



Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Departamento de Engenharia Elétrica  
Programa de Pós-Graduação em Engenharia Elétrica

**Método Dialético de Busca e Otimização  
para Análise de Imagens de Ressonância  
Magnética**

Wellington Pinheiro dos Santos

Tese de Doutorado

Campina Grande  
setembro de 2009

Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Departamento de Engenharia Elétrica  
Programa de Pós-Graduação em Engenharia Elétrica

Wellington Pinheiro dos Santos

## **Método Dialético de Busca e Otimização para Análise de Imagens de Ressonância Magnética**

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da Universidade Federal de Campina Grande como requisito parcial para obtenção do grau de Doutor em Engenharia Elétrica.

Orientador: Prof. Dr. Francisco Marcos de Assis

Campina Grande  
setembro de 2009

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCC

S237m

Santos, Wellington Pinheiro dos

Método Dialético de busca e otimização para análise de imagens de ressonância magnética / Wellington Pinheiro dos Santos.— Campina Grande, 2009.

285 f. : il. color

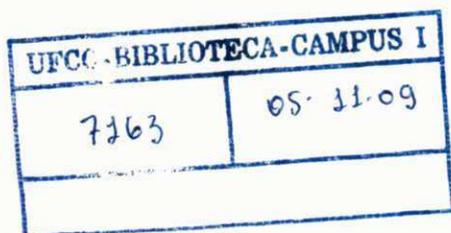
Tese (Doutorado em Engenharia Elétrica)- Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

Referências.

Orientador: Prof. Dr. Francisco Marcos de Assis.

1. Dialética 2. Filosofia da Práxis 3. Reconhecimento de Padrões 4. Imagens de Ressonância Magnética I. Título.

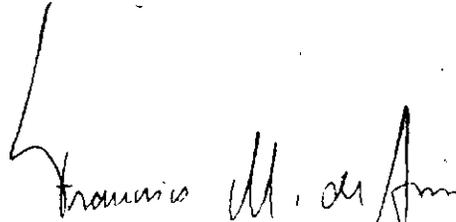
CDU 543.429.2(043)



**MÉTODO DIALÉTICO DE BUSCA E OTIMIZAÇÃO PARA ANÁLISE DE IMAGENS  
DE RESSONÂNCIA MAGNÉTICA**

**WELLINGTON PINHEIRO DOS SANTOS**

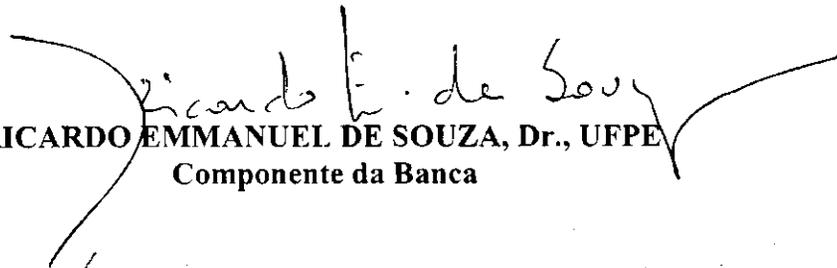
Tese Aprovada em 28.09.2009



**FRANCISCO MARCOS DE ASSIS, Dr., UFCG**  
Orientador



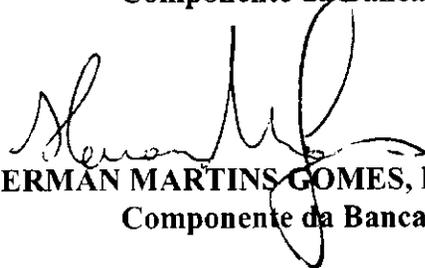
**RONALDO FUMIO HASHIMOTO, Dr., USP**  
Componente da Banca



**RICARDO EMMANUEL DE SOUZA, Dr., UFPE**  
Componente da Banca



**BENEMAR ALENCAR DE SOUZA, D.Sc., UFCG**  
Componente da Banca



**HERMAN MARTINS GOMES, Dr., UFCG**  
Componente da Banca

**CAMPINA GRANDE - PB**  
**SETEMBRO - 2009**

*A meus pais, Minervina e Orlando,  
e a Cecília, minha companheira*

# Agradecimentos

A Deus.

A meus pais Orlando Pinheiro dos Santos e Minervina Ângela Oliveira dos Santos, pelo constante apoio em todos os campos da minha vida.

A Cecília Cordeiro da Silva, minha companheira, pela paciência e estímulo nos momentos difíceis, que não foram poucos.

À minha avó Maria Anita de Oliveira, que nunca se furtou a dar conselhos importantes, mesmo em momentos inesperados, pois o remédio não precisa ser de gosto bom para bem curar.

A Francisco Marcos de Assis, pela cuidadosa e atenciosa orientação, pelo apoio, pela amizade, carinho e atenção que demonstrou para comigo em momentos fundamentais, e pelas divertidas conversas sobre Filosofia, História e assuntos ditos “aleatórios”. Também agradeço pela confiança desde o começo deste trabalho, principalmente quando saí da UFPE e decidi entrar na UFCG. É sabido que não é fácil confiar em um desconhecido, o que eu era naquele momento. Muito obrigado.

Aos colegas do IQuanta, Instituto de Computação Quântica da UFCG, pelo apoio dado ao desenvolvimento deste trabalho.

A Ricardo Emmanuel de Souza, pela co-orientação nos tempos do Doutorado na UFPE, pelo atual apoio e pela amizade que se manifestou em momentos inesperados de minha permanência no Departamento de Física da UFPE. Com ele aprendi mais sobre a vida, sobre a Física e sobre o que se pode esperar além do que entendemos por existência.

A Ascendino Flávio Dias e Silva, pela orientação nos tempos do Doutorado na UFPE e pelos importantes conselhos, tanto no campo profissional quanto no pessoal. Sua presença foi bastante importante para que eu pudesse definir, em todos esses anos, o perfil do pesquisador que hoje sou.

A Ana Lúcia Bezerra Candeias, pela amizade e pelas dicas importantes em momentos chave para o desenvolvimento deste trabalho.

A Antônio Fernando de Sousa Bezerra, pela amizade e pelo apoio onde se fez importante a opinião profissional de um especialista médico nos estudos de caso desenvolvidos neste trabalho.

A Eva Czajkowska, pelos deliciosos patês e pela tradução do russo para o português de vasta documentação de aplicativos de redes neurais para adaptação a este trabalho.

Aos amigos e colegas de trabalho Tiago Massoni, Márcio Lopes, Abel Guilhermino, Cristine Gusmão, Carlos Alexandre de Mello, Sérgio Soares, Adriano Lorena, Ricardo Massa, Carmelo Albanez, Fernando Buarque, e todos os que fizeram o Departamento de Sistemas Computacionais da Escola Politécnica de Pernambuco, Universidade de Pernambuco, de 2006

a 2009, e à direção da Escola Politécnica de Pernambuco, pelo apoio e solidariedade.

Aos meus estudantes de Iniciação Científica, do curso de Graduação em Engenharia da Computação da Escola Politécnica de Pernambuco, Universidade de Pernambuco, Henrique Specht, Priscilla Mendes, Havana Alves, Rafael Galvão e Hélio Bentzen, pelo apoio na geração dos resultados de otimização para as funções de teste utilizadas neste trabalho.

Aos amigos Leonardo Rezende, Katerina Czajkowska, Edilene Moura, Douglas Vitoreti, Sérgio de Lemos Campello, Marcelo Galdino, William Menezes, Amanda Gondim, Thiago Modenesi, Iron Mendes, Álvaro Celso da Silva, Gustavo Gilson, Izaías Torquato, Edson Pimentel, Senomar Teixeira (*in memorian*), José Iba (*in memorian*), Marcílio Feitosa, Rodson Ricardo, Gecionny Pinto, Dom Sebastião Armando Gameleira, David Lira, e tantos outros que, nem este pequeno espaço, nem minha memória muito limitada permitem ser mencionados aqui. A todos estes meus sinceros agradecimentos.

Àqueles que por algum motivo não estão mais presentes em minha vida, meus agradecimentos pelo tempo que estiveram comigo, pois muito do que aprendi com vocês tem me servido de freqüente inspiração para que eu possa cada vez mais me colocar acima de todo individualismo e lutar para fazer de minha vida algo de valor para meu próximo e para mim mesmo, ajudando a transformar as estruturas injustas da sociedade e melhorar o frágil mundo em que vivemos através de meus talentos, dons e conhecimento.

Àqueles a quem Deus decidiu colocar na minha caminhada como pedras e obstáculos. Obrigado por terem me ensinado o valor das vitórias.

E, finalmente, a todos aqueles que já partiram e, de forma direta ou indireta, contribuíram para a conclusão deste trabalho, pois como disse uma vez Isaac Newton, estamos na verdade “montados sobre os ombros de gigantes”.

Muito grato.

*É preciso sonhar,  
mas com a condição de crer em nosso sonho,  
de observar com atenção a vida real,  
de confrontar a observação com nosso sonho,  
de realizar escrupulosamente nossas fantasias.  
Sonhos, acredite neles.*

—V. I. ULIANOV (Que Fazer?, 1902)

# Resumo

A análise de imagens multiespectrais é um campo relativamente promissor, com aplicações em diversas áreas, como sensoriamento remoto e imagens médicas. A Biologia, a Psicologia e as Ciências Sociais estão intrinsecamente ligadas aos fundamentos do desenvolvimento de métodos e algoritmos em Inteligência Computacional, como se pode facilmente notar em estratégias como algoritmos genéticos, programação evolucionária e otimização por enxame de partículas. Contudo, a Filosofia ainda parece ser considerada uma forma de conhecimento enigmática, apesar do poder de generalização e da natureza sistemática de seus métodos investigativos tais como a dialética. Uma premissa fundamental deste trabalho é que a Filosofia pode ser considerada também uma fonte de inspiração. Neste trabalho é proposto um classificador não supervisionado baseado na dialética como definida por Hegel e a escola filosófica da Filosofia da Práxis, com o fim de classificar imagens multiespectrais sintéticas de ressonância magnética. Neste trabalho é demonstrado que esse classificador pode alcançar resultados equivalentes àqueles obtidos pelos mapas auto-organizados de Kohonen. Também neste trabalho foram construídos novos métodos de otimização baseados na dialética: uma versão dita canônica e outra versão baseada na aplicação do Princípio da Máxima Entropia. Esses métodos foram avaliados usando funções de teste e aplicados à otimização do classificador k-médias segundo índices de validade de agrupamento, atingindo bons resultados quando comparados com o k-médias canônico. Os métodos dialéticos de classificação foram também empregados em dois estudos de caso: um deles envolvendo o estudo do progresso da doença de Alzheimer, e o outro abordando a detecção de regiões ativadas em imagens funcionais, mostrando que o classificador dialético pode atingir bons resultados quando o número inicial de classes é desconhecido.

**Palavras-chave:** dialética, Filosofia da Práxis, métodos de otimização e busca, reconhecimento de padrões, imagens de ressonância magnética

# Abstract

Multispectral image analysis is a relatively promising field of research with applications in several areas, such as remote sensing and medical imaging. Biology, Psychology and Social Sciences are intrinsically connected to the very roots of the development of algorithms and methods in Computational Intelligence, as it is easily seen in approaches like genetic algorithms, evolutionary programming and particle swarm optimization. However, Philosophy appears to be still considered a sort of enigmatic knowledge, despite the power of generalization and the systematic nature of investigative methods like dialectics. Here we claim that Philosophy can be also considered as a source of inspiration. In this work we propose a new non-supervised classifier based on dialectics as defined by Hegel's works and the philosophical school of the Philosophy of Praxis, to classify synthetic multispectral magnetic resonance images. This work proved that such a classifier can reach results as good as those obtained by Kohonen's self-organized networks. Herein this work we also generated new optimization methods based on dialectics: a canonical version and a version based on the Principle of Maximum Entropy. These methods were evaluated using benchmark functions and applied to optimize k-means clusters according to cluster validity indexes, getting good results when compared with canonical k-means. Dialectical classification was also performed in two case studies, involving the study of the progress of Alzheimer's disease and detection of activated areas in functional imaging, demonstrating that the dialectical classifier can get good classification results when the initial number of classes is unknown.

**Keywords:** dialectics, philosophy of praxis, optimization methods, pattern recognition, magnetic resonance imaging

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Introdução	1
1.2	Motivação e Objetivos	2
1.2.1	Objetivos Gerais	3
1.2.2	Objetivos Específicos	4
1.3	Metodologia	5
1.4	Organização do Trabalho	6
<b>2</b>	<b>Imagens de Ressonância Magnética Nuclear</b>	<b>9</b>
2.1	Introdução	9
2.2	Princípios de Ressonância Magnética	9
2.2.1	Dinâmica da Magnetização	14
2.2.2	Pulso de Rádio-Frequência	15
2.2.3	Equações de Bloch	16
2.2.4	Detecção de Sinal	17
2.3	Imagens de Ressonância Magnética	21
2.3.1	Equação da Imagem	22
2.3.2	Sequências de Pulsos	23
2.3.2.1	Sequência SE	25
2.3.2.2	Sequência EPI	25
2.3.3	Aplicações	25
2.3.3.1	Imagens de Difusão	28
2.3.3.2	Imagens Funcionais e o Efeito BOLD	29
2.4	Conclusão	31
<b>3</b>	<b>A Concepção Dialética da Realidade</b>	<b>33</b>
3.1	Introdução	33
3.2	História da Concepção Dialética	34
3.3	O Método Dialético Materialista	40
3.3.1	Principais Categorias	42
3.3.2	Modelos e Aplicações	46
3.3.2.1	IR e o Método Dialético	46
3.3.2.2	Dialética e Apoio à Decisão	47
3.3.2.3	Modelagem por Sistemas Dinâmicos Não Lineares	48
3.3.2.4	Dialética e Teoria da Catástrofe	48

3.3.2.5	Dialética e Teoria do Caos	50
3.3.2.6	Dialética e Mapas de K-Médias	51
3.3.2.7	Dialética e Mapas Auto-Organizados de Kohonen	51
3.3.2.8	Dialética e Mapas Fuzzy C-Médias	54
3.3.2.9	Dialética e Sistemas Dinâmicos Complexos Emergentes	55
3.4	O Método Dialético Objetivo	56
3.4.1	Definição Geral	56
3.4.2	Proposta de Definição Específica	57
3.4.3	Classificadores Dialéticos Objetivos	60
3.4.3.1	Treinamento	60
3.4.3.2	Classificação	62
3.5	Conclusão	63
<b>4</b>	<b>O Método Dialético adaptado a Busca e Otimização</b>	<b>65</b>
4.1	Introdução	65
4.2	Princípio da Máxima Entropia	66
4.3	Entropia Fuzzy	69
4.4	Computação Evolucionária	72
4.4.1	Algoritmos Genéticos	73
4.4.2	Programação Evolucionária	76
4.4.2.1	Algoritmo CEP	78
4.4.2.2	Algoritmo FEP	79
4.4.2.3	Algoritmo LEP	80
4.4.2.4	Algoritmo SPMEP	80
4.4.2.5	Comentários	80
4.4.3	Otimização por Enxame de Partículas	81
4.5	Método Dialético de Otimização	84
4.5.1	Definição Geral	84
4.5.2	Algoritmo de Busca e Otimização	86
4.6	Análise pelo Método da Máxima Entropia	90
4.7	Aplicações em Classificação e Reconhecimento de Padrões	95
4.8	Conclusão	96
<b>5</b>	<b>Avaliação do Método Dialético de Otimização</b>	<b>99</b>
5.1	Introdução	99
5.2	Funções de Teste	99
5.2.1	Funções Unimodais	100
5.2.2	Funções Multimodais	101
5.2.3	Função Multiobjetivo	105
5.3	Resultados	106
5.4	Discussão e Conclusões	107

<b>6</b>	<b>Dialética e Quantização Vetorial de Imagens</b>	<b>119</b>
6.1	Introdução	119
6.2	Materiais e Métodos	120
6.2.1	Imagens Multiespectrais Sintéticas	120
6.2.2	Quantização Vetorial	120
6.2.3	Índices de Fidelidade	123
6.2.4	Medidas de Validade de Agrupamento	124
6.2.5	Métodos de Classificação Não Supervisionada	126
6.2.6	Classificadores Otimizados	127
6.3	Resultados e Discussões	127
6.3.1	Classificadores Dialéticos Objetivos	128
6.3.2	Classificadores K-Médias Otimizados pelo Método Dialético	133
6.4	Conclusões	136
<b>7</b>	<b>Classificação de Imagens de Difusão de Alzheimer</b>	<b>143</b>
7.1	Introdução	143
7.2	Materiais e Métodos	144
7.2.1	Imagens Ponderadas em Difusão e Mapas ADC	144
7.2.2	Análise do Mapa ADC usando Fuzzy C-Médias	147
7.2.3	Análise Multiespectral utilizando Redes Neurais	147
7.2.4	Análise Multiespectral pelo Mapa Fuzzy C-Médias	148
7.2.5	Análise Multiespectral pelo Classificador Dialético	148
7.2.6	Ferramentas Computacionais	149
7.3	Resultados	149
7.4	Discussão e Conclusões	157
<b>8</b>	<b>Detecção de Imagens Funcionais</b>	<b>167</b>
8.1	Introdução	167
8.2	Materiais e Métodos	168
8.2.1	Métodos de Detecção Estatísticos Multiespectrais	168
8.2.2	Métodos Multiespectrais Não Paramétricos	170
8.2.3	Método Não Paramétrico Monoespectral Proposto	171
8.3	Resultados	172
8.4	Discussão e Conclusões	178
<b>9</b>	<b>Conclusão e Trabalhos Futuros</b>	<b>183</b>
9.1	Comentários e Considerações	183
9.2	Contribuições	186
9.3	Trabalhos Futuros	187
<b>A</b>	<b>Resultados de Minimização de Funções de Teste</b>	<b>189</b>
A.1	Introdução	189
A.2	Resultados por Método	189
A.2.1	Algoritmo CEP	189
A.2.2	Algoritmo FEP	198

A.2.3	Algoritmo LEP	207
A.2.4	Algoritmo SPMEP	216
A.2.5	Algoritmo PSO	225
A.2.6	Método Dialético Canônico com Passos Variáveis	234
A.2.7	Método Dialético com Máxima Entropia e Passos Variáveis	243
A.2.8	Método Dialético Canônico com Passos Fixos	252
A.2.9	Método Dialético com Máxima Entropia e Passos Fixos	261
<b>Referências Bibliográficas</b>		<b>270</b>

# Lista de Figuras

2.1	Movimento de precessão do momento magnético $\vec{\mu}$ em torno do campo magnético $\vec{B} = B_0\hat{z}$ , com velocidade angular $\omega_0$	10
2.2	Níveis possíveis de energia para um sistema de <i>spin</i> 1/2	12
2.3	<i>Spin</i> paralelo a $\vec{B}$	12
2.4	<i>Spin</i> antiparalelo a $\vec{B}$	13
2.5	Movimento de precessão da magnetização resultante $\vec{M}$ em torno de $\vec{B} = B_0\hat{z}$	14
2.6	Magnetização $\vec{M}$ no referencial girante, no instante imediatamente após o desligamento do pulso de RF de $\alpha$ radianos	16
2.7	Magnetização transversal $M_{xy}(t) = M_{\perp}(t)$ para $M_{\perp}(0) = M_0$	18
2.8	Magnetização longitudinal $M_z(t)$ para $M_z(0) = 0$	18
2.9	Exemplo de FID resultante de um pulso de RF de $\pi/2$ radianos	19
2.10	Diagrama simplificado de um sistema de imagens de RM	24
2.11	Cobertura do plano $k$ para um determinado $k_z$ pela sequência de pulsos SE	26
2.12	Sequência de pulsos SE	27
2.13	Cobertura do plano $k$ para um determinado $k_z$ pela sequência de pulsos EPI	27
2.14	Sequência de pulsos EPI	28
2.15	Sequência de pulsos SE Stejskal-Tanner	29
3.1	Heráclito (540-470 a.C.)	34
3.2	Platão (428-347 a.C.)	35
3.3	“A Escola de Atenas”, de Rafael Sanzio, exposta no Palácio Apostólico do Vaticano (1506-1510)	36
3.4	Sócrates (469-399 a.C.)	36
3.5	Aristóteles (384-322 a.C.)	37
3.6	Yin e Yang: símbolo taoísta da dialética entre opostos	38
3.7	Georg Wilhelm Friederich Hegel (1770-1831)	39
3.8	Karl Marx (1818-1883)	41
3.9	Friederich Engels (1820-1895)	42
3.10	Antonio Gramsci (1891-1937)	44
4.1	Claude Elwood Shannon, engenheiro eletrônico, matemático e pai da Teoria da Informação (1916-2001)	66
4.2	Charles Robert Darwin, um dos pais da Teoria da Evolução (1809-1882)	73
4.3	Gregor Mendel, monge agostiniano, sacerdote e pai da Genética (1822-1884)	74
4.4	Fluxograma geral do GA	75
4.5	Fluxograma geral dos algoritmos EPs	77

4.6	Fluxograma geral do PSO	83
4.7	Um determinado vetor bidimensional $\mathbf{x}$ , representando a tese, e seu oposto $\mathbf{y} = \check{\mathbf{x}}$ , modelando sua antítese absoluta	86
4.8	Caso bidimensional de par de vetores tese e antítese, $\mathbf{w}_p$ and $\mathbf{w}_q$ , e suas respectivas sínteses, $\mathbf{w}_u$ e $\mathbf{w}_v$	87
4.9	Fluxograma geral do ODM adaptado a busca e otimização	91
6.1	Imagem de RM da fatia 97 ponderada em PD	121
6.2	Imagem de RM da fatia 97 ponderada em $T_1$	121
6.3	Imagem de RM da fatia 97 ponderada em $T_2$	121
6.4	Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD, $T_1$ e $T_2$	122
6.5	Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD, $T_1$ e $T_2$ (a) e resultados de classificação usando os métodos KO (b), CM (c), KM (d), ODC-PME (e) e ODC-CAN (f)	129
6.6	Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD, $T_1$ e $T_2$ (a) e resultados de quantização usando os métodos KO (b), CM (c), KM (d), ODC-PME (e) e ODC-CAN (f)	130
6.7	Resultados de $\varepsilon_{ME}$ em função do ruído percentual (ODC)	131
6.8	Resultados de $\varepsilon_{MAE}$ em função do ruído percentual (ODC)	132
6.9	Resultados de $\varepsilon_{RMSE}$ em função do ruído percentual (ODC)	133
6.10	Resultados de $\varepsilon_{PSNR}$ em função do ruído percentual (ODC)	134
6.11	Graus de similaridade do método ODC-PME em relação aos métodos KO, CM e KM, usando o teste de $\chi^2$ , em função do ruído percentual	135
6.12	Graus de similaridade do método ODC-CAN em relação aos métodos KO, CM e KM, usando o teste de $\chi^2$ , em função do ruído percentual	136
6.13	Resultados de $\varepsilon_{ME}$ em função do ruído percentual (ODM)	137
6.14	Resultados de $\varepsilon_{MAE}$ em função do ruído percentual (ODM)	137
6.15	Resultados de $\varepsilon_{RMSE}$ em função do ruído percentual (ODM)	138
6.16	Resultados de $\varepsilon_{PSNR}$ em função do ruído percentual (ODM)	138
6.17	Resultados de $J_e$ em função do ruído percentual	139
6.18	Resultados de $J_o$ em função do ruído percentual	139
6.19	Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD, $T_1$ e $T_2$ (a) e resultados de classificação usando os métodos KO (b), CM (c), KM (d), IC-CAN-KM (e), IC-PME-KM (f), EQ-CAN-KM (g) e EQ-PME-KM (h)	140
6.20	Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD, $T_1$ e $T_2$ (a) e resultados de quantização usando os métodos KO (b), CM (c), KM (d), IC-CAN-KM (e), IC-PME-KM (f), EQ-CAN-KM (g) e EQ-PME-KM (h)	141
7.1	Imagens da fatia 13 (esquerda) e volumes (direita) ponderados em difusão, com expoentes de $0 \text{ s/mm}^2$ (a, e), $500 \text{ s/mm}^2$ (b, f) e $1000 \text{ s/mm}^2$ (c, g), e mapa ADC (d, h)	145

7.2	Conjunto de treinamento (a) e imagem verdade da fatia 13 gerada pela classificação usando a rede polinomial (b)	149
7.3	Fatia 13 da classificação usando PO (a), ADC-CM (b), MLP (c), RBF (d), KO (e), LVQ (f), CM (g), e ODC (h)	151
7.4	Volumes de classificação usando PO (a), ADC-CM (b), MLP (c), RBF (d), KO (e), LVQ (f), CM (g), e ODC (h)	152
7.5	Fatia 13 da classificação pelo classificador dialético objetivo antes (a) e depois (b) da pós-rotulação	153
7.6	Comportamento da acurácia global $\phi$ em função da fatia $s$	154
7.7	Comportamento do índice $\kappa$ em função da variação da fatia $s$	155
7.8	Fatia 13 dos volumes de líquido (a), de substância cinzenta (b), de substância branca (c), e de líquido e substância cinzenta (d) detectados pelo ODC	156
7.9	Espectros morfológicos da fatia 13 do volume de líquido, de substância branca e de substância cinzenta detectados pelo ODC	157
7.10	Distribuições morfológicas acumuladas da fatia 13 do volume de líquido, de substância branca e de substância cinzenta detectados pelo ODC	158
7.11	Volumes da classe líquido detectados pelo CM (a), MLP (b), KO (c), LVQ (d), RBF (e) e PO (f)	159
7.12	Espectros morfológicos da fatia 13 do volume da classe líquido detectados pelos métodos CM, MLP e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC	160
7.13	Distribuições morfológicas acumuladas da fatia 13 do volume da classe líquido detectados pelos métodos CM, MLP e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC	161
7.14	Espectros morfológicos da fatia 13 do volume da classe líquido detectados pelos métodos KO, LVQ, RBF, PO e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC	162
7.15	Distribuições morfológicas acumuladas da fatia 13 do volume da classe líquido detectados pelos métodos KO, LVQ, RBF, PO e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC	163
7.16	Conjunto de treino (b) sobre o resultado do ODC (a) e novas classificações usando MLP (c), RBF (d), LVQ (e) e PO (f)	164
8.1	Mapa temático de ativação do teste $t$ de Student	172
8.2	Composição volumétrica do mapa temático de ativação do teste $t$ de Student	173
8.3	Mapa temático de ativação do teste $t$ emparelhado	173
8.4	Composição volumétrica do mapa temático de ativação do teste $t$ emparelhado	174
8.5	Mapa temático de ativação do teste $t$ combinado	174
8.6	Composição volumétrica do mapa temático de ativação do teste $t$ combinado	175
8.7	Imagem diferença modificada	175
8.8	Composição volumétrica da imagem diferença modificada	176
8.9	Mapa temático de ativação resultante da classificação da imagem normalizada pelo mapa auto-organizado de Kohonen	177

8.10	Composição volumétrica do mapa temático de ativação resultante da classificação da imagem normalizada pelo mapa auto-organizado de Kohonen	178
8.11	Mapa temático de ativação resultante da classificação da imagem diferença modificada pelo mapa auto-organizado de Kohonen	179
8.12	Composição volumétrica do mapa temático de ativação resultante da classificação da imagem diferença modificada pelo mapa auto-organizado de Kohonen	180
8.13	Mapa temático de ativação resultante da classificação da imagem diferença modificada pelo classificador dialético objetivo	181
8.14	Composição volumétrica do mapa temático de ativação resultante da classificação da imagem diferença modificada pelo classificador dialético objetivo	182

## Lista de Tabelas

5.1	Coeficientes da função de Kovalik	103
5.2	Coeficientes da função de Hartman $f_{19}$	104
5.3	Coeficientes da função de Hartman $f_{20}$	104
5.4	Coeficientes de Shekel	105
5.5	Valores limite das funções	106
5.6	Moda do número de avaliações da função objetivo, considerando uma população inicial de 20 indivíduos	107
5.7	Moda do número de avaliações da função objetivo, considerando uma população inicial de 50 indivíduos	108
5.8	Moda do número de avaliações da função objetivo, considerando uma população inicial de 100 indivíduos	109
5.9	Taxa de sucesso (%), considerando uma população inicial de 20 indivíduos	110
5.10	Taxa de sucesso (%), considerando uma população inicial de 50 indivíduos	111
5.11	Taxa de sucesso (%), considerando uma população inicial de 100 indivíduos	112
6.1	Resultados de quantização expressos pelos índices de fidelidade, para 0% de ruído	128
6.2	Graus de similaridade entre o método ODC-PME e os métodos KO, CM e KM, de acordo com o teste de hipótese nula F, para 0% de ruído	129
6.3	Graus de similaridade entre o método ODC-PME e os métodos KO, CM e KM, de acordo com o teste de $\chi^2$ , em função do ruído	130
6.4	Graus de similaridade entre o método ODC-CAN e os métodos KO, CM e KM, de acordo com o teste de hipótese nula F, para 0% de ruído	131
6.5	Graus de similaridade entre o método ODC-CAN e os métodos KO, CM e KM, de acordo com o teste de $\chi^2$ , em função do ruído	132
7.1	Acurácia global $\phi$ (%) e índice $\kappa$ para os métodos de classificação	153
7.2	Volumes percentuais e razão fluido-substância para os métodos de classificação	153
7.3	Índice de generalização dos métodos para as demais bandas	154
7.4	Acurácia global $\phi$ (%) e índice $\kappa$ para os métodos de classificação treinados usando o resultado do ODC	158
7.5	Áreas percentuais e razão fluido-substância para os métodos de classificação treinados usando o resultado do ODC	158
8.1	Avaliação de desempenho usando a taxa de acerto global $\phi$ e o índice $\kappa$	176
8.2	Avaliação de desempenho usando as taxas de detecção	177

A.1	Média, desvio médio e mediana dos valores mínimos (CEP)	190
A.2	Média, desvio médio e mediana dos valores mínimos (CEP)	190
A.3	Média, desvio médio e mediana dos valores mínimos (CEP)	191
A.4	Média, desvio médio e mediana dos valores mínimos (CEP)	191
A.5	Faixa dos valores mínimos encontrados (CEP)	192
A.6	Faixa dos valores mínimos encontrados (CEP)	192
A.7	Faixa dos valores mínimos encontrados (CEP)	193
A.8	Faixa dos valores mínimos encontrados (CEP)	193
A.9	Número de iterações para minimização (CEP)	194
A.10	Número de iterações para minimização (CEP)	194
A.11	Número de iterações para minimização (CEP)	195
A.12	Número de iterações para minimização (CEP)	195
A.13	Número de avaliações da função (CEP)	196
A.14	Número de avaliações da função (CEP)	196
A.15	Número de avaliações da função (CEP)	197
A.16	Número de avaliações da função (CEP)	197
A.17	Média, desvio médio e mediana dos valores mínimos (FEP)	198
A.18	Média, desvio médio e mediana dos valores mínimos (FEP)	199
A.19	Média, desvio médio e mediana dos valores mínimos (FEP)	200
A.20	Média, desvio médio e mediana dos valores mínimos (FEP)	200
A.21	Faixa dos valores mínimos encontrados (FEP)	201
A.22	Faixa dos valores mínimos encontrados (FEP)	201
A.23	Faixa dos valores mínimos encontrados (FEP)	202
A.24	Faixa dos valores mínimos encontrados (FEP)	202
A.25	Número de iterações para minimização (FEP)	203
A.26	Número de iterações para minimização (FEP)	203
A.27	Número de iterações para minimização (FEP)	204
A.28	Número de iterações para minimização (FEP)	204
A.29	Número de avaliações da função (FEP)	205
A.30	Número de avaliações da função (FEP)	205
A.31	Número de avaliações da função (FEP)	206
A.32	Número de avaliações da função (FEP)	206
A.33	Média, desvio médio e mediana dos valores mínimos (LEP)	207
A.34	Média, desvio médio e mediana dos valores mínimos (LEP)	208
A.35	Média, desvio médio e mediana dos valores mínimos (LEP)	209
A.36	Média, desvio médio e mediana dos valores mínimos (LEP)	209
A.37	Faixa dos valores mínimos encontrados (LEP)	210
A.38	Faixa dos valores mínimos encontrados (LEP)	210
A.39	Faixa dos valores mínimos encontrados (LEP)	211
A.40	Faixa dos valores mínimos encontrados (LEP)	211
A.41	Número de iterações para minimização (LEP)	212
A.42	Número de iterações para minimização (LEP)	212
A.43	Número de iterações para minimização (LEP)	213

A.44	Número de iterações para minimização (LEP)	213
A.45	Número de avaliações da função (LEP)	214
A.46	Número de avaliações da função (LEP)	214
A.47	Número de avaliações da função (LEP)	215
A.48	Número de avaliações da função (LEP)	215
A.49	Média, desvio médio e mediana dos valores mínimos (SPMEP)	216
A.50	Média, desvio médio e mediana dos valores mínimos (SPMEP)	217
A.51	Média, desvio médio e mediana dos valores mínimos (SPMEP)	218
A.52	Média, desvio médio e mediana dos valores mínimos (SPMEP)	218
A.53	Faixa dos valores mínimos encontrados (SPMEP)	219
A.54	Faixa dos valores mínimos encontrados (SPMEP)	219
A.55	Faixa dos valores mínimos encontrados (SPMEP)	220
A.56	Faixa dos valores mínimos encontrados (SPMEP)	220
A.57	Número de iterações para minimização (SPMEP)	221
A.58	Número de iterações para minimização (SPMEP)	221
A.59	Número de iterações para minimização (SPMEP)	222
A.60	Número de iterações para minimização (SPMEP)	222
A.61	Número de avaliações da função (SPMEP)	223
A.62	Número de avaliações da função (SPMEP)	223
A.63	Número de avaliações da função (SPMEP)	224
A.64	Número de avaliações da função (SPMEP)	224
A.65	Média, desvio médio e mediana dos valores mínimos (PSO)	225
A.66	Média, desvio médio e mediana dos valores mínimos (PSO)	226
A.67	Média, desvio médio e mediana dos valores mínimos (PSO)	227
A.68	Média, desvio médio e mediana dos valores mínimos (PSO)	227
A.69	Faixa dos valores mínimos encontrados (PSO)	228
A.70	Faixa dos valores mínimos encontrados (PSO)	228
A.71	Faixa dos valores mínimos encontrados (PSO)	229
A.72	Faixa dos valores mínimos encontrados (PSO)	229
A.73	Número de iterações para minimização (PSO)	230
A.74	Número de iterações para minimização (PSO)	230
A.75	Número de iterações para minimização (PSO)	231
A.76	Número de iterações para minimização (PSO)	231
A.77	Número de avaliações da função (PSO)	232
A.78	Número de avaliações da função (PSO)	232
A.79	Número de avaliações da função (PSO)	233
A.80	Número de avaliações da função (PSO)	233
A.81	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)	234
A.82	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)	235
A.83	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)	236
A.84	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)	236
A.85	Faixa dos valores mínimos encontrados (ODM-CAN-V)	237
A.86	Faixa dos valores mínimos encontrados (ODM-CAN-V)	237

A.87	Faixa dos valores mínimos encontrados (ODM-CAN-V)	238
A.88	Faixa dos valores mínimos encontrados (ODM-CAN-V)	238
A.89	Número de iterações para minimização (ODM-CAN-V)	239
A.90	Número de iterações para minimização (ODM-CAN-V)	239
A.91	Número de iterações para minimização (ODM-CAN-V)	240
A.92	Número de iterações para minimização (ODM-CAN-V)	240
A.93	Número de avaliações da função (ODM-CAN-V)	241
A.94	Número de avaliações da função (ODM-CAN-V)	241
A.95	Número de avaliações da função (ODM-CAN-V)	242
A.96	Número de avaliações da função (ODM-CAN-V)	242
A.97	Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)	243
A.98	Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)	244
A.99	Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)	245
A.100	Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)	245
A.101	Faixa dos valores mínimos encontrados (ODM-PME-V)	246
A.102	Faixa dos valores mínimos encontrados (ODM-PME-V)	246
A.103	Faixa dos valores mínimos encontrados (ODM-PME-V)	247
A.104	Faixa dos valores mínimos encontrados (ODM-PME-V)	247
A.105	Número de iterações para minimização (ODM-PME-V)	248
A.106	Número de iterações para minimização (ODM-PME-V)	248
A.107	Número de iterações para minimização (ODM-PME-V)	249
A.108	Número de iterações para minimização (ODM-PME-V)	249
A.109	Número de avaliações da função (ODM-PME-V)	250
A.110	Número de avaliações da função (ODM-PME-V)	250
A.111	Número de avaliações da função (ODM-PME-V)	251
A.112	Número de avaliações da função (ODM-PME-V)	251
A.113	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)	252
A.114	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)	253
A.115	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)	254
A.116	Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)	254
A.117	Faixa dos valores mínimos encontrados (ODM-CAN-F)	255
A.118	Faixa dos valores mínimos encontrados (ODM-CAN-F)	255
A.119	Faixa dos valores mínimos encontrados (ODM-CAN-F)	256
A.120	Faixa dos valores mínimos encontrados (ODM-CAN-F)	256
A.121	Número de iterações para minimização (ODM-CAN-F)	257
A.122	Número de iterações para minimização (ODM-CAN-F)	257
A.123	Número de iterações para minimização (ODM-CAN-F)	258
A.124	Número de iterações para minimização (ODM-CAN-F)	258
A.125	Número de avaliações da função (ODM-CAN-F)	259
A.126	Número de avaliações da função (ODM-CAN-F)	259
A.127	Número de avaliações da função (ODM-CAN-F)	260
A.128	Número de avaliações da função (ODM-CAN-F)	260
A.129	Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)	261

A.130	Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)	262
A.131	Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)	263
A.132	Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)	263
A.133	Faixa dos valores mínimos encontrados (ODM-PME-F)	264
A.134	Faixa dos valores mínimos encontrados (ODM-PME-F)	264
A.135	Faixa dos valores mínimos encontrados (ODM-PME-F)	265
A.136	Faixa dos valores mínimos encontrados (ODM-PME-F)	265
A.137	Número de iterações para minimização (ODM-PME-F)	266
A.138	Número de iterações para minimização (ODM-PME-F)	266
A.139	Número de iterações para minimização (ODM-PME-F)	267
A.140	Número de iterações para minimização (ODM-PME-F)	267
A.141	Número de avaliações da função (ODM-PME-F)	268
A.142	Número de avaliações da função (ODM-PME-F)	268
A.143	Número de avaliações da função (ODM-PME-F)	269
A.144	Número de avaliações da função (ODM-PME-F)	269



## CAPÍTULO 1

# Introdução

### 1.1 Introdução

O imageamento por ressonância magnética nuclear é uma das áreas mais promissoras no campo das ciências aplicadas. Uma das suas principais aplicações é a obtenção de imagens de tecidos moles do corpo humano e de processos metabólicos [1]. Existem diferentes tipos de imagens de ressonância magnética, que se caracterizam tanto pelas sequências de pulsos utilizadas em sua aquisição quanto pelas diferentes aplicações a que se prestam. As imagens de ressonância magnética nuclear (RMN) ponderadas em difusão têm uma grande aplicação na avaliação da doença de Alzheimer, por conjugarem tanto informação anatômica quanto metabólica [1].

Uma área que também está crescendo bastante são as pesquisas em ressonância magnética funcional, onde se procura investigar as áreas do cérebro responsáveis pela execução de determinadas atividades e pelas respostas a estímulos externos específicos, permitindo localizar e mapear o cérebro humano por regiões funcionais. Isto é feito utilizando sequências rápidas para aquisição de imagens tomográficas e considerando o efeito BOLD (*Blood Oxygenation Level Dependent*, nível dependente da oxigenação do sangue), que resulta em artefatos nas imagens de RMN proporcionais aos níveis de desoxi-hemoglobina, ou seja, à taxa de atividade metabólica do cérebro [2, 3, 4, 5, 1, 6].

Assim, experimentos tais como a localização das áreas cerebrais responsáveis pela atividade motora de membros específicos do corpo humano são realizados visando dotar os especialistas de áreas como a Neurologia, a Psicologia e a Fisioterapia, de conhecimentos que melhorem o prognóstico, o diagnóstico e o tratamento de doenças degenerativas e de lesões cerebrais que comprometam os movimentos [5, 1].

A análise multiespectral de imagens é um campo de pesquisa relativamente promissor, com muitas aplicações em áreas as mais diversas, tais como a análise de imagens médicas. Contudo, grande parte dos métodos usualmente empregados são baseados em métodos estatísticos paramétricos. Tais métodos se tornam muito complexos e computacionalmente muito custosos quando aumenta a dimensionalidade do problema. Assim, as redes neurais artificiais e outros métodos de inteligência computacional podem ser utilizadas como métodos não paramétricos na classificação e análise de imagens multiespectrais.

Muitos métodos em inteligência computacional têm inspiração biológica, tais como as redes neurais e os algoritmos genéticos [7, 8, 9]. Contudo, poucos têm inspiração filosófica, ou seja, poucos buscam na Filosofia métodos que possam ser adaptados e transformados em algoritmos de inteligência computacional, muito embora buscar na Filosofia as fontes de inspiração para novos métodos não seja nada mais do que voltar à origem da própria ciência [10].

Este trabalho propõe o Método Dialético Objetivo, inspirado na Filosofia da Práxis [11, 12,

13], como uma ferramenta para busca e otimização e para classificação multiespectral de imagens estruturada na forma de um algoritmo evolucionário baseado na dinâmica das contradições entre pólos dialéticos dentro de um sistema. Como estudo de caso são utilizadas funções de teste para minimização e analisadas imagens multiespectrais compostas por imagens de ressonância magnética sintéticas e reais. Os resultados são comparados com aqueles obtidos usando redes neurais artificiais e mapas de k-médias.

## 1.2 Motivação e Objetivos

A principal motivação deste trabalho é a necessidade de métodos de classificação não supervisionados para problemas de classificação onde o número de classes presentes não é inicialmente conhecido. Uma abordagem na busca de classificadores que se prestem a esse tipo de problema é a investigação de classificadores adaptativos, que possam se ajustar ao problema e buscar as classes que sejam mais estatisticamente significativas. É natural pensar em classificadores ajustados por métodos evolutivos, tais como os algoritmos genéticos, ou em outras abordagens, tais como as redes neurais e outras famílias de técnicas biologicamente inspiradas.

No entanto, nota-se uma certa defasagem entre as fontes de inspiração da inteligência computacional e o que há de mais avançado em campos como a Filosofia. A inteligência computacional predominantemente busca inspiração para a construção de novos métodos inteligentes em sistemas biológicos. Já a Filosofia é dotada de métodos investigativos para modelagem e estudo de sistemas relativamente complexos que nunca tiveram uma formulação algorítmica de forma a despertar o interesse de pesquisadores em Engenharia Elétrica e Computação para potenciais aplicações. Essa lacuna talvez exista por muitas vezes não existirem espaços para o diálogo entre as diversas áreas do conhecimento, espaços livres das demarcações que encapsulam o conhecimento acadêmico em campos bem definidos e isolados. Isso dificulta a pesquisa de natureza multidisciplinar que vá além da simples aplicação das Engenharias nas outras áreas e que não se permite ser ontologicamente transformada pela dialética das interações entre os diferentes grupos de pesquisa.

Este trabalho busca mostrar que é possível desmistificar áreas do conhecimento como a Filosofia frente às Engenharias e à Computação e estudar sistematicamente os métodos de investigação filosófica para construir ferramentas de análise para aplicações em inteligência computacional e reconhecimento de padrões. Assim, este trabalho tenta buscar na Dialética Materialista, como definida pela Filosofia da Práxis, uma ferramenta para o estudo de sistemas dinâmicos adaptativos, sujeitos a processos tais como a evolução e a crise.

No campo das aplicações, uma das principais motivações para este trabalho é a ausência de informação objetiva que permita um melhor acompanhamento do progresso de doenças neurodegenerativas e, assim, permita definir novas estratégias de tratamento, controle e convivência com a doença. Esse acompanhamento pode se dar por meio da classificação anatômica de regiões em imagens de ressonância magnética ponderadas em difusão, como no caso da doença de Alzheimer.

Contudo, as imagens de ressonância magnética ponderadas em difusão não permitem a distinção entre estruturas com coeficientes de difusão muito próximos, o que exige a aquisição de mais imagens, segundo diferentes protocolos de aquisição, para extração de informação

anatômica precisa. O mapa de coeficientes de difusão aparente, montado a partir das imagens ponderadas em difusão, não consegue exibir a diferença entre *voxels* pertencentes à substância branca e *voxels* que fazem parte da substância cinzenta [1].

Desta forma, este trabalho procura também fornecer alternativas à análise do mapa de coeficientes de difusão aparente, uma vez que este mapa é altamente sensível a ruído e é pobre de informação anatômica. Este trabalho mostra que é possível obter mais informação anatômica a partir das imagens de difusão do que os mapas de coeficientes de difusão aparente mostram.

Ainda no campo das aplicações, outra motivação importante vem da parte da detecção de regiões de ativação em imagens funcionais de ressonância magnética: as aplicações de imagens funcionais têm crescido, auxiliando áreas como a Neurologia, a Psicologia e a Fisioterapia a entender melhor a resposta cerebral a estímulos externos e a neurodinâmica de diversos processos físicos e mentais [3, 4, 14, 15, 16, 17, 18]. Este trabalho mostra que é possível gerar novos métodos de detecção de regiões de ativação como uma alternativa aos métodos paramétricos e ao método *fuzzy c*-médias.

### 1.2.1 Objetivos Gerais

Este trabalho tem como objetivos gerais:

1. Estudar os princípios da Filosofia da Práxis e as tentativas de modelar as categorias filosóficas do método dialético e aplicações;
2. Estudar métodos de inteligência computacional para busca e otimização e funções de teste em problemas de minimização de funções, com o objetivo de construir um método evolucionário para busca e otimização baseado nas categorias da dialética;
3. Estudar aplicações de reconhecimento de padrões onde as eventuais características de um método de segmentação baseado no método dialético sejam importantes para detecção de singularidades e busca de classes estatisticamente significativas;
4. Estudar a aplicação de conceitos de Teoria da Informação, em especial o conceito de entropia, na construção de um método dialético de busca e otimização;
5. Estudar diferentes tipos de imagens de ressonância magnética e suas aplicações médicas e biológicas, em especial aquelas aplicações onde seja possível empregar abordagens baseadas na análise multiespectral de imagens, tendo em vista otimizar a extração de informação em aplicações clínicas;
6. Propor métodos dialéticos de busca e otimização que possam ser utilizados na obtenção de métodos de segmentação de imagens multiespectrais de ressonância magnética que sejam ótimos segundo determinados critérios de fidelidade da imagem quantizada em relação à imagem original ou determinadas medidas de validade de agrupamento;
7. Propor métodos dialéticos de segmentação para análise multiespectral de imagens de ressonância magnética, visando fornecer condições objetivas e subjetivas para otimizar a análise por parte do especialista.

### 1.2.2 Objetivos Específicos

Os objetivos específicos deste trabalho são:

1. Propor um método de busca e otimização baseado na dialética materialista, usando analogias geométricas para modelar matematicamente as categorias dialéticas, além de métodos evolucionários como os algoritmos genéticos, a programação evolucionária e a otimização por enxame de partículas, validando este método por meio da otimização de funções padrão de teste;
2. Utilizar o Princípio da Máxima Entropia para gerar uma nova versão do método de busca e otimização proposto e validar essa versão por meio da otimização de funções padrão de teste, comparando-a com a versão canônica;
3. Propor métodos de segmentação baseados no método dialético de busca e otimização, versão canônica e com entropia maximizada, desenvolvido neste trabalho que sejam ótimos segundo alguma medida de fidelidade da imagem quantizada à imagem original ou segundo alguma medida de validade do agrupamento e utilizar esses métodos na segmentação de imagens multiespectrais sintéticas compostas de imagens de ressonância magnética;
4. Propor a classificação multiespectral de imagens para auxiliar na avaliação do progresso da doença de Alzheimer a partir da composição de uma imagem multiespectral sintética cujas bandas correspondam às imagens ponderadas em difusão da fatia em estudo, para que o resultado da classificação seja uma alternativa ao mapa de coeficientes de difusão aparente na avaliação subjetiva efetuada pelo especialista;
5. Propor um novo método de classificação multiespectral baseado no Método Dialético Materialista da Filosofia da Práxis e investigar sua aplicação à análise de imagens de ressonância magnética ponderadas em difusão e de imagens funcionais;
6. Estudar as imagens funcionais de ressonância magnética e seu emprego no mapeamento funcional da atividade motora, tendo como estudo de caso imagens resultantes de um experimento de identificação da região responsável pelos movimentos do dedo indicador direito e da atenção visual direita;
7. Estudar os métodos clássicos de detecção de regiões de ativação em imagens funcionais, estatísticos paramétricos e não paramétricos, utilizando uma abordagem de análise multiespectral de imagens como alternativa à abordagem baseada em sequências temporais usualmente utilizada;
8. Propor novos métodos não paramétricos baseados na classificação não supervisionada pelo método dialético objetivo, proposto neste trabalho, e por mapas auto-organizados, buscando reduzir o problema de classificação multiespectral a um problema de classificação mono-espectral quando possível, ou seja, reduzir o problema à análise de imagem de banda única sempre que for possível.

### 1.3 Metodologia

Para a construção tanto do método dialético de busca e otimização quanto do classificador dialético propostos neste trabalho foi realizada uma extensa revisão bibliográfica contemplando o histórico e as definições da concepção dialética da realidade ao longo da história. Após isso, foi realizado um outro levantamento bibliográfico, só que desta vez contemplando os trabalhos que tivessem buscado algum tipo de inspiração na dialética para resolver problemas específicos ou elaborar modelos matemáticos e computacionais e métodos analíticos.

A revisão bibliográfica também contemplou o estudo de métodos de computação evolucionária, em especial os algoritmos genéticos, a programação evolucionária e a otimização por enxame de partículas, para elaboração do modelo matemático e computacional do método dialético proposto. Depois foi estudado o Princípio da Máxima Entropia para avaliar o algoritmo construído e refiná-lo, gerando mais uma versão do método dialético proposto. Os algoritmos resultantes foram implementados na linguagem orientada a objetos Object Pascal.

