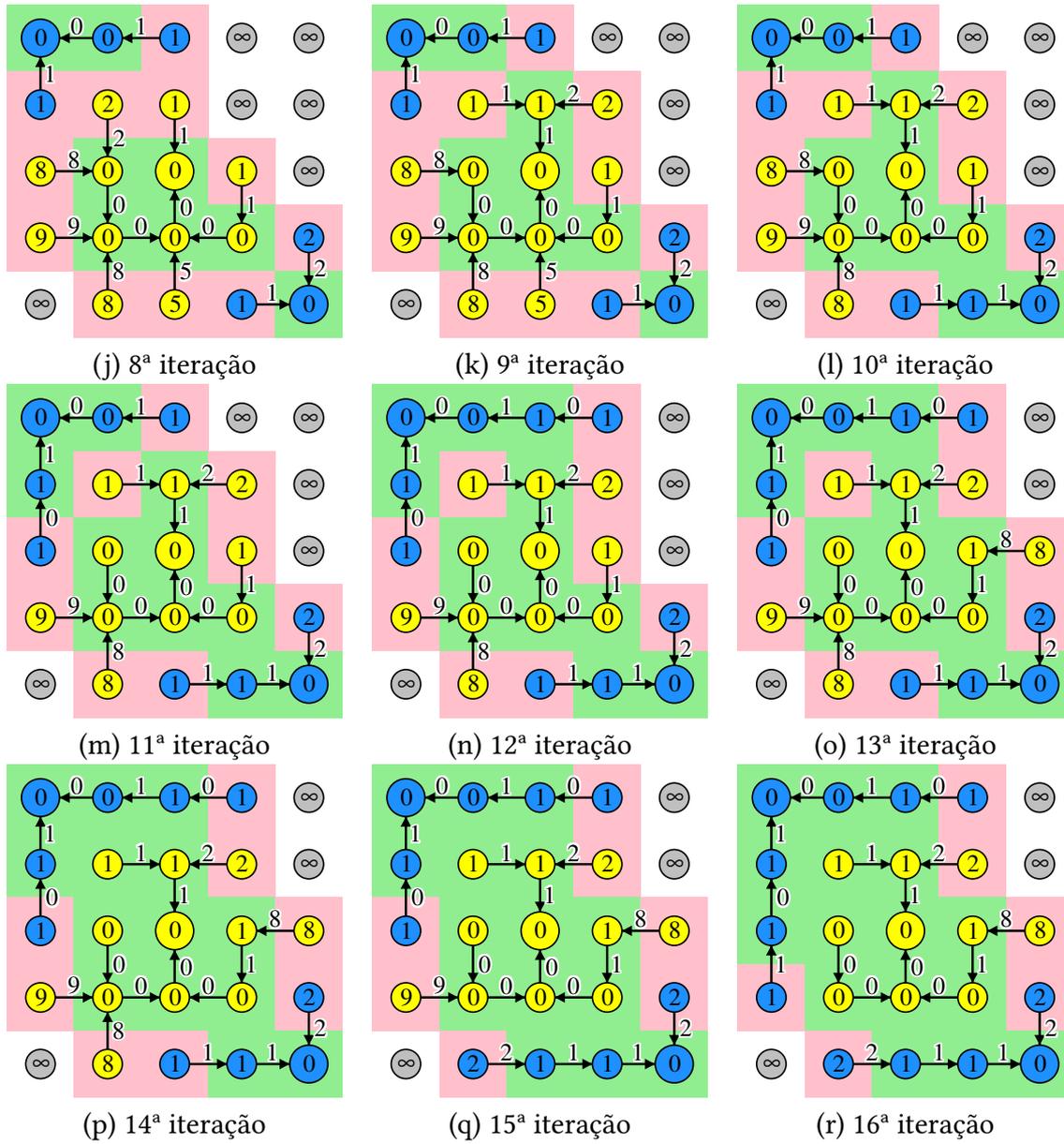


Figura 2.6: Simulação do Algoritmo 1 com função de custo da Equação 2.3. (parte 2) Adaptado de (Paulo A. V. de Miranda, 2020).

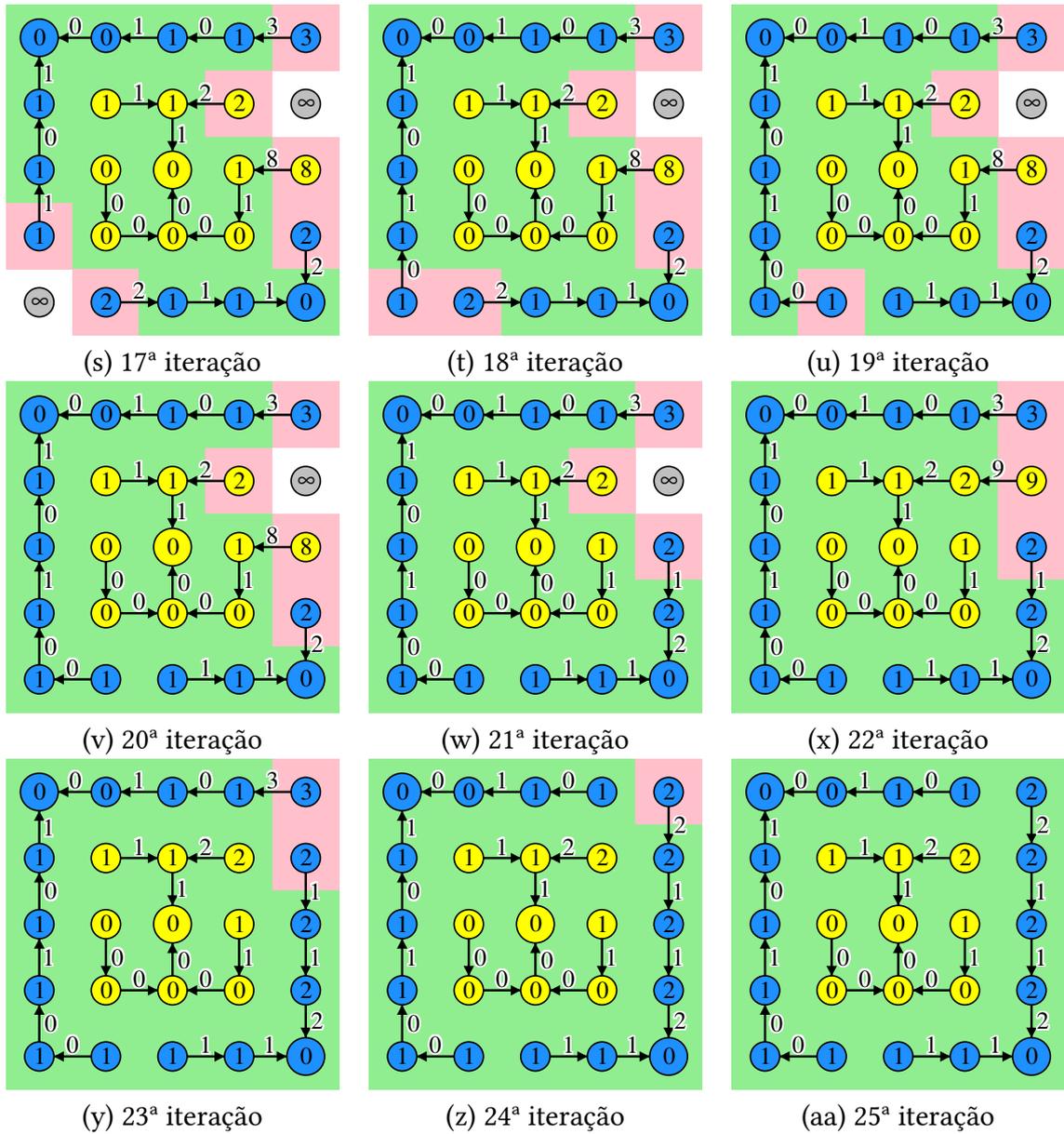


Figura 2.6: Simulação do Algoritmo 1 com função de custo da Equação 2.3. (parte 3) Adaptado de (Paulo A. V. de Miranda, 2020).

2.5 Transformada Imagem-Floresta Orientada

O método da Transformada Imagem-Floresta Orientada (*Oriented Image Foresting Transform* - OIFT) é construído sobre a estrutura da IFT, explorando a orientação/polaridade de borda dos objetos em digrafos, usando funções de conectividade não monotonicamente incrementais (P. Miranda e Mansilla, 2014). Ele ajuda a distinguir entre dois segmentos de borda similares e próximos com orientações opostas de objetos distintos, melhorando os resultados de segmentação.

2.5.1 Polaridade de borda

Polaridade de borda é uma propriedade dos objetos de uma imagem que diz respeito ao gradiente de cor do objeto para o fundo na imagem. Essa propriedade pode ser explorada com conhecimento prévio das imagens em algoritmos de segmentação para produzir melhores resultados. Um exemplo de segmentação de imagem levando em consideração a polaridade de borda é apresentado na Figura 2.7. Note que na polaridade de borda neutra (Figura 2.7a) a borda obtida pela segmentação contém tanto transições de regiões escuras para regiões claras da imagem quanto de regiões claras para regiões escuras. No exemplo, sabendo de antemão que queremos segmentar um objeto escuro de um fundo mais claro, podemos atualizar o parâmetro de polaridade de borda adequadamente e obter um resultado melhor (Figura 2.7b).

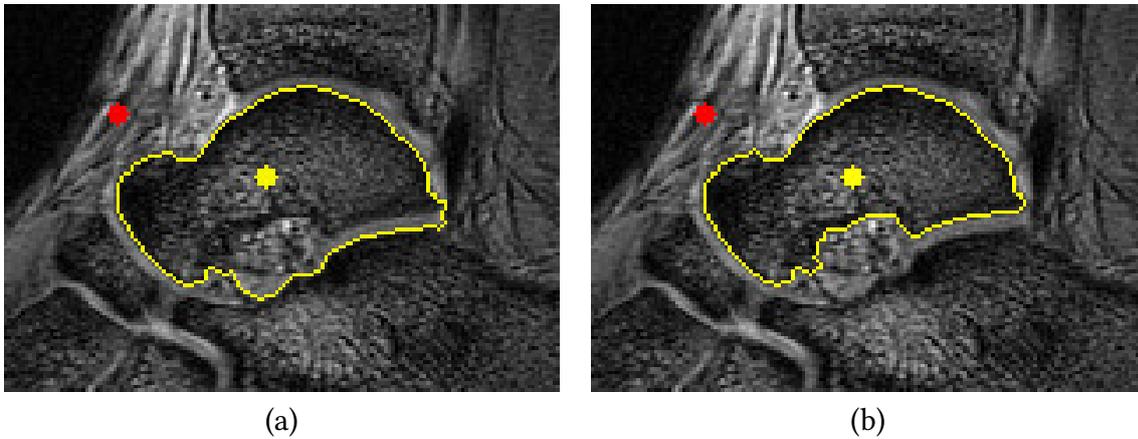


Figura 2.7: Exemplo de resultado de segmentação do osso tálus em uma imagem de ressonância magnética do pé humano utilizando a OIFT, onde é possível definir a restrição de polaridade de borda do objeto a ser segmentado. (a) Polaridade neutra e (b) Polaridade favorecendo objetos escuros em fundos claros.

Uma possível função de peso nos arcos para explorar polaridade de borda na OIFT é dada a seguir:

$$\omega(s, t) = \begin{cases} |I(s) - I(t)| \times (1 + \alpha) & \text{se } I(s) > I(t) \\ |I(s) - I(t)| \times (1 - \alpha) & \text{se } I(s) < I(t) \\ 0 & \text{caso contrário} \end{cases} \quad (2.6)$$

$I(t)$ é o valor de intensidade de t . O parâmetro $\alpha \in [-1, 1]$ é um fator de orientação,

em que $\alpha < 0$ favorece a segmentação de objetos escuros em um fundo mais claro e $\alpha > 0$ favorece a orientação de objetos claros em um fundo mais escuro

2.5.2 Descrição do Algoritmo

Do ponto de vista de cortes em grafos, o algoritmo da OIFT corresponde a um otimizador do tipo Max-Min, isto é, sua segmentação resultante gera uma solução globalmente ótima sujeita às restrições de sementes, ao maximizar a medida de corte em grafo ε_{min} definida por:

$$\varepsilon_{min}(\mathcal{L}) = \min\{\omega(s, t) : \langle s, t \rangle \in \mathcal{A} \ \& \ \mathcal{L}(s) > \mathcal{L}(t) = 0\} \quad (2.7)$$

Por simplicidade, o algoritmo da OIFT é apresentado aqui sem o cálculo explícito da floresta de caminhos, do qual vem o seu nome. Podemos fazer isso, já que o mapa de rótulos \mathcal{L} sem a floresta é suficiente para a tarefa de segmentação. A segmentação da OIFT, indicada por \mathcal{L} , pode ser computada pelo Algoritmo 2, em um digrafo conexo e simétrico G , como descrito em [Mansilla e P. Miranda, 2013b](#). Alternativamente, sua segmentação também pode ser obtida pelo Algoritmo 1, com cálculo explícito da floresta, utilizando a seguinte função não monotonicamente incremental f^* :

$$f^*(\langle t \rangle) = \begin{cases} -\infty & \text{se } t \in S_0 \cup S_1 \\ +\infty & \text{caso contrário} \end{cases}$$

$$f^*(\pi_{r \rightsquigarrow s} \cdot \langle s, t \rangle) = \begin{cases} \omega(s, t) & \text{se } r \in S_1 \\ \omega(t, s) & \text{caso contrário} \end{cases} \quad (2.8)$$

Algoritmo 2. – ALGORITMO DA OIFT

ENTRADA: Grafo de imagem $\langle \mathcal{V}, \mathcal{A}, \omega \rangle$, conjuntos disjuntos não-vazios de sementes S_0 e S_1 .
SAÍDA: O mapa de rótulos $\mathcal{L} : \mathcal{V} \rightarrow \{0, 1\}$.
AUXILIARES: Fila de prioridade Q , variável tmp , a função de custo $V : \mathcal{V} \rightarrow [-\infty, \infty]$, e uma função status $S : \mathcal{V} \rightarrow \{0, 1\}$, em que $S(t) = 1$ para nós processados e $S(t) = 0$ para nós não processados.

1. **Para cada** $t \in \mathcal{V}$, **Faça**
2. $S(t) \leftarrow 0$ e $V(t) \leftarrow \infty$;
3. **Se** $t \in S_0$, **então**
4. $V(t) \leftarrow -\infty$, $\mathcal{L}(t) \leftarrow 0$ e **insira** t em Q ;
5. **Se** $t \in S_1$ **então**
6. $V(t) \leftarrow -\infty$, $\mathcal{L}(t) \leftarrow 1$ e **insira** t em Q .
7. **Enquanto** $Q \neq \emptyset$ **Faça**
8. **Remova** s de Q tal que $V(s)$ é mínimo;
9. $S(s) \leftarrow 1$;
10. **Para cada** $\langle s, t \rangle \in \mathcal{A}$ tal que $S(t) = 0$ **Faça**
11. **Se** $\mathcal{L}(s) = 1$ **então** $tmp \leftarrow \omega(s, t)$
12. **Senão** $tmp \leftarrow \omega(t, s)$;
13. **Se** $tmp < V(t)$ **então**
14. $V(t) \leftarrow tmp$ e $\mathcal{L}(t) \leftarrow \mathcal{L}(s)$
15. **Se** $t \notin Q$ **então** **insira** t em Q .

Note que na linha 12 do Algoritmo 2, o peso $\omega(t, s)$ do arco invertido $\langle t, s \rangle$ é usado (ao invés daquele do $\langle s, t \rangle \in \mathcal{A}$ escolhido). É por isso que um digrafo simétrico é necessário.

2.5.3 Controlando o parâmetro α

Na imagem sintética apresentada na Figura 2.8 temos um exemplo de segmentação utilizando o parâmetro α para controlar a polaridade de borda. Temos duas bordas candidatas para a segmentação: A borda da região de intensidade 8 para a região de intensidade 4 e a borda da região de intensidade 4 para a região de intensidade 10 (Figura 2.8a). Mantendo $\alpha = 0$ (Figura 2.8b), os pesos se mantêm simétricos ($\omega_{st} = \omega_{ts}$) e a segmentação se dá como em um grafo não direcionado. Variando α para um valor menor que 0, aumentamos o peso de arcos de regiões mais escuras para regiões mais claras. Isso favorece que o corte da segmentação seja feito na borda da região de brilho 4 para a região de brilho 10, ou seja, escuro-claro (Figura 2.8c), pela equação 2.7. Da mesma forma, se variamos α para um valor maior que zero, favorecemos um corte da região de brilho 8 para a região de brilho 4, ou seja, claro-escuro (Figura 2.8d).

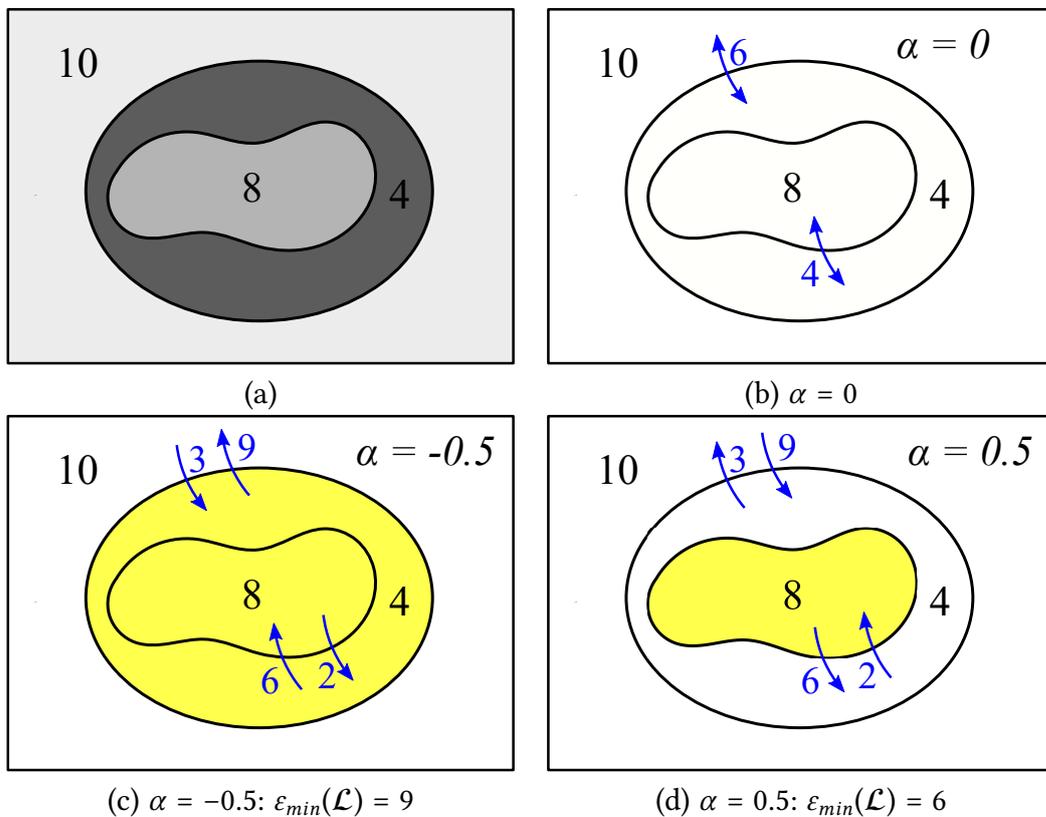


Figura 2.8: Imagem sintética com regiões de brilhos diferentes. Variando α , alteramos o valor de $\varepsilon_{\min}(\mathcal{L})$ e com isso, a borda da segmentação.

2.5.4 Simulação do algoritmo da OIFT

Na Figura 2.9 é apresentada uma simulação do Algoritmo da OIFT, utilizando a IFT com a função de custo da Equação 2.8.

