

KARLA BARBOSA DE FREITAS

Modelos Variacionais em Processamento de Imagens - Formulação Primal e Dual



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA

2011

KARLA BARBOSA DE FREITAS

Modelos Variacionais em Processamento de Imagens - Formulação Primal e Dual

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Uberlândia, como parte dos requisitos para obtenção do título de **MESTRE EM MATEMÁTICA**.

Área de Concentração: Matemática.

Linha de Pesquisa: Análise Numérica.

Orientadora: Profa. Dra. Celia A. Zorzo Barcelos.

UBERLÂNDIA - MG

2011

(Ficha Catalográfica elaborada pela biblioteca da UFU)
(anexar ou “escanear”)



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Av. João Naves de Ávila, 2121, Bloco 1F, Sala 1F 152
Campus Santa Mônica, Uberlândia - MG, CEP 38400-902

ALUNO(A): Karla Barbosa de Freitas.

NÚMERO DE MATRÍCULA: 100089.

ÁREA DE CONCENTRAÇÃO: Matemática.

LINHA DE PESQUISA: Análise Numérica.

PÓS-GRADUAÇÃO EM MATEMÁTICA: Nível Mestrado.

TÍTULO DA DISSERTAÇÃO: Modelos Variacionais em Processamento de Imagens - Formulação Primal e Dual.

ORIENTADORA: Profa. Dra. Celia A. Zorzo Barcelos.

Esta dissertação foi **APROVADA** em reunião pública realizada na Sala Multiuso da Faculdade de Matemática, Bloco 1F, Campus Santa Mônica, em 23 de Agosto de 2011, às 15h00min, pela seguinte Banca Examinadora:

NOME

ASSINATURA

Profa. Dra. Celia A. Zorzo Barcelos

UFU - Universidade Federal de Uberlândia

Prof. Dr. Suetônio de Almeida Meira

UNESP - Universidade Estadual Paulista

Profa. Dra. Ana Maria Amarillo Bertone

UFU - Universidade Federal de Uberlândia

Uberlândia-MG, 23 de Agosto de 2011.

Dedicatória

Dedico este trabalho aos meus pais Carlos e Verônica; pelo esforço, dedicação e compreensão, durante todos os momentos da minha vida.

Ao meu namorado Samir, pelo apoio, compreensão e incentivo durante todo tempo de realização deste trabalho.

Aos meu avós maternos Anna Maria (in memorian) e José Pedro (in memorian), e avós paternos Maria Ubaldina (in memorian) e Oswaldo Borges (in memorian) que infelizmente não puderam estar presentes para viver comigo este momento.

Aos todos os meus amigos que de uma forma ou de outra contribuíram e sempre me incentivaram.

Agradecimentos

Agradeço:

- Primeiramente a Deus.

- Aos colegas do mestrado Carlos Tognon, Daniela Portes, Flávio Fernandes, Lilyane Figueiredo, Thiago Rodrigo e Túlio Guimarães que muito contribuíram para meu crescimento enquanto matemático e mais ainda como pessoa.

- Ao colega Vinícius Ruela Perreira Borges pelo incentivo e apoio durante todo o Mestrado e que muito contribuiu neste trabalho, realizando a maior parte das análises computacionais aqui apresentadas.

- Aos funcionários da FAMAT pelo apoio e incentivo, em especial à Magda Laine e Sandra Valéria.

- Aos professores Suetônio de Almeida Meira e Ana Maria Amarillo Bertone por terem aceito o convite para participarem da banca examinadora e, de mesma forma, agradeço aos professores suplentes, José Roberto Nogueira e César Guilherme de Almeida.

- Aos docentes do Programa de Mestrado-FAMAT que muito contribuíram para a realização deste trabalho e com muito carinho ao professor, Edson Agustini pelas palavras amigas e incentivadoras nos momentos difíceis.

- E de uma maneira muito especial agradeço a minha orientadora, Celia Aparecida Zorzo Barcelos, pela educação, pela paciência e pelo conhecimento transmitido na realização deste trabalho e também a Marcos Aurélio Batista pelas dúvidas tiradas e sugestões dadas que muito contribuíram para a realização deste trabalho.

- A agência financiadora FAPEMIG pelo apoio dado ao longo do curso. Se esqueci de algumas pessoas que de certa forma contribuíram para que este momento fosse alcançado, peço desculpas e agradeço a todos.

Muito obrigado!

FREITAS, K. B. *Modelos Variacionais em Processamento de Imagens - Formulação Primal e Dual*. 2011. 125 p. Dissertação de Mestrado, Universidade Federal de Uberlândia, Uberlândia-MG.

Resumo

Neste trabalho apresentamos alguns problemas de processamento de imagens cujas formulações são variacionais. Para exemplificar estas formulações consideramos o modelo proposto pelos autores Rudin, Osher e Fatemi (ROF) para o problema de remoção de ruídos. Para um melhor entendimento do problema alguns conceitos do Cálculo Variacional, em especial as equações de Euler-Lagrange, Variação Total (TV) em imagens e alguns problemas de processamento de imagens baseados em TV são abordados. Estudaremos a formulação Primal e Dual de um modelo variacional, a equivalência entre as formulações, bem como métodos de resolução. Daremos ainda uma formulação Primal-Dual e o respectivo algoritmo numérico. Aplicações em problema de remoção de ruídos e segmentação de imagens serão apresentados para exemplificar a eficácia da metodologia.

Palavras-chave: restauração de imagens, espaços BV, equação de Euler-Lagrange, Variação Total, formulação Primal e Dual.

FREITAS, K. B. *Variational Models in Image Processing - Primal and Dual Formulation*. 2011. 125 p. M. Sc. Dissertation, Federal University of Uberlândia, Uberlândia-MG.

Abstract

This work presents some problems of image processing whose formulations are variational. To illustrate these formulations, we consider the model proposed by Rudin, Osher and Fatemi (ROF), in which deals with image denoising. For a better understanding, we explore some variational calculus concepts, as Euler-Lagrange equations, Total Variation (TV) regularizing terms and image processing methods based on TV. We discuss primal and dual formulations of a variational model, the equivalence between them and its resolution methods. Furthermore, we study a Primal-Dual formulation and its numerical algorithm. Applications related to denoising and image segmentation are presented to illustrate the effectiveness of the methodology.

Keywords: image restoration, BV spaces, Euler-Lagrange equation, Total Variation, Primal and Dual formulation.

Lista de Figuras

3.1	Problema de Eliminação de Ruído 1. Imagens retiradas de [13]	39
3.2	Problema de Eliminação de Ruído 2. Imagens retiradas de [13]	39
3.3	Forças agindo no contorno: (a) Força de suavização (b) Força estatística em um ponto de fronteira (c) Força estatística em um ponto de junção de regiões. . . .	43
3.4	Imagens sem fronteiras definidas.	44
3.5	Problema de Segmentação 1. Imagens retiradas de [27].	45
3.6	Problema de Segmentação 2. Imagens retiradas de [27].	46
3.7	Problema deblurring. Imagens retiradas de [7].	47
3.8	Problema de retoque Digital 1. Imagens retiradas de [10].	48
3.9	Problema de retoque Digital 2. Imagens retiradas de [10].	49
3.10	Problema de retoque Digital 3. Imagens retiradas de [10].	49
3.11	Problema zoom 1. Imagens retiradas de [6].	51
3.12	Problema zoom 1. Imagens retiradas de [7].	51
3.13	Problema de decomposição em geometria e textura 1. Imagens retiradas de [34].	53
3.14	Problema de decomposição em geometria e textura 2. Imagens retiradas de [34].	53
3.15	Problema de decomposição em geometria e textura 3. Imagens retiradas de [34].	54
4.1	Conjunto de pontos utilizado na convolução.	58
6.1	Teste realizado com o modelo <i>Competição entre Regiões Fuzzy</i> . Imagens retiradas de [32].	86
6.2	Teste realizado com o modelo <i>Competição entre Regiões Fuzzy</i> . Imagens retiradas de [32].	87
6.3	Teste realizado com o modelo <i>Competição entre Regiões Fuzzy</i> . Imagens retiradas de [32].	88
6.4	Problema de remoção de ruídos. Imagem retirada de [6].	90
6.5	Problema de remoção de ruídos 1 ($\sigma = 12$). Imagens retiradas de [6].	91

6.6	Problema de remoção de ruídos 2 ($\sigma = 25$). Imagens retiradas de [6].	91
7.1	Problema de remoção de ruídos. Imagens retiradas de [38].	102
7.2	Problema de remoção de ruídos com diferentes critérios de parada. Imagens retiradas de [38].	104

Lista de Tabelas

7.1	Iterações e tempo gasto para o problema 1, 128×128 , $\lambda = 0,0415$. Tabela retirada de [38].	103
7.2	Iterações e tempo gasto para o problema 2, 256×256 , $\lambda = 0,0415$. Tabela retirada de [38].	103
7.3	Iterações e tempo gasto para o problema 3, 512×512 , $\lambda = 0,0415$. Tabela retirada de [38].	103

Sumário

Resumo	viii
Abstract	ix
Lista de Figuras	xi
Lista de Tabelas	xii
Introdução	1
1 Preliminares e Definições Gerais	4
1.1 Conceitos Básicos	4
1.2 Espaço das Funções Contínuas $C^k(\Omega)$	7
1.3 Espaços L^p	9
1.4 Espaços de Sobolev, $W^{k,p}(\Omega)$	11
1.5 Espaço das funções de variação limitada, $BV(\Omega)$	14
1.6 Condições de Karush-Kuhn-Tucker (KKT)	15
2 Cálculo Variacional	16
2.1 Preliminares	16
2.2 Equação de Euler-Lagrange	21
2.2.1 Primeira variação: equação de Euler-Lagrange	22
2.2.2 Segunda variação	26
2.3 Condições de contorno	27
2.4 Existência de minimizadores	29
2.4.1 Coercividade	29
2.4.2 Semi-continuidade inferior	30
2.4.3 Convexidade	32

3	Variação Total em Imagens	35
3.1	Introdução	35
3.2	Remoção de ruídos	36
3.3	Segmentação de imagens	40
3.3.1	Competição entre Regiões	41
3.3.2	Competição entre Regiões Fuzzy	43
3.4	Outros problemas de processamento de imagens baseados em variação total . . .	46
3.4.1	Deblurring	46
3.4.2	Retoque Digital	48
3.4.3	Zoom	50
3.4.4	Decomposição de uma imagem em Geometria e Textura	52
4	Formulação Primal e Dual de Problemas Variacionais	55
4.1	Introdução	55
4.2	Formulação Primal - Método de Resolução	55
4.2.1	Discretização da equação do fluxo através de diferenças finitas	58
4.3	Equivalência entre as Formulações Primal e Dual	60
5	Algoritmo de Minimização da Variação Total na Formulação Dual	64
5.1	Formulação Dual na Forma Discreta	64
5.1.1	A prova de $\langle -div p, u \rangle_X = \langle p, \nabla u \rangle_Y$	66
5.2	O Algoritmo	70
6	Aplicação do Algoritmo de Minimização	77
6.1	Introdução	77
6.2	Aplicação em Segmentação de Imagens - Modelo de Competição entre Regiões Fuzzy	77
6.2.1	Minimização do Funcional (3.11)	78
6.2.2	Resultados	85
6.3	Aplicação em Eliminação de Ruídos - Modelo ROF	89
7	Sistema Primal-Dual	92
7.1	Introdução	92
7.2	Sistema Primal-Dual	93
7.3	Algoritmo	96

7.4	Conecções Teóricas	99
7.5	Resultados e Comparações	101
8	Conclusão	105
	Referências Bibliográficas	106

Introdução

Modelos de restauração de imagens com base na Variação Total (TV) tem se tornado muito populares desde sua introdução nos trabalhos [28] e [30]. Este é um dos principais problemas no processamento de imagem digital e as técnicas variacionais têm sido bastante utilizadas na formulação de tais modelos. Em geral, estas formulações apresentam dois termos principais: o termo de difusão e o termo de fidelidade. Para melhor entendimento e resolução de tais formulações abordaremos alguns aspectos do Cálculo Variacional, principalmente as Equações de Euler-Lagrange, e alguns problemas de Variação Total em Imagens. Para exemplificar tal formulação, consideraremos um dos modelos mais usados em processamento de imagens, que é devido a Rudin, Osher e Fatemi (ROF) [30], e que consiste em encontrar a solução de

$$\min_u \int_{\Omega} J(u) + \frac{\|u - I\|^2}{2\lambda}, \quad (1)$$

onde u representa a imagem original, I a imagem observada, dada por $I = u + \eta$, η o ruído, λ uma constante positiva e $J(u)$ a variação total da imagem u no seu domínio Ω . Tal modelo, apesar de ser formulado para resolver o problema de remoção de ruídos, pode ser estendido para outros problemas de restauração de imagens, como por exemplo, deblurring, retoque digital, zoom e segmentação de imagens.

Ao longo dos anos, desde o surgimento do modelo ROF, vários algoritmos foram desenvolvidos para resolver ou a formulação primal ou a formulação dual deste modelo. Faremos agora, um breve histórico de alguns desses modelos.

No trabalho original de Rudin, Osher e Fatemi, os autores propuseram resolver a equação de Euler-Lagrange associada ao problema (1) por um método conhecido como “marcha no tempo”. A desvantagem apresentada por este método é o fato de ser muito lento. Em 1996, Vogel e Oman, em seu trabalho [36], propuseram resolver a mesma equação de Euler-Lagrange, porém, agora, através do método de iteração do ponto fixo. Este método requer resolver um sistema

linear a cada iteração, e é mais rápido que o método proposto anteriormente.

A idéia da dualidade e do sistema Primal-Dual foi introduzida por Chan, Golub e Mulet em [9], onde os autores resolveram o sistema Primal-Dual através do método de Newton. Já o problema dual de ROF foi abordado por Chambolle em [6], onde propôs um algoritmo semi-implícito do tipo gradiente descendente com base em algumas observações sobre os multiplicadores de Lagrange.

Depois destes, podemos ainda ressaltar o algoritmo proposto por Goldfarb e Yin detalhado em [19], o método proposto por Wang, Yin e Zhang detalhado em [40] e o modelo de Goldstein e Osher dado em [20].

Neste trabalho, vamos abordar três formas de resolver problemas variacionais de processamento de imagens: resolução na formulação Primal, na formulação Dual e através do sistema Primal-Dual. Apresentaremos a formulação Primal e Dual do modelo ROF, abordaremos um algoritmo proposto por Chambolle para a resolução do problema (1) em sua formulação Dual, bem como sua aplicação na solução de problemas de segmentação de imagens com base no modelo de Competição entre Regiões e também de problemas de remoção de ruídos com base no modelo ROF. Para finalizar, apresentaremos um método que combina as duas formulações, Primal e Dual, tirando vantagens de ambas, como também apresentaremos resultados experimentais para ilustração e comparações com outros métodos existentes.

Este trabalho está dividido em Capítulos do seguinte modo:

No Capítulo 1, serão apresentados alguns conceitos, definições e resultados referentes aos espaços $C^k(\Omega)$, $L^p(\Omega)$, $W^{k,p}(\Omega)$ e $BV(\Omega)$.

No Capítulo 2, será apresentado um estudo sobre Métodos Variacionais para problemas de valor de contorno, juntamente com a equação de Euler-Lagrange para o caso n-dimensional.

No Capítulo 3, será apresentado os trabalhos pioneiros de Variação Total em imagens, realizados pelos autores Rudin, Osher e Fatemi em [30], cujo modelo é proposto para resolver o problema de remoção de ruídos, e Mumford e Shah em [28], proposto para o problema de segmentação de imagens. Apresentamos também outros problemas de processamento de imagens cujas formulações são baseadas em Variação Total.

No Capítulo 4, será estudado as formulações do modelo ROF, dando um método de resolução na formulação primal, a equivalência entre as formulações primal e dual, e um breve estudo sobre a discretização da equação do fluxo utilizada na resolução de ambas formulações.

No Capítulo 5, será estudado o algoritmo proposto por Chambolle em [6] para resolver o problema (1) na sua formulação Dual. Será apresentado a definição de Variação Total e do operador divergente no conjunto discreto, além da prova da convergência do algoritmo.

No Capítulo 6, será apresentado a aplicação do algoritmo proposto por Chambolle no problema de segmentação de imagens e eliminação de ruídos. Especificamente abordaremos a aplicação no modelo de Competição entre Regiões Fuzzy [27] e modelo ROF [30].

No Capítulo 7, será apresentado uma proposta de solução de problemas variacionais usando a formulação Primal-Dual proposto por Zhu e Chan em [38]. Para exemplificar será usado o modelo ROF. Tal formulação tem como motivação sanar as dificuldades de ambas formulações quando tratadas individualmente, de modo que uma formulação sane a dificuldade da outra. Será também apresentado resultados comparativos.

Karla Barbosa de Freitas

Uberlândia-MG, 23 de Agosto de 2011.

Capítulo 1

Preliminares e Definições Gerais

Neste capítulo iremos apresentar os pré-requisitos necessários para a compreensão dos capítulos seguintes, como também, algumas definições, teoremas e notações a serem utilizadas durante este trabalho.

1.1 Conceitos Básicos

Definição 1.1 *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto e $x \in \Omega$, $x = (x_1, \dots, x_n)$. Define-se produto interno e norma no $\mathbb{R}^n$ por:*

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i \quad , \quad x, y \in \mathbb{R}^n \quad (1.1)$$

e

$$|x| = (x \cdot x)^{\frac{1}{2}} \quad , \quad x \in \mathbb{R}^n. \quad (1.2)$$

Definição 1.2 *Seja $\Omega \subset \mathbb{R}^n$. Uma função $f : \Omega \rightarrow \mathbb{R}^n$ é dita ser Lipschitziana se existir uma constante C tal que:*

$$|f(x) - f(y)| \leq C|x - y|,$$

para todo $x, y \in \Omega$.

Definição 1.3 Definimos por $H(x)$ a **função Heaviside** (função degrau unitário) determinada por:

$$H(x) = \begin{cases} 0, & \text{se } x < 0, \\ 1, & \text{se } x \geq 0. \end{cases}$$

A seguir, enunciaremos dois importantes resultados da Teoria de Integração de Lebesgue.

Teorema 1.1 (Lema de Fatou 7.20 de [21]) Suponha que $\{f_k\}_{k=1}^{\infty}$ são não negativas e mensuráveis. Então:

$$\int_{\mathbb{R}^n} \liminf_{k \rightarrow \infty} f_k(x) \, dx \leq \liminf_{k \rightarrow \infty} \int_{\mathbb{R}^n} f_k(x) \, dx.$$

Teorema 1.2 (Teorema de Lebesgue da Convergência Dominada 7.22 de [21]) Suponha que $\{f_k\}_{k=1}^{\infty}$ sejam integráveis e $f_k \rightarrow f$ em quase todo ponto (q.t.p). Suponha também que $|f_k| \leq g$ em quase toda parte, para alguma função integrável g , então f é integrável e:

$$\int_{\mathbb{R}^n} f_k(x) dx \rightarrow \int_{\mathbb{R}^n} f(x) dx, \quad \text{com } k \rightarrow \infty.$$

Relembremos agora a definição e um resultado do Operador Adjunto.

Definição 1.4 O **adjunto** de um operador linear $T : V \rightarrow V$, onde V é um espaço vetorial munido de um produto interno, é o operador T^* tal que:

$$\langle T(u), v \rangle = \langle u, T^*(v) \rangle, \forall u, v \in V.$$

Definição 1.5 Definimos **subdiferencial de J** e denotamos por ∂J o seguinte conjunto:

$$\partial J(u) = \{w \in X; J(v) \geq J(u) + \langle w, v - u \rangle_X \, \forall v \in X\}.$$

onde $\langle u, v \rangle_X = \sum_{i,j} u_{i,j} v_{i,j}$, $\forall u, v \in X$ ou considerando $x_{(j-1)N+i} = u_{i,j}$ e $y_{(j-1)N+i} = v_{i,j}$ $1 \leq i, j \leq N$, temos que $\langle u, v \rangle_X = \sum_{i=1}^{N^2} u_i v_i$, e X é o espaço euclidiano $\mathbb{R}^{N \times N}$.

Proposição 1.1 (Proposição 5.1 de [16]) Seja F uma função de V em $\overline{\mathbb{R}}$ e F^* seu adjunto. Assim $u^* \in \partial F(u)$ se e só se $F(u) + F^*(u^*) = \langle u, u^* \rangle$.

Demonstração

Podemos encontrar a demonstração desta proposição em [16].

Definição 1.6 Consideremos o problema de encontrar $v^* \in K$ tal que:

$$\langle v - v^*, F(v^*) \rangle \geq 0, \quad \forall v \in K. \quad (1.3)$$

O problema acima é chamado de **inequação variacional**, denotado por $VI(K, F)$, com v^* sendo uma solução.

Definição 1.7 Um operador linear $F : H \rightarrow H$, onde H é o espaço de Hilbert, é dito ser

1. monotônica se $\langle u - v, F(u) - F(v) \rangle \geq 0, \quad \forall u, v \in H$.
2. fortemente monotônica se $\exists \nu > 0$ tal que $\langle u - v, F(u) - F(v) \rangle \geq \nu \|u - v\|^2, \quad \forall u, v \in H$.
3. pseudo-monotônica se $\langle u - v, F(v) \rangle > 0 \Rightarrow \langle u - v, F(u) \rangle \geq 0, \quad \forall u, v \in H$.

Definição 1.8 Uma componente conexa em u é um subconjunto de Ω , onde todos os pares (p, q) de pontos são conexos (i.e. existe um caminho de p a q e um caminho de q a p , que não são necessariamente os mesmos).

Definição 1.9 Uma função f de $[a, b]$ em $\mathbb{R}$ é dita convexa se o conjunto:

$$\{(x, y) \in \mathbb{R}^2 \mid y \geq f(x)\}$$

for um conjunto convexo. Isto equivale a afirmar que, para quaisquer x e y pertencentes a $[a, b]$ e para todo $t \in [0, 1]$, tem-se:

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Com isso, dizemos que uma aplicação $x \rightarrow f(x)$ é convexa se, $f(x)$ for uma função convexa.

Definição 1.10 Um conjunto fuzzy pode ser caracterizado por uma função de pertinência que mapeia todos os elementos de um domínio, espaço ou universo de discurso X para um número real em $[0, 1]$, isto é, $\tilde{A} : X \rightarrow [0, 1]$. Um conjunto fuzzy apresenta-se como um conjunto de pares ordenados, em que o primeiro elemento é $x \in X$, e o segundo, $\mu_{\tilde{A}}(x)$, é o grau de pertinência ou a função de pertinência de x em $\tilde{A}$, que mapeia x no intervalo $[0, 1]$, ou seja, $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\}$.

1.2 Espaço das Funções Contínuas $C^k(\Omega)$

Nesta seção, definiremos o espaço das funções contínuas $C^k(\Omega)$, suporte compacto e medida de Radon.

Definição 1.11 *Seja $u : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ então definimos suporte de u como sendo:*

$$\text{supp } u = \Omega \cap \overline{\{x; u(x) \neq 0\}}.$$

Definição 1.12 *Seja $\alpha = (\alpha_1, \dots, \alpha_n)$ uma n -úpla de inteiros não-negativos, então α é chamado de multi-índice e seu comprimento é dado por:*

$$|\alpha| = \sum_{i=1}^n \alpha_i.$$

Por simplicidade, denotaremos os operadores derivadas parciais, gradiente de uma função e a derivada de ordem superior, respectivamente por:

$$u_{x_i} = \frac{\partial u(x)}{\partial x_i}, \quad 1 \leq i \leq n;$$

$$\nabla u(x) = \left(\frac{\partial u(x)}{\partial x_1}, \dots, \frac{\partial u(x)}{\partial x_n} \right);$$

$$D^\alpha u(x) = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}(x).$$

A partir destas definições apresentaremos agora as definições do espaço das funções contínuas.

Definição 1.13 *Sejam $\Omega \subset \mathbb{R}^n$ um conjunto aberto e $u : \Omega \rightarrow \mathbb{R}$ uma função contínua, então:*

i) $C^k(\Omega) = \{u; D^\alpha u \in C(\Omega), \alpha \in A_m \text{ com } 0 < m \leq k\}$, sendo o conjunto A_m dado por:

$$A_m = \left\{ \alpha = (\alpha_1, \dots, \alpha_n); \alpha_i \geq 0 \text{ um inteiro e } \sum_{i=1}^n \alpha_i = m \right\};$$

ii) $C_0(\Omega) = \{u \in C(\Omega); \text{supp } u \subset \Omega \text{ é compacto}\}$;

iii) $C_0^k(\Omega) = C^k(\Omega) \cap C_0(\Omega)$;

iv) $C^k(\overline{\Omega}) = \{u \in C^k(\Omega); D^\alpha u \text{ é uniformemente contínua para todo } |\alpha| \leq k\}$,

onde $C^0(\Omega) = C(\Omega) =$ conjunto de todas as funções contínuas $u : \Omega \rightarrow \mathbb{R}$.

Devemos observar que $C(\overline{\Omega})$ é o conjunto de todas as funções contínuas $u : \Omega \rightarrow \mathbb{R}$ cuja continuidade pode ser estendida para $\overline{\Omega}$. Assim, se $u \in C^k(\overline{\Omega})$ então $D^\alpha u$ pode ser estendida continuamente para $(\overline{\Omega})$ para cada multi-índice α , com $|\alpha| \leq k$.

Definição 1.14 *Seja $\Omega \subset \mathbb{R}^n$ um subconjunto aberto, então define-se $C^{0,1}(\Omega)$ como o conjunto formado pelas funções $u \in C(\Omega)$ tais que:*

$$[u]_{C^{0,1}(K)} = \sup_{x,y \in K, x \neq y} \left\{ \frac{|u(x) - u(y)|}{|x - y|^\alpha} \right\} < \infty,$$

para todo conjunto compacto $K \subset \Omega$.

Definição 1.15 *i) Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto e limitado. Dizemos que Ω tem fronteira C^k , $k \geq 1$ se para todos $x \in \partial\Omega$ existir uma vizinhança $U \subset \mathbb{R}^n$ de x e uma aplicação bijetora $H : Q \rightarrow U$, onde*

$$Q = \{x \in \mathbb{R}^n; |x_j| < 1, j = 1, 2, \dots, n\}$$

e

$$H \in C^k(\overline{Q}), H^{-1} \in C^k(\overline{U}), H(Q_+) = U \cap \Omega, H(Q_0) = U \cap \partial\Omega,$$

com $Q_+ = \{x \in Q, x_n > 0\}$ e $Q_0 = \{x \in Q; x_n = 0\}$, onde $\partial\Omega =$ fronteira de Ω e $\overline{\Omega} = \Omega \cup \partial\Omega$ o fecho de Ω .

ii) Se H está apenas em $C^{0,1}$ dizemos que Ω é um conjunto aberto, limitado e com fronteira Lipschitz.

Definição 1.16 *Se $0 < \gamma \leq 1$, então $C^{k,\gamma}$ é o espaço de funções contínuas u em Ω tal que $|u(x) - u(y)| \leq C|x - y|^\gamma$, para alguma constante C , x e $y \in \Omega$. Isso é chamado de **espaço de Holder** de funções contínuas com expoente γ .*

Definição 1.17 *Definimos:*

$$C_c(X) = \{u : X \subset \mathbb{R}^n \rightarrow \mathbb{R}; u \text{ é contínua e } \text{supp}(u) \text{ é compacto}\},$$

então $C_c(X)$ é um $\mathbb{R}$ -espaço vetorial que coincide com $C(X)$ para X compacto.

Definição 1.18 *Seja μ uma medida de Borel em X e E um subconjunto de Borel de X . A medida μ é chamada de **outer regular** em E se :*

$$\mu(E) = \inf\{\mu(U); U \supset E, U \text{ aberto}\}$$

*e **inner regular** em E se:*

$$\mu(E) = \sup\{\mu(K); K \subset E, K \text{ compacto}\}.$$

Se μ é outer e inner regular em todos os conjuntos de Borel então μ é chamada de regular.

Definição 1.19 *Uma **medida de Radon** em X é uma medida de Borel que é finita em todo conjunto compacto, outer regular em todo conjunto de Borel e inner regular em todo conjunto aberto. (Maiores detalhes ver [21])*

1.3 Espaços L^p

Daremos nesta seção a definição do espaço L^p e alguns resultados importantes sobre estes espaços.

Definição 1.20 *Seja $\Omega \subset \mathbb{R}^n$ mensurável e $1 \leq p < \infty$, definimos $L^p(\Omega)$ como a classe de funções mensuráveis, $u : \Omega \rightarrow \overline{\mathbb{R}}$, tais que:*

$$\int_{\Omega} |u(x)|^p dx < +\infty.$$

Para $u \in L^p(\Omega)$ define-se:

$$\|u\|_{L^p(\Omega)} := \left[\int_{\Omega} |u(x)|^p dx \right]^{\frac{1}{p}}, \quad (1.4)$$

onde $\|f\|_{L^p(\Omega)}$ define uma norma para este espaço (ver [23]).

Definição 1.21 *Seja $\Omega \subset \mathbb{R}^n$ mensurável e $p = \infty$. Assim, define-se $L^\infty(\Omega)$ como sendo o espaço das funções mensuráveis $u : \Omega \rightarrow \overline{\mathbb{R}}$ e limitadas em Ω , ou seja, existe uma constante $\lambda \in \mathbb{R}$ tal que:*

$$|u(x)| \leq \lambda, \text{ q.t.p. } x \in \Omega.$$

Para $u \in L^\infty(\Omega)$ definimos:

$$\|u\|_{L^\infty} = \inf\{\lambda \in \mathbb{R}; |u(x)| \leq \lambda, \text{ q.t.p. } x \in \Omega\} = \inf \text{ ess } \sup |u|$$

Como em $L^p(\Omega)$, podemos demonstrar que o espaço $L^\infty(\Omega)$ é um espaço de Banach. (ver [23])

Agora apresentaremos definições e alguns resultados da convergência forte, fraca e fraca estrela.

Definição 1.22 *Seja $\Omega \subset \mathbb{R}^n$ uma região aberta e $1 \leq p \leq +\infty$. Então:*

i) *Uma sequência $\{u_n\}_{n=1}^\infty$ converge fortemente para u , se u_n e $u \in L^p$ e ainda se:*

$$\lim_{n \rightarrow \infty} \|u_n(x) - u(x)\|_{L^p(\Omega)} = 0.$$

Denotamos a convergência forte em $L^p(\Omega)$ por $u_n \rightarrow u$ em L^p .

ii) *Se $1 \leq p < \infty$, dizemos que a sequência $\{u_n\}_{n=1}^\infty$ converge fracamente para u , se u_n e $u \in L^p$ e ainda se:*

$$\lim_{n \rightarrow \infty} \int_{\Omega} [u_n(x) - u(x)] \varphi(x) dx = 0, \quad \forall \varphi \in L^q(\Omega),$$

onde $\frac{1}{p} + \frac{1}{q} = 1$ e denotamos a convergência fraca em $L^p(\Omega)$ por: $u_n \rightharpoonup u$ em L^p .

iii) *Se $p = \infty$ a sequência u_n é dita convergir fracamente estrela para u se u_n e $u \in L^\infty(\Omega)$ e se:*

$$\lim_{n \rightarrow \infty} \int_{\Omega} [u_n(x) - u(x)] \varphi(x) dx = 0, \quad \forall \varphi \in L^1(\Omega).$$

Denotamos a convergência fraca estrela em $L^\infty(\Omega)$ por: $u_n \xrightarrow{*} u$ em $L^\infty(\Omega)$.