O método dialético de busca e otimização, em sua versão canônica e com entropia maximizada, construído a partir dos estudos anteriormente citados foi inicialmente testado usando 24 funções de teste (*benchmark functions*) que servem para emular diversas situações em problemas de otimização, que no caso modelam problemas unimodais e multimodais, de baixa e alta dimensionalidade, com e sem descontinuidades, para problemas uni-objetivo e multiobjetivo, comumente utilizadas em diversos trabalhos de otimização.

O método dialético de busca e otimização foi então utilizado para gerar métodos de segmentação de imagens ótimos em função de determinados critérios ou objetivos, que podem ser tanto índices de fidelidade da imagem quantizada resultante da classificação quanto medidas de validade de agrupamento. Para tanto, foram utilizadas 1086 imagens multiespectrais sintéticas de ressonância magnética com três bandas, compostas por 3258 imagens sagitais de 1 mm de espessura, resolução de 1 mm<sup>3</sup>, ponderadas em densidade de prótons,  $T_1$  e  $T_2$ , para um cérebro humano normal, obtidas por um sistema tomográfico de ressonância magnética com níveis de ruído de 0%, 1%, 3%, 5%, 7% e 9%, e sem inhomogeneidades de campo. Essas imagens foram geradas pelo simulador BrainWeb [19, 20].

O trabalho também envolveu dois estudos de caso de aplicações do classificador dialético objetivo usando imagens reais de ressonância magnética: análise anatômica de imagens ponderadas em difusão e detecção de regiões do córtex ativadas em imagens de ressonância funcional. Foram utilizadas 80 imagens de ressonância magnética axiais ponderadas em difusão, ou seja, 4 volumes cerebrais de 20 fatias axiais cada, sendo um desses volumes composto por mapas ADC, correspondentes a um paciente do sexo masculino, de 70 anos de idade, portador da doença de Alzheimer. Essas imagens foram obtidas a partir do banco de imagens clínicas do Laboratório de Imagens de Ressonância Magnética do Departamento de Física da Universidade Federal de Pernambuco, composto de imagens clínicas reais adquiridas de um tomógrafo clínico de 1,5 T. As imagens funcionais foram fornecidas pelo Departamento de Física e pelo Laboratório de Percepção Visual do Departamento de Psicologia da Universidade Federal de Pernambuco, adquiridas a partir de um tomógrafo clínico de 1,5 T durante o ano de 2004, no total de 60 imagens: 24 sob estímulo, 24 sem estímulo e 12 imagens de referência para corrigir os efeitos de movimento.

## 1.4 Organização do Trabalho

Este documento de tese está organizado da forma que segue:

O capítulo 2 apresenta noções básicas de ressonância magnética nuclear e do processo de aquisição de imagens, discorrendo sobre os princípios físicos do fenômeno da ressonância magnética nuclear, detecção de sinais de ressonância, sequências de pulsos, imagens ponderadas e aplicações.

O capítulo 3 apresenta uma breve revisão das origens e do desenvolvimento do pensamento dialético ao longo da história do conhecimento, apresentando suas raízes gregas e orientais e seu desenvolvimento até o século XX, passando pelo advento da filosofia alemã do século XIX. Em seguida, os principais conceitos, definições e categorias da dialética são apresentados. O capítulo também discorre pelas tentativas de utilizar conceitos da dialética em problemas de computação e engenharia e as propostas e ensaios de propostas para modelagem matemática e computacional das categorias da dialética. Por fim, é apresentada a proposta de método dialético deste trabalho, com aplicações imediatas em problemas de reconhecimento de padrões.

O capítulo 4 apresenta uma adaptação do método dialético desenvolvido no capítulo 3 a problemas de busca e otimização. Uma primeira adaptação é encapsulada no que se batizou de versão canônica, enquanto uma segunda versão é gerada pela aplicação do Princípio da Máxima Entropia à expressão de ajuste de pesos dos pólos dialéticos, partes integrantes fundamentais de um sistema dialético.

Já o capítulo 5 apresenta os resultados da aplicação das versões canônica e com entropia maximizada do método dialético de busca e otimização desenvolvido no capítulo 4 para otimização de funções, mais especificamente para minimização de 24 funções de teste unimodais e multi-modais, problemas uni-objetivo e multiobjetivo. Através da aplicação do método dialético a um conjunto amplo de problemas modelados pelas funções de teste, busca-se avaliar o comportamento geral do método dialético como método de otimização, para estimar suas qualidades e fraquezas.

O capítulo 6 ilustra a adaptação do método dialético de busca e otimização a problemas de reconhecimento de padrões, sendo utilizado esse método para encontrar os parâmetros de um classificador baseado no mapa de  $k$ -médias que representem um método de segmentação de imagens ótimo segundo determinados critérios objetivos, aqui apresentados como índices de fidelidade da imagem quantizada à imagem original e medidas de validade de agrupamento. Os resultados são gerados usando imagens sintéticas de ressonância magnética ponderadas em  $T_1$ , em  $T_2$  e em densidade de prótons formando imagens multiespectrais sintéticas. Esses resultados são comparados àqueles obtidos por meio do uso de outros classificadores, tais como o mapa de  $k$ -médias clássico, o mapa auto-organizado de Kohonen e o mapa *fuzzy*  $c$ -médias, em função dos índices de fidelidade e das medidas de validade de agrupamento.

No capítulo 7 apresenta-se um estudo de caso de classificação multiespectral de regiões anatômicas em imagens de ressonância magnética ponderadas em difusão, analisando o desempenho de diversos classificadores baseados em redes neurais, supervisionados e não supervisionados, e do classificador dialético objetivo. O intento é mostrar que a utilização da classificação multiespectral baseada em redes neurais, em abordagens híbridas e no classificador dialético objetivo permite o incremento de informação objetiva e subjetiva para o especialista,

podendo servir como uma alternativa a técnicas comumente utilizadas na análise do progresso de doenças neurodegenerativas como o mapa de coeficientes de difusão aparentes, muito utilizado na análise do progresso da doença de Alzheimer por imagens ponderadas em difusão. Neste capítulo também se procura mostrar que o classificador dialético objetivo é a melhor abordagem dentre as utilizadas neste trabalho, quando não se conhece o número de classes estatisticamente significativas presentes na imagem, dado que ele foi capaz de discernir entre a substância branca e a substância cinzenta em imagens de difusão do cérebro, o que não se considerava possível antes, uma vez que seus coeficientes de difusão são muito próximos.

No capítulo 8 apresenta-se um estudo de caso na detecção de regiões de ativação em imagens funcionais utilizando métodos estatísticos paramétricos, não paramétricos e o classificador dialético objetivo. Foram utilizadas imagens sintéticas e imagens reais, obtidas a partir de um experimento onde se desejava detectar as regiões de atividade neuronal responsáveis pela atividade motora do dedo indicador direito e pela percepção visual envolvida no experimento. Embora tal experimento não tenha uma aplicação prática, devido ao fato de já se conhecer com relativa precisão a região do córtex responsável pelas atividades de interesse, este experimento se presta à validação de novas técnicas computacionais de detecção de regiões ativadas, como as técnicas propostas neste trabalho, baseadas em abordagens mono espectrais e multiespectrais.

Por fim, no capítulo 9 são apresentadas as conclusões e comentadas as qualidades e as limitações das abordagens propostas. Também são comentadas as perspectivas de trabalhos futuros relacionados à pesquisa realizada, juntamente com as contribuições teóricas e práticas do autor ao estado da arte dos métodos de inteligência computacional, da computação evolucionária e da análise de imagens, realizadas durante o desenvolvimento deste trabalho e apresentadas na forma de artigos completos em conferências internacionais, capítulos de livro e artigos em periódicos internacionais.



# Imagens de Ressonância Magnética Nuclear

## 2.1 Introdução

O imageamento por ressonância magnética nuclear (MRI, *Magnetic Resonance Imaging*) é uma disciplina relativamente nova no campo das ciências aplicadas, tendo como abordagens principais a aquisição de imagens de tecidos moles do corpo humano e de processos metabólicos que ocupam uma posição chave nas ciências biomédicas [1].

As aplicações baseadas em imagens de ressonância magnética nuclear (RMN, ou simplesmente RM, ressonância magnética) têm crescido principalmente devido a fatores como a flexibilidade e o poder da técnica de lidar com diversas propriedades teciduais e a relativa segurança envolvida nos experimentos, dada a propriedade que tem a RM de ser não invasiva e de os campos magnéticos envolvidos tornarem possível o diagnóstico em indivíduos de quase todas as idades [1].

Contudo, a grande promessa da RM é a sua contribuição para um maior entendimento do corpo humano e de seus diversos sistemas, tanto no que tange à forma quanto à função, o que está relacionado, respectivamente, à aquisição de imagens anatômicas e funcionais.

## 2.2 Princípios de Ressonância Magnética

A ressonância magnética é um fenômeno encontrado em sistemas que possuem momento magnético e momento angular. O termo *ressonância* está relacionado com o fato de um fenômeno estar em sintonia com uma frequência natural do sistema magnético, que neste caso corresponde à frequência  $\omega_0$  de precessão do momento magnético  $\vec{\mu}$  em torno de um campo magnético estático  $\vec{B}$  (ver figura 2.1) [21].

Quando  $\vec{B} = B_0\hat{z}$ , o momento magnético  $\vec{\mu}$  precessiona o vetor  $\vec{B}$  com velocidade angular  $\omega_0$  dada por:

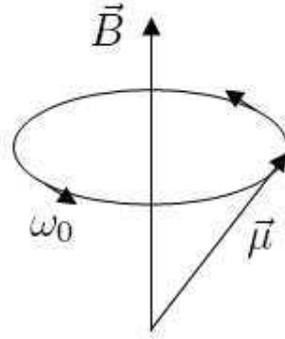
$$\omega_0 = \gamma B_0, \quad (2.1)$$

chamada de *frequência de Larmor* [21], onde o escalar  $\gamma$  é chamado de *fator giromagnético* e varia com o núcleo atômico, havendo também um fator giromagnético para o elétron [1].

O movimento de precessão do momento magnético  $\vec{\mu}$  em torno do campo magnético estático  $\vec{B}$  é governado pela seguinte equação [1]:

$$\frac{d\vec{\mu}}{dt} = \gamma\vec{\mu} \times \vec{B}. \quad (2.2)$$

Um sistema como um núcleo atômico pode consistir em diversas partículas acopladas de



**Figura 2.1** Movimento de precessão do momento magnético  $\vec{\mu}$  em torno do campo magnético  $\vec{B} = B_0\hat{z}$ , com velocidade angular  $\omega_0$

maneira que tenha um *momento magnético total*  $\vec{\mu}$  e um *momento angular total*  $\vec{J}$ , relacionados como segue:

$$\vec{\mu} = \gamma\vec{J}. \quad (2.3)$$

Sabendo que  $\vec{J} = J_x\hat{x} + J_y\hat{y} + J_z\hat{z}$ , a componente  $z$  do momento angular é dada por:

$$J_z = m_j\hbar, \quad (2.4)$$

onde  $\hbar$  é a constante de Planck dividida pelo fator  $2\pi$ .  $m_j$  pode tomar os seguintes  $2j + 1$  valores:

$$m_j = -j, -j + 1, \dots, j - 1, j, \quad (2.5)$$

e  $j$  é um inteiro ou um semi-inteiro positivo relacionado com o módulo  $J$  do momento angular total  $\vec{J}$  como segue:

$$J^2 = j(j + 1)\hbar^2, \quad (2.6)$$

onde

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots \quad (2.7)$$

O momento angular total para sistemas atômicos e nucleares é dado por:

$$\vec{J} = \vec{L} + \vec{S}, \quad (2.8)$$

onde  $\vec{L}$  é o *momento angular orbital* e  $\vec{S}$  é o *momento angular de spin*. Essas grandezas físicas obedecem às seguintes relações, semelhantes às relativas ao momento angular total  $\vec{J}$ , para o momento angular orbital  $\vec{L} = L_x\hat{x} + L_y\hat{y} + L_z\hat{z}$  de módulo  $L$ :

$$L_z = m_l\hbar, \quad (2.9)$$

$$m_l = -l, -l + 1, \dots, l - 1, l, \quad (2.10)$$

$$L^2 = l(l + 1)\hbar^2, \quad (2.11)$$

e para o momento angular de *spin*  $\vec{S} = S_x\hat{x} + S_y\hat{y} + S_z\hat{z}$  de módulo  $S$ :

$$S_z = m_s\hbar, \quad (2.12)$$

$$m_s = -s, -s + 1, \dots, s - 1, s, \quad (2.13)$$

$$S^2 = s(s + 1)\hbar^2, \quad (2.14)$$

onde os números quânticos  $l$  e  $s$  podem assumir os seguintes valores:

$$l = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots, \quad (2.15)$$

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots \quad (2.16)$$

Para um núcleo atômico, o momento angular orbital  $\vec{L}$  é nulo. Assim,  $\vec{J} = \vec{S}$ , o que resulta em:

$$J_z = m_s\hbar. \quad (2.17)$$

O valor que o número quântico de *spin*  $s$  pode assumir para um determinado núcleo atômico é governado pelas seguintes regras [22]:

- Núcleos com número de massa ímpar possuem *spin* semi-inteiro.
- Núcleos com número de massa par e número de prótons par tem *spin* nulo.
- Núcleos com número de massa par mas número de prótons ímpar possuem *spin* inteiro.

Para os núcleos  $^1\text{H}$ ,  $^{13}\text{C}$  e  $^{31}\text{P}$ ,  $s = \frac{1}{2}$  e tais sistemas de *spin* recebem o nome de sistemas de *spin*- $\frac{1}{2}$ .

Em um sistema magnético com  $\vec{B} = B_0\hat{z}$ , a energia potencial magnética  $U$  é dada por:

$$U = -\vec{\mu} \cdot \vec{B} = -\mu_z B_0 = -\gamma m_s \hbar B_0. \quad (2.18)$$

O que equivale a dizer que o sistema pode assumir  $2s + 1$  níveis de energia possíveis  $E(m_s)$ .

O núcleo de hidrogênio, chamado usualmente apenas de próton, é o núcleo mais utilizado em aplicações envolvendo imagens de RM, sendo o padrão em aplicações clínicas.

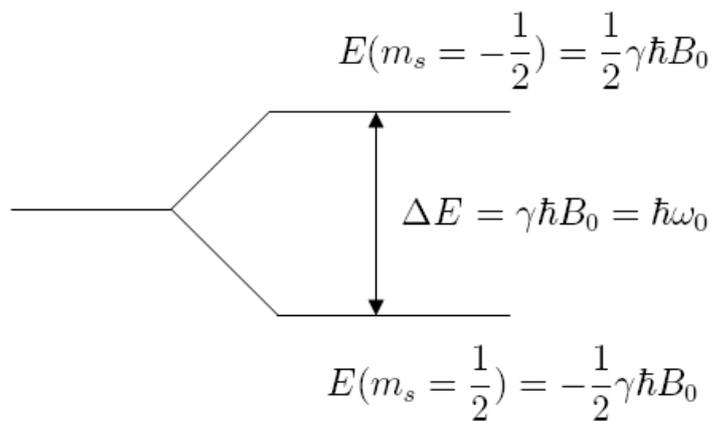
O número quântico de *spin*  $s$  do próton é dado por  $s = \frac{1}{2}$ , resultando  $m_s = \pm\frac{1}{2}$ . Assim, em um sistema de imagens de ressonância de prótons, tem-se 2 níveis possíveis de serem assumidos pelo sistema (ver figura 2.2):

$$E(m_s = \frac{1}{2}) = -\frac{1}{2}\gamma\hbar B_0 \quad (2.19)$$

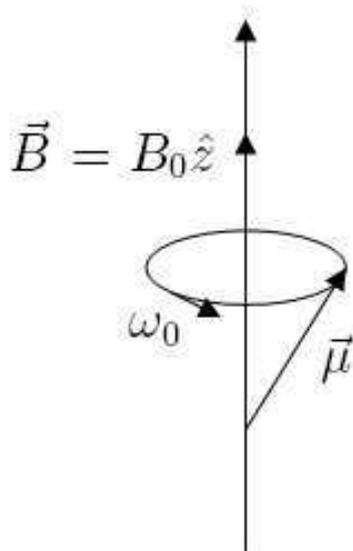
e

$$E(m_s = -\frac{1}{2}) = \frac{1}{2}\gamma\hbar B_0, \quad (2.20)$$

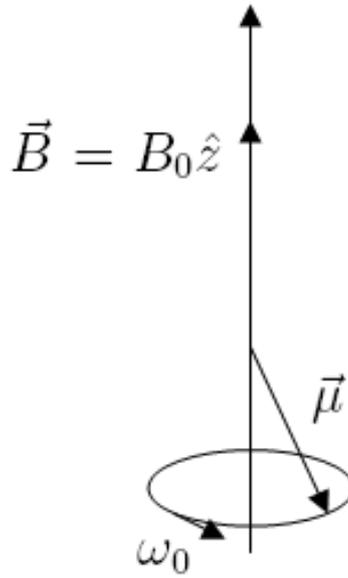
onde  $E(m_s = \frac{1}{2})$  corresponde à precessão de  $\vec{\mu}$  no sentido de  $\vec{B}$  (nível de energia mais baixo, *spin* paralelo), enquanto  $E(m_s = -\frac{1}{2})$  está associado à precessão no sentido oposto ao de  $\vec{B}$  (nível de energia mais alto, *spin* antiparalelo), como se pode ver nas figuras 2.3 e 2.4.



**Figura 2.2** Níveis possíveis de energia para um sistema de *spin* 1/2



**Figura 2.3** *Spin* paralelo a  $\vec{B}$



**Figura 2.4** Spin antiparalelo a  $\vec{B}$

A magnitude da energia de uma transição de níveis no sistema é dada por:

$$\Delta E = E(m_s = -\frac{1}{2}) - E(m_s = \frac{1}{2}) = \gamma \hbar B_0 = \hbar \omega_0. \quad (2.21)$$

Assim, a frequência de absorção ou de emissão de um *quantum* de energia do sistema é a *frequência de Larmor*  $\omega_0$ . Logo, para mudar o movimento de precessão de  $\vec{\mu}$  de paralelo para antiparalelo a  $\vec{B}$ , é preciso fornecer ao sistema um *quantum* de energia  $\Delta E$  de frequência angular  $\omega_0$ . Esses *quanta* de energia  $\Delta E$  são fornecidos ao sistema a partir da aplicação de um campo magnético de rádio-frequência  $\vec{B}_1$ , transversal ao campo magnético estático  $\vec{B}$  e oscilando na frequência de Larmor.

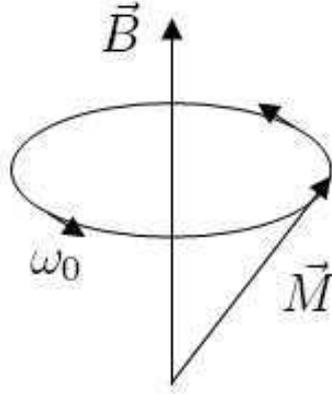
Em um sistema macroscópico, as interações do *spin* dos prótons com os átomos vizinhos levam a importantes modificações no seu comportamento. Os campos locais mudam a frequência de precessão do *spin* e o próton pode trocar energia com a vizinhança [1].

Assim, é preciso introduzir o conceito de momento magnético local por unidade de volume, ou *vetor magnetização* no volume  $V$  de posição  $\vec{r}$  variante no tempo  $t$ ,  $\vec{M}(\vec{r}, t)$ , com  $N_s$  *spins* no volume  $V$ , dado por:

$$\vec{M} = \frac{1}{V} \sum_{i=1}^{N_s} \vec{\mu}_i. \quad (2.22)$$

Sabendo que a quantidade de *spins* paralelos a  $\vec{B} = B_0 \hat{z}$ , dado por  $N_+$ , excede a população de *spins* antiparalelos  $N_-$  da forma que segue:

$$N_+ - N_- \approx N_s \frac{\gamma \hbar B_0}{2k_B T_s}, \quad (2.23)$$



**Figura 2.5** Movimento de precessão da magnetização resultante  $\vec{M}$  em torno de  $\vec{B} = B_0\hat{z}$

a magnetização de equilíbrio  $M_0$  resultante do excesso de *spins* paralelos e no sentido de  $\vec{B} = B_0\hat{z}$  é dada por [21]:

$$M_0 = \frac{\rho_0 \gamma^2 \hbar^2}{4k_B T_s} B_0, \quad (2.24)$$

onde  $\rho_0 = N_s/V$  é a *densidade de spins*, ou número de prótons por unidade de volume;  $k_B$  é a constante de Boltzmann e  $T_s$  é a temperatura.

### 2.2.1 Dinâmica da Magnetização

Considerando que não há interação entre os prótons, a expressão que governa a dinâmica da magnetização do sistema é extraída da seguinte expressão, por sua vez obtida da equação 2.2 [22]:

$$\frac{1}{V} \sum_{i=1}^{N_s} \frac{d\vec{\mu}_i}{dt} = \frac{1}{V} \gamma \sum_{i=1}^{N_s} \vec{\mu}_i \times \vec{B}, \quad (2.25)$$

resultando

$$\frac{d\vec{M}}{dt} = \gamma \vec{M} \times \vec{B}. \quad (2.26)$$

A figura 2.5 mostra o movimento de precessão da magnetização  $\vec{M}$  em torno do campo magnético  $\vec{B}$ .

Resolvendo a equação diferencial para  $\vec{B} = B_0\hat{z}$ , sabendo que  $\vec{M}(t) = M_x(t)\hat{x} + M_y(t)\hat{y} + M_z(t)\hat{z}$ , têm-se as seguintes soluções [1]:

$$\begin{cases} M_x(t) = M_x(0) \cos \omega_0 t + M_y(0) \sin \omega_0 t \\ M_y(t) = M_y(0) \cos \omega_0 t - M_x(0) \sin \omega_0 t \\ M_z(t) = M_z(0) \end{cases} . \quad (2.27)$$

Para facilitar a descrição de experimentos de ressonância magnética, utiliza-se o *sistema de coordenadas girantes*, ou simplesmente *referencial girante*. Assim, a magnetização  $\vec{M}(t)$  será

expressa no referencial girante como

$$\vec{M}(t) = M_{x'}(t)\hat{x}' + M_{y'}(t)\hat{y}' + M_{z'}(t)\hat{z}', \quad (2.28)$$

onde

$$\begin{cases} \hat{x}' = \hat{x} \cos \omega_0 t - \hat{y} \sin \omega_0 t \\ \hat{y}' = \hat{x} \sin \omega_0 t + \hat{y} \cos \omega_0 t \\ \hat{z}' = \hat{z} \end{cases}, \quad (2.29)$$

que resulta, para o referencial girante:

$$\vec{M}(t) = M_x(0)\hat{x}' + M_y(0)\hat{y}' + M_z(0)\hat{z}'. \quad (2.30)$$

Para simplificar as expressões e facilitar o entendimento da dinâmica de  $\vec{M}$ , pode-se representar a magnetização  $\vec{M}$  apenas por sua componente estática *longitudinal*,  $\vec{M}_{||} = M_z\hat{z}$ , e por sua componente dinâmica *transversal*,  $\vec{M}_{\perp}$ , de modo que  $\vec{M} = \vec{M}_{||} + \vec{M}_{\perp}$ , onde

$$\vec{M}_{\perp} = M_x\hat{x} + M_y\hat{y} = M_{x'}\hat{x}' + M_{y'}\hat{y}'. \quad (2.31)$$

### 2.2.2 Pulso de Rádio-Frequência

A aplicação do pulso de rádio-frequência (RF)  $\vec{B}_1 = B_1\hat{x}'$ , onde  $B_1 \ll B_0$ , ao sistema de *spins* durante um intervalo de tempo  $\tau_p$ , gera uma deflexão  $\alpha$  da magnetização  $\vec{M}$  com o eixo  $z$ , no referencial girante, igual a:

$$\alpha = \int_0^{\tau_p} \omega_1(t) dt = \int_0^{\tau_p} \gamma B_1(t) dt, \quad (2.32)$$

onde  $\omega_1(t) = \gamma B_1(t)$  é a velocidade angular de deflexão de  $\vec{M}$  em relação ao eixo  $z$ . No caso de um pulso de RF retangular, ou seja,  $B_1(t) = B_1$  para  $0 \leq t \leq \tau_p$ , tem-se:

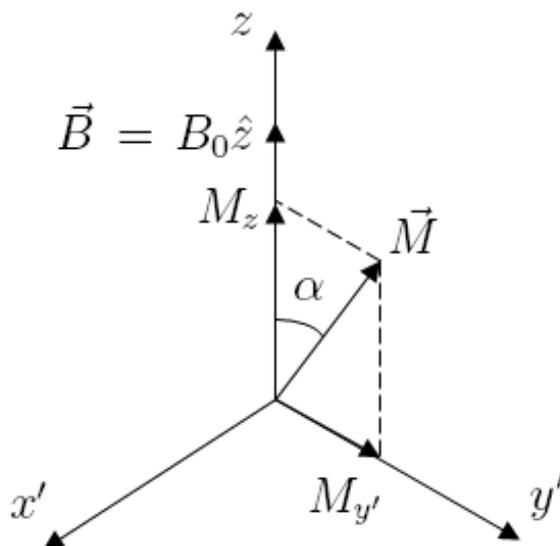
$$\alpha = \gamma B_1 \tau_p. \quad (2.33)$$

Durante a aplicação do pulso de RF, a magnetização  $\vec{M}$  obedece à seguinte dinâmica, ignorando os efeitos das interações entre cada *spin* e seu vizinho e entre cada *spin* e toda a rede de *spins* (ver figura 2.6):

$$\begin{cases} M_{x'}(t) = 0 \\ M_{y'}(t) = M_0 \sin \omega_1 t \\ M_{z'}(t) = M_0 \cos \omega_1 t \end{cases}, \quad (2.34)$$

para  $0 \leq t \leq \tau_p$ .

Os pulsos de RF são utilizados na composição das *sequências de pulsos*, responsáveis pela varredura da amostra, buscando coletar os sinais de cada elemento de volume da amostra, para posterior montagem da imagem correspondente.



**Figura 2.6** Magnetização  $\vec{M}$  no referencial girante, no instante imediatamente após o desligamento do pulso de RF de  $\alpha$  radianos

### 2.2.3 Equações de Bloch

Imediatamente após a aplicação do pulso  $\alpha$ , ou seja, do pulso de RF  $\vec{B}_1 = B_1(t)\hat{x}'$  durante um intervalo de tempo  $\tau_p$  de modo que a magnetização  $\vec{M}$  sofra uma deflexão de  $\alpha$  radianos em relação ao eixo  $z$ , a magnetização  $\vec{M}$  passa a sofrer a influência das interações entre cada *spin* e seu vizinho e entre cada *spin* e toda a rede de *spins* da amostra. À primeira interação dá-se o nome de *interação spin-spin*, e à segunda, *interação spin-rede*. Essas interações são caracterizadas por duas constantes de tempo,  $T_2$  e  $T_1$ , respectivamente, chamadas de *tempos de relaxação*.

Assim, considerando-se a influência das interações entre os *spins* na amostra, após a aplicação do pulso de RF  $\vec{B}_1 = B_1\hat{x}'$ , tem-se a seguinte dinâmica para a magnetização  $\vec{M}$  [1, 22]:

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B} + \frac{1}{T_1}(M_0 - M_z)\hat{z} - \frac{1}{T_2}(M_x\hat{x} + M_y\hat{y}). \quad (2.35)$$

A esta equação dá-se o nome de *equação de Bloch*. É uma equação de natureza empírica [1].

Durante a aplicação do pulso de RF, a magnetização sofre o efeito tanto do campo magnético estático quanto do campo de RF, de forma que  $\vec{B} = B_0\hat{z} + B_1\hat{x}'$ . Entretanto, como o tempo de duração do pulso,  $\tau_p$ , é quase sempre desprezível em relação aos demais tempos envolvidos, só interessa resolver a equação 2.35 para tempos  $t$  após o desligamento do pulso de RF, ou seja,

para  $\vec{B} = B_0\hat{z}$ . Assim, tem-se a seguinte solução:

$$\begin{cases} M_x(t) = e^{-t/T_2}(M_x(0) \cos \omega_0 t + M_y(0) \sin \omega_0 t) \\ M_y(t) = e^{-t/T_2}(M_y(0) \cos \omega_0 t - M_x(0) \sin \omega_0 t) \\ M_z(t) = M_z(0)e^{-t/T_1} + M_0(1 - e^{-t/T_1}) \end{cases}, \quad (2.36)$$

onde

$$\begin{cases} M_x(0) = M_0 \sin \alpha \cos \phi_0 \\ M_y(0) = M_0 \sin \alpha \sin \phi_0 \\ M_z(0) = M_0 \cos \alpha \end{cases}, \quad (2.37)$$

$\phi_0$  é a fase inicial de precessão após o desligamento do pulso de RF.

Considerando a expressão da magnetização em suas componentes longitudinal e transversal, a equação de Bloch assume a seguinte forma [1]:

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B} + \frac{1}{T_1}(M_0 - M_z)\hat{z} - \frac{1}{T_2}\vec{M}_\perp. \quad (2.38)$$

O módulo  $M_\perp$  da componente transversal da magnetização  $\vec{M}_\perp$  passa então a obedecer à seguinte expressão:

$$M_\perp(t) = e^{-t/T_2}M_\perp(0), \quad (2.39)$$

onde

$$M_\perp = \sqrt{M_x^2 + M_y^2}. \quad (2.40)$$

As figuras 2.7 e 2.8 mostram as curvas de relaxação transversal e longitudinal, respectivamente. Ao sinal recebido pela bobina de recepção, induzido pela magnetização transversal  $M_\perp(t)$ , após o desligamento do pulso de RF, dá-se o nome de *decaimento livre de indução* (FID, *Free Induction Decay*) [23]. A figura 2.9 mostra o FID resultante da aplicação de um pulso de RF de  $\pi/2$ .

Na prática, o módulo da componente transversal da magnetização,  $M_\perp$ , obedece à seguinte expressão [22, 1]:

$$M_\perp(t) = e^{-t/T_2^*}M_\perp(0) = e^{-t/T_2'}e^{-t/T_2}M_\perp(0), \quad (2.41)$$

onde  $T_2^*$  é o *tempo de relaxação transversal efetivo*, função do tempo de relaxação *spin-spin*,  $T_2$ , modulado pela relaxação transversal devida às inomogeneidades locais do campo magnético,  $T_2'$ , da forma que segue:

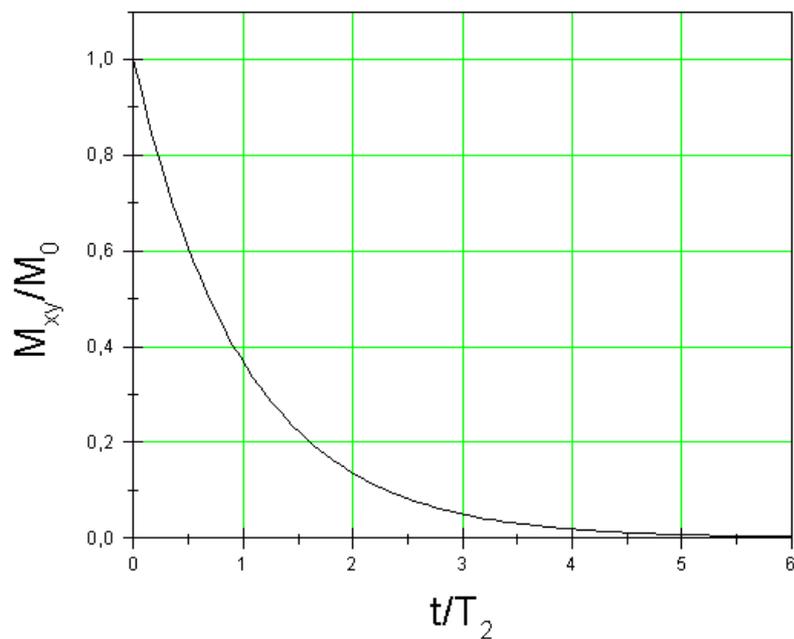
$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'}, \quad (2.42)$$

onde  $T_2' > T_2$ . Na ausência total de inomogeneidades locais no campo magnético,  $T_2' \gg T_2$ .

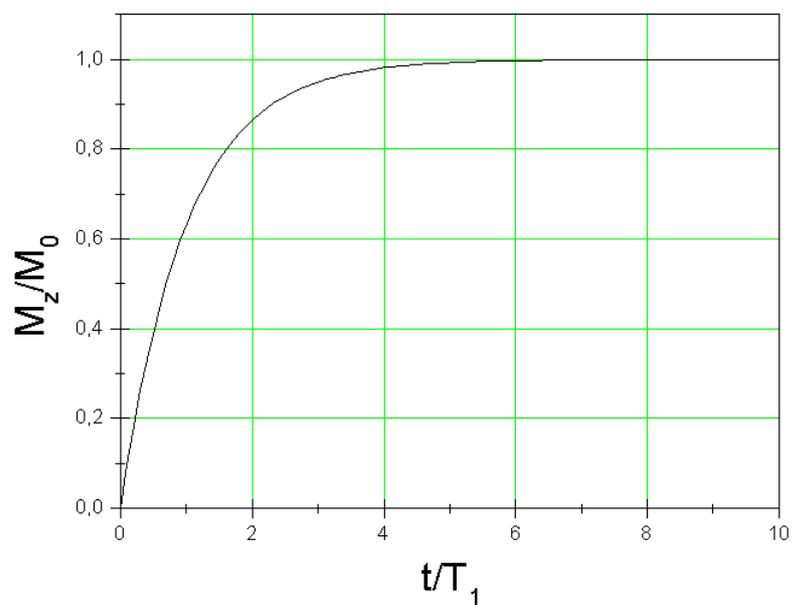
#### 2.2.4 Detecção de Sinal

Nos sistemas de RM, após o desligamento do pulso de RF, o sinal de ressonância começa a ser detectado. O fluxo magnético  $\Phi(t)$  na bobina de recepção do sistema, transversal ao eixo  $z$  do campo magnético estático, é dado por [22]:

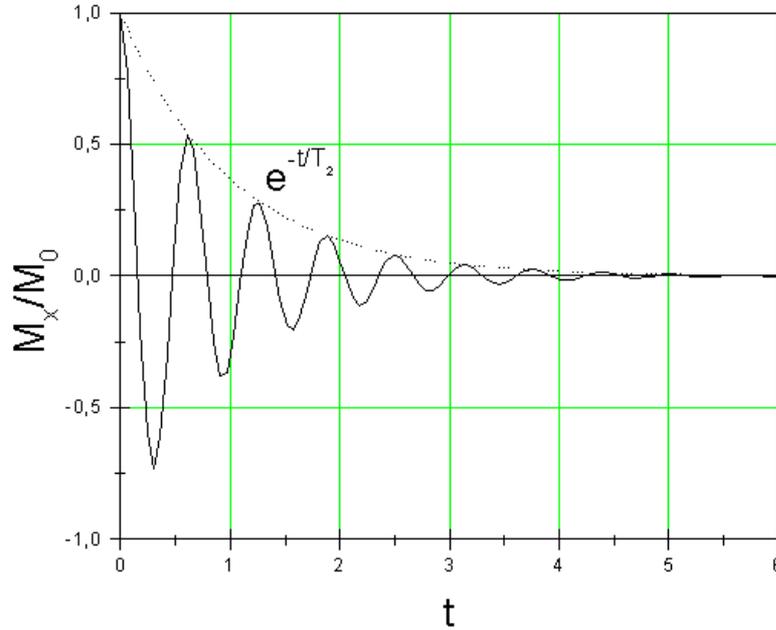
$$\Phi(t) = \int_{V_s} \vec{B}_r(\vec{r}) \cdot \vec{M}(\vec{r}, t) d\vec{r}, \quad (2.43)$$



**Figura 2.7** Magnetização transversal  $M_{xy}(t) = M_{\perp}(t)$  para  $M_{\perp}(0) = M_0$



**Figura 2.8** Magnetização longitudinal  $M_z(t)$  para  $M_z(0) = 0$



**Figura 2.9** Exemplo de FID resultante de um pulso de RF de  $\pi/2$  radianos

onde  $\vec{B}_r(\vec{r})$  é o campo magnético recebido pela bobina de recepção após o desligamento do pulso de RF;  $V_s$  é o volume total da amostra;  $\vec{M}(\vec{r}, t)$  é a magnetização no elemento de volume (*voxel*)  $\Delta V_s$  de posição  $\vec{r} = (x, y, z)$  da amostra.

Pela lei de indução de Faraday, a tensão induzida na bobina,  $V(t)$ , é dada por:

$$V(t) = -\frac{\partial \Phi}{\partial t} = -\frac{\partial}{\partial t} \int_{V_s} \vec{B}_r(\vec{r}) \cdot \vec{M}(\vec{r}, t) d\vec{r}. \quad (2.44)$$

Como o campo magnético recebido pela bobina,  $\vec{B}_r = B_{r,x}\hat{x} + B_{r,y}\hat{y} + B_{r,z}\hat{z}$ , não varia com o tempo, tem-se:

$$V(t) = -\int_{V_s} \vec{B}_r(\vec{r}) \cdot \frac{\partial \vec{M}(\vec{r}, t)}{\partial t} d\vec{r}. \quad (2.45)$$

Considerando a variação da componente longitudinal de  $\vec{M}$  muito mais lenta do que a da componente transversal, pois  $T_2 \ll T_1$  [22, 1, 21], tem-se:

$$V(t) = -\int_{V_s} \left[ B_{r,x}(\vec{r}) \frac{\partial M_x(\vec{r}, t)}{\partial t} + B_{r,y}(\vec{r}) \frac{\partial M_y(\vec{r}, t)}{\partial t} \right] d\vec{r}, \quad (2.46)$$

pois  $\partial M_z / \partial t \approx 0$  para  $t \ll T_1$ .

Logo, a tensão induzida  $V(t)$  varia apenas com a magnetização transversal, já que apenas  $M_x$  e  $M_y$  são variantes no tempo.

Para facilitar a solução e simplificar a notação, pode-se resolver a equação para um único *voxel* de posição  $\vec{r}$ . Assim, considerando as soluções das equações de Bloch da equação 2.36 e que a frequência de ressonância no *voxel* é  $\omega = \omega_0 + \Delta\omega$ , devido a variações no campo magnético estático (inogeneidades), ou seja,  $\vec{B} = (B_0 + \Delta B)\hat{z}$ , chega-se às seguintes expressões:

$$\begin{cases} M_x(t) = e^{-t/T_2} M_{\perp}(0) \cos(\omega t + \phi_0) \\ M_y(t) = e^{-t/T_2} M_{\perp}(0) \sin(\omega t + \phi_0) \end{cases},$$

onde  $\phi_0$  é a fase inicial após o pulso de RF, ou seja,  $\tan \phi_0 = M_y(0)/M_x(0)$ . Como  $\omega \gg 1/T_2$  e  $\omega_0 \gg \Delta\omega$ , obtém-se:

$$\begin{aligned} \frac{\partial M_x}{\partial t} &= -\omega e^{-t/T_2} M_{\perp}(0) \sin(\omega t + \phi_0), \\ \frac{\partial M_y}{\partial t} &= \omega e^{-t/T_2} M_{\perp}(0) \cos(\omega t + \phi_0). \end{aligned}$$

Fazendo  $\vec{B}_r = \vec{B}_{r,\perp} + B_{r,z}\hat{z}$ , onde

$$\begin{aligned} \vec{B}_r &= B_{r,x}\hat{x} + B_{r,y}\hat{y}, \\ B_{r,x} &= B_{r,\perp} \cos \phi_B, \\ B_{r,y} &= B_{r,\perp} \sin \phi_B, \end{aligned}$$

e  $\phi_B$  é a fase de  $\vec{B}_r$  tal que  $\phi_B = 0$  para  $\vec{B}_r$  alinhado com a bobina de recepção no eixo  $x$ , onde  $\tan \phi_B = B_{r,y}/B_{r,x}$ , calculando a expressão:

$$\Delta V = -B_{r,x} \frac{\partial M_x}{\partial t} - B_{r,y} \frac{\partial M_y}{\partial t},$$

tem-se:

$$\Delta V(t) = \omega_0 e^{-t/T_2} B_{r,\perp} M_{\perp}(0) [\cos \phi_B \sin(\omega t + \phi_0) - \sin \phi_B \cos(\omega t + \phi_0)].$$

Simplificando:

$$\Delta V(t) = \omega_0 e^{-t/T_2} B_{r,\perp} M_{\perp}(0) \Re\{e^{-j(\omega t + \phi_0 - \phi_B)}\}.$$

Assim, para toda a amostra:

$$V(t) = \omega_0 \int_{V_s} e^{-t/T_2(\vec{r})} B_{r,\perp}(\vec{r}) M_{\perp}(\vec{r}, 0) \Re\{e^{-j(\omega(\vec{r})t + \phi_0(\vec{r}) - \phi_B(\vec{r}))}\} d\vec{r}, \quad (2.47)$$

onde

$$\omega(\vec{r}) = \omega_0 + \Delta\omega(\vec{r}). \quad (2.48)$$

Outra forma possível para a tensão induzida  $V(t)$  é:

$$V(t) = \omega_0 \int_{V_s} e^{-t/T_2(\vec{r})} B_{r,\perp}(\vec{r}) M_{\perp}(\vec{r}, 0) \Im\{e^{-j(\omega(\vec{r})t + \phi_0(\vec{r}) - \phi_B(\vec{r}))}\} d\vec{r}. \quad (2.49)$$

Após passar por uma etapa de demodulação em amplitude e considerando os dois canais possíveis, obtém-se a seguinte expressão para o sinal  $S(t)$  no domínio dos números complexos [1]:

$$S(t) = \omega_0 \int_{V_s} e^{-t/T_2(\vec{r})} B_{r,\perp}(\vec{r}) M_{\perp}(\vec{r}, 0) e^{-j(\Delta\omega(\vec{r})t + \phi_0(\vec{r}) - \phi_B(\vec{r}))} d\vec{r}. \quad (2.50)$$

Caso a bobina de transmissão seja suficientemente uniforme, a fase  $\phi_0$  passa a não depender da posição  $\vec{r}$ . Caso o mesmo se dê com a bobina de recepção, o que é razoável, uma vez que o mais comum é que a mesma bobina ora seja utilizada para transmissão, ora para recepção, o campo magnético recebido  $\vec{B}_r$  e sua fase  $\phi_B$  também passam a ser independentes da posição  $\vec{r}$ . Assim, o sinal passa a ser representado pela seguinte expressão, anulando as fases  $\phi_0$  e  $\phi_B$  [23, 22, 1]:

$$S(t) = \omega_0 B_{r,\perp} \int_{V_s} e^{-t/T_2(\vec{r})} M_{\perp}(\vec{r}, 0) e^{-j\Delta\omega(\vec{r})t} d\vec{r}. \quad (2.51)$$

Para tempos muito curtos (aquisições rápidas),  $t \ll T_2$ , o que resulta na expressão aproximada:

$$S(t) = \omega_0 B_{r,\perp} \int_{V_s} M_{\perp}(\vec{r}, 0) e^{-j\Delta\omega(\vec{r})t} d\vec{r}. \quad (2.52)$$

Mas como a magnetização de equilíbrio para o *voxel* de posição  $\vec{r}$  é dada por, a partir da equação 2.24:

$$M_{\perp}(\vec{r}, 0) = M_0(\vec{r}) = \frac{\gamma^2 \hbar^2 B_0}{4k_B T_s} \rho_0(\vec{r}), \quad (2.53)$$

onde  $\rho_0(\vec{r})$  é a densidade de *spins* na posição  $\vec{r}$ , o *sinal complexo demodulado*  $S(t)$  pode ser expresso por:

$$S(t) = \int_{V_s} \rho(\vec{r}) e^{-j\Delta\omega(\vec{r})t} d\vec{r}, \quad (2.54)$$

onde  $\rho(\vec{r})$  é a *densidade de spins efetiva*, definida da forma que segue:

$$\rho(\vec{r}) = \omega_0 B_{r,\perp} M_0(\vec{r}) = \frac{\gamma^3 \hbar^2 B_0^2 B_{r,\perp}}{4k_B T_s} \rho_0(\vec{r}). \quad (2.55)$$

### 2.3 Imagens de Ressonância Magnética

Em um sistema de imagens de RM, inhomogeneidades controladas são introduzidas no campo magnético estático para permitir a seleção dos sinais de determinados voxels e, dessa forma, tornar possível a construção de uma imagem a partir da decomposição do sinal adquirido em sinais de frequências e fases variadas. Isso é possível pela introdução de *gradientes lineares* [2, 1].

Os gradientes são campos magnéticos variantes no espaço de maneira que, quando adicionados ao campo magnético estático  $B_0 \hat{z}$ , o campo resultante ausente de pulso de RF,  $\vec{B} = B(\vec{r}, t) \hat{z}$ , seja tal que:

$$B(\vec{r}, t) = B_0 + \vec{G}(t) \cdot \vec{r}, \quad (2.56)$$

onde

$$\vec{G}(t) = G_x(t) \hat{x} + G_y(t) \hat{y} + G_z(t) \hat{z} \quad (2.57)$$

de forma que

$$\frac{\partial B}{\partial x} = G_x(t), \quad (2.58)$$

$$\frac{\partial B}{\partial y} = G_y(t), \quad (2.59)$$

e

$$\frac{\partial B}{\partial z} = G_z(t). \quad (2.60)$$

Logo, para um sistema com gradientes:

$$\Delta B(\vec{r}, t) = G_x(t)x + G_y(t)y + G_z(t)z, \quad (2.61)$$

o que resulta:

$$\Delta \omega(\vec{r}, t) = \Delta \omega(x, y, z, t) = \gamma [G_x(t)x + G_y(t)y + G_z(t)z]. \quad (2.62)$$

Como a densidade de *spins* efetiva é idealmente nula para *voxels* onde não há amostra, os limites da integral na expressão do sinal (equação 2.54) podem se estender a todo o espaço. A expressão do sinal  $S(t)$  se torna então [2, 1]:

$$S(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \rho(x, y, z) e^{-j\phi(x, y, z, t)} dx dy dz, \quad (2.63)$$

onde

$$\phi(x, y, z, t) = \gamma \int_0^t [G_x(t')x + G_y(t')y + G_z(t')z] dt'. \quad (2.64)$$

### 2.3.1 Equação da Imagem

Uma tática muito útil para reduzir a complexidade dos problemas envolvendo imagens de RM é expressar o sinal adquirido em função das coordenadas  $k$ . Esse processo é chamado de *transformação para o espaço  $k$* . Assim, a partir da equação 2.63 chega-se à expressão do sinal em função das coordenadas  $k$ :

$$S(k_x, k_y, k_z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \rho(x, y, z) e^{-j2\pi(k_x x + k_y y + k_z z)} dx dy dz, \quad (2.65)$$

onde

$$k_x = \frac{\gamma}{2\pi} \int_0^t G_x(t') dt' \quad (2.66)$$

é a *codificação de frequência*,

$$k_y = \frac{\gamma}{2\pi} \int_0^t G_y(t') dt' \quad (2.67)$$

é a *codificação de fase* e

$$k_z = \frac{\gamma}{2\pi} \int_0^t G_z(t') dt' \quad (2.68)$$

é a *seleção de fatias*.

Na prática é comum se utilizar a seguinte terminologia para os gradientes:  $G_x = G_R$  é o *gradiente de leitura (Reading, R)*;  $G_y = G_{PE}$  é o *gradiente de codificação de fase (Phase Encoding, PE)*; e  $G_z = G_{SS}$  é o *gradiente de seleção de fatia (Slice Selecting, SS)*.

Pode-se perceber da equação 2.65 que  $S(k_x, k_y, k_z)$  e  $\rho(x, y, z)$  são pares de Fourier de três dimensões. Logo, a *equação da imagem 3D* é dada por:

$$\rho(x, y, z) = \frac{1}{(2\pi)^3} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S(k_x, k_y, k_z) e^{j2\pi(k_x x + k_y y + k_z z)} dk_x dk_y dk_z. \quad (2.69)$$

Nos métodos de aquisição de imagens 2D, a cada variação de  $k_z$  são variados  $k_x$  e  $k_y$  de forma a produzir a imagem 2D  $\rho(x, y)$  da fatia  $z$ . Somente após montada a imagem da fatia  $z$  passa-se à imagem da fatia  $z + \Delta z$ .  $k_z$  é incrementado a cada pulso positivo em  $G_z(t)$ . Para esses métodos, utiliza-se a *equação da imagem 2D*, definida como segue:

$$\rho(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S(k_x, k_y) e^{j2\pi(k_x x + k_y y)} dk_x dk_y, \quad (2.70)$$

onde

$$S(k_x, k_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \rho(x, y) e^{-j2\pi(k_x x + k_y y)} dx dy. \quad (2.71)$$

A *resolução da imagem*, ou *campo de visão (FOV, Field of View)*, é dada por, para cada dimensão:

$$\text{FOV}_x = \frac{1}{\Delta k_{x,\text{max}}}, \quad (2.72)$$

$$\text{FOV}_y = \frac{1}{\Delta k_{y,\text{max}}}, \quad (2.73)$$

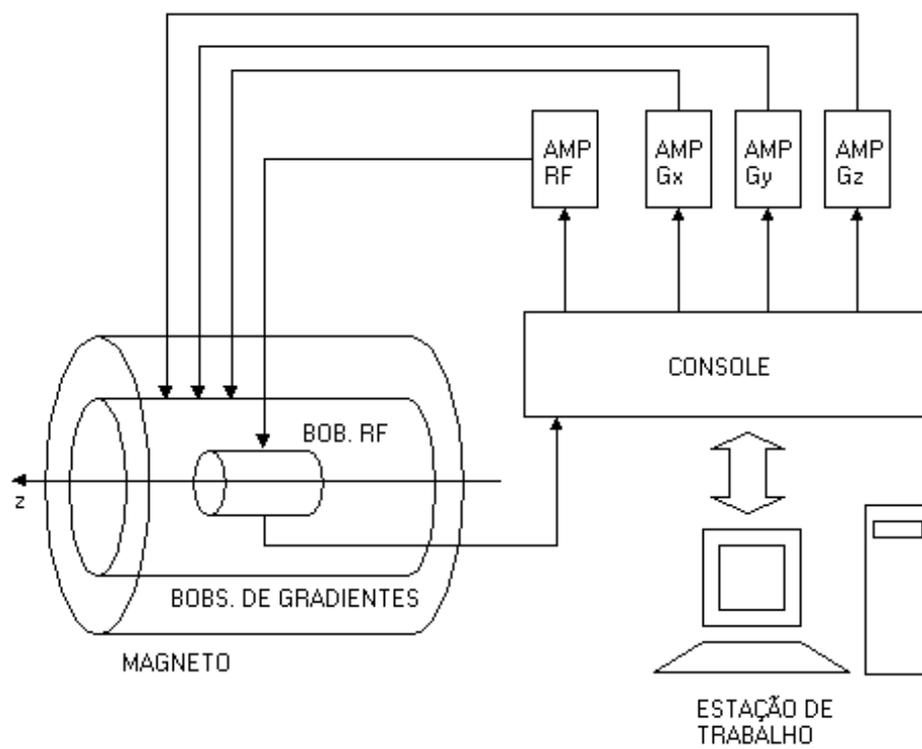
$$\text{FOV}_z = \frac{1}{\Delta k_{z,\text{max}}}, \quad (2.74)$$

onde  $\Delta k_{x,\text{max}}$ ,  $\Delta k_{y,\text{max}}$  e  $\Delta k_{z,\text{max}}$  são os passos máximos efetuados pela sequência de pulsos ao percorrer o espaço  $k$  nas dimensões  $k_x$ ,  $k_y$  e  $k_z$ , respectivamente.