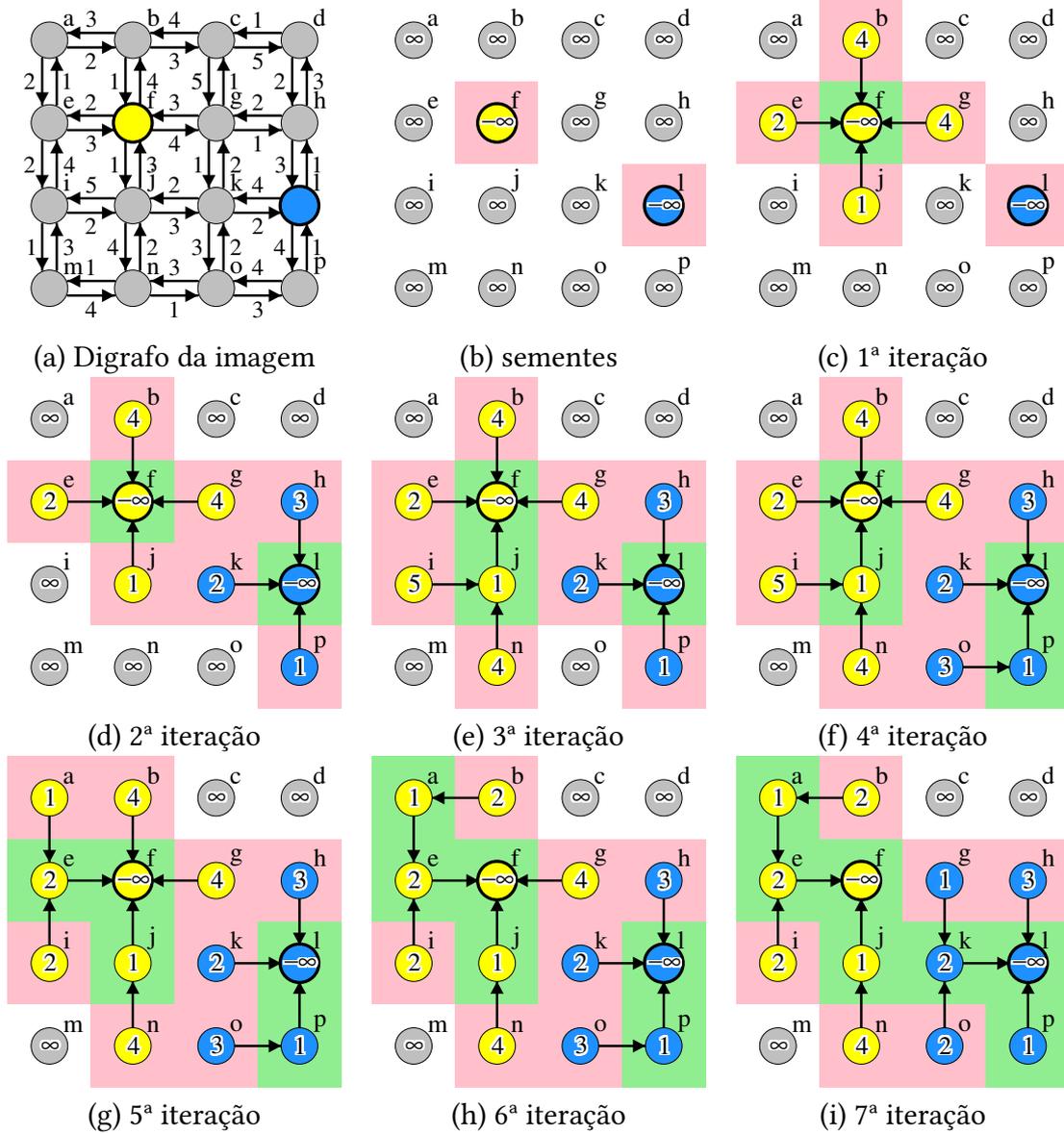


Figura 2.9: Simulação do algoritmo da OIFT com função de custo da Equação 2.8. (parte 1)

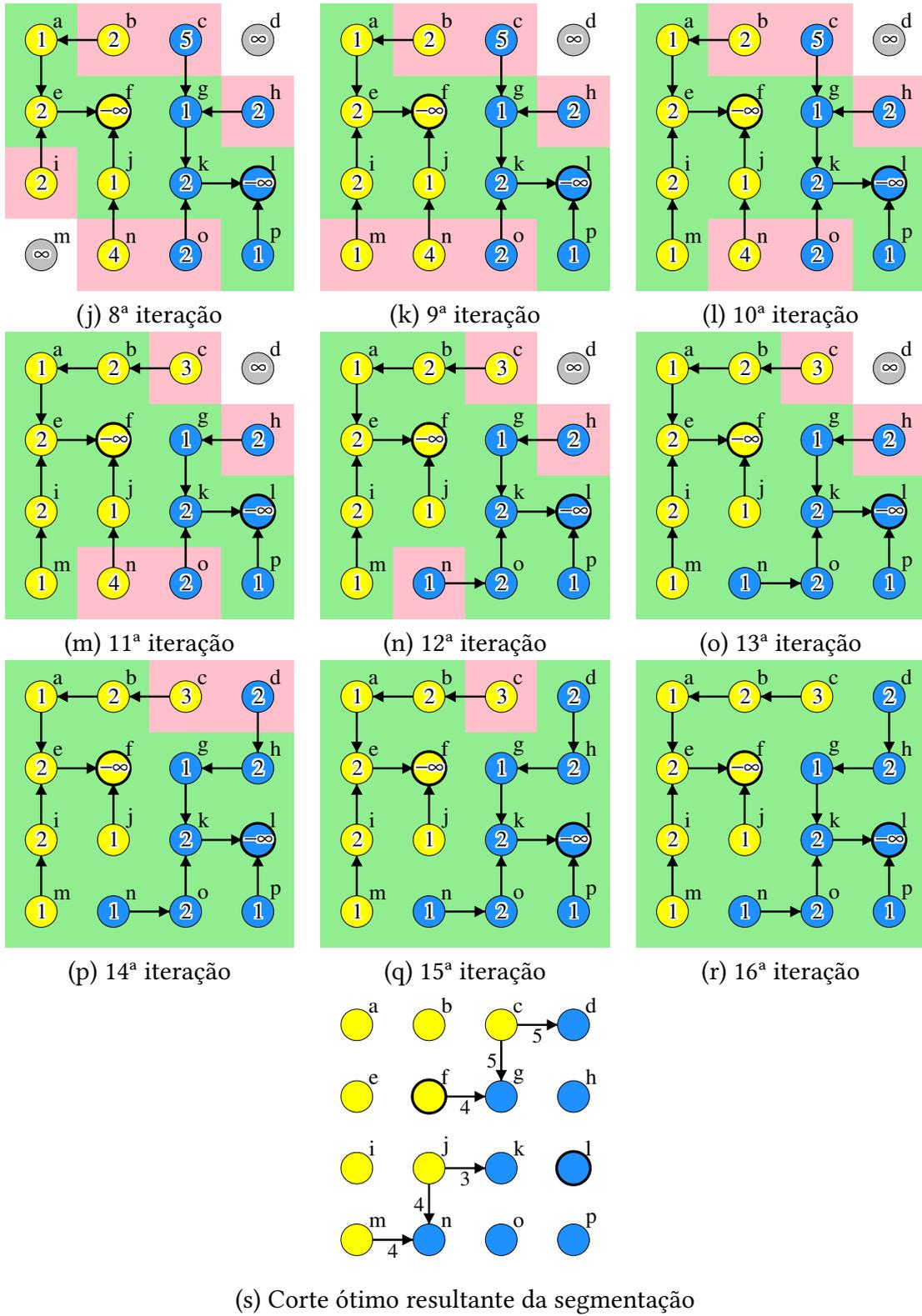


Figura 2.9: Simulação do algoritmo da OIFT com função de custo da Equação 2.8. (parte 2)

2.6 Passeios Aleatórios

2.6.1 Motivação

O algoritmo de passeios aleatórios (*Random Walks* - RW) é motivado pelo seguinte problema:

Dado um caminhante aleatório que, colocado em um ponto entre 0 e N , dá um passo em cada momento em direção a um desses pontos: Ao ponto N com probabilidade $p > 0$, e ao ponto 0 com probabilidade $q = 1 - p > 0$. Qual a probabilidade que ele chegue ao ponto N antes de atingir o ponto 0?

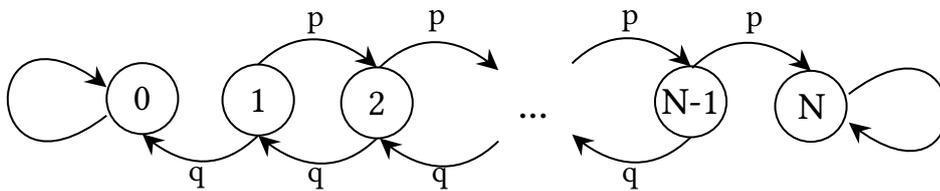


Figura 2.10: Exemplo de passeios aleatórios em um grafo unidimensional.

Seja $p(x)$ a probabilidade do caminhante inicialmente em um ponto x chegar ao nó N antes de chegar ao nó 0, temos:

1. $p(0) = 0$
2. $p(N) = 1$
3. $p(x) = q \cdot p(x - 1) + p \cdot p(x + 1), x \in \{1, 2, \dots, N - 1\}$

Observe que para nós em pontos intermediários com $x \in \{1, 2, \dots, N - 1\}$, a probabilidade $p(x)$ é dada por uma média ponderada das probabilidades de seus vizinhos. Esta relação pode ser estendida para grafos mais genéricos.

Como discutido em [Grady, 2006](#), o problema de calcular os valores de $p(x)$ pode ser resolvido por sucessivas simulações de passeios aleatórios, mas essa é uma solução ineficiente, principalmente considerando que grafos de imagens podem conter milhares de nós. Soluções mais eficientes serão discutidas mais adiante neste documento, mas primeiramente vamos discutir a analogia existente entre passeios aleatórios e circuitos elétricos.

2.6.2 Analogia com circuitos

Um problema análogo ao problema de passeios aleatórios é descrito da seguinte forma:

Dado um circuito com N resistores em série e aplicada uma diferença de potencial de $1V$, qual a voltagem em cada ponto x entre os resistores?

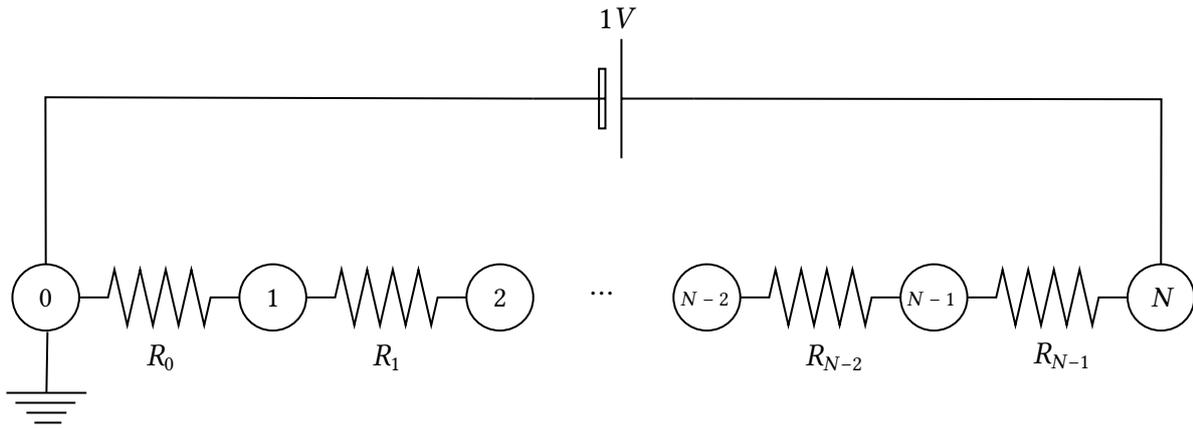


Figura 2.11: Circuito com N resistores em série e diferença de potencial $1V$ entre os nós 0 e N .

Pelas leis de Ohm e Kirchhoff sabemos que a corrente fluindo de um ponto x a y é:

$$i_{xy} = \frac{(v(x) - v(y))}{R_{xy}} \quad (2.9)$$

Em que $v(x)$ representa o potencial em um ponto x e R_{xy} representa a resistência entre os pontos x e y .

E para 3 pontos consecutivos x , y e z , temos:

$$i_{xy} = i_{yz} \quad (2.10)$$

Vamos utilizar a condutância $C_x = \frac{1}{R_x}$ para simplificar a notação a partir de agora. Logo, temos:

$$i_{xy} = (v(x) - v(y)) \cdot C_{xy} \quad (2.11)$$

e:

$$\begin{aligned} (v(x-1) - v(x)) \cdot C_{x-1} &= (v(x) - v(x+1)) \cdot C_x \\ v(x-1) \cdot C_{x-1} - v(x) \cdot C_{x-1} &= v(x) \cdot C_x - v(x+1) \cdot C_x \\ v(x) \cdot (C_x + C_{x-1}) &= v(x-1) \cdot C_{x-1} + v(x+1) \cdot C_x \\ v(x) &= v(x-1) \cdot \frac{C_{x-1}}{C_x + C_{x-1}} + v(x+1) \cdot \frac{C_x}{C_x + C_{x-1}} \end{aligned} \quad (2.12)$$

A analogia se dá assumindo os seguintes valores de p e q para o problema do caminhante aleatório:

$$\begin{aligned}
p &= \frac{C_{x-1}}{C_x + C_{x-1}} \rightarrow p \cdot C_x + p \cdot C_{x-1} = C_{x-1} \\
p \cdot C_x &= C_{x-1} - p \cdot C_{x-1} \rightarrow p \cdot C_x = C_{x-1} \cdot (1 - p) \\
p \cdot C_x &= C_{x-1} \cdot q \\
\frac{p}{q} &= \frac{C_x}{C_{x-1}}
\end{aligned} \tag{2.13}$$

Assim, escolhemos valores de condutância que satisfazem a Equação 2.13 a partir de um C_0 arbitrário, assumindo que $p > 0$ e $q > 0$.

Note que nas duas versões do problema, $p(x)$ e $v(x)$ são calculados como médias ponderadas dos valores de p e v nos nós vizinhos de x .

2.6.3 Minimização de energia

Como proposto por Leo Grady (Grady, 2006), o problema de passeios aleatórios também pode ser descrito como um problema de minimização de energia, assumindo que o grafo é simétrico, conexo e não orientado (i.e., $\omega_{st} = \omega_{ts}$), minimizando a seguinte função:

$$E_1(\mathcal{L}) = \sum_{\langle s,t \rangle \in \mathcal{A}} \omega_{st} \cdot (\mathcal{L}(s) - \mathcal{L}(t))^2 \tag{2.14}$$

em que \mathcal{L} é uma segmentação *fuzzy* que satisfaz as restrições de sementes, tal que $\mathcal{L}(s) = v(s)$ e os valores de ω_{st} correspondem a valores de condutância.

Essa energia é convexa, logo ela possui um minimizador único \mathcal{L}_{opt} , que pode ser obtido diferenciando E_1 e encontrando seu ponto crítico, de modo que para qualquer nó não semente $s \notin S_1 \cup S_0$ temos:

$$\mathcal{L}_{opt}(s) = \frac{\sum_{t \in \mathcal{N}(s)} \omega_{st} \cdot \mathcal{L}_{opt}(t)}{\sum_{t \in \mathcal{N}(s)} \omega_{st}} \tag{2.15}$$

A segmentação rígida final \mathcal{L}_{RW} do método do RW é então obtida atribuindo $\mathcal{L}_{RW}(s) = 1$ para todos os pixels s com $\mathcal{L}_{opt}(s) \geq 0.5$ e $\mathcal{L}_{RW}(s) = 0$ caso contrário.

Como discutido em Grady, 2006, o problema de computar $\mathcal{L}_{opt}(s)$ é equivalente ao problema de determinar a probabilidade de um caminhante aleatório começando em s alcançar primeiro uma das sementes de objeto antes das sementes do fundo. No entanto, a solução via simulação de passeios aleatórios é muito custosa. Uma abordagem mais factível é resolver o sistema de equações lineares esparso resultante da Equação 2.15 para todos os nós sem sementes. Uma terceira solução pelo método da relaxação é explicada na Seção 2.7.

2.7 Método das Relaxações

Como proposto por [Malmberg et al., 2010](#), para uma segmentação inicial calculada \mathcal{L}^0 , uma sequência de segmentações fuzzy $\mathcal{L}^1, \mathcal{L}^2, \dots, \mathcal{L}^N$ por relaxações iterativas é obtida como a seguir:

$$\mathcal{L}^{i+1}(s) = \begin{cases} \frac{\sum_{t \in \mathcal{N}(s)} \omega_{st} \cdot \mathcal{L}^i(t)}{\sum_{t \in \mathcal{N}(s)} \omega_{st}} & \text{se } s \notin \mathcal{S}_1 \cup \mathcal{S}_0 \\ \mathcal{L}^i(s) & \text{caso contrário} \end{cases} \quad (2.16)$$

O \mathcal{L} rígido (*crisp*) final é então obtido, atribuindo $\mathcal{L}(s) = 1$ a todo $s \in \mathcal{V}$ com $\mathcal{L}^N(s) \geq 0.5$ e $\mathcal{L}(s) = 0$ caso contrário. Em [Malmberg et al., 2010](#), \mathcal{L}^0 é tomado como sendo o resultado de segmentação via IFT com o custo de caminho aditivo ([Falcão, Stolfi e Lotufo, 2004](#)) e a segmentação final recebe o nome de Transformada Imagem-Floresta Relaxada (*Relaxed Image Foresting Transform*).

Essa solução converge para o Random Walks para valores de N suficientemente grandes, mas a ideia aqui é considerar uma quantidade menor de relaxações a fim de conseguir um resultado híbrido, conservando os aspectos positivos da segmentação inicial. O parâmetro N é responsável por controlar a suavidade da borda da segmentação, melhorando assim os resultados quando há presença de ruídos e bordas fracas.

2.8 Deep Extreme Cut

Deep Extreme Cut ([Maninis et al., 2018](#)) é um método de segmentação de imagens baseado em *Deep learning* que obtém a segmentação de um objeto a partir de seus quatro pontos extremos (mais acima, mais abaixo, mais a esquerda e mais a direita). A entrada do usuário com base em pontos extremos é geralmente uma boa escolha para algoritmos de disparo único (*single-shot algorithms*), uma vez que a partir dos 4 pontos indicados podemos inferir a caixa delimitadora (*bounding box*) do objeto e, conseqüentemente, todos os pontos fora da caixa são conhecidos por pertencerem ao fundo.

O algoritmo calcula um mapa de probabilidade que diz para cada pixel, seu grau de pertinência ao objeto que desejamos segmentar (sendo esse grau um valor de 0 a 1). As Figuras 1.1a e d ilustram os quatro pontos extremos em objetos de interesse e os mapas de probabilidade resultantes são mostrados nas Figuras 1.1b e e, respectivamente. A partir desses mapas de probabilidade as segmentações são obtidas por limiarização (Figuras 1.1c e f).

O algoritmo se baseia em adicionar um canal extra à imagem (que contém os canais RGB), contendo uma Gaussiana bidimensional centrada em cada um dos quatro pontos extremos. Essa imagem é recortada pela *bounding box* e utilizada como entrada em uma rede neural convolucional (CNN), que aprende a transformar essas informações em um mapa de probabilidade para o objeto que satisfaz as restrições de pontos extremos dada. A

CNN foi treinada para minimizar a perda de entropia cruzada padrão, usando a *Resnet-101* (He *et al.*, 2016) como a espinha dorsal de sua arquitetura, removendo as camadas totalmente conectadas, bem como as camadas de pooling máximo nos dois últimos estágios para preservar a resolução de saída aceitável para previsão densa.

Segundo os autores, uma vantagem desse método é que a utilização de pontos extremos em vez de *bounding boxes* explícitas como entrada do usuário pode diminuir consideravelmente o tempo de interação do usuário e a *bounding box* gerada a partir dos pontos extremos é equivalente em qualidade às que são utilizadas em métodos tradicionais (Maninis *et al.*, 2018).

Capítulo 3

Métodos híbridos via energias quadráticas em grafos direcionados

Este capítulo descreve os métodos híbridos propostos utilizando a extensão do *Random Walks* para grafos direcionados, que é apresentada na Seção 3.1. A OIFT Relaxada é apresentada na Seção 3.2 e o *Deep Extreme Cut* Relaxado é apresentado na Seção 3.3.

3.1 Método das relaxações com pesos assimétricos

A fim de explorar a polaridade de borda, consideramos a seguinte atribuição de pesos de arcos:

$$\omega_{st} = \begin{cases} [K - \kappa_{st} \times (1 + \alpha)]^\beta & \text{se } I(s) > I(t) \\ [K - \kappa_{st} \times (1 - \alpha)]^\beta & \text{se } I(s) < I(t) \\ [K - \kappa_{st}]^\beta & \text{caso contrário} \end{cases} \quad (3.1)$$

em que os pesos ω_{st} são a potência do complemento de uma combinação de uma função de dissimilaridade não direcionada κ_{st} entre nós vizinhos s e t , multiplicada por um fator de orientação para $\alpha \in [-1, 1]$, tal que $\alpha < 0$ favorece a segmentação de objetos escuros em um fundo mais claro e $\alpha > 0$ favorece a orientação oposta, $I(t)$ é o valor de intensidade de t , K é um peso máximo usado para o cálculo do complemento e o expoente $\beta > 1$ realça as diferenças dos pesos.¹

Embora Malmberg et al. afirmem que seu procedimento de relaxação não requer que $\omega_{st} = \omega_{ts}$ (Malmberg et al., 2010), na Seção 4 mostramos que sua direta aplicação nos

¹Transformações crescentes, tal como a potenciação por β não afetam o resultado de métodos tal como a OIFT, porém são importantes no método das relaxações por afetarem o comportamento do Random Walks. Para valores baixos de potência, o Random Walks se torna mais sensível ao posicionamento dos marcadores. Por outro lado, valores de potência muito elevados podem gerar bordas irregulares na segmentação. Portanto, um valor intermediário de β deve ser configurado empiricamente de modo a ter um melhor balanço entre robustez de sementes e suavização de borda do objeto gerado.

pesos direcionados pela Eq. 3.1 gera resultados ruins. A fim de devidamente estender RW para grafos direcionados, Singaraju et al. (Singaraju, Grady e Vidal, 2008) propuseram a substituição da sua função de energia da Eq. 2.14 por:

$$E_3(\mathcal{L}) = \sum_{\langle s,t \rangle \in \mathcal{A}} [\omega_{st} \mathbf{I}(\mathcal{L}_s \geq \mathcal{L}_t) + \omega_{ts} \mathbf{I}(\mathcal{L}_s < \mathcal{L}_t)] \cdot (\mathcal{L}_s - \mathcal{L}_t)^2 \quad (3.2)$$

em que \mathcal{L} é uma segmentação *fuzzy* satisfazendo as sementes, \mathcal{L}_s é uma abreviação de $\mathcal{L}(s)$, e $\mathbf{I}(A)$ uma função indicadora do evento A , tal que $\mathbf{I}(A) = 1$, se A é verdadeiro e $\mathbf{I}(A) = 0$ caso contrário.

Singaraju et al. também descreveram uma rede elétrica equivalente com diodos e resistores (Figura 3.1) que resolve o mesmo problema de segmentação que o seu *framework* pela Eq. 3.2 (Singaraju, Grady e Vidal, 2008). Entretanto, eles observaram que a interpretação de passeios aleatórios de RW deixa de ser válida quando pesos assimétricos são utilizados no digrafo. Portanto, já que a relaxação com pesos assimétricos de Malmberg et al. (Malmberg et al., 2010) corresponde a uma interpretação de passeios aleatórios de RW no digrafo, temos que ela não resolve a minimização da energia E_3 e, conseqüentemente, não é uma abordagem apropriada para digrafos.