Observação 1.1 *Notemos que:*

$$u_n \rightarrow u \text{ em } L^p \Rightarrow \begin{cases} u_n \rightharpoonup u \text{ em } L^p, & 1 \leq p < \infty, \\ u_n \xrightarrow{*} u \text{ em } L^\infty, & p = \infty. \end{cases}$$

Teorema 1.3 *Seja $\Omega \subset \mathbb{R}^n$ um subconjunto aberto e limitado. Então tem-se as seguintes propriedades:*

1. *Se $u_n \xrightarrow{*} u$ em L^∞ , então $u_n \rightharpoonup u$ em L^p , $\forall 1 \leq p < \infty$.*
2. *Se $u_n \rightarrow u$ em L^p , então $\|u_n\|_{L^p} \rightarrow \|u\|_{L^p}$ com $1 \leq p \leq \infty$.*
3. *Se $1 \leq p < \infty$ e se $u_n \rightharpoonup u$ em L^p , então existe uma constante $\delta > 0$ tal que:*

$$\|u_n\|_{L^p} \leq \delta$$

e mais ainda:

$$\|u\|_{L^p} \leq \liminf_{n \rightarrow \infty} \|u_n\|_{L^p}.$$

4. *Se $1 < p < \infty$ e se existir uma constante $\delta > 0$ tal que $\|u_n\|_{L^p} \leq \delta$, então existe uma subsequência $\{u_{n_k}\}$ e $u \in L^p$ tal que $u_{n_k} \rightarrow u$ em L^p .*
5. *Se $1 \leq p < \infty$, então existe uma sequência $u_k \in C_0^\infty$ tal que*

$$\lim_{k \rightarrow \infty} \|u_k - u\| = 0.$$

Demonstração

A demonstração deste teorema pode ser encontrada em [23].

Definição 1.23 *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto e $1 \leq p \leq \infty$. Dizemos que $u \in L_{loc}^p(\Omega)$ se $u \in L^p(K)$ para todo conjunto aberto K , onde $\overline{K} \subset \Omega$ e $\overline{K}$ é compacto.*

1.4 Espaços de Sobolev, $W^{k,p}(\Omega)$

Nesta seção apresentaremos a definição dos espaços de Sobolev e algumas propriedades importantes sobre estes espaços. Antes porém, iremos apresentar uma motivação para as derivadas fracas, chamando as funções $\phi \in C_0^\infty(\Omega)$ de **funções teste**, $\phi : \Omega \rightarrow \mathbb{R}$ com suporte compacto em Ω , e lembrando que $C_0^\infty(\Omega)$ representa o espaço das funções infinitamente diferenciáveis.

Sejam $u \in C^1(\Omega)$ e $\phi \in C_0^\infty(\Omega)$, então ao usarmos integração por partes temos que:

$$\int_{\Omega} u \phi_{x_i} dx = - \int_{\Omega} u_{x_i} \phi dx, \quad (i = 1, \dots, n). \quad (1.5)$$

Não existe condição de contorno, já que ϕ tem suporte compacto em Ω e por isso se anula próximo de $\partial\Omega$. De uma forma mais geral, seja k um inteiro positivo, $u \in C^k(\Omega)$ e $\alpha = (\alpha_1, \dots, \alpha_n)$ um multi-índice de ordem $|\alpha| = \alpha_1 + \dots + \alpha_n = k$, temos então que:

$$\int_{\Omega} u D^{\alpha} \phi dx = (-1)^{|\alpha|} \int_{\Omega} D^{\alpha} u \phi dx. \quad (1.6)$$

A igualdade acima é válida pois:

$$D^{\alpha} \phi(x) = \frac{\partial^{|\alpha|} \phi(x)}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \dots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \phi(x)$$

e assim podemos aplicar (1.5) $|\alpha|$ vezes.

Agora, examinado a igualdade (1.6), válida para $u \in C^k(\Omega)$ podemos nos perguntar: alguma variação desta igualdade poderia ser válida caso u não fosse k vezes diferenciável? Temos que o lado esquerdo de (1.6) faz sentido para u apenas localmente integrável, e o problema é que u não é C^k , assim a expressão $D^{\alpha} u$ do lado direito de (1.6) não tem sentido. Este problema estará resolvido se existir uma função localmente integrável v para que (1.6) seja válida, com v no lugar de $D^{\alpha} u$.

Definição 1.24 *Suponha que $u, v \in L^1_{loc}(\Omega)$ e seja α um multi-índice. Dizemos que v é a α -ésima derivada parcial fraca de u , e escrevemos $v = D^{\alpha} u$, se:*

$$\int_{\Omega} u D^{\alpha} \phi dx = (-1)^{|\alpha|} \int_{\Omega} v \phi dx, \quad (1.7)$$

para toda função teste $\phi \in C_0^{\infty}(\Omega)$.

Definição 1.25 *O espaço de Sobolev, $W^{k,p}(\Omega)$ com $1 \leq p \leq \infty$ e k um inteiro não negativo, consiste de todas as funções $u : \Omega \rightarrow \mathbb{R} \in L^p(\Omega)$, tal que, para cada multi-índice α com $|\alpha| \leq k$, $D^{\alpha} u$ existe no sentido fraco e pertence a $L^p(\Omega)$, ou seja:*

$$W^{k,p} = \{u \in L^p(\Omega); D^{\alpha} u \in L^p(\Omega), \text{ para } 0 \leq |\alpha| \leq k\}.$$

Ilustrando, temos que:

$$W^{1,2} = \{u \in L^2(\Omega); u_{x_i} \in L^2(\Omega)\},$$

onde u_{x_i} é a derivada parcial na variável x_i no sentido fraco. Geralmente escreve-se $H^1(\Omega) = W^{1,2}(\Omega)$. A letra H é usada devido ao fato de $H^1(\Omega)$ ser um espaço de Hilbert.

Definição 1.26 *Seja $u \in W^{k,p}(\Omega)$ então define-se sua norma, como sendo:*

$$\|u\|_{W^{k,p}(\Omega)} = \begin{cases} \left(\sum_{|\alpha| \leq k} \int_{\Omega} |D^{\alpha} u|^p dx \right)^{\frac{1}{p}}, & \text{se } 1 \leq p < \infty, \\ \sum_{|\alpha| \leq k} \text{ess sup}_{\Omega} |D^{\alpha} u|, & \text{se } p = \infty. \end{cases}$$

Depois de definirmos a norma em $W^{k,p}(\Omega)$ podemos definir convergência no espaço de Sobolev do seguinte modo:

Definição 1.27 *Sejam $\{u_k\}_{k=1}^{\infty}$ uma sequência em $W^{k,p}(\Omega)$ e $u \in W^{k,p}(\Omega)$.*

i) Dizemos que u_k converge para u e denotamos por:

$$u_k \rightarrow u \text{ em } W^{k,p}(\Omega)$$

se

$$\lim_{k \rightarrow \infty} \|u_k - u\|_{W^{k,p}(\Omega)} = 0.$$

ii) Escrevemos

$$u_k \rightarrow u \text{ em } W_{loc}^{k,p}(\Omega)$$

quando:

$$u_k \rightarrow u \text{ em } W_{loc}^{k,p}(V)$$

para cada $V \subset \bar{V} \subset U$ e $\bar{V}$ compacto. Dizemos que $u \in W_{loc}^{k,p}(\Omega)$ com $1 \leq p \leq \infty$ e k um inteiro não negativo se $u \in W^{k,p}(K)$, para todo conjunto aberto e compacto K , onde $\bar{K} \subset \Omega$.

Maiores detalhes deste espaço, com propriedades e resultados pode ser encontrado em [25].

1.5 Espaço das funções de variação limitada, $BV(\Omega)$

Nesta seção apresentaremos a definição e alguns conceitos sobre as funções de variação limitada, já que na maior parte deste trabalho iremos procurar soluções (funções) no espaço $BV(\Omega)$.

Definição 1.28 *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto e seja $u \in L^1(\Omega)$, então definimos:*

$$\int_{\Omega} |\nabla u| = \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi dx; \varphi \in C_0^1(\Omega; \mathbb{R}^n), |\varphi(x)| \leq 1, \text{ para } x \in \Omega \right\},$$

onde, $\operatorname{div} \varphi = \sum_{i=1}^n \frac{\partial \varphi_i}{\partial x_i}$.

Definição 1.29 *O espaço $BV(\Omega)$ é definido da seguinte forma:*

$$BV(\Omega) = \left\{ u \in L^1(\Omega); \int_{\Omega} |\nabla u| < \infty \right\},$$

e uma norma no espaço $BV(\Omega)$ será dada por:

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + \int_{\Omega} |\nabla u|.$$

As propriedades de norma são facilmente verificadas a partir da definição da norma de u em $L^1(\Omega)$ e da definição de $\int_{\Omega} |\nabla u|$.

Teorema 1.4 *(Teorema 1.17 de [18]) Seja $u \in BV(\Omega)$. Então existe uma sequência $\{u_j\}$ em $C^\infty(\Omega)$ tal que:*

$$\lim_{j \rightarrow \infty} \int_{\Omega} |u_j - u| dx = 0$$

e

$$\lim_{j \rightarrow \infty} \int_{\Omega} |\nabla u_j| dx = \int_{\Omega} |\nabla u| dx.$$

Demonstração *A demonstração deste teorema pode ser encontrada em [14] e [18].*

□

Observação 1.2 Se $u \in BV(\Omega) \cap L^2(\Omega)$ e $\partial\Omega$ é a fronteira Lipschitz, então existe uma sequência de funções $\{u_n\} \subset C^\infty(\bar{\Omega})$ tal que:

$$u_n \rightarrow u \text{ em } L^2(\Omega)$$

e

$$\int_{\Omega} |\nabla u_n| \, dx \rightarrow \int_{\Omega} |\nabla u| \, dx.$$

A demonstração desta observação pode ser encontrada em [14].

Usando a observação anterior e com uma pequena modificação do Teorema 1.4 podemos ter $u_n \in C^\infty \cap W^{1,1} \cap L^2(\Omega)$ tal que:

$$u_n \rightarrow u \text{ em } L^2(\Omega)$$

e

$$\int_{\Omega} |\nabla u_n| \, dx \rightarrow \int_{\Omega} |\nabla u| \, dx.$$

A modificação citada pode ser encontrada em [14].

1.6 Condições de Karush-Kuhn-Tucker (KKT)

As condições de Karush-Kuhn-Tucker (KKT) são úteis para encontrar soluções de problemas de otimização em programação não-linear, desde que algumas condições de regularidade estejam satisfeitas.

Suponhamos que a função a ser minimizada seja $u : \mathbb{R}^n \rightarrow \mathbb{R}$, e as funções de restrições sejam $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ e $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$. Além disso, suponhamos também que estas funções sejam continuamente diferenciáveis em um ponto x^* . Se x^* é um mínimo local que satisfaz todas as restrições, então existem constantes μ_i ($i = 1, \dots, m$) e λ_j ($j = 1, \dots, l$), chamados de multiplicadores de Lagrange, de tal forma que:

$$-\nabla u(x^*) + \sum_{i=1}^m \mu_i \nabla g_i(x^*) + \sum_{j=1}^l \lambda_j \nabla h_j(x^*) = 0$$

Estas condições são conhecidas como **condições de Karush-Kuhn-Tucker (KKT)**.

Maiores detalhes sobre estas condições e exemplos de sua aplicação pode ser encontrada em [22].

Capítulo 2

Cálculo Variacional

Neste capítulo, vamos introduzir conceitos básicos do Cálculo Variacional com intuito de apresentar resultados necessários para a análise dos modelos propostos nos capítulos subsequentes. Podemos ver o Cálculo Variacional como o cálculo diferencial no espaço das funções onde tentaremos, sobre espaços apropriados, encontrar funções que minimizem certos funcionais. Para encontrar tais funções, consideraremos funções $u : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

2.1 Preliminares

O Cálculo Variacional visa fundamentalmente investigar máximos e mínimos de funcionais. Portanto, nesta seção, vamos apresentar algumas definições e teoremas que irão garantir a existência e unicidade de pontos de mínimo dos funcionais em questão e, conseqüentemente, a existência e unicidade de soluções dos problemas de valor de contorno associados a estes funcionais. Consideraremos funcionais do tipo $I : V \rightarrow \mathbb{R}$, onde $V \subset C[a, b]$ é chamado de **conjunto das funções admissíveis de I** , e se um elemento $v \in V$, então ele poderá ser escrito como $v = v(x)$, $a \leq x \leq b$.

Definição 2.1 Dados $v \in V$ e $\varepsilon > 0$, definimos **vizinhança** de $v \in V$ de raio $\varepsilon > 0$ como:

$$B(v, \varepsilon) = \{w \in V \mid 0 \leq \|v - w\|_{L^2[a, b]} < \varepsilon\}.$$

Definição 2.2 Seja o funcional $I : V \rightarrow \mathbb{R}$. Então se existe $\varepsilon > 0$, tal que, $I(\tilde{v}) \leq I(v)$, $\forall v \in B(\tilde{v}, \varepsilon)$, temos que $\tilde{v} \in V$ é um **mínimo local** de I . Agora, se $I(\tilde{v}) < I(v)$, $\forall v \in B(\tilde{v}, \varepsilon)$, com $v \neq \tilde{v}$, então $\tilde{v}$ é um **mínimo local forte** de I .

Definição 2.3 *Seja o funcional $I : V \rightarrow \mathbb{R}$. Então se $I(\check{v}) \leq I(v)$, $\forall v \in V$, temos que $\check{v} \in V$ é um **mínimo global** de I . Agora, se $I(\check{v}) < I(v)$, $\forall v \in V$, com $v \neq \check{v}$, então $\check{v}$ é um **mínimo global forte** de I .*

Temos que o conjunto V poderá ser, ou não, um espaço linear, porém em ambos os casos será possível encontrar o seguinte conjunto:

$$\tilde{V} = \{\eta \mid \eta = v - w \text{ com } v, w \in V\}, \quad (2.1)$$

que é um espaço linear, chamado de **espaço das funções teste**. Deste modo, temos que o conjunto V poderá ser escrito como $V = \{w \mid w = v^* + \eta, \eta \in \tilde{V}\}$, sendo v^* um elemento arbitrário, porém fixo, de V .

Observemos ainda que, a vizinhança $B(v, \varepsilon)$ dada da Definição 2.1 é equivalente a vizinhança definida por

$$\check{B}(v, \varepsilon) = \{w \in V \mid w = v + \tau\eta; \eta \in \tilde{V}; \|\eta\|_{L^2[a,b]} = 1; \tau \in (-\varepsilon, \varepsilon)\}.$$

De fato, seja $w \in B(v, \varepsilon)$, i.e., $0 \leq \|v - w\|_{L^2[a,b]} < \varepsilon$. Observemos que:

$$w = v + \frac{w - v}{\|w - v\|_{L^2[a,b]}} \cdot \|w - v\|_{L^2[a,b]}.$$

Assim, tomando $\eta = \frac{w-v}{\|w-v\|_{L^2[a,b]}}$, temos que

$$\|\eta\|_{L^2[a,b]} = 1,$$

e desta forma, fazendo $\tau = \|w - v\|_{L^2[a,b]}$ tem-se que $\tau \in (-\varepsilon, \varepsilon)$. Logo, $w = v + \tau\eta$, $\forall v \in V$ com $v \neq w$, ou seja, temos que $w \in \check{B}(v, \varepsilon)$.

Analogamente, seja $w \in V$ com $w = v + \tau\eta$, $v \in V$, $\eta \in \tilde{V}$, $\|\eta\|_{L^2[a,b]} = 1$ e $\tau \in (-\varepsilon, \varepsilon)$. Como $w = v + \tau\eta$, temos que $w - v = \tau\eta$, assim:

$$\|w - v\|_{L^2[a,b]} = \|\tau\eta\|_{L^2[a,b]} = |\tau| \|\eta\|_{L^2[a,b]}.$$

Logo, $0 \leq \|w - v\|_{L^2[a,b]} = \tau$, ou seja, $0 \leq \|w - v\|_{L^2[a,b]} < \varepsilon$. Assim, $w \in B(v, \varepsilon)$. Portanto, concluímos que as duas vizinhanças $B(v, \varepsilon)$ e $\check{B}(v, \varepsilon)$ definidas acima são iguais.

Definição 2.4 Seja o funcional $I : V \rightarrow \mathbb{R}$. Sejam $v \in V$ e $\eta \in \tilde{V}$ dados, onde $\|\eta\|_{L^2[a,b]} = 1$, e suponhamos que para algum $\tau_0 > 0$, a função $I(v + \tau\eta)$ com $|\tau|_{L^2[a,b]} < \tau_0$, tenha derivada de ordem m contínua com relação a τ . Então, a **derivada direcional de ordem m** de I em v na direção de η é:

$$I^{(m)}(v; \eta) = \left. \frac{d^m I(v + \tau\eta)}{d\tau^m} \right|_{\tau=0}.$$

Definição 2.5 Seja o funcional linear $I : V \rightarrow \mathbb{R}$ e suponha que para algum $\tilde{v} \in V$, $I^{(1)}(\tilde{v}; \eta) = 0$, $\forall \eta \in \tilde{V}$ com $\|\eta\|_{L^2[a,b]} = 1$. Então, $\tilde{v}$ é um ponto **estacionário** de I .

Teorema 2.1 Seja o funcional $I : V \rightarrow \mathbb{R}$ e suponha que para algum $\tilde{v} \in V$ a derivada direcional de primeira ordem $I^{(1)}(\tilde{v}; \eta)$ exista para todas as direções de η . Se $\tilde{v}$ é um mínimo local de I então I é estacionário em $\tilde{v}$.

Demonstração:

De acordo com as hipóteses temos a seguinte expansão de Taylor para a função $g(t) = I(\tilde{v} + \tau\eta)$ numa vizinhança de $\tilde{v}$:

$$I(\tilde{v} + \tau\eta) = I(\tilde{v}) + \tau I^{(1)}(\tilde{v}; \eta) + O(\tau),$$

para algum $\eta \in \tilde{V}$, $\|\eta\|_{L^2[a,b]} = 1$, $\tau \in (-\varepsilon, \varepsilon)$, onde $\lim_{\tau \rightarrow 0} \frac{O(\tau)}{|\tau|} = 0$ e $O(\tau)$ denota a ordem da expansão de Taylor.

Assim,

$$I^{(1)}(\tilde{v}; \eta) = \frac{I(\tilde{v} + \tau\eta) - I(\tilde{v})}{\tau} - \frac{O(\tau)}{\tau}.$$

Notemos que $\lim_{\tau \rightarrow 0} \frac{O(\tau)}{\tau} = 0$, pois $\lim_{\tau \rightarrow 0} \frac{O(\tau)}{|\tau|} = 0$ e $I(\tilde{v} + \tau\eta) \geq I(\tilde{v})$, pois $\tilde{v}$ é um ponto de mínimo local. Dessa forma, se $\tau > 0$ então $I^{(1)}(\tilde{v}; \eta) \geq 0$, pois observemos que

$$I^{(1)}(\tilde{v}; \eta) \geq -\frac{O(\tau)}{\tau},$$

já que $\frac{I(\tilde{v} + \tau\eta) - I(\tilde{v})}{\tau} \geq 0$. Assim,

$$\lim_{\tau \rightarrow 0} I^{(1)}(\tilde{v}; \eta) \geq -\lim_{\tau \rightarrow 0} \frac{O(\tau)}{\tau},$$

logo, $I^{(1)}(\check{v}; \eta) \geq 0$. Agora, se $\tau < 0$ então $I^{(1)}(\check{v}; \eta) \leq 0$, pois observemos que

$$I^{(1)}(\check{v}; \eta) \leq -\frac{O(\tau)}{\tau},$$

já que $\frac{I(\check{v} + \tau\eta) - I(\check{v})}{\tau} \leq 0$. Assim,

$$\lim_{\tau \rightarrow 0} I^{(1)}(\check{v}; \eta) \leq -\lim_{\tau \rightarrow 0} \frac{O(\tau)}{\tau},$$

logo, $I^{(1)}(\check{v}; \eta) \leq 0$.

Portanto, temos que $I^{(1)}(\check{v}; \eta) = 0$, ou seja, $\check{v}$ é um ponto estacionário de I .

□

Definição 2.6 Um funcional $I : V \rightarrow \mathbb{R}$ é um funcional quadrático se satisfaz a seguinte identidade:

$$I(v + \tau\eta) = I(v) + \tau I^{(1)}(v, \eta) + \frac{\tau^2}{2} I^{(2)}(v, \eta),$$

$\forall v \in V, \forall \eta \in \tilde{V}$ com $\|\eta\|_{L^2[a,b]} = 1$ e $\forall \tau \in \mathbb{R}$.

Teorema 2.2 Seja o funcional quadrático $I : V \rightarrow \mathbb{R}$. Então, $\check{v} \in V$ é o único mínimo local e único mínimo global forte de I se:

1. $I^{(1)}(\check{v}; \eta) = 0, \forall \eta \in \tilde{V}, \|\eta\|_{L^2[a,b]} = 1$;
2. $I^{(2)}(\check{v}; \eta) > 0, \forall \eta \in \tilde{V}, \|\eta\|_{L^2[a,b]} = 1$.

Demonstração:

Seja $I : V \rightarrow \mathbb{R}$ um funcional quadrático, ou seja, temos a seguinte igualdade:

$$I(\check{v} + \tau\eta) = I(\check{v}) + \tau I^{(1)}(\check{v}, \eta) + \frac{\tau^2}{2} I^{(2)}(\check{v}, \eta), \quad \forall \check{v} \in V, \eta \in \tilde{V}, \|\eta\|_{L^2[a,b]} = 1 \text{ e } \forall \tau \in \mathbb{R}.$$

Por hipótese, temos que $I^{(1)}(\check{v}; \eta) = 0$, assim,

$$I(\check{v} + \tau\eta) = I(\check{v}) + \frac{\tau^2}{2} I^{(2)}(\check{v}, \eta),$$

isto é, $I(\tilde{v} + \tau\eta) > I(\tilde{v})$, $\forall \eta \in \tilde{V}$, $\forall \tau \in \mathbb{R}$, $\tau \neq 0$. Assim, $I(v) > I(\tilde{v})$, $\forall v \in V, v \neq \tilde{v}$, pois

$$V = \{\tilde{v} + \tau\eta \mid \tau \in \mathbb{R}, \eta \in \tilde{V}, \|\eta\|_{L^2[a,b]} = 1\}.$$

Logo, $\tilde{v}$ é um mínimo global forte de I .

Agora, suponhamos que $\tilde{w}$ seja um mínimo local de I . Então temos que

$$I^{(1)}(\tilde{w}; \eta) = 0 \text{ e } I^{(2)}(\tilde{w}; \eta) \geq 0.$$

Como I é quadrático temos que:

$$I(\tilde{w} + \tau\eta) = I(\tilde{w}) + \tau I^{(1)}(\tilde{w}; \eta) + \frac{\tau^2}{2} I^{(2)}(\tilde{w}; \eta),$$

assim,

$$I(\tilde{w} + \tau\eta) = I(\tilde{w}) + \frac{\tau^2}{2} I^{(2)}(\tilde{w}; \eta),$$

isto é, $I(\tilde{w} + \tau\eta) \geq I(\tilde{w})$. Logo, $I(v) \geq I(\tilde{w})$, $\forall v \in V$, $v \neq \tilde{w}$, ou seja, $\tilde{w}$ é mínimo global de I . Porém, observemos que se $\tilde{w}$ for um mínimo global de I , diferente de $\tilde{v}$, então segue que $I(\tilde{w}) \leq I(\tilde{v}) < I(\tilde{w})$, o que seria um absurdo. Portanto, $\tilde{w} = \tilde{v}$ é o único mínimo global forte de I .

□

Lema 2.1 Se $G : [a, b] \rightarrow \mathbb{R}$ é uma função contínua e se $\int_a^b G(x)\eta(x) dx = 0$ para toda função diferenciável $\eta : [a, b] \rightarrow \mathbb{R}$ tal que $\eta(a) = \eta(b) = 0$ então:

$$G(x) = 0, \quad \forall x \in (a, b).$$

Demonstração:

Suponhamos por absurdo que $G(x') \neq 0$ para algum $x' \in (a, b)$. Sem perda de generalidade, vamos supor que $G(x') > 0$. Pela hipótese da continuidade de G , temos que existe uma vizinhança de x' , digamos, $c \leq x' \leq d$ tal que $G(x) > 0, \forall x \in [c, d]$, porém, com isso, temos que a igualdade abaixo não se verifica para toda função diferenciável η

$$\int_a^b \eta(x)G(x) \, dx = 0.$$

Por exemplo, tomando-se a função

$$\eta(x) = \begin{cases} 0 & \text{se } a \leq x \leq c \\ (x-c)^2(x-d)^2 & \text{se } c < x < d \\ 0 & \text{se } d \leq x \leq b \end{cases}$$

temos que η em particular é diferenciável e satisfaz as condições de contorno $\eta(a) = \eta(b) = 0$, então temos que:

$$\int_a^b G(x)\eta(x) \, dx = \int_c^d G(x)(x-c)^2(x-d)^2 \, dx,$$

e como supomos que $G(x) > 0$, $\forall x \in (c, d)$, temos que

$$\int_a^b G(x)\eta(x) \, dx \neq 0,$$

o que contradiz a hipótese. Logo $G(x) = 0$, $\forall x \in (a, b)$. Já o caso $G(x') < 0$ é análogo e assim o lema está provado.

□

Após termos provado as condições para que o funcional I tenha pontos de mínimo, veremos na próxima seção como encontrar tais pontos através da equação de Euler-Lagrange.

2.2 Equação de Euler-Lagrange

Nesta seção, iremos introduzir alguns conceitos supondo que queremos resolver equações diferenciais parciais que, por simplicidade, denotaremos da seguinte forma:

$$A[u] = 0. \tag{2.2}$$

onde $A[\cdot]$ denota um possível operador diferencial parcial não-linear e u é desconhecido.

Sabemos que não existe uma teoria geral para resolver tal EDP, porém, o cálculo variacional identifica uma importante classe de tais problemas não-lineares que podem ser resolvidos usando

técnicas de análise funcional. Esta é a chamada classe de problemas variacionais, ou seja, EDP da forma (2.2), onde o operador não-linear $A[\cdot]$ é a “derivada” de um funcional de “energia” apropriado $I[\cdot]$. Simbolicamente, podemos escrever

$$A[\cdot] = I^{(1)}[\cdot]. \quad (2.3)$$

Assim, o problema (2.2) pode ser reescrito como

$$I^{(1)}[u] = 0. \quad (2.4)$$

Podemos dizer que a vantagem desta nova formulação é que agora é possível reconhecer soluções da equação (2.2) como sendo pontos críticos do funcional $I[\cdot]$. Em certas circunstâncias, este fato pode ser facilmente determinado, como por exemplo, se supormos que o funcional $I[\cdot]$ tenha um mínimo u , então a expressão (2.4) é válida e assim u é uma solução fraca da EDP dada em (2.2). O que devemos observar é que quase sempre é difícil de se resolver a equação (2.2) diretamente, podendo ser mais fácil encontrar um ponto de máximo, mínimo ou outros pontos críticos do funcional $I[\cdot]$.

Temos que inúmeras leis da Ciência originaram-se diretamente de princípios variacionais, que são problemas de valores de contorno nos quais procura-se uma função que satisfaça alguma equação diferencial em uma região Ω , e determinadas condições em $\partial\Omega$. Problemas deste tipo tem a propriedade de que sua solução minimiza um certo funcional $I[\cdot]$, definido em um determinado conjunto de funções, ou em outras palavras, esta solução é um ponto estacionário do funcional $I[\cdot]$.

2.2.1 Primeira variação: equação de Euler-Lagrange

Vamos supor que $\Omega \subset \mathbb{R}^{n-1}$ seja um conjunto aberto, limitado e com fronteira $\partial\Omega$ suave. Consideremos uma função suave $L : \mathbb{R}^n \times \mathbb{R} \times \overline{\Omega} \rightarrow \mathbb{R}$, chamada de **função Lagrangeana** e denotada por:

$$L = L(p, z, x) = L(p_1, \dots, p_n, z, x_1, \dots, x_n)$$

com $p \in \mathbb{R}^n$, $z \in \mathbb{R}$ e $x \in \overline{\Omega}$.

¹Os casos onde $\Omega = [a, b]$ e $\Omega \subset \mathbb{R}^2$ podem ser encontrados em [33].

Substituindo as variáveis “p” por $\nabla w(x)$ e “z” por $w(x)$, temos:

$$L = L(\nabla w(x), w(x), x).$$

Para simplificação das notações definimos:

$$\begin{cases} D_p L = (L_{p_1}, \dots, L_{p_n}) \\ D_z L = L_z \\ D_x L = (L_{x_1}, \dots, L_{x_n}) \end{cases}$$

e consideremos o seguinte funcional $I[\cdot]$:

$$I[w] = \int_{\Omega} L(\nabla w(x), w(x), x) \, dx \quad (2.5)$$

onde a função $w : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é suave e satisfaz a seguinte condição de contorno:

$$w = g \text{ em } \partial\Omega. \quad (2.6)$$

Devemos observar que o funcional (2.5) é construído a partir da função Lagrangeana, onde as variáveis p e z foram substituídas por $\nabla w(x)$ e $w(x)$, respectivamente.

Suponhamos agora que, alguma função u com certa suavidade, satisfaça a condição de contorno exigida (2.6), ou seja, que $u = g$ em $\partial\Omega$ e também que, por sorte, u seja uma função minimizadora do funcional $I[\cdot]$ dentre todas as funções w que também satisfazem (2.6). Iremos mostrar que u é então, automaticamente, solução de uma certa equação diferencial parcial não-linear.

Para confirmarmos este fato, escolhemos uma função qualquer $v \in C_0^\infty(\Omega)$, e consideramos a função i com valores reais, dada por:

$$i(\tau) := I[u + \tau v], \text{ com } \tau \in \mathbb{R}. \quad (2.7)$$

Como por hipótese, u é um minimizador de $I[\cdot]$ e também $u + \tau v = u = g$ em $\partial\Omega$, então temos que a função $i(\cdot)$ atinge seu mínimo quando $\tau = 0$, ou seja,

$$i'(0) = 0. \quad (2.8)$$

A derivada dada pela equação (2.8) é calculada de forma explícita e é chamada de primeira variação. De fato, fazendo substituições adequadas e algumas simplificações temos que a equação (2.7) se reduz a

$$i(\tau) = \int_{\Omega} L(\nabla u + \tau \nabla v, u + \tau v, x) \, dx. \quad (2.9)$$

Assim, derivando a igualdade acima temos:

$$i'(\tau) = \frac{d}{d\tau} \int_{\Omega} L(\nabla u + \tau \nabla v, u + \tau v, x) \, dx \quad (2.10)$$

e, conseqüentemente,

$$i'(\tau) = \int_{\Omega} \sum_{i=1}^n L_{p_i}(\nabla u + \tau \nabla v, u + \tau v, x) v_{x_i} + L_z(\nabla u + \tau \nabla v, u + \tau v, x) v \, dx. \quad (2.11)$$

Agora, fazendo $\tau = 0$ deduz-se de (2.8) que

$$0 = i'(0) = \int_{\Omega} \sum_{i=1}^n L_{p_i}(\nabla u, u, x) v_{x_i} + L_z(\nabla u, u, x) v \, dx,$$

e usando o fato de que v tem suporte compacto (ver definição 1.8) e também fazendo uso de integração por partes, obtemos

$$0 = i'(0) = \int_{\Omega} \left[- \sum_{i=1}^n (L_{p_i}(\nabla u, u, x))_{x_i} + L_z(\nabla u, u, x) \right] v \, dx.$$

Como esta igualdade é válida para todas as funções teste v , podemos concluir que u resolve a seguinte EDP:

$$- \sum_{i=1}^n (L_{p_i}(\nabla u, u, x))_{x_i} + L_z(\nabla u, u, x) = 0 \text{ em } \Omega. \quad (2.12)$$

A equação dada acima é a chamada **equação de Euler-Lagrange**, associada ao funcional de energia $I[\cdot]$, definido em (2.5).

De forma resumida podemos concluir que qualquer função suave que venha minimizar o funcional $I[\cdot]$ é uma solução da equação de Euler-Lagrange dada em (2.12), e reciprocamente, podemos tentar encontrar uma solução de (2.12) procurando por funções que minimizem (2.5).

Veremos agora um exemplo com a finalidade de enfatizar o processo de obtenção da equação de Euler-Lagrange associada a um certo funcional $I[\cdot]$.

Exemplo 2.1 Considere a função

$$L(p, z, x) = (1 + |p|^2)^{\frac{1}{2}}$$

e o funcional

$$I[w] = \int_{\Omega} (1 + |\nabla w|^2)^{\frac{1}{2}} dx.$$

Iremos determinar nos cálculos abaixo a equação de Euler-Lagrange do funcional dado acima.