A figura 2.10 mostra um diagrama simplificado de um sistema de aquisição de imagens de ressonância magnética composto pelos seguintes elementos: (a) o magneto, responsável pela geração do campo magnético estático  $\vec{B} = B_0 \hat{z}$ ; (b) as bobinas de gradiente e seus amplificadores; (c) a bobina de transmissão e recepção de RF e seu amplificador; (d) a estação de trabalho para controle do sistema e visualização das imagens; (e) o console, interface entre a estação de trabalho e os amplificadores e as bobinas, para codificar as sequências de pulsos e capturar o sinal de RF recebido pela bobina de RF, permitindo a montagem da imagem pela estação de trabalho.

### 2.3.2 Sequências de Pulsos

Em linhas gerais, as sequências de pulsos são sequências dos sinais de RF e dos gradientes  $G_x(t)$ ,  $G_y(t)$  e  $G_z(t)$  de forma a preencher o espaço  $k$  e assim montar o sinal  $S(k_x, k_y)$  (caso 2D) a partir do qual será montada a imagem  $\rho(x, y)$ . Como não é possível percorrer todo o



**Figura 2.10** Diagrama simplificado de um sistema de imagens de RM

espaço  $k$ , as sequências passam a corresponder a determinadas trajetórias no espaço  $k$ , resultando em imagens  $\hat{\rho}(x, y)$ , aproximações da imagem  $\rho(x, y)$ , cujo contraste é, por conseguinte, dependente da sequência de pulsos aplicada.

Existem diversas sequências de pulsos, havendo aplicações diversas, uma vez que essas sequências destacam propriedades físicas interessantes da amostra. O projeto de sequências de pulsos para obtenção de imagens é uma área em pleno desenvolvimento [22, 1]. Neste trabalho são apresentadas apenas duas sequências de pulsos, a sequência *SE* e a *EPI*, utilizadas para aquisição das imagens utilizadas nos estudos de caso deste trabalho.

### 2.3.2.1 Sequência SE

A sequência *eco de spin* (*Spin Echo*, SE) é a mais simples e intuitiva de todas. Nesta sequência, uma vez tendo-se selecionado a fatia definindo  $k_z$  e, para cada  $k_y$ , um  $k_x$ , de forma a percorrer uma versão segmentada do plano  $k$  [1], como mostrado na figura 2.11.

A figura 2.12 mostra a sequência de pulsos. A seta de cima para baixo sobre o gradiente  $G_y$  indica que  $k_y$  é decrementado do valor mais positivo ao mais negativo, a cada passo, cada vez que  $k_x$  é totalmente varrido. A linha ADC (*Analog-Digital Converter*) representa a ação do conversor analógico-digital após a bobina receptora ao longo do tempo: o sinal é amostrado no centro do pulso positivo indicado na linha, equivalente a um intervalo de tempo  $T_E$ , denominado de *tempo de eco*, após a aplicação do pulso de  $\pi/2$ .

Versões da sequência *eco de spin* são utilizadas para a aquisição de imagens ponderadas em difusão [1].

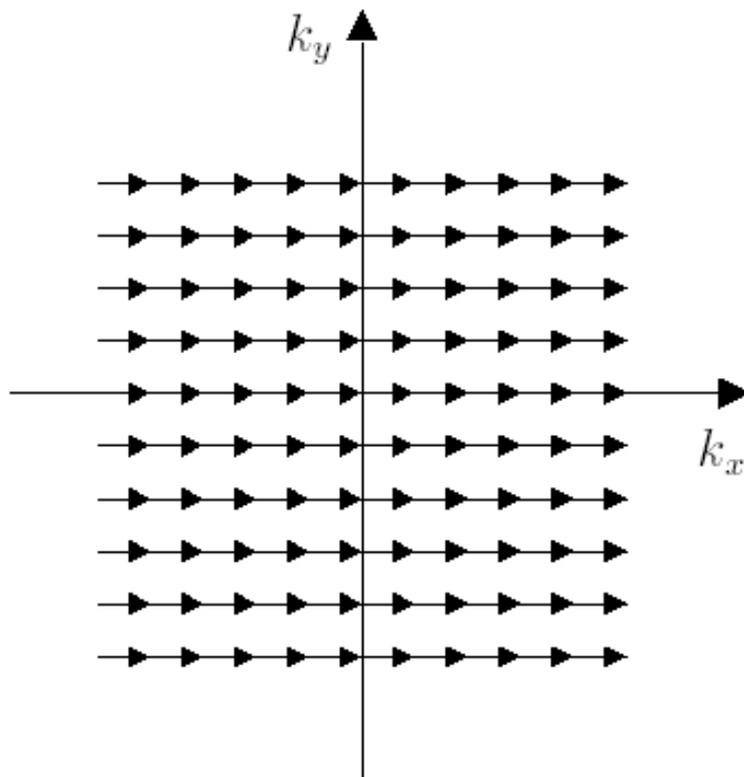
### 2.3.2.2 Sequência EPI

A *sequência de imageamento ecoplanar*, ou simplesmente *EPI* (*Echo Planar Imaging*), é utilizada para a obtenção de imagens funcionais. Ela consiste em, uma vez selecionada a fatia  $k_z$ , varrer continuamente o espaço  $k$  no eixo  $k_x$ , para cada passo de  $k_y$ , após a aplicação do pulso de RF de  $\pi/2$ , seguindo a trajetória mostrada na figura 2.13. Assim, tendo-se selecionado uma fatia, ou seja, dado  $k_z$ , para cada  $k_y$ , é lida continuamente toda uma linha, portanto todos os valores de  $k_x$  [2, 1].

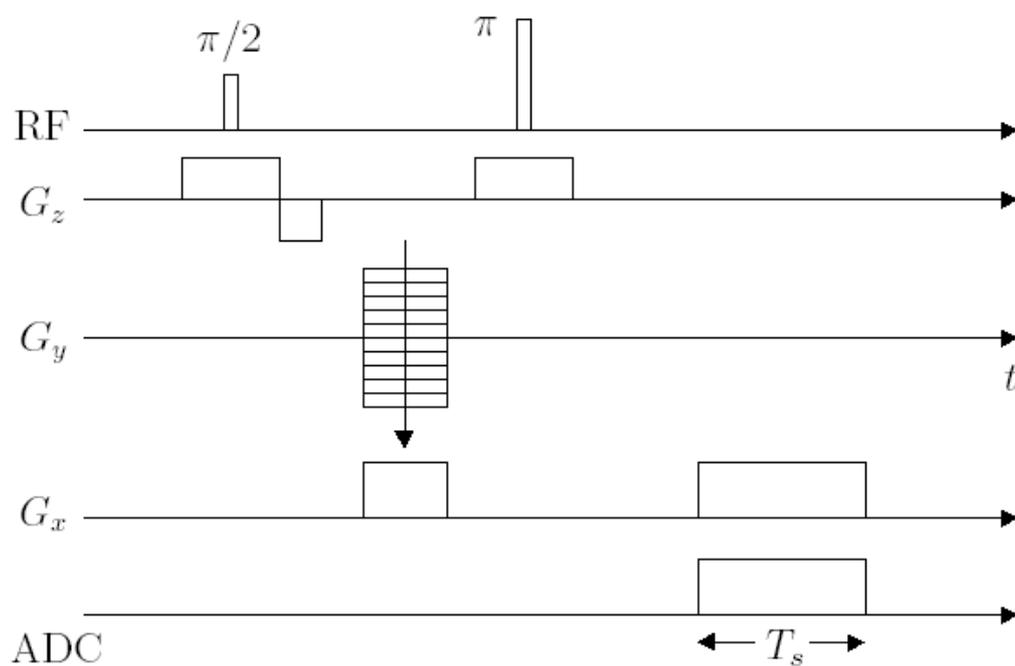
A figura 2.14 mostra a sequência de pulsos. A seta para cima no gradiente  $G_y$  indica que  $k_y$  é incrementado de seu valor mais negativo ao mais positivo. Os pulsos positivos e negativos de  $G_x$  permitem a varredura de  $k_x$ , da esquerda para a direita e da direita para a esquerda, nesta ordem, como mostrado na figura 2.13. Assim, em cada pulso do gradiente  $G_x$  varre-se uma linha inteira do plano  $k$ . Os pulsos de baixa amplitude em  $G_y$  incrementam  $k_y$ , permitindo a varredura de baixo para cima.

## 2.3.3 Aplicações

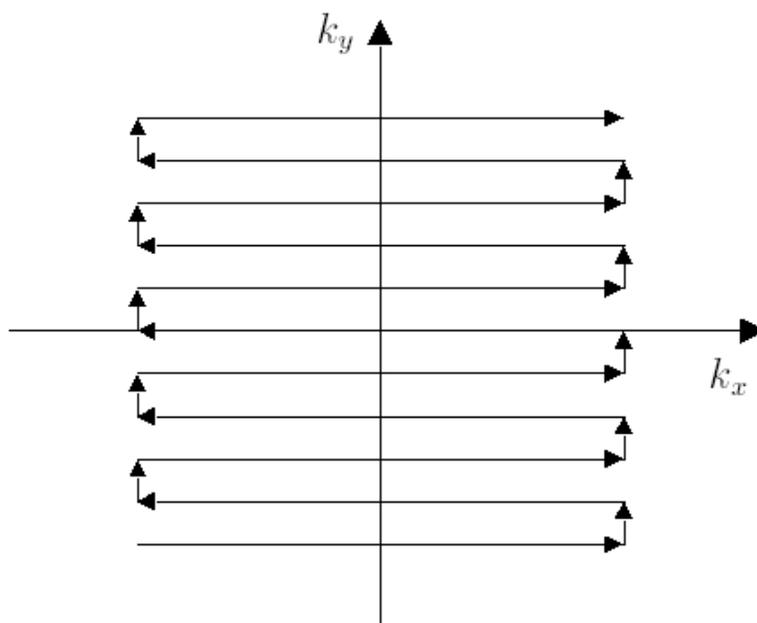
As sequências de pulsos geram diferentes imagens tomográficas, evidenciando diferentes propriedades físicas da amostra e, por conseguinte, têm diversas aplicações. Neste texto são apresentadas apenas as aplicações em *imagens de difusão* e em *imagens funcionais*, por estarem relacionadas com os estudos de caso efetuados neste trabalho.



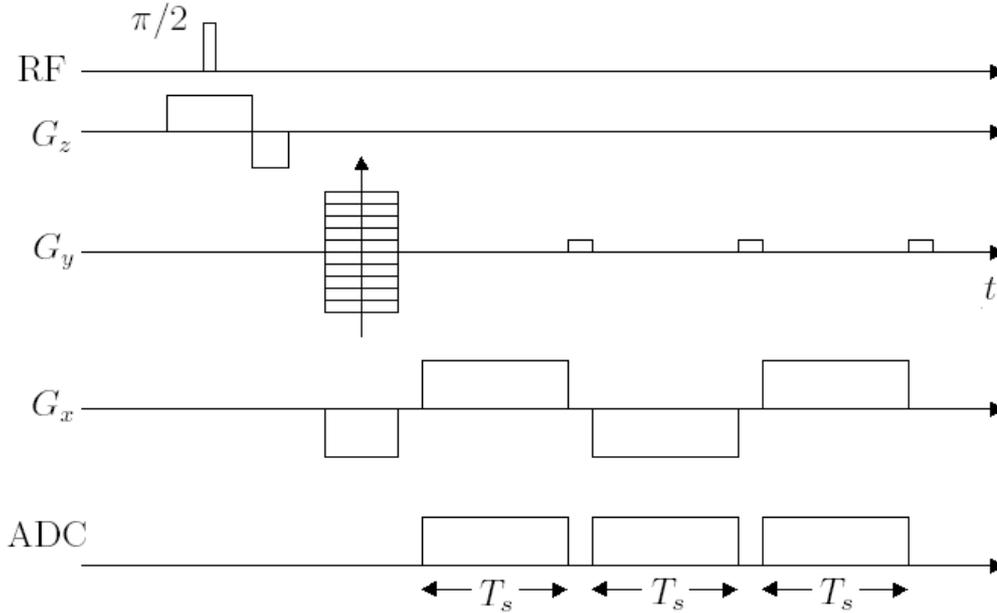
**Figura 2.11** Cobertura do plano  $k$  para um determinado  $k_z$  pela sequência de pulsos SE



**Figura 2.12** Sequência de pulsos SE



**Figura 2.13** Cobertura do plano  $k$  para um determinado  $k_z$  pela sequência de pulsos EPI



**Figura 2.14** Sequência de pulsos EPI

### 2.3.3.1 Imagens de Difusão

As imagens de ressonância magnética ponderadas em difusão, ou simplesmente imagens de difusão, são imagens de ressonância baseadas no movimento aleatório dos prótons nas amostras. Sua aplicação clínica consiste no fato de que a difusão de prótons no cérebro é diminuída em regiões ocupadas por cistos, isquemias, necroses, tumores e outras patologias [24, 25].

A aquisição de imagens ponderadas em difusão é efetuada por meio da sequência de pulsos Stejskal-Tanner, uma sequência derivada da sequência SE [1], como mostrado na figura 2.15.

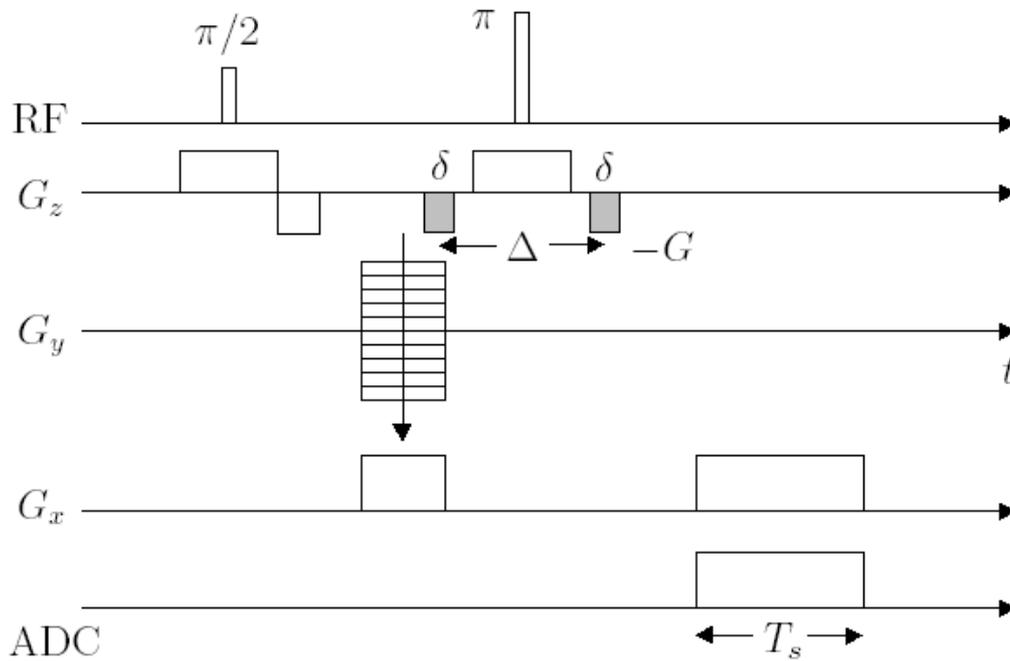
Na sequência Stejskal-Tanner são introduzidas perturbações no gradiente  $G_z(t)$  após a seleção da fatia, na forma de pulsos negativos de amplitude  $G$ , de largura  $\delta$ , no instante  $\Delta/2$  antes do pulso de RF de  $\pi$  radianos e no instante  $\Delta/2$  após o pulso de  $\pi$ , o que provoca uma defasagem dos *spins* tão grande quanto for a sua mobilidade. Assim, a magnetização transversal tem seu valor diminuído, resultando na seguinte expressão para a imagem no *voxel* de posição  $(x, y)$ :

$$\hat{\rho}(x, y) = \rho_0(x, y) e^{-T_E/T_2^*(x, y)} e^{-bD(x, y)}, \quad (2.75)$$

onde  $D(x, y)$  é o *coeficiente de difusão de spins* no *voxel* de posição  $(x, y)$  e  $b$  é o *expoente de difusão*, definido pela expressão [1]:

$$b = \frac{1}{3} \gamma^2 G^2 T_E^3. \quad (2.76)$$

Logo, nos *voxels* onde a difusão de *spins* é maior, a intensidade da imagem é menor. Portanto, em imagens ponderadas em difusão, as regiões do cérebro ocupadas por patologias que diminuem a mobilidade de *spins* serão representadas por *pixels* de maior brilho, considerando



**Figura 2.15** Sequência de pulsos SE Stejskal-Tanner

um mapeamento linear e crescente dos níveis de cinza na imagem resultante, ou seja, considerando que o maior valor de  $\hat{\rho}(x, y)$  é mapeado como branco absoluto, enquanto seu menor valor, como preto absoluto.

### 2.3.3.2 Imagens Funcionais e o Efeito BOLD

As imagens funcionais de RM são, na maioria das aplicações, obtidas a partir do efeito de *nível dependente da oxigenação do sangue* (*Blood Oxygenation Level Dependent*, BOLD). Nesta técnica, o contraste das imagens é dependente da variação da taxa de oxi-hemoglobina ou de desoxi-hemoglobina em vasos sanguíneos que acompanha a atividade neural [2].

O aumento do fluxo de sangue em regiões do cérebro que se tornam ativas parece ser uma consequência do aumento de utilização de energia nas sinapses, apesar de ainda não se conhecer que processos influenciam nessas mudanças metabólicas. Porém, não parece existir uma relação simples entre o aumento do consumo de energia devido à atividade neural e o aumento do fluxo de sangue [2].

Seja  $N_{\text{ox}}$  o número de moléculas de oxi-hemoglobina;  $N_{\text{deox}}$ , o número de moléculas de desoxi-hemoglobina;  $N_0 = N_{\text{ox}} + N_{\text{deox}}$ , o total de moléculas de oxi-hemoglobina e desoxi-hemoglobina;  $\Delta\phi_\alpha$ , a variação relativa do fluxo de sangue; e  $\Delta\phi_\beta$ , a taxa metabólica cerebral, ou seja, a taxa de utilização de oxigênio pelo cérebro, a saturação fracional de oxigênio no sangue,  $Y$ , é dada por [1]:

$$Y = 1 - \frac{N_{\text{deox}}}{N_0}. \quad (2.77)$$

Sabendo que a variação de moléculas de desoxi-hemoglobina é dada por:

$$\Delta N_{\text{deox}} = \frac{\Delta\phi_{\beta} - \Delta\phi_{\alpha}}{1 + \Delta\phi_{\alpha}} N_{\text{deox}}, \quad (2.78)$$

a variação da saturação fracional de oxigênio no sangue,  $\Delta Y$ , é definida pela expressão:

$$\Delta Y = -\frac{\Delta N_{\text{deox}}}{N_0}, \quad (2.79)$$

equivalente a:

$$\Delta Y = \frac{\Delta\phi_{\alpha} - \Delta\phi_{\beta}}{1 + \Delta\phi_{\alpha}} (1 - Y). \quad (2.80)$$

Uma vez que a utilização de oxigênio pelo cérebro não varia significativamente,  $\Delta\phi_{\beta} \approx 0$ , o que resulta [1]:

$$\Delta Y = \frac{\Delta\phi_{\alpha}}{1 + \Delta\phi_{\alpha}} (1 - Y). \quad (2.81)$$

Utilizando uma sequência de pulsos baseada na sequência eco de *spin*, como a sequência EPI, as imagens obtidas são da seguinte forma:

$$\hat{\rho}(x, y) = e^{-T_E/T_2^*(x, y)} \rho_0(x, y) = e^{-T_E R_2^*(x, y)} \rho_0(x, y), \quad (2.82)$$

para

$$R_2^* = R_{2,0}^* + \Delta\phi_{\alpha}(1 - Y) + (1 + \Delta\phi_{\beta})(1 - Y)^2, \quad (2.83)$$

onde  $T_2^* = 1/R_2^*$ ;  $T_{2,0}^* = 1/R_{2,0}^*$  é o tempo de relaxação transversal para o sangue em repouso;  $T_E$  é o *tempo de eco*, um parâmetro da sequência de pulsos [1].

Uma variação  $\Delta Y(x, y)$  na saturação de oxigênio em um determinado *voxel* de posição  $(x, y)$  resulta em uma variação relativa do sinal nesse mesmo *voxel* igual a:

$$\frac{\Delta\hat{\rho}(x, y)}{\hat{\rho}(x, y)} = [2(1 - Y(x, y)) - \Delta\phi_{\alpha}(x, y)] \Delta Y(x, y). \quad (2.84)$$

Para uma saturação de oxigênio típica  $Y = 55\%$ , caso ocorra um aumento de  $\Delta\phi_{\alpha} = 50\%$  no fluxo de sangue em uma determinada região do cérebro, a saturação de oxigênio correspondente a esse elemento de volume aumentará em  $\Delta Y = 15\%$ , gerando uma variação de 6% na intensidade do sinal [1]. Tipicamente, quando ocorre uma ativação de uma determinada região do cérebro, o sinal do *voxel* correspondente aumenta de 1% a 5% [6].

Uma vez que o incremento do nível do sinal nos *voxels* ativados é muito baixo, da ordem do ruído, são utilizados métodos estatísticos e métodos baseados em Inteligência Computacional para detectar as regiões ativadas a partir de diversas imagens de todo o volume cerebral, adquiridas com o paciente sob estímulo e em repouso.

## 2.4 Conclusão

Este capítulo apresentou a teoria básica da obtenção de imagens por ressonância magnética, descrevendo de forma relativamente sucinta os princípios físicos da ressonância magnética nuclear, bem como a aquisição de imagens 2D e 3D, os tipos de imagens e de sequências de pulsos, dando um enfoque especial nas sequências SE e EPI, utilizadas nos estudos de caso deste trabalho.

As imagens de ressonância magnética vêm tomando um espaço cada vez maior como técnica não invasiva de diagnóstico por imagem, tendo também um importante papel no desenvolvimento das ciências cognitivas, em particular no mapeamento funcional do cérebro através do emprego das imagens de ressonância funcional.

Contudo, pode-se afirmar que a extração de informação de imagens de ressonância ainda não é otimizada por motivos diversos, dentre eles a ausência de ferramentas de auxílio ao diagnóstico acessíveis aos especialistas e a relativa deficiência da formação desses mesmos especialistas no que se refere ao domínio das técnicas de imagens [26].

O emprego de técnicas da Inteligência Computacional e o desenvolvimento de sistemas computacionais de análise de imagens de apoio à decisão e de baixo custo pode incrementar a extração de informação na análise das imagens, conseqüentemente aumentando a qualidade dos exames e dos diagnósticos efetuados, por exemplo. Este é um dos objetivos perseguidos nesta pesquisa, cujos resultados empregando análise multiespectral de imagens utilizando as ferramentas de Inteligência Computacional propostas neste trabalho são abordados nos capítulos que seguem.



## A Concepção Dialética da Realidade

### 3.1 Introdução

A concepção dialética da realidade é um tipo de método filosófico investigativo para analisar processos presentes na natureza e em sociedades humanas. Suas origens estão conectadas às filosofias das antigas civilizações da Grécia, da China e da Índia e estão fortemente relacionadas ao pensamento de Heráclito, Platão, e às filosofias de religiões como o Confucionismo, o Budismo e a Escola Zen [10].

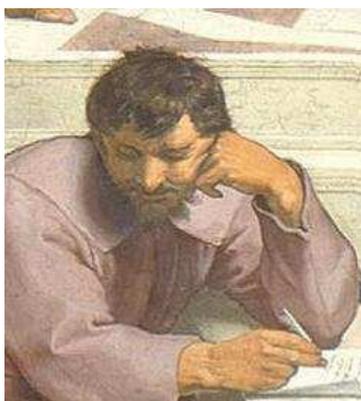
Como um método geral de análise, a dialética tem experimentado um progresso considerável devido ao desenvolvimento da filosofia alemã do século XIX, com a dialética de Hegel, e do século XX, com os trabalhos de Marx, Engels, Gramsci e Lukàcs. Todos esses filósofos produziram trabalhos fundamentais a respeito da dinâmica das contradições na natureza e nas sociedades baseadas em classes sociais, dando origem a o que se convencionou chamar de Materialismo Histórico [27, 28, 29, 30, 31, 32].

O método dialético do Materialismo Histórico é uma ferramenta para o estudo de sistemas que leva em conta a dinâmica de suas contradições como processos dinâmicos com fases intercaladas de *evolução e crise revolucionária*. Isso tem inspirado a concepção de um método computacional inteligente evolutivo para classificação que seja capaz de resolver problemas comumente resolvidos por abordagens baseadas em redes neurais e algoritmos genéticos.

A maioria dos paradigmas de Computação Inteligente, particularmente redes neurais, computação evolucionária e algoritmos de inspiração cultural, são baseados em algum tipo de teoria voltada a aplicações gerais, embora na prática sejam bastante incompletas; como ilustração, têm-se os seguintes fatos: as abordagens baseadas em redes neurais se apóiam sobre determinados modelos do cérebro; já a computação evolucionária é baseada na Teoria da Evolução de Darwin e em seus desenvolvimentos posteriores; por sua vez, algoritmos de inspiração cultural são baseados no estudo de populações, tais como os algoritmos baseados em colônias de formigas.

Contudo, é importante salientar que não é necessário fazer uso de uma teoria completa e de uso geral para se conceber algoritmos nela inspirados. Aliás, se fosse necessário uma teoria completa para tanto, esta seria uma tarefa praticamente impossível. Por exemplo, as redes neurais utilizam um modelo de neurônio bastante conhecido, embora reconhecidamente incompleto [33, 8]. Esta seria uma forte razão para se investigar o uso de Filosofia como uma fonte de inspiração para o desenvolvimento de métodos e modelos computacionais inteligentes para serem aplicados em diversas áreas, tais como classificação não supervisionada e reconhecimento de padrões, por exemplo.

Thornley e Gibb utilizaram uma abordagem dialética para entender mais claramente a natu-



**Figura 3.1** Heráclito (540-470 a.C.)

reza paradoxal e conceitualmente contraditória da disciplina de recuperação de informação (ou mineração de dados) [34], enquanto Rosser Jr. aplicou alguns aspectos e categorias da dialética materialista à dinâmica não linear, comparando alguns aspectos do método dialético de Marx e Engels a conceitos de Teoria da Catástrofe, Complexidade Dinâmica Emergente e Teoria do Caos [35]. Entretanto, ainda se desconhecem trabalhos onde se proponha uma abordagem matemática consistente que estabeleça os fundamentos da dialética como uma ferramenta para construir métodos computacionais inteligentes.

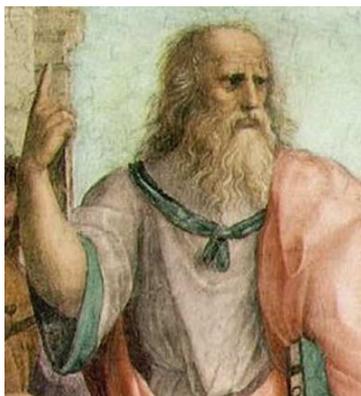
Neste capítulo apresenta-se uma breve síntese da história do método dialético e de seu desenvolvimento, além de se comentar propostas de modelos teóricos matemáticos para a dialética e aplicações. Neste capítulo também é apresentada uma proposta de formulação teórica matemática e algorítmica para o método dialético, com aplicações em Computação Inteligente.

## 3.2 História da Concepção Dialética

O termo *dialética* vem do grego, *dialetiké*, onde o prefixo *dia* significa interação, enquanto *letiké* tem a mesma raiz de *logos*, que pode ser traduzida por saber, palavra, conhecimento. Assim, *dialética* e *diálogo* vêm da mesma origem, o que explica o significado que a dialética tinha entre os filósofos socráticos, bem diferente do significado corrente utilizado no jargão atual da Filosofia: a arte do diálogo. Através da dialética, ou seja, da livre interação entre as diferentes ideias a respeito de um objeto, interação proporcionada pelo diálogo, seria possível chegar à ideia verdade, ou pelo menos ao eixo temático em torno do qual estaria a ideia verdade [11, 36, 37]. O conceito atual de dialética tem muito pouco a ver com esse conceito inicial.

A concepção dialética da realidade não tem uma origem muito bem definida na história do pensamento humano, mas seus antecedentes na história da Filosofia Ocidental podem ser encontrados em Heráclito (ver figura 3.1), um dos filósofos pré-socráticos, embora o termo tenha se popularizado mais tarde com Platão.

Platão (ver figura 3.2) viveu em Atenas, no período de 428-347 a.C. Foi discípulo de Sócrates (ver figura 3.4) e mestre de Aristóteles (ver figura 3.5), sendo considerado o fundador da Academia (representada no quadro “A Escola de Atenas”, de Rafael Senzo, figura 3.3), dada a



**Figura 3.2** Platão (428-347 a.C.)

importância que dava à construção do conhecimento a partir do franco debate de ideias. É à arte de argumentar que Platão dá o nome de dialética, chegando mesmo a afirmar que a dialética é a própria Filosofia por excelência. O termo recebe esse nome por Platão definir a dialética como o movimento das contradições das ideias, ou seja, a dinâmica da argumentação, onde estão presentes uma *tese* e uma *antítese* [38, 10]. Neste contexto se deve entender contradição como sinônimo de diferença.

Tese e antítese *a priori* não estão em contradição antagônica, ou seja, não são totalmente opostas. Da dinâmica entre tese e antítese surge a *síntese*, que vem a se transformar numa nova tese na presença de outra ideia que venha a desempenhar o papel da nova antítese. Assim, a dinâmica da argumentação se dá pelo movimento das contradições entre tese e antítese, ocorrendo sínteses que novamente darão origem a nova tese e nova antítese.

Heráclito, porém, muitos anos antes de Platão, via a dialética não como movimento das contradições entre ideias, mas sim como método para estudar a natureza a partir da dinâmica das contradições entre seus elementos constituintes. Assim, os pólos das contradições na dialética de Heráclito não são ideias, mas podem ser qualquer coisa. Heráclito levava a concepção dialética tão a sério que não conseguia ver nada de constante no mundo, a não ser a mudança, segundo uma frase dele mesmo. Ele afirmava que na base de toda a natureza existe a dialética entre os chamados quatro elementos básicos: a água, o fogo, a terra e o ar [38, 10].

Contudo, embora a escola de pensamento dos quatro elementos tenha se estabelecido bem na Grécia Antiga, as consequências da concepção dialética de Heráclito não foram bem aceitas, uma vez que lançava por terra a segurança de um mundo cíclico, evidenciado por fenômenos como as estações do ano e pelas diversas repetições expressas na ideia de destino e na religião grega [10].

Heráclito chegava a afirmar que, quando um homem colocava um pé dentro de um rio, ele e o rio passavam a ser um só, influenciando um ao outro, mas que, logo no instante seguinte, o pé não era o mesmo pé, muito menos o rio era o mesmo rio. Esse tipo de pensamento trazia uma concepção da natureza como algo continuamente mutável para uma sociedade que sofria de aversão a esse tipo de pensamento, como o era a grega. Para os gregos, a natureza e os seus deuses já eram suficientemente imprevisíveis para se admitir uma corrente de pensamento como a de Heráclito [10]. Assim, apesar de ter sido o pai da dialética, os gregos atribuíram a



**Figura 3.3** “A Escola de Atenas”, de Rafael Sanzio, exposta no Palácio Apostólico do Vaticano (1506-1510)



**Figura 3.4** Sócrates (469-399 a.C.)



**Figura 3.5** Aristóteles (384-322 a.C.)

Heráclito o título póstumo de “O Obscuro” [38, 37].

O pensador grego Aristóteles, nascido mais de um século depois de Heráclito, tendo vivido de 384 a 322 a.C., foi responsável pela reintrodução da concepção dialética de Heráclito no mundo grego, embora numa versão menos radical. Ele foi o primeiro a fundir a noção, comum até aquele momento, de dialética apenas como livre interação e interpenetração de ideias, ao pensamento dialético de Heráclito, incorporando à dialética um sentido próximo ao seu sentido atual [38]. Além disso, Aristóteles foi o primeiro a utilizar a dialética para analisar questões inerentes à essência do ser, também chamadas questões *ontológicas*, e aplicando ao estudo da sociedade, da religião, do Estado e da Ética, e não apenas da natureza, preocupação predominante do pensamento de Heráclito [39].

Para Aristóteles, as coisas, os seres, estão em constante *movimento*. Esse movimento nada mais é do que o desenvolvimento de suas *potencialidades*. Assim, o movimento das coisas é na verdade a atualização das potencialidades no tempo, ou seja, a transformação de possibilidades em realidades efetivas [38]. Um dos principais temas aristotélicos é a Ética, vista essencialmente como originada da dialética entre a História (tradições e costumes) e as leis, sendo esses pólos mediatizados por um terceiro: a razão [39].

É muito provável que as concepções dialéticas da natureza elaboradas no ocidente tenham vindo do oriente diretamente ou, pelo menos, tenham sido influenciadas pelas concepções filosóficas do oriente, em especial das correntes de pensamento monistas orientais, tais como o hinduísmo, o budismo, o taoísmo e filosofias chinesas de menor importância [10].

Para as filosofias monistas, tudo é um, e o movimento desse um é expresso pela dinâmica das contradições entre os pólos que o constituem, mesmo que sejam opostos. Essas ideias se expressam com clareza no símbolo taoísta do *Yin* e *Yang* (ver figura 3.6), onde as formas indicam pólos opostos a partir de suas cores. A relação entre essas formas, o Yin e o Yang, indica movimento, e dentro de cada uma dessas formas tem-se um pequeno círculo de cor oposta à da forma em que está inserido. Esses pequenos círculos expressam a ideia de que o germe do novo já está presente no antigo, e indicam a passagem das mudanças quantitativas para qualitativas, o que está diretamente relacionado com o conceito de *práxis* [10, 11, 37]. Assim, o símbolo Yin e Yang acaba por servir para ilustrar conceitos muito próximos daqueles explorados na concepção contemporânea da dialética [10].



**Figura 3.6** Yin e Yang: símbolo taoísta da dialética entre opostos

Entre o final do século XVIII e começo do século XIX, o filósofo alemão Georg Wilhelm Friedrich Hegel (ver figura 3.7), 1770-1831, a partir das concepções dialéticas orientais trazidas da China pelos jesuítas e resgatando a concepção de Heráclito de dialética como método para análise da realidade e sua visão de realidade como processo em eterna mutação, lança as bases modernas do método dialético, estabelecendo categorias e o aplicando para o estudo de importantes questões para a Filosofia, incluindo a dinâmica social e a origem do Estado [40, 41, 36, 42].

A importância da contribuição de Hegel ao pensamento filosófico ocidental é tão grande que é comum se afirmar que toda a “Geografia” da Filosofia está dividida em dois “continentes”: o continente Kant e o continente Hegel, cabendo ao filósofo escolher a sua “civilização” [41]. Embora Kant alegasse utilizar o método dialético em seus trabalhos distemáticos e até mesmo usasse o termo “dialética”, é o pensamento de Hegel que pode ser considerado uma concepção realmente dialética da realidade, pois ele não apenas contribui para a construção de um novo método dialético, resgatando as ideias de Heráclito e Platão, mas utiliza esse mesmo método renovado como base para todo o seu edifício filosófico [41, 40].

Hegel estabeleceu a sua concepção dialética sobre três categorias ou princípios: a *contradição*, a *totalidade* e o *movimento perpétuo*. Afirmou também que, da mesma forma que Heráclito, a dialética ocorre entre vários pólos, mas ao mesmo tempo afirmou que a dinâmica das contradições será determinada pela contradição principal, e que os pólos que determinam essa contradição principal desempenham papel de tese e antítese, sendo que da sua síntese tem-se um novo pólo a integrar o processo. Definindo pólos da contradição principal como tese e antítese, Hegel faz uma concessão a Platão e às concepções filosóficas orientais [10].

A categoria da *contradição* é definida por Hegel como uma característica que define a identidade dos pólos do sistema. Assim, se dois pólos não são o mesmo, há entre eles contradição, e se há entre eles contradição, é porque não são o mesmo. No movimento das contradições ocorre a *luta de pólos* ou *interpretação de pólos*, ou seja, os pólos integrantes do sistema competem entre si pela hegemonia do sistema e se influenciam mutuamente. Dá-se o nome de *evolução* às mudanças *quantitativas* do sistema, ou seja, aos movimentos das contradições que



**Figura 3.7** Georg Wilhelm Friederich Hegel (1770-1831)

não resultam em mudança dos pólos dominantes. Quando um nóvo pólo passa a hegemonizar o sistema, diferente do pólo dominante anterior, diz-se que aconteceu uma *revolução*, ou *superação dialética*, resultando em mudanças *qualitativas* do sistema [40, 41, 37, 42].

Já o princípio da *totalidade* ilustra que um sistema não pode ser compreendido a partir apenas de seus pólos isoladamente. É necessário compreender como os pólos integrantes estão relacionados e, embora possa existir uma contradição principal, para apreender a dinâmica do sistema é preciso considerar todas as contradições entre os pólos [41, 42].

A categoria do *movimento perpétuo* é expressa por Hegel de forma muito semelhante às ideias de Heráclito de mudança constante, tanto que Hegel se apropria da frase de Heráclito ao expressar que “só a mudança é constante”. Para Hegel, embora ele não tenha utilizado essas palavras, um sistema e um processo são exatamente a mesma coisa: não se pode conceber um sistema totalmente estático [42].

A concepção dialética de Hegel causou um grande impacto na Filosofia Ocidental, marcando o início do predomínio alemão na construção do conhecimento filosófico, que veio a ser sedimentado por nomes como Friederich Nietzsche, Karl Marx e outros. No entanto, a natureza do movimento das contradições e sua origem era uma questão que inquietava a mente de Hegel, que o levou a estabelecer como princípio a existência de um “espírito” que movia o sistema em uma determinada direção. Entenda-se aqui “espírito” como algo que liga os pólos e os move em suas relações em alguma direção, segundo algum sentido. Ou seja, a dialética de Hegel era *subjetiva*, ou *idealista*, porque ligava a origem da dinâmica das contradições à ação dos próprios pólos integrantes do sistema. A ideia de *espírito* está ligada à concepção de que esse movimento, resultante da ação de cada pólo como sujeito, teria um sentido, um fim em si, que para Hegel apontava para algo superior. É a ideia do progresso constante [10].

No entanto, a melhor maneira de entender essas posições de Hegel especificamente é percebê-las como uma espécie de concessão ao pensamento de Platão: em questões ontológicas, a contradição principal no sistema hegeliano se resume à contradição entre sujeito e objeto, entre ideia e matéria, onde sujeito e ideia são *teses*, enquanto objeto e matéria são suas respectivas *antíteses*. Hegel, portanto, coloca o sujeito e a ideia no instante inicial do movimento dialético,

além de submeter a dinâmica das contradições ao desenvolvimento desses pólos em específico, reduzindo os papéis dos pólos objeto e matéria, respectivamente, a papéis secundários, o que parece ser uma influência do pensamento platônico, que submetia o mundo material a modelos ideais, ou seja, a matéria à ideia e o objeto ao sujeito. Assim, nesta concepção, muito embora tanto o sujeito seja influenciado em sua essência pelo objeto quanto a ideia receba influência da matéria ao longo do movimento, é o sujeito que dá origem e sentido ao objeto, e é a ideia que dá sentido e, de certa forma, cria a matéria. Nota-se aqui a razão de a concepção dialética de Hegel ser uma concepção predominantemente *subjetivista*, ou *idealista* [41].

Assim, Hegel considera que a dinâmica das contradições é determinada pela evolução do “espírito do sistema”. A dialética hegeliana é portanto *idealista*, ou *subjetiva*, dada a importância que dá ao sujeito na determinação da dinâmica das contradições, sujeito este representado também como sujeito coletivo, expresso na ideia de espírito do sistema [42].

Ao utilizar seu método dialético para analisar o Estado e a sociedade de sua época, Hegel atribuiu os desenvolvimentos históricos da sociedade ao “espírito do tempo”. Isso teve algumas consequências funestas, tais como o desenvolvimento de uma visão conservadora da sociedade, dado que essa sociedade naturalmente fluiria para uma sociedade superior, sendo os seus problemas apenas “acidentes” na sua evolução. Essa visão também levou a uma aproximação excessiva da Filosofia à Teologia, dada a identificação possível do espírito do tempo à ideia de divindade, o que levou a um consequente distanciamento entre Filosofia e Ciência, dada a influência do racionalismo iluminista que ainda existia e do cientificismo positivista que começava a surgir naquele século.

No entanto, diversos pensadores se sentiram incomodados com esses aspectos da dialética hegeliana e provocados a construir uma concepção dialética que, ao mesmo tempo que aproveitasse as principais ideias e categorias do método de Hegel, reaproximasse Filosofia e Ciência [42, 10]. Uns renunciaram à dialética. Outros passaram a integrar, na Alemanha de então, o grupo dos *hegelianos de esquerda*. Mais adiante os assim chamados hegelianos de esquerda dão lugar a um novo movimento, que propõe uma nova concepção dialética da realidade, a *dialética materialista* [40, 38].

### 3.3 O Método Dialético Materialista

Karl Marx (figura 3.8), 1818-1883, e Friederich Engels (figura 3.9), 1820-1895, preocupavam-se com os rumos que a situação social e política estava tomando em seus países e na Europa como um todo e aderiram ao grupo dos hegelianos de esquerda. Contudo, perceberam que a ideia de espírito do sistema era um entrave para a construção de uma base filosófica para os movimentos progressistas da época, além de ser profundamente anticientífica, dado que, na pior das hipóteses, quase todas as medidas políticas tomadas pelo autoritário Estado alemão da época, além das graves injustiças sociais que acometiam as suas classes trabalhadoras, poderiam ser justificadas como passos necessários na evolução do espírito [42, 43, 44, 45, 46].

Marx ficou bastante impressionado com os trabalhos de Ludwig Feuerbach, que tentavam esboçar uma reação ao idealismo que predominava no pensamento alemão. Ele chegou à conclusão que a Filosofia deveria ser não apenas um conjunto de ferramentas teóricas para estudar e explicar a realidade, mas sim um instrumento para transformá-la [40, 36, 47]. No entanto,



**Figura 3.8** Karl Marx (1818-1883)

Marx percebeu que o pensamento de Feuerbach estava carregado de materialismo vulgar, ou seja, de uma visão que, embora reclamasse a real posição do ser humano dentro da realidade, deixava-o condicionado à natureza e à economia, sem muito espaço para sua ação como sujeito da História [40]. Assim, ao refletir sobre a concepção dialética hegeliana, Marx chegou à conclusão de que o sistema de Hegel estava “de cabeça para baixo”, e que as contribuições do materialismo poderiam ajudar a colocar a dialética em seu lugar adequado [40, 48].

Em questões ontológicas, viu-se que a dialética hegeliana colocava, no caso de dialéticas do tipo sujeito-objeto, o sujeito como tese e o objeto como sua antítese, e em dialéticas do tipo ideia-matéria, a ideia como tese e a matéria como sua respectiva antítese. Além disso, a dinâmica dialética se movia em função dos pólos sujeito e ideia. Havia portanto uma sobrevalorização da influência do sujeito na História, e um predomínio da ideia sobre a matéria, muito embora a dinâmica dialética posterior permitisse períodos de predomínio dos pólos objeto e matéria. Para Marx, tratava-se de “inverter” o método hegeliano, trocando os pólos, dando origem a um sistema onde se tenha, na dialética sujeito-objeto, o objeto como tese e o sujeito como antítese, dado que o sujeito surge antes como objeto; já na dialética ideia-matéria, a matéria surge, no pensamento marxiano, como tese, enquanto a ideia é assumida como antítese, dado que, na concepção materialista, a ideia é antes originada da matéria.

O método dialético concebido e desenvolvido por Marx e Engels parece ter apenas diferenças pontuais em relação ao sistema hegeliano, já que aproveita as categorias definidas por Hegel e apenas inverte a tese e a antítese hegeliana inicial. No entanto, as consequências de tal inversão implicaram profundas diferenças qualitativas: não há mais um espírito do sistema, pois a dinâmica das contradições agora é determinada pelas condições *materiais*, ou *objetivas*, no momento inicial, sendo também afetada por condições *ideais*, ou *subjetivas*, nos momentos posteriores, havendo por vezes um predomínio das condições objetivas ou das condições subjetivas sobre a dinâmica do sistema. O método dialético assim concebido recebe o nome de *dialética materialista*, ou *materialismo dialético* [43, 46].

Particularmente, Marx desenvolveu contribuições que focaram mais áreas como História, Economia, Sociologia e Ciência Política [42, 49], enquanto Engels tinha interesses em História,



**Figura 3.9** Friederich Engels (1820-1895)

Ciência Política, Biologia e suas relações com o Evolucionismo de Darwin e seus desdobramentos. No entanto, Engels não foi apenas uma espécie de coadjuvante na formulação do método dialético materialista: é de Engels a formulação dos princípios da dialética na forma de [50, 28]:

1. transformação da quantidade em qualidade e vice-versa;
2. interpenetração dos opostos (pólos);
3. negação da negação (superação dialética, ou práxis).

Uma das principais aplicações da concepção dialética materialista é a análise de sociedades, onde as sociedades são sistemas cujos pólos integrantes são as classes sociais. Assim, estudam-se as fases históricas pelas quais passa uma determinada sociedade, caracterizadas pela hegemonia de determinadas classes sociais. A ideologia, a cultura, a religião, os modos de produção, o Estado e sua evolução passam a ser vistos a partir das contradições de classe, que por sua vez condicionam a forma e a evolução dessas estruturas e instituições [42, 49, 46].

O trabalho de Marx e Engels foi desenvolvido e aprofundado pelas diversas escolas de pensamento marxistas que surgiram nos anos que se seguiram à publicação de “A Ideologia Alemã” [49], a obra que lançou o pensamento dialético materialista. Dentre os continuadores podem-se destacar figuras como Antonio Gramsci, Rosa Luxemburg, Walter Benjamin, Georg Lukács e a chamada Escola de Frankfurt [32]. Neste trabalho estuda-se a dialética materialista visando construir novos métodos de Computação Inteligente para modelagem e classificação, tendo como principal objetivo construir métodos computacionais para classificação e reconhecimento de padrões.

### 3.3.1 Principais Categorias

A concepção dialética materialista da realidade, surgida da crítica de Marx à concepção dialética idealista de Hegel e ao materialismo vulgar de Feuerbach [42], apesar de todos os ataques que vem sofrendo desde seu início devidos principalmente às concepções políticas e filosóficas possíveis de emergirem da sua aplicação ao estudo da sociedade e à sua superação, ainda

se mostra como uma das mais objetivas ferramentas de interpretação da realidade, ao mesmo tempo capaz de compreendê-la e negá-la [46, 47, 51, 52, 53], visando a sua transformação e superação [54].

A dialética materialista é apresentada nesta seção como uma ferramenta não necessariamente vinculada à ciência política, mas muito mais geral: ferramenta capaz de servir de esqueleto para construção de uma teoria crítica e de ampla aplicação para o entendimento e apreensão da realidade em seus mais diversos aspectos e fenômenos [50], sem perder sua característica mais importante e mais forte, ou seja, a capacidade de aderir destrutivamente ao objeto de estudo, pois ao mesmo tempo que o explica, a dialética lança as bases para a sua superação [46, 54]. Contudo, para compreender o que a dialética materialista entende por superação, é importante definir o conceito de *práxis*.

Embora a palavra grega *práxis* esteja na origem da palavra *prática*, seu significado é bastante diferente, estando relacionado a outros dois termos gregos: *theoria* e *poiésis*. A *theoria* é a reflexão intelectual, a racionalização, não necessariamente voltada para uma aplicação, dando origem à palavra *teoria*. Já *poiésis* significa o próprio trabalho, manual ou mental, com fim produtivo. Assim, *poiésis* está igualmente na origem tanto da palavra *poeta* quanto da palavra grega *poiménis*, que significa *pastor*. *Práxis* assume significados diferentes em Hegel e em Marx, mas em ambos possui significado equivalente a *superação* [11, 36].