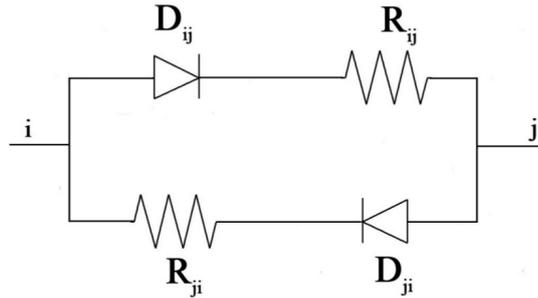


Figura 3.1: Configuração do diodo resistor entre um par de nós do grafo.

Aqui é proposto um procedimento corrigido de relaxação para ser aplicado com pesos assimétricos. Na Equação 3.2, para cada par de nós s e t com $\mathcal{L}_s \neq \mathcal{L}_t$, apenas um dos pesos ω_{st} e ω_{ts} será adicionado no somatório duas vezes. A seleção de ω_{st} ocorre para $\mathcal{L}_s \geq \mathcal{L}_t$ e ω_{ts} caso contrário. Essa direcionalidade corresponde a diodos em cada arco $\langle s, t \rangle$, que são ativados somente quando a queda de potencial $\mathcal{L}_s - \mathcal{L}_t$ através deles é positiva.

Singaraju et al. propuseram um algoritmo iterativo para estimar numericamente o minimizador único da Equação 3.2, em um método descendente de Newton (*Newton descent like framework*), onde o Hessiano é aproximado pelo Laplaciano (Singaraju, Grady e Vidal, 2008). Ele é baseado em tratar a rede elétrica como tendo puramente impedâncias resistivas, definindo os diodos ativos se baseando nos valores de potenciais \mathcal{L} da iteração anterior. A cada iteração, uma otimização por RW é calculada para pesos simétricos. Entretanto, essa solução não é adequada para o nosso propósito, dado que apenas 4-5 iterações foram necessárias nos experimentos realizados em Singaraju, Grady e Vidal, 2008 antes da convergência, como relatado pelos autores, mas para obter um método híbrido, conservando informações importantes da segmentação inicial, precisamos de um algoritmo com convergência muito mais lenta, o qual será interrompido após algumas

iterações.

Em nossa solução, realizamos um passo de relaxação, após definir os diodos ativos, substituindo ω_{st} na Equação 2.16 por $W^i(s, t)$:

$$\mathcal{L}^{i+1}(s) = \begin{cases} \frac{\sum_{t \in \mathcal{N}(s)} W^i(s, t) \cdot \mathcal{L}^i(t)}{\sum_{t \in \mathcal{N}(s)} W^i(s, t)} & \text{se } s \notin \mathcal{S}_1 \cup \mathcal{S}_0 \\ \mathcal{L}^i(s) & \text{caso contrário} \end{cases} \quad (3.3)$$

sendo $W^i(s, t)$ definido como:

$$W^i(s, t) = \begin{cases} [K - \kappa_{st}]^\beta & \text{se } \mathcal{L}^i(s) = \mathcal{L}^i(t) \\ \omega_{st} & \text{se } \mathcal{L}^i(s) > \mathcal{L}^i(t) \\ \omega_{ts} & \text{se } \mathcal{L}^i(s) < \mathcal{L}^i(t) \end{cases} \quad (3.4)$$

Note que quando $\mathcal{L}_s = \mathcal{L}_t$, ω_{st} é escolhido por $\mathcal{I}(\mathcal{L}_s \geq \mathcal{L}_t)$ na Eq. 3.2, mas a energia não é afetada já que $(\mathcal{L}_s - \mathcal{L}_t)^2 = 0$. Contudo, dado que estamos realizando uma quantidade limitada de relaxações, essa seleção arbitrária em favor de ω_{st} quando $\mathcal{L}^i(s) = \mathcal{L}^i(t)$ pode afetar levemente os resultados, portanto, preferimos adotar $[K - \kappa_{st}]^\beta$ como uma solução mais imparcial, que elimina a direcionalidade dada por α nesse caso².

Uma segunda versão, inspirada pela propagação de custo da OIFT, em que um nó do fundo t (i.e., $\mathcal{L}^i(t) < 0.5$) processa ω_{st} de arcos anti-paralelos para seus nós vizinhos s , também foi considerada para análise em nossos experimentos, definindo $W^i(s, t)$ alternativamente como:

$$W^i(s, t) = \begin{cases} \omega_{st} & \text{se } \mathcal{L}^i(t) < 0.5 \\ \omega_{ts} & \text{se } \mathcal{L}^i(t) \geq 0.5 \end{cases} \quad (3.5)$$

Qualquer nó s tal que $\mathcal{L}^i(s) = \mathcal{L}^{i-1}(s)$ é estacionário na iteração i . Ao atualizar \mathcal{L} apenas para nós não-estacionários, mantemos o algoritmo correto e obtemos resultados mais rápidos. Claramente, todos os nós s para os quais $\mathcal{L}^{i-1}(s) = \mathcal{L}^{i-1}(t)$ para todo t em $\mathcal{N}(s)$ são estacionários na i -ésima iteração. Todos os nós em $\mathcal{S} = \mathcal{S}_0 \cup \mathcal{S}_1$ também são estacionários.

Já que \mathcal{L}^0 é rígida, temos $\mathcal{L}^0 = \chi_P$ e os únicos nós não estacionários na primeira iteração devem estar contidos em sua fronteira $F(\chi_P) = \{s \in \mathcal{V} : \exists t \in \mathcal{N}(s) \text{ and } \chi_P(s) \neq \chi_P(t)\}$. Denotamos por $\Omega^1 = F(\mathcal{L}^0) \setminus \mathcal{S}$ o superconjunto inicial de nós não-estacionários na primeira iteração.

Para as próximas iterações, Ω^{i+1} deve ser expandido em relação a Ω^i (Teorema 1 de [Malmberg et al., 2010](#)). Portanto temos $\Omega^{i+1} = (\Omega^i \oplus \mathcal{A}) \setminus \mathcal{S}$, em que o operador de expansão \oplus é definido como $X \oplus \mathcal{A} = X \cup \{t \in \mathcal{V} : \langle s, t \rangle \in \mathcal{A} \text{ para algum } s \in X\}$.

Para uma implementação mais rápida, podemos também considerar $\Omega^{i+1} = (\Omega^i \oplus \mathcal{A}) \setminus \mathcal{S} =$

²Note que a matriz Laplaciana utilizada em [Singaraju, Grady e Vidal, 2008](#) pode não ser simétrica para arcos $\langle s, t \rangle$ e $\langle t, s \rangle$ com $\mathcal{L}(s) = \mathcal{L}(t)$.

$[\Omega^i \cup ((\Omega^i \setminus \Omega^{i-1}) \oplus \mathcal{A})] \setminus \mathcal{S}$, para que o operador de expansão seja aplicado a um conjunto menor de nós³, como proposto abaixo:

Algoritmo 3. – *MÉTODO DAS RELAXAÇÕES*

ENTRADA: Grafo de imagem $\langle \mathcal{V}, \mathcal{A}, \omega \rangle$, conjunto de sementes \mathcal{S} , Segmentação inicial \mathcal{L}^0
AUXILIARES: Vetor de conjuntos de nós não estacionários $\Omega^i, i \in \{0, 1, \dots, N\}$
SAÍDA: Mapa de rótulos $\mathcal{L} : \mathcal{V} \rightarrow \{0, 1\}$.

1. Defina $\Omega^0 \leftarrow \emptyset$ e $\Omega^1 \leftarrow F(\mathcal{L}^0) \setminus \mathcal{S}$.
2. **Para cada** $i = 1, \dots, N$, **Faça**
3. Compute $\mathcal{L}^i(p)$ para todo $p \in \Omega^i$ pela Eq. 3.3 e defina $\mathcal{L}^i(p) = \mathcal{L}^{i-1}(p)$ para todo $p \notin \Omega^i$.
4. $\Omega^{i+1} \leftarrow [\Omega^i \cup ((\Omega^i \setminus \Omega^{i-1}) \oplus \mathcal{A})] \setminus \mathcal{S}$.
5. Defina $\mathcal{L}(p) = 1$ se $\mathcal{L}^N(p) \geq 0.5$ e $\mathcal{L}(p) = 0$ caso contrário, para todo $p \in \mathcal{V}$.
6. **Retorne** \mathcal{L}

3.2 OIFT Relaxada

OIFT (P. Miranda e Mansilla, 2014; Mansilla e P. Miranda, 2013b) é um método que está na intersecção dos *frameworks* da IFT (Falcão, Stolfi e Lotufo, 2004; Ciesielski, Falcão e P. Miranda, 2018) e do Corte em Grafo Generalizado (*Generalized Graph-Cut - GGC*) (Ciesielski, J.K. Udupa *et al.*, 2012; C. Couprie *et al.*, 2010). Como visto na Seção 2.5, do ponto de vista de corte em grafos, a OIFT é um otimizador do tipo Max-Min. Idealmente, os pesos $\omega(s, t)$ devem ser elevados para arcos cruzando a borda do objeto de interesse do seu interior para o seu exterior e baixos caso contrário. Por outro lado, o método dos passeios aleatórios necessita de pesos complementares, dado que os pesos no RW estão associados à condutância do circuito elétrico.

A fim de facilitar a integração entre OIFT e RW no método híbrido proposto, com um esquema de pesos único, ao contrário dos trabalhos (P. Miranda e Mansilla, 2014; Mansilla e P. Miranda, 2013b), aqui consideramos a versão dual da OIFT como um otimizador do tipo Min-Max, via a minimização da energia E_2 (Equação 3.6) para um \mathcal{L} rígido. Em outras palavras, seja $\mathcal{U} = \{\chi_P : \mathcal{S}_1 \subseteq P \subseteq \mathcal{V} \setminus \mathcal{S}_0\}$ o universo de todas as possíveis segmentações separando \mathcal{S}_1 de \mathcal{S}_0 . A OIFT dual calcula uma partição ótima $\mathcal{L}'_{opt} \in \mathcal{U}$ minimizando E_2 em um digrafo simétrico, ou seja, $E_2(\mathcal{L}'_{opt}) = \min_{\mathcal{L} \in \mathcal{U}} E_2(\mathcal{L})$.

$$E_2(\mathcal{L}) = \max_{\langle s, t \rangle \in \mathcal{A} : \mathcal{L}(s)=1} \omega_{st} \cdot |\mathcal{L}(s) - \mathcal{L}(t)| \quad (3.6)$$

A OIFT relaxada pode ser obtida através do seguinte algoritmo:

³Se $\Omega^i = \{o_1, o_2, \dots, o_n\}$ tem seus nós numerados por sua ordem de inserção, na qual $n = |\Omega^i|$, então $\Omega^{i-1} = \{o_1, \dots, o_k\}$, no qual $k = |\Omega^{i-1}|$, e $\Omega^i \setminus \Omega^{i-1}$ pode ser facilmente obtido como $\{o_{k+1}, \dots, o_n\}$.

Algoritmo 4. – ALGORITMO DA OIFT RELAXADA

- ENTRADA:** Grafo de imagem $\langle \mathcal{V}, \mathcal{A}, \omega \rangle$, conjuntos disjuntos não-vazios de sementes S_0 e S_1 .
- SAÍDA:** Mapa de rótulos $\mathcal{L} : \mathcal{V} \rightarrow \{0, 1\}$.
- AUXILIARES:** Fila de prioridade Q , variável tmp , a função de custo $V : \mathcal{V} \rightarrow [-\infty, \infty]$, e uma função status $S : \mathcal{V} \rightarrow \{0, 1\}$, em que $S(t) = 1$ para nós processados e $S(t) = 0$ para nós não processados. Vetor de conjuntos de nós não estacionários Ω^i . Vetor de segmentações parciais $\mathcal{L}^i, i \in \{0, 1, \dots, N\}$
1. **Para cada** $t \in \mathcal{V}$, **Faça**
 2. $S(t) \leftarrow 0$ e $V(t) \leftarrow -\infty$;
 3. **Se** $t \in S_0$, **então**
 4. $V(t) \leftarrow \infty, \mathcal{L}^0(t) \leftarrow 0$ e insira t em Q ;
 5. **Se** $t \in S_1$ **então**
 6. $V(t) \leftarrow \infty, \mathcal{L}^0(t) \leftarrow 1$ e insira t em Q .
 7. **Enquanto** $Q \neq \emptyset$ **Faça**
 8. Remova s de Q tal que $V(s)$ é máximo;
 9. $S(s) \leftarrow 1$;
 10. **Para cada** $\langle s, t \rangle \in \mathcal{A}$ tal que $S(t) = 0$ **Faça**
 11. **Se** $\mathcal{L}^0(s) = 1$ **então** $tmp \leftarrow \omega(s, t)$
 12. **Senão** $tmp \leftarrow \omega(t, s)$;
 13. **Se** $tmp > V(t)$ **então**
 14. $V(t) \leftarrow tmp$ e $\mathcal{L}^0(t) \leftarrow \mathcal{L}^0(s)$
 15. **Se** $t \notin Q$ **então** insira t em Q .
 16. Defina $\Omega^0 \leftarrow \emptyset$ e $\Omega^1 \leftarrow F(\mathcal{L}^0) \setminus S$.
 17. **Para cada** $i = 1, \dots, N$, **Faça**
 18. Compute $\mathcal{L}^i(p)$ para todo $p \in \Omega^i$ pela Eq. 3.3 e defina $\mathcal{L}^i(p) = \mathcal{L}^{i-1}(p)$ para todo $p \notin \Omega^i$.
 19. $\Omega^{i+1} \leftarrow [\Omega^i \cup ((\Omega^i \setminus \Omega^{i-1}) \oplus \mathcal{A})] \setminus S$.
 20. Defina $\mathcal{L}(p) = 1$ se $\mathcal{L}^N(p) \geq 0.5$ e $\mathcal{L}(p) = 0$ caso contrário, para todo $p \in \mathcal{V}$.
 21. **Retorne** \mathcal{L}

A seguir são exemplificadas duas simulações do método das relaxações, executando a partir de uma segmentação inicial por OIFT. Na Figura 3.2 é utilizada a Equação 3.3 com $W^i(s, t)$ tal como definido pela Equação 3.4 (tratando, por questões de simplicidade, o caso $\mathcal{L}^i(s) = \mathcal{L}^i(t)$ como ω_{st}) e na Figura 3.3 é utilizada a definição dada pela Equação 3.5.

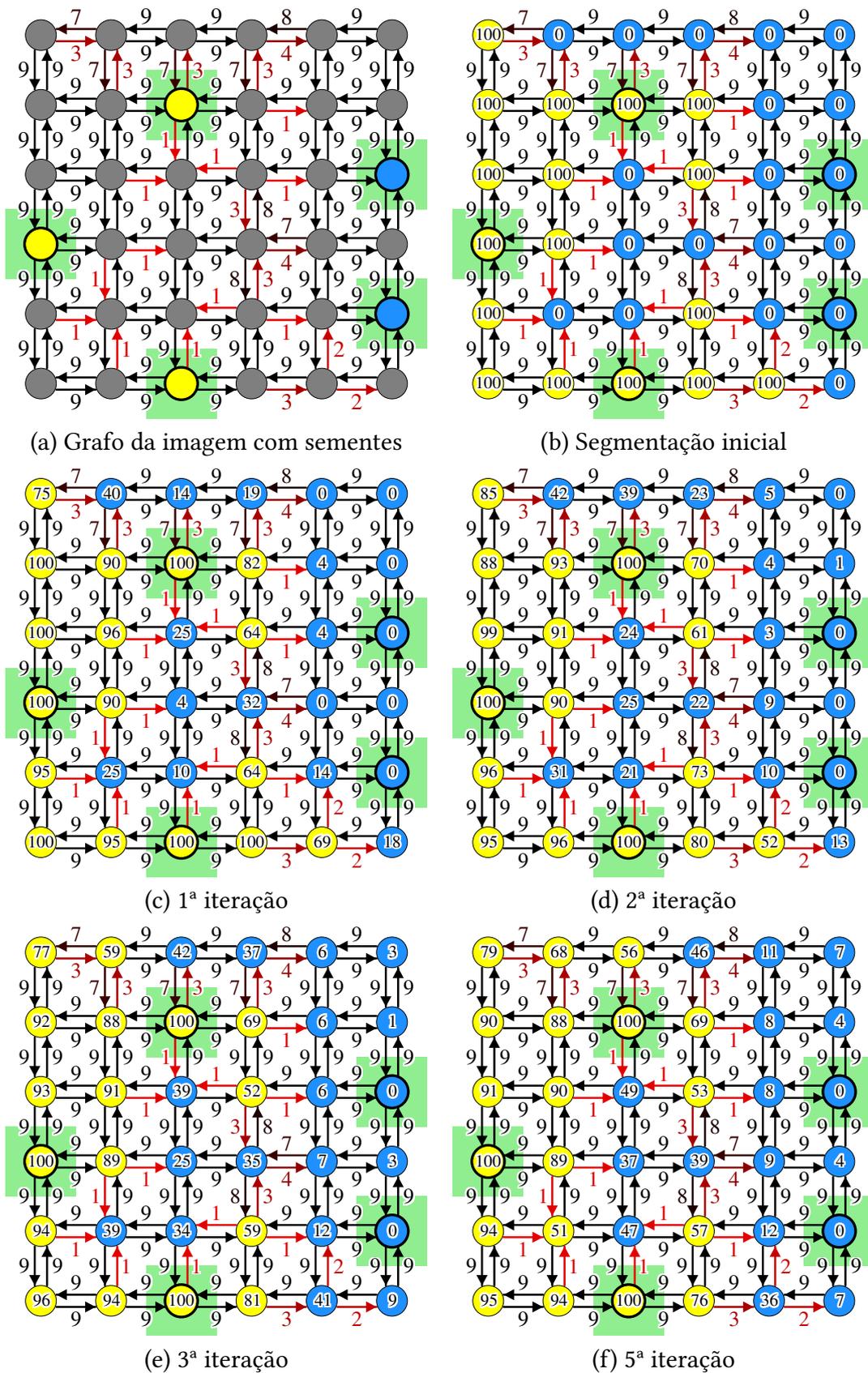


Figura 3.2: Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.4 (Parte 1). Os nós com o fundo em cor verde são sementes e os valores dentro dos nós indicam $100 \cdot \mathcal{L}^i(s)$.

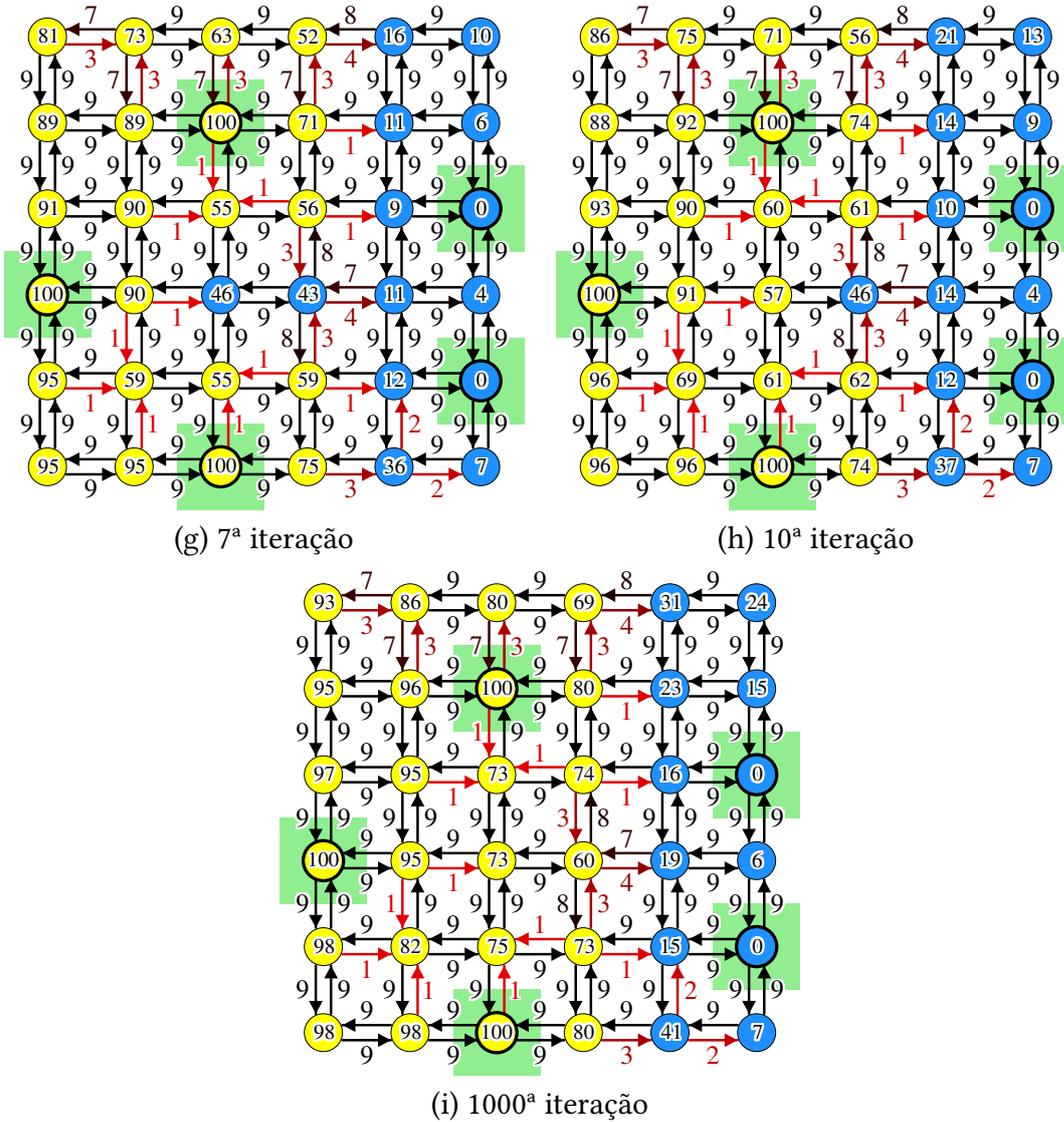


Figura 3.2: Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.4. Os nós com o fundo em cor verde são sementes e os valores dentro dos nós indicam $100 \cdot \mathcal{L}^i(s)$ (Parte 2).