Como $L(p, z, x) = (1 + |p|^2)^{\frac{1}{2}}$, temos que:

$$\begin{cases} L_{p_i} = \frac{\partial}{\partial p_i} \left((1 + |p|^2)^{\frac{1}{2}} \right) = \frac{p_i}{(1 + |p|^2)^{\frac{1}{2}}}; \\ L_z = \frac{\partial}{\partial z} \left((1 + |p|^2)^{\frac{1}{2}} \right) = 0. \end{cases}$$

Assim,

$$L_{p_i} = \frac{w_{x_i}}{(1 + |\nabla w|^2)^{\frac{1}{2}}}. \quad (2.13)$$

Desta forma, considerando que a equação de Euler-Lagrange é dada por:

$$- \sum_{i=1}^n (L_{p_i}(\nabla w, w, x))_{x_i} + L_z(\nabla w, w, x) = 0 \text{ em } \Omega$$

temos que a equação de Euler-Lagrange associada ao funcional $I[w]$ será dada por:

$$- \sum_{i=1}^n \left(\frac{w_{x_i}}{(1 + |\nabla w|^2)^{\frac{1}{2}}} \right)_{x_i} + 0 = 0 \text{ em } \Omega$$

ou, equivalentemente,

$$- \sum_{i=1}^n \left(\frac{w_{x_i}}{(1 + |\nabla w|^2)^{\frac{1}{2}}} \right)_{x_i} = 0 \text{ em } \partial\Omega.$$

Outros exemplos pode ser encontrados em [17].

2.2.2 Segunda variação

Continuando com a análise feita na subseção anterior, onde obtivemos a primeira variação de $I[\cdot]$, determinaremos agora a segunda variação deste funcional para a função u . Esta segunda variação será determinada considerando que já temos u como minimizador de $I[\cdot]$, ou seja,

$$i'' \geq 0.$$

Como definimos em (2.7), temos que $i(\cdot)$ é dado por:

$$i(\tau) := I[u + \tau v] \text{ com } \tau \in \mathbb{R},$$

onde $v \in C_0^\infty(\Omega)$ e $u = g$ em $\partial\Omega$. Assim, de (2.11) podemos calcular $i''(\tau)$, que será dado por:

$$i''(\tau) = \int_{\Omega} \sum_{i,j=1}^n L_{p_i p_j}(\nabla u + \tau \nabla v, u + \tau v, x) v_{x_i} v_{x_j} + 2 \sum_{i=1}^n L_{p_i z}(\nabla u + \tau \nabla v, u + \tau v, x) v_{x_i} v + L_{zz}(\nabla u + \tau \nabla v, u + \tau v, x) v^2 dx.$$

Fazendo $\tau = 0$ nesta última igualdade, temos que:

$$0 \leq i''(0) = \int_{\Omega} \sum_{i,j=1}^n L_{p_i p_j}(\nabla u, u, x) v_{x_i} v_{x_j} + 2 \sum_{i=1}^n L_{p_i z}(\nabla u, u, x) v_{x_i} v + L_{zz}(\nabla u, u, x) v^2 dx, \quad (2.14)$$

sendo esta desigualdade válida para todas as funções teste $v \in C_0^\infty(\Omega)$.

Notemos que a desigualdade (2.14) também é válida para qualquer função v lipschitziana contínua que se anula em $\partial\Omega$. Fixando $\xi \in \mathbb{R}^n$ e definindo:

$$v(x) = \epsilon \rho\left(\frac{x \cdot \xi}{\epsilon}\right) \zeta(x), \quad x \in \Omega \quad (2.15)$$

sendo $\zeta \in C_0^\infty(\Omega)$ e $\rho : \mathbb{R} \rightarrow \mathbb{R}$ uma função periódica dada por:

$$\rho(x) := \begin{cases} x, & \text{se } 0 \leq x \leq \frac{1}{2}, \\ 1 - x, & \text{se } \frac{1}{2} \leq x \leq 1, \end{cases} \quad (2.16)$$

com $\rho(x+1) = \rho(x)$, obtemos,

$$|\rho'| = 1 \quad \text{quase todo ponto.} \quad (2.17)$$

Notemos ainda que

$$v_{x_i}(x) = \rho' \left(\frac{x \cdot \xi}{\epsilon} \right) \xi_i \zeta + O(\epsilon),$$

com $\epsilon \rightarrow 0$. Substituindo (2.15) em (2.14) obtemos a seguinte desigualdade:

$$0 \leq \int_{\Omega} \sum_{i,j=1}^n L_{p_i p_j}(\nabla u, u, x) (\rho')^2 \xi_i \xi_j \zeta^2 dx + O(\epsilon).$$

Usando agora (2.17) e fazendo $\epsilon \rightarrow 0$ obtemos:

$$0 \leq \int_{\Omega} \sum_{i,j=1}^n L_{p_i p_j}(\nabla u, u, x) \xi_i \xi_j \zeta^2 dx. \quad (2.18)$$

Com isso, a desigualdade acima é válida para toda função $\zeta \in C_0^\infty(\Omega)$, então deduzimos que

$$\sum_{i,j=1}^n L_{p_i p_j}(\nabla u, u, x) \xi_i \xi_j \geq 0, \quad \xi \in \mathbb{R}^n, x \in \Omega. \quad (2.19)$$

Na próxima seção veremos como determinar as condições de contorno (2.6) quando estas não são impostas.

2.3 Condições de contorno

Nesta seção, analisaremos as condições de contorno para o problema (2.5), isto é, quando escolhermos o espaço V , das funções admissíveis para o funcional I em (2.5), exigimos que toda função $v \in V$ satisfaça as condições de contorno (2.6). Isto sugere que, todo ponto estacionário do funcional I também satisfaça tais condições.

O que devemos nos perguntar neste momento é o seguinte: caso não seja imposta nenhuma condição de contorno para as funções admissíveis, qual será a condição de contorno que um ponto estacionário deverá satisfazer? Para responder tal questão, faremos uma análise sobre as condições de contorno essenciais e naturais que definimos a seguir.

Chamamos de *condições de contorno naturais* para o funcional I , as condições não impostas para as funções admissíveis, e de *condições de contorno essenciais*, as condições impostas para o funcional I , como feito em (2.6). Assim, se analisarmos o funcional $\int_a^b L(v'(x), v(x), x) dx$ com $v \in V = C^2[a, b]$, $v(a) = \alpha$ e $v(b) = \beta$, sendo as condições de contorno essenciais para as

funções admissíveis $v \in V$, temos que $\eta(a) = 0$ e $\eta(b) = 0$ serão as condições de contorno para as funções teste η e, $\left. \frac{\partial L}{\partial v'} \right|_{x=a} = 0$ e $\left. \frac{\partial L}{\partial v'} \right|_{x=b} = 0$ serão as condições de contorno naturais para o ponto estacionário.

Agora, observando as condições acima, vemos que se uma condição de contorno essencial é imposta para as funções admissíveis, então as funções teste também deverão satisfazer as condições impostas correspondentes. Podemos verificar também que se alguma condição de contorno essencial não for imposta às funções admissíveis, então um ponto estacionário deverá satisfazer a condição de contorno natural correspondente. Esta verificação e mais exemplos sobre essas condições podem ser encontradas em [4].

Vejam agora um exemplo sobre essas condições. Se apenas a segunda condição de contorno essencial for imposta, temos que:

$$V = \{v \in C^2[a, b] : v(b) = \beta\} \text{ e } \tilde{V} = \{\eta \in C^2[a, b] : \eta(b) = 0\}, \quad (2.20)$$

e então

$$\left[\frac{\partial L}{\partial v'} \eta \right] \Big|_a^b = 0, \quad \forall \eta \in \tilde{V},$$

o que implica

$$\eta(a) \left[\frac{\partial L}{\partial v'} \right] \Big|_{x=a} = 0, \quad \forall \eta \in \tilde{V}.$$

Assim, como existem $\eta \in \tilde{V}$ tais que $\eta(a) \neq 0$, temos que:

$$\left[\frac{\partial L}{\partial v'} \right] \Big|_{x=a} = 0. \quad (2.21)$$

Portanto, (2.21) é a condição de contorno natural correspondente à condição de contorno essencial omitida.

Na próxima seção identificaremos condições do Lagrangeano L que garantem a existência de um mínimo para o funcional $I[\cdot]$, pelo menos em um espaço de Sobolev apropriado.

2.4 Existência de minimizadores

Nesta seção apresentaremos idéias para determinar quando o funcional

$$I[w] = \int_{\Omega} L(\nabla w(x), w(x), x) \, dx, \quad (2.22)$$

definido para funções apropriadas $w : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e, satisfazendo a condição de contorno

$$w = g \text{ em } \partial\Omega, \quad (2.23)$$

deve ter um mínimo.

2.4.1 Coercividade

Sabemos que nem toda função suave $f : \mathbb{R} \rightarrow \mathbb{R}$ precisa necessariamente atingir seu ínfimo, como por exemplo as seguintes funções:

$$f(x) = e^x \text{ ou } f(x) = \frac{1}{1+x^2}.$$

Estas funções sugerem, em geral, que sejam necessárias algumas hipóteses de forma que possamos controlar o funcional $I[w]$ para determinadas funções w . Uma boa maneira de se obter tal controle seria a hipótese de que $I[w]$ cresça rapidamente quando $|w| \rightarrow \infty$.

Especificamente, vamos assumir que:

$$1 < q < \infty \quad (2.24)$$

seja fixo, e que existem constantes $\alpha > 0$, $\beta \geq 0$, tais que:

$$L(p, z, x) \geq \alpha|p|^q - \beta, \quad (2.25)$$

para todo $p \in \mathbb{R}^n$, $z \in \mathbb{R}$ e $x \in \Omega$.

Como consequência temos que:

$$I[w] \geq \alpha \|\nabla w\|_{L^q(\Omega)}^q - \gamma, \quad (2.26)$$

para $\gamma = \beta|\Omega|$ e alguma constante $\alpha > 0$. Portanto, $I[w] \rightarrow \infty$ quando $\|\nabla w\|_{L^q(\Omega)}^q \rightarrow \infty$. A condição dada em (2.26) é chamada de *condição de coersividade* do funcional I .

Agora, voltando nossa atenção para o objetivo de encontrar funções minimizantes para o funcional I , observamos que, da desigualdade dada em (2.26), nos parece razoável definir $I[w]$ não apenas para funções suaves w , mas também para funções $w \in W^{1,q}(\Omega)$ que satisfaçam a condição de contorno dado em (2.23), no sentido do seu traço.

Ao ampliarmos a classe de funções w para as quais o funcional $I[w]$ está definido, ampliamos também o número de candidatos as funções minimizantes de $I[w]$, portanto agora, o conjunto das funções admissíveis w , denotado por V será dado por:

$$V := \{w \in W^{1,q}(\Omega); w = g \text{ em } \partial\Omega \text{ no sentido do seu traço}\}. \quad (2.27)$$

2.4.2 Semi-continuidade inferior

Nesta subseção observaremos que, embora uma função contínua $f : \mathbb{R} \rightarrow \mathbb{R}$ satisfaça a condição de coercividade (2.26) e atinja seu ínfimo, o funcional $I[\cdot]$ pode em geral não estar satisfazendo tal condição. Para entendermos melhor o que foi descrito vamos definir

$$m = \inf_{w \in V} I[w], \quad (2.28)$$

e tomando funções $u_k \in V$ com $(k = 1, 2, \dots)$ temos que

$$I[u_k] \rightarrow m \text{ quando } k \rightarrow \infty. \quad (2.29)$$

A sequência $\{u_k\}_{k=1}^\infty$ é chamada de *sequência minimizante* do funcional.

Neste momento torna-se interessante mostrarmos que alguma subsequência de $\{u_k\}_{k=1}^\infty$ converge, de fato, para uma função minimizante. Para isto, deveremos usar algum tipo de compacidade, e isto é um problema, pois o espaço $W^{1,q}(\Omega)$ é de dimensão infinita.

Na verdade, se utilizarmos a desigualdade da coercividade dada em (2.26) será possível mostrar apenas que a sequência minimizante pertence a um subconjunto limitado de $W^{1,q}(\Omega)$, porém isso não implicará na existência de uma subsequência que converge em $W^{1,q}(\Omega)$.

Contudo, vamos dirigir nossa atenção para a compacidade fraca dos espaços reflexivos de Banach, e o fato de que $L^q(\Omega)$ é reflexivo, nos permite concluir a existência de uma subsequência $\{u_{k_j}\}_{j=1}^\infty \subset \{u_k\}_{k=1}^\infty$ e de uma função $u \in W^{1,q}(\Omega)$ tal que:

$$\begin{cases} u_{k_j} \rightharpoonup u \text{ fracamente em } L^q(\Omega) \\ \nabla u_{k_j} \rightharpoonup \nabla u \text{ fracamente em } L^q(\Omega). \end{cases} \quad (2.30)$$

Daqui por diante, a condição dada acima será abreviada, e diremos apenas que

$$u_{k_j} \rightharpoonup u \text{ fracamente em } W^{1,q}(\Omega). \quad (2.31)$$

Além disso, $u = g$ em $\partial\Omega$ se verifica no sentido do traço, e então $u \in V$.

Como consequência de usar a topologia fraca, recuperamos a compacidade necessária na desigualdade (2.26) para deduzirmos (2.31) para uma subsequência apropriada. Com tudo, nos surge uma outra dificuldade visto que, praticamente em todos os casos, o funcional I não é contínuo com relação a convergência fraca. Em outras palavras, não podemos deduzir de (2.29) e (2.31) que:

$$I[u] = \lim_{j \rightarrow \infty} I[u_{k_j}], \quad (2.32)$$

e assim, u é um minimizador.

O problema que temos é que $\nabla u_{k_j} \rightharpoonup \nabla u$ não implica que $\nabla u_{k_j} \rightarrow \nabla u$ (q.t.p), e é muito possível que, por exemplo, os gradientes ∇u_{k_j} embora estejam limitados em $L^q(\Omega)$, possam estar oscilando muito quando $k_j \rightarrow \infty$.

O que nos salva e pode garantir que u venha a ser uma função minimizante, é observarmos que o funcional $I[u]$ não precisa necessariamente satisfazer (2.32), mas sim que

$$I[u] \leq \liminf_{j \rightarrow \infty} I[u_{k_j}], \quad (2.33)$$

e então de (2.29) podemos deduzir que $I[u] \leq m$. Porém, de (2.28) temos que $m \leq I[u]$. Assim, u é de fato um minimizador.

Definição 2.7 Dizemos que o funcional I é de **fraca semi-continuidade inferior** em $W^{1,q}(\Omega)$, se:

$$I[u] \leq \liminf_{k \rightarrow \infty} I[u_k],$$

sempre que $u_k \rightharpoonup u$ fracamente em $W^{1,q}(\Omega)$.

Nosso próximo passo, portanto, será identificar as condições razoáveis sobre a não-linearidade da função L para que possamos garantir que $I[\cdot]$ seja fracamente semi-contínuo inferiormente.

2.4.3 Convexidade

A análise que iremos fazer nesta subseção será baseada na segunda variação obtida na seção 2.2.2, onde encontramos a seguinte desigualdade:

$$\sum_{i,j=1}^n L_{p_i} L_{p_j}(\nabla u(x), u(x), x) \xi_i \xi_j \geq 0 \quad (\xi \in \mathbb{R}^n, x \in \Omega),$$

que é válida como uma condição necessária sempre que u for uma função minimizante com certa suavidade. Esta desigualdade nos sugere que é razoável assumirmos que a função L seja convexa, pelo menos inicialmente.

Teorema 2.3 (*Fraca semi-continuidade inferior*) (Teorema 1, parte 3 de [17]) Suponhamos que L seja limitada inferiormente e que a aplicação

$$p \mapsto L(p, z, x)$$

seja convexa (ver Definição 1.9) para cada $z \in \mathbb{R}$, $x \in \Omega$. Então o funcional I é fracamente semi-contínuo inferiormente em $W^{1,q}(\Omega)$.

Demonstração

A demonstração deste teorema pode ser encontrada em [17].

□

Teorema 2.4 (*Existência do minimizador*) (Teorema 2, parte 3 de [17]) Suponhamos que L satisfaça a desigualdade de coercividade dada em (2.26) e seja convexa em relação a variável p . Suponhamos também que o conjunto admissível V seja não vazio.

Então existe pelo menos uma função $u \in V$ tal que:

$$I[u] = \min_{w \in V} I[w].$$

Demonstração

A demonstração deste teorema pode ser encontrada em [17].

□

Agora que conseguimos estabelecer a existência de uma função minimizante u , devemos como de praxe, tentar impor certas condições sobre o funcional L , com o intuito de garantirmos sua unicidade. Para isto, vamos supor que:

$$L = L(p, x) \text{ não dependa de } z, \quad (2.34)$$

e que exista um $\theta > 0$ tal que

$$\sum_{i,j=1}^n L_{p_i p_j}(p, x) \xi_i \xi_j \geq \theta |\xi|^2. \quad (2.35)$$

Temos que a condição (2.35) nos diz que a aplicação $p \rightarrow L(p, x)$ é uniformemente convexa para cada $x \in \Omega$.

Teorema 2.5 (Unicidade do minimizador) (Teorema 3, parte 3 de [17]) *Suponhamos que as condições (2.34) e (2.35) sejam satisfeitas. Então temos que a função minimizante $u \in V$, de $I[\cdot]$ é única.*

Demonstração:

Suponhamos que $u, \tilde{u} \in V$ sejam minimizadores de $I[\cdot]$ e definimos $v := \frac{u+\tilde{u}}{2} \in V$. Assim, temos que

$$I[v] \leq \frac{I[u] + I[\tilde{u}]}{2}, \quad (2.36)$$

com a desigualdade estrita a menos que $u = \tilde{u}$.

Para provarmos esta desigualdade, observemos que a partir da suposição da convexidade uniforme temos que:

$$L(p, x) \geq L(q, x) + D_p L(q, x) \cdot (p - q) + \frac{\theta}{2} |p - q|^2 \quad (x \in \Omega, p, q \in \mathbb{R}^n). \quad (2.37)$$

Agora, definindo $q = \frac{Du + D\tilde{u}}{2}$, $p = Du$ em (2.37), em Ω , temos:

$$I[v] + \int_{\Omega} D_p L\left(\frac{Du + D\tilde{u}}{2}, x\right) \cdot \left(\frac{Du - D\tilde{u}}{2}\right) dx + \frac{\theta}{8} \int_{\Omega} |Du - D\tilde{u}|^2 dx \leq I[u]. \quad (2.38)$$

De maneira análoga, seja $q = \frac{Du + D\tilde{u}}{2}$, $p = D\tilde{u}$ em (2.37) e integrando em Ω temos:

$$I[v] + \int_{\Omega} D_p L\left(\frac{Du + D\tilde{u}}{2}, x\right) \cdot \left(\frac{D\tilde{u} - Du}{2}\right) dx + \frac{\theta}{8} \int_{\Omega} |Du - D\tilde{u}|^2 dx \leq I[\tilde{u}]. \quad (2.39)$$

Somando (2.38) com (2.39) e dividindo por 2, temos que:

$$I[v] + \frac{\theta}{8} \int_{\Omega} |Du - D\tilde{u}|^2 dx \leq \frac{I[u] + I[\tilde{u}]}{2}. \quad (2.40)$$

Isto prova (2.36).

Agora, como $I[u] = I[\tilde{u}] = \min_{w \in \Omega} I[w] \leq I[v]$, deduzimos que $Du = D\tilde{u}$ em Ω . Como $u = \tilde{u} = g$ em $\partial\Omega$ no sentido do traço, segue que $u = \tilde{u}$.

□

Maiores detalhes sobre a equação de Euler-Lagrange, exemplos e sua generalização podem ser encontrados em [4] e [17].

Capítulo 3

Variação Total em Imagens

3.1 Introdução

Neste Capítulo apresentaremos o problema de restauração de imagens baseado em Variação Total (TV), abordando principalmente nos trabalhos pioneiros de Rudin, Osher, Fatemi (ROF) em 1992 [30], e Mumford e Shah (MS) em 1989 [28]. Apresentaremos também algumas aplicações de Variação Total em problemas de processamento de imagens.

Modelos de restauração de imagens com base na Variação Total (TV) têm sido muito populares desde a sua introdução por (ROF) e (MS). Este é um dos problemas fundamentais no processamento de imagem digital e os modelos variacionais têm sido extremamente bem sucedidos em uma grande variedade de problemas deste tipo, e ainda continuam a ser uma das áreas mais ativas de pesquisa. O problema mais comum de restauração de imagens é, talvez, remoção de ruídos. Ele é uma importante tarefa no processamento de imagem, como um processo em si, e como um componente em outros processos. A principal propriedade de um modelo com tal objetivo, é que ele irá remover o ruído, preservando as bordas da imagem. Tradicionalmente, utiliza-se modelos lineares para resolução de tal problema, e uma das grandes vantagens desses modelos de remoção de ruído é a velocidade. Porém, um inconveniente para tais modelos é que eles não são capazes de preservar as bordas de uma forma satisfatória. Já modelos não-lineares, podem lidar com as bordas de uma forma mais eficiente. Um modelo popular para a filtragem não-linear da imagem é a Variação Total (TV), introduzido por Rudin, Osher e Fatemi [30].

Na prática, os métodos mais utilizados para estimar ruído são baseados em mínimos quadrados. A justificativa vem do argumento estatístico que, pelo menos a estimativa pelos mínimos quadrados é melhor sobre o conjunto de todas as imagens possíveis, e este procedimento é dado

pela norma L^2 . No entanto existem discussões garantindo que a norma adequada quando se trabalha com imagens é a norma da variação total, e não a norma L^2 . Temos que normas de TV são, essencialmente, normas L^1 de derivadas, e portanto os autores mostram que procedimentos baseados nesta norma são mais apropriados para o objetivo de restauração de imagens. Temos também que o espaço das funções de variação total limitada (BV) (ver definição 1.29) desempenha um papel importante na estimativa precisa de descontinuidades nas soluções.

Em comparação com os métodos de mínimos quadrados, onde as soluções lineares são bem conhecidas e facilmente calculadas, a estimativa baseada na norma L^1 é não-linear e computacionalmente complexa. Recentemente, o assunto da estimativa de dados estatísticos pela norma L^1 tem recebido atenção redobrada.

Temos que a Variação Total pode ser definida para toda função $u \in L^1(\Omega)$ ($\Omega \subset \mathbb{R}^2$) como na definição 1.28, ou seja:

$$J(u) = \sup \left\{ \int_{\Omega} u(x) \operatorname{div} \xi(x) dx; \xi \in \mathcal{C}_c^1(\Omega; \mathbb{R}^2), |\xi(x)| \leq 1 \ \forall x \in \Omega \right\}. \quad (3.1)$$

A expressão acima nos diz que J é finito se e só se a distribuição Du de derivadas de u for uma medida de Radon finita em Ω (ver definição 1.15), e neste caso temos $J(u) = |Du|$. Agora, se u tem gradiente $\nabla u \in L^1(\Omega, \mathbb{R}^2)$, então $J(u) = \int_{\Omega} |\nabla u(x)| dx$, onde $|\nabla u| = |(u_x, u_y)| = \sqrt{u_x^2 + u_y^2}$.

Um outro problema fundamental no processamento de imagens é a Segmentação, que consiste em fragmentar uma imagem, em unidades homogêneas, considerando algumas de suas características intrínsecas, como por exemplo cor, textura, forma, contraste e etc. Tal problema foi abordado primeiramente pelos autores Mumford e Shah em [28], e será visto com maiores detalhes em uma das próximas seções. Dentre as técnicas de segmentação temos a **Segmentação Hard**, onde um ponto $x \in \Omega$, apresenta apenas duas opções, pertencer ou não a uma dada região, e também a **Segmentação Soft**, onde cada ponto apresenta grau de pertinência à cada região segmentada, e este é representado por uma função *fuzzy* (ver Definição 1.10).

3.2 Remoção de ruídos

Temos que modelos de restauração de imagens baseados em Variação Total (TV) foram introduzidos por Rudin, Osher e Fatemi (ROF) no seu trabalho pioneiro [30]. O modelo proposto pelos autores foi projetado com o objetivo explícito de preservação de descontinuidades abruptas (bordas) em uma imagem durante a remoção de ruído e detalhes indesejados. Com base em

experiências anteriores, os autores propuseram resolver o problema de remoção de ruídos minimizando a norma da variação total de uma solução estimada. Foi obtido assim, um algoritmo de minimização restrita, com um fator não-linear e dependente do tempo, onde as limitações são determinadas pelas estatísticas do ruído. Os métodos tradicionais tentavam reduzir ou remover o ruído antes das operações de transformação da imagem, e esta é a abordagem adotada em [30]. No entanto, a mesma idéia da minimização de TV pela norma L^1 pode ser usada para o desenvolvimento de algoritmos híbridos, que combinam a filtragem de ruído com outras tarefas de processamento de imagem.

Temos que se o ruído é aditivo, uma expressão genérica para o processo de remoção de ruídos é:

$$I = u + \eta,$$

onde u denota a imagem original, η o ruído e $I : \Omega \rightarrow \mathbb{R}$ representa a imagem observada.

Sejam então, σ o desvio padrão do ruído na imagem I , Ω o domínio da imagem u e $|\Omega|$ a área de Ω . Temos assim o seguinte modelo proposto pelos autores Rudin, Osher e Fatemi (ROF):

$$\min_{u \in BV} \int_{\Omega} |\nabla u|, \text{ sujeito a } \|u - I\|_2^2 \leq |\Omega| \sigma^2, \quad (3.2)$$

onde $\int_{\Omega} |\nabla u|$ é a *variação total* (TV) de u que pode ser denotado por $TV[u]$, $|\nabla u| = |(u_x, u_y)| = \sqrt{u_x^2 + u_y^2}$, $I \in L^2(\Omega)$ e $\|\cdot\|$ representa a norma usual em $L^2(\Omega)$. O conjunto $BV(\Omega)$ denota a coleção de todas as funções em $L^1(\Omega)$ com variação total finita, como definido no Capítulo 1, ou seja:

$$BV(\Omega) \equiv \{u \mid u \in L^1(\Omega), TV[u] < \infty\}. \quad (3.3)$$

Como $W^{1,1}(\Omega) = \{u \in L^1(\Omega) \mid u_{x_i} \in L^1(\Omega)\}$, temos que $W^{1,1}(\Omega) \subseteq BV(\Omega) \subseteq L^1(\Omega)$, e assim, a definição de TV em (3.1) é equivalente a função objetivo (3.2), quando $u \in W^{1,1}(\Omega)$ (ver definição 1.21).

Assim, temos que o modelo ROF restrito (3.2) é equivalente ao seguinte modelo irrestrito:

$$\min_{u \in BV} \int_{\Omega} |\nabla u| + \frac{\beta}{2} \|u - I\|_2^2, \quad (3.4)$$

para o multiplicador de Lagrange β apropriado. Este modelo foi estudado de forma extensiva, como por exemplo, nas referências [2] e [8].

Temos que, tanto os autores ROF quanto pesquisadores posteriores, centraram seus objetivos na resolução do modelo sem restrições (3.4) ao invés do modelo original restrito (3.2), pois a otimização irrestrita geralmente é mais fácil de resolver. O modelo ROF original foi introduzido para resolver problemas de remoção de ruídos em imagem, mas a metodologia pode ser estendida naturalmente para restaurar imagem borrada e outras, bastando incluir o operador núcleo, do seguinte modo:

$$\min_{u \in BV} \int_{\Omega} |\nabla u| + \frac{\beta}{2} \|Ku - I\|_2^2. \quad (3.5)$$

Aqui, K é um dado operador linear, conhecido como operador núcleo, e todos os outros termos são definidos como em (3.2). Neste modelo, I é assumida como sendo a soma de um ruído gaussiano η e uma imagem borrada Ku , resultante de um operador linear K agindo na imagem original u , ou seja, $I = Ku + \eta$.

Baseado neste mesmo problema, os autores L. Alvarez, P.L. Lions e J.M. Morel [3], propuseram minimizar o seguinte funcional para remoção de ruídos de uma dada imagem:

$$\min_{u \in BV} \frac{1}{2} \int_{\Omega} |\nabla u|^2 + \frac{\beta}{2} \int_{\Omega} (u - I)^2,$$

onde u representa a imagem original e I a imagem observada com ruídos. Este funcional, apesar de remover melhor os ruídos, não apresenta boa reconstrução da imagem, pois não possui nenhum mecanismo para preservação das bordas.

Posteriormente aos trabalhos [3] e [30], os autores Chen, Levine e Rao [13] propuseram um novo funcional a ser minimizado para resolução do mesmo problema, porém agora o expoente da norma do gradiente de u não é mais uma constante, e sim uma variável denominada de $q(x)$, onde $1 < q(x) \leq 2$. A vantagem deste modelo é que o funcional utiliza a norma L^1 para pontos próximos da borda e a norma L^2 para pontos longe da borda, conseguindo com isso preservar as bordas da imagem e remover melhor o ruído.

O funcional proposto a ser minimizado é:

$$\min_{u \in BV \cap L^2(\Omega)} \int_{\Omega} \phi(x, \nabla u) + \frac{\beta}{2} (u - I)^2$$

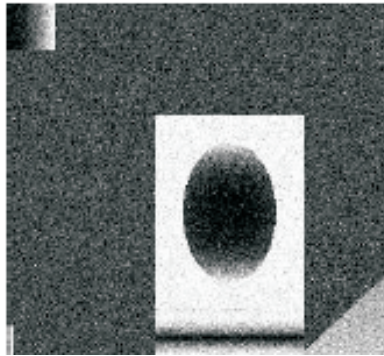
sendo,

$$\phi(x, \nabla u) := \begin{cases} \frac{1}{q(x)} |\nabla u|^{q(x)}, & \text{se } |\nabla u| \leq \delta, \\ |\nabla u| - \frac{\beta q(x) - \beta q(x)}{q(x)}, & \text{se } |\nabla u| > \delta, \end{cases}$$

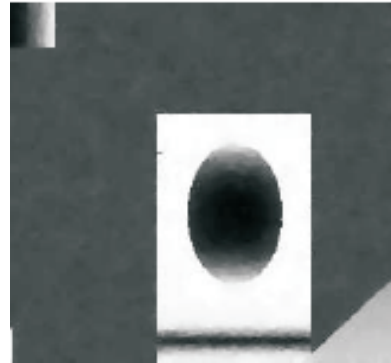
onde, $\delta > 0$ é fixo.

Vejamos agora, a aplicação do funcional acima em algumas imagens que apresentam ruídos:

Exemplo 1- Em uma imagem suave com ruído, composta por figuras geométricas, foi aplicado o método proposto em [13] obtendo assim a imagem limpa, preservando as fronteiras e sem introdução de bordas falsas.



(a) Imagem com ruído.



(b) Resultado da remoção de ruído.

Figura 3.1: Problema de Eliminação de Ruído 1. Imagens retiradas de [13]

Exemplo 2- Em uma imagem suave com ruído aditivo gaussiano foi aplicado o método proposto em [13], obtendo assim uma imagem limpa que preserva suas bordas.



(a) Imagem com ruído.



(b) Resultado da remoção de ruído.

Figura 3.2: Problema de Eliminação de Ruído 2. Imagens retiradas de [13]

Ao longo dos anos, o modelo ROF foi estendido para solução de muitos outros problemas de processamento de imagens, incluindo retoque digital, deblurring e decomposição em geometria e textura. Veremos alguns desses problemas em seções subsequentes.

3.3 Segmentação de imagens

Um outro trabalho pioneiro no uso de Variação Total (TV) para restauração de imagens foi realizado pelos autores Mumford e Shah em 1989 [28], onde propuseram um método de segmentação de imagens, baseado em um funcional de energia. A idéia é que a minimização de tal funcional dê como resultado uma imagem próxima à primeira, composta de várias regiões, onde cada região apresenta intensidade quase constante. A dificuldade apresentada na minimização deste funcional é que ele envolve duas variáveis: a função intensidade e o conjunto Γ de bordas, que é o conjunto de descontinuidades da função intensidade.