Para Hegel, a *práxis* é o movimento *theoria-poiésis-theoria*, sujeito-objeto-sujeito ou ideia-matéria-ideia. Portanto, no pensamento hegeliano, quando a ideia se movimenta rumo à matéria, que por sua vez modifica a ideia, diz-se ter acontecido uma *superação dialética*, ou *mudança qualitativa*. Assim, no conceito hegeliano de *práxis*, a matéria está, na prática, subordinada à ideia e existe em função da mudança qualitativa da ideia [11].

Para Marx, Engels e os filósofos dialéticos materialistas, a *práxis* é o movimento *poiésis-theoria-poiésis*, objeto-sujeito-objeto ou matéria-ideia-matéria. Assim, dá-se uma *superação dialética* quando, a partir da matéria, surge a ideia que modifica a matéria. Portanto, na dialética materialista, a ideia está subordinada à matéria, e uma mudança qualitativa se dá quando essa ideia modifica a matéria [11].

A diferença qualitativa entre os dois conceitos de *práxis* fica evidente no seguinte exemplo: caso se queira entender o processo de aprendizado como um processo dialético envolvendo teoria e prática, segundo a perspectiva hegeliana o aprendizado partiria sempre de conhecimento *a priori* a ser aperfeiçoado e testado na prática, gerando conhecimento qualitativamente diferente. Já na perspectiva dialética materialista seria possível conceber o processo de aprendizado sem nenhum conhecimento *a priori*, tendo apenas a apreensão sensorial do objeto do conhecimento e, em seguida, a reflexão a respeito dessas informações sensoriais e a posterior atuação sobre o objeto do conhecimento. Nota-se que as duas visões possíveis de aprendizado, a dialética hegeliana e a dialética materialista, são bastante diferentes [55, 56].

Dada a importância do conceito de *práxis* para a dialética materialista, o pensador italiano Antonio Gramsci, (figura 3.10), 1891-1937, batizou o edifício teórico que tem suas bases no método dialético materialista com o termo *Filosofia da Práxis* [12]. Quando empregada como concepção dialética da história, a Filosofia da Práxis também recebe o nome de *materialismo histórico*, onde o adjetivo *histórico* pressupõe a história como dinâmica dialética [13, 57, 58]. No entanto, é comum se utilizar os termos *materialismo dialético*, *dialética materialista*, *ma-*



**Figura 3.10** Antonio Gramsci (1891-1937)

*terialismo histórico, Teoria Crítica e Filosofia da Práxis* como sinônimos, quando na verdade a dialética materialista, ou materialismo dialético, é o método que embasa essas construções filosóficas.

O método dialético consiste em conceber partes da realidade, ou *fenômenos*, como *sistemas*. Esses sistemas são compostos por diversos *pólos*. A cada pólo está associada uma potência (caso se queira aqui fazer uma concessão à dialética de Aristóteles [39]) ou *força*. O sistema é então caracterizado pela sua *correlação de forças*, ou seja, pelo conjunto das forças dos pólos que o compõem [59].

A dialética materialista é composta por três *categorias* básicas: a *contradição*, o *princípio da totalidade* e o *movimento perpétuo* [54]. Os pólos estão interrelacionados de forma que a força de um pólo no instante imediatamente futuro é função da correlação de forças presente. Isto se dá entre pólos distintos e recebe o nome de *contradição*. Assim, a contradição é a base da dialética, e a própria dialética é o movimento das contradições, ou seja, é a dinâmica entre os diferentes pólos que compõem o sistema [60].

Entretanto, um pólo não pode ser analisado isolado dos outros pólos que compõem o sistema. Ele também está sob influência de forças externas, de *condições*. Essas condições podem ser *externas* ou *internas* ao sistema. Haverá condições externas ao sistema quando esse sistema integrar um sistema maior, abstraído pelas condições que o afetam. As condições internas podem ser *objetivas* ou *subjetivas*. Ao princípio de não se poder conceber um sistema totalmente isolado, mas sim em suas várias relações e dependências com outros sistemas, o mesmo valendo para os seus pólos constituintes, dá-se o nome de *princípio da totalidade*, sendo este princípio talvez a categoria mais importante da dialética materialista [54, 36].

Uma *condição subjetiva* é função da força do pólo ao qual está associada e das condições externas e internas. Uma condição subjetiva está associada à propriedade do pólo associado de ser *sujeito*. É expressão da *vontade* do pólo. É *ideia*. Assim, as condições subjetivas são *controláveis* pelos pólos associados [13] por meio de *fatores subjetivos* inerentes ao pólo, resultantes da *ação* do pólo. Uma *condição objetiva* é função da correlação de forças e das condições

externas e internas. Uma condição objetiva está associada à propriedade do pólo associado de ser *objeto* em relação aos outros pólos e a ele mesmo. Representa o que é *espontâneo* no pólo. É *matéria*. Assim, as condições objetivas *não são controláveis* pelos pólos associados [13].

Michael Löwy dá uma definição bastante interessante para mais uma das três categorias constituintes do método dialético da Filosofia da Práxis, tirada do Fausto de Goethe: “Eu sou o espírito que sempre nega, e isso com razão, porque tudo que existe merece acabar” [54]. É o diabo de Goethe que usa esta expressão na primeira vez em que aparece para Fausto. Esta citação expressa de forma original a *categoria dialética do movimento perpétuo*, que como o nome indica, expressa a hipótese fundamental da dialética de que não existe nada de eterno, nada fixo, nada absoluto. Tudo está em constante mudança, em perpétua transformação, sujeito à superação [54].

O método dialético considera os sistemas como sendo sistemas *dinâmicos*, ou seja, sistemas variantes no tempo. Assim, toda análise dialética é ao mesmo tempo uma análise dinâmica, pois se baseia no estudo dos estados passados do sistema, do seu estado presente, e estima seu estado futuro, considerando determinadas condições mais prováveis [54]. Desta forma, um *estado* é caracterizado pela *hegemonia* de um dos pólos do sistema. Quando em um sistema um pólo tem uma força associada maior do que as dos outros pólos, diz-se que este pólo *hegemoniza* o sistema. O pólo hegemônico, no entanto, influencia e é influenciado pelos outros pólos integrantes do sistema. Assim, a dinâmica do sistema é a tentativa de resolução da contradição, ou seja, é a tentativa de um pólo obter a hegemonia absoluta sobre os outros pólos do sistema [13]. Este mecanismo é chamado de *luta de pólos*, ou *interpenetração de pólos*, e é uma consequência do princípio da totalidade [60].

Assim, a definição de *estado* na concepção dialética é uma definição estritamente *qualitativa*. De acordo com as condições externas e internas e com a correlação de classes anterior, vão se procedendo mudanças *quantitativas* dentro do sistema, caracterizadas pelas variações dos valores das forças dos pólos. A estas mudanças quantitativas dá-se o nome de *evolução*. Quando um novo pólo hegemoniza o sistema, dá-se o nome de *revolução* e diz-se que ocorreu uma mudança *qualitativa* [60]. Toda revolução é precedida por uma *crise revolucionária*. A crise revolucionária acontece quando a hegemonia do sistema é compartilhada, ou seja, quando duas ou mais forças disputam a hegemonia. Se, após o estado de crise revolucionária, o sistema retornar ao estado anterior, diz-se que houve uma *contra-revolução* [46].

A respeito do conceito de materialismo, é importante realçar que se deve entender, neste trabalho, o materialismo tal como é definido na dialética materialista concebida por Gramsci [13], com o sentido de propriedade de ser objeto em relação a um sujeito, ou seja, *objetivismo*, e não como materialismo vulgar, o que leva muitos autores “marxistas” a considerar que as condições materiais *determinam* todas as outras relações no sistema, o que, além de ser um equívoco, reduz a dialética materialista a mero determinismo, na prática a destituindo, portanto, da ferramenta dialética e retornando ao materialismo vulgar. Mesmo teóricos como Plekhanov chegaram a confundir a concepção de materialismo de Marx com o materialismo vulgar em suas obras, carregando-as de certa dubiedade [57, 58, 61], quando o próprio Marx renegou tanto o materialismo vulgar [42, 54] quanto o idealismo expresso no voluntarismo de alguns pensadores, que consideravam que a vontade humana poderia fazer tudo, inclusive acelerar etapas do desenvolvimento histórico e social sem considerar as condições objetivas [47].

Gramsci adverte para o perigo de se confundir o materialismo da Filosofia da Práxis com o materialismo vulgar dos positivistas [12], que viam o mundo objetivo como o único existente e o conhecimento científico rigorosamente construído como o único conhecimento válido, ou seja, os positivistas eram fetichistas da ciência, o que levava a posições extremamente naturalistas (no sentido de conceber o ser humano reduzido a um sistema que oferece respostas quando devidamente estimulado, praticamente negando ou reduzindo sua subjetividade e realçando apenas sua objetividade), posições estas que frequentemente levavam os positivistas a posturas conservadoras e, por vezes, reacionárias [54].

Pode-se dizer que a confusão entre o materialismo da Filosofia da Práxis, ou seja, o materialismo dialético, e o materialismo vulgar dos positivistas foi responsável em parte pela estagnação do desenvolvimento da Filosofia da Práxis, dada a sua redução a um “marxismo ortodoxo”, alijado de uma maior compreensão dialética e reduzido ao determinismo vulgar e ao mecanicismo, ignorando a dinâmica dos processos reais e reduzindo a concepção dialética a uma forma mais sofisticada de positivismo.

### 3.3.2 Modelos e Aplicações

A dialética pode ser uma ferramenta teórica de cunho muito mais geral do que o que veio a ser chamado de *marxismo* e outras concepções filosóficas dialéticas classicamente concebem. Ela tem aplicações potenciais nos mais diversos campos do saber, que vão das ciências humanas, particularmente a Filosofia, onde nasceu e classicamente vem sendo utilizada, à Engenharia e às Ciências Exatas, como proposto neste trabalho, em áreas como a Inteligência Computacional. Contudo, deve ser adotado um modelo matemático adequado, o que se mostra uma tarefa bastante difícil, dado que modelos baseados em equações diferenciais lineares e não lineares, ou seja, modelos clássicos para sistemas dinâmicos, não conseguem modelar todos os aspectos do método dialético [35].

Entretanto, a abordagem dialética possui diversas aplicações potenciais, o que gera uma necessidade de se estabelecer modelos adequados para a representação desse método filosófico investigativo [35]. Uma dessas aplicações potenciais está no auxílio à Recuperação de Informação (*Information Retrieval*, IR) [34].

#### 3.3.2.1 IR e o Método Dialético

IR é a ciência da busca de informação em documentos, da busca dos próprios documentos, de metadados que descrevem os documentos, ou da busca de informação em bases de dados, que podem ser desde bases de dados propriamente ditas até mesmo um conjunto de hipertextos interconectados através de uma rede, como a World Wide Web [62].

Contudo, IR é uma área do conhecimento ainda bastante problemática, caracterizada por conflitos conceituais, paradoxos, e pela falta de clareza conceitual e de coerência teórica. Contudo, o uso da dialética, emprestado da Filosofia da Práxis, pode ajudar no estabelecimento de conceitos, definições e metodologias que possam lidar com as questões candentes da IR [34].

A natureza problemática da IR como área do conhecimento está relacionada com a natureza problemática dos seus próprios conceitos centrais, dentre eles o conceito de *significado*. Uma abordagem do conceito de significado como um processo dialético pode ajudar no desenvolvi-

mento da IR, ao mesmo tempo em que expõe os porquês de seus problemas. Uma das razões para a natureza conflituosa da IR é a forma como a IR manifesta e exemplifica as relações dialéticas que existem no conceito de significado. Assim, uma disciplina como a IR, ainda com muitas dificuldades para discutir e analisar seus conflitos sem sucumbir às suas próprias contradições e confusões conceituais, pode ser beneficiada pelas ferramentas fornecidas pela abordagem dialética [34].

A Recuperação de Informação tem falhado ao tentar incorporar o conceito de significado desenvolvido pelas diversas áreas da Filosofia [63] e por diversas ciências, como a Linguística [55]. No entanto, a ausência de um conceito de significado não tem afetado apenas a assim chamada tradição qualitativa da IR: sua ausência também afeta sua tradição objetiva ou matemática. Entretanto, a natureza e a razão da existência das dificuldades teóricas em IR também são causadas, pelo menos parcialmente, pelas contradições inerentes ao conceito de significado. Logo, o desenvolvimento de um conceito de significado como um processo dialético e o seu uso para revelar quais conflitos dentro da IR são um produto ou manifestação desse mesmo processo dialético pode esclarecer a relação entre a IR e o conceito de significado, além de poder ajudar na definição de representações adequadas para o problema [34].

Thornley e Gibb [34] propõem que o conceito de significado deve ser considerado como um processo dialético entre os pólos objetividade e subjetividade, aliando dialeticamente as tradições quantitativa e qualitativa da IR, respectivamente, uma vez que a tradição quantitativa, da sua parte, exige métodos de validação e métricas de confiabilidade que carregam muita subjetividade, apesar da representação do objeto como um vetor de atributos, enquanto a tradição qualitativa parte do pressuposto de que é o contexto individual que importa, o que torna difícil uma abordagem mais especializada, além de dificultar a apresentação de resultados na forma de estatísticas, por exemplo [34]. Thornley e Gibb propõem ainda que, quando da definição da representação do significado de um objeto, pode ser muito mais importante representar características pelo que elas não são, tirando proveito da ignorância a respeito de sua natureza, do que tentar estabelecer definições mais rígidas para essas características [34].

### 3.3.2.2 Dialética e Apoio à Decisão

Muitas decisões complexas e não estruturadas são prejudicadas pela falta de clareza a respeito dos pressupostos e perspectivas envolvidos no processo de decisão. Os Sistemas de Apoio à Decisão (*Decision Support Systems*, DSS) tradicionais em geral não levam em conta esses pressupostos e perspectivas no processo decisório quando envolve questões de grande complexidade [64].

Assim, Jarupathirun e Zahedi propõem uma abordagem baseada na dialética platônica. Muito embora o texto fale em dialética socrática, Sócrates não deixou escritos, e tudo que dele se conhece é formulação de seu discípulo Platão, daí muitos filósofos e historiadores alegarem que muitas vezes se lê Platão quando Platão cita Sócrates, o que pode ser o caso da dialética platônica [10].

A abordagem dialética para apoio à decisão proposta por Jarupathirun e Zahedi foi chamada pelos autores de Sistemas Dialéticos de Suporte à Decisão (*Dialectic Decision Support Systems*, DDSS), e consiste em um conjunto de ferramentas de Inteligência Artificial baseadas em agentes, sendo uma tentativa de modelar tese e antítese, simulando o diálogo crítico entre

as partes. Essa forma de modelagem recebe muitas vezes o nome de “inclusão do advogado do diabo” [64]. Entretanto, dado serem baseados em agentes e em regras, os DDSS carecem de um modelo matemático sólido que possa ser utilizado em outras aplicações fora do apoio à decisão, sem contar que implementam um modelo muito primário de dialética, como é o método dialético de Platão.

### 3.3.2.3 Modelagem por Sistemas Dinâmicos Não Lineares

Um dos principais problemas na área da Economia Política é a transformação qualitativa de um sistema econômico em outro. Uma longa tradição, baseada em escritos seminais de Marx, argumenta que essa transformação qualitativa pode ser explicada por uma interpretação materialista do método dialético de Hegel [35].

No entanto, muito embora Marx possa ser considerado o primeiro economista realmente rigoroso do ponto de vista matemático, nessa parte da análise se percebe claramente a ausência de formulações matemáticas, o que faz com que diversos autores considerem o método dialético em conflito com qualquer abordagem quantitativa matematicamente precisa, dada a natureza nebulosa de muitas definições dialéticas e mesmo dos resultados da análise dialética [35].

Rosser Jr. argumenta que é possível construir modelos matemáticos precisos para modelar determinados aspectos do método dialético, em particular aqueles relativos à transformação qualitativa de sistemas. Embora o aspecto da transformação de quantidade em qualidade e vice-versa, para usar uma formulação dialética muito comum nos trabalhos de Engels [50, 28], não constitua a totalidade do método dialético materialista, este é um dos principais aspectos para discutir certos elementos da Teoria do Caos, por exemplo, e de outras teorias que possam ser consideradas casos especiais da dinâmica não linear [35].

Na maioria dos modelos lineares, mudanças contínuas nas entradas não levam a mudanças descontínuas nas saídas. São essas mudanças descontínuas nas saídas, provocadas por mudanças contínuas nas entradas, e vice-versa, que Rosser Jr. considera sua interpretação matemática da transformação dialética de mudanças quantitativas em mudanças qualitativas, ou seja, do conceito de superação dialética, ou práxis [35].

### 3.3.2.4 Dialética e Teoria da Catástrofe

O conceito de *bifurcação* é um conceito chave para analisar discontinuidades em sistemas dinâmicos não lineares [65, 35]. Esse conceito foi descoberto por Henri Poincaré (1854-1912), que desenvolveu uma teoria qualitativa de equações diferenciais para explicar a mecânica celeste quando há mais de dois astros envolvidos. Essa teoria qualitativa envolve a tentativa de encontrar as soluções de equações diferenciais ao redor de pontos de singularidade, sem realmente resolver as equações e utilizando critérios geométricos [66].

Seja a família geral de  $n$  equações diferenciais, cujo comportamento é controlado por um parâmetro de controle  $\mu$  de dimensão  $m$ , da forma que segue:

$$\frac{d}{dt}\mathbf{x} = f_{\mu}(\mathbf{x}), \quad (3.1)$$

onde  $\mathbf{x} \in \mathbb{R}^n$  e  $\mu \in \mathbb{R}^m$ . O ponto de equilíbrio é atingido quando:

$$f_{\mu}(\mathbf{x}) = 0. \quad (3.2)$$

As bifurcações acontecem nas singularidades onde é válida a seguinte condição:

$$\frac{d}{dt}f_{\mu}(\mathbf{x}) = \frac{d^2}{dt^2}f_{\mu}(\mathbf{x}) = 0. \quad (3.3)$$

Nesses pontos podem ocorrer mudanças estruturais, caso o equilíbrio se bifurque em dois equilíbrios estáveis e um instável. A Teoria da Catástrofe consiste no estudo das singularidades estáveis de uma função potencial  $f_{\mu} : \mathbb{R}^n \rightarrow \mathbb{R}$ , assumindo que existe um gradiente. Essas singularidades podem ser entendidas como pontos nos quais os equilíbrios perdem sua estabilidade com a possibilidade de uma mudança descontínua das variáveis de estado, mudança esta proveniente de mudanças contínuas das variáveis de controle [65, 35, 67]. A Teoria da Catástrofe tem diversas aplicações e desenvolvimentos [68, 67, 69], apesar de ser uma teoria relativamente jovem (meados de 1960) e sofrer diversas críticas quanto à sua aplicabilidade [65].

Para René Thom (1923), matemático francês criador da Teoria da Catástrofe, com o formalismo fornecido por esta teoria é possível criar um modelo matemático geral para a morfogênese, ou seja, da transformação qualitativa de uma coisa em outra, da emergência de órgãos e estruturas do desenvolvimento de um determinado organismo, por exemplo. O próprio Thom faz uma ligação entre a sua teoria e a dialética, embora o faça a partir das ideias da dialética idealista de Heráclito [35].

Uma crítica que pode ser feita à modelagem dialética baseada na Teoria da Catástrofe é que, além do fato de servirem para modelar apenas o aspecto da superação dialética, ou mudança de quantidade em qualidade e vice-versa, é que tudo se reduz a mudanças descontínuas em variáveis ou funções, e não verdadeiras mudanças qualitativas. Entretanto, pode-se argumentar que o que emerge das mudanças pode ser considerado variáveis e funções totalmente diferentes, uma vez que dão origem a dinâmicas totalmente novas. Contudo, não se tem uma mudança na quantidade de variáveis envolvidas no sistema dinâmico [35]. Além do mais, não está clara a relação entre o mapeamento variáveis de controle *versus* variáveis de estado e os pólos do processo dialético, o que dificulta a modelagem; sem contar que o conceito de equilíbrio não faz parte do método dialético, embora alguns marxistas o tenham incluído em suas análises, como Bukharin, teórico soviético dos anos 1930 [35].

A Teoria da Catástrofe não se presta como base para construção de um formalismo matemático para o método dialético, dado que não permite a modelagem explícita dos pólos dialéticos, do conceito de contradição, nem do conceito de totalidade. No entanto, como um subconjunto da teoria dos sistemas dinâmicos não lineares, modela a categoria do movimento perpétuo e se pode dizer que é capaz de modelar a práxis, de certa forma, considerando as mudanças contínuas nas variáveis de entrada que provocam mudanças descontínuas das variáveis de controle e vice-versa. Contudo, o máximo que se pode afirmar a respeito da Teoria da Catástrofe é que ela é passível de uma interpretação dialética materialista, que no máximo pode ajudar a entender sua estrutura, possíveis aplicações e resultados, mas que não a torna base para um formalismo matemático adequado para a análise dialética.

## 3.3.2.5 Dialética e Teoria do Caos

O estudo da dinâmica caótica também se origina nas definições da mecânica celeste de Poincaré. A Teoria da Catástrofe e a Teoria do Caos têm em comum o fato de lidarem com transformação de quantidade em qualidade, dialeticamente falando, ao mesmo tempo em que compartilham a ideia de bifurcação do equilíbrio de sistemas dinâmicos não lineares em pontos críticos [35, 66].

Ainda existem controvérsias a respeito da definição de sistema caótico, mas uma definição amplamente aceita é a de que sistemas caóticos são aqueles onde pequenas mudanças nas variáveis de entrada podem levar a grandes mudanças na dinâmica do sistema. Esse efeito, chamado de Dependência Sensível das Condições Iniciais (*Sensitive Dependence on Initial Conditions*, SDIC) é conhecido como “efeito borboleta”, por causa da ideia da parábola, literariamente falando, do bater das asas de uma borboleta em um determinado ponto do planeta que dá origem a um furacão em outro ponto [35].

Uma das equações de tempo discreto geradoras de sistemas caóticos mais estudadas em modelos econômicos é a *equação de diferença logística*, definida a seguir:

$$x(t+1) = \alpha x(t)(k - x(t)), \quad (3.4)$$

onde  $\alpha$  é o *parâmetro de ajuste*, cuja variação muda qualitativamente a dinâmica do sistema. O aumento do valor de  $\alpha$  leva à duplicação da cascata de bifurcações a partir de um único equilíbrio, levando a uma dinâmica caótica e culminando em um comportamento explosivo [35].

Fazendo uma generalização para outros sistemas caóticos a partir da equação 3.4, pode-se dizer que existem duas interpretações dialéticas possíveis: a primeira é de que a cascata de bifurcações pode ser encarada como uma possível representação da emergência de mudanças qualitativas das mudanças quantitativas, a transformação de quantidade e qualidade e vice-versa, como definida por Engels [28]. Contudo, aqui vale a mesma crítica feita à representação desse princípio dialético pela Teoria da Catástrofe [35].

A segunda interpretação possível envolve o conceito dialético de interpenetração dos opostos. Esta interpretação emerge do papel dual da variável  $x$  na equação 3.4:  $x$  pode ao mesmo tempo tanto fazer o valor futuro aumentar quanto diminuir. Isso parece óbvio e simples de observar, mas pode servir para modelar o conflito entre dois pólos, onde o aumento de  $\alpha$  faz a intensidade do conflito aumentar, ou seja, as contradições se intensificam [35].

Contudo, pode-se perceber que nesse modelo os pólos não estão explícitos, embora a equação 3.4 possa se perceber claramente que existem dois pólos, e a quantidade de pólos não varia com as mudanças qualitativas, o que implica fortes limitações na representação do conceito de práxis. Assim, embora estejam representados o princípio da contradição e do movimento perpétuo, o princípio da totalidade está representado de forma limitada, com os pólos integrantes implícitos e numa quantidade fixa, além da práxis também estar representada com limitações. Assim, o máximo que se pode afirmar a respeito da Teoria do Caos é que ela possui uma interpretação dialética materialista, que no máximo pode ajudar a entender sua estrutura, possíveis aplicações e resultados, mas que também não a torna base para um formalismo matemático adequado para a análise dialética.

## 3.3.2.6 Dialética e Mapas de K-Médias

O *mapa de k-médias* é um tipo de rede neural baseado no algoritmo de agrupamento k-médias [33], sendo portanto uma rede não supervisionada. O mapa de k-médias utiliza neurônios iguais aos empregados nas redes SOM. Entretanto, não há nenhuma interação entre estes neurônios e nem a possibilidade de organizá-los em estruturas matriciais: o mapa de k-médias é simplesmente uma rede de camada única.

O procedimento de treinamento de um mapa de k-médias é em tudo semelhante ao treinamento de uma rede SOM, excluindo os passos que se referem à entrada das posições dos neurônios, ao cálculo das distâncias entre os neurônios e ao cálculo da função de vizinhança, dado que não há possibilidade de se organizar os neurônios em estruturas matriciais nem fazê-los interagir entre si. O ajuste de pesos é dado da forma que segue [8]:

$$w_{i,j}(t+1) = \begin{cases} w_{i,j}(t) + \eta(t)(x_j(t) - w_{i,j}(t)), & i = k \\ w_{i,j}(t), & i \neq k \end{cases}, \quad (3.5)$$

onde  $k$  é o índice do neurônio vencedor.

Da expressão da atualização dos pesos fica evidente que o mapa de k-médias é capaz de implementar a categoria do movimento perpétuo, uma vez que os pesos são atualizados dinamicamente. Além disso, é possível associar os neurônios aos pólos do sistema dialético, considerando que os pólos podem ser modelados usando vetores de pesos. O princípio da contradição também está presente, embora não explícito, dado que fica evidente a importância da concorrência entre os pólos (neurônios). No entanto, não se pode afirmar em momento algum que seja possível modelar a categoria da totalidade usando os mapas de k-médias, dado que o comportamento dos neurônios é puramente individual, dependendo apenas dos seus valores passados e das entradas da rede, e não dos pesos dos outros neurônios. Não há interpenetração dos pólos, portanto.

Além do mais, no que se refere à categoria da práxis, só se pode falar na sua existência com algumas restrições; aliás, as mesmas presentes nos modelos baseados em sistemas dinâmicos não lineares: a interpretação da práxis como mudança qualitativa dada por uma descontinuidade nas saídas provocada por variações contínuas nas entradas, o que acontece quando da convergência do algoritmo de agrupamento k-médias. Contudo, não há verdadeira passagem das mudanças quantitativas para as qualitativas, dado que não há variação do número de neurônios do mapa de k-médias.

## 3.3.2.7 Dialética e Mapas Auto-Organizados de Kohonen

Os *mapas auto-organizáveis de Kohonen* foram originalmente desenvolvidos para fazer agrupamento de dados [33, 70]. Também chamados de redes SOM (*Self-Organizing Maps*) [8], os mapas auto-organizáveis possuem uma arquitetura bastante diferente das redes neurais usuais: neles os neurônios podem ser dispostos matricialmente, tentando emular a distribuição dos neurônios biológicos no cérebro.

Nas redes SOM os neurônios podem ser dispostos de forma linear ou matricial, em todas as dimensões possíveis, sendo que a quantidade de vizinhos influencia apenas na etapa de aprendizado, estabelecendo dependências entre o ajuste dos pesos de um neurônio e os pesos

de seus vizinhos, segundo uma determinada função de vizinhança, como se mostra na descrição do processo de aprendizado.

No processo de aprendizado, áreas específicas de neurônios vão sendo ativadas. Ao final do processo, a rede SOM fica dividida em áreas especialistas, responsáveis pela classificação de padrões específicos, semelhante ao cérebro, onde cada região é responsável por uma atividade específica.

Os mapas auto-organizáveis utilizam neurônios cuja saída é a que segue:

$$y_i = e^{-\|\mathbf{x} - \mathbf{w}_i\|^2}, \quad (3.6)$$

onde  $\mathbf{w}_i = (w_{i,1}, w_{i,2}, \dots, w_{i,n})^T$  é o vetor de pesos do  $i$ -ésimo neurônio, para  $1 \leq i \leq m$  e  $1 \leq j \leq n$ .

Os pesos das redes SOM podem ser ajustados pelo seguinte procedimento modificado, baseado no procedimento clássico de ajuste de redes SOM, mas tendo como critérios de parada o número máximo de iterações e a soma dos ajustes dos pesos:

1. Inicializar os pesos  $w_{i,j}(0)$ , onde  $1 \leq i \leq m$  e  $1 \leq j \leq n$ , com valores aleatórios e necessariamente diferentes.
2. Definir os vetores posição  $\mathbf{r}_i$ , correspondentes à posição do  $i$ -ésimo neurônio na grade, onde  $1 \leq i \leq m$ .
3. Inicializar a taxa de aprendizado inicial  $\eta_0$ , onde  $0 < \eta_0 < 1$ .
4. Inicializar a largura topológica  $\sigma_0$ . A largura topológica é a largura da função de vizinhança gaussiana  $h_{i,k}(t)$ , utilizada para definir a vizinhança do neurônio vencedor a ser treinada.
5. Repetir até  $\delta(t) = 0$  ou um máximo de  $N_{\max}$  iterações:

(a) Para cada padrão  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  do conjunto de entrada  $\Psi = \{\mathbf{x}^{(l)}\}_{l=1}^L$ , repetir:

i. Calcular as saídas  $y_i(t)$  dos neurônios, onde  $1 \leq i \leq m$ .

ii. Calcular  $y_{\max}(t)$ :

$$y_{\max}(t) = \max(y_1(t), y_2(t), \dots, y_m(t)).$$

iii. Calcular o índice  $k$  do neurônio vencedor:

$$y_i = y_{\max} \Rightarrow k = i.$$

iv. Calcular o ajuste  $\Delta w_{i,j}(t)$ :

$$\Delta w_{i,j}(t) = \eta(t) h_{i,k}(t) (x_j(t) - w_{i,j}(t)),$$

onde  $1 \leq i \leq m$  e  $1 \leq j \leq n$ .

v. Atualizar os pesos:

$$w_{i,j}(t+1) = w_{i,j}(t) + \Delta w_{i,j}(t),$$

onde:

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau_1}\right),$$

$$h_{i,k}(t) = \exp\left(-\frac{d_{i,k}^2}{2\sigma^2(t)}\right),$$

$$\eta(t) = \eta_0 \exp\left(-\frac{t}{\tau_2}\right),$$

onde  $\tau_1$  e  $\tau_2$  são *constantes de tempo* e:

$$d_{i,k} = \|\mathbf{r}_i - \mathbf{r}_k\|,$$

para  $1 \leq i \leq m$  e  $1 \leq j \leq n$ .

(b) Calcular  $\delta(t)$  da forma que segue:

$$\delta(t) = \sum_{i=1}^m \sum_{j=1}^n |\Delta w_{i,j}(t)|.$$

As redes SOM podem ser utilizadas também como etapa de classificação ou agrupamento de dados para uma RNA supervisionada [71, 9, 8], tendo diversas aplicações potenciais em reconhecimento de padrões na Engenharia Biomédica como, por exemplo, o auxílio ao diagnóstico da epilepsia a partir de sintomas [72].

É possível interpretar as características e funcionalidades das redes SOM à luz das categorias e princípios do método dialético, tentando entender como seria possível modelar as categorias dialéticas a partir do modelo SOM de Kohonen. As expressões matemáticas do treinamento das redes SOM e o suporte ao treinamento em tempo real já evidenciam a possibilidade de representação da categoria do movimento perpétuo, dado que descrevem a atualização dinâmica dos pesos dos neurônios (pólos).

Caso se considere cada neurônio como um pólo dialético, também é possível mostrar que as redes SOM implementam o princípio da contradição, dada a competição entre os neurônios, o que também configura a presença da luta de pólos. No entanto, a categoria da totalidade só é modelada adequadamente quando são utilizadas funções de vizinhança contínuas para toda a grade de neurônios, como é o caso das funções gaussianas, dado que essas funções modelam a influência do comportamento “coletivo” do sistema dialético (mapa de neurônios) sobre o comportamento “individual” do pólo (neurônio visto isoladamente), evidenciando também a presença do aspecto de interpenetração dos pólos. Além disso, a implementação do princípio da totalidade também é limitada pela arquitetura da rede, que faz com que a interpenetração dos pólos dependa bastante da configuração dos neurônios, e não apenas de seus pesos, como evidencia a ideia de grade de neurônios.

Entretanto, é bastante difícil afirmar que as redes SOM possam servir para implementar a categoria da superação dialética: apesar de surgirem mudanças qualitativas quando da convergência do algoritmo, a partir de mudanças quantitativas nas entradas, e de isso poder ser considerado como uma espécie de práxis, como argumentado em favor dos sistemas dinâmicos não lineares, o fato de não se dar origem a novos neurônios e nem se ter o desaparecimento de neurônios existentes faz com que se possa argumentar que não há de fato a passagem das mudanças quantitativas para as qualitativas. Ou seja: as redes SOM estão sujeitas ao mesmo tipo de crítica feito às propostas baseadas em sistemas dinâmicos não lineares, apesar de se ter uma modelagem explícita dos pólos dialéticos a partir dos neurônios das redes SOM.

### 3.3.2.8 Dialética e Mapas Fuzzy C-Médias

O *mapa fuzzy c-médias* é uma rede neuro-fuzzy não supervisionada projetada para a tarefa de agrupamento de dados [73], tendo sido bastante utilizada em aplicações em classificação de regiões em imagens de RMN [74, 75, 76] e na detecção de regiões ativadas em RMN funcional [77, 78, 79].

Dado um conjunto de dados  $\Psi = \{\mathbf{x}_l\}_{l=1}^L$ , onde  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ , e um universo de  $m$  classes  $\Omega = \{C_1, C_2, \dots, C_m\}$ , a tarefa de agrupar os elementos de  $\Psi$  em  $\Omega$  consiste em minimizar a função objetivo  $J$  definida como segue [74, 78]:

$$J = \sum_{i=1}^m \sum_{l=1}^L \mu_i^q(\mathbf{x}_l) \|\mathbf{x}_l - \mathbf{w}_i\|^2, \quad (3.7)$$

onde  $\mu_i(\mathbf{x}_l)$  é a função de pertinência de  $\mathbf{x}_l$  à classe  $C_i$ ,  $\mathbf{w}_i$  é o vetor de pesos do  $i$ -ésimo neurônio e o inteiro  $q > 1$  é o *expoente de nebulosidade* [74].

Os neurônios possuem a seguinte expressão de saída [74]:

$$y_i = \mu_i(\mathbf{x}) = \left( \sum_{k=1}^m \left( \frac{\|\mathbf{x} - \mathbf{w}_i\|}{\|\mathbf{x} - \mathbf{w}_k\|} \right)^{2/(q-1)} \right)^{-1}, \quad (3.8)$$

onde  $1 \leq i \leq m$ .

Os pesos são ajustados da seguinte forma: se o  $k$ -ésimo neurônio é o neurônio vencedor, ou seja:

$$\mu_k(\mathbf{x}) = \bigvee_{i=1}^m \mu_i(\mathbf{x}), \quad (3.9)$$

então:

$$w_{i,j}(t+1) = \begin{cases} w_{i,j}(t) + \eta(t) \mu_i^q(\mathbf{x}(t)) (x_j(t) - w_{i,j}(t)), & i = k \\ w_{i,j}(t), & i \neq k \end{cases}, \quad (3.10)$$

onde  $k \in \{1, 2, \dots, m\}$ , e  $\eta(t)$  é a taxa de aprendizado e pode ser reduzida a cada iteração de forma semelhante à utilizada no treinamento das redes SOM. O operador  $\bigvee$  é o operador “ou” da Lógica Fuzzy como definida por Takagi e Sugeno:  $c = a \bigvee b = \max\{a, b\}$ ,  $\forall a, b \in [0, 1]$  [80].

Também é possível interpretar as características e funcionalidades dos mapas *fuzzy c-médias* usando as categorias e princípios do método dialético, buscando modelar as categorias dialéticas a partir do modelo *fuzzy c-médias*. As expressões matemáticas de treinamento e o suporte

ao treinamento em tempo real mostram que é possível se falar na presença da categoria do movimento perpétuo.

Considerando cada neurônio um pólo dialético, também é possível mostrar que os mapas *fuzzy* c-médias implementam o princípio da contradição, tendo em vista a competição entre os neurônios, ou seja, a luta de pólos. A categoria da totalidade é modelada adequadamente, dado que a influência do comportamento individual do neurônio (pólo) é afetada pelo comportamento da rede, ao mesmo tempo em que afeta a mesma rede, o que pode ser visto na expressão da saída de um neurônio (expressão 3.8), que leva em conta os pesos dos outros neurônios presentes no mapa.

Contudo, não se pode afirmar que os mapas *fuzzy* c-médias possam servir para implementar a categoria da superação dialética: semelhante aos mapas SOM, apesar de surgirem mudanças qualitativas quando da convergência do algoritmo, a partir de mudanças quantitativas nas entradas, e de isso poder ser considerado como uma espécie de práxis, como argumentado em favor dos sistemas dinâmicos não lineares, o fato de não surgirem novos neurônios e nem se ter o desaparecimento de neurônios existentes faz com que se possa argumentar que não há de fato a passagem das mudanças quantitativas para as qualitativas. Logo, os mapas *fuzzy* c-médias estão sujeitos ao mesmo tipo de crítica feito às propostas baseadas em sistemas dinâmicos não lineares, apesar de se ter uma modelagem explícita dos pólos dialéticos.

### 3.3.2.9 Dialética e Sistemas Dinâmicos Complexos Emergentes

Ao contrário da Teoria da Catástrofe e da Teoria do Caos, não existe um único critério ou modelo que defina a dinâmica complexa. Na verdade não há nem mesmo um critério único para definir o que venha a ser complexidade. Alguns modelos de dinâmica complexa envolvem atratores estranhos não caóticos, fractais ou outros fenômenos complexos para além dos modelos de catástrofe e de caos. Entretanto, muitos modelos incluem esses últimos. Todos esses modelos podem ser encarados como dotados de algum tipo de dinâmica dialética associada à dinâmica caótica, no que se refere a pontos de bifurcação gerando mudanças dinâmicas qualitativas e conflitos entre elementos opostos envolvidos na dinâmica [35].

Muitos desses modelos envolvem sistemas de equações em larga escala e pesadas simulações de fenômenos auto-organizados emergindo da dinâmica de forças em conflito. Essa auto-organização é reconhecida pelos mais rigorosos dialéticos como sendo uma verdadeira implementação da transformação da quantidade em qualidade, dada a capacidade de surgimento de novas variáveis, a eliminação de variáveis antigas e a transformação qualitativa das variáveis existentes, além da variação do número de variáveis, e por conseguinte do número de pólos [35]. No entanto, cabe aqui o comentário que, mesmo com esses sistemas dinâmicos, a modelagem do método dialético ainda não é perfeita na sua essência, dado que os pólos ainda não são explicitamente representados, mesmo que esteja claro como fica a implementação da categoria da contradição, do movimento perpétuo e da práxis.

Um ponto importante para a discussão é que as abordagens dialéticas em geral não são bem aceitas no mundo anglo-saxão, por basicamente três razões: a dialética é vista como não científica, nebulosa e sujeita a manipulação ideológica e política por parte de alguns grupos. A não cientificidade e o caráter nebuloso atribuídos ao método dialético se devem à dificuldade de se estabelecer um formalismo matemático adequado para a modelagem das principais ca-

tegorias e aspectos da dialética, além dos vícios reducionistas, positivistas e habituados a uma versão estreita da lógica aristotélica, das estatísticas comparativas e das linhas de pensamento newtonianas e laplacianas [35].

Os acadêmicos anglo-saxões também alegam a sujeição da dialética a manipulações ideológicas ao fato de parte considerável dos pensadores dialéticos serem marxistas. No continente europeu há maior aceitação do método dialético. Mas quando descolados do método dialético materialista, e portanto de qualquer relação com o marxismo, os trabalhos baseados em outras abordagens dialéticas tendem a ter maior circulação do que aqueles mais próximos à dialética materialista. Além disso, há uma maior aceitação da interpretação dialética do conceito de sistemas auto-organizados de ordem emergente em sistemas dinâmicos complexos da parte dos europeus continentais do que da parte do mundo anglo-saxão [35].

Os modelos baseados em sistemas dinâmicos complexos emergentes, apesar de modelarem os princípios da totalidade, da contradição, do movimento perpétuo e da práxis, ainda não modelam explicitamente os pólos integrantes do sistema, o que, aliado ao fato de ainda estarem em desenvolvimento, carecendo de conceitos mais formais e de consensos para as definições dentro dessa área do conhecimento, torna-os ainda inviáveis como formalismo matemático base para implementação do método dialético, sem contar que a citada complexidade e custo computacional do modelo são impecilhos adicionais [35].

### 3.4 O Método Dialético Objetivo

Tendo em vista as limitações das propostas de modelos apresentadas anteriormente e evidenciadas nas aplicações, fica evidente a necessidade de um modelo mais refinado para modelagem das principais categorias da dialética. Neste trabalho se propõe o *Método Dialético Objetivo* (*Objective Dialectical Method*, ODM), um método de Computação Inteligente baseado em evolução e revolução, para ser adaptado às tarefas de modelagem de sistemas dinâmicos, classificação e reconhecimento de padrões e busca e otimização, sendo potencialmente uma alternativa às redes neurais e aos algoritmos genéticos.

#### 3.4.1 Definição Geral

Em linhas gerais, o ODM é definido da seguinte forma:

1. As entradas do sistema são representadas adequadamente como um *vetor de condições*;
2. Define-se a quantidade inicial de *pólos* integrantes do sistema, juntamente com o *número de fases históricas* e a *duração de cada fase histórica*, podendo tanto o número de fases históricas quanto a duração de cada fase ser aleatórios;
3. A cada pólo é associado um *vetor de pesos*, de mesma dimensionalidade do vetor de condições, uma *função de anticontradição* e uma *medida de força*. Esses vetores de pesos podem ser aleatórios ou escolhidos a partir dos vetores de condições;
4. Iniciam-se as fases históricas. Cada fase histórica está dividida em duas etapas:

**Evolução:** Nesta etapa as condições são apresentadas ao sistema. Os pólos competem entre si, ao mesmo tempo em que transformam quantitativamente uns aos outros, em um processo que recebe o nome de *luta de pólos* e modela o princípio da interpenetração dos pólos. As funções de anticontração de cada pólo são avaliadas e, dado um determinado vetor de condições, o pólo vencedor, ou seja, o que apresentar menor contração com relação à entrada, tem seus pesos ajustados e sua força incrementada. Esse processo continua até o final da fase histórica;

**Crise Revolucionária:** Acontece no final da fase histórica. Na etapa de crise revolucionária são executados os seguintes passos:

- (a) As medidas de força de cada pólo são comparadas. Aqueles pólos cujas medidas de força forem menores do que uma força mínima são marcados;
- (b) Avalia-se a contração entre cada par de pólos. Caso a contração entre dois pólos seja menor do que uma contração mínima, um dos dois pólos é marcado;
- (c) A partir das contrações avaliadas na etapa anterior, encontra-se a maior delas. Essa é a contração principal do sistema e é a partir do par de pólos envolvidos nessa contração (tese e antítese) que pode ser gerado um novo pólo, cujo vetor de pesos é montado a partir dos vetores de pesos dos pólos do par, como forma de permitir elementos do estado antigo nos pólos do novo estado, modelando a existência de elementos do antigo no novo, quando do início da superação dialética. É possível também escolher, ao invés de uma contração principal, contrações principais, e a partir delas gerar mais de um novo pólo. Contudo, a geração de um novo pólo dialético é aleatória;
- (d) Todos os pólos marcados são eliminados e um novo conjunto de pólos integrantes do sistema é gerado;
- (e) Os vetores de pesos dos pólos restantes são alterados aleatoriamente, representando o impacto da crise revolucionária sobre os pólos remanescentes e os novos pólos.

Como se pode perceber da definição anterior, o ODM implementa a superação dialética de forma bastante clara, uma vez que não só os pólos são qualitativamente diferentes após a passagem para uma nova fase histórica, mas também o número de pólos do processo dialético pode aumentar ou diminuir em relação à fase histórica anterior.

### 3.4.2 Proposta de Definição Específica

A partir da definição geral do ODM, na seção anterior, pode-se propor diversas definições específicas para o ODM. Uma proposta de definição que aproveita muito do mapa *fuzzy* c-médias é a seguinte:

Seja  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  um vetor de  $n$  condições que influencia um sistema com  $n_C(t)$  pólos no instante  $t$ , de forma que  $\Omega(t) = \{C_1, C_2, \dots, C_{n_C(t)}\}$  seja o conjunto de pólos integrantes do sistema, cada pólo  $C_i$  está associado a uma função de anticontração  $g_i$  e a um vetor de pesos

$\mathbf{w}_i$  definidos como segue:

$$g_i(\mathbf{x}) = \left( \sum_{l=1}^{n_C(t)} \left( \frac{\|\mathbf{x} - \mathbf{w}_i(t)\|}{\|\mathbf{x} - \mathbf{w}_l(t)\|} \right)^{2/(q-1)} \right)^{-1} \quad (3.11)$$

$$\mathbf{w}_i = (w_{i,1}, w_{i,2}, \dots, w_{i,n})^T, \quad (3.12)$$

para  $1 \leq i \leq n_C(t)$  e  $q > 1$ .

Sejam  $n_P$  o número máximo de fases históricas,  $n_H(t)$  a duração de uma fase histórica, e  $\eta(t)$  o passo, onde  $0 < \eta(t) < 1$ , o algoritmo segue até que o número de pólos atinja um certo número desejado, ou que se atinja o número máximo de fases históricas, ou mesmo que os vetores de pesos dos pólos ao final de uma determinada fase histórica minimizem uma determinada função de custo.

A etapa de *evolução* ou de *luta de pólos* pode ser implementada da seguinte forma:

$$w_{i,j}(t+1) = \begin{cases} w_{i,j}(t) + \eta(t) g_i^q(\mathbf{x})(x_j(t) - w_{i,j}(t)), & i = k(t) \\ w_{i,j}(t), & i \neq k(t) \end{cases}, \quad (3.13)$$

$$f_i(t+1) = \begin{cases} f_i(t) + 1, & i = k(t) \\ f_i(t), & i \neq k(t) \end{cases}, \quad (3.14)$$

para  $1 \leq i \leq n_C(t)$  e  $1 \leq j \leq n$ , onde  $f_i$  é a medida de força associada ao pólo  $C_i$  e  $k$  é o índice do pólo vencedor, dado que

$$g_k(\mathbf{x}) = \max(g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_{n_C}(\mathbf{x})). \quad (3.15)$$

Na etapa de *crise revolucionária* define-se um *vetor binário de marcações* da forma que segue:

$$\mathbf{m} = (m_1, m_2, \dots, m_{n_C(t)})^T,$$

onde  $m_i = 1$  quando o pólo  $C_i$  deve ser absorvido por outro pólo ou eliminado, e  $m_i = 0$  caso contrário, para  $1 \leq i \leq n_C(t)$ . Esse vetor é inicialmente nulo. É importante ressaltar que esse vetor está representado apenas para fins didáticos, pois pode não ser computacionalmente eficiente utilizá-lo, já que é possível eliminar os pólos marcadas sem precisar definir um vetor de marcação.

Na etapa de crise revolucionária, primeiro são avaliadas as medidas de força de cada pólo. Para tanto, as medidas de força são normalizadas da forma que segue:

$$\bar{f}_i(t) = \frac{f_i(t)}{\max\{f_l(t)\}_{l=1}^{n_C(t)}}, \quad (3.16)$$

para  $1 \leq i \leq n_C(t)$ . Caso a medida de força normalizada seja menor que a *medida de força mínima* para a sobrevivência de um pólo após o processo de luta de pólos,  $f_{\min}$ , ou seja,  $\bar{f}_i(t) < f_{\min}$ , tem-se  $m_i = 1$ .

Após a avaliação das forças de cada pólo, são avaliadas as contradições. A contradição do pólo  $C_i$  em relação ao pólo  $C_j$ , representada por  $\delta_{i,j}$ , é determinada por:

$$\delta_{i,j} = 1 - g_i(\mathbf{w}_j), \quad (3.17)$$

onde  $1 \leq i, j \leq n_C(t)$ . Caso a contradição seja menor que a *contradição mínima*,  $\delta_{\min}$ , ou seja,  $\delta_{i,j} < \delta_{\min}$ , considera-se que na verdade  $C_i$  e  $C_j$  são efetivamente o mesmo pólo, e uma das duas é marcada. Caso se escolha  $C_i$ , faz-se  $m_i = 1$ .

Após a fase de avaliação das contradições, ainda na etapa de crise revolucionária, gera-se o *conjunto parcial de pólos integrantes do novo sistema*,  $\Omega'(t)$ , da seguinte maneira:

$$m_i(t) = 0 \Leftrightarrow C_i \in \Omega'(t), \quad (3.18)$$

para  $1 \leq i \leq n_C(t)$ .

A busca das *contradições principais* é feita a partir da ordenação das contradições  $\delta_{i,j}$ , desconsiderando as contradições  $\delta_{j,i}$ , para  $i \neq j$  e  $m_i = m_j = 0$ . As contradições principais serão as  $p \geq 1$  maiores do conjunto ordenado. A partir delas são gerados os novos pólos  $C_k$ , onde  $k > n_C(t)$ , desde a contradição que envolve o respectivo par de pólos  $C_i$  e  $C_j$  seja maior do que uma contradição máxima, ou seja,  $\delta_{i,j} \geq \delta_{\max}$ , onde  $\delta_{\max}$  é a contradição máxima para geração de síntese, estando relacionada ao conceito de práxis, quando este implica não apenas na mudança qualitativa dos pólos integrantes do processo dialético, mas quando há geração de um novo pólo. Assim, através desse procedimento, modela-se o conceito dialético de que as maiores contradições entre pólos tendem a se resolver gerando sínteses dos pares tese-antítese geradores; essas sínteses, por sua vez, dão origem a novos pólos.