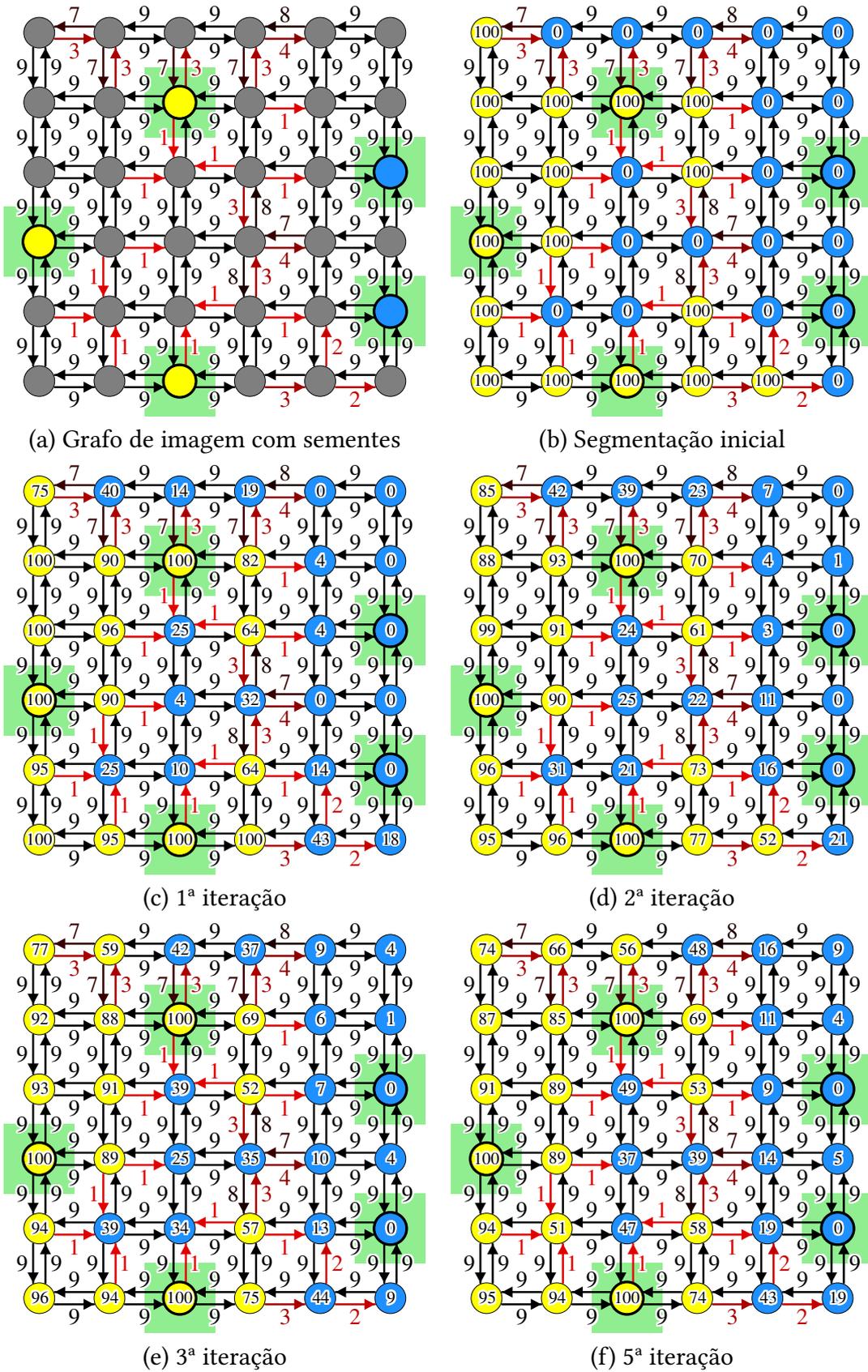


Figura 3.3: Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.5. Os nós com o fundo em cor verde são sementes e os valores dentro dos nós indicam $100 \cdot \mathcal{L}^i(s)$ (Parte 1).

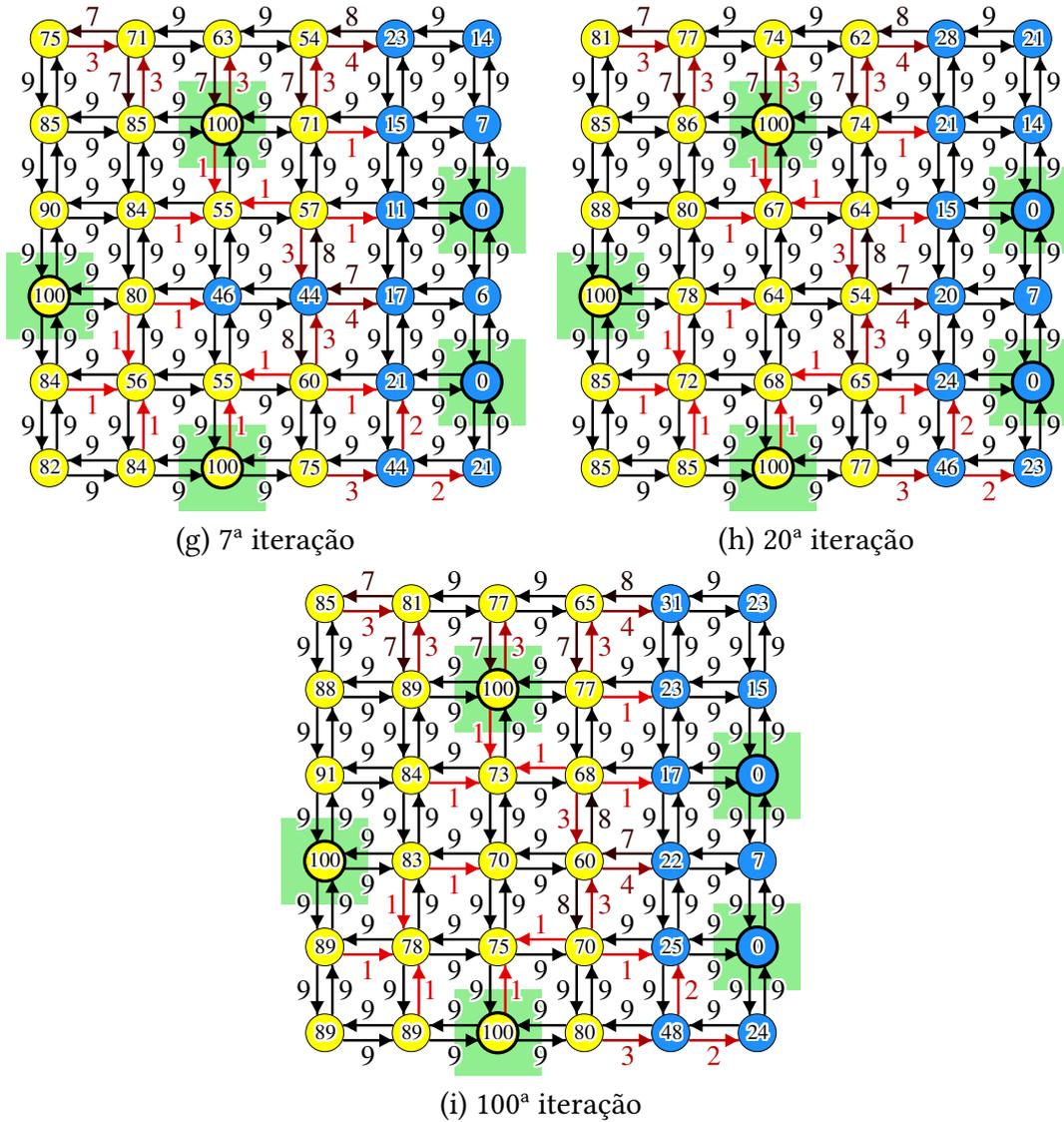


Figura 3.3: Exemplo do procedimento de relaxação para grafos direcionados usando a Equação 3.5. Os nós com o fundo em cor verde são sementes e os valores dentro dos nós indicam $100 \cdot \mathcal{L}^i(s)$ (Parte 2).

3.3 Relaxed Deep Extreme Cut

O *Deep Extreme Cut* (DEXTR) pode apresentar algumas deficiências em relação a aderência à borda. Erros residuais consideráveis em todo o limite são bastante comuns em *Deep Extreme Cut* (Figuras 3.4e e f). A correção interativa de todos esses erros residuais espalhados ao longo de toda a borda, a fim de gerar um resultado de segmentação verdadeiramente correto, exigiria um grande esforço de interação por parte do usuário, reduzindo ou até anulando completamente os ganhos do método. Para melhorar os resultados do método em benchmarks de comparação, normalmente um limiar (*threshold*) alto é considerado para eliminar falsos positivos, mas isso pode resultar na perda de partes importantes do objeto para algumas imagens (por exemplo, a asa esquerda nas Figuras 3.4c e d). Observe que não há um limiar único que funcione corretamente para todas as partes do objeto.

Neste trabalho, visamos atenuar os problemas de aderência à borda do DEXTR, utilizando um método híbrido entre o *Random Walks* e o *Deep Extreme Cut*, via o processo de relaxação em grafos direcionados proposto na Seção 3.1 (Figuras 3.4g e h). No entanto, para a correção de erros mais pronunciados, tais como os apresentados na asa esquerda das Figuras 3.4c e d, uma abordagem mais complexa com etapas adicionais se faz necessária (Oliveira, Caio L. Demario e P. A. Miranda, 2021), estando além do escopo da presente dissertação.

A fim de utilizar a saída do *Deep Extreme Cut* como segmentação inicial para o método das relaxações apresentado na Seção 3.1, é necessário definir os conjuntos de sementes de objeto S_1 e de fundo S_0 . Para o método híbrido proposto, o conjunto de sementes de objeto é definido como o conjunto dos 4 pontos extremos inicialmente selecionados pelo usuário e o conjunto de sementes de fundo como o conjunto de pixels fora da caixa delimitadora (*bounding box*) do objeto B definida a partir dos 4 pontos extremos:

$$B = \{t = (t_x, t_y) \in \mathcal{V} : x_{min} - \delta \leq t_x \leq x_{max} + \delta \wedge y_{min} - \delta \leq t_y \leq y_{max} + \delta\} \quad (3.7)$$

tal que:

$$(x_{min}, y_{min}) = \left(\min_{(x,y) \in S_1} x, \min_{(x,y) \in S_1} y \right) \quad (3.8)$$

$$(x_{max}, y_{max}) = \left(\max_{(x,y) \in S_1} x, \max_{(x,y) \in S_1} y \right) \quad (3.9)$$

Na Equação 3.7, o valor de δ define uma pequena margem de segurança para evitar a definição de sementes de fundo dentro do objeto no caso de seleção imprecisa dos pontos extremos por parte do usuário. Nos experimentos consideramos $\delta = 3$ pixels.

Para o cálculo da relaxação, usamos um grafo de vizinhança-8, com medida de dissimilaridade κ_{st} na Equação 3.1 dada por $\kappa_{st} = G(s) + G(t)$, onde $G(s)$ representa a magnitude do gradiente de Sobel das cores no espaço do CIE Lab, $\beta = 8$, $\alpha = 0,5$ e a orientação da polaridade de borda é definida a partir do mapa de probabilidade resultante do *Deep Extreme Cut*, que é usado no lugar de $I(s)$ e $I(t)$ nas condições da Equação 3.1.

Algoritmo 5. – *RELAXED DEEP EXTREME CUT*

ENTRADA: Imagem \hat{I} , conjunto de 4 pontos extremos S_1

AUXILIARES: Vetor de conjuntos de nós não estacionários $\Omega^i, i \in \{0, 1, \dots, N\}$

SAÍDA: Mapa de rótulos $\mathcal{L} : \mathcal{V} \rightarrow \{0, 1\}$.

1. Defina $S_0 = \{t \in \mathcal{V} : t \notin \mathcal{B}\}$ como o conjunto de pixels fora da caixa delimitadora.
2. Compute DEXTR usando os pixels em S_1 como pontos extremos para obter \mathcal{L}^0 .
3. Defina $S = S_0 \cup S_1$.
4. Defina $\Omega^0 \leftarrow \emptyset$ e $\Omega^1 \leftarrow F(\mathcal{L}^0) \setminus S$.
5. **Para cada** $i = 1, \dots, N$, **Faça**
 6. $\left[\begin{array}{l} \text{Compute } \mathcal{L}^i(p) \text{ para todo } p \in \Omega^i \text{ pela Eq. 3.3 e defina } \mathcal{L}^i(p) = \mathcal{L}^{i-1}(p) \text{ para todo } p \notin \Omega^i. \\ \Omega^{i+1} \leftarrow [\Omega^i \cup ((\Omega^i \setminus \Omega^{i-1}) \oplus \mathcal{A})] \setminus S. \end{array} \right.$
 7. $\Omega^{i+1} \leftarrow [\Omega^i \cup ((\Omega^i \setminus \Omega^{i-1}) \oplus \mathcal{A})] \setminus S$.
 8. Defina $\mathcal{L}(p) = 1$ se $\mathcal{L}^N(p) \geq 0.5$ e $\mathcal{L}(p) = 0$ caso contrário, para todo $p \in \mathcal{V}$.
 9. **Retorne** \mathcal{L}

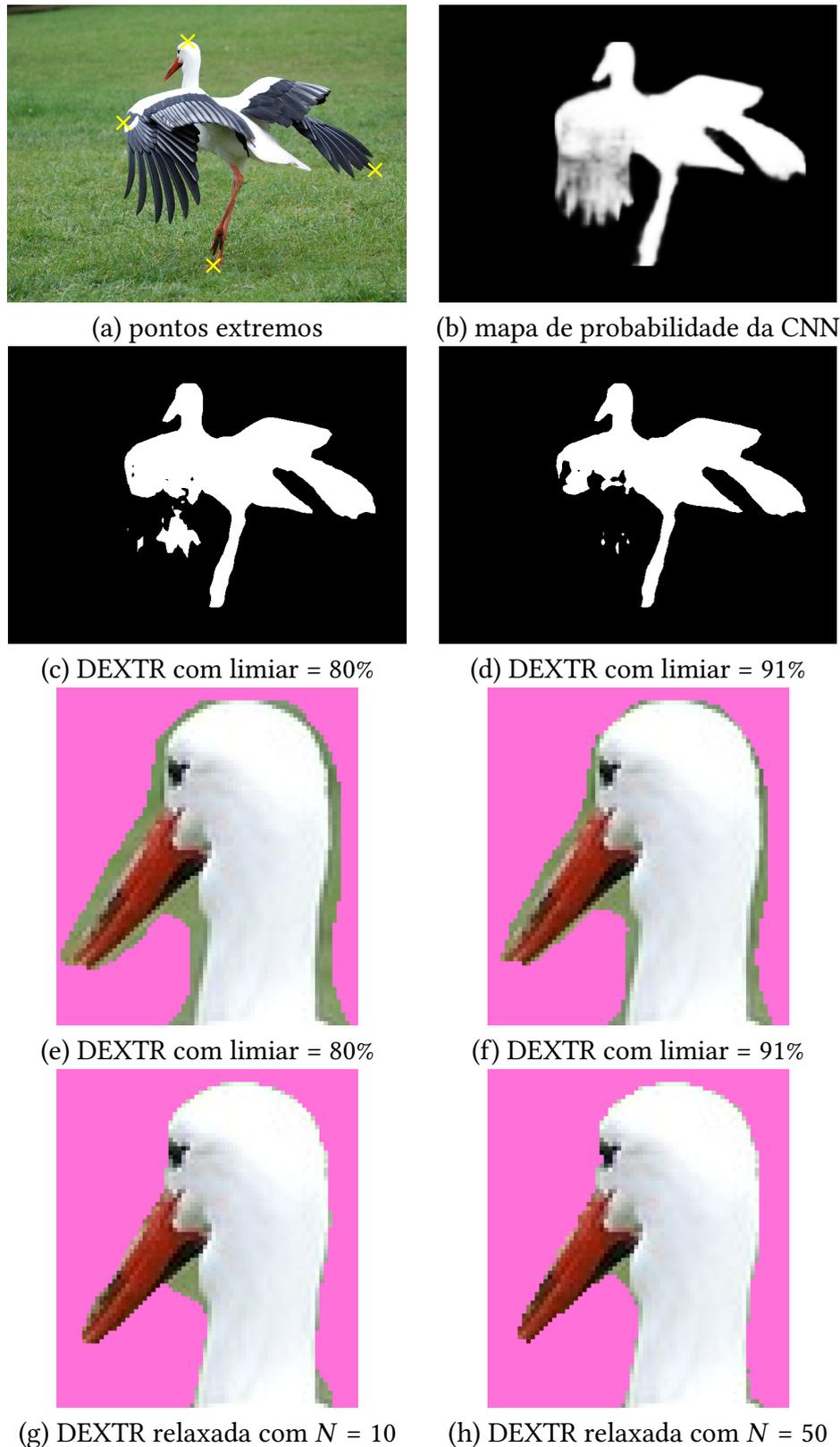


Figura 3.4: Exemplo da segmentação de uma imagem natural por Deep Extreme Cut. (a) A seleção de pontos extremos com 4 cliques. (b) O mapa de probabilidade para o objeto produzido pela CNN com pesos ajustados na base de treinamento SBD + PASCAL. (c) O resultado por Deep Extreme Cut com seu valor limiar padrão de 80%. (d) O resultado por Deep Extreme Cut com valor limiar de 91% conforme necessário para um desempenho ideal no conjunto de dados do GrabCut (Rother, Kolmogorov e Blake, 2004). (e-f) Observe que o Deep Extreme Cut tem baixa aderência à borda. (g-h) Uma melhor aderência à borda é obtida via o procedimento proposto de relaxação.

Capítulo 4

Resultados Experimentais

Neste capítulo apresentamos os resultados experimentais dos métodos propostos, dados pelos Algoritmos 4 e 5.

4.1 Medidas de acurácia

A seguir apresentamos as medidas de acurácia usadas nos experimentos.

4.1.1 Coeficiente de Dice

O coeficiente de Dice (Dice, 1945) é uma estatística que mede a similaridade entre duas amostras. A similaridade é medida por um valor em $[0, 1]$ sendo 1 o valor mais alto de similaridade e 0 o mais baixo. O cálculo do coeficiente se dá pela fórmula a seguir:

$$dice(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|} \quad (4.1)$$

Onde $|\cdot|$ indica a cardinalidade do conjunto e A e B são as amostras.

Para o uso como medida de acurácia de segmentação de imagens, medimos a similaridade entre o conjunto dos pixels rotulados como objetos pela segmentação \mathcal{O} e o conjunto dos pixels rotulados como objeto no gabarito da segmentação \mathcal{G} . Então temos:

$$dice(\mathcal{O}, \mathcal{G}) = \frac{2 \cdot |\mathcal{O} \cap \mathcal{G}|}{|\mathcal{O}| + |\mathcal{G}|} \quad (4.2)$$

Comparando a segmentação da imagem com o gabarito podemos descrever quatro conjuntos de pixels:

- Pixels rotulados como objeto pela segmentação e pelo gabarito, denominados verdadeiros positivos, ou *true positive*, TP.
- Pixels rotulados como objeto pela segmentação e como fundo pelo gabarito, denominados falsos positivos, ou *false positive*, FP.

- Pixels rotulados como fundo pela segmentação e como objeto pelo gabarito, denominados falsos negativos, ou *false negative*, FN.
- Pixels rotulados como fundo pela segmentação e pelo gabarito, denominados verdadeiros negativos, ou *true negative*, TN.

Esses conjuntos podem ser apresentados na forma de uma matriz de confusão, como na tabela abaixo:

		Gabarito	
		Objeto	Fundo
Segmentação	Objeto	TP Pixels rotulados corretamente como objeto	FP Pixels rotulados incorretamente como objeto
	Fundo	FN Pixels rotulados incorretamente como fundo	TN Pixels rotulados corretamente como fundo

Tabela 4.1: Matriz de confusão para segmentação e gabarito da segmentação.

E baseado nos conjuntos da matriz de confusão podemos reescrever o coeficiente de Dice como:

$$dice(\mathcal{O}, \mathcal{G}) = \frac{2 \cdot |TP|}{2 \cdot |TP| + |FN| + |FP|} \quad (4.3)$$

4.1.2 Coeficiente de Jaccard

O coeficiente de Jaccard (Jaccard, 1901), assim como o coeficiente de Dice, é uma estatística usada para medir a similaridade de dois conjuntos, e é dado pela seguinte equação:

$$Jaccard(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (4.4)$$

Também podemos descrever o coeficiente de Jaccard com base nos conjuntos da matriz de confusão como:

$$Jaccard(A, B) = \frac{|TP|}{|TP| + |FN| + |FP|} \quad (4.5)$$

4.1.3 Métrica Balanceada Geral (GB)

A acurácia pode também ser medida utilizando a métrica *Balanceada Geral* (GB) (F. A. M. Cappabianco *et al.*, 2019) dada pela Equação 4.6, com $\delta = 1$, onde TP, FP e FN representam os conjuntos de pixels verdadeiro positivo, falso positivo e falso negativo, respectivamente e \mathcal{G} é o conjunto de pixels que representa o objeto de interesse no gabarito. GB com

$\delta = 1$ é uma versão balanceada e geral da métrica de Jaccard (F. A. M. Cappabianco *et al.*, 2019).

$$GB = \frac{\delta|TP|}{\delta|TP| + |FN| + FP_S}, \quad FP_S = \sum_{i=0}^{+\infty} \frac{|FP|^{i+1}}{|G|^i} \quad (4.6)$$

4.1.4 Quociente Isoperimétrico

O quociente isoperimétrico é uma medida que indica circularidade e é definido como a razão da área do objeto A pela área do círculo com o mesmo perímetro p . O quociente Q é um valor menor ou igual a 1, em que o valor 1 indica que o objeto é um círculo.

$$Q = \frac{A}{\pi r^2} \leq 1$$

e sendo o perímetro do círculo definido como $2\pi r$, temos:

$$Q = \frac{4\pi A}{p^2} \leq 1 \quad (4.7)$$

O perímetro de um objeto segmentado em uma imagem é a quantidade de pixels que estão na borda do objeto, ou seja, que possuem vizinhos rotulados como fundo.