Em visão computacional, segmentação se refere ao processo de dividir uma imagem digital em múltiplas regiões ou objetos, com o objetivo de simplificar e/ou mudar a representação de uma imagem, para facilitar a sua análise. Segmentação de imagens é tipicamente usada para localizar objetos e formas (linhas, curvas, etc) em imagens. O resultado da segmentação é um conjunto de regiões/objetos ou um conjunto de contornos extraídos da imagem, onde regiões adjacentes devem possuir diferenças significativas com respeito a uma mesma característica.

Como vimos no início deste capítulo, dentre as técnicas de segmentação temos a segmentação hard. O funcional proposto pelos autores Mumford e Shah a ser minimizado para resolução deste problema é um exemplo de tal segmentação e é dado por:

$$E(u, \Gamma) = \int_{\Omega} \|u - I\|^2 dx + \alpha \int_{\Omega - \Gamma} |\nabla u|^2 dx + \nu |\Gamma|, \quad (3.6)$$

onde $\Gamma \subset \Omega$ é o conjunto de descontinuidades, u é a função diferenciável intensidade e α é uma constante não-negativa. O terceiro termo é o tamanho da borda, Ω é um conjunto aberto e limitado de $\mathbb{R}^n$, $n=2,3$ que denota o domínio da imagem u , e I é a imagem inicial.

Observações:

1. o primeiro termo mede quanto u se aproxima de I ;
2. o segundo termo (termo de suavização) calcula a variação mínima de u dentro de cada região, sem considerar a fronteira;

3. no terceiro termo, Γ deve ser o menor possível;
4. α serve para balancear os termos de suavização e fidelidade;

Em um trabalho dos mesmos autores, foi provado a seguinte conjectura para o funcional (3.6):

Conjectura: Existe um minimizador de E tal que as bordas (conjunto Γ de descontinuidade) é a união de conjuntos finitos de curvas $C^{1,1}$, onde $C^{k,\gamma}$ é o espaço de funções k vezes continuamente diferenciável, cuja k -ésima derivada parcial pertence a $C^{k,\gamma}(\Omega)$ (ver Definição 1.16).

Agora, querendo provar a existência de soluções para o funcional (3.6), os autores mostraram que o par (u, Γ) que é solução de (3.6), satisfaz a seguinte conjectura:

Conjectura de Mumford-Shah:

1. Γ consiste de um número finito de curvas $\gamma_i (\in C^{1,1})$, união com $\partial\Omega$ e cada uma das suas extremidades;
2. u é C^1 em cada componente conexa (ver Definição 1.8) de $\Omega - \Gamma$.

Uma outra maneira de resolver problemas de segmentação de imagens através da segmentação hard, pode ser considerando o modelo denominado Competição entre Regiões, que detalhamos na próxima subseção.

3.3.1 Competição entre Regiões

O método variacional criado para resolver o problema de segmentação de imagens, denominado Competição entre Regiões [39], foi desenvolvido por Zhu e Yuille em 1995, e consiste basicamente em segmentar uma imagem inicial fazendo com que regiões desta imagem passem a competir entre si através de seus pontos, e esses por sua vez se juntam às regiões que lhes são estatisticamente mais semelhantes.

Neste modelo, uma região de uma dada imagem é dita ser homogênea se os valores de intensidade de níveis de cinza forem gerados por uma distribuição de probabilidade especificada previamente e dada por $P(I(x, y)|\alpha)$, onde α representa o conjunto de parâmetros da distribuição. No caso da distribuição Gaussiana, com média μ e desvio padrão σ , o conjunto de parâmetros de uma região da imagem será representado por $\alpha(\mu, \sigma)$, e a distribuição de probabilidade Gaussiana será calculada da seguinte maneira:

$$P(I(x, y)|(\mu, \sigma)) = \frac{1}{\alpha\sqrt{2\pi}} \left(-\frac{(I(x, y) - \mu)^2}{2\sigma^2} \right). \quad (3.7)$$

Vamos supor que toda imagem de domínio Ω seja inicialmente dividida em M regiões seccionalmente homogêneas e denominadas por Ω_i , $i = 1, 2, \dots, M$. Pelo critério de Bayes/MDL (ver [39]), o funcional de energia para uma escolha adequada de famílias de distribuição de probabilidade é dado por:

$$E[\Gamma, \{\alpha_i\}] = \sum_{i=1}^M \left\{ \frac{\mu}{2} \int_{\partial\Omega_i} ds - \log P(\{I(x, y) : (x, y) \in \Omega_i\}|\alpha_i) + \lambda \right\}, \quad (3.8)$$

onde Ω é o domínio da imagem, $\Omega = \bigcup_{i=1}^M \Omega_i$, $\Omega_i \cap \Omega_j = \emptyset$ se $i \neq j$, $\partial\Omega_i$ é a fronteira da região Ω_i e $\Gamma = \bigcup_{i=1}^M \Gamma_i$ são bordas da imagem com $\Gamma_i = \partial\Omega_i$.

O primeiro termo da equação (3.8) descreve o comprimento da borda da imagem I , já o segundo termo é a soma do custo para calcular a intensidade de cada ponto (x, y) dentro da região Ω_i , de acordo com a distribuição de probabilidade P . O fator λ é usado para descrever a distribuição de probabilidade. Temos então que se $v = (x, y) \in \mathbb{R}^2$ é um ponto de I , então $P_i(v|\{\alpha_i\})$ descreve a probabilidade de v pertencer a região Ω_i .

O funcional (3.8) é minimizado em duas etapas. Na primeira etapa, fixa-se Γ , ou seja, deixa-se fixo Ω_i e $\{I(x, y); \forall (x, y) \in \Omega_i\}$, e os parâmetros $\{\alpha_i\}$ são estimados pela maximização das probabilidades condicionais. Na segunda etapa, fixa-se $\{\alpha_i\}$ e aplicamos o método da descida íngreme com relação a Γ , movimentando o contorno Γ para cada ponto $v = (x, y) \in I$ de acordo com a equação:

$$\frac{\partial v}{\partial t} = -\frac{\partial E[\Gamma, \{\alpha_i\}]}{\partial v} = \sum_{k \in Q(v)} \left(-\frac{\mu}{2} \kappa_{k(v)} n_{k(v)} + \log P(I(v)|\alpha_k) n_{k(v)} \right) \quad (3.9)$$

onde $Q(v) = \{k | v \in \Gamma_k\}$, $\kappa_{k(v)}$ é a curvatura de Γ_k no ponto v , e $n_{k(v)}$ é o vetor normal de Γ_k também no ponto v . Pela equação anterior, temos que existem duas forças apontando para a direção da normal: o primeiro termo representando a força de suavização e o segundo a força estatística, como ilustrado na figura abaixo:

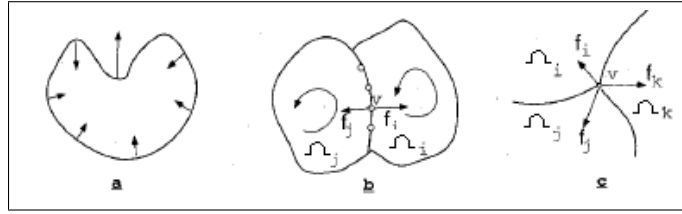


Figura 3.3: Forças agindo no contorno: (a) Força de suavização (b) Força estatística em um ponto de fronteira (c) Força estatística em um ponto de junção de regiões.

Por exemplo, a figura central ilustra o vetor v comum à bordas das regiões Ω_i e Ω_j . Como as curvas Γ_i e Γ_j tem vetores normais inversos em v então $n_i = -n_j$ e $\kappa_i n_i = \kappa_j n_j$. Portanto a equação (3.9) pode ser reescrita como:

$$\begin{aligned} \frac{\partial v}{\partial t} &= -\mu \kappa_i(v) n_i(v) + (\log P(I_v | \alpha_i) - \log P(I_v | \alpha_j)) n_i(v) \\ &= -\mu \kappa_i(v) n_i(v) + \log \left(\frac{P(I_v | \alpha_i)}{P(I_v | \alpha_j)} \right) n_i(v). \end{aligned} \quad (3.10)$$

Se α_i e α_j são os parâmetros das regiões Ω_i e Ω_j , respectivamente e se $P(I_v | \alpha_i) > P(I_v | \alpha_j)$, isto é, se a intensidade em v se adapta melhor à distribuição da região Ω_i em relação a região Ω_j , então a borda se move em direção a n_i .

Maiores detalhes do algoritmo pode ser encontrado em [39].

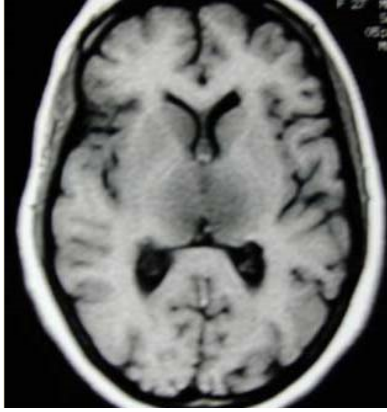
Neste modelo tem-se apenas a garantia da convergência para um mínimo local, e a desvantagem apresentada é a alta dependência da amostragem realizada em cada uma das regiões.

Uma outra maneira de resolver o problema de segmentação é através da segmentação soft, em particular, por um modelo conhecido como Competição de Regiões Fuzzy [27], proposto pelos autores Mory e Ardon, baseado no trabalho de Zhu e Yuille [39] e que veremos alguns detalhes na próxima subseção.

3.3.2 Competição entre Regiões Fuzzy

É comum, em imagens do dia-a-dia, nos depararmos com situações onde não conseguimos determinar a localização exata das bordas de um objeto, como exemplo, a Figura 3.4 apresenta situações deste tipo. Neste caso, na Figura 3.4(a) não é possível determinar com precisão a fronteira entre o pássaro e o fundo, já a Figura 3.4(b), exibe uma tomografia cerebral sem

bordas nítidas, e por fim, a Figura 3.4(c) mostra uma foto tirada da região Amazônia por um satélite onde não conseguimos identificar as regiões de florestas.



(a) Tomografia cerebral (*Retida de www.cerebromente.org.br*).



(b) Praia (*Retirada de [31]*).

Figura 3.4: Imagens sem fronteiras definidas.

Com isso, para segmentar imagens deste tipo vem sido desenvolvidos modelos vantajosos de segmentação *soft*, e buscando alternativas para resolver este problema, os autores Mory e Ardon propuseram um novo modelo variacional de segmentação para estas imagens baseado no modelo de Competição de Regiões. Este novo modelo variacional é denominado Competição entre Regiões Fuzzy [27], e a única diferença entre este algoritmo e o modelo apresentado no início deste capítulo é a existência de uma função de pertinência definida sobre um conjunto limitado no intervalo $[0,1]$. O método baseia-se em dividir a imagem em duas regiões através de suas distribuições de intensidade.

Quando particionamos uma imagem I em duas regiões, o problema de minimização pode ser escrito por:

$$\min_{\substack{\Omega_1 \subset \Omega \\ \alpha \in A}} \left\{ F_0(\Omega_1, \alpha) = \int_{\Gamma} g(\Gamma(s)) ds + \int_{\Omega_1} r_1^{\alpha_1}(\mathbf{x}) d\mathbf{x} + \int_{\Omega \setminus \Omega_1} r_2^{\alpha_2}(\mathbf{x}) d\mathbf{x} \right\}, \quad (3.11)$$

com $\mathbf{x} \in R^n$, $\Omega \subset R^n$ o domínio da imagem, $\Omega_1 \subset \Omega$ o *foreground* (imagem), $\Gamma = \partial\Omega_1$ são as bordas e $r_i^{\alpha_i} : \Omega \rightarrow \mathbb{R}$ são as funções que medem o erro da probabilidade de $\mathbf{x}$ estar em uma região da imagem. A função positiva e decrescente g , é um detector de bordas do gradiente da imagem definida como:

$$g(s) = \frac{1}{1 + \gamma s^2}, \quad (3.12)$$

tal que γ é uma constante positiva. Maiores detalhes deste funcional será visto no capítulo 6.

Na literatura são propostos basicamente três funções de erro $r_i^{\alpha_i}$ que são:

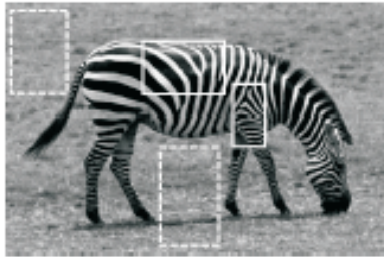
- $r_i = \delta(I - c_i)^2$, onde os parâmetros $\alpha_i = c_i$ são constantes; [12].
- $r_i = -\delta \log(P_i(I|\alpha_i))$, onde $P_i(I|\alpha_i)$ é uma distribuição de probabilidade, com parâmetros $\alpha_i = (\mu_i, \sigma_i)$ escalares; [39], [29].
- $r_i = \delta(I - s_i)^2 + \mu|\nabla s_i|^2$, onde os parâmetros $\alpha_i = s_i$ são funções; [28], [35],

onde δ é um parâmetro regularizador positivo, que serve para balancear os termos competição e suavização. Detalhes sobre as possíveis funções de erro podem ser encontradas em [32].

Se o ótimo α é conhecido a priori, temos que (3.11) é um problema de segmentação supervisionado. Caso contrário, a segmentação é não supervisionada e α tem que ser otimizado, executando a minimização alternada em Ω_i (variável de partição) e nos componentes de α (parâmetros da região).

Abaixo vemos dois exemplos de imagens segmentadas pelo modelo proposto em [27].

Exemplo 1- Na imagem original (zebra), foi aplicado o método de competição entre regiões, proposto em [27], obtendo assim a imagem segmentada, i.e., a zebra sem o fundo.



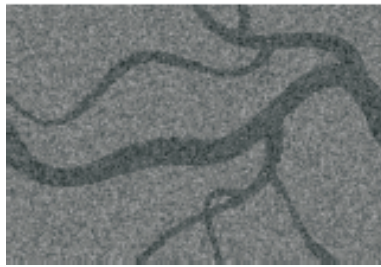
(a) Imagem original.



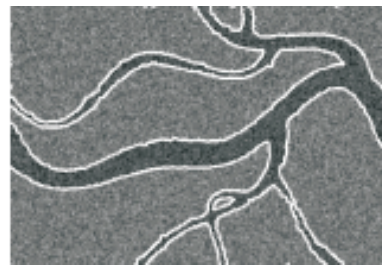
(b) Imagem segmentada.

Figura 3.5: Problema de Segmentação 1. Imagens retiradas de [27].

Exemplo 2- Na imagem original, foi aplicado o método de competição entre regiões, proposto em [27], obtendo assim a imagem segmentada.



(a) Imagem original.



(b) Imagem segmentada.

Figura 3.6: Problema de Segmentação 2. Imagens retiradas de [27].

3.4 Outros problemas de processamento de imagens baseados em variação total

Nesta seção apresentaremos outros problemas de processamento de imagens baseado na Variação Total, onde daremos uma breve introdução a cada um destes problemas e algumas de suas formulações.

3.4.1 Deblurring

Deblurring é um velho problema de processamento de imagens, porém continua a atrair a atenção de pesquisadores e profissionais afins. Em uma série de problemas do mundo real, como por exemplo da astronomia, podemos encontrar aplicações para os algoritmos de restauração de imagens baseado no problema deblurring. Este problema consiste em restaurar uma imagem borrada ou degradada (fora de foco) que pode ser a causa de diversos fatores, dentre eles:

1. Movimento durante o processo de captura de imagem, pela câmara ou pelo sujeito;
2. Lente fora de foco, uso de lente inadequada e turbulência atmosférica.

Assim, uma imagem borrada pode ser descrita pela equação genérica $I = Ku + \eta$, onde:

I : imagem borrada (imagem observada);

K : operador de distorção, também chamado de função de espalhamento pontual (PSF);

u : imagem original (limpa);

η : ruído aditivo, introduzido durante a aquisição das imagens.

Querendo resolver este problema, os autores Chambolle, Caselles, Novaga, Cremers e Pock [7] propuseram uma extensão do funcional ROF (3.4) do seguinte modo:

$$\min_u \int_{\Omega} |\nabla u| + \frac{\beta}{2} \int_{\Omega} (Ku - I)^2 dx, \quad (3.13)$$

onde $\Omega \subset \mathbb{R}^2$ é o domínio da imagem e K é um operador linear núcleo (ou operador distorção) definido acima.

Abaixo vemos um exemplo da aplicação do funcional acima na resolução de um problema deblurring. Temos à esquerda uma imagem desfocada, onde foi aplicado o método proposto por Chambolle, Caselles, Novaga, Cremers e Pock em [7], obtendo assim a imagem à esquerda que apresenta bordas mais nítidas.



(a) Imagem borrada.



(b) Resultado deblurring.

Figura 3.7: Problema deblurring. Imagens retiradas de [7].

3.4.2 Retoque Digital

Retoque Digital (ou Inpainting) refere-se a aplicação de algoritmos sofisticados para recuperar partes perdidas ou danificadas de dados de uma imagem (principalmente as pequenas regiões ou para remover pequenos defeitos). Por exemplo, no caso de uma pintura valiosa, esta tarefa seria realizada por um artista plástico de restauração de imagens.

Esta técnica pode ser empregada para remover ruído, melhorar cor, brilho e detalhes. Na fotografia e cinema, é usado para restauração de filmes, e também remoção de deterioração (por exemplo, rachaduras nas fotografias e manchas de poeira em vídeo). É também utilizado para remoção de olhos vermelhos, de data estampada em fotos e de objetos para efeito criativo, e em vídeos, pode ser usado para remover logotipos.

Este problema foi introduzido pela primeira vez em processamento digital de imagens na obra de Bertalmio [5] e mais tarde popularizada em outros trabalhos. Com tudo, a formulação que iremos apresentar do problema de retoque digital deve-se aos autores Tony Chan e J. Shen, no trabalho [11]. Assim, a formulação de Chan e Shen para resolver tal problema é dada por:

$$\min_u \alpha \int_{\Omega} |\nabla u| + \beta \int_{\Omega-D} (Ku - I)^2 dx,$$

onde K é o operador núcleo, I é a imagem observada, D é um subconjunto onde a informação da imagem está ausente ou inacessível e Ω o domínio da imagem original u.

Vejamos a seguir exemplos de aplicação do modelo acima em retoque digital:

Exemplo 1- Em uma imagem desfocada com partes obscuras, foi aplicado o método proposto em [11], obtendo assim uma imagem com bordas mais nítidas e recuperando as partes não visíveis da imagem.



(a) Imagem com partes obscuras.



(b) Resultado do retoque digital.

Figura 3.8: Problema de retoque Digital 1. Imagens retiradas de [10].

Exemplo 2- Em uma imagem com partes obscuras devido aos escritos, foi aplicado o método proposto em [11], obtendo assim uma imagem limpa com a preservação de suas bordas.



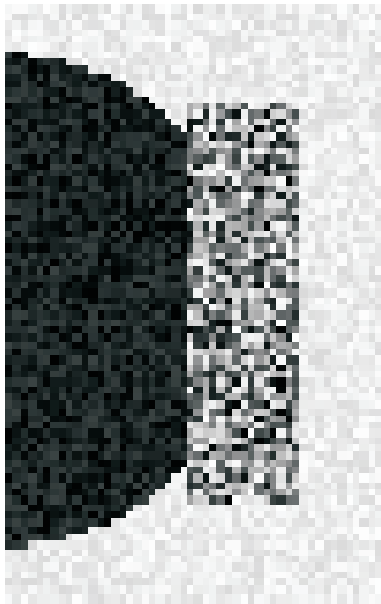
(a) Imagem com escritos.



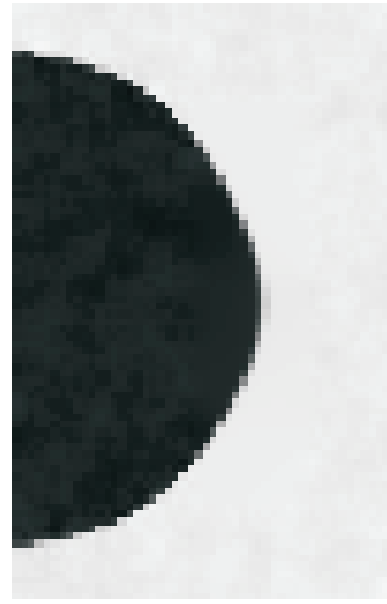
(b) Resultado do retoque digital.

Figura 3.9: Problema de retoque Digital 2. Imagens retiradas de [10].

Exemplo 3- Em uma imagem ruidosa com uma parte obscura, foi aplicado o método proposto em [11], obtendo assim uma imagem limpa, i.e., sem ruídos com a recuperação da parte não visível da imagem.



(a) Imagem com uma parte obscura.



(b) Resultado do retoque digital.

Figura 3.10: Problema de retoque Digital 3. Imagens retiradas de [10].

3.4.3 Zoom

Zoom é o método usado para aproximar ou afastar imagens, sejam elas em 3D, 2D ou digital. O zoom é muito usado em jogos computacionais e em softwares que fazem edição de vídeos e imagens. Ele apresenta grande utilidade para quem deseja ver uma imagem ou qualquer outro objeto em um software ou na vida real mais de perto, porém no caso de imagens e vídeos, a nitidez é perdida ao realizar o zoom.

Para resolver este problema, Antonin Chambolle propôs em seu artigo [6] a minimização do seguinte funcional:

$$\min_{u \in X, w \in Z^\perp} \int_{\Omega} |\nabla u| + \frac{\|u - (I + w)\|^2}{2\lambda}, \quad (3.14)$$

onde I é a imagem observada, $w = \frac{I-u}{\lambda}$ e $\lambda = \frac{1}{\beta}$.

Posteriormente, os autores Chambolle, Caselles, Novaga, Cremers e Pock [7] propuseram uma extensão do funcional ROF (3.4) para o problema de zoom, do seguinte modo:

$$\min_u \int_{\Omega} |\nabla u| + \frac{\beta}{2} \int_{\Omega} (Ku - I)^2 dx, \quad (3.15)$$

onde $\Omega \subset \mathbb{R}^2$ é o domínio da imagem. O que difere este funcional das outras extensões do funcional ROF (3.4) é o operador K , que aqui descreve o procedimento de redução da resolução.

A seguir, mostramos dois exemplos de zoom aplicando os modelos mostrados acima.

Exemplo 1- Na imagem original de uma mulher (esquerda) foi aplicado o método proposto em [6], obtendo assim a imagem reduzida. A partir desta imagem reduzida, foi aplicado novamente o mesmo método obtendo assim a imagem ampliada à direita.



(a) Imagem original e imagem reduzida.



(b) Imagem expandida pelo algoritmo proposto por Chambolle.

Figura 3.11: Problema zoom 1. Imagens retiradas de [6].

Exemplo 2- Na imagem original (olho humano) à esquerda foi aplicado o método proposto em [7], obtendo assim a imagem reduzida. A partir desta imagem reduzida, foi aplicado novamente o mesmo método obtendo assim a imagem ampliada à direita.



(a) Imagem original e imagem reduzida.



(b) Imagem expandida pelo algoritmo proposto por Chambolle, Caselles, Novaga, Cremers e Pock.

Figura 3.12: Problema zoom 1. Imagens retiradas de [7].

3.4.4 Decomposição de uma imagem em Geometria e Textura

O problema de remoção de ruídos pode ser visto como a separação de uma imagem dada I com ruído em dois componentes para formar a decomposição: $I = u + \eta$, onde u é a imagem original e $\eta = I - u$ é o ruído. Com base nesta decomposição Meyer [26] utiliza este mesmo procedimento para a extração de textura de uma imagem. Nestas condições η capta o ruído e também a textura. Assim o modelo proposto de decomposição é:

$$\inf_{(u,v) \in BV \times G(\mathbb{R}^2)} \left\{ \int_{\mathbb{R}^2} |\nabla u| + \frac{1}{2\lambda} \|v\|_{G(\mathbb{R}^2)}^2; \ I = u + v \right\}, \quad (3.16)$$

onde, v é a textura, $\lambda = \frac{1}{\beta}$, $\|v\|_{G(\mathbb{R}^2)}^2 = \inf_{g=(g_1,g_2)} \{ \|\sqrt{g_1^2 + g_2^2}\|_{L^\infty(\mathbb{R}^2,\mathbb{R}^2)} \mid v = \partial_x g_1 + \partial_y g_2 \}$ e o espaço G é dado por: $G(\mathbb{R}^2) = \{v \mid v = \operatorname{div} g, g \in L^\infty(\mathbb{R}^2, \mathbb{R}^2)\}$.

Já em um trabalho relacionado ao mesmo problema, os autores Aujol, Aubert, Blanc-Feraud e Chambolle [1], estudaram um modelo análogo ao do trabalho de Meyer [26], porém em um domínio limitado $\Omega \subset \mathbb{R}^2$. Para isto, substituíram $G(\mathbb{R}^2)$ pelo espaço $G(\Omega) = \{v \in L^2(\Omega) \mid v = \operatorname{div} g, g \in L^\infty(\Omega, \mathbb{R}^2)\}$. Neste método os autores consideram que a imagem I seja decomposta em $I = u + \eta + v$, onde u , η , e v são geometria, ruído, e textura respectivamente.

Posteriormente, no trabalho de L. A. Vese e S. J. Osher [34] foi proposto um modelo de decomposição, onde se separa uma imagem I em geometria e duas componentes de textura, g_1 e g_2 do seguinte modo:

$$\inf_{u,g_1,g_2} \left\{ \int_{\Omega} |\nabla u| + \beta \int_{\Omega} |I - u - \partial_x g_1 - \partial_y g_2|^2 dx dy + \mu \left[\int_{\Omega} (\sqrt{g_1^2 + g_2^2})^p dx dy \right]^{1/p} \right\}, \quad (3.17)$$

onde β, μ são parâmetros não negativos e $p \rightarrow \infty$.

Vejamos alguns exemplos da decomposição em geometria e textura:

Exemplo 1- Na imagem original à esquerda foi aplicado o método proposto em [34], obtendo assim a decomposição da imagem em textura (imagem central) e desenho (imagem à direita).

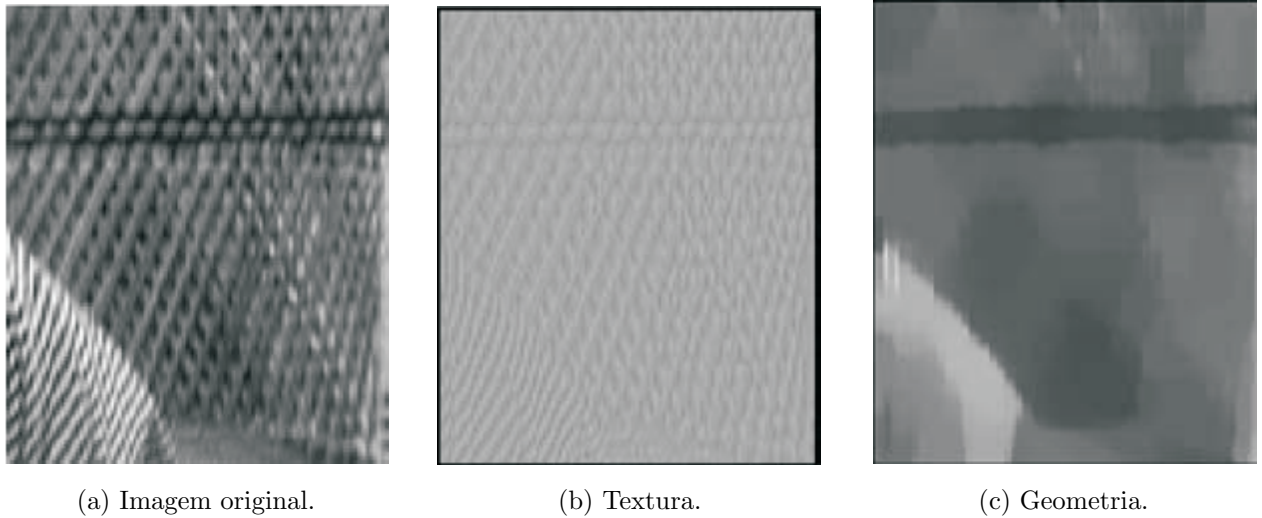


Figura 3.13: Problema de decomposição em geometria e textura 1. Imagens retiradas de [34].

Exemplo 2- Na imagem original (digital) à esquerda foi aplicado o método proposto em [34], obtendo assim a decomposição da imagem em textura (imagem central) e desenho (imagem à direita).

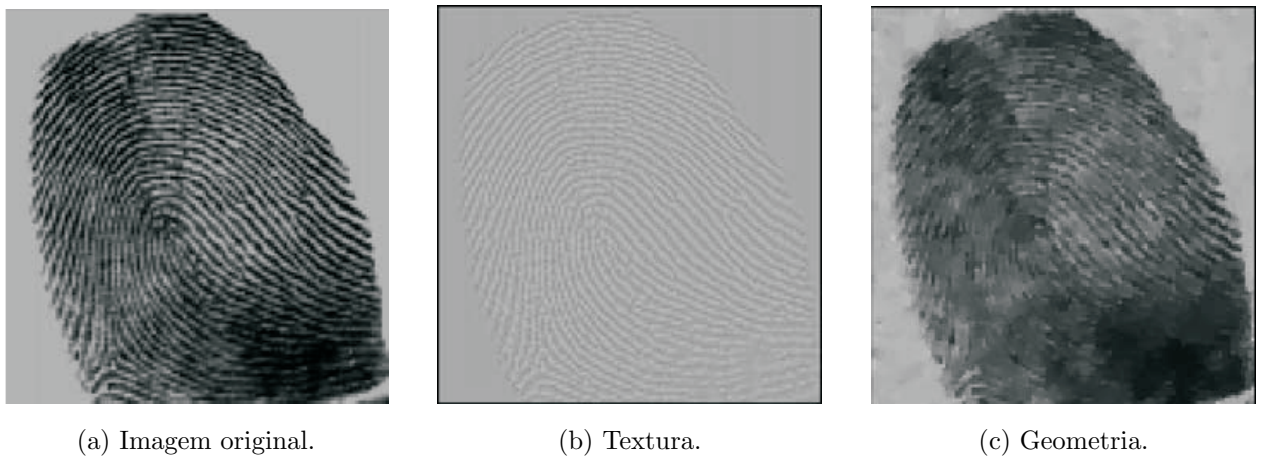
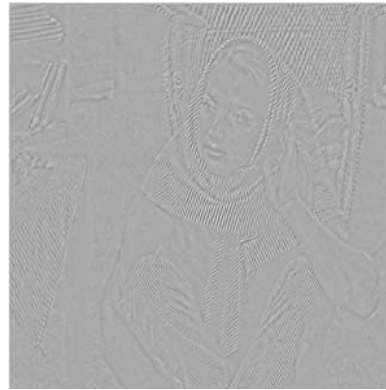


Figura 3.14: Problema de decomposição em geometria e textura 2. Imagens retiradas de [34].

Exemplo 3- Na imagem original (mulher) à esquerda foi aplicado o método proposto em [34], obtendo assim a decomposição da imagem em textura (imagem central) e desenho (imagem à direita).



(a) Imagem original.



(b) Textura.



(c) Geometria.

Figura 3.15: Problema de decomposição em geometria e textura 3. Imagens retiradas de [34].

Capítulo 4

Formulação Primal e Dual de Problemas Variacionais

4.1 Introdução

Dentre as maneiras de resolver problemas variacionais, a literatura apresenta método na formulação Primal, método na formulação Dual ou no chamado sistema Primal-Dual. Veremos um algoritmo para este último no capítulo 7. Para exemplificar, vamos considerar o modelo ROF e, neste capítulo, trataremos das resoluções nas formulações Primal e Dual, que são respectivamente:

$$\min_u \int_{\Omega} J(u) + \frac{\beta}{2} \|u - I\|^2 \quad e \quad \min_w \int_{\Omega} \frac{1}{\lambda} J^*(w) + \frac{\|w - I/\lambda\|^2}{2}.$$

Mostraremos um método de resolução na formulação Primal e a equivalência entre as duas formulações, deixando o algoritmo para resolução na formulação dual para o capítulo seguinte, onde falaremos sobre o método proposto por Chambolle para tal resolução.