Os vetores de pesos associados aos novos pólos  $C_u$  e  $C_v$ , especificamente,  $\mathbf{w}_u$  e  $\mathbf{w}_v$ , podem ser calculados combinando-se os pesos dos pólos  $C_p$  e  $C_q$  envolvidos na contradição, de forma semelhante à dos operadores de cruzamento em algoritmos genéticos, como segue:

$$w_{u,r}(t+1) = \begin{cases} w_{p,r}(t+1), & r \bmod 2 = 1 \\ w_{q,r}(t+1), & r \bmod 2 = 0 \end{cases}, \quad (3.19)$$

$$w_{v,r}(t+1) = \begin{cases} w_{p,r}(t+1), & r \bmod 2 = 0 \\ w_{q,r}(t+1), & r \bmod 2 = 1 \end{cases}, \quad (3.20)$$

onde  $1 \leq r \leq n$ , enquanto  $\mathbf{w}_p$  e  $\mathbf{w}_q$  são os vetores de pesos dos pólos  $C_p$  e  $C_q$ , respectivamente, componentes da contradição principal.

Os novos pólos assim gerados,  $C_u$  e  $C_v$ , são parte do *conjunto de novos pólos*,  $\Omega''(t)$ . Assim, o *conjunto de pólos remanescentes*,  $\Omega'''(t)$ , é obtido da seguinte forma:

$$\Omega'''(t) = \Omega'(t) \cup \Omega''(t). \quad (3.21)$$

A influência qualitativa da crise revolucionária sobre os pólos é modelada pela *função de crise*,  $\chi(t)$ , que pode ser como a função seguinte:

$$\chi(t) = \chi_{\max}(t)G(0,1), \quad (3.22)$$

onde  $\chi(t)$  é a *função de crise*,  $\chi_{\max}(t)$  é a *máxima crise* e  $G(0,1)$  é uma função que retorna um número aleatório de distribuição gaussiana com esperança 0 e variância 1. O conjunto dos novos pólos integrantes do sistema é,

$$\Omega(t+1) = \{C_1(t+1), C_2(t+1), \dots, C_{n_C(t+1)}(t+1)\}, \quad (3.23)$$

onde

$$C_i(t+1) = C_i(t) \in \Omega'''(t), \quad (3.24)$$

e

$$w_{i,j}(t+1) = w_{i,j}(t) + \chi(t), \quad (3.25)$$

para  $1 \leq i \leq n_C(t+1)$  e  $1 \leq j \leq n$ . A etapa de crise revolucionária então se encerra e, com ela, a fase histórica. O sistema passa então para uma nova fase histórica.

### 3.4.3 Classificadores Dialéticos Objetivos

Os classificadores dialéticos objetivos são baseados no método dialético objetivo. Seu funcionamento é relativamente simples: os vetores de atributos são montados e considerados vetores de condições que, ao serem apresentados ao sistema dialético, condicionarão a dinâmica das contradições entre os pólos integrantes que, por sua vez, modelam os pólos reconhecidos na tarefa de classificação não supervisionada.

Assim, um classificador dialético objetivo (*Objective Dialectical Classifier*, ODC) é na verdade um classificador não supervisionado adaptável onde, ao invés de se pressupor um determinado número de classes (pólos do sistema dialético inicial), pode-se optar por estabelecer uma quantidade de classes inicial e, ao longo das fases históricas, como resultado da luta de pólos e das crises revolucionárias, classes (pólos) são eliminadas ou absorvidas por outras, podendo-se, ao final do processo de treinamento, obter o número de classes estatisticamente significativo presente no conjunto de treinamento, e um classificador adequado, associado ao estado final do sistema dialético.

Para acelerar a convergência do classificador dialético, pode-se eliminar o operador de geração de novos pólos, presente na etapa de crise revolucionária. Contudo, isso pode viciar o método, dado que o operador de geração de novos pólos é um tipo de operador aleatório gerador de diversidade, quando se compara o classificador dialético com classificadores evolutivos. No entanto, isso pode ser compensado de certa forma pela alteração dos pesos dos pólos efetuada pela função de crise. O importante a ressaltar é que, dada a quantidade de parâmetros iniciais, é de se esperar que os classificadores dialéticos objetivos dependam bastante dos parâmetros iniciais.

As seções seguintes apresentam uma proposta de definição específica para os procedimentos de treinamento e classificação.

#### 3.4.3.1 Treinamento

O treinamento do classificador dialético objetivo pode ser realizado segundo o algoritmo que segue:

1. Definir o número de fases históricas  $n_P$ .
2. Definir a duração de cada fase histórica  $n_H$ .
3. Definir o número de pólos de parada  $n_{C,f}$ .
4. Definir a força mínima,  $0 \leq f_{\min} \leq 1$ .

5. Definir a contradição mínima,  $\delta_{\min}$ .
6. Definir a contradição máxima,  $\delta_{\max}$ .
7. Definir a máxima crise,  $0 \leq \chi_{\max} \leq 1$ .
8. Inicializar o passo de cada fase histórica,  $0 < \eta(0) < 1$ .
9. Definir o número inicial de pólos  $\#\Omega(0) = n_C(0)$ , o que define também o conjunto de pólos inicial:

$$\Omega(0) = \{C_1(0), C_2(0), \dots, C_{n_C(0)}(0)\}.$$

10. Inicializar os pesos  $w_{i,j}(0)$ , onde  $1 \leq i \leq n_C(0)$  e  $1 \leq j \leq n$ .
11. Seja  $\#\Omega(t)$  a cardinalidade de  $\Omega(t)$ , repetir até  $n_P$  iterações ou  $\#\Omega(t) = n_{C,f}$ :

(a) Repetir até  $n_H$  iterações:

- i. Inicializar as medidas de força  $f_i = 0$ , para  $1 \leq i \leq n_C(t)$ .
- ii. Para cada vetor de condições  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  do conjunto de entrada  $\Psi = \{\mathbf{x}^{(l)}\}_{l=1}^L$ , repetir:
  - A. Computar o valor das funções de anticontração  $g_i(\mathbf{x})$ , onde  $1 \leq i \leq n_C(t)$ .
  - B. Calcular  $g_{\max}$ :

$$g_{\max} = \max(g_1(\mathbf{x}), g_2(t)(\mathbf{x}), \dots, g_{n_C(t)}(\mathbf{x})).$$

C. Calcular o índice  $k(t)$  do pólo vencedor:

$$g_i = g_{\max} \Rightarrow k(t) = i.$$

D. Atualizar os pesos do pólo vencedor:

$$w_{i,j}(t+1) = \begin{cases} w_{i,j}(t) + \eta(t)g_i^q(\mathbf{x})(x_j(t) - w_{i,j}(t)), & i = k \\ w_{i,j}(t), & i \neq k \end{cases},$$

onde  $q > 1$ .

E. Atualizar as medidas de força dos pólos:

$$f_i(t+1) = \begin{cases} f_i(t) + 1, & i = k(t) \\ f_i(t), & i \neq k(t) \end{cases}.$$

iii. Mudança quantitativa:  $\Omega(t+1) = \Omega(t)$ .

(b) Calcular as medidas de força normalizadas:

$$\tilde{f}_i(t) = \frac{f_i(t)}{\max\{f_l(t)\}_{l=1}^{n_C(t)}},$$

para  $1 \leq i \leq n_C(t)$ .

(c) Calcular as contradições:

$$\delta_{i,j} = 1 - g_i(\mathbf{w}_j),$$

onde  $2 \leq j \leq n_C(t)$  e  $1 \leq i < j$ .

(d) Mudança qualitativa: computar o novo conjunto de pólos,  $\Omega(t+1)$ , caso  $\delta_{i,j} \geq \delta_{\max}$ :

$$\bar{f}_i(t) > f_{\min} \Rightarrow C_i(t) \in \Omega(t+1),$$

onde  $1 \leq i \leq n_C(t)$  e

$$\delta_{i,j} \geq \delta_{\min} \Rightarrow C_i(t), C_j(t) \in \Omega(t+1),$$

$$\delta_{i,j} < \delta_{\min} \Rightarrow C_i(t) \in \Omega(t+1),$$

$$\delta_{i,j} \geq \delta_{\max} \Rightarrow C_u(t), C_v(t) \in \Omega(t+1),$$

onde  $2 \leq j \leq n_C(t)$ ,  $1 \leq i < j$ ,  $u = n_C(t) + 1$ ,  $v = n_C(t) + 2$ , e

$$w_{u,r}(t+1) = \begin{cases} w_{i,r}(t+1), & r \bmod 2 = 1 \\ w_{j,r}(t+1), & r \bmod 2 = 0 \end{cases},$$

$$w_{v,r}(t+1) = \begin{cases} w_{i,r}(t+1), & r \bmod 2 = 0 \\ w_{j,r}(t+1), & r \bmod 2 = 1 \end{cases},$$

onde  $1 \leq r \leq n$ .

(e) Adicionar o efeito de crise aos pesos dos novos pólos integrantes do sistema:

$$w_{i,j}(t+2) = w_{i,j}(t+1) + \chi_{\max} G(0, 1),$$

para  $1 \leq i \leq n_C(t+1)$ ,  $1 \leq j \leq n$  e  $\Omega(t+2) = \Omega(t+1)$ .

### 3.4.3.2 Classificação

Uma vez treinado, o classificador dialético objetivo se comporta como um método de classificação não supervisionada. Da mesma forma, isso também fica evidente no treinamento do classificador, quando se faz  $n_P = n_H = 1$ , o que transforma o classificador ODC em um mapa de k-médias, por exemplo.

A classificação é feita portanto da seguinte forma: dado um conjunto de condições de entrada

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T,$$

se o sistema se estabilizou com  $\Omega = \{C_1, C_2, \dots, C_{n_C}\}$ , então se calcula:

$$g_{\max} = \max(g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_{n_C}(\mathbf{x})).$$

A regra de classificação portanto é:

$$g_k(\mathbf{x}) = g_{\max} \Rightarrow \mathbf{x} \in C_k,$$

onde  $1 \leq k \leq n_C$ .

### 3.5 Conclusão

O método dialético pode ser utilizado no estudo de diversos fenômenos e processos, sejam eles sociais, econômicos ou naturais, por exemplo, desde que estes fenômenos ou processos possam ser modelados por meio de uma união dialética e se tenha como conhecer seus pólos integrantes principais e suas relações mútuas. No entanto, para aplicar o método filosófico investigativo da dialética em áreas ligadas às Ciências Exatas, como a Economia, as Engenharias e a Ciência da Computação, também é necessária uma formulação matemática adequada, que possa instanciar as principais categorias do método dialético, sem se preocupar com a modelagem matemática de todos os aspectos da dialética, que podem não ser de interesse para a maioria das aplicações nas áreas do conhecimento citadas. Nesse sentido, o método dialético objetivo foi projetado para ser uma formulação matemática e algorítmica para o método dialético materialista.

O método dialético objetivo como ferramenta matemática para uma teoria crítica aproveita o caráter filosófico investigativo da dialética na construção de uma ferramenta para modelagem comportamental de sistemas e reconhecimento de padrões, apresentando uma proposta de implementação do método dialético objetivo muito semelhante a sistemas híbridos baseados em redes neurais artificiais.

O método dialético objetivo é uma tentativa de formular o método dialético da Filosofia da Práxis em uma forma algorítmica e consistente, com um formalismo próprio, evitando utilizar abordagens baseadas em heurísticas e sistemas dinâmicos modelados por equações diferenciais não lineares, dedicada a tarefas de modelagem e classificação. O classificador dialético objetivo é uma versão do método dialético objetivo para classificação não supervisionada e reconhecimento de padrões. A partir da análise dos procedimentos de treinamento e classificação do classificador dialético objetivo pode-se perceber que, quando se limita o problema a apenas uma única fase histórica, o método se reduz a um classificador não supervisionado baseado em agrupamento, de certa forma semelhante a métodos como os mapas auto-organizados de Kohonen, mapas de *k*-médias e *fuzzy c*-médias, o que é de se esperar, dado que em uma única fase histórica não há superação dialética, não há práxis. Aliás, o próprio conceito de fase histórica é dependente da práxis.

Assim, o classificador dialético objetivo pode ser utilizado em problemas onde o número de classes estatisticamente significativas não é bem conhecido, dada a sua capacidade de gerar novas classes e de eliminar classes existentes ao longo do processo de treinamento não supervisionado, ou ainda em problemas onde se deseja encontrar um mapa de agrupamento ótimo para uma determinada tarefa de classificação. Contudo, a tarefa de encontrar um mapa de agrupamento ótimo é empírica, pois é preciso analisar o comportamento do treinamento em função dos diversos parâmetros do método, como a força mínima, a contradição máxima e a contradição mínima, o número de pólos integrantes inicial, o número de fases históricas e a duração de cada fase histórica, a probabilidade de geração de novos pólos, bem como o passo de cada fase histórica.



# O Método Dialético adaptado a Busca e Otimização

## 4.1 Introdução

A concepção dialética da realidade é um tipo de método filosófico investigativo para analisar processos presentes na natureza e em sociedades humanas. Suas origens estão conectadas às filosofias das antigas civilizações da Grécia, da China e da Índia e estão fortemente relacionadas ao pensamento de Heráclito, Platão, e às filosofias de religiões como o Confucionismo, o Budismo e a Escola Zen [10].

Como um método geral de análise, a dialética tem experimentado um progresso considerável devido ao desenvolvimento da filosofia alemã do século XIX, com a dialética de Hegel, e do século XX, com os trabalhos de Marx, Engels, Gramsci e Lukács. Todos esses filósofos produziram trabalhos fundamentais a respeito da dinâmica das contradições na natureza e nas sociedades baseadas em classes sociais, dando origem a o que se convencionou chamar de Materialismo Histórico [27, 28, 29, 30, 31, 32].

O método dialético do Materialismo Histórico é uma ferramenta para o estudo de sistemas que leva em conta a dinâmica de suas contradições como processos dinâmicos com fases intercaladas de *evolução* e *crise revolucionária*. Isso tem inspirado a concepção de um método computacional inteligente evolutivo para classificação que seja capaz de resolver problemas comumente resolvidos por abordagens baseadas em redes neurais e algoritmos genéticos.

A maioria dos paradigmas de Inteligência Computacional, particularmente redes neurais, computação evolucionária e algoritmos de inspiração sócio-cultural (comportamental), são baseados em algum tipo de teoria voltada a aplicações gerais, embora na prática sejam teorias bastante incompletas; como ilustração, têm-se os seguintes fatos: as abordagens baseadas em redes neurais se apóiam sobre determinados modelos do cérebro [33]; já a computação evolucionária é baseada na Teoria da Evolução de Darwin e em seus desenvolvimentos posteriores, dentre eles a Genética de Mendel [33, 81]; por sua vez, algoritmos de inspiração sócio-cultural são baseados no estudo de populações, tais como os algoritmos baseados em colônias de formigas e em otimização por enxame de partículas (*Particle Swarm Optimization*, PSO) [81].

O Método Dialético Objetivo (*Objective Dialectical Method*, ODM) é um método de Inteligência Computacional baseado na dialética materialista para as tarefas de classificação supervisionada, classificação não supervisionada, busca e otimização; sua versão para classificação não supervisionada tem sido empregada com sucesso na classificação de imagens multiespectrais sintéticas de ressonância magnética anatômica, como ferramenta para auxiliar a análise do especialista [82, 83]. Neste capítulo se busca apresentar o ODM como uma proposta de algoritmo dialético para busca e otimização, aproveitando as formalizações da Computação Evolucionária e da Lógica Fuzzy.



**Figura 4.1** Claude Elwood Shannon, engenheiro eletrônico, matemático e pai da Teoria da Informação (1916-2001)

Este capítulo também busca mostrar que é possível usar o Princípio da Máxima Entropia para se obter uma versão modificada do algoritmo através da obtenção de novas expressões para as funções de pertinência, uma vez que a maximização da entropia assim definida tanto garante a convergência do algoritmo quanto garante a máxima exploração da vizinhança do ponto ótimo. Ao final, as expressões canônicas do método dialético objetivo para busca e otimização são comparadas às expressões obtidas pela maximização da entropia, sendo tecidos comentários a respeito de suas semelhanças e diferenças.

## 4.2 Princípio da Máxima Entropia

A Teoria da Informação tem suas origens no clássico artigo de Claude Elwood Shannon (ver figura 4.1) de 1948 que, junto com desenvolvimentos posteriores realizados por outros pesquisadores, serviu de resposta direta à necessidade da Engenharia Elétrica de projetar sistemas de comunicações que fossem ao mesmo tempo eficientes e confiáveis [84, 33]. Para tanto, a Teoria da Informação destaca a necessidade de que os projetos de sistemas sejam construídos sobre as bases de modelos teóricos da informação. Dentre esses modelos, os mais utilizados são o Princípio da Máxima Informação Mútua e o Princípio da Máxima Entropia [33]. Nesta seção será abordado apenas o segundo modelo, dada a sua aplicabilidade à proposta de análise dinâmica e projeto de um método de busca e otimização baseado na dialética materialista.

Considerando uma variável aleatória discreta  $X$  modelada da forma que segue:

$$X = \{x_k : k = 0, \pm 1, \pm 2, \dots, \pm(K-1), \pm K\}, \quad (4.1)$$

onde  $x_k$  é um número discreto e  $2K + 1$  é a quantidade de níveis discretos, assumindo que o

evento  $X = x_k$  ocorre com probabilidade

$$p_k = P(X = x_k), \quad (4.2)$$

desde que

$$\sum_{k=-K}^K p_k = 1, \quad (4.3)$$

para  $0 \leq p_k \leq 1$ , a *quantidade de informação*  $I(x_k)$  associada ao evento  $X = x_k$  é definida como [33]:

$$I(x_k) = \log \left( \frac{1}{p_k} \right) = -\log p_k, \quad (4.4)$$

onde a base do logaritmo é arbitrária. No contexto da Teoria da Informação, a quantidade de informação está associada à incerteza sobre um evento. Assim, de um ponto de vista intuitivo, quanto maior a “surpresa”, ou seja, a “novidade” associada a um evento, menor a probabilidade de ocorrer o referido evento e, portanto, maior a quantidade de informação associada, dado que  $\lim_{p_k \rightarrow 0} I(x_k) = +\infty$ . Essa definição também é consistente do ponto de vista etimológico quando se tem em mente a maioria das palavras utilizadas nas línguas ocidentais como sinônimo de informação, tais como *novas* e *novidades* (em português), *news* (inglês) e *euangelion* (grego clássico, de onde vem a palavra *evangelho*, que originalmente tinha o sentido de novidades da frente de guerra), sendo essa característica parte do poder de generalização da Teoria da Informação como definida por Shannon.

Uma vez que a quantidade de informação  $I(x_k)$  é uma variável aleatória discreta com probabilidade associada  $p_k$ , o valor médio estatístico de  $I(x_k)$  sobre todo o intervalo de  $2K + 1$  níveis discretos pode ser encarado como a seguinte estatística de  $X$  [33]:

$$H(X) = \langle I(x_k) \rangle = \sum_{k=-K}^K p_k I(x_k) = - \sum_{k=-K}^K p_k \log p_k, \quad (4.5)$$

onde a estatística  $H(X)$  é chamada de *entropia* da variável aleatória  $X$ , pela analogia de sua expressão algébrica à expressão da entropia como definida pela Termodinâmica Estatística [33]. Assim, a entropia é na verdade a medida na quantidade de informação média, ou quantidade esperada de informação, associada ao sistema.

Caso  $X$  seja uma variável aleatória contínua com função densidade de probabilidade  $f_X$ , utiliza-se a *entropia diferencial*, definida como [33]:

$$h(X) = - \int_{-\infty}^{+\infty} f_X(x) \log f_X(x) dx. \quad (4.6)$$

Supondo um sistema estocástico onde a distribuição de probabilidade dos estados possíveis de serem assumidos pelo sistema seja desconhecida, sabendo-se que algumas restrições sobre os valores dos estados possíveis são conhecidas, é necessário escolher um modelo probabilístico que seja ótimo em algum sentido. O *Princípio da Máxima Entropia* tenta resolver esse problema afirmando que, quando uma inferência é feita com base em informação incompleta,

essa inferência deve ser realizada a partir da distribuição de probabilidade que maximiza a entropia, desde que sujeita às restrições sobre a distribuição [85, 86, 87, 33].

Algebricamente, considerando um sistema com  $n$  estados possíveis, ou seja,

$$X = \{x_1, x_2, \dots, x_n\},$$

o Princípio da Máxima Entropia consiste na resolução de um problema de otimização restrito [85], onde se deseja otimizar a expressão:

$$H = - \sum_{i=1}^n p_i \ln p_i, \quad (4.7)$$

para  $p_i = P(X = x_i)$ , de acordo com as seguintes restrições:

$$\sum_{i=1}^n p_i = 1, \quad (4.8)$$

$$\langle g(X) \rangle = \sum_{i=1}^n p_i g(x_i), \quad (4.9)$$

onde  $\langle g(X) \rangle$  é o valor esperado da função  $g(X)$ , sendo este, portanto, um valor conhecido. Esse problema de maximização pode ser resolvido de acordo com o Método dos Multiplicadores de Lagrange, admitindo que:

$$p_i = \exp[-\lambda - \mu g(x_i)], \quad (4.10)$$

onde  $\lambda$  e  $\mu$  são parâmetros de Lagrange para o caso unidimensional, e substituindo nas expressões 4.8 e 4.9, resultando nas seguintes expressões:

$$\langle g(X) \rangle = - \frac{\partial}{\partial \mu} \ln Z(\mu), \quad (4.11)$$

$$\lambda = \ln Z(\mu), \quad (4.12)$$

onde

$$Z(\mu) = \sum_{i=1}^n \exp[-\mu g(x_i)] \quad (4.13)$$

é chamada de *função de partição* [85].

O resultado pode ser generalizado para um conjunto de  $m$  funções  $g_k(X)$ , dados os valores esperados:

$$\langle g_k(X) \rangle = \sum_{i=1}^n p_i g_k(x_i), \quad (4.14)$$

para  $1 \leq k \leq m$ , e a função de partição [85]:

$$Z(\lambda_1, \lambda_2, \dots, \lambda_m) = \sum_{i=1}^n \exp[- \sum_{k=1}^m \lambda_k g_k(x_i)], \quad (4.15)$$

dando como resultado as distribuições [85]:

$$p_i = \exp[-\lambda_0 - \sum_{k=1}^m \lambda_k g_k(x_i)]. \quad (4.16)$$

Portanto, para o caso de múltiplas funções de caracterização com valores esperados conhecidos, a máxima entropia é dada por [85]:

$$H_{\max} = \lambda_0 + \sum_{k=1}^m \lambda_k \langle g_k(X) \rangle, \quad (4.17)$$

onde as constantes  $\lambda_k$  são os *multiplicadores de Lagrange* [85, 33], obtidos pela resolução numérica do sistema composto pelas seguintes equações:

$$\lambda_0 = \ln Z(\lambda_1, \lambda_2, \dots, \lambda_m), \quad (4.18)$$

$$\langle g_k(X) \rangle = -\frac{\partial}{\partial \lambda_k} \ln Z(\lambda_1, \lambda_2, \dots, \lambda_m), \quad (4.19)$$

para  $k = 1, 2, \dots, m$ . Substituindo 4.18 e 4.15 em 4.16, tem-se:

$$p_i = \exp[-\lambda_0] \exp[-\sum_{k=1}^m \lambda_k g_k(x_i)] = \frac{\exp[-\sum_{k=1}^m \lambda_k g_k(x_i)]}{Z(\lambda_1, \lambda_2, \dots, \lambda_m)},$$

o que resulta na distribuição de Gibbs [85]:

$$p_i = \frac{\exp[-\sum_{k=1}^m \lambda_k g_k(x_i)]}{\sum_{j=1}^n \exp[-\sum_{k=1}^m \lambda_k g_k(x_j)]}, \quad (4.20)$$

para  $k = 1, 2, \dots, m$ .

As variâncias das funções  $g_k(X)$  são dadas por [85]:

$$\Delta^2 g_k(X) = \langle g_k^2(X) \rangle - \langle g_k(X) \rangle^2 = \frac{\partial^2}{\partial \lambda_k^2} \ln Z(\lambda_1, \lambda_2, \dots, \lambda_m), \quad (4.21)$$

para  $k = 1, 2, \dots, m$ .

Assim, para uma única função ( $m = 1$ ), a expressão 4.20 se torna:

$$p_i = \frac{\exp[-\lambda g(x_i)]}{\sum_{j=1}^n \exp[-\lambda g(x_j)]}. \quad (4.22)$$

### 4.3 Entropia Fuzzy

A necessidade de representar conhecimento disponível de variáveis cujos valores são desconhecidos tem se tornado uma necessidade crescente em sistemas de informação [88]. Existem diversos tipos de medida de incerteza. O primeiro e mais conhecido deles é a medida de probabilidade. Entretanto, outras medidas de incerteza são possíveis utilizando medidas *fuzzy*, tais

como incertezas probabilísticas e incertezas possibilísticas [88]. Assim, é possível construir diversas formas de medir a incerteza e a informação associadas a uma variável, o que torna possível construir diversas formas de medida de entropia, a depender da medida de incerteza utilizada [88, 89, 90, 91].

Seja  $V$  uma variável que pode assumir os  $n$  estados possíveis:

$$X = \{x_1, x_2, \dots, x_n\},$$

onde  $X$  é o conjunto de estados, uma medida *fuzzy*  $\mu$  é um mapeamento de um subconjunto de  $X$  no intervalo  $[0, 1]$ , ou seja,  $\mu : 2^X \rightarrow [0, 1]$ , obedecendo às seguintes propriedades [88]:

1.  $\mu(X) = 1$ ;
2.  $\mu(\emptyset) = 0$ ;
3.  $A \subseteq B \Rightarrow \mu(A) \leq \mu(B), \forall A, B \subseteq 2^X$ .

No caso de se ter a incerteza expressa na forma probabilística, o conhecimento a respeito de uma variável desconhecida pode ser representado por uma medida *fuzzy*  $\mu$  onde  $\mu(\{x_i\}) = p_i$  e conseqüentemente [88]:

$$\mu(E) = \sum_{x_i \in E} \mu(\{x_i\}), \quad (4.23)$$

onde  $\mu(E)$  nesta situação acaba sendo uma medida de probabilidade de  $E$ . Uma vez que  $\mu(X) = 1$ , tem-se:

$$\mu(X) = \sum_{x_i \in X} \mu(\{x_i\}) = \sum_{i=1}^n p_i = 1. \quad (4.24)$$

Uma definição possibilística da incerteza do evento  $E$  é a que segue:

$$\mu(E) = \max_{x_i \in E} \mu(\{x_i\}). \quad (4.25)$$

Em diversas aplicações da Teoria da Decisão e da Teoria dos Jogos, medidas não aditivas são necessárias para expressar incerteza [88]. Combarro e Miranda (2008) afirmam que as medidas *fuzzy* são medidas não aditivas por excelência, dado que obedecem às três condições citadas nos parágrafos anteriores, ou seja,  $\mu(X) = 1$ ,  $\mu(\emptyset) = 0$  e  $A \subseteq B \Rightarrow \mu(A) \leq \mu(B), \forall A, B \subseteq 2^X$ , e não respeitam a aditividade, uma vez que, caso  $A \cap B = \emptyset$ , tem-se  $\mu(A \cup B) \neq \mu(A) + \mu(B)$  [92]. Na verdade,  $\mu(A \cup B) = \max\{\mu(A), \mu(B)\}$ .

Assim, para definir uma medida *fuzzy*, uma vez que as pertinências *fuzzy* são medidas não aditivas, caso se queira medir a quantidade de informação média, é preciso incluir também a informação complementar [89]. Por conseguinte, Fan e Ma (2002) apresentam diversas entropias *fuzzy* possíveis, definidas pela inclusão da informação complementar, como é o caso da seguinte entropia obtida a partir da generalização de um sistema com dois estados [89]:

$$H(A) = - \sum_{i=1}^n [\mu_A(x_i) \ln \mu_A(x_i) + (1 - \mu_A(x_i)) \ln(1 - \mu_A(x_i))], \quad (4.26)$$

onde  $\mu_A(x_i)$  é o grau de pertinência de  $x_i$  ao conjunto  $A$ . Diversas fórmulas de entropia são apresentadas por Fan e Ma (2002), mas todas envolvem de alguma maneira os graus de pertinência complementares  $\mu_A(x_i)$  e  $1 - \mu_A(x_i)$ , mas a apresentação de todas as fórmulas de entropia *fuzzy* foge do escopo deste trabalho [89]. Contudo, a tentativa de se aplicar o Princípio da Máxima Entropia usando as entropias apresentadas por Fan e Ma (2002) resulta em equações do segundo grau que podem não ter solução real, o que dificulta a obtenção de expressões analíticas para as funções de pertinência.

Uma vez que a solução das equações envolvidas na aplicação do Princípio da Máxima Entropia usando a Entropia de Shannon já é algo bem estabelecido dentro da Teoria da Informação, Yager (2000) propõe uma fórmula de entropia *fuzzy* similar à Entropia de Shannon, definida para a medida *fuzzy*  $\mu$  da forma que segue [88]:

$$H(\mu) = - \sum_{i=1}^n V_i \ln V_i, \quad (4.27)$$

onde  $V_i$  é o índice de Shapley, uma medida de incerteza que inclui os valores de pertinência  $\mu$  e seus complementos, definido por [88]:

$$V_i = \sum_{k=0}^{n-1} \gamma_k \sum_{E \subseteq F_i, |E|=k} [\mu(E \cup \{x_i\}) - \mu(E)], \quad (4.28)$$

onde  $F_i = X - \{x_i\}$ ,  $|E|$  é a cardinalidade do conjunto  $E$ , e

$$\gamma_k = \frac{k!(n-k-1)!}{n!}, \quad (4.29)$$

onde  $V_i \in [0, 1]$  e  $\sum_{i=1}^n V_i = 1$ .

Em problemas de agrupamento (*clustering*), as variáveis envolvidas assumem apenas um único estado, dado que agora não faz sentido afirmar que um objeto pertence igualmente a dois grupos (*clusters*) ao mesmo tempo, por exemplo; assim, as medidas de incerteza *fuzzy* passam a ser medidas probabilísticas. Logo, segundo Yager (2000), tem-se [88]:

$$\mu(E \cup \{x_i\}) - \mu(E) = \mu(\{x_i\}) = p_i,$$

o que resulta

$$V_i = \sum_{k=0}^{n-1} \gamma_k \sum_{E \subseteq F_i, |E|=k} p_i = p_i \sum_{k=0}^{n-1} \gamma_k \sum_{E \subseteq F_i, |E|=k} 1.$$

Contudo, para qualquer valor de  $k$ , existem  $\frac{(n-1)!}{k!(n-k-1)!}$  subconjuntos  $E$  tais que  $E \subseteq X - \{x_i\}$  e  $|E| = k$ , o que resulta [88]:

$$\sum_{E \subseteq F_i, |E|=k} 1 = \frac{(n-1)!}{k!(n-k-1)!},$$

de onde vem que

$$V_i = p_i \sum_{k=0}^{n-1} \gamma_k \frac{(n-1)!}{k!(n-k-1)!}.$$

Mas como

$$\gamma_k = \frac{k!(n-k-1)!}{n!},$$

tem-se

$$V_i = p_i \sum_{k=0}^{n-1} \frac{k!(n-k-1)!}{n!} \frac{(n-1)!}{k!(n-k-1)!} = p_i \sum_{k=0}^{n-1} \frac{(n-1)!}{n!} = p_i \sum_{k=0}^{n-1} \frac{1}{n} = p_i,$$

de onde se conclui que, para casos onde a medida de incerteza *fuzzy* coincide com uma medida probabilística, ou seja,  $\mu_i = \mu(\{x_i\}) = p_i$ , como no caso de problemas de agrupamento, a Entropia de Shapley coincide com a Entropia de Shannon, de forma que [88, 90]:

$$H(\mu) = - \sum_{i=1}^n V_i \ln V_i = - \sum_{i=1}^n p_i \ln p_i, \quad (4.30)$$

sendo a Entropia de Shannon um caso particular da Entropia de Shapley [88].

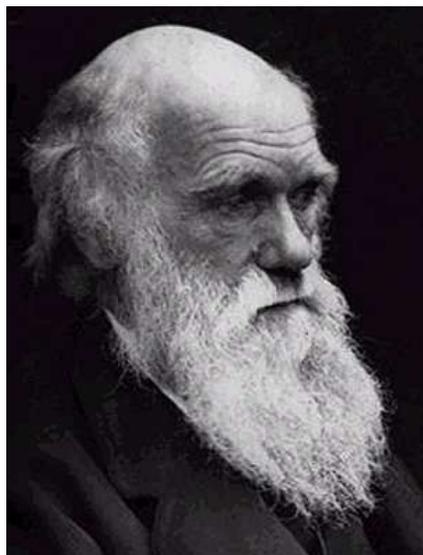
#### 4.4 Computação Evolucionária

Uma das principais metodologias componentes da Inteligência Computacional é a Computação Evolucionária, que lida com formalismos e estratégias para resolução de problemas baseando-se em mecanismos evolutivos provenientes da Genética, de elementos da Teoria da Evolução de Darwin (ver figura 4.2) [93], tais como os mecanismos de seleção natural, cruzamento e mutações, além dos paradigmas baseados em comportamento emergente adaptativo [81]. O principal objetivo da Computação Evolucionária é prover ferramentas para a construção de sistemas inteligentes para modelar comportamento inteligente [81].

O campo da Computação Evolucionária pode ser dividido em basicamente cinco famílias de paradigmas [81]:

1. Algoritmos genéticos;
2. Programação evolucionária;
3. Estratégias evolutivas;
4. Programação genética;
5. Otimização por enxame de partículas.

Dessas, serão tratadas neste trabalho apenas as seguintes abordagens: algoritmos genéticos, programação evolucionária e otimização por enxame de partículas, devido a formarem a base principal de inspiração para o formalismo do método dialético objetivo como método de busca e otimização heurística, além da própria dialética materialista.



**Figura 4.2** Charles Robert Darwin, um dos pais da Teoria da Evolução (1809-1882)

#### 4.4.1 Algoritmos Genéticos

Os algoritmos genéticos (*Genetic Algorithms*, GAs) buscam inspiração na Teoria da Evolução de Darwin e na Genética de Mendel (ver figura 4.3), assimilando e adaptando seus principais conceitos e definições, como os conceitos de fenótipo, genótipo, cromossomo, gene, cruzamento, seleção natural, aptidão, geração, mutação, clonagem e elitismo [81, 8, 93].

Os GAs consistem basicamente em processos iterativos para busca e otimização heurísticas, onde se tem a geração de um conjunto inicial de candidatos à solução do problema, no caso, candidatos à maximização de uma determinada função objetivo ou, semelhantemente, à minimização de uma determinada função custo; esse conjunto inicial de candidatos recebe o nome de *população inicial*, enquanto cada candidato à solução, membro desse conjunto inicial, é um *indivíduo*, inicialmente caracterizado por um *fenótipo*; para resolução do problema, os indivíduos passam a ser representados por *genótipos*, caracterizados por *cromossomos*, vetores tipicamente binários, onde cada coordenada corresponde a uma característica do indivíduo, ou *gene*, que pode estar ativado ou não ativado [8].

Um GA pode ser definido pelo seguinte algoritmo generalizado, ilustrado pela figura 4.4:

1. Geração da população inicial: é gerado um conjunto  $X(0)$  de  $m(0)$  indivíduos (cromossomos) na forma de vetores binários aleatórios  $\mathbf{x}_i(0) \in \{0, 1\}^n$ , para  $1 \leq i \leq m(0)$ , ou seja,  $X(0) = \{\mathbf{x}_1(0), \mathbf{x}_2(0), \dots, \mathbf{x}_{m(0)}(0)\}$ , onde  $n$  é a dimensionalidade do problema, ou *número de genes* (onde  $x_{i,j} \in \{0, 1\}$  é o  $j$ -ésimo gene do  $i$ -ésimo indivíduo, para  $1 \leq j \leq n$ );
2. Repita até que seja atingido um determinado número máximo de gerações (iterações) ou que se atinja uma boa estimativa,  $f(\tilde{\mathbf{x}}(t))$ , do valor máximo verdadeiro,  $f(\mathbf{x}^*)$ , da função objetivo  $f : \{0, 1\}^n \rightarrow \mathbb{R}$ :
  - (a) Exposição dos indivíduos ao ambiente: avaliação da função objetivo em todos os indivíduos,  $f(\mathbf{x}_i(t))$ , para  $1 \leq i \leq m(t)$ ;



**Figura 4.3** Gregor Mendel, monge agostiniano, sacerdote e pai da Genética (1822-1884)

- (b) Seleção natural: cada indivíduo recebe um determinado *grau de aptidão*,  $n_i(t)$ , proporcional ao valor da função objetivo correspondente, de forma que [94]:

$$n_i(t) = \frac{f(\mathbf{x}_i(t))}{\sum_{j=1}^{m(t)} f(\mathbf{x}_j(t))}, \quad (4.31)$$

onde  $1 \leq i \leq m(t)$ . No *Método da Roleta*, números aleatórios  $r_k(t)$  uniformemente distribuídos no intervalo  $[0, 1]$  são gerados em uma determinada quantidade  $m_r(t)$ , onde  $1 \leq k \leq m_r(t) \leq m(t)$  [8]; caso a condição

$$\sum_{j=1}^{i-1} n_j(t) < r_k(t) < \sum_{j=1}^i n_j(t) \quad (4.32)$$

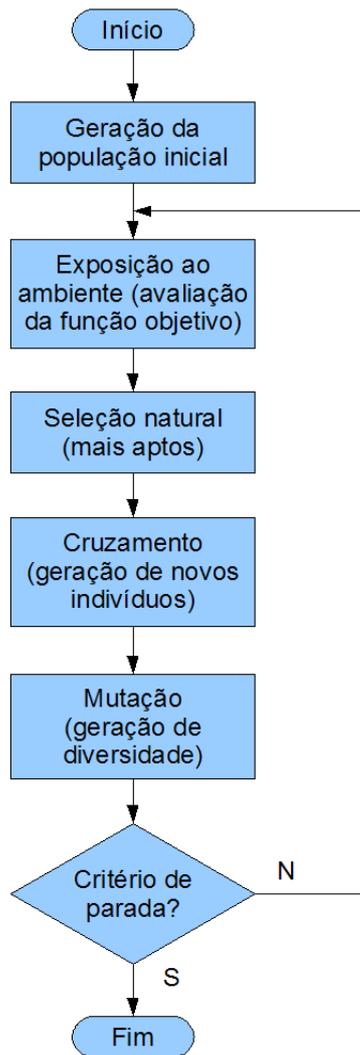
seja satisfeita, o  $i$ -ésimo indivíduo é considerado *apto* para passar à fase seguinte e realizar o cruzamento, e passa a integrar o conjunto dos indivíduos aptos a cruzar,  $A(t)$ , ou seja,  $\mathbf{x}_i(t) \in A(t)$ , para  $1 \leq i \leq m(t)$ ;

- (c) Cruzamento: Define-se uma máscara  $\mathbf{m} = (m_1, m_2, \dots, m_n)^T \in \{0, 1\}^n$ ; a partir dessa máscara são gerados os indivíduos da geração seguinte, da forma que segue [95, 8, 94]:

$$x'_{u,i}(t+1) = \begin{cases} x_{r,i}(t), & m_i = 1 \\ x_{s,i}(t), & m_i = 0 \end{cases}, \quad (4.33)$$

$$x'_{v,i}(t+1) = \begin{cases} x_{s,i}(t), & m_i = 1 \\ x_{r,i}(t), & m_i = 0 \end{cases}, \quad (4.34)$$

$$(4.35)$$



**Figura 4.4** Fluxograma geral do GA

onde  $\mathbf{x}_r(t)$  e  $\mathbf{x}_s(t)$  são os *pais*, ou *antepassados*, enquanto  $\mathbf{x}_u(t+1)'$  e  $\mathbf{x}_v(t+1)$  são os *filhos*, ou *descendentes*, antes de sofrerem mutações, para  $1 \leq r, s \leq m(t)$  e  $1 \leq u, v \leq m(t+1)$ ;

- (d) *Mutação*: nesta etapa, tipicamente um gene de cada indivíduo filho é aleatoriamente escolhido e modificado, ou seja,

$$x_i(t+1) = \varphi(x'_i(t+1)) = \begin{cases} 1 - x'_i(t+1), & i = k \\ x'_i(t+1), & i \neq k \end{cases}, \quad (4.36)$$

onde  $\varphi(\cdot)$  é o *operador de mutação*, que define o valor do índice  $k$  e a negação do  $k$ -ésimo gene. Uma estratégia bastante comum é fazer com que indivíduos muito aptos passem para a geração seguinte automaticamente, sem serem submetidos à operação aleatória de seleção natural; a essa estratégia dá-se o nome de *elitismo* [8, 81]. Já quando esses indivíduos também não são submetidos à mutação, tem-se uma situação em que indivíduos filhos são exatamente iguais aos pais, o que recebe o nome de *clonagem* [8, 81].

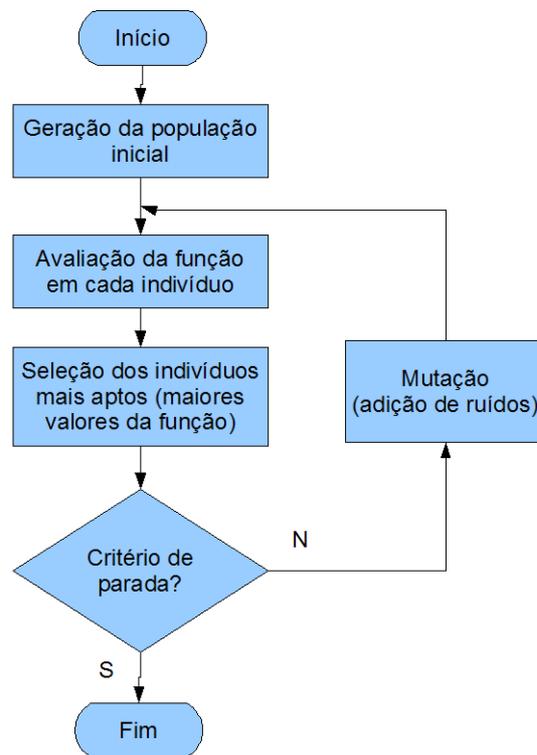
Os algoritmos genéticos podem ser utilizados com sucesso na solução de diversos problemas, como a determinação da arquitetura de rede neural artificial mais adequada a um determinado problema (otimização de arquitetura neural), como problemas de controle de processos, classificação e reconhecimento de padrões [8]. GAs também podem ser utilizados na solução de problemas de otimização multiobjetivo [95]. No entanto, um problema crítico é a modelagem dos candidatos a soluções na forma de cromossomos [81, 8].

No caso da busca de redes neurais ótimas ou sub-ótimas para determinados problemas, as arquiteturas de redes, e possivelmente até seus pesos, devem ser modelados na forma de vetores binários, de acordo com a definição usual de algoritmos genéticos, enquanto a função a ser otimizada deve ser alguma medida de qualidade a ser atingida, como por exemplo a taxa de acerto global ou o índice  $\kappa$ , em problemas de classificação e reconhecimento de padrões. Assim, embora o uso de GAs possa ajudar na busca da solução ótima ou sub-ótima, as soluções usando GAs demandam um custo computacional relativamente alto, além de serem muito sensíveis à modelagem dos fenótipos em genótipos [81, 8].

#### 4.4.2 Programação Evolucionária

A Programação Evolucionária (*Evolutionary Programming*, EP) é uma das principais famílias de técnicas da Computação Evolucionária. A EP foi inicialmente desenvolvida como uma abordagem de Inteligência Artificial, mas que passou a ser largamente empregada na resolução de problemas numéricos, dentre eles os problemas de busca e otimização [96, 81, 97].

Os algoritmos de EP são muito semelhantes aos algoritmos genéticos, tendo como principais diferenças a ausência da representação na forma de cromossomos e de operadores de cruzamento: nos EPs os indivíduos são representados por vetores de pesos reais ao invés de cromossomos; os indivíduos filhos são gerados a partir da mutação dos pais, simplesmente, sem cruzamento, podendo parte dos indivíduos filhos serem gerados a partir de clonagem; a mutação consiste na adição de algum tipo de ruído aos pesos dos indivíduos; já a seleção é feita a partir dos melhores indivíduos da união das populações de pais e filhos [97].



**Figura 4.5** Fluxograma geral dos algoritmos EPs

Em problemas de busca e otimização, os EPs são definidos pelo seguinte algoritmo generalizado, ilustrado pela figura 4.5:

1. Geração da população inicial de indivíduos, na forma de vetores de pesos, onde cada indivíduo é um candidato à solução;
2. Repetir até que o critério de parada seja atingido:
  - (a) Exposição dos indivíduos ao ambiente, ou seja, avaliação da função objetivo em todos os indivíduos da população;
  - (b) Busca do indivíduo mais apto, ou seja, do indivíduo associado ao maior valor da função objetivo;
  - (c) Geração dos novos indivíduos por mutação dos indivíduos mais aptos escolhidos dentre os indivíduos anteriores.

Assim, o problema de otimização de uma função objetivo  $f : S \rightarrow \mathbb{R}$ , onde  $S \subseteq \mathbb{R}^n$ , sendo  $n$  a dimensionalidade do problema, considerando um conjunto inicial de  $m$  indivíduos candidatos à solução e representados pelos vetores  $\mathbf{x}_i$ , seria resolvido heurísticamente pela atualização dos vetores segundo expressões semelhantes à que segue:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{r}_i(t), \quad (4.37)$$

para  $1 \leq i \leq m$ , onde  $\mathbf{r}_i(t)$  é um vetor aleatório  $n$ -dimensional cujas coordenadas  $r_{i,j}(t)$ , para  $1 \leq j \leq n$ , são tipicamente números aleatórios com distribuição gaussiana, podendo obedecer a outras distribuições probabilísticas, como a de Cauchy, ou mesmo ser obtidos por processos iterativos próprios envolvendo números aleatórios de distribuições quaisquer [97]. As iterações continuam até que se atinja uma determinada quantidade máxima de iterações ou que um dos seguintes critérios seja atingido:

$$\|\mathbf{x}^* - \tilde{\mathbf{x}}(t)\| < \varepsilon, \quad (4.38)$$

$$|f(\mathbf{x}^*) - f(\tilde{\mathbf{x}}(t))| < \varepsilon, \quad (4.39)$$

para

$$f(\tilde{\mathbf{x}}(t)) = \max_{1 \leq j \leq m} f(\mathbf{x}_j(t)), \quad (4.40)$$

onde  $\mathbf{x}^*$  é o ponto de máximo da função objetivo  $f$  e  $f(\mathbf{x}^*)$  é o valor máximo de  $f$ , enquanto  $\tilde{\mathbf{x}}(t)$  e  $f(\tilde{\mathbf{x}}(t))$  são suas respectivas estimativas na  $t$ -ésima geração;  $\varepsilon$  é a tolerância da solução. Os critérios das expressões 4.38 e 4.39 são critérios mutuamente excludentes, não devendo, portanto, ser utilizados juntos.

Nas subseções a seguir são apresentados os principais algoritmos de programação evolucionária.

#### 4.4.2.1 Algoritmo CEP

O algoritmo CEP (*Classical Evolutionary Programming*) é baseado na estratégia de mutação gaussiana e funciona da seguinte maneira [97, 98]:

1. Geração de uma população inicial de  $m$  indivíduos  $n$ -dimensionais, representados pelo par  $(\mathbf{x}_i, \mathbf{s}_i)$ ,  $\forall i \in \{1, 2, \dots, m\}$ , onde  $\mathbf{x}_i$  são variáveis objetivas, ou seja, candidados à solução que otimiza a função objetivo  $f : S \rightarrow \mathbb{R}$ , onde  $S \subseteq \mathbb{R}^n$  e  $n$  é a dimensionalidade do problema, enquanto  $\mathbf{s}_i = (s_{i,1}, s_{i,2}, \dots, s_{i,n})^T$  são os desvios padrões para as mutações gaussianas, também chamados de parâmetros estratégicos em algoritmos evolucionários auto-adaptativos;
2. Repetir até que um determinado critério de parada seja satisfeito:

- (a) Avaliação da função objetivo nos indivíduos pais,  $(\mathbf{x}_i, \mathbf{s}_i)$ ,  $\forall i \in \{1, 2, \dots, m\}$ , gerando  $f(\mathbf{x}_i)$ ;
- (b) Geração de  $m$  indivíduos filhos por mutação,  $(\mathbf{x}'_i, \mathbf{s}'_i)$ ,  $\forall i \in \{1, 2, \dots, m\}$ , de forma que:

$$s'_{i,j} = s_{i,j} \exp(\tau_a G(0, 1) + \tau_b G_j(0, 1)), \quad (4.41)$$

$$x'_{i,j} = x_{i,j} + s'_{i,j} G_j(0, 1), \quad (4.42)$$

$\forall i \in \{1, 2, \dots, m\}$  e  $\forall j \in \{1, 2, \dots, n\}$ , onde  $G(0, 1)$  é uma variável aleatória gaussiana de média 0 e variância 1 gerada para um  $i$  fixo, enquanto  $G_j(0, 1)$  é gerada para cada  $j$ ; já os parâmetros de controle são definidos como  $\tau_a = 1/\sqrt{2n}$  e  $\tau_b = 1/\sqrt{2\sqrt{n}}$ ;

- (c) Avaliação da função objetivo nos indivíduos filhos,  $(\mathbf{x}'_i, \mathbf{s}'_i)$ ,  $\forall i \in \{1, 2, \dots, m\}$ , gerando  $f(\mathbf{x}'_i)$ ;
- (d) Seleção dos  $m$  indivíduos mais aptos dentre a população de  $2m$  indivíduos composta da junção das populações de pais e de filhos, para compor a nova geração de pais: em problemas de minimização isso significa selecionar os  $m$  indivíduos  $(\mathbf{x}_i, \mathbf{s}_i)$  ou  $(\mathbf{x}'_i, \mathbf{s}'_i)$  com menores avaliações da função objetivo, ou seja, com os  $m$  menores valores de  $f(\mathbf{x}_i)$  ou  $f(\mathbf{x}'_i)$ . Idem para problemas de maximização.