4.2 OIFT Relaxada

Experimentos foram conduzidos no grafo definido pela Equação 3.1, comparando os métodos híbridos:

- GC e ORFC (Bejar e P. Miranda, 2015) (*ORFC + GC*),
- Malmberg et al. (Malmberg *et al.*, 2010) (*Relax*), com OIFT como \mathcal{L}^0 em vez da IFT, e
- as duas versões propostas da OIFT Relaxada, utilizando as Equações 3.4 e 3.5 (*ORelax₁* e *ORelax₂*).

Como métodos de linha de base, também incluímos nas comparações:

- OIFT (*Normal*),
- OIFT com pós-processamento pelo filtro moda (*Filtro Moda*), e
- RW (Singaraju, Grady e Vidal, 2008) como um caso limite de N .

Nesta seção, apresentamos os resultados de acurácia obtidos em experimentos envolvendo conjuntos de dados 2D, compostos por 40 cortes de imagens de RM do pé. As 40 fatias de tamanho 256×256 pixels foram selecionadas a partir de dados de volumes de ressonância magnética 3D pertencentes ao pé esquerdo ou direito de 20 indivíduos diferentes. Vários cortes (1-3) foram selecionados aleatoriamente em locais anatômicos semelhantes de cada volume 3D. O objeto de interesse nessas fatias são os dois grandes

ossos do complexo peritalar, a saber, o tálus e o calcâneo. Finalmente, 40 imagens de corte de tamanho 512×512 pixels de estudos torácicos de tomografia (TC) de 10 indivíduos foram usadas para segmentar o fígado. Os dados de imagens e *ground truth* foram obtidos de um especialista do departamento de radiologia da Universidade da Pensilvânia.

As mesmas sementes foram usadas para todos os métodos, as quais foram manualmente selecionadas para entregar uma boa indicação aproximada de objeto, gerando resultados razoáveis para os métodos de linha de base.

Na Eq. 3.1, utilizamos $\kappa_{st} = G(s) + G(t)$, na qual $G(s)$ é o gradiente de Sobel e $\beta = 8$ para melhores resultados, caso contrário, o componente RW é sensível ao posicionamento de semente. O valor de suavidade obtido foi calculado pelo quociente isoperimétrico, a acurácia pelo coeficiente de Dice e o tempo foi obtido em um processador Intel Core i3-5005U a 2GHz $\times 4$.

As Figuras 4.1, 4.2 e 4.3 mostram as curvas experimentais, variando o parâmetro de suavidade, correspondente:

- à quantidade de relaxações N nos métodos *Relax*, *ORelax₁* e *ORelax₂*, ou
- ao raio do filtro moda em pixels utilizado no método de linha de base da OIFT com pós-processamento pelo filtro moda (*Filtro Moda*).

Note que os métodos *ORFC + GC* e OIFT (*Normal*) não são afetados por variações no parâmetro de suavidade, dado que eles não possuem tal parâmetro. As curvas dos métodos *Relax*, *ORelax₁* e *ORelax₂* são apresentadas para diferentes valores de α que controla a polaridade de borda. Em todos os casos, ambas as versões propostas (*ORelax₁* e *ORelax₂*) melhoraram a qualidade percebida dos resultados da segmentação, aumentando a circularidade e acurácia em relação à OIFT, e com tempos computacionais próximos a *ORFC+GC* para $N < 100$, que poderiam ser ainda reduzidos, já que são facilmente paralelizáveis. Note que Dice piora para N grande, indicando que os métodos híbridos com $80 < N < 150$ levam a melhores resultados do que suas contrapartes por RW. Por exemplo, *ORelax₁* tem apenas 0,970 de Dice para $N = 1000$ na imagem do fígado. Alguns resultados com $N = 100$ são apresentados nas Figuras 4.4 e 4.5.

Na Figura 4.6, a OIFT regular tem um coeficiente de Dice de 98,05% (Figura 4.6b), a OIFT relaxada (*Malmberg et al., 2010*), quando aplicada ao digrafo da OIFT, com 100 iterações tem 97,25% (Figura 4.6c), e o filtro da moda tem 97,75%, 96,17% e 84,90% para raios de 30, 50 e 100 pixels (Figuras 4.6d e 4.6e), respectivamente, enquanto a OIFT relaxada orientada, tal como proposto neste trabalho, tem 98,27% (Figura 4.6f).

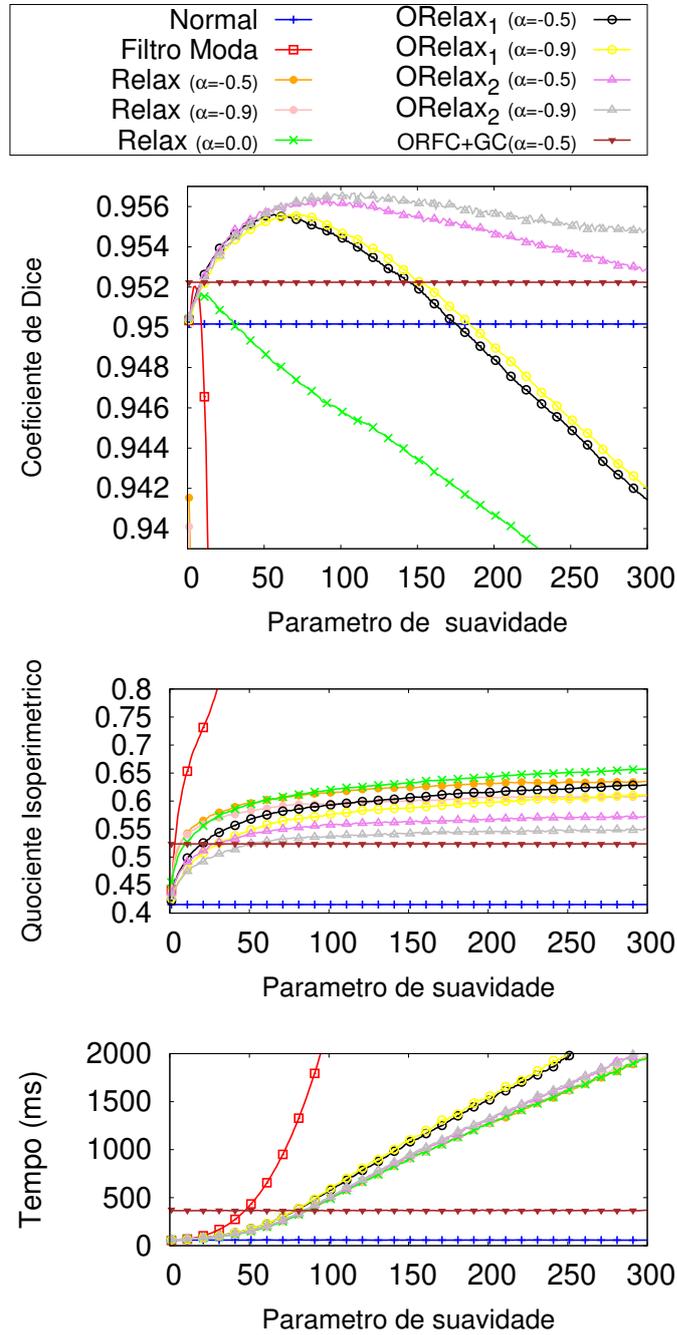


Figura 4.1: Resultados experimentais de segmentação do osso tálus. Primeira linha: Curvas médias de acurácia de Dice para diferentes valores do parâmetro de suavidade e métodos. Segunda linha: Curvas médias de circularidade pelo quociente isoperimétrico. Terceira linha: Curvas médias de tempo de execução.

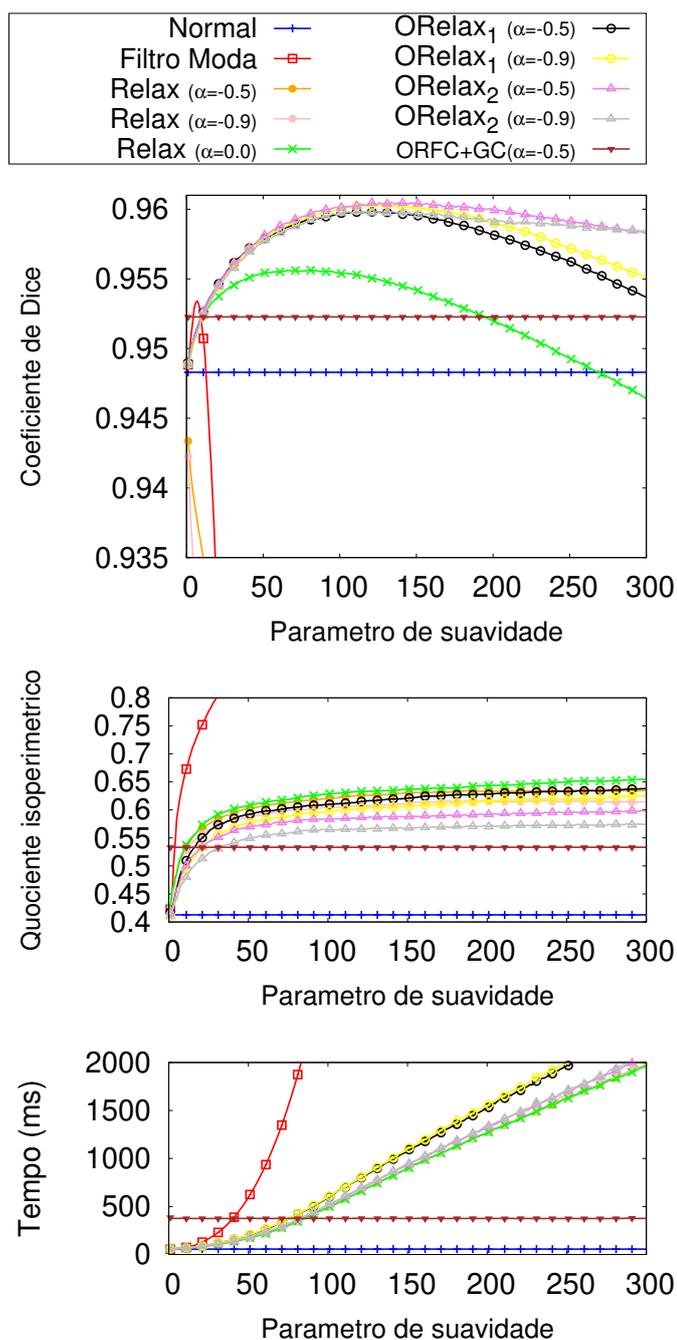


Figura 4.2: Resultados experimentais de segmentação do osso calcâneo. Primeira linha: Curvas médias de acurácia de Dice para diferentes valores do parâmetro de suavidade e métodos. Segunda linha: Curvas médias de circularidade pelo quociente isoperimétrico. Terceira linha: Curvas médias de tempo de execução.

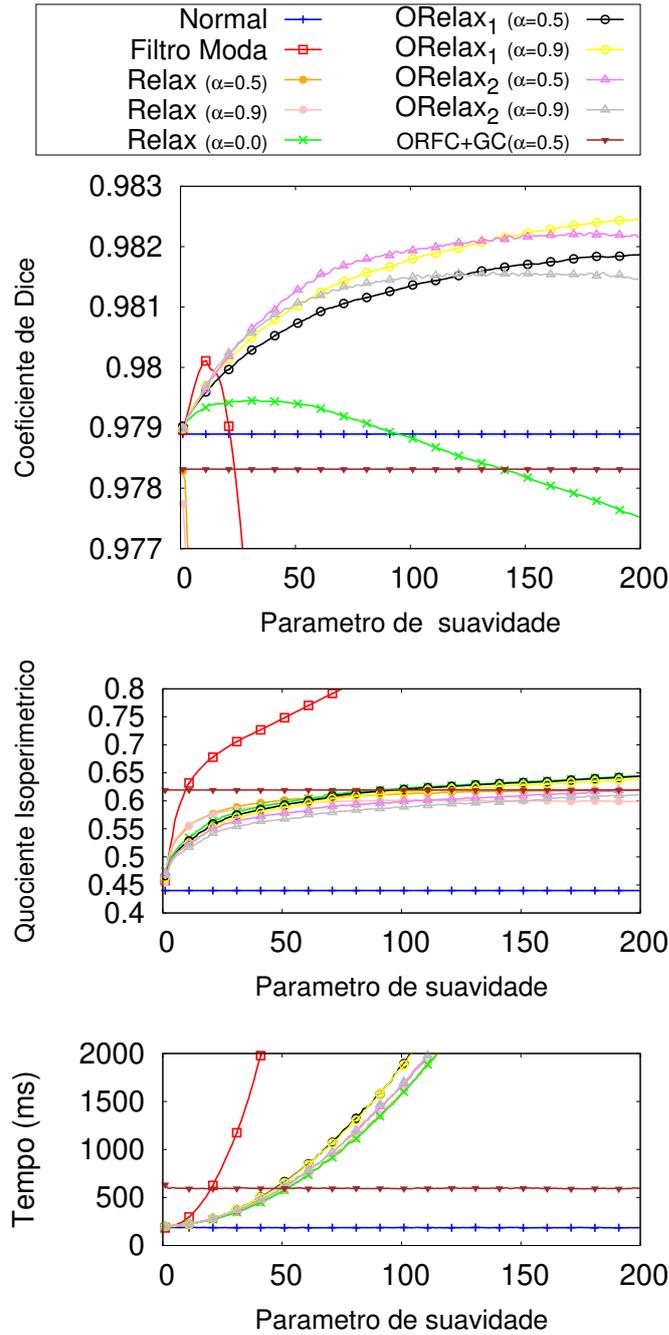


Figura 4.3: Resultados experimentais de segmentação do fígado. Primeira linha: Curvas médias de acurácia de Dice para diferentes valores do parâmetro de suavidade e métodos. Segunda linha: Curvas médias de circularidade pelo quociente isoperimétrico. Terceira linha: Curvas médias de tempo de execução.

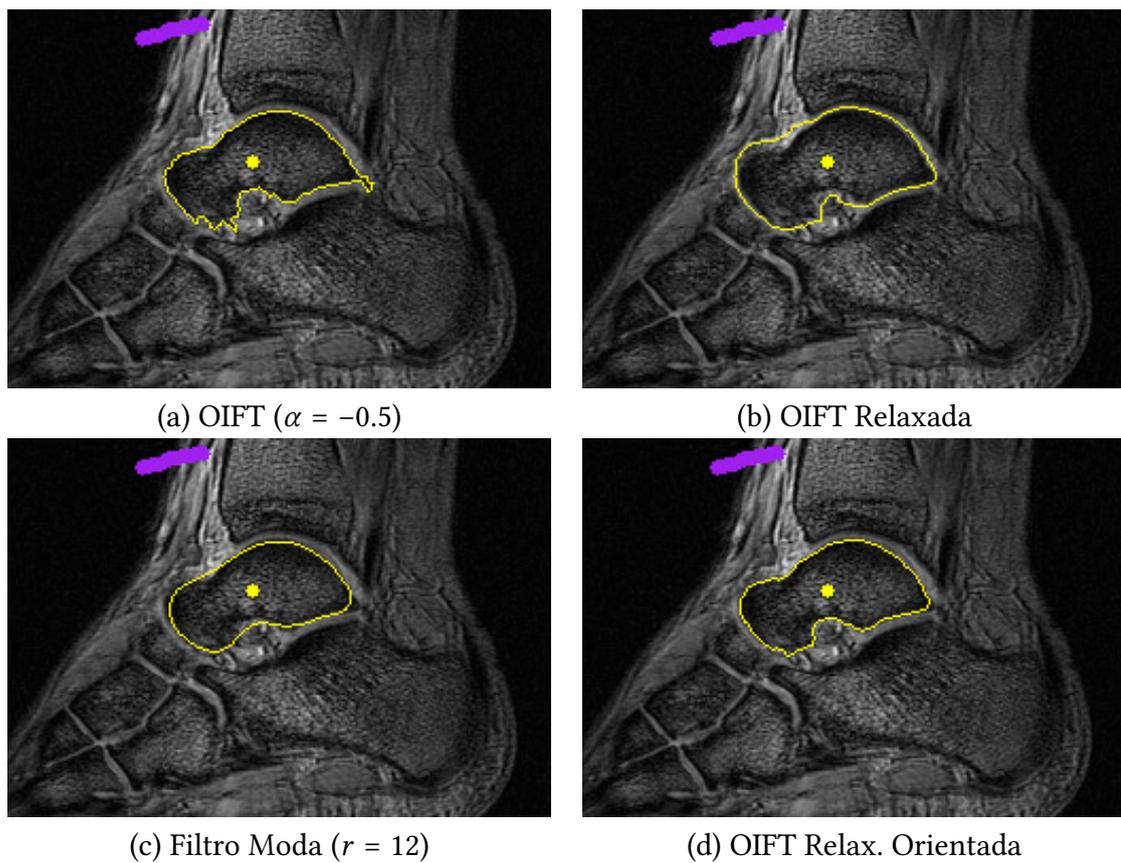


Figura 4.4: Segmentação do osso tálus. (a) OIFT sem relaxação. (b) OIFT com relaxação, como proposto em *Malmberg et al., 2010*, aplicada em seu grafo direcionado. (c) OIFT com um filtro de moda com raio de 12 pixels. (d) OIFT proposta com relaxação orientada.

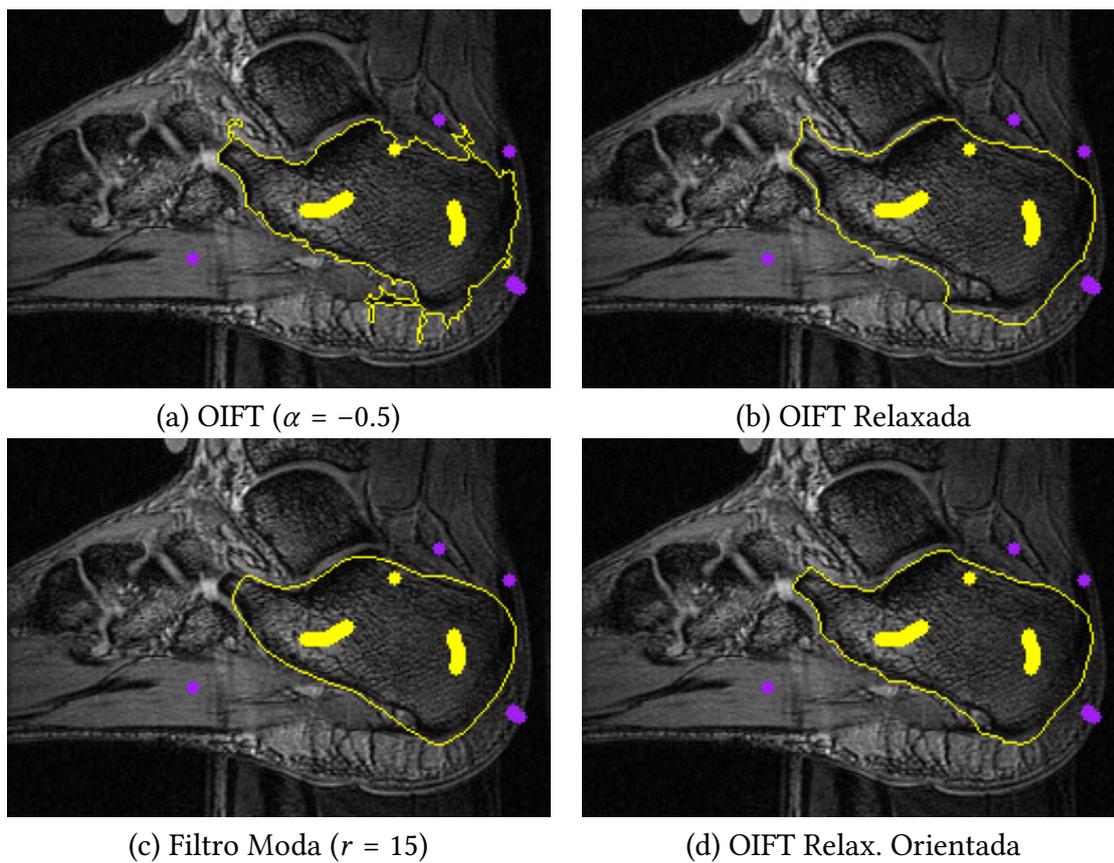


Figura 4.5: Segmentação do osso calcâneo. (a) OIFT sem relaxação. (b) OIFT com relaxação, como proposto em *Malmberg et al., 2010*, aplicada em seu grafo direcionado. (c) OIFT com um filtro de moda com raio de 15 pixels. (d) OIFT proposta com relaxação orientada.