4.2 Formulação Primal - Método de Resolução

Podemos resolver problemas variacionais na sua formulação Primal através das equações de Euler-Lagrange, vistas no Capítulo 2. Nesta seção encontraremos as equações de Euler-Lagrange do modelo ROF que, como vimos, pode ser entendido para vários outros problemas de processamento de imagens.

Vimos que o modelo ROF é dado por (3.4), ou seja,

$$\min_u \int_{\Omega} J(u) + \frac{\beta}{2} \|u - I\|^2,$$

onde $I \in X$, X é o espaço euclidiano $\mathbb{R}^{N \times N}$, N^2 é o número de total de pontos da imagem, $\beta > 0$ e $\|\cdot\|$ é a norma euclidiana.

Assim, a Equação de Euler-Lagrange do funcional (3.4) em relação a u é dada por:

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial u_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial u_y} \right) = 0, \quad (4.1)$$

onde

$$\begin{aligned} L = J(u) + \frac{\beta}{2} \|u - I\|^2 &= |\nabla u| + \frac{\beta}{2} \langle u - I, u - I \rangle \\ &= |\nabla u| + \frac{\beta}{2} (u - I)^2 \\ &= |\nabla u| + \frac{\beta}{2} (u^2 - 2uI + I^2). \end{aligned}$$

Portanto,

$$\frac{\partial L}{\partial u} = \frac{\partial}{\partial u} \left(\frac{\beta(u^2 - 2uI + I^2)}{2} \right) + \frac{\partial}{\partial u} (|\nabla u|) = \frac{\beta(2u - 2I)}{2} = \beta(u - I), \quad (4.2)$$

$$\left(\frac{\partial L}{\partial u_x} \right) = \frac{\partial}{\partial u_x} ([u_x^2 + u_y^2]^{\frac{1}{2}}) = \frac{1}{2} [u_x^2 + u_y^2]^{-\frac{1}{2}} 2u_x,$$

e

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial u_x} \right) &= \frac{\partial}{\partial x} \left(\frac{1}{2} [u_x^2 + u_y^2]^{-\frac{1}{2}} 2u_x \right) \\ &= u_{xx} [u_x^2 + u_y^2]^{-\frac{1}{2}} + u_x \left(-\frac{1}{2} \right) [u_x^2 + u_y^2]^{-\frac{3}{2}} (2u_x u_{xx} + 2u_y u_{yx}) \\ &= \frac{u_{xx}}{|\nabla u|} - \frac{u_{xx} u_x^2 + u_x u_y u_{yx}}{|\nabla u|^3}. \end{aligned}$$

Assim,

$$\frac{\partial}{\partial x} \left(\frac{\partial L}{\partial u_x} \right) = \frac{u_{xx}|\nabla u|^2 - u_{xx}u_x^2 - u_x u_y u_{yx}}{|\nabla u|^3}. \quad (4.3)$$

Da mesma forma,

$$\left(\frac{\partial L}{\partial u_y} \right) = \frac{\partial}{\partial u_y} ([u_x^2 + u_y^2]^{\frac{1}{2}}) = \frac{1}{2} [u_x^2 + u_y^2]^{-\frac{1}{2}} 2u_y,$$

e

$$\begin{aligned} \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial u_y} \right) &= \frac{\partial}{\partial y} \left(\frac{1}{2} [u_x^2 + u_y^2]^{-\frac{1}{2}} 2u_y \right) \\ &= u_{yy} [u_x^2 + u_y^2]^{-\frac{1}{2}} + u_y \left(-\frac{1}{2} \right) [u_x^2 + u_y^2]^{-\frac{3}{2}} (2u_y u_{yy} + 2u_x u_{xy}) \\ &= \frac{u_{yy}}{|\nabla u|} - \frac{u_{yy}u_y^2 + u_x u_y u_{xy}}{|\nabla u|^3}. \end{aligned}$$

Assim,

$$\frac{\partial}{\partial y} \left(\frac{\partial L}{\partial u_y} \right) = \frac{u_{yy}|\nabla u|^2 - u_{yy}u_y^2 - u_x u_y u_{yx}}{|\nabla u|^3}. \quad (4.4)$$

Substituindo os termos (4.2), (4.3) e (4.4) na equação (4.1) tem-se:

$$\begin{aligned} \beta(u - I) - \left(\frac{u_{xx}|\nabla u|^2 - u_{xx}u_x^2 - u_x u_y u_{yx}}{|\nabla u|^3} + \frac{u_{yy}|\nabla u|^2 - u_{yy}u_y^2 - u_x u_y u_{yx}}{|\nabla u|^3} \right) &= 0 \\ \beta(u - I) - \left(\frac{u_{xx}u_x^2 + u_{xx}u_y^2 - u_{xx}u_x^2 - u_x u_y u_{yx} + u_{yy}u_x^2 + u_{yy}u_y^2 - u_{yy}u_y^2 - u_x u_y u_{yx}}{|\nabla u|^3} \right) &= 0 \\ \beta(u - I) - \left(\frac{u_{xx}u_y^2 - 2u_x u_y u_{yx} + u_{yy}u_x^2}{|\nabla u|^3} \right) &= 0, \end{aligned}$$

que é a equação de Euler-Lagrange da formulação primal do modelo ROF (3.4).

Agora, podemos resolver esta equação através da equação do fluxo, do seguinte modo:

$$u_t = -\beta(u - I) + \left(\frac{u_{xx}u_y^2 - 2u_x u_y u_{yx} + u_{yy}u_x^2}{|\nabla u|^3} \right). \quad (4.5)$$

Para resolver a equação (4.5) usaremos diferenças finitas que sera visto a seguir.

4.2.1 Discretização da equação do fluxo através de diferenças finitas

Esta seção tem por finalidade apresentar resultados matemáticos necessários para discretização da equação do fluxo, que será usada durante este trabalho. Serão suavizadas determinadas imagens, as quais apresentam algum tipo de interferência, como por exemplo ruído e segmentação. Representamos estas imagens por funções u , com $u : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, e $n = 2$ ou $n = 3$. Procuramos então, uma solução $u(x)$, com $x \in \mathbb{R}^n$, de uma determinada equação diferencial parcial. Assim, devemos discretizar o domínio de $u(x, y)$, isto é, o domínio de região Ω será discretizado em uma malha de pontos bidimensionais, onde a malha de passos h e k que será construída, associada a (x_i, y_j) , é dada por:

$$(x_i, y_j) = (x_0 \pm ih, y_0 \pm jk), \text{ para } i, j = 1, 2, \dots, \quad (4.6)$$

sendo (x_0, y_0) um ponto de referência e h e k números positivos. Consideramos $h = k$ para que a malha possa se tornar regular em (x, y) , sendo a vizinhança deste ponto representada como na Figura abaixo.



Figura 4.1: Conjunto de pontos utilizado na convolução.

Como as imagens que são consideradas neste trabalho apresentam dimensão $m \times n$, a região Ω será discretizada em uma malha de pontos $m \times n$, os quais estarão igualmente espaçados, ou seja, $h = 1$.

Como nosso objetivo é aproximar uma função de duas variáveis, e estamos considerando o passo de tamanho constante $h = 1$, obtemos as seguintes aproximações relativas às derivadas parciais de primeira e segunda ordens da função $u(x, y)$:

$$\frac{\partial u}{\partial x} = u_x \approx \frac{u(x+h, y) - u(x, y)}{h} \Rightarrow u_x(x_i, y_i) \approx u_{i+1,j} - u_{i,j}, \quad \text{Fórmula Avançada.}$$

$$\frac{\partial u}{\partial x} = u_x \approx \frac{u(x, y) - u(x-h, y)}{h} \Rightarrow u_x(x_i, y_i) \approx u_{i,j} - u_{i-1,j}, \quad \text{Fórmula Atrasada.}$$

$$\frac{\partial u}{\partial x} = u_x \approx \frac{u(x+h, y) - u(x-h, y)}{2h} \Rightarrow u_x(x_i, y_i) \approx \frac{u_{i+1,j} - u_{i-1,j}}{2}, \quad \text{Fórmula Centrada.}$$

$$\frac{\partial u}{\partial y} = u_y \approx \frac{u(x, y+h) - u(x, y)}{h} \Rightarrow u_y(x_i, y_i) \approx u_{i,j+1} - u_{i,j}, \quad \text{Fórmula Avançada.}$$

$$\frac{\partial u}{\partial y} = u_y \approx \frac{u(x, y) - u(x, y-h)}{h} \Rightarrow u_y(x_i, y_i) \approx u_{i,j} - u_{i,j-1}, \quad \text{Fórmula Atrasada.}$$

$$\frac{\partial u}{\partial y} = u_y \approx \frac{u(x, y+h) - u(x, y-h)}{2h} \Rightarrow u_y(x_i, y_i) \approx \frac{u_{i,j+1} - u_{i,j-1}}{2}, \quad \text{Fórmula Centrada.}$$

$$\frac{\partial u}{\partial x \partial y} = u_{x,y} \approx \frac{u_{i+1,j+1} + u_{i-1,j-1} - u_{i-1,j+1} - u_{i+1,j-1}}{4}.$$

$$\frac{\partial^2 u}{\partial x^2} = u_{xx} \approx u_{i+1,j} - 2u_{i,j} + u_{i-1,j}.$$

$$\frac{\partial^2 u}{\partial y^2} = u_{yy} \approx u_{i,j+1} - 2u_{i,j} + u_{i,j-1}.$$

Por não ser nosso objetivo, não apresentamos detalhes da obtenção das fórmulas acima, porém maiores detalhes sobre tais deduções podem ser encontradas em [15].

Como a malha que utilizamos para discretização de Ω apresenta certa regularidade, temos que para pontos interiores serão usados nas aproximações para a primeira derivada os operadores diferenças centradas e percebemos que tais pontos estão bem definidos na malha usada na discretização. Já para os pontos nas regiões de contorno, serão utilizados operadores diferenças avançadas e atrasadas, dependendo da posição destes pontos.

Para terminar a discretização da equação do fluxo (4.5), que é dada por:

$$\frac{\partial u}{\partial t}(x, t) = -\beta(u - I) + \left(\frac{u_{xx}u_y^2 - 2u_xu_yu_{yx} + u_{yy}u_x^2}{|\nabla u|^3} \right) = -\beta(u - I) + \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right),$$

denotamos $L(u) = -\beta(u - I) + \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right)$. Assim, podemos escrever o problema (4.5) da seguinte forma:

$$\frac{\partial u}{\partial t} = \mathbb{L}(u), \quad (4.7)$$

com $u(x, y, 0) = I(x, y)$.

Portanto, a solução deste problema é obtida substituindo as derivadas presentes em $\mathbb{L}(u)$ por diferenças finitas e a equação diferencial parcial dada em (4.7) é então aproximada pela seguinte equação:

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\tau} = -\beta(u^n - I)_{i,j} + \left(\frac{u_{xx}u_y^2 - 2u_xu_yu_{yx} + u_{yy}u_x^2}{|\nabla u^n|^3} \right)_{i,j},$$

ou equivalentemente,

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\tau} = -\beta(u^n - I)_{i,j} + \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right)_{i,j}, \quad \text{com } u_{i,j}^0 = I(x_i, y_j). \quad (4.8)$$

Após apresentarmos a discretização da equação do fluxo, veremos na próxima seção a equivalência entre as duas formulações primal e dual.

4.3 Equivalência entre as Formulações Primal e Dual

Uma outra maneira de resolver problemas variacionais, citada anteriormente, é na sua formulação Dual. Porém, primeiramente, mostraremos nesta seção a equivalência entre as formulações primal e dual do modelo ROF e deixaremos o método de resolução no dual para o próximo capítulo.

Sabemos que $J(u)$ é definido por (3.1), e assim pela transformação de Legendre-Fenchel temos que, o fato de J ser homogêneo, isto é, $J(\lambda u) = \lambda J(u)$ para cada u , e $\lambda > 0$, se transforma em:

$$J^*(v) = \sup_u \langle u, v \rangle_X - J(u), \quad (4.9)$$

com $\langle u, v \rangle_X = \sum_{i,j} u_{ij} v_{ij}$, que é a “função característica” do conjunto fechado convexo K dado por:

$$J^*(v) = \chi_K(v) = \begin{cases} 0, & \text{se } v \in K, \\ +\infty, & \text{caso contrário.} \end{cases} \quad (4.10)$$

Assim, como vimos anteriormente, a equação de Euler-Lagrange de (3.4) é dada por:

$$\beta(u - I) - \left(\frac{u_{xx}u_y^2 - 2u_xu_yu_{yx} + u_{yy}u_x^2}{|\nabla u|^3} \right) = 0,$$

que é equivalente a,

$$\beta(u - I) + \underbrace{\left(\frac{-u_{xx}u_y^2 + 2u_xu_yu_{yx} - u_{yy}u_x^2}{|\nabla u|^3} \right)}_{\in \partial J(u)} = 0 \Rightarrow \beta(u - I) + \partial J(u) \ni 0, \quad (4.11)$$

onde ∂J é o “subdiferencial” de J (ver definição 1.5).

Com isso, considerando $\beta = \frac{1}{\lambda}$, a equação de Euler-Lagrange (4.11) pode ser reescrita do seguinte modo:

$$\frac{1}{\lambda}(u - I) + \partial J(u) \ni 0.$$

Agora, subtraindo $\frac{u}{\lambda}$ de ambos os lados da expressão acima, temos:

$$-\frac{I}{\lambda} + \partial J(u) \ni -\frac{u}{\lambda},$$

e de maneira análoga, subtraindo $\frac{I}{\lambda}$ de ambos os lados, temos:

$$\partial J(u) \ni \frac{I}{\lambda} - \frac{u}{\lambda} \Rightarrow \partial J(u) \ni \frac{I - u}{\lambda},$$

ou seja,

$$\frac{I - u}{\lambda} \in \partial J(u). \quad (4.12)$$

Assim, como $J^{**} = J$ temos que se $\underbrace{\frac{I - u}{\lambda}}_{\mu} \in \partial J(u)$ então:

$$\frac{I - u}{\lambda} \in \partial J(u) \Rightarrow u \in \partial J^* \left(\frac{I - u}{\lambda} \right). \quad (4.13)$$

De fato, pela Proposição 1.1 temos que

$$J(u) + J^*(\mu) = \langle u, \mu \rangle \Rightarrow J^*(u) + J^{**}(\mu) = \langle u, \mu \rangle \Rightarrow u \in \partial J^*(\mu).$$

Assim temos (4.13), ou seja:

$$\frac{I - u}{\lambda} \in \partial J(u) \Rightarrow u \in \partial J^*\left(\frac{I - u}{\lambda}\right).$$

Agora, se multiplicarmos por $\frac{1}{\lambda}$ de ambos os lados, a expressão acima pode ser reescrita do seguinte modo:

$$u \in \partial J^*\left(\frac{I - u}{\lambda}\right) \Rightarrow \frac{u}{\lambda} \in \frac{1}{\lambda} \partial J^*\left(\frac{I - u}{\lambda}\right).$$

Agora, subtraindo $\frac{u}{\lambda}$ de ambos os lados, temos:

$$0 \in \frac{-u}{\lambda} + \frac{1}{\lambda} \partial J^*\left(\frac{I - u}{\lambda}\right),$$

e de maneira análoga, somando $\frac{I}{\lambda}$ de ambos os lados, obtemos:

$$0 + \frac{I}{\lambda} \in \frac{I}{\lambda} - \frac{u}{\lambda} + \frac{1}{\lambda} \partial J^*\left(\frac{I - u}{\lambda}\right) \Rightarrow \frac{I}{\lambda} \in \frac{I - u}{\lambda} + \frac{1}{\lambda} \partial J^*\left(\frac{I - u}{\lambda}\right), \quad (4.14)$$

e assim, temos que $w = \frac{I - u}{\lambda}$ é minimizador de $\frac{\|w - I/\lambda\|^2}{2} + \frac{1}{\lambda} J^*(w)$ pois, pela expressão (4.14), nomeando $\frac{I - u}{\lambda} = w$ e fazendo $J^* = G$, temos que:

$$\frac{1}{\lambda} \in w + \frac{1}{\lambda} \partial G(w) \Rightarrow w - I + \beta \partial G(w) \ni 0,$$

que, como em (4.11), a expressão acima implica que w é minimizador de

$$\int_{\Omega} G(w) + \frac{\|w - I/\lambda\|^2}{2} = \int_{\Omega} \frac{1}{\lambda} J^*(w) + \frac{\|w - I/\lambda\|^2}{2}. \quad (4.15)$$

A expressão (4.15) é conhecida como **formulação dual** de (3.4).

Agora, como J^* é definido pela expressão (4.9), temos que $w = \pi_K(I/\lambda)$ (projeção não linear). Assim, a solução do problema (3.4) é dado por:

$$u = I - \pi_{\lambda K}(I). \quad (4.16)$$

Portanto, para resolvermos o problema (3.4) precisamos calcular a projeção $\pi_{\lambda K}$. No próximo capítulo descreveremos um método para calcular esta projeção (no conjunto discreto) em dimensão 2, e para outras dimensões é feito de maneira análoga com poucas modificações.

Capítulo 5

Algoritmo de Minimização da Variação Total na Formulação Dual

5.1 Formulação Dual na Forma Discreta

Neste Capítulo mostraremos um algoritmo proposto por Chambolle em [6] que minimiza a variação total de uma imagem e é baseado em sua formulação dual. Também estudaremos a sua convergência.

No capítulo anterior verificamos a equivalência entre as formulações primal e dual, onde considerávamos imagens como matrizes bidimensionais de tamanho $N \times N$ e denotamos por X o espaço euclidiano $\mathbb{R}^{N \times N}$. Definimos também, a variação total em um conjunto contínuo pela expressão (3.1). Agora daremos uma definição discreta para a variação total, pois iremos desenvolver o algoritmo no conjunto discreto e, para isso, vamos utilizar o operador gradiente discreto. Seja $u \in X$ e ∇u um vetor em $Y = X \times X$ dado por

$$(\nabla u)_{i,j} = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2),$$

com

$$(\nabla u)_{i,j}^1 := \begin{cases} u_{i+1,j} - u_{i,j}, & \text{se } i < N, \\ 0, & \text{se } i = N. \end{cases}$$

e

$$(\nabla u)_{i,j}^2 := \begin{cases} u_{i,j+1} - u_{i,j}, & \text{se } j < N, \\ 0, & \text{se } j = N. \end{cases}$$

para cada $i, j = 1, \dots, N$. Temos que $(\nabla u)_{i,j}^1$ se aproxima da derivada de u em relação a x , e $(\nabla u)_{i,j}^2$ da derivada de u em relação a y .

A partir da definição de ∇u , temos que a variação total de u pode ser dada por:

$$J(u) = \sum_{1 \leq i, j \leq N} |(\nabla u)_{i,j}|, \quad (5.1)$$

com $|y| := \sqrt{y_1^2 + y_2^2}$, para cada $y = (y_1, y_2) \in \mathbb{R}^2$.

Temos que (5.1) é a discretização da variação total e, como vimos, pode também ser definido no conjunto contínuo pela expressão (3.1).

Com visto também no capítulo anterior, temos pela transformação de Legendre-Fenchel que, o fato de J ser homogêneo implica em (4.9) e (4.10). Assim, como $J^{**} = J$ e $J^*(v) = 0$ se $v \in K$, temos que:

$$J^*(v) = \sup_{v \in K} \langle u, v \rangle_X - J(u),$$

que é equivalente a:

$$0 = \sup_{v \in K} \langle u, v \rangle_X - J(u) \Rightarrow J(u) = \sup_{v \in K} \langle u, v \rangle_X, \quad (5.2)$$

e em termos de conjuntos contínuos, vemos facilmente, de (3.1), que K é o fecho do conjunto

$$\{div \xi; \xi \in \mathcal{C}_c^1(\Omega; \mathbb{R}^2), |\xi(x)| \leq 1 \forall x \in \Omega\}.$$

Encontraremos agora uma caracterização semelhante na definição discreta. Para isso, usaremos no conjunto $Y = X \times X$, o produto escalar euclidiano padrão, dado por:

$$\langle p, q \rangle_Y = \sum_{1 \leq i, j \leq N} (p_{i,j}^1 q_{i,j}^1 + p_{i,j}^2 q_{i,j}^2),$$

onde $p = (p^1, p^2), q = (q^1, q^2) \in Y$.

Assim, para cada $u \in X$ temos que:

$$J(u) = \sup_p \langle p, \nabla u \rangle_Y, \quad (5.3)$$

onde $\sup$ é tomado para todo $p \in Y$ tal que $|p_{i,j}| \leq 1$, para cada i, j .

Por fim, introduzimos a definição de divergência discreta, $\text{div} : Y \rightarrow X$ pela analogia com o contínuo, dada por $\text{div} = -\nabla^*$ (onde ∇^* é o adjunto de ∇). Isto é, para cada $p \in Y$ e $u \in X$ temos que:

$$\langle -\text{div } p, u \rangle_X = \langle p, \nabla u \rangle_Y,$$

onde div é dado por:

$$(\text{div } p)_{i,j} := \begin{cases} p_{i,j}^1 - p_{i-1,j}^1, & \text{se } 1 < i < N \\ p_{i,j}^1, & \text{se } i = 1 \\ -p_{i-1,j}^1, & \text{se } i = N \end{cases} + \begin{cases} p_{i,j}^2 - p_{i,j-1}^2, & \text{se } 1 < j < N \\ p_{i,j}^2, & \text{se } j = 1 \\ -p_{i,j-1}^2, & \text{se } j = N \end{cases}$$

para cada $p = (p^1, p^2) \in Y$.

Portanto, da definição do operador divergente e de (5.3), temos a caracterização do conjunto K na definição discreta dada por:

$$K = \{\text{div } p \mid p \in Y, |p_{i,j}| \leq 1, \forall i, j = 1, \dots, N\}.$$

Na próxima seção veremos a prova da igualdade $\langle -\text{div } p, u \rangle_X = \langle p, \nabla u \rangle_Y$ no conjunto discreto.

5.1.1 A prova de $\langle -\text{div } p, u \rangle_X = \langle p, \nabla u \rangle_Y$

Faremos a prova começando pelo termo $\langle p, \nabla u \rangle_Y$. Desenvolveremos este termo usando as definições dadas na seção anterior e chegaremos em $\langle -\text{div } p, u \rangle_X$.

Temos portanto a seguinte igualdade:

$$\langle p, \nabla u \rangle_Y = \sum_{i,j} [p_{i,j}^1 (\nabla u)_{i,j}^1 + p_{i,j}^2 (\nabla u)_{i,j}^2].$$

Separando o somatório em $i = N$, $j = N$, $i < N$ e $j < N$ temos:

$$\begin{aligned} \langle p, \nabla u \rangle_Y &= p_{N,N}^1 (\nabla u)_{N,N}^1 + p_{N,N}^2 (\nabla u)_{N,N}^2 + \sum_{i < N} [p_{i,N}^1 (\nabla u)_{i,N}^1 + p_{i,N}^2 (\nabla u)_{i,N}^2] + \\ &+ \sum_{j < N} [p_{N,j}^1 (\nabla u)_{N,j}^1 + p_{N,j}^2 (\nabla u)_{N,j}^2] + \sum_{i,j < N} [p_{i,j}^1 (\nabla u)_{i,j}^1 + p_{i,j}^2 (\nabla u)_{i,j}^2], \end{aligned}$$

e usando a definição de $(\nabla u)^1$ e $(\nabla u)^2$ obtemos:

$$\begin{aligned} \langle p, \nabla u \rangle_Y &= \sum_{i < N} p_{i,N}^1 (u_{i+1,N} - u_{i,N}) + \sum_{j < N} p_{N,j}^2 (u_{N,j+1} - u_{N,j}) + \sum_{i,j < N} [p_{i,j}^1 (u_{i+1,j} - u_{i,j}) + \\ &+ p_{i,j}^2 (u_{i,j+1} - u_{i,j})]. \end{aligned}$$

A seguir, aplicando a propriedade distributiva concluímos:

$$\begin{aligned} \langle p, \nabla u \rangle_Y &= \sum_{i < N} p_{i,N}^1 u_{i+1,N} - \sum_{i < N} p_{i,N}^1 u_{i,N} + \sum_{j < N} p_{N,j}^2 u_{N,j+1} - \sum_{j < N} p_{N,j}^2 u_{N,j} + \\ &+ \sum_{i,j < N} p_{i,j}^1 u_{i+1,j} - \sum_{i,j < N} p_{i,j}^1 u_{i,j} + \sum_{i,j < N} p_{i,j}^2 u_{i,j+1} - \sum_{i,j < N} p_{i,j}^2 u_{i,j}. \end{aligned}$$

Reescrevendo os somatórios obtemos:

$$\begin{aligned} \langle p, \nabla u \rangle_Y &= \left(\sum_{i=1}^{N-1} p_{i,N}^1 u_{i+1,N} \right) - \left(\sum_{\substack{1 < i < N \\ j=N}} p_{i,N}^1 u_{i,N} + p_{1,N}^1 u_{1,N} \right) + \left(\sum_{j=1}^{N-1} p_{N,j}^2 u_{N,j+1} \right) - \\ &- \left(\sum_{\substack{1 < j < N \\ i=N}} p_{N,j}^2 u_{N,j} + p_{N,1}^2 u_{N,1} \right) + \left(\sum_{\substack{1 < i < N-2 \\ j=1}} p_{i,j}^1 u_{i+1,j} + \sum_{i=1}^{N-2} \left(\sum_{j=2}^{N-1} p_{i,j}^1 u_{i+1,j} \right) + \right. \\ &+ p_{N-1,1}^1 u_{N,1} + \left. \sum_{\substack{1 < j < N-2 \\ i=N-1}} p_{i,j}^1 u_{i+1,j} \right) - \left(\sum_{i < N} \left(\sum_{j < N} p_{i,j}^1 u_{i,j} \right) \right) + \left(\sum_{\substack{j < N-1 \\ i=1}} p_{i,j}^2 u_{i,j+1} + \right. \\ &+ p_{i,N-1}^2 u_{1,N} + \left. \sum_{i=2}^{N-1} \left(\sum_{j=1}^{N-2} p_{i,j}^2 u_{i,j+1} \right) + \sum_{\substack{1 < i < N-2 \\ j=N-1}} p_{i,j}^2 u_{i,j+1} \right) - \left(\sum_{i < N} \left(\sum_{j < N} p_{i,j}^2 u_{i,j} \right) \right). \end{aligned}$$

$$\begin{aligned}
\langle p, \nabla u \rangle_Y &= \left(\sum_{i=2}^{N-1} p_{i-1,N}^1 u_{i,N} + p_{N-1,N}^1 u_{N,N} \right) - \left(\sum_{\substack{1 \leq i \leq N \\ j=N}} p_{i,N}^1 u_{i,N} + p_{1,N}^1 u_{1,N} \right) + \\
&+ \left(\sum_{j=2}^{N-1} p_{N,j-1}^2 u_{N,j} + p_{N,N-1}^2 u_{N,N} \right) - \left(\sum_{\substack{1 \leq j \leq N \\ i=N}} p_{N,j}^2 u_{N,j} + p_{N,1}^2 u_{N,1} \right) + \\
&+ \left(\sum_{1 \leq i \leq N} p_{i-1,1}^1 u_{i,1} + \sum_{1 \leq i,j \leq N} p_{i-1,j}^1 u_{i,j} + p_{N-1,1}^1 u_{N,1} + \sum_{1 \leq j \leq N} p_{N-1,j}^1 u_{N,j} \right) - \\
&- \left(\sum_{i \leq N} \left(\sum_{1 \leq j \leq N} p_{i,j}^1 u_{i,j} + p_{i,1}^1 u_{i,1} \right) \right) + \left(\sum_{1 \leq j \leq N} p_{1,j-1}^2 u_{1,j} + p_{1,N-1}^2 u_{1,N} + \right. \\
&+ \sum_{1 \leq i,j \leq N} p_{i,j-1}^2 u_{i,j} + \sum_{1 \leq i \leq N} p_{i,N-1}^2 u_{i,N} \left. \right) - \left(\sum_{i \leq N} \left(\sum_{1 \leq j \leq N} p_{i,j}^2 u_{i,j} + p_{i,1}^2 u_{i,1} \right) \right) \\
\langle p, \nabla u \rangle_Y &= \left(\sum_{1 \leq i \leq N} p_{i-1,N}^1 u_{i,N} + p_{N-1,N}^1 u_{N,N} \right) - \left(\sum_{\substack{1 \leq i \leq N \\ j=N}} p_{i,N}^1 u_{i,N} + p_{1,N}^1 u_{1,N} \right) + \\
&+ \left(\sum_{1 \leq j \leq N} p_{N,j-1}^2 u_{N,j} + p_{N,N-1}^2 u_{N,N} \right) - \left(\sum_{\substack{1 \leq j \leq N \\ i=N}} p_{N,j}^2 u_{N,j} + p_{N,1}^2 u_{N,1} \right) + \\
&+ \left(\sum_{1 \leq i \leq N} p_{i-1,1}^1 u_{i,1} + \sum_{1 \leq i,j \leq N} p_{i-1,j}^1 u_{i,j} + p_{N-1,1}^1 u_{N,1} + \sum_{1 \leq j \leq N} p_{N-1,j}^1 u_{N,j} \right) - \\
&- \left(\sum_{1 \leq i \leq N} \left(\sum_{1 \leq j \leq N} p_{i,j}^1 u_{i,j} + p_{i,1}^1 u_{i,1} \right) + \sum_{\substack{1 \leq j \leq N \\ i=1}} p_{1,j}^1 u_{1,j} + p_{1,1}^1 u_{1,1} \right) + \\
&+ \left(\sum_{1 \leq j \leq N} p_{1,j-1}^2 u_{1,j} + p_{1,N-1}^2 u_{1,N} + \sum_{1 \leq i,j \leq N} p_{i,j-1}^2 u_{i,j} + \sum_{1 \leq i \leq N} p_{i,N-1}^2 u_{i,N} \right) - \\
&- \left(\sum_{1 \leq i \leq N} \left(\sum_{1 \leq j \leq N} p_{i,j}^2 u_{i,j} + p_{i,1}^2 u_{i,1} \right) + \sum_{\substack{1 \leq j \leq N \\ i=1}} p_{1,j}^2 u_{1,j} + p_{1,1}^2 u_{1,1} \right) \\
\langle p, \nabla u \rangle_Y &= \left(\sum_{1 \leq i \leq N} p_{i-1,N}^1 u_{i,N} + p_{N-1,N}^1 u_{N,N} \right) - \left(\sum_{\substack{1 \leq i \leq N \\ j=N}} p_{i,N}^1 u_{i,N} + p_{1,N}^1 u_{1,N} \right) + \\
&+ \left(\sum_{1 \leq j \leq N} p_{N,j-1}^2 u_{N,j} + p_{N,N-1}^2 u_{N,N} \right) - \left(\sum_{\substack{1 \leq j \leq N \\ i=N}} p_{N,j}^2 u_{N,j} + p_{N,1}^2 u_{N,1} \right) + \\
&+ \left(\sum_{1 \leq i \leq N} p_{i-1,1}^1 u_{i,1} + \sum_{1 \leq i,j \leq N} p_{i-1,j}^1 u_{i,j} + p_{N-1,1}^1 u_{N,1} + \sum_{1 \leq j \leq N} p_{N-1,j}^1 u_{N,j} \right) - \\
&- \left(\sum_{1 \leq i,j \leq N} p_{i,j}^1 u_{i,j} + \sum_{\substack{1 \leq i \leq N \\ j=1}} p_{i,1}^1 u_{i,1} + \sum_{\substack{1 \leq j \leq N \\ i=1}} p_{1,j}^1 u_{1,j} + p_{1,1}^1 u_{1,1} \right) + \\
&+ \left(\sum_{1 \leq j \leq N} p_{1,j-1}^2 u_{1,j} + p_{1,N-1}^2 u_{1,N} + \sum_{1 \leq i,j \leq N} p_{i,j-1}^2 u_{i,j} + \sum_{1 \leq i \leq N} p_{i,N-1}^2 u_{i,N} \right) - \\
&- \left(\sum_{1 \leq i,j \leq N} p_{i,j}^2 u_{i,j} + \sum_{\substack{1 \leq i \leq N \\ j=1}} p_{i,1}^2 u_{i,1} + \sum_{\substack{1 \leq j \leq N \\ i=1}} p_{1,j}^2 u_{1,j} + p_{1,1}^2 u_{1,1} \right).
\end{aligned}$$