#### 4.4.2.2 Algoritmo FEP

O algoritmo FEP (*Fast Evolutionary Programming*) é uma modificação do algoritmo CEP proposta por Yao *et al.* (1999) [97] e se diferencia do CEP pelo uso da mutação de Cauchy, definida como segue [97, 98]:

$$s'_{i,j} = s_{i,j} \exp(\tau_a G(0, 1) + \tau_b G_j(0, 1)), \quad (4.43)$$

$$x'_{i,j} = x_{i,j} + s'_{i,j} \delta_j, \quad (4.44)$$

$\forall i \in \{1, 2, \dots, m\}$  e  $\forall j \in \{1, 2, \dots, n\}$ , onde  $\delta_j$  é uma variável aleatória de Cauchy padrão gerada para cada  $j$ . A função densidade de probabilidade de Cauchy padrão é dada por:

$$f_C(x) = \frac{1}{\pi(1+x^2)}. \quad (4.45)$$

## 4.4.2.3 Algoritmo LEP

O algoritmo LEP (*Lévy-Type Evolutionary Programming*) é uma modificação dos algoritmos CEP e FEP proposta por Iwamatsu (2002) [99] e se diferencia do CEP e do FEP pelo uso da mutação baseada na distribuição de probabilidade de Lévy, definida como segue [99, 98]:

$$s'_{i,j} = s_{i,j} \exp(\tau_a G(0, 1) + \tau_b G_j(0, 1)), \quad (4.46)$$

$$x'_{i,j} = x_{i,j} + s'_{i,j} L_j(\beta), \quad (4.47)$$

$\forall i \in \{1, 2, \dots, m\}$  e  $\forall j \in \{1, 2, \dots, n\}$ , onde  $L_j(\beta)$  é uma variável aleatória de Lévy com parâmetro de escala  $0 < \beta \leq 3$  gerada para cada  $j$ .

## 4.4.2.4 Algoritmo SPMEP

O algoritmo SPMEP (*Single-Point Mutation Evolutionary Programming*) possui uma filosofia diferente dos demais algoritmos apresentados: ao invés de simplesmente mudar a distribuição das variáveis aleatórias presentes as mutações, o SPMEP busca alterar apenas um dos  $n$  componentes dos pares  $(\mathbf{x}_i, \mathbf{s}_i)$  em cada mutação, gerando os indivíduos filhos da forma que segue [100, 98]:

$$s'_{i,j} = s_{i,j} \exp(-\alpha), \quad (4.48)$$

$$x'_{i,j} = x_{i,j} + s'_{i,j} G_j(0, 1), \quad (4.49)$$

$\forall i \in \{1, 2, \dots, m\}$ , onde  $j$  é fixo para cada mutação e escolhido de forma que  $j \sim U[1, n]$ .

## 4.4.2.5 Comentários

Os algoritmos EPs variam basicamente de acordo com a estratégia de mutação. Várias estratégias de mutação têm sido propostas na última década. A mutação gaussiana é a mutação clássica em EPs (CEP), mas não é muito boa para otimização de funções multimodais [98, 97]. A mutação de Cauchy foi proposta no algoritmo FEP e pode gerar melhores soluções do que o CEP em funções multimodais, mas é menos eficiente do que o CEP em problemas unimodais [98]. A mutação baseada na distribuição de Lévy, do algoritmo LEP, foi proposta como um algoritmo mais flexível do que CEP e FEP, e é equivalente ao FEP quando o parâmetro de escala  $\beta = 1$ , e se reduz ao CEP quando  $\beta = 2$ , mas na prática é difícil determinar o melhor valor de  $\beta$  para um determinado problema [98]. Além do mais, o algoritmo LEP não consegue otimizar todas as funções de teste utilizadas neste trabalho eficientemente [98, 99]. Já o algoritmo SPMEP, baseado na mutação de ponto fixo, é superior ao CEP e ao FEP para muitas funções multimodais de alta dimensão, mas é inferior ao CEP para funções de baixa dimensão e poucos mínimos locais [98].

Os algoritmos EPs são relativamente simples de implementar. No entanto, em problemas de busca e otimização de funções, resultam em muitas iterações antes de atingir resultados aceitáveis para determinadas classes de funções, além de exigir uma população inicialmente grande, o que resulta, aliado a uma quantidade razoável de iterações, um grande número de avaliações da função antes de se atingir uma solução aproximada razoável, o que implica em alto custo computacional caso a função seja de avaliação computacionalmente custosa ou

matematicamente complexa e se encaixe na classe de funções para a qual o algoritmo não é ótimo [97].

Na verdade, Wolpert e Macready mostraram que isso é verdade para qualquer algoritmo de otimização de uso geral, pois embora cada algoritmo obtenha resultados razoáveis ou mesmo ótimos para determinadas classes de funções, para outras funções ocorrerá uma espécie de compensação no desempenho, de forma que, na média, seus resultados se equivalem [101].

#### 4.4.3 Otimização por Enxame de Partículas

Os algoritmos de Otimização por Enxame de Partículas (*Particle Swarm Optimization*, PSO) foram criados por James Kennedy e Russel Eberhart, em 1995, respectivamente um psicólogo social e um engenheiro eletricitista [102]. Os PSOs são baseados no comportamento e movimento de bandos de animais, tais como peixes e pássaros, sendo portanto algoritmos baseados em teorias que descrevem comportamento social animal, possuindo elementos em comum com os algoritmos genéticos e com programação evolucionária [103, 102].

De forma semelhante aos algoritmos genéticos, o PSO é inicializado com uma população inicial aleatória. Entretanto, enquanto nos algoritmos genéticos os indivíduos dessa população inicial são representados por cromossomos, no PSO estão associados a cada indivíduo um vetor posição e um vetor de velocidades. Além disso, no PSO não ocorrem mutações nem seleção de indivíduos. Assim, a cada iteração, apenas são ajustadas posições e velocidades dos diferentes indivíduos na direção da *melhor posição global* e da *melhor posição individual*, de acordo com uma determinada *função objetivo*, segundo a seguinte expressão canônica [81, 104, 105, 106, 107, 108]:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1), \quad (4.50)$$

desde que

$$\mathbf{v}_i(t+1) = w\mathbf{v}_i(t) + c_1r_1(t)(\mathbf{p}_i(t) - \mathbf{x}_i(t)) + c_2r_2(t)(\mathbf{p}_g(t) - \mathbf{x}_i(t)), \quad (4.51)$$

para  $1 \leq i \leq m$ , onde  $m$  é o número de partículas do enxame;  $w$  é o fator de inércia, onde  $0 < w < 1$ ;  $r_1(t)$  e  $r_2(t)$  são números aleatórios uniformemente distribuídos no intervalo  $[0, 1]$ ;  $c_1$  e  $c_2$  são constantes de constrição, também chamados de coeficientes de aceleração, de forma que  $c_1 + c_2 = 4$  (tipicamente,  $c_1 = 2 + \Delta$  e  $c_2 = 2 - \Delta$ , onde  $\Delta \approx 0$ ), sendo que  $c_1$  é o peso devido à consciência própria da partícula, consciência individual ou consciência local, dependendo da implementação, enquanto  $c_2$  é o peso devido à consciência global;  $\mathbf{x}_i$  é a posição, enquanto  $\mathbf{v}_i$  é a velocidade da  $i$ -ésima partícula;  $\mathbf{p}_g$  é a melhor posição global, ou seja

$$f(\mathbf{p}_g) = \max_{1 \leq j \leq m} f(\mathbf{x}_j), \quad (4.52)$$

enquanto  $\mathbf{p}_i$  é a melhor posição individual ou local em relação à  $i$ -ésima partícula. Caso a implementação seja baseada na melhor posição individual, a definição utilizada será a que segue:

$$f(\mathbf{x}_i(t'')) = \max_{0 \leq t' \leq t} f(\mathbf{x}_i(t')) \Rightarrow \mathbf{p}_i(t) = \mathbf{x}_i(t''). \quad (4.53)$$

Caso a implementação utilize a definição de melhor posição local, na vizinhança  $N_i$  da  $i$ -ésima partícula  $f(\mathbf{p}_i)$  é máximo, ou seja:

$$f(\mathbf{p}_i) = \max_{j \in N_i} f(\mathbf{x}_j), \quad (4.54)$$

onde  $f : S \rightarrow \mathbb{R}$  é a função objetivo a ser maximizada, desde que  $S \subseteq \mathbb{R}^n$ , sendo que  $n$  é a dimensionalidade do problema. Assim, os vetores  $\mathbf{x}_i$ ,  $\mathbf{v}_i$ ,  $\mathbf{p}_i$  e  $\mathbf{p}_g$  são  $n$ -dimensionais, onde  $1 \leq i \leq m$ .  $\mathbf{p}_i$  pode ainda ser a melhor posição individual da  $i$ -ésima partícula, ou seja, a posição que modela a experiência individual de cada partícula, de forma que

$$f(\mathbf{p}_i(t)) = \max_{0 \leq t' \leq t} f(\mathbf{x}_i(t')). \quad (4.55)$$

É importante destacar que, quando se fala em vizinhança nos algoritmos PSO, não se trata de vizinhança levando em conta a posição da partícula, e sim o seu índice, que pode ser explorado na montagem de diferentes topologias. Assim, existem diversas versões do algoritmo PSO: há versões que exploram a ideia de se utilizar diferentes topologias locais [109]; umas são dedicadas à resolução de problemas de otimização multiobjetivo, utilizando o conceito de co-evolução, onde um grupo de enxames evolui em conjunto de forma interdependente, onde cada um é responsável por um determinado objetivo a ser atingido, dependendo dos demais [81, 110, 111, 108, 112]. Também existem versões que introduzem melhorias na expressão canônica por meio da introdução de conceitos físicos como a velocidade de escape, resultando na seguinte expressão de ajuste de velocidades [113]:

$$\mathbf{v}_i(t+1) = w\mathbf{v}_e(t) + c_1r_1(t)(\mathbf{p}_i(t) - \mathbf{x}_i(t)) + c_2r_2(t)(\mathbf{p}_g(t) - \mathbf{x}_i(t)), \quad (4.56)$$

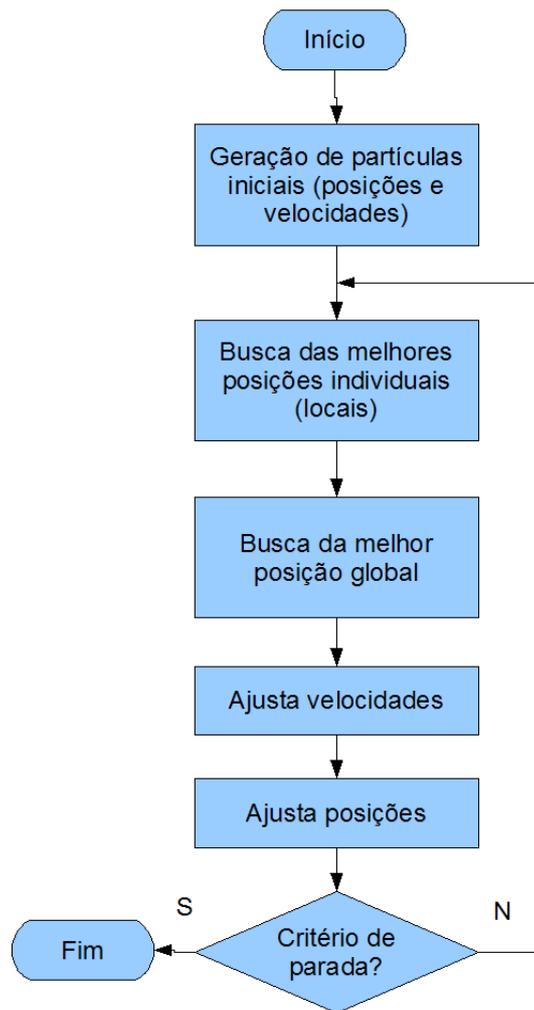
para

$$v_{e,j}(t) = \begin{cases} v_{i,j}(t), & |v_{i,j}(t)| > e_c(t) \\ \frac{1}{\rho} v_{\max} r_j(t), & |v_{i,j}(t)| \leq e_c(t) \end{cases}, \quad (4.57)$$

onde  $\mathbf{v}_e(t) = (v_{e,1}(t), v_{e,2}(t), \dots, v_{e,n}(t))^T$  é a *velocidade de escape*, para  $1 \leq i \leq m$  e  $1 \leq j \leq n$ ;  $v_{\max}$  é a velocidade escalar máxima, assumindo tipicamente o maior valor do domínio da função objetivo;  $r_j(t)$  é uma sequência de números aleatórios uniformemente distribuídos no intervalo  $[-1, 1]$  no instante  $t$ ;  $\rho$  é o fator de escala para controlar o domínio da velocidade de escape; e  $e_c(t)$  é o limiar ajustável que determina a condição de escape.

Existem ainda versões do PSO inspiradas em teorias como a Teoria do Caos, que consiste no ajuste do fator de inércia  $w(t)$  por meio de uma equação caótica,  $w(t+1) = 4w(t)(1-w(t))$ , onde  $1 < w(t) < 1$  [104]. Também há versões híbridas que misturam conceitos de PSOs e Lógica Fuzzy [81, 114]. Outras versões são dedicadas à resolução de problemas de otimização no domínio discreto [115]. Utilizando o PSO como inspiração, também é possível construir novos algoritmos baseados em processos como, por exemplo, processos químicos [116].

Como não possui mecanismos de mutação e seleção de indivíduos, o PSO é uma classe de algoritmos relativamente simples de serem implementados, já que se reduzem a um processo iterativo aliado à busca dos melhores indivíduos por meio da avaliação da função objetivo em todos os indivíduos (partículas), conforme mostra a figura 4.6. Diversos resultados mostram que, para uma classe relativamente grande de problemas, inclusive problemas multiobjetivo (tipicamente resolvidos utilizando conjuntos de PSOs em cooperação), o PSO atinge bons resultados [81].



**Figura 4.6** Fluxograma geral do PSO

## 4.5 Método Dialético de Otimização

É possível adaptar o método dialético objetivo para problemas de busca e otimização considerando uma abordagem onde cada pólo dentro de uma fase histórica corresponde a um candidato à solução do problema. Assim, a união dialética dos diversos pólos corresponde ao universo de soluções possíveis do problema.

A ideia básica por trás do método dialético de otimização é associar a função objetivo do problema de otimização à força social de cada pólo, sendo a dinâmica do método de otimização governada pela luta de pólos: o movimento dos pólos se dá em função do *pólo hegemônico presente* e do *pólo hegemônico histórico*, onde o pólo hegemônico é aquele dotado da maior força social do conjunto das forças de todos os pólos em um determinado momento histórico, ou seja, é a maior força social dentro do conjunto correlação de forças, sendo o pólo hegemônico presente o pólo associado à maior força social da correlação de forças atual, enquanto o pólo hegemônico histórico é aquele que detém a maior força histórica até o instante atual.

A busca das soluções possíveis passa a ser ligada intrinsecamente ao movimento dos pólos e de suas contradições nas diversas fases históricas, o que inclui, nos períodos de crise revolucionária, a fusão de pólos com baixos níveis de contradição entre si e a geração de novos pólos tanto como síntese entre os pólos envolvidos em altos níveis de contradição, quanto como antítese a pólos que passam a ser pólos tese, além da adição de perturbações aos pólos sobreviventes, processo que por sua vez se assemelha às mutações nas abordagens baseadas em computação evolucionária e tem por fim adicionar diversidade ao processo de busca.

### 4.5.1 Definição Geral

Para se compreender a proposta de adaptação do método dialético objetivo para busca e otimização, é importante antes definir alguns conceitos importantes que serão utilizados ao longo deste texto:

**Pólo:** É a menor unidade integrante de um sistema dialético. Dado um conjunto de pólos  $\Omega = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ , um *pólo*  $\mathbf{w}_i$ , da mesma forma que na definição do classificador dialético objetivo [82, 83], está associado a um vetor de pesos  $\mathbf{w}_i = (w_{i,1}, w_{i,2}, \dots, w_{i,n})^T$ , onde  $\mathbf{w}_i \in S$ ,  $m$  é o número de pólos e  $n$  é a dimensionalidade do sistema. Na proposta de método dialético de otimização, um pólo também executa o papel de um candidato à solução. A dimensionalidade do problema coincide com a dimensionalidade da função objetivo  $f : S \rightarrow \mathbb{R}$  que se deseja maximizar, onde  $S \subseteq \mathbb{R}^n$  e  $\Omega \subseteq S$ .

**Força social:** A cada pólo  $\mathbf{w}_i$  está associada uma *força social* cujo cálculo coincide com a avaliação da função objetivo  $f$  no pólo  $\mathbf{w}_i$ , ou seja, a força social no  $i$ -ésimo pólo é dada pelo termo  $f(\mathbf{w}_i)$ .

**Hegemonia:** No processo de luta de pólos, diz-se que um pólo  $k$  exerce a *hegemonia* no momento  $t$  quando:

$$f(\mathbf{w}_k(t)) = f_C(t) = \max_{1 \leq j \leq m(t)} f(\mathbf{w}_j(t)), \quad (4.58)$$

onde  $1k \in \{1, 2, \dots, m(t)\}$ . O vetor  $\mathbf{w}_C(t) = \mathbf{w}_k(t)$  é chamado de *pólo hegemônico presente*, ou *pólo hegemônico contemporâneo*, enquanto  $f_C(t)$  é a *força hegemônica presente*, ou *força hegemônica contemporânea*. A *força hegemônica histórica* no instante  $t$ ,  $f_H(t)$ , é dada por:

$$f_H(t) = \max_{0 \leq t' \leq t} f_C(t'), \quad (4.59)$$

onde  $\mathbf{w}_H(t) = \mathbf{w}_C(t')$ , para  $f(\mathbf{w}_C(t')) = f_H(t)$  e  $0 \leq t' \leq t$ .

**Antítese absoluta:** Dado  $x$  tal que  $a \leq x \leq b$ , onde  $a, b \in \mathbb{R}$ , o *oposto* de  $x$  é dado por [117]:

$$\check{x} = b - x + a. \quad (4.60)$$

Assumindo  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$  e que  $\mathbf{x} \in S \Rightarrow r_i \leq x_i \leq s_i, \forall i = 1, 2, \dots, n$ , onde  $r_i$  e  $s_i$  são os limites inferior e superior da  $i$ -ésima dimensão de  $S$ , o *vetor oposto* associado,  $\check{\mathbf{x}} = (\check{x}_1, \check{x}_2, \dots, \check{x}_n)^T$  tem suas coordenadas calculadas da forma que segue [117]:

$$\check{x}_i = s_i - x_i + r_i, \quad (4.61)$$

onde  $i = 1, 2, \dots, n$ . O *vetor antítese absoluta* do pólo  $\mathbf{w}$  é definido pelo seu vetor oposto  $\check{\mathbf{w}}$ . Rahnamayan *et al* afirmam que, em programação evolucionária e em otimização por enxame de partículas, a presença de pares opostos na população inicial tipicamente acelera em 10% a convergência dos algoritmos [117]. Ao mesmo tempo, a *antítese absoluta*, ou *antítese antagônica*, é um conceito muito importante para a dialética materialista, pois expressa qualitativamente a máxima contradição entre dois pólos, sendo importante para a dinâmica do sistema dialético [11, 60, 36]. A figura 4.7 ilustra um exemplo bidimensional de vetor tese e sua antítese absoluta.

**Contradição:** A contradição entre os pólos  $\mathbf{w}_p$  e  $\mathbf{w}_q$  é definida por:

$$\delta_{p,q} = d(\mathbf{w}_p, \mathbf{w}_q), \quad (4.62)$$

onde  $d : S^2 \rightarrow \mathbb{R}_+$  é uma *função de distância*, de forma que  $\delta_{p,q} = \delta_{q,p}, \forall p, q$ . Uma função de distância típica é a função de distância euclidiana.

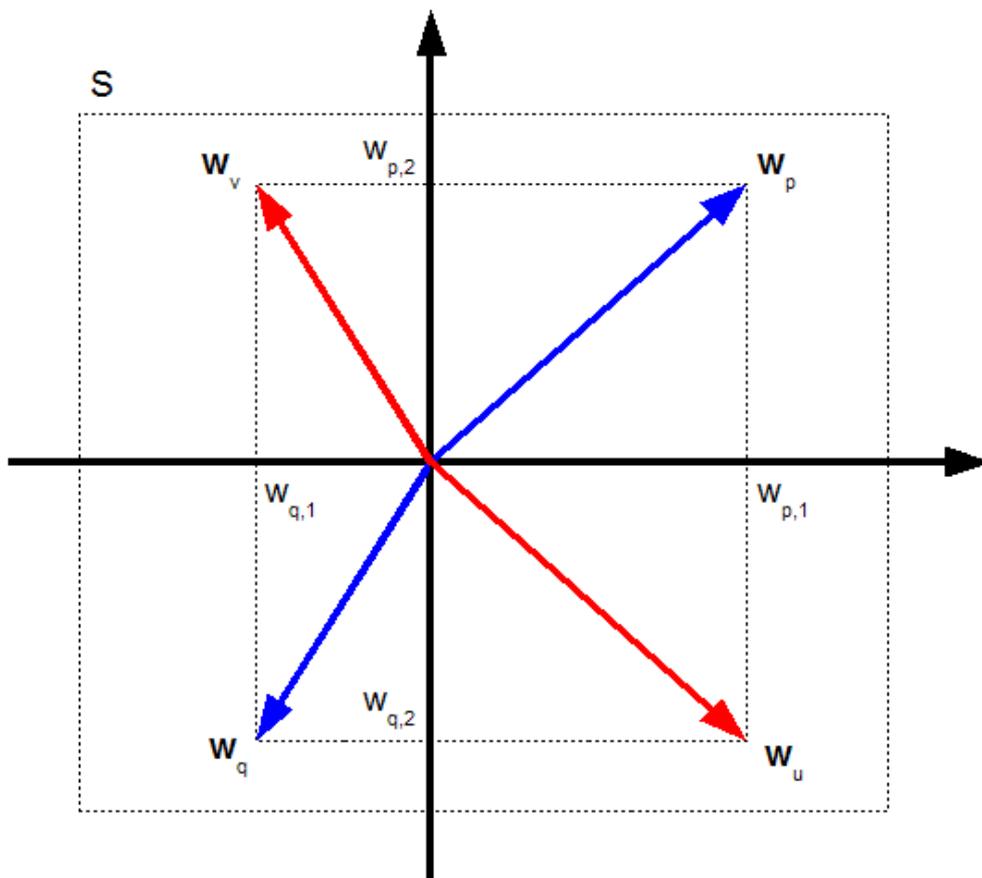
**Síntese:** De acordo com a concepção dialética, a síntese é a resolução da contradição entre dois pólos, onde um exerce o papel de tese e o outro, de antítese [36]. Os pólos  $\mathbf{w}_u, \mathbf{w}_v \in S$  são sínteses entre os pólos  $\mathbf{w}_p$  e  $\mathbf{w}_q$  se forem obtidos da forma

$$(\mathbf{w}_u, \mathbf{w}_v) = g(\mathbf{w}_p, \mathbf{w}_q), \quad (4.63)$$

onde  $g : S^2 \rightarrow S^2$ . Uma abordagem intuitiva a partir da concepção dialética seria considerar que toda síntese herda características da tese e da antítese [11, 38, 37, 13], havendo duas sínteses possíveis, como segue:

$$w_{u,i} = \begin{cases} w_{p,i}, & i \bmod 2 = 0, \\ w_{q,i}, & i \bmod 2 = 1, \end{cases}, \quad (4.64)$$





**Figura 4.8** Caso bidimensional de par de vetores tese e antítese,  $w_p$  and  $w_q$ , e suas respectivas sínteses,  $w_u$  e  $w_v$

dialética, têm-se um conjunto inicial de pólos composto de pares tese-antítese em contradição antagônica (um dos pólos é uma tese, enquanto o outro é sua antítese antagônica), gerando maior dinâmica inicial, ou seja, a luta de pólos é mais intensa [11, 38]. Já do ponto de vista da computação evolucionária, essa estratégia acelera tipicamente em 10% a convergência do algoritmo de otimização, uma vez que os algoritmos de busca e otimização baseados em computação evolucionária são altamente dependentes da população inicial e, portanto, a probabilidade de encontrar os pontos ótimos utilizando menos iterações pode ser aumentada ao se avaliar tanto um candidato à solução quanto o seu oposto [117]. Assim:

$$w_{i,j}(0) = \begin{cases} U(r_j, s_j), & 1 \leq i \leq \frac{1}{2}m(0), \\ \check{w}'_{i',j}(0), & 1 + \frac{1}{2}m(0) \leq i \leq m(0), \end{cases} \quad (4.66)$$

para  $i' = i - \frac{1}{2}m(0)$ ,  $1 \leq i \leq m(0)$  e  $1 \leq j \leq n$ , onde  $n$  é a dimensionalidade do problema de otimização,  $U(r_j, s_j)$  é um número aleatório uniformemente distribuído no intervalo  $[r_j, s_j]$  e  $S = [r_1, s_1] \times [r_2, s_2] \times \cdots \times [r_n, s_n]$ , desde que  $s_j > r_j$  e  $s_j, r_j \in \mathbb{R}$ .

2. Enquanto não se atinge um máximo de  $n_P$  fases históricas e a força hegemônica histórica não é maior do que um dado limiar superior de força (estimativa inicial do valor máximo da função objetivo),  $f_H(t) < f_{\text{sup}}$  (critério para se considerar o máximo da função objetivo atingido), repete-se:

**Evolução:** Enquanto não se atinge um máximo de  $n_H$  iterações e  $f_H(t) < f_{\text{sup}}$ , os pólos são ajustados segundo a seguinte expressão:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \Delta \mathbf{w}_{C,i}(t) + \Delta \mathbf{w}_{H,i}(t), \quad (4.67)$$

para

$$\Delta \mathbf{w}_{C,i}(t) = \eta_C(t)(1 - \mu_{C,i}(t))^2(\mathbf{w}_C(t) - \mathbf{w}_i(t)), \quad (4.68)$$

$$\Delta \mathbf{w}_{H,i}(t) = \eta_H(t)(1 - \mu_{H,i}(t))^2(\mathbf{w}_H(t) - \mathbf{w}_i(t)), \quad (4.69)$$

$$0 < \eta_C(t) < 1, \quad (4.70)$$

$$0 < \eta_H(t) < 1, \quad (4.71)$$

onde  $\eta_C(0) = \eta_H(0) = \eta_0$  e  $0 < \eta_0 < 1$ . Os termos  $\Delta \mathbf{w}_{C,i}(t)$  e  $\Delta \mathbf{w}_{H,i}(t)$  modelam as influências das hegemonias presente e histórica, nesta ordem, sobre o  $i$ -ésimo pólo, enquanto  $\eta_C(t)$  e  $\eta_H(t)$  são os respectivos passos de atualização dos pólos, atualizados a cada iteração e a cada fase histórica, respectivamente, de forma que

$$\eta_C(t+1) = \alpha \eta_C(t), \quad (4.72)$$

ao final de cada iteração e

$$\eta_H(t+1) = \alpha \eta_H(t), \quad (4.73)$$

ao final de cada fase histórica, para  $\alpha < 1$  (tipicamente,  $\alpha = 0,9999$ ). O decréscimo dos passos ao longo do tempo, embora lento, é efetuado para incrementar a

busca e, portanto, facilitar a convergência do algoritmo, tal como em algumas redes neurais artificiais, garantindo assim a convergência assintótica do algoritmo [33]. Os termos  $\mu_{C,i}$  e  $\mu_{H,i}$  são chamados de *pertinência presente* e *pertinência histórica*, respectivamente, sendo definidos da forma que segue, baseada nas funções de pertinência da versão clássica do classificador *fuzzy* c-médias [74, 75, 79, 77, 76]:

$$\mu_{C,i}(t) = \left( \sum_{j=1}^{m(t)} \frac{|f(\mathbf{w}_i(t)) - f_C(t)|}{|f(\mathbf{w}_j(t)) - f_C(t)|} \right)^{-1}, \quad (4.74)$$

$$\mu_{H,i}(t) = \left( \sum_{j=1}^{m(t)} \frac{|f(\mathbf{w}_i(t)) - f_H(t)|}{|f(\mathbf{w}_j(t)) - f_H(t)|} \right)^{-1}, \quad (4.75)$$

onde  $1 \leq i \leq m(t)$ . Assim, quando  $f(\mathbf{w}_i(t))$  se aproxima de  $f_C(t)$ , o termo  $\mu_{C,i}(t)$  se aproxima de 1, o que aproxima  $\Delta \mathbf{w}_{C,i}(t)$  de 0 e, portanto, torna a influência da correlação de forças presente praticamente nula, evitando alterações no peso devidas à hegemonia presente. Semelhantemente, quando  $f(\mathbf{w}_i(t))$  se aproxima de  $f_H(t)$ , o termo  $\mu_{H,i}(t)$  se aproxima de 1, o que aproxima  $\Delta \mathbf{w}_{H,i}(t)$  de 0, torna a influência da hegemonia histórica praticamente nula.

**Crise Revolucionária:** Na etapa de crise revolucionária são executados os seguintes passos:

- (a) Todas as contradições  $\delta_{i,j}$  são avaliadas; as contradições menores do que uma *contradição mínima*  $\delta_{\min}$  implicam a fusão entre os pólos, de forma que

$$\delta_{i,j}(t) > \delta_{\min} \Rightarrow \mathbf{w}_i(t), \mathbf{w}_j(t) \in \Omega(t+1), \quad (4.76)$$

$$\delta_{i,j}(t) \leq \delta_{\min} \Rightarrow \mathbf{w}_i(t) \in \Omega(t+1). \quad (4.77)$$

$i \neq j, \forall i, j$  onde  $1 \leq i, j \leq m(t)$  e  $\Omega(t+1)$  é o novo conjunto de pólos.

- (b) A partir das contradições avaliadas na etapa anterior, encontram-se aquelas maiores do que uma *contradição máxima*  $\delta_{\max}$ ; essas contradições serão consideradas as *contradições principais* do sistema dialético, sendo os pares de pólos envolvidos considerados como *pares tese-antítese*, cujos *pólos síntese* passam também a pertencer ao novo conjunto de pólos, ou seja:

$$\delta_{i,j}(t) > \delta_{\max} \Rightarrow g(\mathbf{w}_i(t), \mathbf{w}_j(t)) \in \Omega(t+1), \quad (4.78)$$

para  $i \neq j, \forall i, j$  onde  $1 \leq i, j \leq m(t)$ .

- (c) Adiciona-se o *efeito de crise*, dada a *máxima crise*,  $\chi_{\max}$ , a todos os pólos do sistema dialético  $\Omega(t+1)$ , gerando o novo conjunto de pólos,  $\Omega(t+2)$ , de forma que  $\mathbf{w}_k(t+2) \in \Omega(t+2)$ , desde que

$$\mathbf{w}_{k,i}(t+2) = \mathbf{w}_{k,i}(t+1) + \chi_{\max} G(0, 1), \quad (4.79)$$

para  $1 \leq k \leq m(t+1)$  e  $1 \leq i \leq n$ , onde  $G(0, 1)$  é um número aleatório de distribuição gaussiana com esperança 0 e variância 1.

- (d) Caso o critério de parada ainda não tenha sido atingido (número máximo de fases históricas atingido ou outro critério de parada a ser definido), é gerado um novo conjunto de pólos, de forma que:

$$\mathbf{w}_i(t+2) \in \Omega(t+2) \Rightarrow \check{\mathbf{w}}_i(t+2) \in \Omega(t+2), \quad (4.80)$$

para  $1 \leq i \leq m(t+2)$ , onde  $m(t+2) = 2m(t+1)$ . Logo, o conjunto de pólos é ampliado por meio da adição dos pólos em antítese antagônica aos pólos já existentes. Tal procedimento se dá como uma forma de modelar a concepção dialética de que, ao passar a uma nova fase histórica, o sistema carrega em si também os seus opostos, o que corresponde ao germe de sua potencial transformação em algo novo. Além do mais, esse procedimento repete o que foi assumido quando da inicialização do sistema dialético (geração do conjunto de pólos inicial,  $\Omega(0)$ ), podendo acelerar a convergência do algoritmo à solução ótima pela ideia de acrescentar elementos opostos à busca [117].

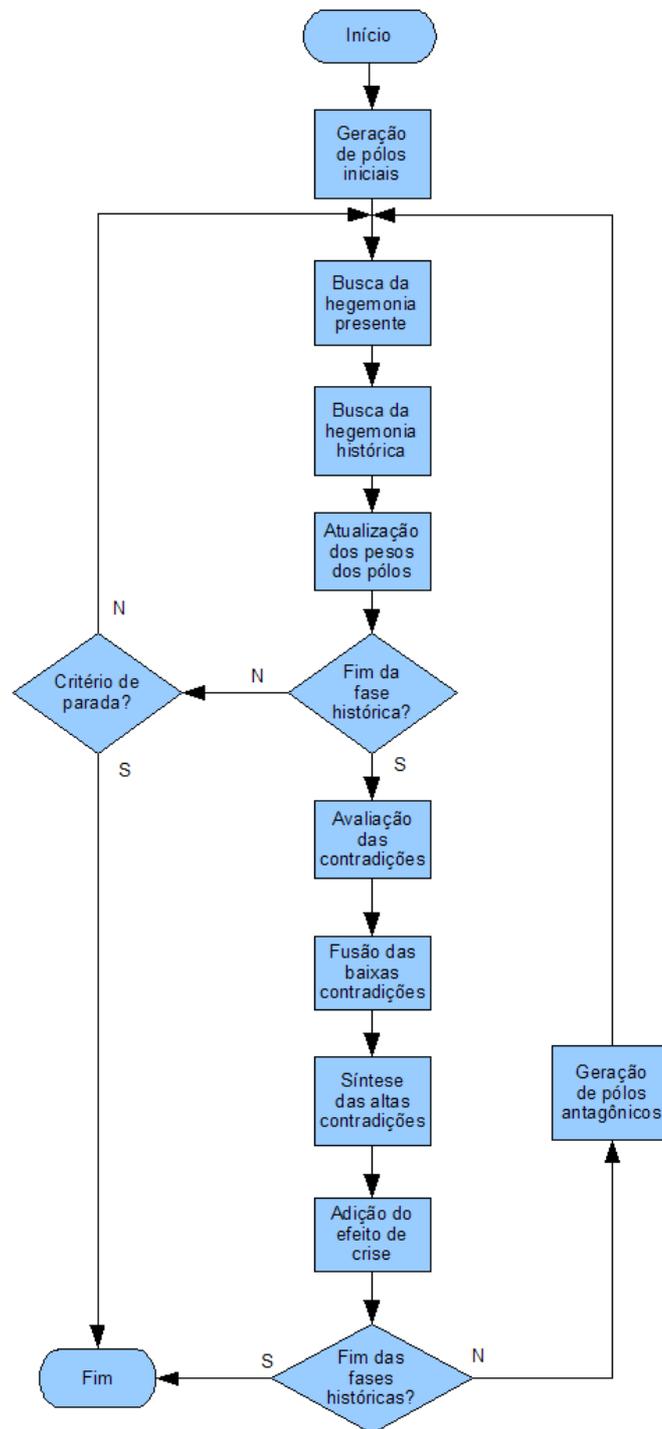
A figura 4.9 exibe o fluxograma geral do método dialético objetivo (ODM) adaptado a busca e otimização de forma simplificada. Embora seja relativamente mais complexo do que o fluxograma correspondente aos algoritmos de otimização por enxame de partículas (PSO), o ODM possui a propriedade de ter custo computacional decrescente em cada fase histórica, uma vez que, à medida que a solução se aproxima da solução ótima, os pólos se aproximam e, portanto, as contradições entre si diminuem, o que faz com que diversos pólos sejam fundidos.

## 4.6 Análise pelo Método da Máxima Entropia

Um método de busca e otimização será tão bom quanto boa for a sua capacidade de descobrir pontos no espaço de busca onde o valor da função objetivo é melhor do que o melhor antes obtido. Outra característica importante é que um algoritmo de busca e otimização será também tão bom quanto for garantida a sua convergência, que deve de preferência convergir para um ponto onde a função objetivo é ótima ou quase ótima (um máximo/mínimo global ou um máximo/mínimo local). Tem-se assim dois critérios básicos de avaliação de um algoritmo de busca e otimização:

1. Capacidade de exploração, ou capacidade de descobrir de novos pontos ótimos (pontos de máximo ou de mínimo, dependendo do problema);
2. Estabilidade e convergência.

Diversos trabalhos buscam analisar a estabilidade e a convergência de um algoritmo de busca (algoritmos genéticos, programação evolucionária e otimização por enxame de partículas, por exemplo) usando desde critérios de estabilidade assintótica derivados do Controle de Processos [118, 119, 107], dado que um algoritmo pode ser considerado um processo dinâmico, até a modelagem do algoritmo de busca como uma cadeia de Markov [120, 121, 122, 123, 124, 125, 126, 94], uma vez que o espaço de busca é finito e contável, dadas as limitações



**Figura 4.9** Fluxograma geral do ODM adaptado a busca e otimização

de representação numérica em sistemas computacionais, sendo o ponto de máximo/mínimo considerado um estado da cadeia.

Neste trabalho buscou-se usar uma abordagem um tanto diferente: foi utilizado o Princípio da Máxima Entropia para gerar uma versão modificada do algoritmo, que será comparada com a versão canônica do método dialético de busca e otimização aqui apresentada. A análise foi focada nos estágios de evolução, o que *a priori* considera fases históricas relativamente grandes. A escolha de focar as etapas de evolução e não de crise revolucionária se deu baseada no fato de que a etapa de crise revolucionária não é qualitativamente diferente do que ocorre nos métodos de busca e otimização baseados em algoritmos genéticos e em programação evolucionária, a não ser pelo processo de fusão de pólos semelhantes (baixa contradição entre si) e de geração de novos pólos a partir das contradições principais.

Para aplicar o Princípio da Máxima Entropia no método dialético de busca e otimização é necessário partir da expressão de atualização dos pesos dos pólos. Para simplificar, aqui se utilizará uma versão onde o passo histórico e o passo contemporâneo são constantes.

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \Delta \mathbf{w}_{C,i}(t) + \Delta \mathbf{w}_{H,i}(t), \quad (4.81)$$

para

$$\Delta \mathbf{w}_{C,i}(t) = \eta_C(t)(1 - \mu_{C,i}(t))^2(\mathbf{w}_C(t) - \mathbf{w}_i(t)), \quad (4.82)$$

$$\Delta \mathbf{w}_{H,i}(t) = \eta_H(t)(1 - \mu_{H,i}(t))^2(\mathbf{w}_H(t) - \mathbf{w}_i(t)), \quad (4.83)$$

$$\eta_C(t) = \eta_H(t) = \eta_0, \quad (4.84)$$

$$0 < \eta_0 < 1. \quad (4.85)$$

Quanto à convergência do algoritmo dialético de busca, é necessário que o processo minimize as seguintes quantidades:

$$E_H = \sum_{i=1}^m \mu_{H,i} |f(\mathbf{w}_i) - f_H|, \quad (4.86)$$

$$E_C = \sum_{i=1}^m \mu_{C,i} |f(\mathbf{w}_i) - f_C|. \quad (4.87)$$

Sabendo-se que as pertinências histórica e contemporânea obedecem à seguintes propriedades:

$$\sum_{i=1}^m \mu_{H,i} = 1, \quad (4.88)$$

$$0 \leq \mu_{H,i} \leq 1, \quad (4.89)$$

$$\sum_{i=1}^m \mu_{C,i} = 1, \quad (4.90)$$

$$0 \leq \mu_{C,i} \leq 1, \quad (4.91)$$

é possível perceber aqui um caso onde as medidas de incerteza expressas pelas funções de pertinência são na verdade probabilísticas, conforme a seção 4.3, cabendo portanto o uso da

Entropia de Shannon, que neste caso é equivalente à Entropia de Shapley, dado que na verdade  $\mu_{H,i}$  e  $\mu_{C,i}$  expressam as probabilidades  $p_{H,i}$  e  $p_{C,i}$  de o  $i$ -ésimo pólo coincidir com o pólo hegemônico histórico e com o pólo hegemônico contemporâneo, respectivamente. Assim, com base em funções de pertinência, pode-se definir uma *entropia histórica*  $H(\mu_H)$  e uma *entropia contemporânea*  $H(\mu_C)$ , que podem ser maximizadas em paralelo, para maximizar a capacidade exploratória do algoritmo:

$$H(\mu_H) = - \sum_{i=1}^m \mu_{H,i} \ln \mu_{H,i}, \quad (4.92)$$

$$H(\mu_C) = - \sum_{i=1}^m \mu_{C,i} \ln \mu_{C,i}. \quad (4.93)$$

Assim, o trabalho se reduz a dois problemas de maximização restritos: a) maximizar  $H(\mu_H)$  para  $E_H$  pequeno; b) maximizar  $H(\mu_C)$  para  $E_C$  pequeno. Isso se assemelha à construção de um algoritmo de agrupamento *fuzzy* c-médias usando o Princípio da Máxima Entropia, utilizando o Método dos Multiplicadores de Lagrange [127]. A partir da expressão 4.22, são obtidas as seguintes expressões para as funções de pertinência históricas e contemporâneas:

$$\mu_{H,i} = \frac{\exp(-\lambda_H g_{H,i})}{\sum_{j=1}^m \exp(-\lambda_H g_{H,j})}, \quad (4.94)$$

$$\mu_{C,i} = \frac{\exp(-\lambda_C g_{C,i})}{\sum_{j=1}^m \exp(-\lambda_C g_{C,j})}, \quad (4.95)$$

ou melhor:

$$\mu_{H,i} = \frac{\exp(-\lambda_H |f(\mathbf{w}_i) - f_H|)}{\sum_{j=1}^m \exp(-\lambda_H |f(\mathbf{w}_j) - f_H|)}, \quad (4.96)$$

$$\mu_{C,i} = \frac{\exp(-\lambda_C |f(\mathbf{w}_i) - f_C|)}{\sum_{j=1}^m \exp(-\lambda_C |f(\mathbf{w}_j) - f_C|)}, \quad (4.97)$$

para

$$g_{H,i} = |f(\mathbf{w}_i) - f_H|,$$

$$g_{C,i} = |f(\mathbf{w}_i) - f_C|,$$

onde  $\lambda_H > 0$  e  $\lambda_C > 0$  são os multiplicadores de Lagrange para as funções de pertinência históricas e contemporâneas, respectivamente. Quando  $\lambda_H \rightarrow 0$ , as funções de pertinência histórica tendem a ser iguais; quando  $\lambda_H \rightarrow +\infty$ , a tendência é de que uma das funções de pertinência histórica tenda a quase 1, enquanto as outras tendem a 0, o que significa que um dos pólos tende a convergir rapidamente para o pólo hegemônico histórico [128, 129, 130]. O mesmo vale para  $\lambda_C$  e as funções de pertinência contemporâneas. Uma vez que os pólos hegemônicos não são fixos, dado que se trata de um processo de busca, e nem a quantidade de pólos é fixa, devido à etapa de crise revolucionária, as expressões das funções de pertinência se tornam:

$$\mu_{H,i}(t) = \frac{\exp(-\lambda_H(t) |f(\mathbf{w}_i(t)) - f_H(t)|)}{\sum_{j=1}^{m(t)} \exp(-\lambda_H(t) |f(\mathbf{w}_j(t)) - f_H(t)|)}, \quad (4.98)$$

$$\mu_{C,i}(t) = \frac{\exp(-\lambda_C(t)|f(\mathbf{w}_i(t)) - f_C(t)|)}{\sum_{j=1}^{m(t)} \exp(-\lambda_C(t)|f(\mathbf{w}_j(t)) - f_C(t)|)}. \quad (4.99)$$

Uma vez que  $f_H(t)$  e  $f_C(t)$  não são fixos, não é possível determinar algebricamente os multiplicadores de Lagrange  $\lambda_H(t)$  e  $\lambda_C(t)$ . Uma alternativa interessante é fazer  $\lambda_H(t)$  e  $\lambda_C(t)$  inicialmente pequenos, para iniciar o algoritmo com uma boa capacidade de exploração, dado que as funções de pertinência passam a ter uma característica mais *fuzzy* (ou seja, com transições mais suaves, por exemplo,  $e^{-\lambda_1 y} > e^{-\lambda_2 y}$ , para  $\lambda_1 < \lambda_2$ ,  $\forall y \in \mathbb{R}$ , e  $\lambda_1, \lambda_2 > 0$ ) e, à medida em que os pólos são fundidos, resultado da convergência para um determinado ponto de mínimo,  $\lambda_H(t)$  e  $\lambda_C(t)$  aumentam. Assim, uma boa escolha poderia ser:

$$\lambda_H(t) = \lambda_C(t) = \frac{1}{m(t)}, \quad (4.100)$$

o que dá origem às seguintes expressões:

$$\mu_{H,i}(t) = \frac{\exp\left(-\frac{1}{m(t)}|f(\mathbf{w}_i(t)) - f_H(t)|\right)}{\sum_{j=1}^{m(t)} \exp\left(-\frac{1}{m(t)}|f(\mathbf{w}_j(t)) - f_H(t)|\right)}, \quad (4.101)$$

$$\mu_{C,i}(t) = \frac{\exp\left(-\frac{1}{m(t)}|f(\mathbf{w}_i(t)) - f_C(t)|\right)}{\sum_{j=1}^{m(t)} \exp\left(-\frac{1}{m(t)}|f(\mathbf{w}_j(t)) - f_C(t)|\right)}. \quad (4.102)$$

Comparando as expressões 4.98 e 4.99 com as expressões canônicas 4.75 e 4.74, nesta ordem, é possível notar o seguinte: quanto mais próximo  $f(\mathbf{w}_i)$  for de  $f_H$ , maior a pertinência, ou seja:  $\mu_{H,i} \propto 1/d(f(\mathbf{w}_i), f_H)$ , onde  $d: \mathbb{R}^2 \rightarrow \mathbb{R}_+$  é uma medida de distância, que pode também vir a ser simplesmente o módulo da diferença, caso os argumentos sejam escalares. Assim, um caminho intuitivo é simplesmente normalizar  $\mu_{H,i} = K/d(f(\mathbf{w}_i), f_H)$ , onde  $K \in \mathbb{R}_+$ , por  $\sum_{j=1}^m \mu_{H,j} = \sum_{j=1}^m K/d(f(\mathbf{w}_j), f_H)$ , resultando na expressão canônica 4.75, para  $d(x, y) = |x - y|$ , onde  $x, y \in \mathbb{R}$ . Logo:

$$\lim_{f(\mathbf{w}_i) \rightarrow f_H} \mu_{H,i} = \lim_{f(\mathbf{w}_i) \rightarrow f_H} \frac{1/d(f(\mathbf{w}_i), f_H)}{\sum_{j=1}^m 1/d(f(\mathbf{w}_j), f_H)},$$

que pela Regra de L'Hôpital resulta:

$$\lim_{f(\mathbf{w}_i) \rightarrow f_H} \mu_{H,i} = \lim_{f(\mathbf{w}_i) \rightarrow f_H} \frac{1/d(f(\mathbf{w}_i), f_H)}{1/d(f(\mathbf{w}_i), f_H)} = 1.$$

Uma outra maneira de formular o problema de obter as expressões das funções de pertinência seria fazer  $\mu_{H,i} \propto \exp[-d(f(\mathbf{w}_i), f_H)]$ , o que é outra maneira de expressar que o aumento da distância entre o valor da função objetivo no  $i$ -ésimo pólo e o valor da função no pólo hegemônico histórico corresponde à diminuição da pertinência histórica do  $i$ -ésimo pólo, evitando que o valor exploda para infinito. Assim, normaliza-se

$$\mu'_{H,i} = K \exp[-d(f(\mathbf{w}_i), f_H)]$$

por

$$\sum_{j=1}^m \mu'_{H,j} = \sum_{j=1}^m K \exp[-d(f(\mathbf{w}_j), f_H)],$$

resultando

$$\mu_{H,i} = \frac{\mu'_{H,i}}{\sum_{j=1}^m \mu'_{H,j}},$$

onde  $K \in \mathbb{R}_+$ , resultando na expressão 4.96, para  $\lambda_H = 1$ . Análises semelhantes podem ser feitas para as expressões das funções de pertinência contemporâneas  $\mu_{C,i}$ , para  $1 \leq i \leq m$ , sendo obtidos resultados idênticos.

Logo, existem semelhanças intuitivas entre as expressões canônicas das funções de pertinência históricas e contemporâneas e aquelas expressões obtidas pelo uso do Princípio da Máxima Entropia, o que levanta a necessidade de gerar resultados experimentais que possam comparar as duas abordagens usando uma boa quantidade de funções de teste, conforme Wolpert e Macready [101].

## 4.7 Aplicações em Classificação e Reconhecimento de Padrões

Os métodos de busca e otimização baseados em computação evolucionária podem ser utilizados em diversas aplicações onde é necessário minimizar uma determinada função custo. Uma dessas aplicações é o uso de computação evolucionária para geração de métodos de agrupamento que, por sua vez, podem ser utilizados para implementar classificadores não supervisionados em aplicações como reconhecimento de padrões e classificação de imagens multiespectrais [131].