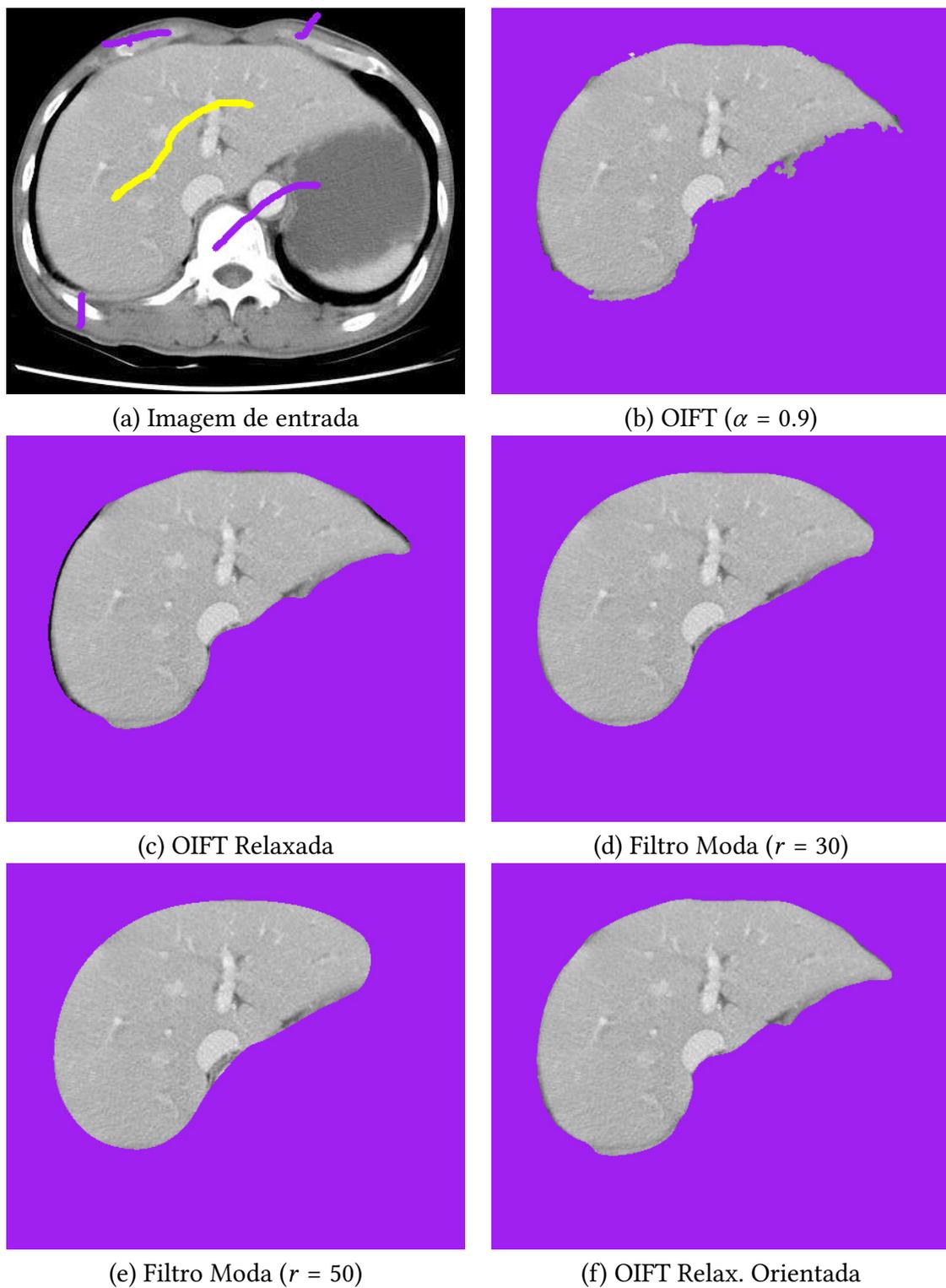


Figura 4.6: (a) imagem de CT torácica para segmentar o fígado. (b) OIFT sem relaxação. (c) OIFT com relaxação, como proposto em *Malmberg et al., 2010*, sendo aplicada no grafo direcionado da OIFT. (d-e) OIFT com pós-processamento por um filtro de moda, com raios 30 e 50, respectivamente. (f) OIFT proposta com relaxação orientada.

4.2.1 Implementação em 3D

Avaliamos também o método da OIFT relaxada (Algoritmo 4) em conjuntos de dados 3D de ressonância magnética de indivíduos normais, na faixa etária de 25 a 49 anos. As imagens foram adquiridas com um scanner 2T Elscint e tamanho de voxel de $0,98 \times 0,98 \times 1,00 \text{ mm}^3$. A implementação do Algoritmo 4 em volumes 3D é direta, bastando o uso de um grafo de imagens tridimensionais de vizinhança-6, tal como descrito na Seção 2.3.1. Para o cálculo da relaxação, usamos a medida de dissimilaridade $\kappa_{st} = G(s) + G(t)$ na Equação 3.1, onde $G(s)$ representa a magnitude do gradiente de Sobel, e usamos $\beta = 8$ e $\alpha = 0.5$.

As Figuras 4.7, 4.8, 4.9, 4.10, 4.11, 4.12 e 4.13 mostram a segmentação 3D do cérebro (tarefa conhecida como *skull stripping*) em volumes de ressonância magnética (RM) do crânio, com os resultados qualitativos da OIFT e da OIFT relaxada orientada (com $N = 100$ iterações de relaxação) nos planos sagital, coronal e axial, incluindo também as renderizações 3D dos resultados. Observe o maior grau de suavidade dos contornos obtidos pela OIFT relaxada, evitando vazamentos comumente presentes nos resultados da OIFT regular. Em especial, nas Figuras 4.10, 4.11, 4.12 e 4.13, é fácil observar a eliminação de erros da segmentação por OIFT na região do nervo óptico via o procedimento proposto de relaxação orientada. Outro ponto a ser observado é o aumento da simetria da máscara de segmentação em relação aos hemisférios cerebrais, conforme pode ser observado nos cortes axial e coronal. Por outro lado, o aumento de suavidade afeta negativamente a segmentação dos sulcos no córtex cerebral. Logo, uma ideia a ser explorada seria configurar de modo adaptativo o número de iterações de relaxação de modo otimizado para as diferentes regiões da imagem.

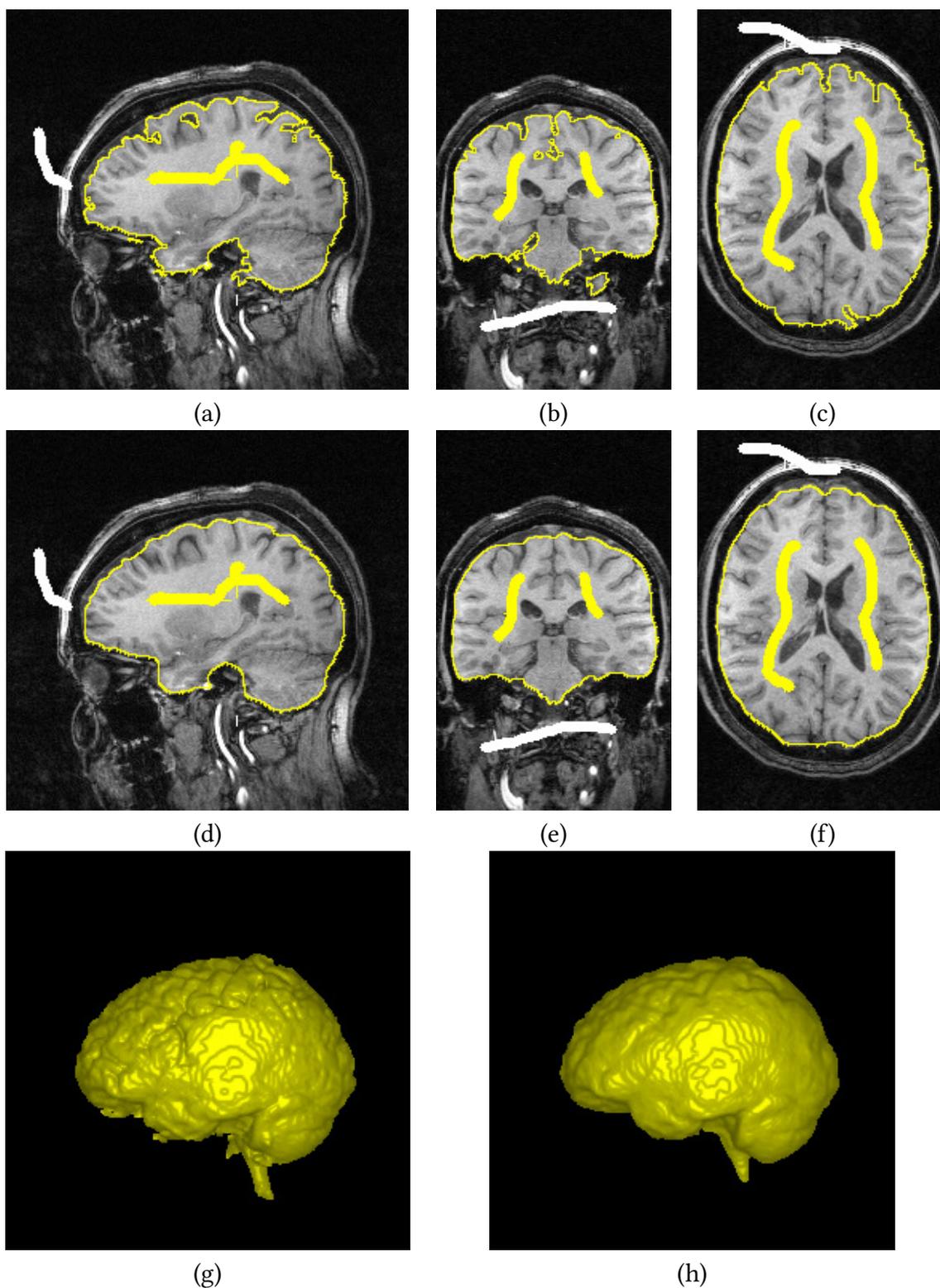


Figura 4.7: Exemplo #1 de segmentação 3D do cérebro em volume de ressonância magnética (RM) do crânio. (a-c) Resultado da OIFT nos planos sagital, coronal e axial. (d-f) Resultado da OIFT Relax. Orientada nos planos sagital, coronal e axial. (g-h) Renderizações 3D dos resultados da OIFT e OIFT Relax. Orientada, respectivamente.

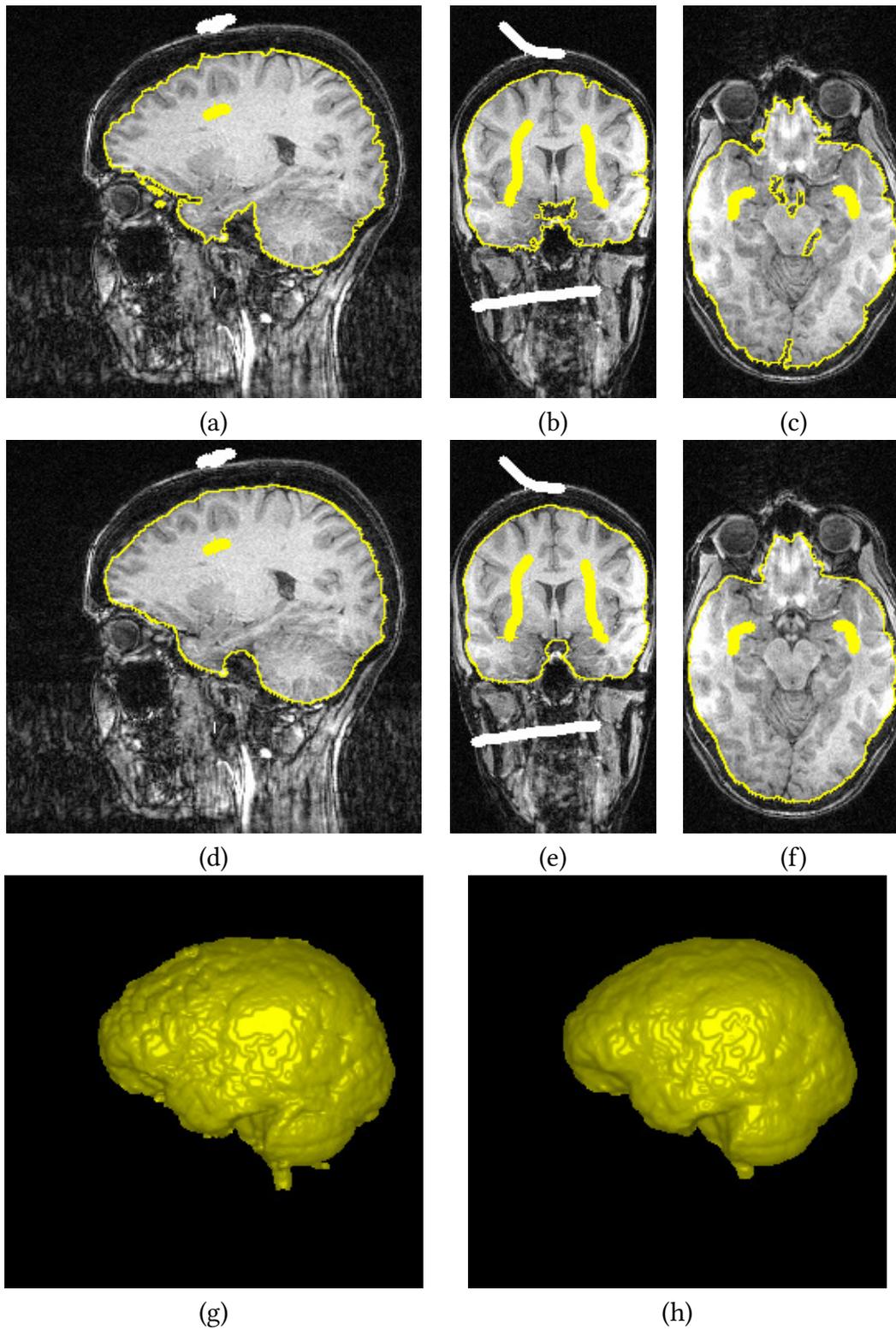


Figura 4.8: Exemplo #2 de segmentação 3D do cérebro em volume de ressonância magnética (RM) do crânio. (a-c) Resultado da OIFT nos planos sagital, coronal e axial. (d-f) Resultado da OIFT Relax. Orientada nos planos sagital, coronal e axial. (g-h) Renderizações 3D dos resultados da OIFT e OIFT Relax. Orientada, respectivamente.

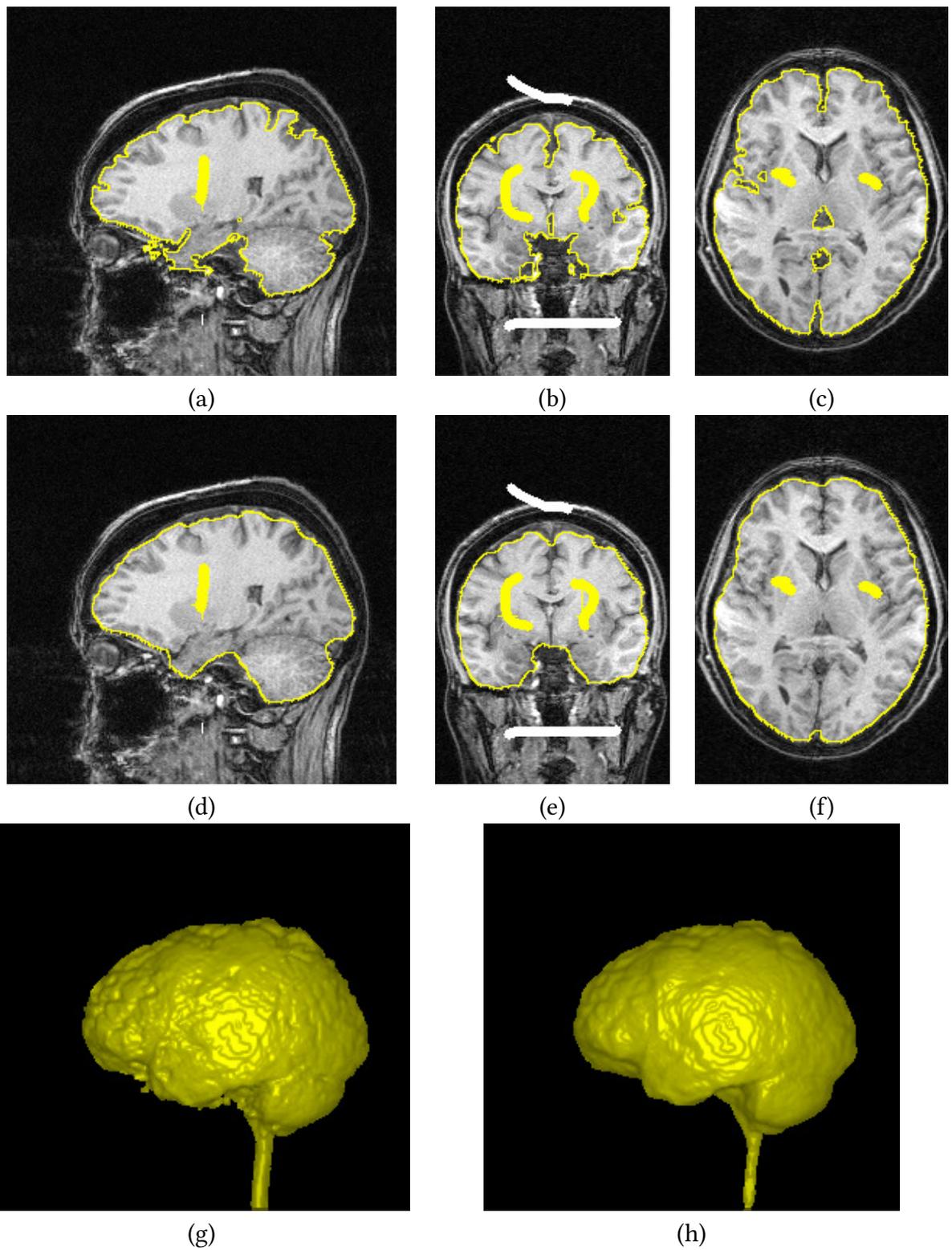


Figura 4.9: Exemplo #3 de segmentação 3D do cérebro em volume de ressonância magnética (RM) do crânio. (a-c) Resultado da OIFT nos planos sagital, coronal e axial. (d-f) Resultado da OIFT Relax. Orientada nos planos sagital, coronal e axial. (g-h) Renderizações 3D dos resultados da OIFT e OIFT Relax. Orientada, respectivamente.

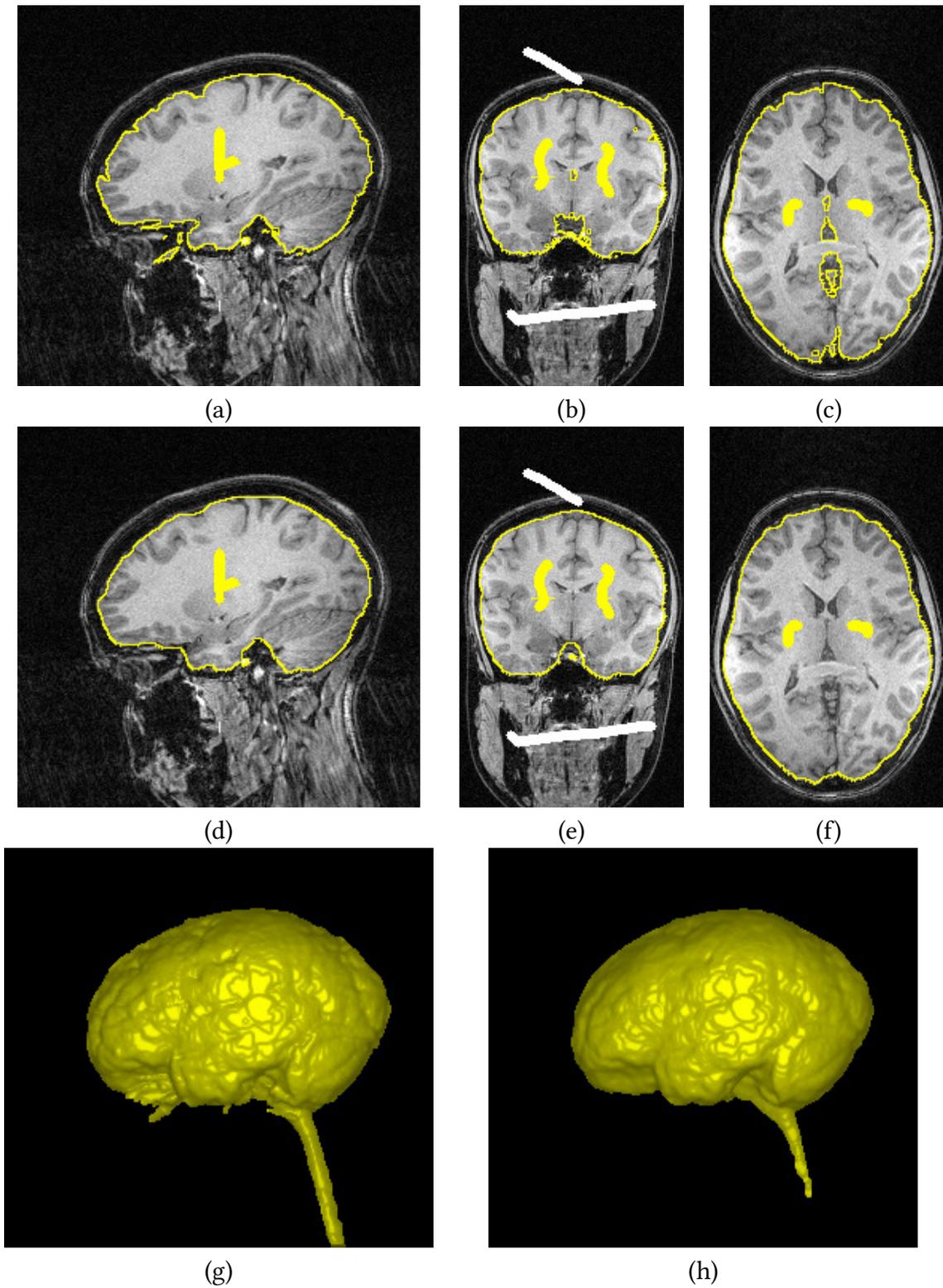


Figura 4.10: Exemplo #4 de segmentação 3D do cérebro em volume de ressonância magnética (RM) do crânio. (a-c) Resultado da OIFT nos planos sagital, coronal e axial. (d-f) Resultado da OIFT Relax. Orientada nos planos sagital, coronal e axial. (g-h) Renderizações 3D dos resultados da OIFT e OIFT Relax. Orientada, respectivamente.