Reordenado os termos chegamos a:

$$\begin{aligned}
\langle p, \nabla u \rangle_Y = & -[p_1^{1,1} u_{1,1} + p_{1,1}^2 u_{1,1} + p_{1,N}^1 u_{1,N} - p_{1,N-1}^2 u_{1,N} - p_{N-1,1}^1 u_{N,1} + p_{N,1}^2 u_{N,1} - \\
& - p_{N-1,N}^1 u_{N,N} - p_{N,N-1}^2 u_{N,N} + \sum_{1 < i < N} p_{i,1}^1 u_{i,1} - \sum_{1 < i < N} p_{i-1,1}^1 u_{i,1} + \\
& + \sum_{1 < i < N} p_{i,1}^2 u_{i,1} + \sum_{1 < i < N} p_{i,N}^1 u_{i,N} - \sum_{1 < i < N} p_{i-1,N}^1 u_{i,N} - \sum_{1 < i < N} p_{i,N-1}^2 u_{i,N} + \\
& + \sum_{1 < j < N} p_{1,j}^1 u_{1,j} + \sum_{1 < j < N} p_{1,j}^2 u_{1,j} - \sum_{1 < j < N} p_{1,j-1}^2 u_{1,j} - \sum_{1 < j < N} p_{N-1,j}^1 u_{N,j} + \\
& + \sum_{1 < j < N} p_{N,j}^2 u_{N,j} - \sum_{1 < j < N} p_{N,j-1}^2 u_{N,j} + \sum_{1 < i,j < N} p_{i,j}^1 u_{i,j} - \sum_{1 < i,j < N} p_{i-1,j}^1 u_{i,j} + \\
& + \sum_{1 < i,j < N} p_{i,j}^2 u_{i,j} - \sum_{1 < i,j < N} p_{i,j-1}^2 u_{i,j}],
\end{aligned}$$

e juntando termos semelhantes temos que:

$$\begin{aligned}
\langle p, \nabla u \rangle_Y = & -[p_{1,1}^1 u_{1,1} + p_{1,1}^2 u_{1,1} + p_{1,N}^1 u_{1,N} - p_{1,N-1}^2 u_{1,N} - p_{N-1,1}^1 u_{N,1} + p_{N,1}^2 u_{N,1} - \\
& - p_{N-1,N}^1 u_{N,N} - p_{N,N-1}^2 u_{N,N} + \sum_{1 < i < N} [(p_{i,1}^1 - p_{i-1,1}^1 + p_{i,1}^2) u_{i,1} + \\
& + (p_{i,N}^1 - p_{i-1,N}^1 - p_{i,N-1}^2) u_{i,N}] + \sum_{1 < j < N} [(p_{1,j}^1 + p_{1,j}^2 - p_{1,j-1}^2) u_{1,j} + \\
& + (-p_{N-1,j}^1 + p_{N,j}^2 - p_{N,j-1}^2) u_{N,j}] + \sum_{1 < i,j < N} (p_{i,j}^1 - p_{i-1,j}^1 + p_{i,j}^2 - p_{i,j-1}^2) u_{i,j}].
\end{aligned}$$

Agora, usando a definição do operador divergente concluímos:

$$\begin{aligned}
\langle p, \nabla u \rangle_Y = & -[(\operatorname{div} p \cdot u)_{1,1} + (\operatorname{div} p \cdot u)_{1,N} + (\operatorname{div} p \cdot u)_{N,1} + (\operatorname{div} p \cdot u)_{N,N} + \\
& + \sum_{1 < i < N} ((\operatorname{div} p \cdot u)_{i,1} + (\operatorname{div} p \cdot u)_{i,N}) + \sum_{1 < j < N} ((\operatorname{div} p \cdot u)_{1,j} + (\operatorname{div} p \cdot u)_{N,j}) + \\
& + \sum_{1 < i,j < N} (\operatorname{div} p \cdot u)_{i,j}] \\
\langle p, \nabla u \rangle_Y = & - \sum_{i,j} (\operatorname{div} p \cdot u)_{i,j} = \sum_{i,j} (-\operatorname{div} p)_{i,j} u_{i,j} = \langle -\operatorname{div} p, u \rangle_X.
\end{aligned}$$

Assim, como queríamos demonstrar temos que

$$\langle p, \nabla u \rangle_Y = \langle -\operatorname{div} p, u \rangle_X.$$

5.2 O Algoritmo

Nesta seção vamos mostrar o algoritmo proposto por Chambolle em [6] para resolver o problema de ROF (3.4) na sua formulação dual (4.15). Com tudo, vimos no capítulo anterior que a solução do problema primal ROF é:

$$u = I - \pi_{\lambda K}(I).$$

Agora, resolver $\pi_{\lambda K}(I)$ é o mesmo que encontrar a solução do seguinte problema:

$$\min\{\|\lambda \operatorname{div} p - I\|^2 : p \in Y = X \times X, |p_{i,j}|^2 - 1 \leq 0, \forall i, j = 1, \dots, N\}. \quad (5.4)$$

Assim, pelas condições de Kuhn-Tucker (ver seção 1.6) temos que existe um multiplicador de Lagrange $\alpha_{i,j} \geq 0$, associado ao problema (5.4), tal que para cada i, j temos:

$$-(\nabla(\lambda \operatorname{div} p - I))_{i,j} + \alpha_{i,j} p_{i,j} = 0, \quad (5.5)$$

com $\alpha_{i,j} > 0$ e $|p_{i,j}| = 1$ ou $|p_{i,j}| < 1$ e $\alpha_{i,j} = 0$. No ultimo caso temos $(\nabla(\lambda \operatorname{div} p - I))_{i,j} = 0$, e em ambos temos que:

$$\alpha_{i,j} = |(\nabla(\lambda \operatorname{div} p - I))_{i,j}|. \quad (5.6)$$

Substituindo (5.6) em (5.5) temos:

$$-(\nabla(\lambda \operatorname{div} p - g))_{i,j} + |(\nabla(\lambda \operatorname{div} p - g))_{i,j}| p_{i,j} = 0 \quad (5.7)$$

e nomeando a função lagrangeana $L = -(\nabla(\lambda \operatorname{div} p - g))_{i,j} + |(\nabla(\lambda \operatorname{div} p - g))_{i,j}| p_{i,j}$, Chambolle propôs o seguinte algoritmo, chamado de *gradiente descendente semi-implícito* (ou ponto fixo):

Escolhendo $\tau > 0$ e sendo $p^0 = 0$ temos pela equação do fluxo que, para qualquer $n \geq 0$,

$$p_t = (\nabla(\lambda \operatorname{div} p - g))_{i,j} - |(\nabla(\lambda \operatorname{div} p - g))_{i,j}| p_{i,j}.$$

Agora, discretizando a igualdade acima usando diferenças finitas obtemos:

$$\frac{p_{i,j}^{n+1} - p_{i,j}^n}{\tau} = \nabla(\lambda \operatorname{div} p^n - I)_{i,j} - |\nabla(\lambda \operatorname{div} p^n - I)|_{i,j} p_{i,j}^{n+1}.$$

Multiplicando de ambos os lados na igualdade acima por τ , temos:

$$p_{i,j}^{n+1} = p_{i,j}^n + \tau \left(\nabla \left(\lambda \operatorname{div} p^n - I \right) \right)_{i,j} - \tau \left| \nabla \left(\lambda \operatorname{div} p^n - I \right) \right|_{i,j} p_{i,j}^{n+1},$$

e agrupando os termos que apresentam $p_{i,j}^{n+1}$, chegamos a:

$$0 = p_{i,j}^n + \tau \left(\nabla \left(\lambda \operatorname{div} p^n - I \right) \right)_{i,j} - p_{i,j}^{n+1} \left(1 + \tau \left| \nabla \left(\lambda \operatorname{div} p^n - I \right) \right|_{i,j} \right).$$

Por fim, colocando em evidência $p_{i,j}^{n+1}$, concluímos:

$$\Rightarrow p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau(\nabla(\lambda \operatorname{div} p^n - I))_{i,j}}{1 + \tau|\nabla(\lambda \operatorname{div} p^n - I)|_{i,j}}. \quad (5.8)$$

Teorema 5.1 (Teorema 3.1 de [6]) *Seja $\tau \leq \frac{1}{8}$. Então $\lambda \operatorname{div} p^n$ converge para $\pi_{\lambda K}(I)$ com $n \rightarrow \infty$.*

Demonstração: *Temos, por indução, que para cada $n \geq 0$, $|p_{i,j}^n| \leq 1$ para todo i, j . De fato, sabemos que para $n = 0$ é válido $|p_{i,j}^0| \leq 1$, pois $p^0 = 0$. Supomos válido para $n - 1$, ou seja, $|p_{i,j}^{n-1}| \leq 1$ e provamos que vale para n . Como $|p_{i,j}^{n-1}| \leq 1$, então*

$$p^{n-1} + \tau(\nabla(\operatorname{div} p^{n-1} - \frac{I}{\lambda})) \leq 1[1 + \tau(\nabla(\operatorname{div} p^{n-1} - \frac{I}{\lambda}))],$$

dai,

$$p^n = \frac{p^{n-1} + \tau(\nabla(\operatorname{div} p^{n-1} - \frac{I}{\lambda}))}{1 + \tau(\nabla(\operatorname{div} p^{n-1} - \frac{I}{\lambda}))} \leq 1.$$

Agora, fixemos $n \geq 0$ e seja $\eta = \frac{(p^{n+1} - p^n)}{\tau}$. Chamando $\left(\operatorname{div} p^n - \frac{I}{\lambda} \right) = A$, temos que:

$$\|A\|^2 = \left\| \operatorname{div} p^{n+1} - \frac{I}{\lambda} \right\|^2 = \sum_{1 \leq i, j \leq N} \left(\operatorname{div} p^{n+1} - \frac{I}{\lambda} \right)^2$$

$$\|A\|^2 = \sum_{1 \leq i, j \leq N} \left[(\operatorname{div} p^{n+1})^2 - 2(\operatorname{div} p^{n+1}) \frac{I}{\lambda} + \frac{I^2}{\lambda^2} \right].$$

Substituindo $\operatorname{div} p^{n+1}$ por $\operatorname{div}(p^n + \tau[(\nabla A) - |\nabla A| p^{n+1}])$ obtemos:

$$\|A\|^2 = \sum_{1 \leq i, j \leq N} \left[(\operatorname{div}(p^n + \tau[(\nabla A) - |\nabla A| p^{n+1}]))^2 - 2(\operatorname{div}(p^n + \tau[(\nabla A) - |\nabla A| p^{n+1}])) \frac{I}{\lambda} + \frac{I^2}{\lambda^2} \right],$$

aplicando o operador divergente na soma $p^n + \tau[(\nabla A) - |\nabla A| p^{n+1}]$, temos que:

$$\|A\|^2 = \sum_{1 \leq i, j \leq N} \left[(\operatorname{div} p^n + \operatorname{div} (\tau[(\nabla A) - |\nabla A| p^{n+1}]))^2 - 2(\operatorname{div} p^n + \operatorname{div} (\tau[(\nabla A) - |\nabla A| p^{n+1}])) \frac{2I}{\lambda} + \frac{I^2}{\lambda^2} \right].$$

Agora, desenvolvendo o quadrado perfeito e $\operatorname{div} (\tau[(\nabla A) - |\nabla A| p^{n+1}])$, temos:

$$\|A\|^2 = \sum_{1 \leq i, j \leq N} \left[(\operatorname{div} p^n)^2 + 2\operatorname{div} p^n \operatorname{div} (\tau[(\nabla A) - |\nabla A| p^{n+1}]) + (\operatorname{div} [\tau(\nabla A) - |\nabla A| p^{n+1}])^2 - \frac{2I}{\lambda} (\operatorname{div} p^n) - \frac{2I}{\lambda} \operatorname{div} [\tau(\nabla A)] + \frac{2I}{\lambda} \operatorname{div} [\tau|\nabla A| p^{n+1}] + \frac{I^2}{\lambda^2} \right],$$

e aplicando o operador divergente na soma $\tau[(\nabla A) - |\nabla A| p^{n+1}]$, obtemos:

$$\|A\|^2 = \sum_{1 \leq i, j \leq N} \left[(\operatorname{div} p^n)^2 + 2(\operatorname{div} p^n)(\operatorname{div} [\tau(\nabla A)]) - 2(\operatorname{div} p^n)(\operatorname{div} [\tau|\nabla A| p^{n+1}]) + (\operatorname{div} [\tau(\nabla A)])^2 - 2(\operatorname{div} [\tau(\nabla A)])(\operatorname{div} [\tau|\nabla A| p^{n+1}]) + (\operatorname{div} [\tau|\nabla A| p^{n+1}])^2 - \right.$$

$$-\frac{2I}{\lambda} (\operatorname{div} p^n) - \frac{2I}{\lambda} \operatorname{div} [\tau(\nabla A)] + \frac{2I}{\lambda} \operatorname{div}[\tau|\nabla A| p^{n+1}] + \frac{I^2}{\lambda^2} \Big].$$

Reordenando os termos e colocando τ e τ^2 em evidência, chegamos a:

$$\begin{aligned} \|A\|^2 = \sum_{1 \leq i, j \leq N} \Big[& (\operatorname{div} p^n)^2 - \frac{2I}{\lambda} (\operatorname{div} p^n) + \frac{I^2}{\lambda^2} + 2\tau[(\operatorname{div} [\tau(\nabla A)]) - (\operatorname{div} [\tau|\nabla A| p^{n+1}])]A + \\ & + \tau^2[(\operatorname{div} [\tau(\nabla A)])^2 - 2(\operatorname{div} [\tau(\nabla A)]) (\operatorname{div} [\tau|\nabla A| p^{n+1}]) + (\operatorname{div} [\tau|\nabla A| p^{n+1}])^2] \Big]. \end{aligned}$$

Simplificando os termos e substituindo $\operatorname{div} \eta$, obtemos:

$$\begin{aligned} \|A\|^2 = \sum_{1 \leq i, j \leq N} \Big[& (\operatorname{div} p^n)^2 - 2(\operatorname{div} p^n) \frac{I}{\lambda} + \frac{I^2}{\lambda^2} + 2\tau(\operatorname{div} \eta)A + \\ & + \tau^2[(\operatorname{div} [\tau(\nabla A)])(\operatorname{div} [\tau|\nabla A| p^{n+1}])]^2 \Big]. \end{aligned}$$

Agora, substituindo $(\operatorname{div} p^n)^2 - 2(\operatorname{div} p^n) \frac{I}{\lambda} + \frac{I^2}{\lambda^2}$ por A^2 , temos que:

$$\|A\|^2 = \sum_{1 \leq i, j \leq N} [A^2 + 2\tau (\operatorname{div} \eta) A + \tau^2 (\operatorname{div} \eta)^2],$$

e aplicando a definição de $\|\cdot\|$ e de $\langle \cdot, \cdot \rangle$, concluímos:

$$\|A\|^2 = \|A\| + 2\tau \langle \operatorname{div} \eta, A \rangle + \tau^2 \|\operatorname{div} \eta\|^2.$$

Portanto:

$$\|A\|^2 = \left\| \operatorname{div} p^{n+1} - \frac{I}{\lambda} \right\|^2 = \|A\|^2 + 2\tau \langle \operatorname{div} \eta, A \rangle + \tau^2 \|\operatorname{div} \eta\|^2. \quad (5.9)$$

Como

$$2\tau \langle \operatorname{div} \eta, A \rangle = \tau(2 \langle \operatorname{div} \eta, A \rangle) = -\tau(2 \langle \eta, \nabla A \rangle), \text{ pois } \langle -\operatorname{div} p, u \rangle = \langle p, \nabla u \rangle,$$

e $\tau^2 \|\operatorname{div} \eta\|^2 \leq -K^2 \tau \|\eta\|_Y^2$, pois $K^2 \tau \|\eta\|_Y^2 > 0$ e $\tau \|\operatorname{div} \eta\|^2 > 0$, temos que:

$$\|A\|^2 + 2\tau \langle \operatorname{div} \eta, A \rangle + \tau^2 \|\operatorname{div} \eta\|^2 \leq \|A\|^2 - \tau(2\langle \eta, \nabla A \rangle - K^2 \tau \|\eta\|_Y^2).$$

Assim,

$$\begin{aligned} \left\| \operatorname{div} p^{n+1} - \frac{I}{\lambda} \right\|^2 &= \|A\|^2 + 2\tau \langle \operatorname{div} \eta, A \rangle + \tau^2 \|\operatorname{div} \eta\|^2 \\ &\leq \|A\|^2 - \tau(2\langle \eta, \nabla A \rangle - K^2 \tau \|\eta\|_Y^2). \end{aligned}$$

Denotemos $\|\eta\|_Y^2 = \langle \eta, \eta \rangle_Y$ e K a norma do operador divergente, $\operatorname{div} : Y \rightarrow X$. Agora, como $2\langle \eta, \nabla A \rangle = \sum 2\eta_{i,j}(\nabla A)_{i,j}$, temos que:

$$2\langle \eta, \nabla A \rangle - K^2 \tau \|\eta\|_Y^2 = \sum 2\eta_{i,j}(\nabla A)_{i,j} - K^2 \tau |\eta_{i,j}|^2.$$

Uma vez que $\eta_{i,j}$ é da forma $(\nabla A)_{i,j} - \rho_{i,j}$, com $\rho_{i,j} = |(\nabla A)_{i,j}| p^{n+1}$, temos que para cada i, j se verifica:

$$2\eta_{i,j}(\nabla A)_{i,j} - K^2 \tau |\eta_{i,j}|^2 = 2((\nabla A)_{i,j} - |(\nabla A)_{i,j}| p^{n+1})(\nabla A)_{i,j} -$$

$$-K^2 \tau \sum_{i,j} [((\nabla A)_{i,j} - |(\nabla A)_{i,j}| p^{n+1})^2]_{i,j}$$

desenvolvendo a multiplicação $((\nabla A)_{i,j} - |(\nabla A)_{i,j}| p^{n+1})(\nabla A)_{i,j}$ e o quadrado perfeito

$$= 2(\nabla A)_{i,j}^2 - 2(\nabla A)_{i,j}|(\nabla A)_{i,j}| p^{n+1} - K^2 \tau [(\nabla A)_{i,j}^2 - 2(\nabla A)_{i,j}|(\nabla A)_{i,j}| p^{n+1} +$$

$$+ |(\nabla A)_{i,j}|^2 (p^{n+1})^2]$$

substituindo $(\nabla A)_{i,j}^2 - 2(\nabla A)_{i,j}|(\nabla A)_{i,j}| p^{n+1} + |(\nabla A)_{i,j}|^2 (p^{n+1})^2$ por $|\eta_{i,j}|^2$

$$= (\nabla A)_{i,j}^2 - 2(\nabla A)_{i,j}|(\nabla A)_{i,j}| p^{n+1} - K^2 \tau |\eta_{i,j}|^2 + (\nabla A)_{i,j}^2$$

somando e subtraindo o termo $|(\nabla A)_{i,j}|^2 (p^{n+1})^2$

$$\begin{aligned} &= (\nabla A)_{i,j}^2 - 2(\nabla A)_{i,j}|(\nabla A)_{i,j}| p^{n+1} + |(\nabla A)_{i,j}|^2 (p^{n+1})^2 - K^2\tau|\eta_{i,j}|^2 + (\nabla A)_{i,j}^2 - \\ &\quad - |(\nabla A)_{i,j}|^2 (p^{n+1})^2 \end{aligned}$$

substituindo $(\nabla A)_{i,j}^2 - 2(\nabla A)_{i,j}|(\nabla A)_{i,j}| p^{n+1} + |(\nabla A)_{i,j}|^2 (p^{n+1})^2$ por $|\eta_{i,j}|^2$

$$= |\eta_{i,j}|^2 - K^2\tau|\eta_{i,j}|^2 + (\nabla A)_{i,j}^2 - |(\nabla A)_{i,j}|^2 (p^{n+1})^2$$

colocando $|\eta_{i,j}|^2$ em evidência

$$= (1 - K^2\tau)|\eta_{i,j}|^2 + (|(\nabla A)_{i,j}|^2 - |\rho_{i,j}|^2).$$

Portanto,

$$2\eta_{i,j}(\nabla A)_{i,j} - K^2\tau|\eta_{i,j}|^2 = (1 - K^2\tau)|\eta_{i,j}|^2 + (|(\nabla A)_{i,j}|^2 - |\rho_{i,j}|^2).$$

Uma vez que $|p_{i,j}^{n+1}| \leq 1$, temos que $|\rho_{i,j}| \leq |A_{i,j}|$, pois, $|\rho| = |\nabla A| p^{n+1}$. Portanto, se $\tau \leq 1/K^2$, observamos que $\|A\|^2$ diminui com n , exceto quando $\eta = 0$, ou seja, $p^{n+1} = p^n$. De fato, se $\|\operatorname{div} p^{n+1} - I/\lambda\| = \|A\|$, temos que $|\rho_{i,j}| = |A_{i,j}|$, para cada i, j , de modo que $|A_{i,j}| = 0$ ou $|p_{i,j}^{n+1}| = 1$. Em ambos os casos, a equação (5.8) resulta em $p^{n+1} = p^n$.

Seja agora $m = \lim_{n \rightarrow \infty} \|A\|$, e $\bar{p}$ o limite de uma subsequência convergente (p^{n_k}) de (p^n) . Seja também $\bar{p}'$ o limite de p^{n_k+1} , então temos que

$$\bar{p}'_{i,j} = \frac{\bar{p}_{i,j} + \tau(\nabla(\operatorname{div} \bar{p} - I/\lambda))_{i,j}}{1 + \tau|(\nabla(\operatorname{div} \bar{p} - I/\lambda))_{i,j}|},$$

e repetindo os mesmos cálculos anteriores vemos que

$$m = \lim_{n \rightarrow \infty} \|\operatorname{div} \bar{p} - I/\lambda\| = \lim_{n \rightarrow \infty} \|\operatorname{div} \bar{p}' - I/\lambda\|,$$

pois $m = \lim_{n \rightarrow \infty} \|A\| = \lim_{n \rightarrow \infty} \|\operatorname{div} p^{n+1} - I/\lambda\|$. Temos também que $\bar{\eta}_{i,j} = \frac{(\bar{p}'_{i,j} - \bar{p}_{i,j})}{\tau} = 0$, para todo i, j , ou seja, $\bar{p} = \bar{p}'$.

Assim, como $-(\nabla(\lambda \operatorname{div} p - I))_{i,j} + |(\nabla(\lambda \operatorname{div} p - I))_{i,j}| p_{i,j} = 0$ temos que

$$-(\nabla(\lambda \operatorname{div} \bar{p} - I))_{i,j} + |(\nabla(\lambda \operatorname{div} \bar{p} - I))_{i,j}| \bar{p}_{i,j} = 0$$

é a equação de Euler-Lagrange para uma solução de (5.4).

Com isso, deduzimos que $\bar{p}$ resolve (5.4) e que $\lambda \operatorname{div} \bar{p}$ é a projeção $\pi_K(I)$. Uma vez que essa projeção é única, temos que toda sequência $\lambda \operatorname{div} p^n$ converge para $\pi_K(I)$. Portanto, o teorema está provado se pudermos mostrar que $K^2 \leq 8$.

Por definição temos que $K = \sup_{\|p\|_Y \leq 1} \|\operatorname{div} p\|$. Agora, vamos adotar por convenção que $p_{0,j} = p_{i,0} = p_{i,N} = 0$, para todo i, j . Então como $(a+b)^2 \leq (a+b)^2 + (a-b)^2 = 2a^2 + 2b^2$ temos que:

$$\begin{aligned} \|\operatorname{div} p\|^2 &= \sum_{1 \leq i,j \leq N} (p_{i,j}^1 - p_{i-1,j}^1 + p_{i,j}^2 - p_{i,j-1}^2)^2 = \sum_{1 \leq i,j \leq N} ((p_{i,j}^1 - p_{i-1,j}^1)^2 + (p_{i,j}^2 - p_{i,j-1}^2)^2) \\ &\leq 2(p_{i,j}^1 - p_{i-1,j}^1)^2 + 2(p_{i,j}^2 - p_{i,j-1}^2)^2 \leq 2[2(p_{i,j}^1)^2 + 2(p_{i-1,j}^1)^2] + 2[2(p_{i,j}^2)^2 + 2(p_{i,j-1}^2)^2] \\ &= 4(p_{i,j}^1)^2 + 4(p_{i-1,j}^1)^2 + 4(p_{i,j}^2)^2 + 4(p_{i,j-1}^2)^2 = 4[(p_{i,j}^1)^2 + (p_{i-1,j}^1)^2 + (p_{i,j}^2)^2 + (p_{i,j-1}^2)^2] \\ &= 4[(p_{i,j}^1)^2 + (p_{i,j}^2)^2] + 4[(p_{i-1,j}^1)^2 + (p_{i,j-1}^2)^2] = 4\|p\|_Y^2 + 4[(p_{i,j}^1)^2 + (p_{i,j}^2)^2 + (p_{N,j}^1)^2 + \\ &\quad + (p_{i,N}^2)^2 - (p_{0,j}^1)^2 - (p_{i,0}^2)^2] = 4\|p\|_Y^2 + 4\|p\|_Y^2 = 8\|p\|_Y^2. \end{aligned}$$

Dai, temos que $K^2 \leq 8$, provando o teorema.

□

No próximo capítulo, veremos duas aplicações do algoritmo visto aqui em problemas de processamento de imagens: segmentação de imagens baseado em competição de regiões fuzzy e remoção de ruídos.

Capítulo 6

Aplicação do Algoritmo de Minimização

6.1 Introdução

Para ilustrar a aplicação do algoritmo, vamos considerar o modelo de Competição entre Regiões Fuzzy e o modelo ROF para o problema de Segmentação de Imagens e Eliminação de Ruídos, respectivamente.

Serão mostrados neste capítulo, além do algoritmo, alguns resultados da aplicação dos métodos citados.

6.2 Aplicação em Segmentação de Imagens - Modelo de Competição entre Regiões Fuzzy

Como visto no capítulo 3, uma das maneiras de segmentar imagens pode ser minimizando o funcional (3.11). Uma forma comum de resolver (3.11) é reescrever F_0 como um funcional que contenha apenas integrais sobre o conjunto de bordas, ou então, reescrever F_0 como um funcional que só contenha integrais sobre o domínio Ω . Para isso, a idéia é substituir em F_0 o conjunto Ω_1 pela função de *Heaviside* $H(\Psi)$ (ver definição 1.3) e derivar as equações de Euler-Lagrange em relação a Ψ . A dificuldade apresentada é o fato de que o espaço de otimização - conjunto das funções características - é não convexo.

Nesta seção detalharemos todos os cálculos da equação de Euler-Lagrange e da minimização do funcional (3.11), e alguns resultados. Tal método foi proposto pelos autores Mory e Ardon e detalhes deste processo pode ser visto em [27].

6.2.1 Minimização do Funcional (3.11)

Queremos resolver o funcional (3.11), porém ele é não-convexo, pois o conjunto de subdomínios $\Omega_1 \subset \Omega$ não o é. Com tudo, podemos expressá-lo como um problema de otimização no conjunto das funções características não-convexas do seguinte modo:

$$\min_{\chi, \alpha} \left\{ \bar{F}_0(\chi, \alpha) = \int_{\Omega} g(\mathbf{x}) |\nabla \chi(\mathbf{x})| d\mathbf{x} + \int_{\Omega} \chi(\mathbf{x}) r_1^{\alpha_1}(\mathbf{x}) d\mathbf{x} + \int_{\Omega} (1 - \chi(\mathbf{x})) r_2^{\alpha_2}(\mathbf{x}) d\mathbf{x} \right\}. \quad (6.1)$$

Assim podemos estender (3.11) a um problema que é convexo, bastando substituir em (6.1) a função característica χ por uma função *fuzzy* nomeada de u (ver Definição 1.10), que pertence a algum conjunto convexo. Uma escolha apropriada para este conjunto é o espaço das funções de variação limitada $BV(\Omega)_{[0,1]}$ (ver definição (1.25)), com seus valores em $[0, 1]$. Assim, esta extensão nos leva a um novo problema de *Competição entre Regiões Fuzzy* com $u \in BV(\Omega)_{[0,1]}$ dado por:

$$\min_{u \in BV(\Omega)_{[0,1]}, \alpha} \left\{ \check{F}(u, \alpha) = \int_{\Omega} g(\mathbf{x}) |\nabla u(\mathbf{x})| d\mathbf{x} + \int_{\Omega} \underbrace{u(\mathbf{x})}_{P((\mathbf{x}) \in \Omega_1)} r_1^{\alpha_1}(\mathbf{x}) d\mathbf{x} + \int_{\Omega} \underbrace{(1 - u(\mathbf{x}))}_{P((\mathbf{x}) \in \Omega \setminus \Omega_1)} r_2^{\alpha_2}(\mathbf{x}) d\mathbf{x} \right\}, \quad (6.2)$$

onde $\mathbf{x} = (x, y)$.

Uma interpretação física que podemos fazer neste momento sobre a variável u , é que para qualquer $\mathbf{x} \in \Omega$, $u(\mathbf{x})$ representa a probabilidade de $\mathbf{x}$ pertencer a Ω_1 . Agora, o fato do problema (6.2) ser convexo em u nos diz que ele apresenta apenas soluções globais que minimizam $\check{F}$, independentes das condições iniciais e pode ser resolvida (mantendo-se α fixo) com eficientes algoritmos numéricos.

Proposição 6.1 (*Proposição 1 de [27]*) *Considerando o parâmetro da região α fixo e se u^* é um minimizador global de $\check{F}$ em $BV_{[0,1]}(\Omega)$, então para quase todo $t \in [0, 1]$, a função característica $\mathbf{x} \rightarrow \mathcal{X}_{u^*}(\mathbf{x}, t)$ do conjunto $\Omega_t = \{\mathbf{x} \in \Omega; u^*(\mathbf{x}) > t\}$ é também um minimizador global de $\check{F}$. Além disso, Ω_t é um minimizador global de F_0 .*

Demonstração: Para toda função $u \in BV_{[0,1]}(\Omega)$ e $r \in L^1(\Omega)$, temos que:

$$\int_{\Omega} g |\nabla u| = \int_0^1 \int_{\Omega} g(\mathbf{x}) |\nabla \mathcal{X}_u(\mathbf{x}, t)| d\mathbf{x} dt, \quad (6.3)$$

$$\int_{\Omega} ur = \int_{\Omega} \left(\int_0^{u(\mathbf{x})} dt \right) r(\mathbf{x}) d\mathbf{x} = \int_0^1 \int_{\Omega} \mathcal{X}_u(\mathbf{x}, t) r(\mathbf{x}) d\mathbf{x} dt. \quad (6.4)$$

De agora em diante, para simplificar a notação vamos omitir a variável fixa α . Usando as equações (6.3) e (6.4) para minimizar u^* , temos que $\check{F}(u^*) = \int_0^1 \check{F}(\mathcal{X}_{u^*}(\bullet, t)) dt$, que é o mesmo que $\int_0^1 \{\check{F}(\mathcal{X}_{u^*}(\bullet, t)) - \check{F}(u^*)\} dt = 0$. Se para todo t , $\check{F}(u^*) \leq \check{F}(\mathcal{X}_{u^*}(\bullet, t))$, então $\check{F}(\mathcal{X}_{u^*}(\bullet, t)) = \check{F}(u^*)$ para quase todo $t \in [0, 1]$. Assim, temos que a função $\mathbf{x} \rightarrow \mathcal{X}_{u^*}(\mathbf{x}, t)$ é também minimizador de $\check{F}(\bullet, \alpha)$ para quase todo $t \in [0, 1]$.

Além disso, se $\Omega_1 \subset \Omega$ é tal que $F_0(\Omega_1) < \check{F}(u^*)$, então a função característica também satisfaz $F_0(\Omega_1) = \check{F}(\mathcal{X}^{\Omega_1}) < \check{F}(u^*)$, que é uma contradição. Portanto, temos que para todo $\Omega_1 \subset \Omega$, $F_0(\Omega_1) \leq F_0(\Omega_1)$.