Assim, o problema de agrupar os elementos do conjunto  $Z = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{n_Z}\}$ , com  $n_Z$  elementos, em  $n_G$  grupos com centróides  $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n_G}\}$  se reduz a minimizar a função

$$J_e = \sum_{i=1}^{n_G} \sum_{\mathbf{z} \in G_i} \frac{d(\mathbf{z}, \mathbf{v}_i)}{n_G n_{G,i}}, \quad (4.103)$$

onde  $d(\mathbf{z}, \mathbf{v}_i)$  é uma medida de distância entre um vetor  $\mathbf{z}$  da amostra e o centróide do  $i$ -ésimo grupo  $\mathbf{v}_i$ , podendo vir a ser a distância euclidiana, por exemplo, enquanto  $n_{G,i}$  é o número de elementos de  $Z$  agrupados no  $i$ -ésimo grupo,  $G_i$ , e  $J_e$  é uma medida do erro de quantização [131, 105]. Os vetores candidatos a soluções são definidos da forma que segue:

$$\mathbf{x} = (\mathbf{v}_1^T, \mathbf{v}_2^T, \dots, \mathbf{v}_{n_G}^T)^T. \quad (4.104)$$

Logo, modelando o problema para a solução utilizando o método dialético objetivo, os pólos podem assumir a seguinte forma, advinda da modelagem para agrupamento e classificação não supervisionada por otimização por enxame de partículas (PSO) [131, 105]:

$$\mathbf{w}_j = (\mathbf{v}_{j,1}^T, \mathbf{v}_{j,2}^T, \dots, \mathbf{v}_{j,n_G}^T)^T, \quad (4.105)$$

para  $1 \leq j \leq m$ . Quanto à função a ser minimizada, evidentemente outras funções objetivo podem ser adotadas, uma vez que há diversas formas de se estimar o erro de quantização, como utilizar índices de fidelidade, por exemplo [132, 131].

A escolha da função objetivo a ser minimizada depende de fatores como a velocidade do processo de agrupamento, que no caso do PSO canônico costuma ser muito mais baixa do que a de algoritmos de agrupamento como o k-médias [33], por exemplo, muito embora a qualidade do agrupamento, e portanto do resultado da classificação não supervisionada, utilizando o PSO, é significativamente maior [131, 105].

Também é possível montar métodos de classificação supervisionada tendo por base o conjunto de treinamento e a otimização de funções obtidas a partir da matriz de confusão, tais como a taxa de acerto global e o índice  $\kappa$  de correlação estatística [133, 134].

## 4.8 Conclusão

O ODM adaptado a busca e otimização é uma proposta de família de algoritmos evolucionários que, apesar de sua inspiração filosófica baseada no método dialético materialista, possui muito em comum com outras abordagens evolucionárias, mas contém também importantes diferenças. Em comum com os algoritmos genéticos, têm-se nas etapas de crise revolucionária diversos operadores que guardam alguma semelhança com as ideias de seleção natural, cruzamento e mutação, tais como a fusão dos pólos de baixa contradição, a geração de pólos síntese e de pólos em contradição antagônica com os pólos existentes, e a adição do efeito de crise. Contudo, ao contrário dos algoritmos genéticos clássicos (pois existem versões de algoritmos genéticos com cromossomos baseados em vetores reais), não é necessário discretizar os vetores candidatos à solução, tal como no processo de formação de cromossomos, permitindo-se a modelagem na forma de vetores de variáveis contínuas.

Enquanto no PSO cada indivíduo (partícula) é representado por dois vetores, um vetor posição e um vetor velocidade, onde o vetor velocidade é atualizado em função da contribuição da melhor posição global e da melhor posição individual ou local (melhor em função do maior valor da função objetivo), tendo por sua vez o vetor posição ajustado em função do vetor velocidade, no ODM cada indivíduo (pólo) é representado por um vetor de pesos, ajustado em função do vetor de pesos que resulta no maior valor histórico da função objetivo (maior valor considerando-se todos os valores passados) e do vetor de pesos associado ao maior valor presente da função objetivo; ou seja, enquanto no PSO os ajustes das propriedades dos indivíduos são feitos em função das contribuições globais e individuais ou locais, o ODM tem as propriedades dos indivíduos ajustadas em função das contribuições histórica e presente. Além disso, a maior diferença entre o ODM e o PSO está no fato de que, no PSO, o tamanho da população permanece constante ao longo de todo o processo, enquanto no ODM novos indivíduos são gerados a cada fase histórica, da mesma forma que indivíduos semelhantes são fundidos em um único indivíduo.

Este capítulo também mostrou que o Princípio da Máxima Entropia pode ser aplicado em situações onde não há modelos probabilísticos conhecidos, mas onde outras entidades fazem o papel de medidas de incerteza, como no caso das funções de pertinência histórica e contemporânea, mostrando que esse princípio pode ser utilizado para construir métodos de busca e

otimização que possam ter sua capacidade exploratória maximizada e sua convergência garantida, e que é possível analisar o desempenho de um algoritmo a partir da comparação de sua versão canônica com a versão obtida a partir da aplicação do Princípio da Máxima Entropia.

No entanto, as expressões obtidas para as funções de pertinência são dependentes de um parâmetro experimental, um multiplicador de Lagrange que não pode ser definido algebricamente, tendo em vista que os valores de hegemonia histórica e contemporânea são dinâmicos e variantes no tempo, embora se estabilizem no tempo, dado que o método converge. Esses valores são variáveis porque se trata aqui de um método de busca e otimização.

Assim, para que possam ser tiradas melhores conclusões, é necessário realizar experimentos de minimização com funções padrão de teste, para que o desempenho da versão do algoritmo construída com o Princípio da Máxima Entropia possa ser comparado com os diversos valores possíveis do multiplicador de Lagrange e com a versão canônica, mantendo-se constantes os outros parâmetros, tais como número inicial de pólos, contradição mínima, contradição máxima, passo contemporâneo, passo histórico, número de fases históricas, tamanho de cada fase histórica e valor limite da função objetivo como critério de parada.



# Avaliação do Método Dialético de Otimização

## 5.1 Introdução

Para testar métodos de busca e otimização é comum se utilizar funções de teste (*benchmark functions*) que servem para emular diversas situações em problemas de otimização. Para se ter uma boa avaliação do método de otimização proposto é importante usar muitas funções nos estudos experimentais, uma vez que muitos estudos empíricos pecam por não experimentar suas propostas de métodos utilizando uma quantidade considerável de funções de teste com comportamentos diferentes [97]. Além disso, sempre há o risco de, ao se utilizar poucas funções de teste, acabar por se otimizar o algoritmo para aquelas funções em específico, e não para um amplo conjunto de problemas [97].

Wolpert e Macready provaram que, sob certas circunstâncias, nenhum algoritmo de busca é, na média, melhor do que os outros em todos os problemas [101]. Assim, se o número de funções de teste é pequeno, torna-se muito difícil fazer uma conclusão geral a respeito do desempenho do algoritmo [101, 97].

Neste trabalho, o Método Dialético de Otimização é avaliado usando 24 funções de teste, que por sua vez modelam problemas unimodais e multimodais, de baixa e alta dimensionalidade, com e sem descontinuidades, para problemas uni-objetivo e multiobjetivo, comumente utilizadas em diversos trabalhos [117, 81, 114, 131, 105, 115, 97]. Para comparação também foram gerados resultados para as mesmas funções usando os algoritmos de programação evolucionária CEP, FEP, LEP e SPMEP, e o algoritmo de otimização por enxame de partículas (PSO) baseado em consciência individual e consciência global.

## 5.2 Funções de Teste

As 24 funções de teste são descritas nas subseções seguintes. As funções de  $f_1$  a  $f_{13}$  são problemas de alta dimensão, sendo que as funções de  $f_1$  a  $f_7$  são unimodais [114, 105, 97]. A função  $f_6$  é a função passo, uma função descontínua com apenas um mínimo e diversos platôs [97]. A função  $f_7$  é a função quártica, ou função de ruído [97]. As funções de  $f_8$  a  $f_{13}$  são funções multimodais de alta dimensão, onde o número de mínimos locais aumenta exponencialmente com a dimensão do problema, consistindo na classe de problemas mais difíceis de serem resolvidos [97]. As funções de  $f_{14}$  a  $f_{23}$  são funções multimodais de baixa dimensão com poucos mínimos locais [97]. Já a função  $f_{24}$  é um problema de otimização multiobjetivo [108].

### 5.2.1 Funções Unimodais

As funções de teste unimodais de alta dimensão são as seguintes, para  $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ :

1. Modelo esférico [117, 119, 107, 135, 97]:

$$f_1(\mathbf{x}) = \sum_{i=1}^{30} x_i^2, \quad (5.1)$$

onde  $-100 \leq x_i \leq 100$ , para  $1 \leq i \leq 30$ , e  $\min(f_1) = f_1(0, \dots, 0) = 0$ .

2. Problema de Schwefel 1 [117, 114, 97]:

$$f_2(\mathbf{x}) = \sum_{i=1}^{30} |x_i| + \prod_{i=1}^{30} |x_i|, \quad (5.2)$$

onde  $-10 \leq x_i \leq 10$ , para  $1 \leq i \leq 30$ , e  $\min(f_2) = f_2(0, \dots, 0) = 0$ .

3. Problema de Schwefel 2 [117, 114, 135, 97]:

$$f_3(\mathbf{x}) = \sum_{i=1}^{30} \left( \sum_{j=1}^i x_j \right)^2, \quad (5.3)$$

onde  $-100 \leq x_i \leq 100$ , para  $1 \leq i \leq 30$ , e  $\min(f_3) = f_3(0, \dots, 0) = 0$ .

4. Problema de Schwefel 3 [117, 97]:

$$f_4(\mathbf{x}) = \max_{1 \leq i \leq 30} \{|x_i|\}, \quad (5.4)$$

onde  $-100 \leq x_i \leq 100$ , para  $1 \leq i \leq 30$ , e  $\min(f_4) = f_4(0, \dots, 0) = 0$ .

5. Função de Rosenbrock generalizada [117, 114, 113, 116, 119, 107, 97]:

$$f_5(\mathbf{x}) = \sum_{i=1}^{29} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2], \quad (5.5)$$

onde  $-30 \leq x_i \leq 30$ , para  $1 \leq i \leq 30$ , e  $\min(f_5) = f_5(1, \dots, 1) = 0$ .

6. Função passo [117, 97]:

$$f_6(\mathbf{x}) = \sum_{i=1}^{30} \left[ \mu \left( x_i - \frac{1}{2} \right) \right]^2, \quad (5.6)$$

onde

$$\mu(x - \Delta) = \begin{cases} -1, & x < -\Delta \\ x, & |x| \leq \Delta \\ 1, & x > \Delta \end{cases}, \quad (5.7)$$

onde  $-100 \leq x_i \leq 100$ , para  $1 \leq i \leq 30$ , e  $\min(f_6) = f_6(0, \dots, 0) = 0$ .

7. Função quártica, ou função de ruído [117, 97]:

$$f_7(\mathbf{x}) = \sum_{i=1}^{30} [ix_i^4 + r_i], \quad (5.8)$$

onde  $r_i$  é um número aleatório uniformemente distribuído no intervalo  $[0, 1]$  e  $-100 \leq x_i \leq 100$ , para  $1 \leq i \leq 30$ , e  $\min(f_7) = f_7(0, \dots, 0) = 0$ .

### 5.2.2 Funções Multimodais

As funções de teste multimodais de alta dimensão são as que seguem:

1. Problema de Schwefel 4 [97]:

$$f_8(\mathbf{x}) = - \sum_{i=1}^{30} x_i \sin(\sqrt{|x_i|}), \quad (5.9)$$

onde  $-500 \leq x_i \leq 500$ , para  $1 \leq i \leq 30$ , e  $\min(f_8) = f_8(420,9687; \dots; 420,9687) = -12569,5$ .

2. Função de Rastrigin generalizada [117, 114, 113, 119, 107, 135, 97]:

$$f_9(\mathbf{x}) = \sum_{i=1}^{30} [x_i^2 - 10 \cos 2\pi x_i + 10], \quad (5.10)$$

onde  $-5,12 \leq x_i \leq 5,12$ , para  $1 \leq i \leq 30$ , e  $\min(f_9) = f_9(0, \dots, 0) = 0$ .

3. Função de Ackley [117, 114, 113, 135, 97]:

$$f_{10}(\mathbf{x}) = -20 \exp \left( -0,2 \sqrt{\frac{1}{30} \sum_{i=1}^{30} x_i^2} \right) - \exp \left( \frac{1}{30} \sum_{i=1}^{30} \cos 2\pi x_i \right) + 20 + e, \quad (5.11)$$

onde  $-32 \leq x_i \leq 32$ , para  $1 \leq i \leq 30$ , e  $\min(f_{10}) = f_{10}(0, \dots, 0) = 0$ .

4. Função de Griewank generalizada [117, 114, 113, 119, 107, 97]:

$$f_{11}(\mathbf{x}) = \frac{1}{4000} \sum_{i=1}^{30} x_i^2 - \prod_{i=1}^{30} \cos \left( \frac{x_i}{\sqrt{i}} \right) + 1, \quad (5.12)$$

onde  $-600 \leq x_i \leq 600$ , para  $1 \leq i \leq 30$ , e  $\min(f_{11}) = f_{11}(0, \dots, 0) = 0$ .

5. Funções penalizadas generalizadas [97]:

$$f_{12}(\mathbf{x}) = \frac{\pi}{30} (v_1 + v_2 + v_3) + \sum_{i=1}^{30} u(x_i, 10, 100, 4), \quad (5.13)$$

$$v_1 = 10 \sin^2(\pi y_1), \quad (5.14)$$

$$v_2 = \sum_{i=1}^{29} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})], \quad (5.15)$$

$$v_3 = (y_{30} - 1)^2, \quad (5.16)$$

$$f_{13}(\mathbf{x}) = 0,1(v_4 + v_5 + v_6) + \sum_{i=1}^{30} u(x_i, 5, 100, 4), \quad (5.17)$$

$$v_4 = \sin^2(3\pi x_1), \quad (5.18)$$

$$v_5 = \sum_{i=1}^{29} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})], \quad (5.19)$$

$$v_6 = (x_{30} - 1)^2 [1 + \sin^2(2\pi x_{30})], \quad (5.20)$$

para

$$y_i = 1 + \frac{1}{4}(x_i + 1), \quad (5.21)$$

$$u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m, & x_i > a \\ 0, & |x_i| \leq a \\ k(-x_i - a)^m, & x_i < -a \end{cases}, \quad (5.22)$$

onde  $-50 \leq x_i \leq 50$ , para  $1 \leq i \leq 30$ , e

$$\min(f_{12}) = f_{12}(1, \dots, 1) = 0$$

e

$$\min(f_{13}) = f_{13}(1, \dots, 1) = 0.$$

As funções multimodais de baixa dimensionalidade e com poucos pólos são as que seguem:

1. Função de Shekel [97]:

$$f_{14}(\mathbf{x}) = \left[ \frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{i,j})^6} \right]^{-1}, \quad (5.23)$$

para

$$[a_{i,j}]_{2 \times 25} = \begin{bmatrix} -32 & -16 & 0 & 16 & 32 & -32 & \dots & 0 & 16 & 32 \\ -32 & -32 & -32 & -32 & -32 & -16 & \dots & 32 & 32 & 32 \end{bmatrix}, \quad (5.24)$$

onde  $-65,536 \leq x_i \leq 65,536$ ,  $1 \leq i \leq 2$  e  $\min(f_{14}) = f_{14}(-32, -32) \approx 1$ .

2. Função de Kovalik [117, 97]:

$$f_{15}(\mathbf{x}) = \sum_{i=1}^{11} \left[ a_i - \frac{x_1(b_i^2 + b_i x_2)}{(b_i^2 + b_i x_3 + x_4)} \right]^2, \quad (5.25)$$

para  $-5 \leq x_i \leq 5$ ,  $1 \leq i \leq 4$ , e  $\min(f_{15}) \approx f_{15}(0, 1928; 0, 1908; 0, 1231; 0, 1358) \approx 0$ . Os coeficientes  $a_i$  e  $b_i$  são dados pela tabela 5.1.

$i$	$a_i$	$b_i^{-1}$
1	0,1957	0,25
2	0,1947	0,5
3	0,1735	1
4	0,1600	2
5	0,0844	4
6	0,0627	6
7	0,0456	8
8	0,0342	10
9	0,0323	12
10	0,0235	14
11	0,0246	16

**Tabela 5.1** Coeficientes da função de Kovalik

3. Função Six-Hump Camel-Back [117, 97]:

$$f_{16}(\mathbf{x}) = 4x_1^2 - 2, 1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4, \quad (5.26)$$

para  $-5 \leq x_1, x_2 \leq 5$ , onde  $\min(f_{16}) = f_{16}(\mathbf{x}^*) = -1,0316285$  e

$$\mathbf{x}^* \in \{(0,08983; -0,7126)^T, (-0,08983; 0,7126)^T\}.$$

4. Função Branin [117, 97, 136]:

$$f_{17}(\mathbf{x}) = \left(x_2 - \frac{5,1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10 \left(1 - \frac{1}{8\pi}\right) \cos x_1 + 10, \quad (5.27)$$

para  $-5 \leq x_1 \leq 10$  e  $0 \leq x_2 \leq 15$ , onde  $\min(f_{17}) = f_{17}(\mathbf{x}^*) = 0,398$  e

$$\mathbf{x}^* \in \{(-3,142; 12,275)^T, (3,142; 2,275)^T, (9,425; 2,425)^T\}.$$

5. Função Goldstein-Price [97, 136]:

$$f_{18}(\mathbf{x}) = v_1(\mathbf{x})v_2(\mathbf{x}), \quad (5.28)$$

$$v_1(\mathbf{x}) = 1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2), \quad (5.29)$$

$$v_2(\mathbf{x}) = 30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2), \quad (5.30)$$

para  $-2 \leq x_1, x_2 \leq 2$ , onde  $\min(f_{18}) = f_{18}(0, -1) = 3$ .

6. Família Hartman [117, 97, 136]:

$$f(\mathbf{x}) = -\sum_{i=1}^4 c_i \exp\left(-\sum_{j=1}^n a_{i,j}(x_j - p_{i,j})^2\right), \quad (5.31)$$

$i$	$a_{i,1}$	$a_{i,2}$	$a_{i,3}$	$c_i$	$p_{i,1}$	$p_{i,2}$	$p_{i,3}$
1	3	10	30	1	0,3689	0,1170	0,2673
2	0,1	10	35	1,2	0,4699	0,4387	0,7470
3	3	10	30	3	0,1091	0,8732	0,5547
4	0,1	10	35	3,2	0,0382	0,5743	0,8828

**Tabela 5.2** Coeficientes da função de Hartman  $f_{19}$

para  $0 \leq x_j \leq 1$ , onde  $1 \leq j \leq n$ .

Para  $n = 3$ , tem-se  $f(\mathbf{x}) = f_{19}(\mathbf{x})$  e  $\min(f_{19}) = f_{19}(\mathbf{x}^*) = -3,86$ , para

$$\mathbf{x}^* = (0,114;0,556;0,852)^T,$$

tendo seus coeficientes  $c_i$ ,  $a_{i,j}$  e  $p_{i,j}$  definidos na tabela 5.2.

Já para  $n = 6$ , tem-se  $f(\mathbf{x}) = f_{20}(\mathbf{x})$  e  $\min(f_{20}) = f_{20}(\mathbf{x}^*) = -3,32$ , para

$$\mathbf{x}^* = (0,201;0,150;0,447;0,275;0,311;0,657)^T,$$

tendo seus coeficientes  $c_i$ ,  $a_{i,j}$  e  $p_{i,j}$  definidos na tabela 5.3.

$i$	$a_{i,1}$	$a_{i,2}$	$a_{i,3}$	$a_{i,4}$	$a_{i,5}$	$a_{i,6}$	$c_i$	$p_{i,1}$	$p_{i,2}$	$p_{i,3}$	$p_{i,4}$	$p_{i,5}$	$p_{i,6}$
1	10	3	17	3,5	1,7	8	1	0,1312	0,1696	0,5569	0,0124	0,8283	0,5886
2	0,05	10	17	0,1	8	14	1,2	0,2329	0,4135	0,8307	0,3736	0,1004	0,9991
3	3	3,5	1,7	10	17	8	3	0,2348	0,1415	0,3522	0,2883	0,3047	0,6650
4	17	8	0,05	10	0,1	14	3,2	0,4047	0,8828	0,8732	0,5743	0,1091	0,0381

**Tabela 5.3** Coeficientes da função de Hartman  $f_{20}$

7. Família Shekel [117, 97, 136]:

$$f(\mathbf{x}) = -\sum_{i=1}^m \left( \sum_{j=1}^4 (x_j - a_{i,j})^2 \right)^{-1}, \quad (5.32)$$

para  $0 \leq x_i \leq 10$ , onde  $1 \leq i \leq 4$ . A função apresenta  $m$  mínimos locais nos pontos

$$\mathbf{x}_i \approx (a_{i,1}, a_{i,2}, a_{i,3}, a_{i,4})^T,$$

que assumem os valores

$$f(\mathbf{x}_i) \approx -\frac{1}{c_i},$$

de forma que

$$\min(f) = \min_{1 \leq i \leq m} \{f(\mathbf{x}_i)\} \approx f(\mathbf{x}^*) \approx -\frac{1}{c_1}$$

e

$$\mathbf{x}^* \approx (a_{1,1}, a_{1,2}, a_{1,3}, a_{1,4})^T.$$

Para as funções  $f_{21}$ ,  $f_{22}$  e  $f_{23}$ ,  $m$  assume os valores 5, 7 e 10, respectivamente. Os coeficientes  $c_i$  e  $a_{i,j}$  são definidos conforme a tabela 5.4.

$i$	$a_{i,1}$	$a_{i,2}$	$a_{i,3}$	$a_{i,4}$	$c_i$
1	4	4	4	4	0,1
2	1	1	1	1	0,2
3	8	8	8	8	0,2
4	6	6	6	6	0,4
5	3	7	3	7	0,4
6	2	9	2	9	0,6
7	5	5	3	3	0,3
8	8	1	8	1	0,7
9	6	2	6	2	0,5
10	7	3,6	7	3,6	0,5

Tabela 5.4 Coeficientes de Shekel

### 5.2.3 Função Multiobjetivo

Eberhart [81] apresenta como uma função de teste para problemas de otimização multiobjetivo o seguinte problema, apresentado por Shi e Krohling [108]: otimizar a seguinte função [111, 112]:

$$f_{26}(\mathbf{x}) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 + 7x_6^2 - 4x_6x_7 - 10x_6 - 8x_7, \quad (5.33)$$

sujeita às condições

$$c_1(\mathbf{x}) = 127 - 2x_1^2 - 3x_2^4 - x_3 - 4x_4^4 - 5x_5^5 \geq 0, \quad (5.34)$$

$$c_2(\mathbf{x}) = 282 - 7x_1 - 3x_2 - 10x_3^2 - x_4 + x_5 \geq 0, \quad (5.35)$$

$$c_3(\mathbf{x}) = 196 - 23x_1 - x_2^2 - 6x_6^2 + 8x_7 \geq 0, \quad (5.36)$$

$$c_4(\mathbf{x}) = -4x_1^2 - x_2^2 + 3x_1x_2 - 2x_3^2 - 5x_6 + 11x_7 \geq 0, \quad (5.37)$$

para  $f_{26} : [-10, 10]^7 \rightarrow \mathbb{R}$ , onde o mínimo global corresponde a:

$$\mathbf{x}^* = (2, 330499; 1, 951372; -0, 477541; 4, 365726; -0, 624487; 1, 038131; 1, 594227)^T,$$

onde  $\min(f_{26}) = f_{26}(\mathbf{x}^*) = 680, 6300573$ .

O problema pode ser modelado como um problema de otimização multiobjetivo, assumindo que se deve encontrar o mínimo das funções  $f_{26}(\mathbf{x})$  e  $g_j(\mathbf{x})$ , para  $1 \leq j \leq 4$  e

$$g_j(\mathbf{x}) = \begin{cases} f_{\inf}, & c_j(\mathbf{x}) \geq 0 \\ f_{\sup}, & c_j(\mathbf{x}) < 0 \end{cases}, \quad (5.38)$$

onde  $f_{\inf} < f_{26}(\mathbf{x}) < f_{\sup}$ ,  $\forall \mathbf{x} \in [-10, 10]^7$ . Neste trabalho,  $f_{\inf} = -1 \times 10^{50}$  e  $f_{\sup} = 1 \times 10^{50}$ .

Apesar de diversos trabalhos proporem a solução de problemas desse tipo usando estratégias de co-evolução [81, 111, 112, 108], neste trabalho se propõe uma abordagem diferente e, acredita-se, mais simples: a redução do problema à minimização de uma única função  $\hat{f}_{26} : [-10, 10]^7 \rightarrow \mathbb{R}$  definida como segue:

$$\hat{f}_{26}(\mathbf{x}) = \max\{f_{26}(\mathbf{x}), \max_{1 \leq i \leq 4} \{g_i(\mathbf{x})\}\}, \quad (5.39)$$

para  $\mathbf{x} \in [-10, 10]^7$ .

$f$	$f_{lim}$	$f$	$f_{lim}$	$f$	$f_{lim}$	$f$	$f_{lim}$
$f_1$	0,001	$f_7$	0,001	$f_{13}$	0,001	$f_{19}$	-3,8
$f_2$	0,001	$f_8$	-12500	$f_{14}$	1,01	$f_{20}$	-3,2
$f_3$	0,001	$f_9$	0,001	$f_{15}$	0,001	$f_{21}$	-9,5
$f_4$	0,001	$f_{10}$	0,001	$f_{16}$	-1,02	$f_{22}$	-9,5
$f_5$	0,001	$f_{11}$	0,001	$f_{17}$	0,40	$f_{23}$	-9,5
$f_6$	0,001	$f_{12}$	0,001	$f_{18}$	3,01	$f_{26}$	716,1

**Tabela 5.5** Valores limite das funções

### 5.3 Resultados

Os resultados foram gerados utilizando os seguintes parâmetros iniciais, para cada uma das 24 funções de teste e repetindo 100 vezes cada experimento: 20, 50 e 100 pólos iniciais; máximo de 500 fases históricas, com duração de 100 gerações cada fase histórica; contradição mínima de 0,1; contradição máxima de 0,9, e efeito de crise máximo de 0,9. Foram testadas as seguintes versões do método dialético de busca e otimização:

- ODM-CAN-V: Versão canônica do método dialético, com passos histórico e contemporâneo decrescentes, com valores iniciais de 0,99;
- ODM-PME-V: Versão do método dialético otimizada pelo Princípio da Máxima Entropia, com passos histórico e contemporâneo decrescentes, com valores iniciais de 0,99;
- ODM-CAN-F: Versão canônica do método dialético, com passos histórico e contemporâneo fixos em 0,5;
- ODM-PME-F: Versão do método dialético otimizada pelo Princípio da Máxima Entropia, com passos histórico e contemporâneo fixos em 0,5.

Para efeito de comparação também foram gerados resultados usando algoritmos de otimização por programação evolucionária, para 20, 50 e 100 pais e um máximo de 50000 gerações. Também foram gerados resultados usando o algoritmo de otimização por enxame de partículas (PSO) com consciência individual e consciência global, com  $\Delta = 0,001$  e fator de inércia  $w = 0,8$ , para 20, 50 e 100 partículas e um máximo de 50000 iterações. Cada experimento foi executado de 30 a 100 vezes. Esses parâmetros foram escolhidos para que esses resultados possam ser comparados àqueles gerados com as versões do método dialético objetivo, tendo sido definidos empiricamente.

Para avaliar o desempenho do método dialético objetivo, são utilizadas diversas estatísticas, a saber: o valor mínimo médio,  $\bar{y}$ ; o desvio médio do valor mínimo,  $\sigma_y$ ; a mediana do valor mínimo,  $m_d(y)$ ; a moda do valor mínimo,  $m_o(y)$ ; o menor valor mínimo,  $y_{min}$ ; o maior valor mínimo,  $y_{max}$ ; o número de iterações médio,  $\bar{k}$ ; o desvio médio do número de iterações,  $\sigma_k$ ; a mediana do número de iterações,  $m_d(k)$ ; a moda do número de iterações,  $m_o(k)$ ; o menor número de iterações,  $k_{min}$ ; o maior número de iterações,  $k_{max}$ ; o número de avaliações da função médio,  $\bar{q}$ ; o desvio médio do número de avaliações,  $\sigma_q$ ; a mediana do número de avaliações,  $m_d(q)$ ; a moda do número de avaliações,  $m_o(q)$ ; o menor número de avaliações,  $q_{min}$ ; o maior número de avaliações,  $q_{max}$ , e a taxa de sucesso,  $\phi$ , definida como o percentual de experimentos

	CEP	FEP	LEP	SPMEP	ODM-CAN-V	ODM-PME-V	PSO
$f_1$	2000000	2000000	2000000	2000000	400	<b>160</b>	1000000
$f_2$	2000000	2000000	2000000	2000000	<b>520</b>	101800	1000000
$f_3$	2000000	2000000	2000000	2000000	<b>540</b>	101800	1000000
$f_4$	2000000	2000000	2000000	2000000	<b>500</b>	101800	1000000
$f_5$	2000000	2000000	2000000	2000000	<b>18752</b>	101800	1000000
$f_6$	2000000	2000000	2000000	2000000	420	<b>40</b>	1000000
$f_7$	2000000	2000000	2000000	80	<b>20</b>	<b>20</b>	1000000
$f_8$	2000000	1120	120	2000000	<b>20</b>	<b>20</b>	200
$f_9$	2000000	2000000	2000000	2000000	<b>400</b>	101800	1000000
$f_{10}$	2000000	2000000	2000000	2000000	500	<b>60</b>	1000000
$f_{11}$	2000000	2000000	2000000	2000000	420	<b>60</b>	1000000
$f_{12}$	2000000	2000000	2000000	2000000	<b>3260</b>	5006	1000000
$f_{13}$	2000000	2000000	2000000	2000000	9408	<b>7004</b>	1000000
$f_{14}$	2000000	1760	2200	13120	<b>160</b>	103000	41740
$f_{15}$	<b>47640</b>	141360	2000000	57600	101800	101800	1000000
$f_{16}$	1200	3720	920	800	160	<b>20</b>	80
$f_{17}$	1120	3960	2680	26960	200	<b>20</b>	20060
$f_{18}$	145520	59920	41880	120	220	<b>80</b>	1360
$f_{19}$	40	78680	71880	1360	80	<b>20</b>	<b>20</b>
$f_{20}$	2000000	2000000	2000000	14960	120	<b>20</b>	480
$f_{21}$	2000000	473440	436400	2000000	101800	<b>20</b>	1000000
$f_{22}$	337120	746440	464600	2000000	180	<b>20</b>	1000000
$f_{23}$	70800	94640	314960	1262200	180	<b>20</b>	1000000
$f_{26}$	<b>3560</b>	10720	10680	10520	101800	101800	1000000

**Tabela 5.6** Moda do número de avaliações da função objetivo, considerando uma população inicial de 20 indivíduos

onde o valor mínimo atingido é menor do que o valor limite,  $f_{lim}$ , estabelecido como precisão do processo de busca e otimização e, portanto, como critério de parada, ou seja,  $f(\mathbf{x}) < f_{lim}$ . Nas tabelas 5.6, 5.7 e 5.8 são apresentados os resultados da moda do número de avaliações da função objetivo,  $m_o(k)$ , para populações iniciais de 20, 50 e 100 indivíduos, nesta ordem; nas tabelas 5.9, 5.10 e 5.11 são mostrados os resultados da taxa de sucesso,  $\phi$ , para populações iniciais de 20, 50 e 100 candidatos à solução, respectivamente. Os melhores valores estão em negrito. Essas tabelas desconsideram os resultados obtidos com os métodos dialéticos com passos fixos. As tabelas com os resultados completos se encontram no apêndice A.

## 5.4 Discussão e Conclusões

Para o caso da minimização de funções unimodais de alta dimensão ( $f_1$  a  $f_7$ ), observando o resultado quanto à moda do número de iterações e a taxa percentual de sucesso (tabelas A.89, A.105, A.121 e A.137), além do número de avaliações da função no processo de minimização (tabelas A.93, A.109, A.125 e A.141), pode-se perceber que o método ODM-CAN-V apresentou resultados muito bons para 20, 50 e 100 pólos iniciais, atingindo os limiares de mínimo com poucas iterações a não ser pela função de Rosenbrock ( $f_5$ ), que demanda boa capacidade de busca local, mas que mesmo assim atingiu 100% de sucesso, sendo os únicos casos onde

	CEP	FEP	LEP	SPMEP	ODM-CAN-V	ODM-PME-V	PSO
$f_1$	5000000	5000000	5000000	5000000	450	<b>150</b>	2500000
$f_2$	5000000	5000000	5000000	5000000	<b>550</b>	104800	5000000
$f_3$	5000000	5000000	5000000	5000000	<b>550</b>	104800	5000000
$f_4$	5000000	5000000	5000000	5000000	550	<b>150</b>	5000000
$f_5$	5000000	5000000	5000000	5000000	<b>26632</b>	104800	5000000
$f_6$	5000000	5000000	5000000	5000000	450	<b>100</b>	5000000
$f_7$	5000000	5000000	5000000	200	<b>50</b>	<b>50</b>	5000000
$f_8$	5000000	1500	700	5000000	<b>50</b>	<b>50</b>	400
$f_9$	5000000	5000000	5000000	5000000	<b>400</b>	104800	5000000
$f_{10}$	5000000	5000000	5000000	5000000	500	<b>150</b>	5000000
$f_{11}$	5000000	5000000	5000000	5000000	<b>450</b>	5202	5000000
$f_{12}$	5000000	5000000	5000000	5000000	<b>7604</b>	9408	5000000
$f_{13}$	5000000	5000000	5000000	5000000	11466	<b>7418</b>	5000000
$f_{14}$	5000000	3100	4500	3800	150	<b>50</b>	150
$f_{15}$	102100	358700	539300	<b>39500</b>	104800	104800	5000000
$f_{16}$	2300	2600	1400	2300	100	<b>50</b>	150
$f_{17}$	3800	4800	5400	14500	200	<b>50</b>	500
$f_{18}$	11400	18100	85200	42500	200	<b>50</b>	1350
$f_{19}$	100	100	100	1300	<b>50</b>	<b>50</b>	150
$f_{20}$	5000000	5000000	5000000	22500	100	<b>50</b>	19600
$f_{21}$	719800	108900	729800	751300	200	<b>50</b>	5000000
$f_{22}$	40200	367800	300900	100600	200	<b>50</b>	5000000
$f_{23}$	139100	142000	96900	1775700	200	<b>50</b>	5000000
$f_{26}$	<b>7500</b>	15800	18200	14800	104800	104800	5000000

**Tabela 5.7** Moda do número de avaliações da função objetivo, considerando uma população inicial de 50 indivíduos

foram atingidas taxas de sucesso menores do que 94% para a minimização da função  $f_2$  com 20 pólos, tendo sido atingida a taxa de sucesso de 88%, o que pode ser considerado também bastante bom; quanto ao método ODM-PME-V, enquanto os resultados para as funções  $f_1$  e  $f_6$  foram muito piores do que aqueles obtidos com o ODM-CAN-V, ao mesmo tempo em que o resultado para  $f_7$  foi igualmente muito bom, os resultados para as funções  $f_2$ ,  $f_3$  e  $f_5$  foram muito piores para todas as configurações utilizadas (20, 50 e 100 pólos iniciais), o que inclui a função  $f_4$  para 20 pólos iniciais.

No entanto, quando se observam as tabelas A.97 e A.101, pode-se perceber que há uma tendência a se atingir os mínimos das citadas funções, dada a proximidade dos valores mínimos obtidos dos mínimos absolutos das funções, o que indica que à maximização da entropia corresponde uma mais lenta convergência para o caso dessas funções unimodais de alta dimensão, mas há uma tendência que pode indicar a garantia da convergência; já os resultados com o método ODM-CAN-F indicam que houve em geral uma grande piora no desempenho em relação ao ODM-CAN-V quanto ao uso de 20 pólos iniciais, tanto em relação à moda do número de iterações e do número de avaliações da função objetivo, quanto à taxa de sucesso, como resultado da fixação dos passos histórico e contemporâneo, embora tenha ocorrido melhora sensível quando se usam 50 e 100 pólos iniciais.

Já com o ODM-PME-F, todos os resultados para as funções unimodais de alta dimensão em geral pioraram, principalmente os das funções  $f_2$  a  $f_5$ , onde o máximo de iterações foi atingido

	CEP	FEP	LEP	SPMEP	ODM-CAN-V	ODM-PME-V	PSO
$f_1$	1000000	1000000	1000000	1000000	500	<b>200</b>	500000
$f_2$	1000000	1000000	1000000	1000000	<b>600</b>	109800	500000
$f_3$	1000000	1000000	1000000	1000000	600	<b>300</b>	500000
$f_4$	1000000	1000000	1000000	1000000	600	<b>200</b>	500000
$f_5$	1000000	1000000	1000000	1000000	<b>14380</b>	109800	500000
$f_6$	1000000	1000000	1000000	1000000	500	<b>200</b>	500000
$f_7$	1000000	1000000	1000000	400	<b>100</b>	<b>100</b>	500000
$f_8$	1000000	400	600	1000000	<b>100</b>	<b>100</b>	600
$f_9$	1000000	1000000	1000000	1000000	<b>400</b>	109800	500000
$f_{10}$	1000000	1000000	1000000	1000000	600	<b>200</b>	500000
$f_{11}$	1000000	1000000	1000000	1000000	<b>500</b>	10202	500000
$f_{12}$	1000000	1000000	1000000	1000000	<b>400</b>	14206	500000
$f_{13}$	1000000	1000000	1000000	1000000	109800	<b>10812</b>	500000
$f_{14}$	3800	3800	3200	13000	200	<b>100</b>	195900
$f_{15}$	121600	381000	<b>58200</b>	141200	109800	109800	500000
$f_{16}$	2600	4200	5000	2600	<b>100</b>	<b>100</b>	300
$f_{17}$	9200	3800	2600	1000	200	<b>100</b>	1200
$f_{18}$	16600	85000	10200	400	200	<b>100</b>	800
$f_{19}$	200	200	16400	1800	<b>100</b>	<b>100</b>	200
$f_{20}$	1000000	1000000	1000000	19000	<b>100</b>	<b>100</b>	18200
$f_{21}$	1111200	421800	593400	4991400	200	<b>100</b>	500000
$f_{22}$	116400	151000	215000	261600	200	<b>100</b>	119800
$f_{23}$	478200	137400	683400	1476200	200	<b>100</b>	500000
$f_{26}$	15400	13400	42200	10800	<b>100</b>	<b>100</b>	20800

**Tabela 5.8** Moda do número de avaliações da função objetivo, considerando uma população inicial de 100 indivíduos

na maioria dos experimentos. Assim, o método ODM-CAN-V teve um desempenho superior aos outros métodos para o caso de funções unimodais de alta dimensão, mas o ODM-PME-V indica que o ODM-PME-V introduz uma maior garantia de convergência, apesar de diminuir o desempenho quanto à velocidade de convergência. Contudo, o uso de todos os métodos mostrou ser capaz de atingir os limiares de valores mínimos pelo menos em 6% dos casos.

Observando as tabelas, pode-se perceber que o método ODM-CAN-V possui um excelente desempenho, tanto em relação à taxa de sucesso (100%) quanto ao número de iterações, para o caso de funções unimodais de alta dimensionalidade (funções  $f_1$  a  $f_7$ ). O resultado também é bom para as funções multimodais de alta dimensionalidade e muitos mínimos locais (funções  $f_8$  a  $f_{13}$ ), com taxas de sucesso não inferiores a 85% para 20 pólos iniciais, sendo que esse limite inferior sobe para cerca de 95% quando o número de pólos inicial sobe para 100 pólos, o que evidencia que o aumento do número inicial de pólos pode incrementar a capacidade da busca, podendo gerar um conjunto inicial de candidatos a soluções mais rico.

Das tabelas pode-se perceber que o número de iterações da maioria dos experimentos do método ODM-CAN-V com funções unimodais e multimodais de alta dimensão só é alto para as funções  $f_5$ ,  $f_{12}$  e  $f_{13}$ , respectivamente a função de Rosenbrock (unimodal) e as funções penalizadas generalizadas. A função de Rosenbrock é uma função unimodal caracterizada por ter a vizinhança do ponto de mínimo relativamente fácil de ser alcançada, ou seja, há poucas exigências quanto à busca global, embora o ponto de mínimo em si seja relativamente difícil de

	CEP	FEP	LEP	SPMEP	ODM-CAN-V	ODM-PME-V	PSO
$f_1$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_2$	0	0	0	0	<b>88</b>	38	0
$f_3$	0	0	0	0	<b>94</b>	58	0
$f_4$	0	0	0	0	<b>100</b>	89	0
$f_5$	0	0	0	0	<b>100</b>	59	0
$f_6$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_7$	0	0	0	73,33	<b>100</b>	<b>100</b>	0
$f_8$	0	<b>100</b>	<b>100</b>	0	<b>99</b>	<b>99</b>	<b>100</b>
$f_9$	0	0	0	3,33	<b>96</b>	79	0
$f_{10}$	0	0	0	0	88	<b>92</b>	0
$f_{11}$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_{12}$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_{13}$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_{14}$	59	<b>100</b>	<b>100</b>	<b>100</b>	35	24	<b>100</b>
$f_{15}$	<b>100</b>	<b>96,66</b>	90	<b>100</b>	35	18	30
$f_{16}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{17}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{18}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>97</b>	<b>100</b>
$f_{19}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>98</b>	77	<b>100</b>
$f_{20}$	0	0	0	<b>100</b>	81	62	<b>96,66</b>
$f_{21}$	90	<b>96,66</b>	<b>100</b>	63,33	77	<b>100</b>	80
$f_{22}$	<b>100</b>	<b>100</b>	<b>100</b>	86,66	78	<b>100</b>	83,33
$f_{23}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>96,66</b>	<b>100</b>	93	73,33
$f_{26}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	15	14	66,66

**Tabela 5.9** Taxa de sucesso (%), considerando uma população inicial de 20 indivíduos

ser atingido, exigindo uma boa capacidade de busca local. Isso mostra que a etapa de evolução tem um papel fundamental na busca da solução, podendo o ponto de mínimo ser encontrado já nessa etapa, sem a necessidade de outras fases históricas.

Já as funções penalizadas generalizadas possuem muitos mínimos locais razoavelmente fortes quanto ao seu poder de atração da solução buscada. Os bons resultados de taxa de sucesso com essas funções, aliados ao número de iterações total utilizado, mostra que a fase de crise revolucionária é importante para gerar a diversidade sobre os pólos necessária para que a busca realizada pela etapa de evolução seja eficaz. Contudo, observando os resultados de todos os métodos para a função penalizada generalizada  $f_{13}$  quanto ao número inicial de pólos, pode-se observar que o aumento de pólos pode vir a ter um resultado pernicioso quanto à capacidade da busca, que pode vir a ser mais custosa, como pode ser observado para o uso de 100 pólos iniciais.

As tabelas A.81, A.82, A.83, e A.85, A.86, A.87, em particular o menor valor da faixa de valores mínimos encontrados, mostram que, para os experimentos com 20, 50 e 100 pólos, o mínimo é atingido em pelo menos uma das 100 vezes em que cada experimento é rodado, para funções unimodais de alta dimensão ou multimodais de alta dimensão e uma quantidade exponencial de mínimos locais. Além disso, a convergência para o ponto de mínimo é evidenciada pelas médias amostrais expressas nas referidas tabelas, que mostram a tendência do valor mínimo esperado ao valor mínimo exato.

As funções de  $f_{14}$  a  $f_{23}$  têm como características serem funções de baixa dimensionalidade

	CEP	FEP	LEP	SPMEP	ODM-CAN-V	ODM-PME-V	PSO
$f_1$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_2$	0	0	0	0	<b>96</b>	45	0
$f_3$	0	0	0	0	<b>100</b>	66	0
$f_4$	0	0	0	0	<b>100</b>	<b>95</b>	0
$f_5$	0	0	0	0	<b>100</b>	37	0
$f_6$	0	0	0	0	<b>100</b>	<b>98</b>	0
$f_7$	0	0	0	<b>96,66</b>	<b>100</b>	<b>100</b>	0
$f_8$	0	<b>100</b>	<b>100</b>	0	<b>100</b>	<b>99</b>	<b>100</b>
$f_9$	0	0	0	3,33	<b>99</b>	88	0
$f_{10}$	0	0	0	0	<b>98</b>	<b>96</b>	0
$f_{11}$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_{12}$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_{13}$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_{14}$	<b>94</b>	<b>100</b>	<b>100</b>	<b>100</b>	74	28	<b>100</b>
$f_{15}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	38	24	20
$f_{16}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{17}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{18}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{19}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{20}$	1	0	3,33	<b>100</b>	<b>94</b>	63	<b>100</b>
$f_{21}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>96,66</b>	87	<b>100</b>	<b>100</b>
$f_{22}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	84	<b>100</b>	86,66
$f_{23}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>99</b>	<b>94</b>	90
$f_{26}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	45	40	86,66

**Tabela 5.10** Taxa de sucesso (%), considerando uma população inicial de 50 indivíduos

e poucos mínimos locais, porém fortemente atrativos. Observando as tabelas, pode-se perceber o comportamento do método ODM-CAN-V na minimização dessas funções, com uma grande queda da taxa de sucesso das funções  $f_{14}$  e  $f_{15}$ , ou seja, para a função de Shekel e de Kovalik, nesta ordem, e uma razoável queda da taxa de sucesso das funções  $f_{20}$ ,  $f_{21}$  e  $f_{22}$ , ou seja, para a função de Hartman e para os dois primeiros membros da família Shekel, respectivamente.

Para a função  $f_{14}$ , nota-se que, com 20 pólos iniciais, a maioria dos experimentos com o método ODM-CAN-V atingiu o máximo de iterações e a taxa de sucesso foi inferior a 40%; no entanto, a taxa de sucesso praticamente dobra com 50 pólos, e chega a passar de 90%, ao mesmo tempo em que o número de iterações da maioria dos experimentos cai bruscamente, mostrando que a minimização da função  $f_{14}$  pelo método ODM-CAN-V é altamente dependente da população inicial de pólos, particularmente de sua quantidade e, portanto, da capacidade de busca inicial do algoritmo. Já as baixas taxas de sucesso (pouco mais de 30%) e o uso da quantidade máxima de iterações (50000) mostram que o método ODM-CAN-V não conseguiu minimizar a função  $f_{15}$  para a grande maioria dos casos; contudo, os resultados das tabelas A.83 e A.87 mostram que existe uma tendência ao ponto de mínimo, dada a proximidade dos valores esperados e dos valores mínimos encontrados a valores abaixo do valor limite, o que exigiria o aumento do número máximo de iterações, ou seja, o aumento da duração das fases históricas e o aumento da quantidade das mesmas fases.

O desempenho do ODM-CAN-V para minimizar a função de Hartman os primeiros dois membros da família Shekel pode ser evidenciado pelo sensível aumento da taxa de sucesso de

	CEP	FEP	LEP	SPMEP	ODM-CAN-V	ODM-PME-V	PSO
$f_1$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_2$	0	0	0	0	<b>100</b>	30	0
$f_3$	0	0	0	0	<b>100</b>	75	0
$f_4$	0	0	0	0	<b>100</b>	<b>98</b>	0
$f_5$	0	0	0	0	<b>100</b>	32	0
$f_6$	0	0	0	6,66	<b>100</b>	<b>100</b>	0
$f_7$	0	0	0	<b>100</b>	<b>100</b>	<b>100</b>	3,33
$f_8$	0	<b>100</b>	<b>100</b>	0	<b>100</b>	<b>100</b>	<b>100</b>
$f_9$	0	0	0	13,33	<b>100</b>	<b>94</b>	0
$f_{10}$	0	0	0	3,33	<b>100</b>	<b>95</b>	0
$f_{11}$	0	0	0	3,33	<b>100</b>	<b>100</b>	0
$f_{12}$	0	0	0	0	<b>100</b>	<b>100</b>	0
$f_{13}$	0	0	0	0	<b>97</b>	<b>100</b>	0
$f_{14}$	<b>99</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>97</b>	32	<b>100</b>
$f_{15}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	43	39	20
$f_{16}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{17}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{18}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{19}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$f_{20}$	3,33	3,33	0	<b>100</b>	<b>99</b>	86	<b>100</b>
$f_{21}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>96,66</b>	<b>91</b>	<b>100</b>	76,66
$f_{22}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	89	<b>100</b>	<b>96,66</b>
$f_{23}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>99</b>	<b>98</b>	90
$f_{26}$	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	62	54	<b>100</b>

**Tabela 5.11** Taxa de sucesso (%), considerando uma população inicial de 100 indivíduos

acordo com o aumento do número de pólos, o que resulta em uma grande queda do número de iterações para um número bastante baixo. No entanto, a taxa de sucesso para 20, 50 e 100 pólos oscila em torno de 80%, ao mesmo tempo em que o número de iterações cai bruscamente, o que mostra que o ODM-CAN-V é bastante sensível a mínimos locais fortemente atratores, tais como os presentes na função de Hartman e na família Shekel. A tese de que a taxa de sucesso é reduzida por causa da atração por mínimos locais fica evidente pelos resultados da tabela A.87, onde os valores máximos acabam por coincidir com o maior valor da função objetivo em mínimos locais.

Da mesma forma que para as funções unimodais de alta dimensão e multimodais de alta dimensão e número de mínimos locais exponencial (funções de  $f_1$  a  $f_7$ , e de  $f_8$  a  $f_{13}$ , respectivamente), os resultados das tabelas A.83 e A.87 mostram que os valores limites são atingidos para todas as funções multimodais de baixa dimensão e poucos mínimos locais (funções de  $f_{14}$  a  $f_{23}$ ), em pelo menos uma das 100 vezes em que cada experimento de 20, 50 e 100 pólos é rodado.

Quanto à minimização das funções unimodais de alta dimensão (de  $f_1$  a  $f_7$ ), o desempenho do método ODM-PME-V foi bastante diferente daquele obtido com o ODM-CAN-V: enquanto o resultado com a função de modelo esférico  $f_1$  foi similar para os dois métodos, para a função  $f_2$  o desempenho caiu para menos da metade, no que se refere à taxa de sucesso, enquanto a moda do número de iterações atingiu o máximo; contudo, as tabelas A.97 e A.101 mostram que os mínimos encontrados tendem ao valor mínimo absoluto da função  $f_2$ , evidenciando que

a convergência para o ponto de mínimo se tornou mais lenta no método ODM-PME-V, o que parece mostrar que a aplicação do Princípio da Máxima Entropia reduziu a capacidade de busca global, mas apenas com esses dados nada pode ainda ser afirmado a respeito da busca local. O mesmo pode ser afirmado a respeito da minimização das funções  $f_3$  e  $f_5$ . Um fato importante a ser notado da tabela A.101 é que os pontos de mínimo de todas as funções unimodais de alta dimensão foram atingidos para 20, 50 e 100 pólos, como atestam os menores pontos de mínimo nas tabelas.