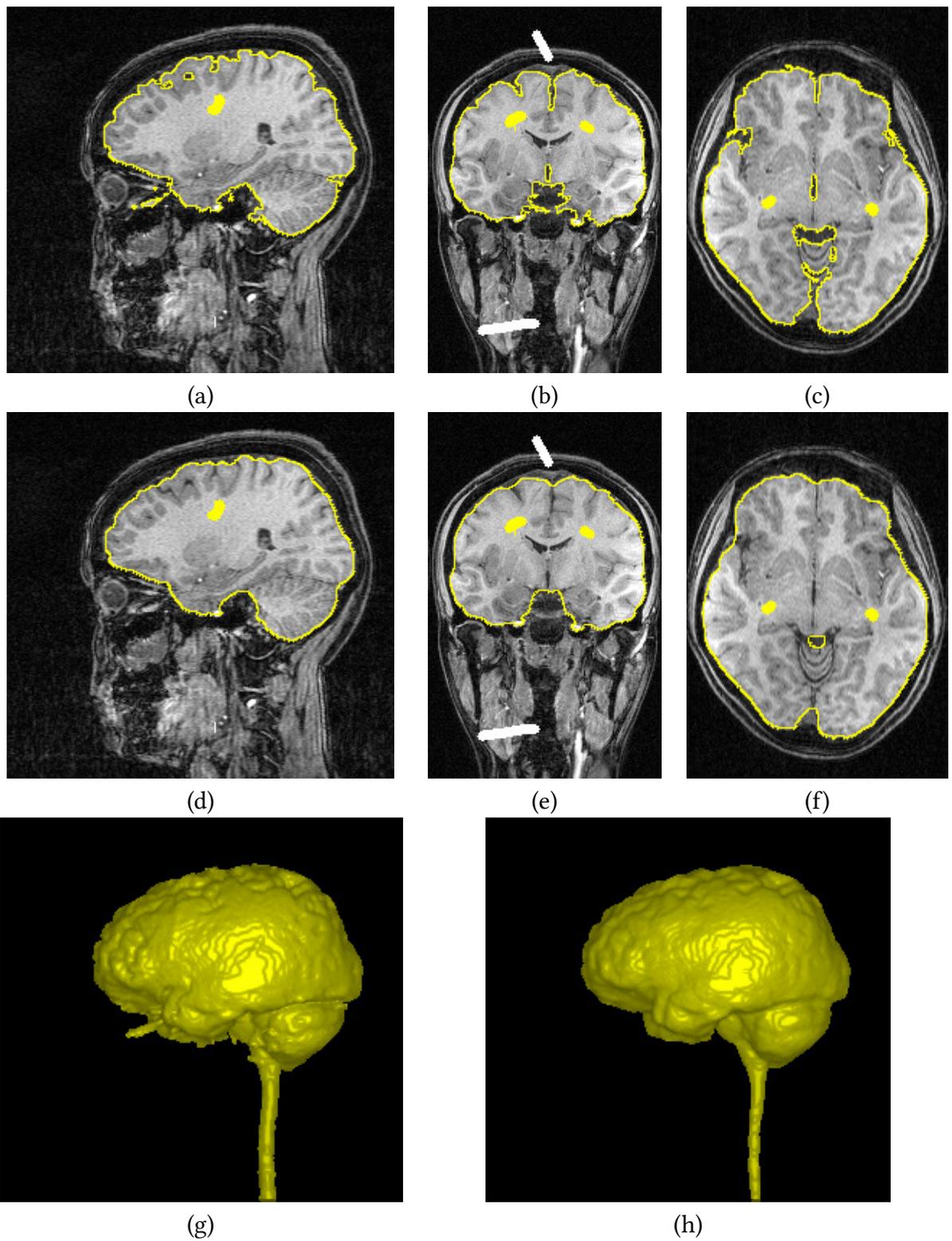


Figura 4.11: Exemplo #5 de segmentação 3D do cérebro em volume de ressonância magnética (RM) do crânio. (a-c) Resultado da OIFT nos planos sagital, coronal e axial. (d-f) Resultado da OIFT Relax. Orientada nos planos sagital, coronal e axial. (g-h) Renderizações 3D dos resultados da OIFT e OIFT Relax. Orientada, respectivamente.

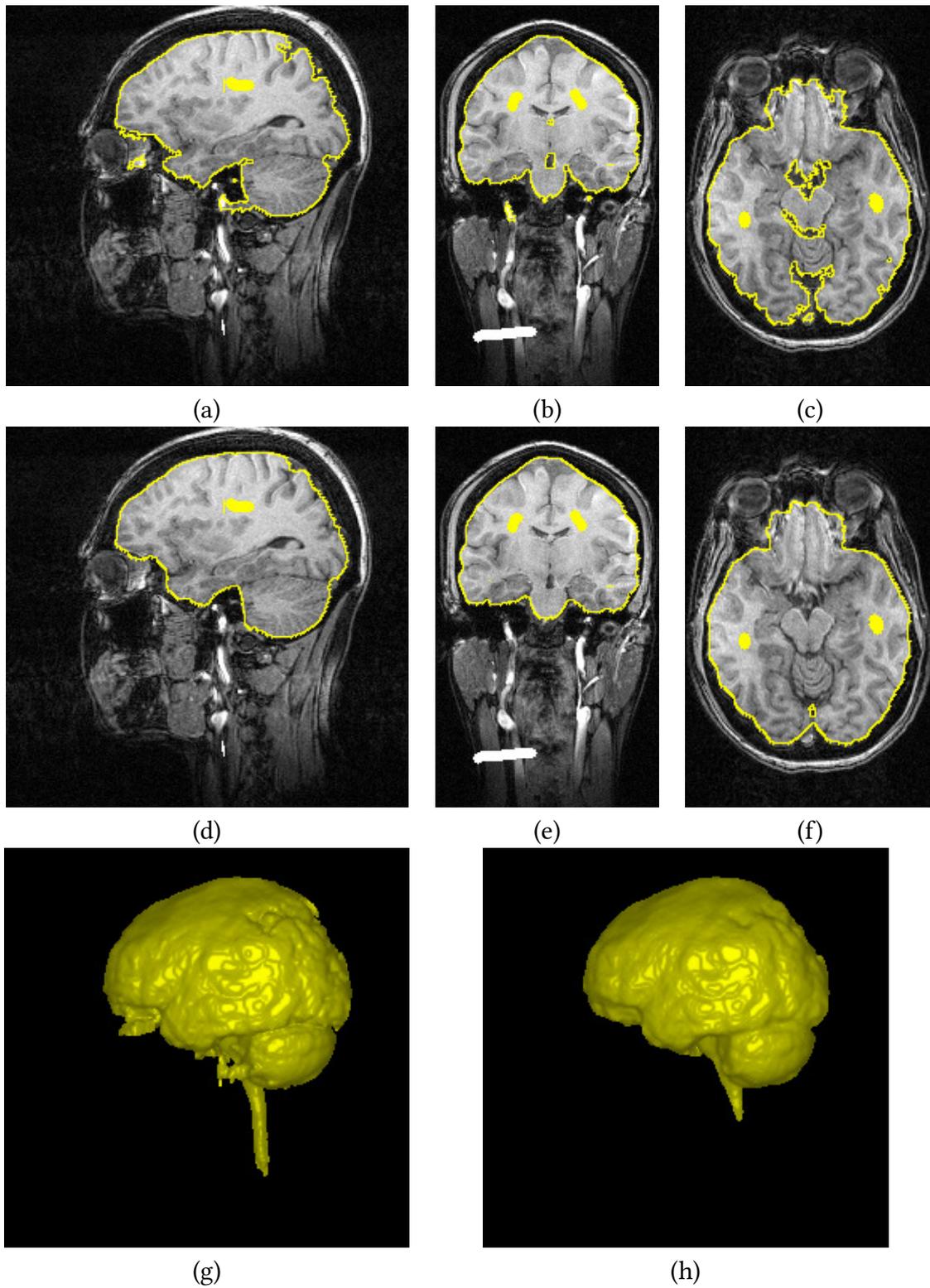


Figura 4.12: Exemplo #6 de segmentação 3D do cérebro em volume de ressonância magnética (RM) do crânio. (a-c) Resultado da OIFT nos planos sagital, coronal e axial. (d-f) Resultado da OIFT Relax. Orientada nos planos sagital, coronal e axial. (g-h) Renderizações 3D dos resultados da OIFT e OIFT Relax. Orientada, respectivamente.

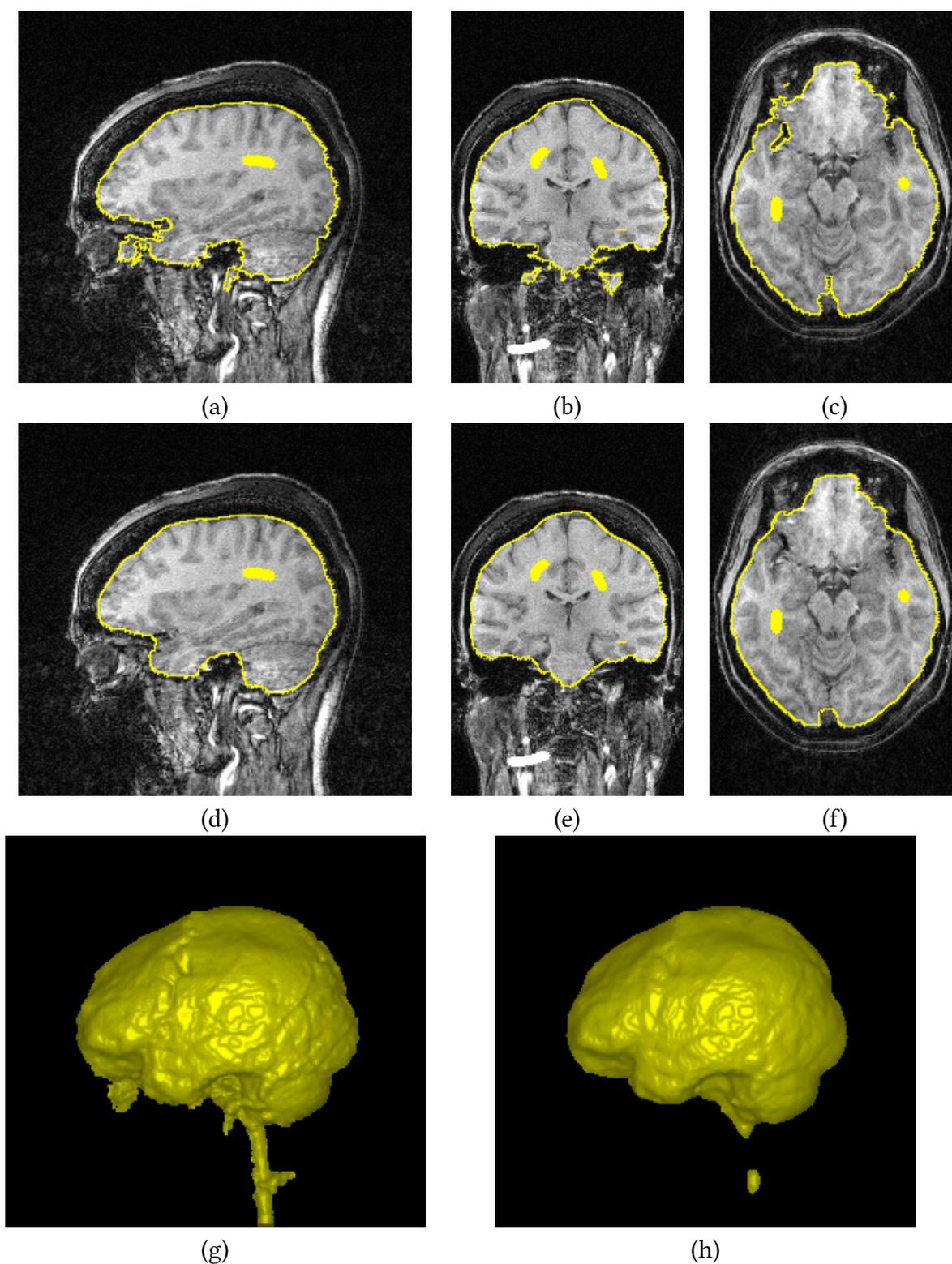


Figura 4.13: Exemplo #7 de segmentação 3D do cérebro em volume de ressonância magnética (RM) do crânio. (a-c) Resultado da OIFT nos planos sagital, coronal e axial. (d-f) Resultado da OIFT Relax. Orientada nos planos sagital, coronal e axial. (g-h) Renderizações 3D dos resultados da OIFT e OIFT Relax. Orientada, respectivamente.

4.3 Relaxed Deep Extreme Cut

Nesta seção apresentamos os resultados do método híbrido, que combina a segmentação por Deep Extreme Cut de imagens coloridas 2D, baseada na seleção dos quatro pontos extremos, com o procedimento de relaxação em grafos direcionados proposto (Algoritmo 3), tal como discutido na Seção 3.3. Para os experimentos, usamos o conjunto de dados GrabCut de 50 imagens naturais (Rother, Kolmogorov e Blake, 2004).

A taxa de erro é a porcentagem de pixels erroneamente classificados dentro das caixas delimitadoras. Em relação à medida de taxa de erro, Deep Extreme Cut (Maninis *et al.*, 2018) já demonstrou superar vários métodos da literatura, incluindo Deconvolution (Noh, Hong e Han, 2015), Deep GrabCut (Xu *et al.*, 2017), MILCut (Wu *et al.*, 2014), Box Prior (Lempitsky *et al.*, 2009), Extreme Clicking (Papadopoulos *et al.*, 2017), One Cut (Tang, Gorelick *et al.*, 2013) e Kernel cut (Tang, Ayed *et al.*, 2015). Nosso método híbrido tem resultados ainda melhores, aumentando os ganhos em relação a todas essas abordagens.

Além de calcular a taxa de erro, também calculamos a similaridade pelos coeficientes de Dice, Jaccard e GB entre o gabarito de segmentação e os resultados segmentados para os diferentes métodos. Para o método Grabcut (Rother, Kolmogorov e Blake, 2004), utilizamos sua implementação disponível no OpenCV em Python com diferentes números de iterações (5, 10, 20 e 40). Para Deep Extreme Cut (DEXTR) (Maninis *et al.*, 2018), usamos seus resultados pré-calculados, conforme disponíveis na sua página dos autores da web para o conjunto de dados GrabCut e calculamos os pontos extremos automaticamente a partir das imagens dos gabaritos. Com relação ao método proposto (DEXTR Relaxada), calculamos seus resultados para diferentes números de iterações N de relaxação (5, 10, 15, 20 e 30).

Os resultados são apresentados na Tabela 4.2. Exemplos de segmentação da base do Grabcut são apresentados nas Figuras 4.14, 4.15 e 4.16, mostrando a melhor aderência à borda do método híbrido proposto. O método proposto superou os demais métodos em todas medidas de avaliação consideradas. A melhor configuração do método proposto se deu para $N = 10$ iterações de relaxação. Porém, dado que a base do Grabcut é uma base heterogênea de dados, composta por objetos das mais diversas naturezas, observamos na prática que o valor ideal do parâmetro N varia de imagem para imagem.

Em (Oliveira, Caio L. Demario e P. A. Miranda, 2021), usamos uma estimativa de N adaptativa, a fim de melhor se ajustar a cada caso. As contribuições adicionais presentes em (Oliveira, Caio L. Demario e P. A. Miranda, 2021) geram valores ainda mais elevados de acurácia, alcançando um valor médio de Dice de 98,32% e taxa de erro de 1,31%, porém essas contribuições estão além do escopo do presente trabalho.

	Coef. Dice	Coef. Jaccard	Jac. Gen. (GB $\delta = 1.0$)	Taxa de erro
Grabcut (5)	91,01%	85,39%	79,34%	8,09%
Grabcut (10)	91,60%	86,29%	80,97%	7,40%
Grabcut (20)	91,61%	86,32%	80,98%	7,39%
Grabcut (40)	91,65%	86,38%	81,10%	7,34%
DEXTR	96,66%	94,36%	94,05%	2,30%
DEXTR Relaxada (5)	96,98%	94,97%	94,69%	2,03%
DEXTR Relaxada (10)	97,03%	95,06%	94,78%	2,00%
DEXTR Relaxada (15)	97,03%	95,05%	94,78%	2,00%
DEXTR Relaxada (20)	97,02%	95,03%	94,76%	2,01%
DEXTR Relaxada (30)	96,99%	94,98%	94,72%	2.03%

Tabela 4.2: Resultados experimentais na base do Grabcut.

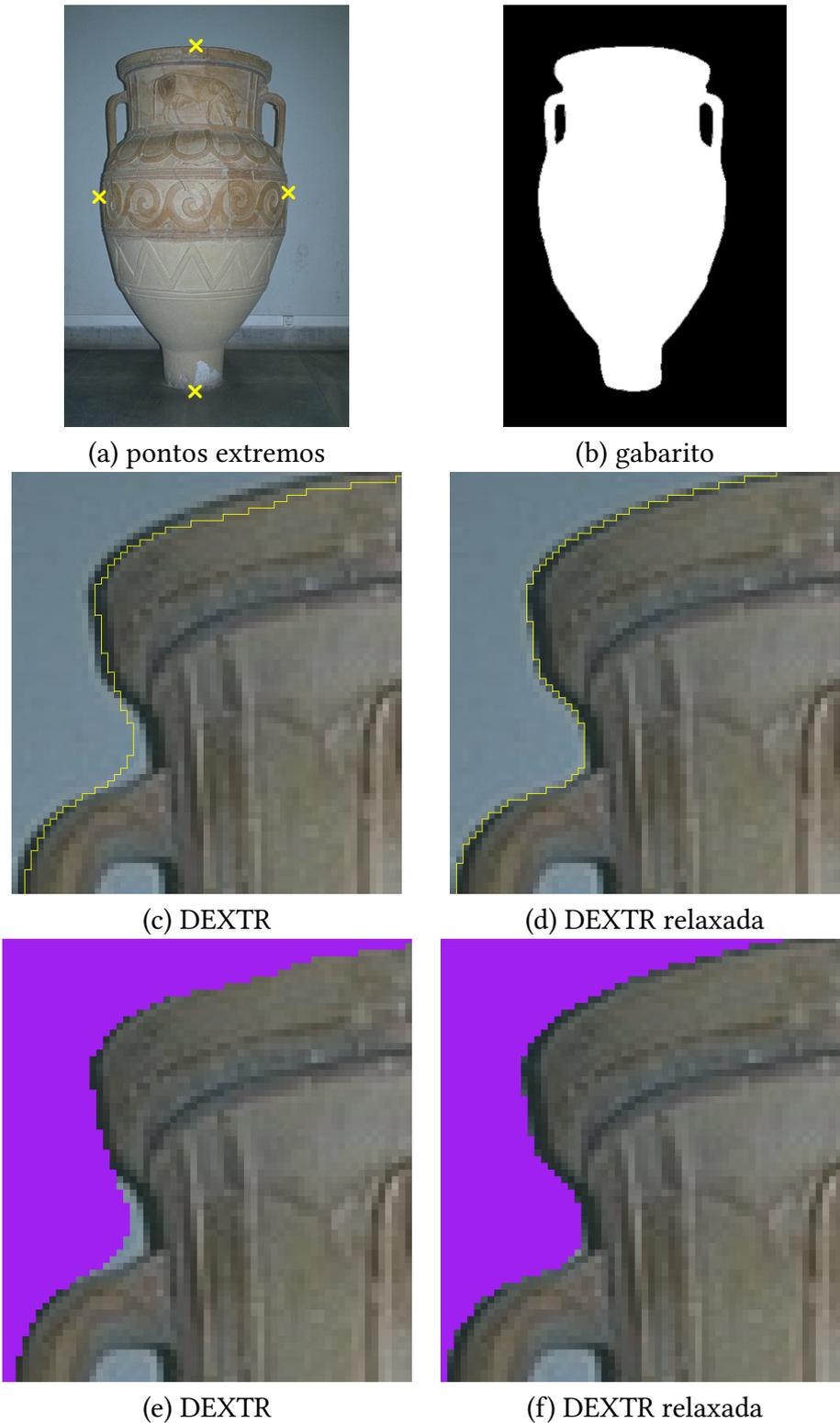


Figura 4.14: Exemplo de segmentação da imagem 227092.jpg da base do grabcut. (a) Imagem original com os quatro pontos extremos selecionados. (b) Gabarito de segmentação. (c-d) Borda de separação em região ampliada (Zoom) das segmentações por Deep Extreme Cut e sua versão com relaxação proposta, respectivamente. (e-f) As seleções de foreground correspondentes de c e d, respectivamente, com acurácias de Dice de 98,76% e 99,06%. Observe a melhor aderência à borda do método híbrido proposto.