□

Portanto, devemos escolher uma estratégia eficiente de otimização que seja capaz de resolver qualquer problema de segmentação de imagens em duas regiões, e que possa ser expressa pela formulação geral (3.11), utilizando as funções de erro mostradas anteriormente. Assim (6.2) poderá ser minimizada iterativamente, alternando entre as duas seguintes etapas:

- a.** Mantendo-se u fixo, calcula-se os parâmetros α , de acordo com a função de erro escolhida;
- b.** Mantendo-se α fixo, minimiza-se (6.2) em relação a variável de partição (Ω_1) e calcula-se a função de pertinência u ;

Quando o estado estacionário u^* for alcançado, faz-se uma limiarização em u , obtendo a segmentação final.

Temos que a etapa **a.** depende apenas da função de erro escolhida, pois ela determina as equações que encontram valores ótimos para α_i . Já na etapa **b.**, fixando-se α , temos que a função de pertinência $u(x)$ é atualizada, calculando as equações de Euler-Lagrange do funcional (6.2) em relação a função u . Para isso, reescrevemos o funcional $\check{F} = \int_{\Omega} L(u, \alpha) d\mathbf{x}$ com $L(u_x, u_y, \nabla u) = g(\mathbf{x})|\nabla u(\mathbf{x})| + u(\mathbf{x})r_1^{\alpha_1}(\mathbf{x}) + (1 - u(\mathbf{x}))r_2^{\alpha_2}(\mathbf{x})$, e a equação de Euler-Lagrange do funcional (6.2) em relação a u , para todo $\mathbf{x} = (x, y) \in \Omega$ é dada por:

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial u_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial u_y} \right) = 0. \quad (6.5)$$

Lembrando que $L = g(\mathbf{x})|\nabla u(\mathbf{x})| + u(\mathbf{x})r_1^{\alpha_1}(\mathbf{x}) + (1 - u(\mathbf{x}))r_2^{\alpha_2}(\mathbf{x})$ e derivando L com relação a u_x temos que:

$$\frac{\partial L}{\partial u} = \frac{\partial(g|\nabla u|)}{\partial u} + \frac{(ur_1^{\alpha_1})}{\partial u} + \frac{((1-u)r_2^{\alpha_2})}{\partial u} = r_1^{\alpha_1} - r_2^{\alpha_2}, \quad (6.6)$$

e ainda:

$$\left(\frac{\partial L}{\partial u_x} \right) = g \left(\frac{u_x}{\sqrt{u_x^2 + u_y^2}} \right), \quad (6.7)$$

e

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial u_x} \right) &= g_x \left(\frac{u_x}{\sqrt{u_x^2 + u_y^2}} \right) + g \left(\frac{u_{xx}|\nabla u|^2 - u_x^2 u_{xx} - u_x u_y u_{xy}}{|\nabla u|^3} \right) \\ &= \frac{g_x u_x}{|\nabla u|} + \frac{g u_{xx} |\nabla u|^2 - g u_x^2 u_{xx} - g u_x u_y u_{xy}}{|\nabla u|^3} \\ &= \frac{g_x u_x}{|\nabla u|} + \frac{g u_{xx} u_x^2 + g u_{xx} u_y^2 - g u_{xx} u_x^2 - g u_{xy} u_y u_x}{|\nabla u|^3}. \end{aligned}$$

Portanto, simplificando a expressão acima obtemos:

$$\frac{\partial}{\partial x} \left(\frac{\partial L}{\partial u_x} \right) = \frac{g_x u_x}{|\nabla u|} + \frac{g u_{xx} u_y^2 - g u_{xy} u_y u_x}{|\nabla u|^3}. \quad (6.8)$$

Da mesma forma, derivando o L com relação a u_y temos:

$$\left(\frac{\partial L}{\partial u_y} \right) = g \left(\frac{u_y}{\sqrt{u_x^2 + u_y^2}} \right), \quad (6.9)$$

e ainda:

$$\begin{aligned} \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial u_y} \right) &= g_y \left(\frac{u_y}{\sqrt{u_x^2 + u_y^2}} \right) + g \left(\frac{u_{yy}|\nabla u|^2 - u_y^2 u_{yy} - u_x u_y u_{xy}}{|\nabla u|^3} \right) \\ &= \frac{g_y u_y}{|\nabla u|} + \frac{g u_{yy} |\nabla u|^2 - g u_y^2 u_{yy} - g u_x u_y u_{xy}}{|\nabla u|^3} \\ &= \frac{g_y u_y}{|\nabla u|} + \frac{g u_{yy} u_x^2 + g u_{yy} u_y^2 - g u_{yy} u_y^2 - g u_{xy} u_y u_x}{|\nabla u|^3}. \end{aligned}$$

Portanto, simplificando a expressão acima obtemos:

$$\frac{\partial}{\partial y} \left(\frac{\partial L}{\partial u_y} \right) = \frac{g_y u_y}{|\nabla u|} + \frac{g u_{yy} u_x^2 - g u_{xy} u_y u_x}{|\nabla u|^3}. \quad (6.10)$$

Assim, substituindo os termos (6.6), (6.8) e (6.10) na equação (6.5) tem-se:

$$r_1^{\alpha_1} - r_2^{\alpha_2} - \frac{1}{|\nabla u|} (\nabla g \cdot \nabla u) - \frac{g u_{xx} u_y^2 + g u_{yy} u_x^2 - 2g u_{xy} u_y u_x}{|\nabla u|^3} = 0,$$

que é equivalente a:

$$r_1^{\alpha_1} - r_2^{\alpha_2} - \frac{1}{|\nabla u|} (\nabla g \cdot \nabla u) - g \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right). \quad (6.11)$$

Devemos observar que minimizar $\check{F}$ com respeito a u em $BV_{[0,1]}$ é o mesmo que minimizar $\int_{\Omega} g |\nabla u| + \int_{\Omega} u r$ em BV com $0 \leq u \leq 1$, onde a função de competição é $r = r_1^{\alpha_1} - r_2^{\alpha_2}$. Este problema tem o mesmo conjunto de minimizadores do problema $\bar{F}(u) = \int_{\Omega} g |\nabla u| + \int_{\Omega} u r + \delta \nu(u)$, com $\nu(\varepsilon) = \max(0, |2\varepsilon - 1| - 1)$ e $\delta > \frac{1}{2}|r|_{\infty}$, que pode ser resolvido pelo método do gradiente descendente baseado nas equações de Euler-Lagrange. Porém, nenhuma vantagem da convexidade do funcional $\bar{F}$ poderia ser aproveitada.

Assim, querendo utilizar o algoritmo da projeção dual de Chambolle [6], visto no capítulo 5, adicionamos uma variável auxiliar v , que se aproxima de u , e obtemos a seguinte aproximação:

$$\min_{(u,v) \in BV(\Omega)_{[0,1]}} \left\{ \bar{F}(u,v) = \int_{\Omega} g |\nabla u| + \frac{1}{2\theta} \int_{\Omega} |u - v|^2 + \int_{\Omega} r v + \delta \nu(v) \right\}, \quad (6.12)$$

onde $\bar{F}(u,v)$ é convexa em u e v , e θ é um escalar suficientemente pequeno, para que os minimizadores (u^*, v^*) sejam praticamente idênticos na norma L^2 . Com este fato e com o método da projeção dual de *Chambolle*, o funcional (6.12) pode ser minimizado de maneira eficiente por este modelo seguindo duas etapas.

a. Mantendo-se u fixo o valor v é dado por:

$$v = \max(\min(u - \theta r, 1), 0), \quad (6.13)$$

onde a função de competição é $r = r_1^{\alpha_1} - r_2^{\alpha_2}$.

A expressão (6.13) é obtida a partir da equação de Euler-Lagrange de (6.12) com relação a v , que é dado por:

$$\frac{\partial L}{\partial v} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial v_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial v_y} \right) = 0$$

onde $L = g|\nabla u| + \frac{1}{2\theta}|u - v|^2 + rv + \delta\nu(v)$. Assim, temos que a igualdade acima pode ser escrita do seguinte modo:

$$\frac{v - u}{\theta} + r - 0 - 0 = 0,$$

ou equivalentemente,

$$v = u - r\theta.$$

b. Mantendo-se v fixo e considerando $p = (p^1, p^2)$ calcula-se u via Algoritmo da projeção do seguinte modo:

$$u = v - \theta \operatorname{div} p. \tag{6.14}$$

A expressão (6.14) é obtida através da equação de Euler-Lagrange de (6.12) com relação a u , que é dada por:

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x} \left(\frac{\partial L}{\partial u_x} \right) - \frac{\partial}{\partial y} \left(\frac{\partial L}{\partial u_y} \right) = 0,$$

onde $L = g|\nabla u| + \frac{1}{2\theta}|u - v|^2 + rv + \delta\nu(v)$. Assim, temos que a igualdade acima pode ser escrita do seguinte modo:

$$\frac{u - v}{\theta} + g \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) = 0.$$

Queremos resolver (6.12) baseado em [6] e a expressão (6.12) pode ser reescrita com a variável dual $p = (p_1, p_2)$ do seguinte modo:

$$\min_u \max_{|p| \leq g} \int_{\Omega} u \operatorname{div} p + \frac{1}{2\theta} (u - v)^2 dx, \quad (6.15)$$

que é equivalente a:

$$\max_{|p| \leq g} \min_u \int_{\Omega} u \operatorname{div} p + \frac{1}{2\theta} (u - v)^2 dx. \quad (6.16)$$

Agora, calculando a equação de Euler-Lagrange de (6.16) temos:

$$\operatorname{div} p + \frac{1}{\theta} (u - v) = 0 \Rightarrow u = v - \theta \operatorname{div} p. \quad (6.17)$$

Substituindo (6.17) em (6.16) obtemos:

$$\max_{|p| \leq g} \int_{\Omega} (v - \theta \operatorname{div} p) \operatorname{div} p + \frac{1}{2\theta} (v - \theta \operatorname{div} p - v)^2 dx,$$

que é o mesmo que:

$$\max_{|p| \leq g} \int_{\Omega} (v - \theta \operatorname{div} p) \operatorname{div} p + \frac{\theta}{2} (\operatorname{div} p)^2 dx. \quad (6.18)$$

Simplificando (6.18), obtemos:

$$\max_{|p| \leq g} \int_{\Omega} v \operatorname{div} p - \frac{\theta}{2} (\operatorname{div} p)^2 dx, \quad (6.19)$$

e a variação de energia em (6.19) com respeito ao campo de vetores p é dada pela seguinte expressão:

$$\int_{\Omega} (-\nabla v + \theta \nabla \operatorname{div} p) \cdot \delta p \, dx. \quad (6.20)$$

Agora, pelas condições de Kuhn-Tucker e juntamente com a condição $|p|^2 - g^2 \leq 0$, temos que:

$$-\nabla(\theta \operatorname{div} p - v) + \lambda p = 0, \quad (6.21)$$

onde $\lambda \geq 0$ é o multiplicador de Lagrange.

Como queremos resolver (6.12) pelo algoritmo proposto no Capítulo 5, podemos determinar neste momento λ do seguinte modo:

Na equação (6.21) temos que, se $|p(x)|^2 \leq g^2(x)$ então $\lambda = 0$. Caso contrário, se $|p(x)|^2 = g^2(x)$, então

$$|\nabla(\theta \operatorname{div} p - v)|^2 + \lambda^2 g^2(x) = 0, \quad (6.22)$$

que em ambos os casos temos:

$$\lambda = \frac{1}{g(x)} |\nabla(\theta \operatorname{div} p - v)|. \quad (6.23)$$

Substituindo (6.23) em (6.21) obtemos:

$$\nabla(\theta \operatorname{div} p - v) - \frac{1}{g(x)} |\nabla(\theta \operatorname{div} p - v)| p = 0. \quad (6.24)$$

Agora, nomeando a função lagrangeana $L = \nabla(\theta \operatorname{div} p - v) - \frac{1}{g(x)} |\nabla(\theta \operatorname{div} p - v)| p$, escolhendo $\tau > 0$ e sendo $p^0 = 0$ temos, pela equação do fluxo, que para qualquer $n \geq 0$

$$p_t = -\nabla(\theta \operatorname{div} p - v) + \frac{1}{g(x)} |\nabla(\theta \operatorname{div} p - v)| p.$$

Agora, a discretização da igualdade acima é dada por

$$\frac{p_{i,j}^{n+1} - p_{i,j}^n}{\tau} = \nabla(\theta \operatorname{div} p^n - v)_{i,j} - \frac{1}{g} |\nabla(\theta \operatorname{div} p^n - v)|_{i,j} p_{i,j}^{n+1}.$$

Multiplicando de ambos os lados por τ temos:

$$p_{i,j}^{n+1} = p_{i,j}^n + \tau \left(\nabla \left(\operatorname{div} p^n - \frac{v}{\theta} \right) \right)_{i,j} - \frac{1}{g} \tau \left| \nabla \left(\operatorname{div} p^n - \frac{v}{\theta} \right) \right|_{i,j} p_{i,j}^{n+1},$$

e agrupando os termos p^{n+1} obtemos:

$$0 = p_{i,j}^n + \tau \left(\nabla \operatorname{div} p^n - \frac{v}{\theta} \right)_{i,j} - p_{i,j}^{n+1} \left(1 + \frac{\tau}{g} \left| \nabla \left(\operatorname{div} p^n - \frac{v}{\theta} \right) \right|_{i,j} \right).$$

Isolando p^{n+1} , concluímos que:

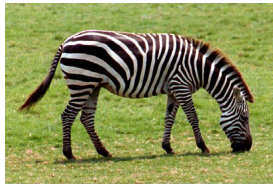
$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \tau (\nabla \operatorname{div} p^n - \frac{v}{\theta})_{i,j}}{1 + \frac{\tau}{g} |\nabla(\operatorname{div} p^n - \frac{v}{\theta})|_{i,j}}, \quad (6.25)$$

onde τ é escolhido de maneira adequada para garantir a estabilidade do modelo. Temos que no caso 2D, $\tau \leq \frac{1}{8}$.

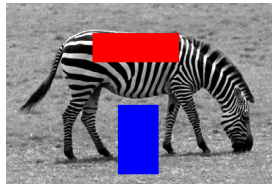
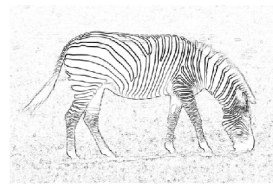
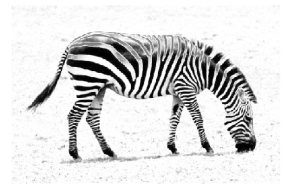
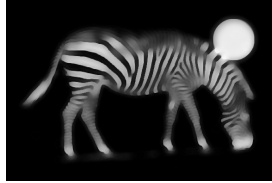
Na próxima seção veremos alguns resultados do modelo proposto em três imagens, onde detalharemos suas segmentações.

6.2.2 Resultados

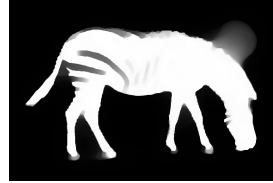
Nesta subseção, serão ilustrados resultados da execução do modelo *Competição de Regiões Fuzzy*. No primeiro experimento representado na Figura 6.1, mostramos uma imagem que deve ser segmentada em 2 partes, a zebra e o fundo. Utilizando o modelo analisado neste capítulo, temos na Figura 6.1 (b) a ilustração de 2 partes da imagem, numa representando o fundo - background - (região azul) e a outra representando a zebra - foreground - (região vermelha). Tais regiões são usadas para definir, a priori, r_1 e r_2 , que são as funções erro da zebra e do fundo respectivamente, onde $r = r_1 - r_2$ é a função competição dada por $r = \lambda \log(P_2/P_1)$. As funções P_1 e P_2 de distribuição de probabilidade do *foreground* (animal) e do *background* (fundo) das amostras destas regiões e está representado na Figura 6.1 (d). Na Figura 6.1 (c) temos a ilustração da função g de detecção de bordas. Já as Figuras 6.1 (e), (i) e (m) ilustram diferentes funções de pertinência u no instante inicial, e nas Figuras 6.1 (f)-(g), (j)-(k), (n)-(o) vemos essas funções de pertinência em instantes de tempo intermediário. Por fim, as Figuras 6.1 (h), (l) e (p) nos mostram os respectivos estados finais da função u , que independe da função de pertinência no instante inicial do tempo. Já as Figuras 6.1 (q) e (r) mostram o resultado final desta segmentação.



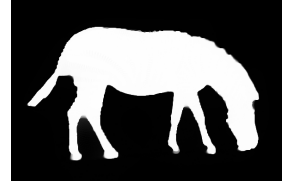
(a) Imagem original.

(b) Amostragem vermelho *foreground* e azul *background*.(c) Função g de detecção de bordas.(d) Imagem da Competição entre Regiões r .(e) Função de pertinência u no instante inicial.

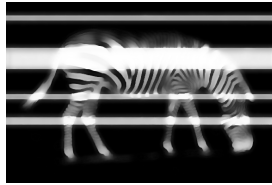
(f) Tempo intermediário (iter. 10).



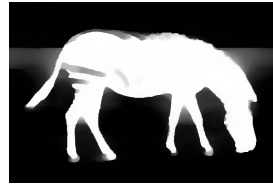
(g) Tempo intermediário (iter. 60).



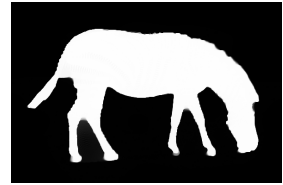
(h) Estado estacionário (iter. 350).

(i) Função de pertinência u no instante inicial.

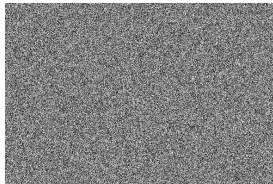
(j) Tempo intermediário (iter. 10).



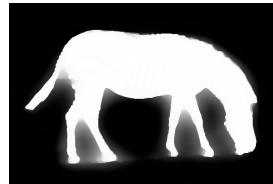
(k) Tempo intermediário (iter. 60).



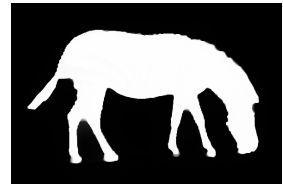
(l) Estado estacionário (iter. 350).

(m) Função de pertinência u no instante inicial.

(n) Tempo intermediário (iter. 10).



(o) Tempo intermediário (iter. 60).



(p) Estado estacionário (iter. 350).

(q) Segmentação da imagem uI .(r) Segmentação da imagem $I(1 - u)$.

Figura 6.1: Teste realizado com o modelo *Competição entre Regiões Fuzzy*. Imagens retiradas de [32].

Também no segundo experimento apresentamos na Figura 6.2 (a) a imagem que deve ser segmentada em 2 partes (avião e céu), utilizando o modelo proposto neste capítulo. Na Figura 6.2 (b) ilustramos as 2 partes da imagem, uma representando o céu - background - (região azul) e a outra representando o avião - foreground - (região vermelha). Tais regiões são definidas como na Figura 6.1. Já as Figuras 6.2 (d), (e) e (f) representam a função de pertinência u nos instantes inicial, intermediário e estacionário, respectivamente. Por fim, a imagem segmentada é mostrada na Figura 6.2 (h).



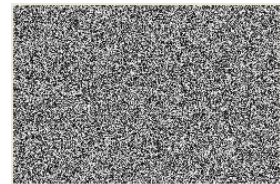
(a) Imagem original.



(b) Amostragem: vermelho (*foreground*) e azul (*background*).



(c) Imagem da Competição entre Regiões, função r .



(d) Função de pertinência u no instante inicial.



(e) Tempo intermediário (iter. 10).



(f) Tempo intermediário (iter. 30).



(g) Estado estacionário (iter. 1000).



(h) Segmentação da imagem a partir da função u .

Figura 6.2: Teste realizado com o modelo *Competição entre Regiões Fuzzy*. Imagens retiradas de [32].

De maneira análoga, no terceiro experimento apresentamos na Figura 6.3 (a) a imagem que deve ser segmentada em 2 partes (pássaro e fundo), utilizando o modelo proposto neste capítulo. Na Figura 6.3 (b) ilustramos as 2 partes da imagem, uma representando o fundo - background - (região azul) e a outra representando o pássaro - foreground - (região vermelha). Tais regiões são definidas como na Figura 6.1. Já as Figuras 6.3 (d), (e) e (f) representam a função de pertinência u nos instantes inicial, intermediário e estacionário, respectivamente. Por fim, em 6.3 (h) temos a imagem segmentada.

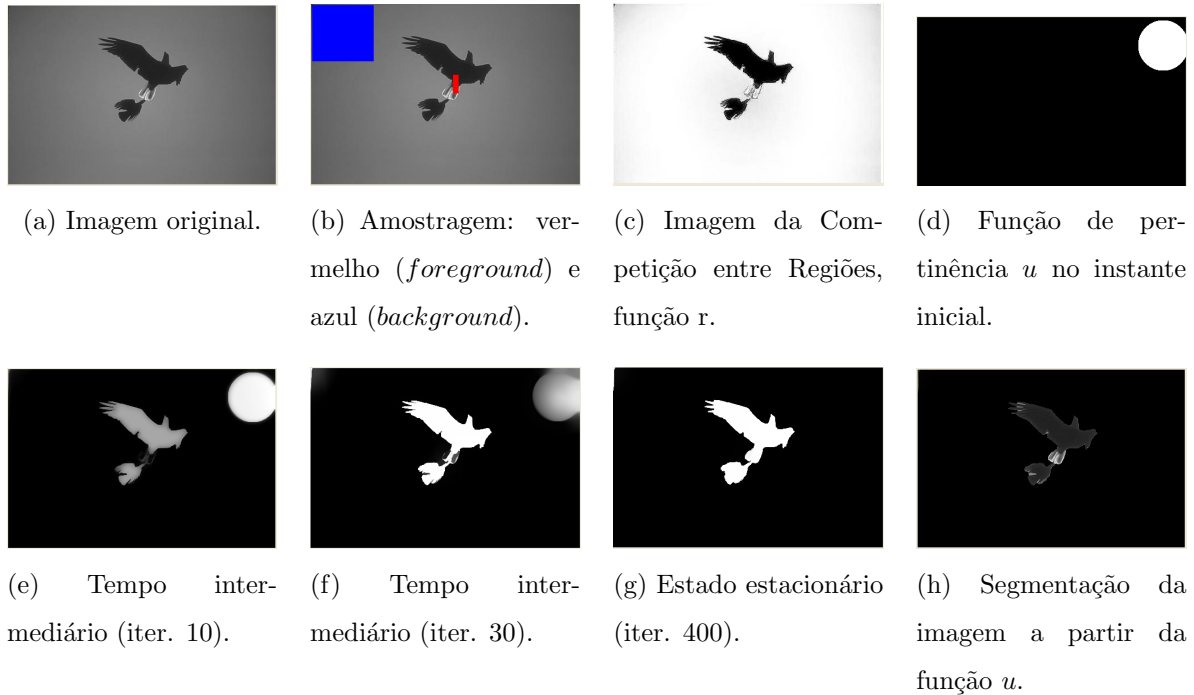


Figura 6.3: Teste realizado com o modelo *Competição entre Regiões Fuzzy*. Imagens retiradas de [32].

Maiores detalhes das imagens, parâmetros utilizados e outros testes podem ser encontrados em [32].

Na próxima seção, daremos uma outra aplicação do algoritmo mostrado no Capítulo 5. Trataremos do problema de processamento de imagens denoising (eliminação de ruídos).

6.3 Aplicação em Eliminação de Ruídos - Modelo ROF

Como vimos anteriormente, a idéia de minimizar a variação total (TV) para resolver o problema de remoção de ruído de uma imagem, foi sugerido pelos autores Rudin, Osher e Fatemi (ROF) [30], onde é considerado que a imagem observada $I = (I_{i,j})_{1 \leq i,j \leq N}$, com N definido como no capítulo 5, é a adição, a priori, de uma imagem seccionalmente suave $u = (u_{i,j})$ e um ruído aleatório gaussiano, de variância estimada σ^2 . Portanto, isto sugere recuperar a imagem original u , a partir da imagem observada I , tentando resolver o seguinte problema:

$$\min\{J(u); \|u - I\|^2 = N^2\sigma^2\}, \quad (6.26)$$

onde $J(u)$ é a variação total de u , definido em (5.1), e N^2 o número total de pontos. Sendo assim, como em (3.4), o problema acima é equivalente a

$$\min_{u \in X} \frac{N^2\sigma^2}{2\lambda} + |\nabla u|. \quad (6.27)$$

Podemos então mostrar que existe (tanto no conjunto contínuo, quanto no discreto) um multiplicador de Lagrange $\lambda > 0$ tal que, o problema $\|I - \langle I \rangle\|^2 \geq N^2\sigma^2$ tenha uma única solução que será dada pelo problema equivalente (3.4), onde $\beta = \frac{1}{\lambda}$ e $\langle I \rangle$ é o valor médio dos pontos $I_{i,j}$.

No capítulo 5 abordamos a resolução numérica do problema (3.4) na sua formulação dual, no entanto, uma vez que σ é em geral, mais fácil de se estimar do que λ , Chambolle propôs um algoritmo que aborda diretamente a resolução de (6.27). O problema é encontrar $\lambda > 0$ tal que $\|\pi_{\lambda K}(I)\|^2 = N^2\sigma^2$. Para isto, definimos para todo $s > 0$, $f(s) = \|\pi_{sK}(I)\|$.

Temos então o seguinte algoritmo para resolver o problema (6.27). Assumindo $0 < N\sigma < \|I - \langle I \rangle\|$, precisamos encontrar um valor $\bar{\lambda}$ para os quais $f(\bar{\lambda}) = N\sigma$.

Primeiramente, escolhemos um valor inicial arbitrário $\lambda_0 > 0$, e calculamos $v_0 = \pi_{\lambda_0 K}(I)$ pelo algoritmo descrito no capítulo 5, e calculamos também $f_0 = f(\lambda_0) = \|v_0\|$. Assim, dado λ_n e f_n , seja $\lambda_{n+1} = \left(\frac{N\sigma}{f_n}\right)\lambda_n$ e com isso calculamos $v_{n+1} = \pi_{\lambda_{n+1} K}(I)$ e $f_{n+1} = \|v_{n+1}\|$. Portanto, temos o seguinte teorema:

Teorema 6.1 Para $n \rightarrow \infty$, $f_n \rightarrow N\sigma$, quando $I - v_n$ converge para a única solução de (6.27).

Demonstração: A demonstração deste Teorema pode ser encontrado em [6].

□

Maiores detalhes deste algoritmo pode ser encontrado em [6].

Mostraremos agora alguns exemplos da remoção de ruídos em imagens processadas com esse algoritmo. Na prática, o autor observou que λ pode ser substituído por um novo valor $\frac{N\sigma}{\|div p^n\|}$ em (5.8) depois de cada iteração, e assim, obtemos uma convergência muito rápida para o limite de u , resolvendo (6.27).

Temos a imagem original representada na Figura 6.4. Já nos exemplos das Figuras 6.5 (direita) e 6.6 (direita), temos a imagem com a presença de ruídos, e nas Figuras 6.5 (esquerda) e 6.6 (esquerda) o resultado da aplicação do algoritmo na remoção de ruídos.



(a) Imagem original.

Figura 6.4: Problema de remoção de ruídos. Imagem retirada de [6].



(a) Imagem com ruído.



(b) Aplicação do algoritmo na remoção de ruídos.

Figura 6.5: Problema de remoção de ruídos 1 ($\sigma = 12$). Imagens retiradas de [6].



(a) Imagem com ruído.



(b) Aplicação do algoritmo na remoção de ruídos.

Figura 6.6: Problema de remoção de ruídos 2 ($\sigma = 25$). Imagens retiradas de [6].

Apresentamos neste capítulo duas aplicações do algoritmo de Chambolle para solução de um problema variacional na sua formulação Dual. Veremos no próximo capítulo como explorar as vantagens das formulações Primal e Dual, abordando um algoritmo que contempla ambas formulações.

Capítulo 7

Sistema Primal-Dual

7.1 Introdução

Até agora vimos duas maneiras de resolver problemas de processamento de imagens baseados na minimização da variação total. A primeira, na sua formulação primal, pelas equações de Euler-Lagrange, e a segunda na sua formulação dual, pelo algoritmo do gradiente descendente proposto por Chambolle. Porém, temos que as duas formulações apresentam algumas dificuldades, que serão destacadas neste capítulo. Na tentativa de sanar tais dificuldades, podemos usar as vantagens de ambas formulações numa formulação Primal-Dual, que se alterna entre as duas formando assim um sistema de equações conhecido como *sistema Primal-Dual*. Abordaremos neste capítulo as dificuldades de ambas formulações e a motivação para a formulação deste sistema. Apresentaremos também as notações a serem usadas neste capítulo, um algoritmo para o sistema Primal-Dual e alguns resultados e comparações com outros algoritmos já existentes.

A idéia da dualidade e do sistema Primal-Dual foi introduzida primeiramente em [9], onde os autores Chan, Golub e Mulet usam o método de Newton para resolver o sistema primal-dual do modelo ROF. Neste capítulo, falaremos sobre o algoritmo de resolução do mesmo sistema Primal-Dual, proposto pelos autores Zhu e Chan em [38].

Temos do ponto de vista computacional que, as formulações primal e dual do modelo de ROF, vistas nos capítulos 3 e 4, apresentam diferentes desafios para o cálculo de suas soluções. A formulação primal (3.4) apresenta o termo $|\nabla u|$ não-suave, o que torna os métodos baseados em derivadas impossíveis de serem calculados sem um parâmetro suavizador artificial. Já a formulação dual (5.4), impõe restrições que exigem mais esforços computacionais, além de apresentar energia dual quadrática que, embora seja mais suave, pode não apresentar uma única

solução dependendo do rank da função ∇ . Além disso, as duas formulações compartilham o problema de rigidez espacial. Baseado nestes problemas, os autores Zhu e Chan propuseram um novo algoritmo que alterna entre as duas formulações fazendo com que uma sane as dificuldades da outra. Este algoritmo tem por objetivo resolver um sistema composto pelas duas formulações, chamado de sistema Primal-Dual.

Na próxima seção veremos como é obtido o sistema Primal-Dual do modelo ROF.

7.2 Sistema Primal-Dual

Neste capítulo também iremos restringir nossa atenção para o caso discreto, ou seja, assumir que o domínio da imagem é um quadrado de $N \times N$ pontos, nomeados como (i, j) para $i = 1, 2, \dots, N$ e $j = 1, 2, \dots, N$. Representaremos as imagens como sendo matrizes bidimensionais de dimensão $N \times N$, onde $u(i, j)$ representa o valor da função u no ponto (i, j) . O operador gradiente discreto e variação total discreta é como definido no Capítulo 5, ou seja,

$$(\nabla u)_{i,j}^1 := \begin{cases} u_{i+1,j} - u_{i,j}, & \text{se } i < N, \\ 0, & \text{se } i = N, \end{cases}$$

e

$$(\nabla u)_{i,j}^2 := \begin{cases} u_{i,j+1} - u_{i,j}, & \text{se } j < N, \\ 0, & \text{se } j = N, \end{cases}$$

e a variação total (TV) de u dada por:

$$TV(u) = J(u) = \sum_{1 \leq i,j \leq N} |(\nabla u)_{i,j}|,$$

com $|y| := \sqrt{y_1^2 + y_2^2}$, para cada $y = (y_1, y_2) \in \mathbb{R}^2$.

Para apresentarmos o algoritmo proposto pelos autores Zhu e Chan reordenaremos a imagem u (resp. I) em um vetor y (resp. z), associando o elemento (i, j) da estrutura bidimensional com o elemento $(j-1)N+i$ na estrutura vetorial, da seguinte maneira:

$$y_{(j-1)N+i} = u_{i,j}, \quad 1 \leq i, j \leq N.$$

Temos que $y \in \mathbb{R}^M$, onde $M = N^2$. Agora, o componente (i, j) do gradiente ∇u pode ser representado como uma multiplicação do vetor $y \in \mathbb{R}^M$ por uma matriz $A_k \in \mathbb{R}^{M \times 2}$, para $k = 1, 2, \dots, M$, onde:

$$A_l^T y = \begin{cases} (y_{l+1} - y_l, y_{l+N} - y_l)^T, & \text{se } l \bmod N \neq 0 \text{ e } l \leq M - N, \\ (0, y_{l+N} - y_l)^T, & \text{se } l \bmod N = 0 \text{ e } l \leq M - N, \\ (y_{l+1} - y_l, 0)^T, & \text{se } l \bmod N \neq 0 \text{ e } l > M - N, \\ (0, 0)^T, & \text{se } l \bmod N = 0 \text{ e } l > M - N. \end{cases}$$

A versão discreta do modelo ROF (3.4) pode então ser escrita como

$$\min_y P(y) := \sum_{l=1}^M \|A_l^T y\| + \frac{\lambda}{2} \|y - z\|^2. \quad (7.1)$$

Usando esta notação, encontraremos a formulação dual, em termos da matriz A_k e do vetor y , do modelo ROF (7.1). Para isto, temos por consequência direta da desigualdade de Cauchy-Schwartz que, para qualquer vetor b ,

$$\|a\| = \max_{\|b\| \leq 1} a^T b. \quad (7.2)$$

Usando então a igualdade acima, podemos escrever a variação total discreta de y como:

$$\begin{aligned} \sum_{l=1}^M \|A_l^T y\| &= \max_{\|x_l\| \leq 1} \sum_{l=1}^M (A_l^T y)^T x_l \\ &= \max_{\|x_l\| \leq 1} y^T \sum_{l=1}^M A_l x_l \\ &= \max_{x \in X} y^T A x \end{aligned} \quad (7.3)$$

onde,

$$A = [A_1, A_2, \dots, A_M] \in \mathbb{R}^{M \times 2M}, \quad x_l = \begin{bmatrix} x_l^1 \\ x_l^2 \end{bmatrix} \in \mathbb{R}^2, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} \in \mathbb{R}^{2M}, \quad e$$

$$X = \{x : x \in \mathbb{R}^{2M}, \|x_l\| \leq 1 \text{ para } l = 1, 2, \dots, M\}. \quad (7.4)$$

Portanto, de (7.3) podemos reformular o modelo ROF primal (7.1) para os problemas max-min ou min-max do seguinte modo:

$$\begin{aligned}
\min_y \max_{x \in X} \phi(y, x) &= y^T Ax + \frac{\lambda}{2} \|y - z\|^2 \\
&= \max_{x \in X} \min_y y^T Ax + \frac{\lambda}{2} \|y - z\|^2,
\end{aligned} \tag{7.5}$$

onde a igualdade decorre do teorema max-min (ver [16]).

Calculando agora a equação de Euler-Lagrange de (7.5) temos que:

$$\lambda(y - z) + Ax = 0, \tag{7.6}$$

ou equivalentemente:

$$y = z - \frac{1}{\lambda} Ax.$$

Substituindo (7.6) em (7.5) temos o seguinte problema dual:

$$\max_{x \in X} D(x) := \frac{\lambda}{2} \left[\|z\|^2 - \left\| \frac{1}{\lambda} Ax - z \right\|^2 \right], \tag{7.7}$$

ou equivalentemente,

$$\min_{x \in X} \|Ax - \lambda z\|^2. \tag{7.8}$$

Assim, o problema (7.8) é equivalente ao problema discreto (5.4), lembrando que a matriz A é a versão discreta do operador divergente negativo $-\nabla \cdot$ e x é a variável dual discreta.

Vamos definir agora o Gap da dualidade $G(y, x)$ para o par ótimo primal-dual (y, x) , como a diferença entre as funções objetivos do primal e dual, ou seja, (7.1) e (7.7)

$$\begin{aligned}
G(y, x) &= P(y) - D(x) \\
&= \sum_{l=1}^M \|A_l^T y\| + \frac{\lambda}{2} \|y - z\|^2 - \frac{\lambda}{2} \left[\|z\|^2 - \left\| \frac{1}{\lambda} Ax - z \right\|^2 \right] \\
&= \sum_{l=1}^M (\|A_l^T y\| - x_l^T A_l^T y) + \frac{\lambda}{2} \left\| y - z + \frac{1}{\lambda} Ax \right\|^2.
\end{aligned} \tag{7.9}$$

O Gap da dualidade pode ser entendido como uma medida de proximidade do par ótimo primal-dual (y, x) para a solução primal-dual. Assim, $G(y, x)$ pode ser usada como critério de parada para algoritmos numéricos.

Se y e x são solução viáveis, temos de (7.9) que $G(y, x) \geq 0$, e $G(y, x) = 0$ se, e só se

$$\begin{cases} \|A_l^T y\| - x_l^T A_l^T y = 0, & \text{se } l = 1, \dots, N, \\ y - z + \frac{1}{\lambda} Ax = 0, \end{cases} \quad (7.10)$$

que é equivalente a:

$$\begin{cases} \|A_l^T y\| x_l - A_l^T y = 0, & \text{se } l = 1, \dots, N, \\ y - z + \frac{1}{\lambda} Ax = 0. \end{cases} \quad (7.11)$$

Damos então a expressão (7.11) o nome de sistema *Primal-Dual* e veremos na próxima seção um algoritmo para resolver tal sistema, proposto por Chan, Golub e Mulet em [9].

7.3 Algoritmo

Temos que a maioria dos algoritmos existentes para resolver os modelos ROF (7.5) ou (7.11) podem ser divididos em duas categorias: aqueles que precisam resolver um sistema linear em cada iteração (implícito) e aqueles que requer apenas a multiplicação de uma matriz por um vetor em um conjunto discreto (explícito). De modo geral, os métodos implícitos como por exemplo o proposto pelos autores Chan, Golub e Mulet, tem convergência rápida, no entanto, métodos explícitos são preferidos em muitas situações, pela sua simplicidade, sua rápida convergência inicial e resultados visualmente satisfatórios. Nesta seção vamos nos referir a dois métodos explícitos: método da marcha do tempo, como o *algoritmo primal gradiente descendente* e o método da dualidade de Chambolle, baseado no algoritmo semi-implícito tipo gradiente de descida, como o *algoritmo dual gradiente descendente*, que são representados do seguinte modo:

Primal de ROF (7.1):

$$\min_{y \in \mathbb{R}^M} \sum_{l=1}^M \|A_l^T y\| + \frac{\lambda}{2} \|y - z\|^2.$$

Algoritmo Primal gradiente descendente com suavizador β , obtido através de cálculos análogos aos feitos no capítulo 4 para obtenção de u^{n+1} :

$$y^{k+1} = y^k - \theta_k \left(\frac{1}{\lambda} \sum_{l=1}^M \frac{A_l A_l^T y^k}{\sqrt{\|A_l^T y^k\|^2 + \beta}} + y^k - z \right). \quad (7.12)$$

Algoritmo Primal gradiente descendente sem suavização, obtido apenas pela substituição do fator $\frac{A_l^T y_l^k}{\sqrt{\|A_l^T y_l^k\|}}$ por x_l^k :

$$y^{k+1} = y^k - \theta_k \left(\frac{1}{\lambda} \sum_{l=1}^M A_l x_l^k + y^k - z \right), \quad (7.13)$$

onde,

$$x_l^k = \begin{cases} \frac{A_l^T y_l^k}{\|A_l^T y_l^k\|}, & \text{se } A_l^T y_l^k \neq 0, \\ \text{qualquer elemento na bola unitária } B(0, 1) \subset \mathbb{R}^2, & \text{caso contrário.} \end{cases} \quad (7.14)$$

Dual ROF (7.3):

$$\max_{x \in X} \|Ax - \lambda z\|^2,$$

onde $X = \{x : x \in \mathbb{R}^{2M}, \|x_l\| \leq 1 \text{ para } l = 1, 2, \dots, M\}$.

Algoritmo Dual gradiente descendente (Chambolle): obtido através de cálculos análogos aos feitos no capítulo 5 para obtenção de p^{n+1} :

$$x_l^{k+1} = \frac{x_l^k - \tau_k A_l^T (Ax^k - \lambda z)}{1 + \tau_k \|A_l^T (Ax^k - \lambda z)\|}. \quad (7.15)$$

Observemos que os métodos acima são baseados exclusivamente na formulação primal (7.1) ou na formulação dual (7.7). A proposta de unificar as formulações primal e dual e apresentar um método do tipo gradiente descendente com base em ambas as formulações é dos autores Zhu e Chan e está detalhado em [38]. Assim, em cada passo, as atualizações irão explorar informações de ambas formulações, primal e dual, e com isso obtém-se uma melhora na velocidade da convergência.

Notemos mais uma vez que os problemas primal e dual apresentam diferentes desafios computacionais. O primal é de difícil resolução e tem convergência lenta quando $\|A_l^T y\| = 0$. Já

o dual é de difícil resolução e tem convergência lenta onde as restrições acontecem, isto é, nos pontos onde $\|x_i\| = 1$. Portanto, a formulação que combina estes dois problemas em um único algoritmo poderia ser capaz de resolver as dificuldades de um com a ajuda do outro.

O método proposto considera a atualização de uma solução (y^k, x^k) obtida no passo k em 2 etapas.

1. PASSO 1 (DUAL)

Fixando $y = y^k$, aplica-se o passo 1 do método do gradiente ascendente para maximizar $\phi(y^k, x)$, ou seja, calcula-se

$$\max_{x \in X} \phi(y^k, x). \quad (7.16)$$

A direção de busca é a ascendente $\nabla_x \phi(y^k, x) = A^T y^k$. Desta forma, atualizamos x do seguinte modo:

$$x^{k+1} = P_X(x^k + \tau_k \lambda A^T y^k), \quad (7.17)$$

onde τ_k é o tamanho do passo do dual e P_X denota a projeção sobre o conjunto X dado por:

$$P_X(z) = \arg \min_{x \in X} \|z - x\|.$$

O fator λ é usado em (7.17) para que o tamanho do passo τ_k não seja sensível aos problemas de diferentes níveis de cinza.

2. PASSO 2 (PRIMAL)

Fixando $x = x^{k+1}$, aplica-se o passo 1 do método do gradiente descendente para minimizar $\phi(y, x^{k+1})$, ou seja, calcula-se

$$\min_{y \in \mathbb{R}^N} \phi(y, x^{k+1}). \quad (7.18)$$

A direção de busca é a ascendente $\nabla_y \phi(y, x^{k+1}) = Ax^{k+1} + \lambda(y^k - z)$. Desta forma, atualizamos y do seguinte modo:

$$y^{k+1} = y^k - \theta_k \left(\frac{1}{\lambda} Ax^{k+1} + y^k - z \right), \quad (7.19)$$

onde θ_k é o tamanho do passo (primal).

O método proposto neste capítulo está relacionado aos métodos do tipo projeção existentes na literatura para encontrar pontos de sela e, mais geralmente, soluções para inequação variacional. Na próxima seção, vamos discutir brevemente sobre a estrutura dos métodos de projeção para resolver as desigualdades variacionais e apontar as ligações e diferenças entre o método citado neste capítulo e alguns trabalhos anteriores.

7.4 Conecções Teóricas

Será dado nesta seção algumas definições e comparações com outros métodos já existentes, que podem também ser encontradas em [37].

Seja H um espaço de Hilbert real (no nosso caso, $\mathbb{R}^n$), cujo produto interno e norma são denotados respectivamente por $\langle \cdot, \cdot \rangle$, e $\| \cdot \|$. Seja K um conjunto convexo fechado em H , e $F : H \rightarrow H$ (um mapeamento de H em si mesmo). Consideremos o problema de inequação variacional (ver definição 1.6), temos que na maioria das aplicações reais, K é convexo e F satisfaz a propriedade de continuidade Lipschitz (ver definição 1.3) e alguma monotonicidade (ver definição 1.7). Assim, encontrar um ponto de sela (y, x) para o problema max-min (7.5), pode ser visto como um caso especial do seguinte problema de inequação variacional:

$$\text{encontrar } v^* \in K \text{ tal que } \langle v - v^*, F(v^*) \rangle \geq 0, \quad \forall v \in K, \quad (7.20)$$

onde

$$v = \begin{bmatrix} y \\ x \end{bmatrix}, \quad F(v) = \begin{bmatrix} \phi_y(x, y) \\ -\phi_x(x, y) \end{bmatrix} \quad e \quad K = Y \times X.$$

Em particular, temos que o problema (7.5) pode ser transformado em um problema de inequação variacional $VI(K, F)$ em (7.20), com F e K definidos do seguinte modo para ROF irrestrito (7.5):

$$F(v) = \begin{bmatrix} Ax + \lambda(y - z) \\ -A^T y \end{bmatrix} \quad e \quad K = \mathbb{R}^M \times X.$$

O problema de inequação variacional está intimamente relacionado ao problema do ponto-fixo, e a teoria do ponto-fixo tem desempenhado um papel muito importante no desenvolvimento de vários algoritmos para resolver as inequações variacionais. Na verdade, temos o seguinte resultado:

Lema 7.1 *O elemento v^* é uma solução de $VI(K, F)$ se, e só se*

$$v^* = P_K(v^* - \alpha F(v^*)), \text{ para qualquer } \alpha > 0.$$

A formulação do ponto fixo no lema acima sugere um simples algoritmo iterativo para solução de u^* :

$$v^{k+1} = P_K(v^k - \alpha_k F(v^k)). \quad (7.21)$$

A convergência do algoritmo acima necessita que F seja fortemente monotônica e Lipschitziana, o que é uma condição demasiadamente restritiva em muitos casos. Uma alternativa para esse problema é considerar o seguinte algoritmo implícito:

$$v^{k+1} = P_K(v^k - \alpha_k F(v^{k+1})). \quad (7.22)$$

A convergência deste novo algoritmo requer somente a monotonicidade de F , porém resolver a atualização implícita em cada iteração não é uma tarefa fácil, tornando-se então um método não muito prático. Para superar as desvantagens dos métodos do tipo projeção definido em (7.21) e (7.22), Korpelevich [24] propôs um método modificado chamado de *algoritmo extragradiente*. O método consiste em calcular duas projeções em cada iteração: um passo preditor e um passo corretor, ou seja,

$$\bar{v}^k = P_K(v^k - \alpha_k F(v^k)), \quad (7.23)$$

$$v^{k+1} = P_K(v^k - \alpha_k F(\bar{v}^k)). \quad (7.24)$$

Se F for pseudo-monotônica ou Lipschitziana, temos provado a convergência global para o algoritmo acima, desde que o tamanho do passo α_k seja suficientemente pequeno para satisfazer

$$\alpha_k \|F(v^k) - F(\bar{v}^k)\| \leq \mu \|v^k - \bar{v}^k\|, \text{ para algum fixo } \mu \in (0, 1).$$

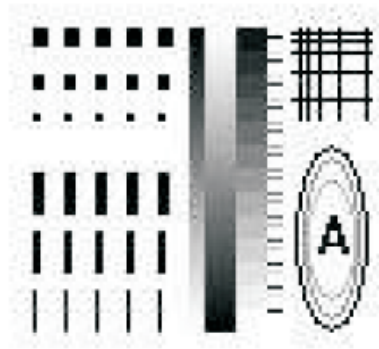
Existem muitos outros variantes do algoritmo original extragradiente com diferentes preditores e variando o tamanho do passo corretor, com o objetivo de melhorar o desempenho do par.

No nosso caso, temos que o problema (7.5) pode ser transformado em uma inequação variacional $VI(K, F)$, onde F é monotônica e Lipschitziana. A solução de ROF pode ser obtida por diversos métodos, no entanto o algoritmo (7.23; 7.24) tem um desempenho bem superior aos encontrados na literatura. Existem muitas explicações possíveis para isso: primeiramente no conjunto das inequações variacionais, as variáveis y e x são combinadas em uma única variável u , e são atualizadas em cada passo com um único tamanho de passo. Enquanto que no algoritmo apresentado neste capítulo, o primal y e o dual x são atualizados alternativamente, com liberdade para escolher seus próprios tamanhos de passo. Além disso, todos os algoritmos existentes foram desenvolvidos para solucionar as inequações variacionais como uma classe geral, enquanto o método mostrado neste capítulo explora a informação particular do problema, incluindo a função bilinear F e a estrutura especial do conjunto K , isto nos permite escolher o tamanho ideal do passo para melhorar o desempenho. Por outro lado, os autores não apresentam a prova da convergência global do algoritmo, o qual seria útil para nos ajudar a melhor entender como o algoritmo funciona.

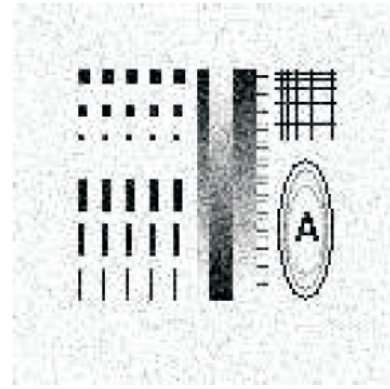
7.5 Resultados e Comparações

Mostraremos nesta seção alguns resultados da aplicação do algoritmo proposto neste capítulo e retiradas de [38], em 3 problemas de remoção de ruídos. As imagens originais e as ruidosas estão representadas na Figura 7.1, e seus tamanhos são 128×128 , 256×256 e 512×512 , respectivamente. O ruído das imagens é gerado pela adição de ruídos gaussianos de desvio padrão $\sigma = 20$ nas imagens originais, e os parâmetros λ usados pelos autores foram 0,0415, 0,053 e 0,0485 respectivamente. Os algoritmos que serão comparados são:

1. Método do tipo gradiente descendente semi-implícito de Chambolle (Capítulo 4)[6];
2. Método CGM de Chan, Golub e Mulet [9]; e
3. Primal-Dual (PDHG) proposto neste capítulo [38].



(a) Imagem original do problema 1, 128×128 .



(b) Imagem com ruído gaussiano ($\sigma = 20$) do problema 1.



(c) Imagem original do problema 2, 256×256 .



(d) Imagem com ruído gaussiano ($\sigma = 20$) do problema 2.



(e) Imagem original do problema 3, 512×512 .



(f) Imagem com ruído gaussiano ($\sigma = 20$) do problema 3.

Figura 7.1: Problema de remoção de ruídos. Imagens retiradas de [38].

As tabelas 7.1, 7.2 e 7.3 a seguir, apresentam o número de iterações e o tempo gasto pelos três algoritmos (PDHG, Chambolle e CGM) para resolver os problemas 1, 2 e 3. Em todos os algoritmos foram usados o mesmo ponto de partida $(y^0, x^0) = (z, 0)$, onde z é o vetor que representa a matriz I como definido na seção 7.2. Observamos que o método proposto neste capítulo é melhor para todos os diferentes critérios de parada e significativamente mais rápido

que o método de Chambolle e CGM.

Algorithms	TOL = 10^{-2}		TOL = 10^{-4}		TOL = 10^{-6}	
	Iter	CPU (s)	Iter	CPU (s)	Iter	CPU (s)
Chambolle	37	0.19	2074	5.8	36876	107
PDHG	14	0.11	106	0.33	456	1.2
CGM	5	1.20	14	3.5	19	4.7

Tabela 7.1: Iterações e tempo gasto para o problema 1, 128×128 , $\lambda = 0,0415$. Tabela retirada de [38].

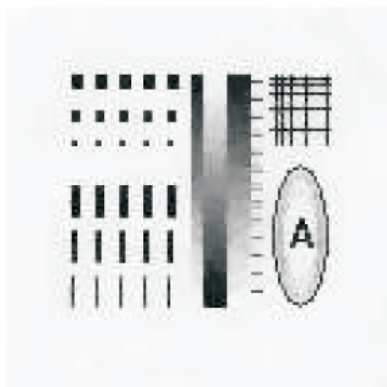
Algorithms	TOL = 10^{-2}		TOL = 10^{-4}		TOL = 10^{-6}	
	Iter	CPU (s)	Iter	CPU (s)	Iter	CPU (s)
Chambolle	45	1.2	1213	31	22597	579
PDHG	14	0.28	73	1.5	328	6.6
CGM	6	7.9	14	19	19	26

Tabela 7.2: Iterações e tempo gasto para o problema 2, 256×256 , $\lambda = 0,0415$. Tabela retirada de [38].

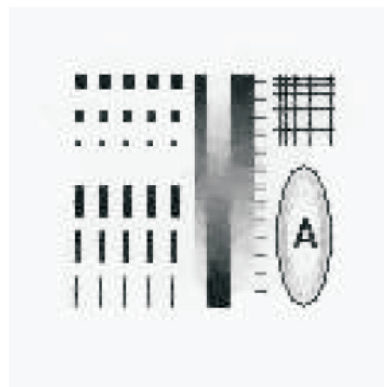
Algorithms	TOL = 10^{-2}		TOL = 10^{-4}		TOL = 10^{-6}	
	Iter	CPU (s)	Iter	CPU (s)	Iter	CPU (s)
Chambolle	61	7.5	1218	150	21925	2715
PDHG	16	1.5	72	6.8	320	30
CGM	7	51	14	101	20	143

Tabela 7.3: Iterações e tempo gasto para o problema 3, 512×512 , $\lambda = 0,0415$. Tabela retirada de [38].

Na Figura 7.2 temos o problema de remoção de ruídos para diferentes critérios de parada (TOL). Porém observamos que os menores valores de parada não produzem diferenças visuais.



(a) Problema 1 com critério de parada 10^{-2} .



(b) Problema 1 com critério de parada 10^{-4} .



(c) Problema 2 com critério de parada 10^{-2} .



(d) Problema 2 com critério de parada 10^{-4} .



(e) Problema 3 com critério de parada 10^{-2} .



(f) Problema 3 com critério de parada 10^{-4} .

Figura 7.2: Problema de remoção de ruídos com diferentes critérios de parada. Imagens retiradas de [38].

Capítulo 8

Conclusão

Neste trabalho abordamos a formulação variacional de alguns problemas de processamento de imagens. Estudamos alguns conceitos do Cálculo Variacional, em especial as equações de Euler-Lagrange e também condições para que o funcional $I[w] = \int_{\Omega} L(\nabla w(x), w(x), x) \, dx$ sujeito a condição de contorno $w = g$ tenha um mínimo, pelo menos em um espaço de Sobolev apropriado.

Abordamos o conceito de Variação Total em imagens e os trabalhos pioneiros nesta área com aplicação no problema de remoção de ruídos e segmentação, bem como em outros problemas de processamento de imagens, como por exemplo, deblurring, zoom, retoque digital e decomposição em geometria e textura.

Para exemplificar a metodologia, usamos o modelo clássico de Rudin, Osher e Fatemi (ROF) e apresentamos as formulações Primal e Dual de tal modelo, juntamente com um método de resolução para cada formulação, encontrando o funcional de energia e as equações de Euler-Lagrange de ambas. O método de resolução abordado na formulação Dual foi dado por Chambolle e detalhado no Capítulo 5.

Estudamos duas aplicações do método de resolução na formulação Dual, uma para o problema de segmentação de imagens baseado no modelo de Competição entre Regiões retirada de [27], e outra para o problema de remoção de ruídos baseado no método ROF retirada de [6].

Abordamos também os problemas encontrados na minimização dos funcionais em ambas formulações, onde a formulação Primal (3.4) apresenta o termo $|\nabla u|$ não-suave e a formulação dual (5.4) impõe restrições que exigem mais esforços computacionais, além de apresentar energia dual quadrática que, embora seja mais suave, pode não apresentar uma única solução dependendo do rank da função ∇ . Além disso, as duas formulações compartilham o problema de rigidez espacial.

Motivados para sanar tais dificuldades, apresentamos um outro método alternativo para a minimização de um funcional utilizando as formulações Primal e Dual denominado sistema Primal-Dual. Tal método explora vantagens de ambas formulações e tem uma performance superior aos métodos propostos por Chambolle e por Chan, Golub e Mulet quando se trata do tempo gasto para resolver os problemas de remoção de ruídos.

Observamos que a abordagem deste trabalho está nos cálculos numéricos apresentados, já a implementação numérica da Seção 6.2 deve-se à Vinícius R. P. Borges e as demais implementações, aos autores que referimos em cada resultado apresentado.

Conclui-se então que as equações de Euler-Lagrange, as formulações Primal e Dual e o sistema Primal-Dual são ferramentas poderosas quando se trata da resolução de problemas de processamento de imagens com formulações variacionais. Tais assuntos abordados neste trabalho estão sendo cada vez mais utilizados por pesquisadores, e estes obtendo resultados vantajosos na resolução de tais problemas.

Referências Bibliográficas

- [1] AUJOR, J.F., AUBERT, G., BLANC-FERAUD, L. & CHAMBOLLE, A. *Image Decomposition: Applications to Textured Images and SAR Images*. Technical Report, v. 4704, INRIA, France, 2002.
- [2] ACAR, R. & VOGEL, C. R. *Analysis of bounded variation penalty methods for ill-posed problems*. Inverse Problems, v.10(6), p. 1217-1229, 1994.
- [3] ALVAREZ, L., LIONS, P.L. & MOREL, J.M. *Image selective smoothing and edge detection by nonlinear diffusion II*. SIAM J. Numer. Anal., v. 29, p. 845-866, 1992.
- [4] AXELSSON, O. & BARKER, V. A. *Finite element solution of boudary value problems. Theory and computation*. Academic Press, Orlando, Flórida, 1984.
- [5] BERTALMIO, M., SAPIRO, G., CASELLES, V. & BALLESTER, C. In Proc. 27th Ann. Conf. on Comput. Graph. & Interactive Tech., p. 417-427, 2000.
- [6] CHAMBOLLE, A. *An Algorithm for Total Variation Minimization and Applications*. J. Math. Imaging Vision, v. 20, p. 89-97, 2004.
- [7] CHAMBOLLE, A., CASELLES, V., NOVAGA. M., CREMERS, D. & POCK, T. *An introduction to Total Variation for Image Analysis*. Theoretical Foundations and Numerical Methods for Sparse Recovery (M. Fornasier, ed.), Radon Series on Computational and Applied Methematics, De Gruyter Verland, p. 263-340, 2009.
- [8] CHAMBOLLE, A. & LIONS, P. L. *Image recovery via total variation minimization and related problems*. Numer. Math., v. 76(2), p. 167-188, 1997.
- [9] CHAN, T., GOLUB, G., & MULET, P. *A Nonlinear Primal-dual Method for Total Variation-based Image Restoration*. SIAM J. Sci. Comp., v. 20, p. 1964-1977, 1999.
- [10] CHAN, T. F., SHEN, J. & VESE, L. *Variational PDE models in image processing*. Notices Amer. Math. Soc., v. 50, p. 14-26, 2002.

- [11] CHAN, T. & SHEN, J. *Mathematical Models of Local Non-texture Inpaintings*. SIAM J. Appl. Math., v. 62, p. 1019-1043, 2001.
- [12] CHAN, T.F. & VESE, L.A. *Active contours without edges*. IEEE Trans. on Image Processing, v. 10(2), p. 266-277, 2001.
- [13] CHEN, Y., LEVINE, S. & RAO, M. *Variable exponent, linear growth functionals in image restoration*. SIAM Journal of Applied Mathematics, Vol. 66, No. 4, p. 1383-1406, 2006.
- [14] CHEN, Y. & WUNDERLI, T. *Adaptive total variation for image restoration in BV space*. J. Math. Anal. Appl. v. 272, p. 117-137, 2002.
- [15] D'IPPÓLITO, K. M. *Estudo de métodos numéricos para eliminação de ruídos em imagens digitais*. Dissertação de Mestrado. Departamento de Ciências da Computação e Estatística, Unesp, São José do Rio Preto, 2006.
- [16] EKELAND, I. & TEMAM, R. *Convex Analysis and Variational Problems*. Amsterdam: North Holland, 1976.
- [17] EVANS, L. C. *Partial differential equations* CRC Press, Boca Raton, 1992.
- [18] GIUSTI, E. *Minimal surfaces and functions of bounded variation*, v. 80 of Monographs in Mathematics. Birkhauser Verlag. Basel, 1984.
- [19] GOLDFARD, D. & YIN, W. *Second-order cone programming methods for total variation-based image restoration*. SIAM J. Sci. Comput., v. 27, p. 622-645, 2005.
- [20] GOLDSTEIN, T. & OSHER, S. *The split Bregman algorithm for L1 Regularized problems*. UCLA CAM Report, p. 08-29, 2008.
- [21] ISNARD, C. *Introdução á Medida e Integração*. 1. ed. Rio de Janeiro: IMPA, 2007.
- [22] IZMAILOV, A. & SOLODOV, M. *Otimização*. IMPA, Rio de Janeiro, 2009.
- [23] JOST, J. & LI-JOST, X. *Calculus of variations (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, 1998.
- [24] KORPELEVICH, G. *The extragradient method for finding saddle points and other problems*. Ekonomika i Matematicheskie Metody, v. 12, p. 747-756, 1976.
- [25] MEDEIROS, L.A. & MIRANDA, M.M. *Espaços de Sobolev : iniciação aos problemas elíticos não homogêneos*. Rio de Janeiro: UFRJ. IM, 2000. 151p.

- [26] MEYER, Y. *Oscillating Patterns in Image Processing and in some Non- linear Evolution Equations*. AMS, 2001. (Lewis Memorial Lectures).
- [27] MORY, B. & ARDON, R. *Fuzzy Region Competition: A Convex Two-Phase Segmentation Framework*. Lecture Notes in Computer Science, v. 4485, p. 214-226, 2007.
- [28] MUMFORD, D. & SHAH, J. *Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems*. Communications on Pure and Applied Mathematics, v. 42, No. 5., p. 577-685, 1989.
- [29] PARAGIOS, N. & DERICHE, N. *Geodesic active regions and level set methods for supervised texture segmentation*. Int. Journal of Computer Vision, v. 46(3), p. 223-247, 2002.
- [30] RUDIN, L. I.; OSHER, S.; FATEMI, E. *Nonlinear total variation based noise removal algorithms*. Physica D, North-Holland, v. 60, p. 259-268, 1992.
- [31] SHEN, J. *A Stochastic-Variational Model for Soft Mumford-Shah Segmentation*. Int. Journal of Biomedical Imaging, v. 2006, p. 1-14, 2006.
- [32] SILVA, V. *Segmentação de Imagens via Competição entre Regiões Fuzzy - Abordagens Paramétricas e Não-Paramétricas*. 2011. 161 p. Dissertação (Mestrado em Ciências da Computação) - Faculdade de Computação, Universidade Federal de Uberlândia, Uberlândia-MG. 2011.
- [33] SILVA, D. H. *Um estudo sobre a minimização de funcionais de expoente variável aplicados à restauração de imagens digitais*. 2009. 78 f. Dissertação (Mestrado em Matemática) - Faculdade de Matemática, Universidade Federal de Uberlândia, Uberlândia-MG. 2009.
- [34] VESE, L. & OSHER, S. *Modeling Textures with Total Variation Minimization and Oscillating Patterns In Image Processing*. J. Math. Imaging Vision, v. 20, p. 7-18, 2004.
- [35] VESE, L.A. & CHAN, T.F. *A multiphase level set framework for image segmentation using the mumford and shah model*. Int. Journal of Computer Vision, v. 50(3), p. 271-293, 2002.
- [36] VOGEL, C. R. & OMAN, M. E. *Iterative methods for total variation denoising*. Iterative methods for total variation denoising, SIAM J. Sci. Stat. Comput., v. 17, p. 227-238, 1996.
- [37] XIU, N. & ZHANG, J. *Some recent advances in projection-type methods for variational inequalities*. J. of Comp. and Applied Math, v. 152, p. 559-585, 2003.

- [38] ZHU, M., & CHAN, T. *An Efficient Primal-Dual Hybrid Gradient Algorithm For Total Variation Image Restoration*. CAM Reports 08-34, UCLA, Center for Applied Math, 2008.
- [39] ZHU, S. & YUILLE, A. *Region Competition: unifying snakes, region growing, and bayes/MDL for multiband image segmentation*. IEEE Trans. On Pattern Analysis and Machine Intelligence, v. 18(9), p. 884-900, 1996.
- [40] WANG, Y., YIN, W. & ZHANG, Y. *A Fast Algorithm for Image Deblurring with Total Variation Regularization*. Rice University CAAM Technical Report TR07-10, 2007.