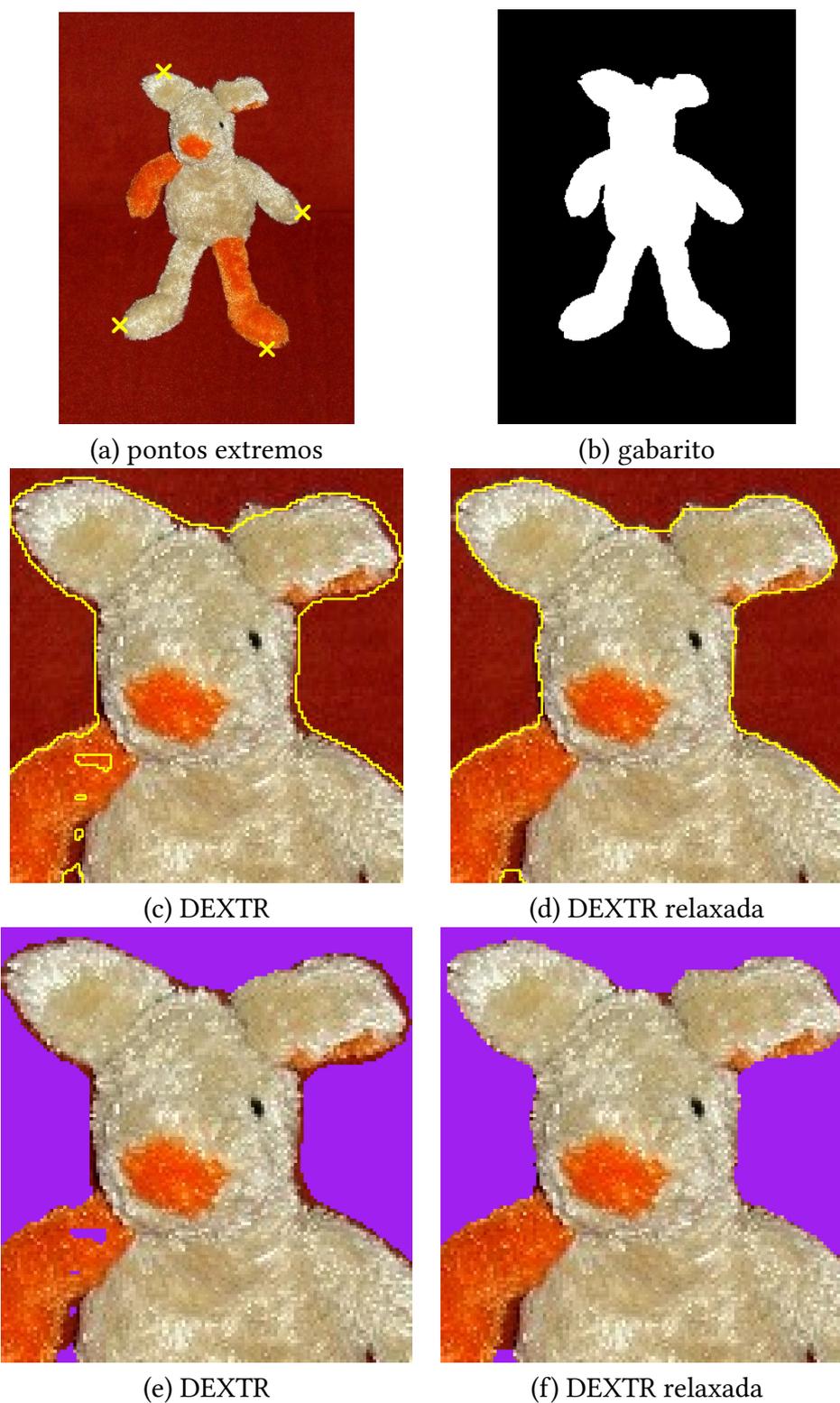


Figura 4.15: Exemplo de segmentação da imagem *teddy.jpg* da base do *grabcut*. (a) Imagem original com os quatro pontos extremos selecionados. (b) Gabarito de segmentação. (c-d) Borda de separação em região ampliada (Zoom) das segmentações por Deep Extreme Cut e sua versão com relaxação proposta, respectivamente. (e-f) As seleções de foreground correspondentes de c e d, respectivamente, com acurácias de Dice de 97,12% e 98,01%. Observe a melhor aderência à borda do método híbrido proposto.



Figura 4.16: Exemplo de segmentação da imagem *bool.jpg* da base do *grabcut*. (a) Imagem original com os quatro pontos extremos selecionados. (b) Gabarito de segmentação. (c-d) Borda de separação em região ampliada (Zoom) das segmentações por Deep Extreme Cut e sua versão com relaxação proposta, respectivamente. (e-f) As seleções de foreground correspondentes de c e d, respectivamente, com acurácias de Dice de 97,80% e 98,68%. Observe a melhor aderência à borda do método híbrido proposto.

Capítulo 5

Conclusão

Neste trabalho foram apresentados dois novos métodos híbridos baseados na extensão do *Random Walks* em grafos direcionados e foram apresentadas suas vantagens, atuando em alguns pontos fracos de importantes métodos da literatura, também sendo possível incorporar a restrição de polaridade de borda como conhecimento prévio dos objetos de interesse na imagem. No caso da OIFT, foi possível melhorar a regularidade de forma da segmentação e no caso do *Deep Extreme Cut* foi possível reduzir a sua deficiência na aderência à borda. Os resultados apresentados no Capítulo 4 mostram os experimentos realizados e as curvas de acurácia, exibindo a melhora de qualidade das segmentações utilizando os métodos híbridos em relação aos métodos puros.

Para adaptar os métodos propostos em uma nova aplicação, deve-se proceder com uma seleção adequada de seus parâmetros sensíveis por meio de uma avaliação empírica. Os parâmetros em questão são:

- $\alpha \in [-1, 1]$ na Equação 3.1, que controla a penalização para bordas com polaridade indesejada e realce para as bordas com polaridade desejada. Em aplicações em que o objeto de interesse apresenta transições bem definidas (claro para escuro ou escuro para claro em uma dada imagem de referência), devemos usar $|\alpha| \approx 1$, sendo $\alpha > 0$ usado para indicar transições na borda de claro para escuro e $\alpha < 0$ para indicar transições de escuro para claro.
- $\beta > 1$ na Equação 3.1, que realça as diferenças dos pesos, favorecendo uma maior robustez por parte da componente do Random Walks do método híbrido proposto, em relação ao posicionamento das sementes. Por outro lado, valores muito elevados de β podem reduzir a suavidade dos contornos gerados. Portanto, devemos buscar por um valor intermediário de equilíbrio que seja adequado para a aplicação.
- N que define a quantidade de relaxações.
- pesos κ_{st} na Equação 3.1, que medem a dissimilaridade não direcionada entre nós vizinhos s e t . Idealmente, deve-se configurar estes pesos de modo a termos valores elevados na borda do objeto de interesse e baixos no seu interior e exterior.
- sementes S_1 e S_0 , no caso do primeiro método proposto da OIFT Relaxada, indicando pixels pertencentes ao objeto e fundo, respectivamente.

Durante esta pesquisa, foram publicados dois artigos em conferências internacionais, o primeiro descrevendo a OIFT Relaxada (C. L. Demario e P. A. V. Miranda, 2019) e o segundo descrevendo o *Relaxed Deep Extreme Cut*, incluindo também outras melhorias não descritas nesta dissertação (Oliveira, Caio L. Demario e P. A. Miranda, 2021).

5.1 Trabalhos futuros

Como trabalhos futuros, pretendemos:

- Paralelizar o método de relaxação proposto.
- Estender os resultados anteriores para a segmentação hierárquica de múltiplos objetos, tomando como base o algoritmo da *Hierarchical Layered Oriented Image Foresting Transform* (HLOIFT) (Leon e P. Miranda, 2017).
- Estender o método *Relaxed Deep Extreme Cut* para atuar em imagens multidimensionais.
- Aplicar o *Relaxed Deep Extreme Cut* em alguma tarefa automática de segmentação, via a seleção automática de seus 4 pontos extremos.

Referências

- [Allène *et al.* 2007] C. Allène *et al.* “Some Links Between Min-cuts, Optimal Spanning Forests and Watersheds”. Em: *Proc. of the 8th Int. Symposium on Mathematical Morphology*. Rio de Janeiro, Brazil, 2007, pp. 253–264.
- [Audigier e Lotufo 2005] R. Audigier e R.A. Lotufo. “The Tie-zone Watershed: Definition, Algorithm and Applications”. Em: *IEEE Intl. Conf. on Image Processing*. 2005, pp. 654–657.
- [Bai e Sapiro 2007] X. Bai e G. Sapiro. “Distance Cut: Interactive Segmentation and Matting of Images and Videos”. Em: *Proc. of the IEEE Intl. Conf. on Image Processing*. Vol. 2. 2007, pp. II - 249-II –252.
- [Bejar, F. Cappabianco e P. Miranda 2017] H.H.C. Bejar, F.A.M. Cappabianco e P.A.V. Miranda. “Efficient Image Segmentation in Graphs with Localized Curvilinear Features”. Em: *Image Analysis and Processing*. Cham: Springer Intl. Publishing, 2017, pp. 718–728. ISBN: 978-3-319-68560-1.
- [Bejar e P. Miranda 2015] H.H.C. Bejar e P.A.V. Miranda. “Oriented relative fuzzy connectedness: theory, algorithms, and its applications in hybrid image segmentation methods”. Em: *EURASIP Jnl. on Image and Video Processing* 2015.1 (jul. de 2015), p. 21. ISSN: 1687-5281.
- [Y. Boykov e Funka-Lea 2006] Y. Boykov e G. Funka-Lea. “Graph Cuts and Efficient N-D Image Segmentation”. Em: *Intl. Journal of Computer Vision* 70.2 (2006), pp. 109–131. ISSN: 0920-5691.
- [Y.Y. Boykov e Jolly 2001] Y.Y. Boykov e M.-P. Jolly. “Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images”. Em: *International Conference on Computer Vision (ICCV)*. Vol. 1. 2001, pp. 105–112.
- [Braz *et al.* 2019] C.M. Braz *et al.* “Graph-based Segmentation with Local Band Constraints”. Em: *21st Intl. Conf. on Discrete Geometry for Computer Imagery*. accepted. Paris, FR, mar. de 2019.
- [C.Couprie *et al.* 2010] C.Couprie *et al.* “Power Watersheds: A Unifying Graph-Based Optimization Framework”. Em: *IEEE Trans. on PAMI* 99 (2010). ISSN: 0162-8828.
- [F. A. M. Cappabianco *et al.* 2019] F. A. M. Cappabianco *et al.* “A General and Balanced Region-Based Metric for Evaluating Medical Image Segmentation Algorithms”. Em: *2019 IEEE International Conference on Image Processing (ICIP)*. Set. de 2019, pp. 1525–1529. DOI: [10.1109/ICIP.2019.8803083](https://doi.org/10.1109/ICIP.2019.8803083).
- [Ciesielski, Falcão e P. Miranda 2018] K.C. Ciesielski, A.X. Falcão e P.A.V. Miranda. “Path-Value Functions for Which Dijkstra’s Algorithm Returns Optimal Mapping”. Em: *Journal of Mathematical Imaging and Vision* 60.7 (2018), pp. 1025–1036. ISSN: 1573-7683.

- [Ciesielski, P. Miranda *et al.* 2013] K.C. Ciesielski, P.A.V. Miranda *et al.* “Joint graph cut and relative fuzzy connectedness image segmentation algorithm”. Em: *Medical Image Analysis* 17.8 (2013), pp. 1046–1057.
- [Ciesielski, J.K. Udupa *et al.* 2012] K.C. Ciesielski, J.K. Udupa *et al.* “A unifying graph-cut image segmentation framework: algorithms it encompasses and equivalences among them”. Em: *Proc.SPIE*. Vol. 8314. 2012, p. 12.
- [Cousty *et al.* 2010] J. Cousty *et al.* “Watershed Cuts: Thinnings, Shortest Path Forests, and Topological Watersheds”. Em: *Trans. on Pattern Analysis and Machine Intelligence* 32 (2010), pp. 925–939. ISSN: 0162-8828.
- [Delong e Y. Boykov 2009] A. Delong e Y. Boykov. “Globally optimal segmentation of multi-region objects”. Em: *Computer Vision, 12th Intl. Conf. on. IEEE*. 2009, pp. 285–292.
- [C. L. Demario e P. A. V. Miranda 2019] C. L. Demario e P. A. V. Miranda. “Relaxed Oriented Image Foresting Transform for Seeded Image Segmentation”. Em: *2019 IEEE International Conference on Image Processing (ICIP)*. Set. de 2019, pp. 1520–1524. DOI: [10.1109/ICIP.2019.8803080](https://doi.org/10.1109/ICIP.2019.8803080).
- [Dice 1945] L.R. Dice. “Measures of The Amount of Ecologic Association Between Species”. Em: *Ecology* 26 (1945), pp. 297–302.
- [Dijkstra 1959] E.W. Dijkstra. “A Note on Two Problems in Connexion with Graphs”. Em: *Numerische Mathematik* 1.1 (1959), pp. 269–271.
- [Falcão, Stolfi e Lotufo 2004] A.X. Falcão, J. Stolfi e R.A. Lotufo. “The Image Foresting Transform: Theory, Algorithms, and Applications”. Em: *Trans. on PAMI* 26.1 (2004), pp. 19–29.
- [Grady 2006] L. Grady. “Random Walks for Image Segmentation”. Em: *Trans. on PAMI* 28.11 (2006), pp. 1768–1783. ISSN: 0162-8828.
- [Gulshan *et al.* 2010] V. Gulshan *et al.* “Geodesic Star Convexity for Interactive Image Segmentation”. Em: *Proc. of Computer Vision and Pattern Recognition*. San Francisco, CA, 2010, pp. 3129–3136.
- [He *et al.* 2016] K. He *et al.* “Deep Residual Learning for Image Recognition”. Em: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Jun. de 2016, pp. 770–778. DOI: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [J.Liu 2006] J.Liu. “Synergistic hybrid image segmentation: combining model and image-based strategies”. advisor: J.K. Udupa. Tese de dout. Dept. of Bioengineering, Univ. of Pennsylvania, nov. de 2006.
- [Jaccard 1901] Paul Jaccard. “Etude de la distribution florale dans une portion des Alpes et du Jura”. Em: *Bulletin de la Societe Vaudoise des Sciences Naturelles* 37 (jan. de 1901), pp. 547–579. DOI: [10.5169/seals-266450](https://doi.org/10.5169/seals-266450).
- [Lempitsky *et al.* 2009] V. Lempitsky *et al.* “Image segmentation with a bounding box prior”. Em: *2009 IEEE 12th International Conference on Computer Vision*. 2009, pp. 277–284.
- [Leon e P. Miranda 2017] L.M.C. Leon e P.A.V. Miranda. “Multi-Object Segmentation by Hierarchical Layered Oriented Image Foresting Transform”. Em: *Conference on Graphics, Patterns and Images (SIBGRAPI)*. Out. de 2017, pp. 79–86.
- [Madabhushi e J.K. Udupa 2005] A. Madabhushi e J.K. Udupa. “Interplay Between Intensity Standardization and Inhomogeneity Correction in MR Image Processing”. Em: *IEEE Transactions on Medical Imaging* 24.5 (2005), pp. 561–576.

- [Malmberg *et al.* 2010] F. Malmberg *et al.* “Relaxed image foresting transforms for interactive volume image segmentation”. Em: *Proc.SPIE*. Vol. 7623. 2010, pp. 7623 - 7623 -11.
- [Maninis *et al.* 2018] K.-K. Maninis *et al.* “Deep extreme cut: From extreme points to object segmentation”. Em: *In IEEE Conf. on Computer Vision and Pattern Recognition*. 2018, pp. 616–625.
- [Mansilla e P. Miranda 2013a] L.A.C. Mansilla e P.A.V. Miranda. “Image Segmentation by Oriented Image Foresting Transform with Geodesic Star Convexity”. Em: *Computer Analysis of Images and Patterns*. Vol. 8047. Ago. de 2013, pp. 572–579.
- [Mansilla e P. Miranda 2013b] L.A.C. Mansilla e P.A.V. Miranda. “Image Segmentation by Oriented Image Foresting Transform: Handling Ties and Colored Images”. Em: *18th Intl. Conf. on Digital Signal Processing*. Santorini, GR, jul. de 2013, pp. 1–6.
- [P. Miranda e Falcão 2009] P.A.V. Miranda e A.X. Falcão. “Links Between Image Segmentation Based on Optimum-Path Forest and Minimum Cut in Graph”. Em: *Journal of Mathematical Imaging and Vision* 35.2 (2009), pp. 128–142. ISSN: 1573-7683.
- [P. Miranda e Falcão 2011] P.A.V. Miranda e A.X. Falcão. “Elucidating the relations among seeded image segmentation methods and their possible extensions”. Em: *XXIV Conference on Graphics, Patterns and Images*. Maceió, AL, ago. de 2011.
- [P. Miranda e Mansilla 2014] P.A.V. Miranda e L.A.C. Mansilla. “Oriented Image Foresting Transform Segmentation by Seed Competition”. Em: *Image Processing, IEEE Trans. on* 23.1 (jan. de 2014), pp. 389–398. ISSN: 1057-7149.
- [Paulo A. V. de Miranda 2020] Paulo A. V. de Miranda. *MAC6903 - 2s20 - Processamento de Imagens usando Grafos*. Jul. de 2020. URL: http://www.vision.ime.usp.br/~pmiranda/mac6903_2s20/page/mac6903_2s20.html.
- [Moya 2015] Nikolas Moya. “Interactive segmentation of multiple 3D objects in medical images by optimum graph cuts”. Dissertação (mestrado) disponível em: <<http://repositorio.unicamp.br/jspui/handle/REPOSIP/275554>>. Diss. de mest. Universidade Estadual de Campinas, Instituto de Computação, 2015.
- [Noh, Hong e Han 2015] H. Noh, S. Hong e B. Han. “Learning Deconvolution Network for Semantic Segmentation”. Em: *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015, pp. 1520–1528.
- [Oliveira, Caio L. Demario e P. A. Miranda 2021] Débora E. C. Oliveira, Caio L. Demario e Paulo A.V. Miranda. “Image Segmentation by Relaxed Deep Extreme Cut with Connected Extreme Points”. Em: *IAPR International Conference on Discrete Geometry and Mathematical Morphology (DGMM)*. submetido, em revisão. 2021.
- [Papadopoulos *et al.* 2017] D. P. Papadopoulos *et al.* “Extreme Clicking for Efficient Object Annotation”. Em: *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017, pp. 4940–4949.
- [Price, Morse e Cohen 2010] B. L. Price, B. Morse e S. Cohen. “Geodesic graph cut for interactive image segmentation”. Em: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Jun. de 2010, pp. 3161–3168. DOI: [10.1109/CVPR.2010.5540079](https://doi.org/10.1109/CVPR.2010.5540079).
- [Protiere e Sapiro 2007] A. Protiere e G. Sapiro. “Interactive Image Segmentation via Adaptive Weighted Distances”. Em: *IEEE Transactions on Image Processing* 16.4 (abr. de 2007), pp. 1046–1057.

- [Rother, Kolmogorov e Blake 2004] C. Rother, V. Kolmogorov e A. Blake. “GrabCut: Interactive Foreground Extraction using Iterated Graph Cuts”. Em: *ACM Transactions on Graphics* 23.3 (2004), pp. 309–314. ISSN: 0730-0301.
- [Singaraju, Grady e Vidal 2008] D. Singaraju, L. Grady e R. Vidal. “Interactive image segmentation via minimization of quadratic energies on directed graphs”. Em: *Intl. Conf. on Computer Vision and Pattern Recognition*. Jun. de 2008, pp. 1–8.
- [Sinop e L.Grady 2007] A.K. Sinop e L.Grady. “A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm”. Em: *Proc. of the 11th Intl. Conf. on Computer Vision (ICCV)*. IEEE, 2007, pp. 1–8.
- [Spina, de Miranda e Xavier Falcão 2014] T. V. Spina, P. A. V. de Miranda e A. Xavier Falcão. “Hybrid Approaches for Interactive Image Segmentation Using the Live Markers Paradigm”. Em: *IEEE Transactions on Image Processing* 23.12 (dez. de 2014), pp. 5756–5769. DOI: [10.1109/TIP.2014.2367319](https://doi.org/10.1109/TIP.2014.2367319).
- [Strand et al. 2013] R. Strand et al. “The Minimum Barrier Distance”. Em: *Computer Vision and Image Understanding* 117 (2013), pp. 429–437.
- [Tang, Ayed et al. 2015] M. Tang, I. B. Ayed et al. “Secrets of GrabCut and Kernel K-Means”. Em: *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015, pp. 1555–1563.
- [Tang, Gorelick et al. 2013] M. Tang, L. Gorelick et al. “GrabCut in One Cut”. Em: *2013 IEEE International Conference on Computer Vision*. 2013, pp. 1769–1776.
- [Tavares, Bejar e P. Miranda 2017] A.C.M. Tavares, H.H.C. Bejar e P.A.V. Miranda. “Seed Robustness of Oriented Image Foresting Transform: Core Computation and the Robustness Coefficient”. Em: *Mathematical Morphology and Its Applications to Signal and Image Processing*. Springer, 2017, pp. 119–130. ISBN: 978-3-319-57240-6.
- [JK Udupa et al. 2014] JK Udupa et al. “Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images”. Em: *Med Image Anal* 18.5 (jul. de 2014), pp. 752–71. DOI: [10.1016/j.media.2014.04.003](https://doi.org/10.1016/j.media.2014.04.003).
- [Vezhnevets e Konouchine 2005] V. Vezhnevets e V. Konouchine. “GrowCut- Interactive Multi-Label N-D Image Segmentation By Cellular Automata”. Em: *Proc. Graphicon*. 2005, pp. 150–156.
- [Wu et al. 2014] J. Wu et al. “MILCut: A Sweeping Line Multiple Instance Learning Paradigm for Interactive Image Segmentation”. Em: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 256–263.
- [Xu et al. 2017] Ning Xu et al. “Deep GrabCut for Object Selection”. Em: *Proceedings of the British Machine Vision Conference (BMVC)*. Ed. por Gabriel Brostow Tae-Kyun Kim Stefanos Zafeiriou e Krystian Mikolajczyk. BMVA Press, set. de 2017, pp. 182.1–182.12. ISBN: 1-901725-60-X.