Para as funções multimodais de alta dimensão e quantidade exponencial de mínimos locais, os resultados dos dois métodos, ODM-CAN-V e ODM-PME-V, foram similares, à exceção daqueles obtidos da minimização das funções  $f_9$  e  $f_{13}$ , respectivamente a função de Rastrigin generalizada e a função penalizada generalizada 2. Para a função de Rastrigin, o método ODM-PME-V encontrou resultados muito próximos do mínimo da função, mas a convergência foi bem mais lenta do que aquela obtida com o método ODM-CAN-V, o que fez com que o máximo de iterações fosse atingido para grande parte dos experimentos, fazendo com que ao final houvesse valores ainda maiores do que o limiar da função objetivo, embora já tendendo ao valor de mínimo.

Já em relação à função penalizada generalizada 2, percebe-se das tabelas A.106 e A.90 que, para 20, 50 e 100 pólos iniciais, a convergência foi atingida para a maioria dos experimentos com 2602, 1309 e 506 iterações, nesta ordem, usando o método ODM-PME-V, enquanto o número de iterações foi bem mais alto para o ODM-CAN-V: 3804, 3333 e 50000, o que mostra que a capacidade exploratória do ODM-PME-V é muito maior do que a do ODM-CAN-V, dado que a função  $f_{13}$  possui mínimos locais razoavelmente fortes. No entanto, o ODM-PME-V atingiu os pontos de mínimo de todas as funções multimodais de alta dimensão e muitos mínimos locais, para 20, 50 e 100 pólos, como se nota da tabela A.102.

A grande diferença entre os métodos ODM-CAN-V e ODM-PME-V reside na minimização de funções multimodais de baixa dimensionalidade e mínimos locais em pequena quantidade, mas fortemente atrativos (funções de  $f_{14}$  a  $f_{23}$ ). Para as funções de  $f_{16}$  a  $f_{23}$ , os resultados dos dois métodos foram praticamente idênticos. Contudo, há grande vantagem do método ODM-PME-V sobre o ODM-CAN-V quanto ao número de iterações, conforme tabelas A.107 e A.91: para as funções  $f_{16}$ ,  $f_{17}$  e de  $f_{19}$  a  $f_{23}$ , o ponto de mínimo é atingido com uma única iteração já com apenas 20 pólos iniciais, enquanto o mínimo de  $f_{18}$  é atingido com 4 iterações para 20 pólos, usando o método ODM-PME-V; para 50 e 100 pólos, os pontos de mínimo de todas as funções citadas são atingidos com apenas uma única iteração. Com o ODM-CAN-V, são dispendidas de 6 a 11 iterações para 20 pólos, ocorrendo um pico de 50000 iterações para a função  $f_{21}$ , e o número de iterações se reduz para 1 ou duas iterações apenas quando se usam 100 pólos iniciais.

Além disso, da tabela A.87, observando a coluna de maiores valores mínimos encontrados, pode-se perceber que, usando o método ODM-CAN-V, acabou-se por cair em mínimos locais, o que não acontece com o método ODM-PME-V, cuja convergência ocorreu para a maioria dos experimentos com uma única iteração, conforme tabelas A.103 e A.107. Isso mostra que o aumento da capacidade exploratória do método ODM-PME-V ficou evidente em relação ao ODM-CAN-V, uma vez que o primeiro conseguiu escapar dos mínimos locais, ao contrário do segundo. Das tabelas A.103 e A.87, além das tabelas A.107 e A.91 pode-se perceber também

que, em relação à minimização das funções  $f_{14}$  a  $f_{15}$ , as diferenças do método ODM-PME-V em relação ao ODM-CAN-V se devem na verdade à convergência mais lenta do ODM-PME-V em relação ao ODM-CAN-V, mais evidente comparando as tabelas A.103 e A.87, e A.99 e A.83, dado que essas duas funções representam problemas onde excelência na busca global é preferível a uma alta capacidade de busca local. Aqui também é importante notar que, como para as funções anteriores, para 20, 50 e 100 pólos, os mínimos absolutos foram atingidos pelo menos uma vez, para todas as funções multimodais com poucos e fortes mínimos locais, usando o método ODM-PME-V.

Considerando os resultados obtidos com os métodos dialéticos de passos fixos (ODM-CAN-F e ODM-PME-F), não houve na média melhorias consideráveis em relação às versões com passos decrescentes, conforme as respectivas tabelas, a não ser pelo fato de que os resultados dos experimentos com 20 pólos foram quase sempre bem piores do que aqueles obtidos com suas versões com passos decrescentes.

De todos os resultados obtidos pode-se concluir que a versão do método dialético de busca e otimização gerada pela introdução do Princípio da Máxima Entropia tem sua capacidade de busca global reduzida em relação à versão canônica do método dialético de busca. Contudo, os resultados indicaram também que a convergência é garantida para uma quantidade finita de iterações, uma vez que, mesmo nas situações onde a quantidade máxima de iterações é atingida, os valores mínimos obtidos já são bem próximos dos mínimos globais. Além disso, a introdução do Princípio da Máxima Entropia fez com que a capacidade exploratória do algoritmo fosse incrementada, o que tornou o método dialético de busca muito mais robusto em relação a mínimos locais fortemente atrativos, permitindo não só que se escapasse deles, mas também que a convergência fosse atingida com apenas uma única iteração para a maioria dos experimentos.

Comparando as versões canônica e com entropia maximizada do método dialético com passos variáveis com os métodos baseados em programação evolucionária, pode-se perceber que, quanto à taxa de sucesso, para uma população inicial de 20 indivíduos, o método ODM-CAN-V atingiu os melhores resultados para as funções de  $f_1$  a  $f_6$ , tendo empatado com o método ODM-PME-V para  $f_7$  com sucesso de 100% para as funções de  $f_1$  a  $f_7$ , exceto para as funções  $f_2$  (88%) e  $f_3$  (94%). Dos métodos de programação evolucionária o único que consegue atingir um certo sucesso para a função  $f_7$  é o SPMEP, com 73,33%. Para a função  $f_8$  os resultados dos métodos ODM-CAN-V, ODM-PME-V, FEP e LEP foram aproximadamente equivalentes (de 99% a 100%). Para a função  $f_9$  o ODM-CAN-V foi o que atingiu o melhor resultado, 96%, seguido do ODM-PME-V, com 79%. Para a função  $f_{10}$  os resultados do ODM-CAN-V e do ODM-PME-V foram próximos (88% e 92%, respectivamente, contra 0% para os métodos de programação evolucionária), mas das funções  $f_{11}$  a  $f_{13}$  as versões do método dialético atingiram 100% de sucesso contra 0% dos métodos CEP, FEP, LEP e SPMEP. Para a função  $f_{14}$  os métodos FEP, LEP e SPMEP atingiram resultados muito superiores aos outros métodos. Para a função  $f_{15}$  todos os métodos de programação evolucionária foram melhores que os métodos dialéticos. Já para as funções de  $f_{16}$  a  $f_{18}$  todos os métodos retornaram praticamente 100% de sucesso. Já na função  $f_{19}$  apenas o ODM-PME-V teve um resultado distante de 100%: sucesso de 77%. Para a função  $f_{20}$  o método SPMEP foi o melhor, seguido do ODM-CAN-V. Para as funções  $f_{21}$  e  $f_{22}$  os melhores resultados foram atingidos pelos métodos CEP, FEP, LEP e

ODM-PME-V. Já para as função  $f_{23}$  os resultados foram idênticos para todos os métodos (de 93% a 100%). Para a função  $f_{26}$  os resultados foram muito bons usando programação evolucionária, mas pífios com métodos dialéticos.

Comparando os métodos dialéticos com o PSO para uma população inicial de 20 indivíduos, considerando a taxa de sucesso, da função  $f_1$  até a  $f_{14}$  o desempenho do PSO foi semelhante ao dos métodos FEP e LEP, cabendo aqui as mesmas considerações feitas quando da comparação entre os métodos dialéticos e esses métodos. Contudo, o resultado para a função  $f_{15}$  (30%) foi praticamente tão ruim quanto aquele obtido com o ODM-CAN-V (35%), embora um pouco melhor do que o resultado do ODM-PME-V (18%). Entretanto, os métodos ODM-CAN-V e PSO se equiparam para as funções de  $f_{16}$  a  $f_{19}$  (praticamente 100%), sendo o PSO um pouco melhor do que o ODM-CAN-V e bem melhor do que o ODM-PME-V para a função  $f_{20}$  (96,66% contra 81% e 62%, respectivamente). No entanto, tanto o ODM-CAN-V quanto o PSO se mostram mais sensíveis à existência de pontos de mínimo fortemente atrativos em funções multimodais de baixa dimensionalidade do que o ODM-PME-V, como mostram os resultados para as funções de Shekel, de  $f_{21}$  a  $f_{23}$ . O resultado do PSO para a função multiobjetivo  $f_{26}$ , contudo, se mostra muito melhor (66%) do que os resultados obtidos com os métodos dialéticos, embora ainda não possa ser considerado satisfatório.

Para uma população inicial de 50 indivíduos, considerando a taxa de sucesso, para as funções de  $f_1$  a  $f_{13}$  o ODM-CAN-V se mostrou superior a todos os outros métodos, seguido do ODM-PME-V (muito inferior ao ODM-CAN-V para as funções  $f_2$ ,  $f_3$  e  $f_5$  e um pouco inferior para a função  $f_9$ ), embora o SPMEP tenha obtido resultado equivalente para a função  $f_7$ , enquanto os métodos FEP e LEP se equiparam ao ODM-CAN-V para a função  $f_8$ . Para as funções  $f_{14}$  e  $f_{15}$ , os métodos de programação evolucionária foram bem superiores aos métodos dialéticos, embora o ODM-CAN-V tenha atingido um resultado razoável para  $f_{14}$  (74%). De  $f_{16}$  a  $f_{19}$  todos os métodos atingiram resultados equivalentes e muito bons. Para a função  $f_{20}$  apenas os métodos ODM-CAN-V e SPMEP atingiram altas taxas de sucesso, enquanto os outros tiveram resultados muito ruins, a exceção do ODM-PME-V, que atingiu um resultado entre ruim e mediano (63%). Para as funções de  $f_{21}$  a  $f_{23}$  todos os métodos atingiram resultados muito bons, sendo os menores os obtidos com o ODM-CAN-V para as funções  $f_{21}$  e  $f_{22}$  (87% e 84%, respectivamente). Já para a função  $f_{26}$  os métodos dialéticos tiveram resultados um pouco melhores do que os obtidos usando 20 indivíduos, mas ainda bem inferiores ao sucesso de 100% obtidos pelos métodos de programação evolucionária (45% e 40% para ODM-CAN-V e ODM-PME-V, nesta ordem).

Considerando as taxas de sucesso para uma população inicial de 100 indivíduos, observa-se que o comportamento dos métodos de computação evolucionária expressos com 50 indivíduos se mantém basicamente o mesmo, mas o dos métodos dialéticos melhora, de onde resulta que o ODM-CAN-V tem o melhor desempenho para todos os casos, exceto para as funções  $f_{15}$  (onde é bem inferior aos métodos de programação evolucionária: 43%),  $f_{22}$  (ligeiramente inferior: 89%), e  $f_{26}$  (62%, superior ao resultado de 45% obtido com 50 pólos iniciais). Análise semelhante pode ser feita com o ODM-PME-V, de onde se pode concluir que os métodos dialéticos são bastante sensíveis à população inicial de indivíduos candidatos à solução.

Observando os resultados da taxa de sucesso obtidos com populações de 50 e 100 indivíduos observa-se que o comportamento do PSO da função  $f_1$  até a função  $f_{14}$  continua idêntico

ao dos métodos FEP e LEP. Contudo, o resultado do PSO não melhora para a função  $f_{15}$ , mostrando uma certa estagnação, pois os métodos dialéticos sofreram sensível melhora com o aumento da população inicial. No entanto, pode-se perceber claramente que o PSO vai ficando mais robusto em relação aos mínimos locais fortemente atrativos em funções multimodais de baixa dimensão, onde a taxa de sucesso vai aumentando com o aumento do número inicial de indivíduos até chegar próxima a 100%. O mesmo se observa na minimização da função multiobjetivo  $f_{26}$ , onde o PSO supera os métodos dialéticos e se equipara os métodos de programação evolucionária.

Esses resultados mostram que, levando em consideração apenas a taxa de sucesso, o método ODM-CAN-V é o mais adequado para funções unimodais e funções multimodais de alta dimensão e muitos mínimos locais, seguido do método ODM-PME-V, enquanto os métodos de programação evolucionária são no mínimo inadequados para esses tipos de problemas. Pode-se concluir também que o aumento do número de pólos pode tornar os métodos dialéticos também bastante eficientes para lidar com funções multimodais de baixa dimensão e poucos mínimos locais, tornando-os praticamente equivalentes aos métodos de programação evolucionária quanto à taxa de sucesso. Contudo, para a abordagem multiobjetivo como formulada neste trabalho, representada pela função  $f_{26}$ , os métodos dialéticos se mostraram limitados, o que pode sugerir que o problema multiobjetivo seja abordado de maneira diferente, seja reformulando a função multiobjetivo, seja usando conjuntos de otimizadores, como sugere a abordagem de Eberhart e Shi (2007) para resolução de problemas multiobjetivo com múltiplos otimizadores por enxames de partículas [81]. Contudo, isso foge ao escopo deste trabalho.

Considerando que o custo computacional da avaliação da função é muito maior do que o custo computacional associado aos outros procedimentos dos algoritmos de busca e otimização, e que os algoritmos são muito diversos e sua comparação usando apenas iterações não é possível, já que o custo computacional associado a uma iteração vai depender do algoritmo em estudo, é relativamente comum avaliar o custo computacional a partir do número de avaliações da função até que o critério de parada comum seja atingido [101].

Uma maneira relativamente pragmática de se comparar o custo computacional entre algoritmos é usar não a média do número de avaliações da função, e sim a moda, que dá uma noção maior do comportamento dinâmico dos algoritmos comparados. Assim, observando a moda do número de avaliações de cada algoritmo, para 20, 50 e 100 indivíduos iniciais, pode-se perceber claramente que os métodos dialéticos são menos custosos do que os métodos de programação evolucionária para as funções estudadas, a exceção da função  $f_{15}$ , para populações de 20, 50 e 100 indivíduos iniciais, e para a função  $f_{26}$ , para 20 e 50 indivíduos iniciais. Comparando os métodos dialéticos com o PSO, pode-se notar claramente que a situação é bem semelhante: o custo computacional dos métodos dialéticos é tipicamente menor do que o do PSO, à exceção da função  $f_{19}$ , para uma população inicial de 20 indivíduos, onde o desempenho do ODM-PME-V e do PSO são iguais (20 avaliações da função objetivo).

Um fator que influencia no fato de a moda do número de avaliações da função objetivo ser tipicamente menor nos métodos dialéticos do que nos métodos baseados em programação evolucionária e no PSO é que, enquanto a população dos métodos de programação evolucionária na prática é o dobro da população definida no início do algoritmo em cada geração, e a população do PSO é constante e igual à população inicial em cada geração, a população de

candidatos à solução nos métodos dialéticos pode até vir a aumentar de uma geração para a outra, mas tem a tendência de se reduzir até o mínimo de 2 indivíduos, como bem expressam todos os experimentos deste trabalho onde o máximo de iterações (fases históricas e gerações) foi atingido pelos métodos dialéticos: todos esses experimentos resultaram em apenas 2 pólos na última iteração da última fase histórica. Isso acontece porque o método dialético materialista adaptado a busca e otimização é *teleológico*, ou seja, suas fases de evolução são guiadas por um sentido, no caso, a busca do ponto de mínimo da função objetivo e, segundo a Filosofia, nos sistemas dialéticos teleológicos a tendência é de que as contradições do sistema dialético se reduzam a uma única contradição principal, e por conseguinte os pólos integrantes se reduzam a apenas dois [42, 137, 138, 37, 53].

Para as funções multimodais de baixa dimensão e poucos mínimos locais, o ODM-PME-V se mostrou o de mais rápida convergência, o que, comparado ao ODM-CAN-V, pode ser resultado do aumento da capacidade exploratória esperado da introdução do Princípio da Máxima Entropia à definição do Método Dialético Objetivo. Para as funções unimodais e as funções multimodais de alta dimensão e muitos mínimos locais, os métodos ODM-CAN-V e ODM-PME-V apresentaram custo computacional muito próximo, sendo o ODM-CAN-V sutilmente menos custoso. O destaque ficou por conta da função multiobjetivo: os resultados da moda do número de avaliações da função aliados à taxa de sucesso mostraram que, embora a taxa de sucesso seja entre baixa e mediana, os métodos dialéticos, quando convergem para a solução, o fazem com poucas iterações (uma única iteração para  $f_{26}$ ) e, portanto, a um baixo custo computacional.



# Dialética e Quantização Vetorial de Imagens

## 6.1 Introdução

Nos capítulos anteriores foi visto que é possível construir um algoritmo para classificação não supervisionada de imagens baseado no método dialético. A principal vantagem dos assim chamados classificadores dialéticos objetivos está na possibilidade de se realizar uma classificação sem saber previamente o número real de classes presente na imagem, característica compartilhada com diversas redes neurais construtivas. No entanto, também foi visto que é possível usar o método dialético como inspiração para construir métodos de busca e otimização, podendo-se alcançar resultados interessantes quanto à precisão da otimização e ao número de iterações e de avaliações da função dispendidos até se alcançar um determinado valor limiar de otimização.

Contudo, aproveitando o mesmo princípio da otimização de funções, é possível construir algoritmos de classificação não supervisionada e agrupamento usando métodos evolucionários de otimização, onde os vetores candidatos à solução são construídos com parâmetros do classificador (no caso do mapa de k-médias, esses parâmetros seriam compostos pelas coordenadas dos pesos dos centróides); assim, avaliar uma determinada função objetivo em um determinado ponto equivaleria a realizar uma classificação usando os parâmetros do classificador passados pelo vetor candidato à solução e, após isso, avaliar a qualidade da classificação segundo algum item de avaliação da validade do agrupamento, no caso de se otimizar um método de agrupamento, como o k-médias, por exemplo. Assim, é possível usar otimização em substituição ao algoritmo de treinamento do método de agrupamento a otimizar, resultando classificadores com determinadas características otimizadas.

Neste capítulo se busca avaliar o uso dos classificadores dialéticos objetivos quanto à qualidade da quantização, usando para tanto índices de fidelidade como medidas indiretas da distorção de quantização, além de índices de validade de agrupamento. Os resultados são comparados com resultados semelhantes obtidos com classificadores baseados no mapa de k-médias, no mapa *fuzzy* c-médias e no mapa auto-organizado de Kohonen.

Também é proposto um novo método de classificação não supervisionada e agrupamento baseado na otimização do mapa de k-médias usando o método dialético de otimização em função de índices de avaliação da validade do agrupamento. Todos os resultados são gerados a partir da classificação de imagens de ressonância magnética sintéticas e comparados entre si, considerando as diversas versões canônica e com entropia maximizada, e com resultados gerados com classificadores baseados no mapa de k-médias padrão, no mapa auto-organizado de Kohonen, e no mapa *fuzzy* c-médias, usando testes estatísticos.

## 6.2 Materiais e Métodos

O objetivo principal deste capítulo é mostrar que tanto o classificador dialético objetivo (ODC) quanto o método dialético de busca e otimização (ODM) podem ser utilizados para segmentar imagens multiespectrais, sendo o primeiro um método de classificação e segmentação em si, enquanto o último é utilizado para minimizar uma função de custo de um algoritmo de agrupamento padrão aplicado à segmentação de imagens.

A ideia do uso do ODM na classificação é substituir o algoritmo de treinamento do classificador não supervisionado por um processo de otimização usando o ODM para otimizar o resultado de algum índice de avaliação da qualidade do processo de agrupamento (ou *clustering*) e, em seguida, avaliar o resultado utilizando índices de fidelidade que medem a semelhança entre a imagem quantizada obtida da segmentação e a imagem original.

O método de agrupamento otimizado neste trabalho foi o mapa de k-médias, enquanto as imagens adotadas neste estudo de caso consistem em imagens multiespectrais sintéticas de ressonância magnética. Os resultados em seguida são comparados com aqueles obtidos por outros algoritmos de agrupamento, como o k-médias padrão, o *fuzzy* c-médias e o mapa auto-organizado de Kohonen, quanto à distorção de quantização. Para comparação são utilizados testes estatísticos, como o teste de  $\chi^2$ , para avaliar quão próximos são os resultados de quantização do ponto de vista global, ou seja, considerando diversos resultados para índices de fidelidade diferentes.

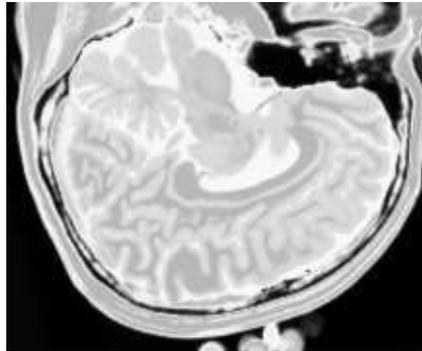
### 6.2.1 Imagens Multiespectrais Sintéticas

No estudo de caso deste capítulo foram utilizadas 1086 imagens multiespectrais sintéticas de ressonância magnética com três bandas, compostas por 3258 imagens sagitais de 1 mm de espessura, resolução de 1 mm<sup>3</sup>, ponderadas em densidade de prótons,  $T_1$  e  $T_2$ , para um cérebro humano normal, obtidas por um sistema tomográfico de ressonância magnética com níveis de ruído de 0%, 1%, 3%, 5%, 7% e 9%, e sem inhomogeneidades de campo. Essas imagens foram geradas pelo simulador de RM BrainWeb, que permite a geração de imagens sintéticas de RM ponderadas em densidade de prótons,  $T_1$  e  $T_2$  variando o nível de ruído e o percentual de inhomogeneidade de campo [19, 20].

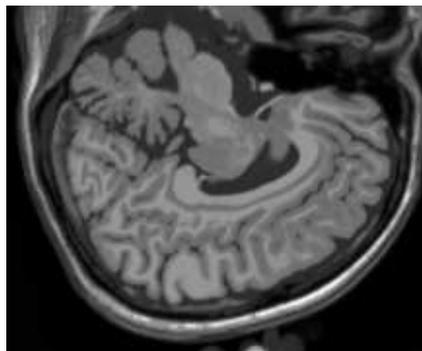
As figuras 6.1 (banda 0), 6.2 (banda 1) e 6.3 (banda 2) mostram a fatia 97 de um volume de imagens sagitais com 181 fatias e 0% de ruído, ponderadas em PD (densidade de próton),  $T_1$  e  $T_2$ , enquanto a figura 6.4 ilustra a composição colorida R0-G1-B2 da mesma fatia. Pode-se notar no topo do crânio, na parte inferior das imagens, a presença de artefatos, que podem ser resultantes de erros no simulador, mas que não são prejudiciais à análise, uma vez que, neste trabalho, não é dada ênfase à análise anatômica.

### 6.2.2 Quantização Vetorial

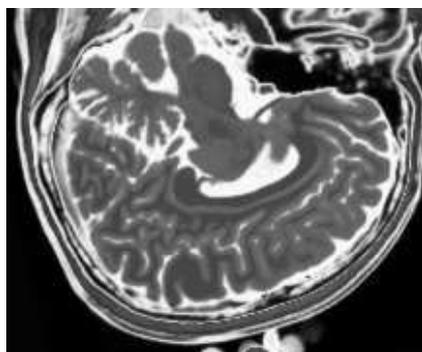
A motivação da Teoria da Quantização Vetorial consiste na redução da dimensionalidade da representação dos dados de entrada, visando a compressão de dados. Assim, dado um vetor de entrada  $\mathbf{x}$ , a função  $\mathbf{c}(\mathbf{x})$  é dita um *codificador* de  $\mathbf{x}$ , enquanto  $\mathbf{x}'(\mathbf{c})$  é um *decodificador* de  $\mathbf{c}$ . Na quantização vetorial,  $\mathbf{c}(\mathbf{x})$  transforma  $\mathbf{x}$  de forma a poder ser representado com menos



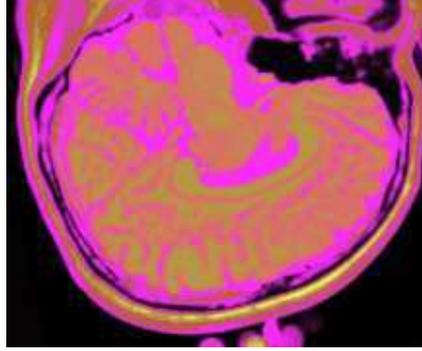
**Figura 6.1** Imagem de RM da fatia 97 ponderada em PD



**Figura 6.2** Imagem de RM da fatia 97 ponderada em  $T_1$



**Figura 6.3** Imagem de RM da fatia 97 ponderada em  $T_2$



**Figura 6.4** Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD,  $T_1$  e  $T_2$

dimensões, ou por meio de uma palavra menor, enquanto  $\mathbf{x}'(\mathbf{c})$  retorna o vetor  $\mathbf{c}$  à mesma base de representação de  $\mathbf{x}$  [33].

Ou seja, dada a transformação não linear  $\Phi : \mathbb{X} \rightarrow \mathbb{A}$ , onde  $\mathbf{x} \in \mathbb{X}$  e  $\mathbb{A}$  é necessariamente um espaço de características discreto, as transformações  $\mathbf{c} = \Phi(\mathbf{x})$  e  $\mathbf{x}' = \Phi^{-1}(\mathbf{c})$  correspondem às tarefas de codificação e decodificação, nesta ordem, em um esquema qualquer de quantização vetorial [33].

Considerando que a entrada  $\mathbf{x}$  possui uma função densidade de probabilidade  $p_{\mathbf{x}}(\mathbf{x})$ , a *distorção esperada* do sistema de quantização será dada pela expressão:

$$D = \frac{1}{2} \int_{\mathbf{x} \in \mathbb{X}} p_{\mathbf{x}}(\mathbf{x}) d(\mathbf{x}, \mathbf{x}') d\mathbf{x}, \quad (6.1)$$

onde  $d(\mathbf{x}, \mathbf{x}')$  é uma medida de distância entre o vetor de entrada  $\mathbf{x}$  e o vetor reconstruído  $\mathbf{x}'$ , que pode ser tanto a distância euclidiana quanto outra medida de distância, como Mahalanobis, por exemplo. Uma escolha bastante comum é o quadrado da distância euclidiana:  $d(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^2 = (\mathbf{x} - \mathbf{x}')^T (\mathbf{x} - \mathbf{x}')$  [33]. Assim:

$$D = \frac{1}{2} \int_{\mathbf{x} \in \mathbb{X}} p_{\mathbf{x}}(\mathbf{x}) \|\mathbf{x} - \mathbf{x}'\|^2 d\mathbf{x}. \quad (6.2)$$

Um sistema de quantização é tão bom quanto menor for a *distorção esperada* e menor for a dimensionalidade de  $\mathbb{A}$ . Obviamente este é um compromisso difícil de atender, pois quanto menor a *distorção*, mais a dimensionalidade de  $\mathbb{A}$  se aproxima da dimensionalidade de  $\mathbb{X}$ , e portanto menor a capacidade de compressão, ao mesmo tempo que, quanto menor a dimensionalidade de  $\mathbb{A}$ , maior a capacidade de compressão sobre o dado reconstruído, mas também maior é a *distorção esperada*.

A quantização *pixel a pixel* de imagens multiespectrais também tem um importante papel quando se deseja comprimir imagens. Na abordagem utilizada neste trabalho, a quantização de imagens multiespectrais coincide com o uso de classificadores não supervisionados baseados em centróides, ou seja, classificadores baseados em métodos de agrupamento de dados (*clustering*). Assim, dado um classificador com  $m$  classes tal que o conjunto de classes seja dado por:

$$\Omega = \{c_1, c_2, \dots, c_m\},$$

onde a cada classe  $c_j$  estão associados um vetor de pesos  $\mathbf{w}_j$  e uma função discriminante  $g_j(\mathbf{x})$ , onde  $1 \leq j \leq m$ , tem-se a seguinte regra, inspirada no critério de decisão de Bayes [133, 134]:

$$g_k(\mathbf{x}) = \max\{g_j(\mathbf{x})\}_{j=1}^n \Rightarrow \mathbf{x}' = \mathbf{w}_k, \quad (6.3)$$

para  $\mathbf{x} = f(\mathbf{u})$ , onde  $f : S \rightarrow [0, 1]^n$  é uma imagem multiespectral normalizada de  $n$  bandas, e  $f' : S \rightarrow [0, 1]^n$  é a versão reconstruída de  $f$ , dado que  $f'(\mathbf{u}) = \mathbf{x}'$ ,  $\mathbf{u} \in S$  e  $1 \leq k \leq m$ .

Como quantizadores vetoriais foram utilizados neste trabalho: mapas de k-médias, mapas auto-organizados de Kohonen com função de vizinhança gaussiana e com função de vizinhança retangular, mapas *fuzzy* c-médias e classificadores dialéticos objetivos. A dimensionalidade do espaço de características codificadas  $\mathbb{A}$  foi avaliada indiretamente pelo número de classes encontradas nos mapas de classificação, enquanto a distorção foi avaliada também indiretamente usando índices de fidelidade global *pixel a pixel*. A classificação não supervisionada também foi avaliada usando medidas de validade de agrupamento.

### 6.2.3 Índices de Fidelidade

Os principais índices de fidelidade de imagem *pixel a pixel* abordados são o *erro máximo*  $\varepsilon_{ME}$  (*Maximum Error*, ME), o *erro médio absoluto*  $\varepsilon_{MAE}$  (*Mean Absolute Error*, MAE), o *erro médio quadrático*  $\varepsilon_{MSE}$  (*Mean Square Error*, MSE), a *raiz do erro médio quadrático*  $\varepsilon_{RMSE}$  (*Root Mean Square Error*, RMSE), o *erro médio quadrático normalizado*  $\varepsilon_{NMSE}$  (*Normalized Mean Square Error*, NMSE), a *relação sinal-ruído de pico*  $\varepsilon_{PSNR}$  (*Peak Signal to Noise Ratio*, PSNR) e a *relação sinal-ruído*  $\varepsilon_{SNR}$  (*Signal to Noise Ratio*, SNR), descritos da forma que segue [132]:

$$\varepsilon_{ME} = \max\{\|f(\mathbf{u}) - f'(\mathbf{u})\|\}_{\mathbf{u} \in S}, \quad (6.4)$$

$$\varepsilon_{MAE} = \frac{1}{\#S} \sum_{\mathbf{u} \in S} \|f(\mathbf{u}) - f'(\mathbf{u})\|, \quad (6.5)$$

$$\varepsilon_{MSE} = \frac{1}{\#S} \sum_{\mathbf{u} \in S} \|f(\mathbf{u}) - f'(\mathbf{u})\|^2, \quad (6.6)$$

$$\varepsilon_{RMSE} = \sqrt{\varepsilon_{MSE}}, \quad (6.7)$$

$$\varepsilon_{NMSE} = \frac{\sum_{\mathbf{u} \in S} \|f(\mathbf{u}) - f'(\mathbf{u})\|^2}{\sum_{\mathbf{u} \in S} \|f(\mathbf{u})\|^2}, \quad (6.8)$$

$$\varepsilon_{PSNR} = 20 \log_{10} \frac{L_{\max}}{\varepsilon_{RMSE}}, \quad (6.9)$$

$$\varepsilon_{SNR} = 10 \log_{10} \frac{1}{\varepsilon_{NMSE}}, \quad (6.10)$$

para  $f : S \rightarrow \{0, 1, \dots, L_{\max}\}^n$ . Para imagens multiespectrais normalizadas  $f : S \rightarrow [0, 1]^n$ , tem-se  $L_{\max} = 1$ .

Existem diversos outros índices para medição de fidelidade de imagens [139, 140, 141], como o índice de Wang e Bovik [142]. No entanto, neste trabalho o principal interesse ao usar índices de fidelidade é aproximar a medida da distorção média associada ao processo de

quantização. Logo, o principal interesse está em índices baseados em diferenças *pixel a pixel*, o que não é o caso do índice de Wang e Bovik e de outros índices baseados na comparação entre estatísticas globais das imagens [142, 132].

#### 6.2.4 Medidas de Validade de Agrupamento

Os índices de validade de agrupamento (*cluster validity indexes*) correspondem a funções matemáticas e estatísticas para avaliar quantitativamente um algoritmo de agrupamento. Em geral, índices de validade de agrupamento servem para determinar experimentalmente tanto o melhor número de agrupamentos (ou, usando a terminologia de classificação de imagens, o número ideal de classes presentes na imagem) quanto a melhor configuração de agrupamento, ou seja, o melhor particionamento dos dados de entrada. Neste caso, os dados de entrada são os *pixels* da imagem a ser classificada ou segmentada. Assim, uma prática comum é que se façam diversos experimentos até que se encontre o melhor número de classes e a melhor configuração de particionamento [131, 139, 140].

Assim, a classificação ou segmentação de uma imagem é tão boa quanto maior for a *coesão* do processo de agrupamento, ou seja, quanto menor for a maior distância dos pixels ao centróide do grupo ao qual foram associados; o processo de agrupamento também tem sua qualidade avaliada pela *separação* entre os centróides dos grupos: quanto maior for a menor distância entre centróides, melhor o agrupamento. Além disso, quanto menor for o *erro de distorção de quantização*, melhor o processo de agrupamento [139, 140, 141, 131].

A seção 4.7 ilustra um índice de avaliação de agrupamento de aplicação geral, que tenta aproximar o erro de distorção, definido da forma que segue, já adaptado para o contexto de segmentação de imagens: o problema de agrupar os *pixels* da imagem  $f : S \rightarrow W^n$ , com altura  $n_H$ , largura  $n_W$  e  $n_B = n$  bandas, resultando  $n_f = n_H \times n_W$  *pixels*, em  $n_G$  grupos (ou classes) com centróides  $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n_G}\}$  se reduz a minimizar a função

$$J_e = \sum_{i=1}^{n_G} \sum_{\forall f(\mathbf{u}) \in G_i} \frac{\|f(\mathbf{u}) - \mathbf{v}_i\|}{n_{G,i}}, \quad (6.11)$$

onde  $\mathbf{u} \in S$  e  $\|f(\mathbf{u}) - \mathbf{v}_i\|$  é a distância entre o *pixel*  $f(\mathbf{u})$  da imagem  $f : S \rightarrow W^n$  e o centróide do  $i$ -ésimo grupo  $\mathbf{v}_i$ , podendo vir a ser a distância euclidiana, por exemplo, enquanto  $n_{G,i}$  é o número de elementos de  $f$  agrupados no  $i$ -ésimo grupo e  $J_e$  é uma medida do erro de quantização [131, 105]. Os vetores candidatos a soluções são definidos da forma que segue:

$$\mathbf{x} = (\mathbf{v}_1^T, \mathbf{v}_2^T, \dots, \mathbf{v}_{n_G}^T)^T. \quad (6.12)$$

A coesão do processo de agrupamento é inversamente proporcional à distância máxima intra-agrupamento (*maximum intra-clusters distance*), ou seja, a maior distância entre os *pixels* e os respectivos centróides dos grupos aos quais foram alocados, sendo dada por [131, 139]:

$$d_{\max}(V) = \max_{j=1,2,\dots,n_G} \left\{ \sum_{\forall f(\mathbf{u}) \in G_j} \frac{\|f(\mathbf{u}) - \mathbf{v}_j\|}{n_{G,j}} \right\} = \max_{j=1,2,\dots,n_G} S_j, \quad (6.13)$$

onde

$$S_j = \sum_{\forall f(\mathbf{u}) \in G_j} \frac{\|f(\mathbf{u}) - \mathbf{v}_j\|}{n_{G,j}}, \quad (6.14)$$

de onde vem que:

$$J_e = \frac{1}{n_G} \sum_{i=1}^{n_G} S_i. \quad (6.15)$$

A menor separação dos centróides no processo de agrupamento, ou distância mínima inter-centróides (*minimum inter-clusters distance*) é importante para avaliar a presença de grupos irrelevantes, que em uma análise posterior podem ser eliminados, sendo fundidos a outros mais próximos. Esse índice é dado pela expressão [131, 139]:

$$d_{\min} = \min_{\forall i,j,i \neq j} \{\|\mathbf{v}_i - \mathbf{v}_j\|\}. \quad (6.16)$$

Pode-se perceber que ao agrupamento ótimo correspondem a maximização da coesão e da separação, junto com a minimização do erro de quantização, o que equivaleria a minimizar  $d_{\max}(V)$ , maximizar  $d_{\min}$  e minimizar  $J_e$  [131]. Omran *et al.* (2005) propõem que é possível construir um novo índice para avaliação do processo de agrupamento de imagens, que pode servir como função objetivo  $f_o : W^{n \times n_G} \rightarrow W$  a ser minimizada por um algoritmo de Computação Evolucionária, a fim de obter um classificador otimizado de acordo com a minimização dessa função objetivo, tendo sido utilizado o índice  $J_o$  que segue, otimizado para gerar um classificador não supervisionado k-médias [131]:

$$f_o(\mathbf{x}_i) = J_o(\mathbf{x}_i) = w_1 d_{\max}(V_i) + w_2 (L_{\max} - d_{\min}(\mathbf{x}_i)) + w_3 J_e(\mathbf{x}_i), \quad (6.17)$$

onde  $\mathbf{x}_i$  é um candidato à solução e  $V_i$  é o conjunto de centróides associados ao respectivo candidato. Já as grandezas  $w_1$ ,  $w_2$  e  $w_3$  são pesos, que podem ser considerados, em uma primeira abordagem, iguais, ou seja,  $w_1 = w_2 = w_3 = 1/3$ , não privilegiando nenhum aspecto da medida quantitativa da qualidade do agrupamento.

Há diversos outros índices de avaliação do agrupamento, tais como o índice DB, definido por [139]:

$$\text{DB}(f, V) = \frac{1}{n_G} \sum_{i=1}^{n_G} \max_{i \neq j} \left\{ \frac{S_i + S_j}{\|\mathbf{v}_i - \mathbf{v}_j\|} \right\}, \quad (6.18)$$

e o índice de Xie-Beni, dado por [141]:

$$\text{XB}(f, V) = \frac{\sum_{i=1}^{n_G} \sum_{\forall \mathbf{u} \in S} \mu_i^2(f(\mathbf{u})) \|f(\mathbf{u}) - \mathbf{v}_i\|^2}{n \min_{i \neq j} \|\mathbf{v}_i - \mathbf{v}_j\|^2}, \quad (6.19)$$

onde

$$\mu_i(f(\mathbf{u})) = \frac{\|f(\mathbf{u}) - \mathbf{v}_i\|^{2/(1-q)}}{\sum_{j=1}^{n_G} \|f(\mathbf{u}) - \mathbf{v}_j\|^{2/(1-q)}}, \quad (6.20)$$

para  $q > 1$ .

Omran *et al.* (2005) desenvolveram um algoritmo de agrupamento baseado na otimização do mapa de k-médias segundo a função objetivo definida na expressão 6.17 utilizando a estratégia de otimização por enxame de partículas (PSO) [131]. Neste trabalho foram gerados resultados utilizando como função objetivo tanto o índice combinado de Omran  $J_o$  quanto apenas o erro de quantização  $J_e$ , otimizando os parâmetros do mapa de k-médias usando o método dialético objetivo de busca e otimização.

### 6.2.5 Métodos de Classificação Não Supervisionada

Os métodos de agrupamento, como os mapas auto-organizados de Kohonen, os mapas de k-médias e *fuzzy* c-médias, também podem ser utilizados como métodos de classificação não supervisionada de imagens multiespectrais. Aliás, é fácil mostrar que as tarefas de quantização vetorial e classificação multiespectral são coincidentes, a diferença é que, na primeira, o interesse está nos pesos dos centróides, enquanto na segunda, nos seus índices.

As imagens multiespectrais sintéticas obtidas pelas composições coloridas R0-G1-B2 foram classificadas usando os seguintes métodos, também utilizados para avaliar o desempenho da quantização vetorial:

1. *Mapa auto-organizado de Kohonen (KO)*: 3 entradas, 13 saídas, máximo de 200 iterações, taxa de aprendizado inicial  $\eta_0 = 0,1$ , arquitetura circular, função de distância gaussiana;
2. *Mapa fuzzy c-médias (CM)*: 3 entradas, 13 saídas, máximo de 200 iterações, taxa de aprendizado inicial  $\eta_0 = 0,1$ ;
3. *Mapa de k-médias (KM)*: 3 entradas, 13 saídas, máximo de 200 iterações, taxa de aprendizado inicial  $\eta_0 = 0,1$ ;
4. *Classificador dialético objetivo canônico (ODC-CAN)*: 14 pólos iniciais, 2 fases históricas de 150 iterações cada fase, passo histórico inicial  $\eta_0 = 0,1$ , força mínima de 5%, contradição mínima de 0,01, contradição máxima de 0,98, máxima crise de 35%, até 12 pólos. Depois de todas as fases históricas, o processo de treinamento finalizou com 13 pólos;
5. *Classificador dialético objetivo com entropia máxima (ODC-PME)*: 14 pólos iniciais, 2 fases históricas de 150 iterações cada fase, passo histórico inicial  $\eta_0 = 0,1$ , força mínima de 5%, contradição mínima de 0,01, contradição máxima de 0,98, máxima crise de 35%, até 12 pólos. Depois de todas as fases históricas, o processo de treinamento finalizou com 13 pólos. Aqui foi aplicado o Princípio da Máxima Entropia para obter novas expressões para as funções de anticontração, semelhantemente ao *fuzzy* c-médias com máxima entropia, resultando praticamente as mesmas expressões das funções de pertinência do *fuzzy* c-médias com máxima entropia [127].

O número de entradas é igual ao número de bandas da imagem, uma vez que a classificação é *pixel a pixel*, sendo portanto o *pixel* a entrada da rede. Já o número de saídas foi definido como sendo o número de classes presentes na imagem, que por sua vez é o número de elementos

anatômica e fisiologicamente diferentes presentes na imagem (11 classes) mais duas classes para alocar diferentes populações de ruído [20, 19]. Os demais parâmetros foram definidos empiricamente.

### 6.2.6 Classificadores Otimizados

As imagens multiespectrais sintéticas obtidas pelas composições coloridas R0-G1-B2 foram classificadas usando os seguintes métodos baseados na otimização do mapa de k-médias, considerando 13 classes, 20 pólos iniciais, máximo de 10 fases históricas, com duração máxima de 20 gerações cada fase histórica, passos iniciais (histórico e contemporâneo) de 0,99, contradição mínima de 0,1, contradição máxima de 0,9, efeito de crise máximo de 0,9, valor limite de 0,01:

1. *EQ-CAN-KM*: k-médias otimizado pelo método dialético canônico, de acordo com o erro de quantização  $J_e$ ;
2. *IC-CAN-KM*: k-médias otimizado pelo método dialético canônico, de acordo com o índice combinado de Omran  $J_o$ ;
3. *EQ-PME-KM*: k-médias otimizado pelo método dialético com entropia maximizada, de acordo com o erro de quantização  $J_e$ ;
4. *IC-PME-KM*: k-médias otimizado pelo método dialético com entropia maximizada, de acordo com o índice combinado de Omran  $J_o$ .

Os parâmetros iniciais foram definidos de forma semelhante ao que foi descrito na seção anterior. Os parâmetros do método dialético foram definidos empiricamente.

Uma vez que a velocidade de convergência e os resultados de qualquer algoritmo de busca e otimização são altamente dependentes da inicialização do algoritmo, foram utilizadas 5 iterações do algoritmo k-médias padrão para geração do conjunto inicial de pólos, de acordo com o critério de geração da população inicial proposto por Saha e Bandyopadhyay (2005), que evita buscas iniciais totalmente aleatórias, uma vez que, embora apenas 5 iterações não sejam suficientes para garantir a convergência do algoritmo k-médias, e nem essa convergência é desejada, uma vez que se está utilizando um algoritmo de otimização em substituição ao algoritmo de treinamento, essas 5 iterações iniciais ajudam o algoritmo de otimização a iniciar com pólos que podem estar na vizinhança da solução ótima, sendo este portanto um critério melhor do que inicializar aleatoriamente, evitando buscas totalmente aleatórias e variações quanto à convergência do algoritmo [143].

## 6.3 Resultados e Discussões

Para facilitar a avaliação dos resultados obtidos, estes foram divididos em duas categorias: resultados gerados usando classificadores dialéticos objetivos; e resultados gerados usando mapa de k-médias otimizado pelo método dialético de otimização, em suas versões canônica e com maximização da entropia, em função do índice combinado de Omran e do erro de quantização.

Tanto a qualidade da quantização quanto a da segmentação são avaliadas usando índices de validade de agrupamento e índices de fidelidade da imagem quantizada em relação à imagem original.

### 6.3.1 Classificadores Dialéticos Objetivos

A figura 6.5 mostra os resultados de classificação, enquanto a figura 6.6 exibe os resultados de quantização para a imagem sem ruído da fatia 97, figura 6.4, usando os métodos KO, CM, KM, ODC-PME e ODC-CAN. Esses resultados ilustram qualitativamente as diferenças entre os métodos de classificação e quantização, dado que a fatia 97 possui todas as 13 classes presentes na análise [19].

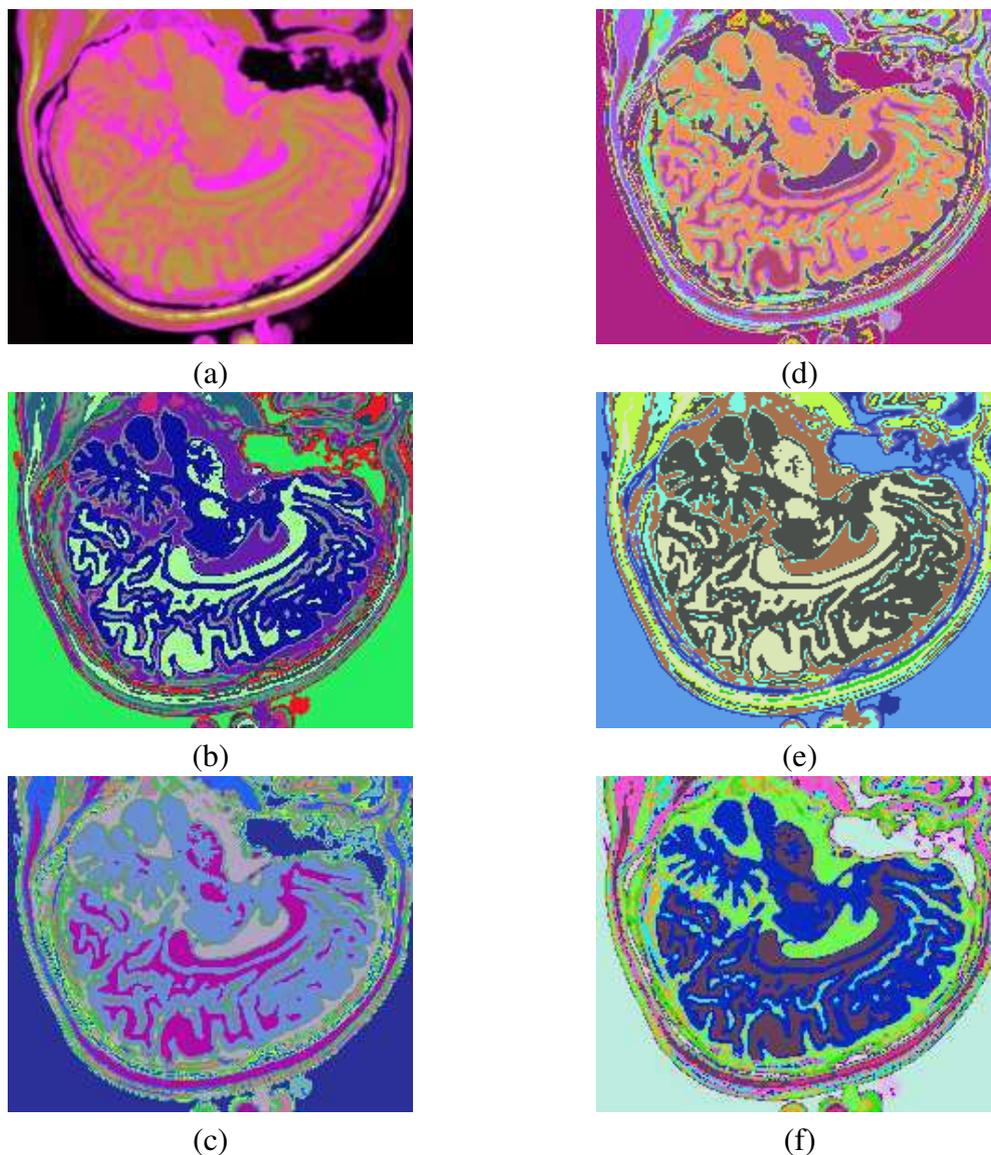
A tabela 6.1 mostra os resultados da avaliação dos métodos de classificação não supervisionada quanto à quantização vetorial, usando os índices de fidelidade  $\epsilon_{ME}$ ,  $\epsilon_{MAE}$ ,  $\epsilon_{MSE}$ ,  $\epsilon_{RMSE}$  e  $\epsilon_{PSNR}$ , considerando todas as 181 fatias com 3 bandas ( $DP$ ,  $T_1$  and  $T_2$ ), para os métodos KO, CM, KM, ODC-PME e ODC-CAN e 0% de ruído. Já os gráficos das figuras 6.7, 6.8, 6.9 e 6.10 mostram os resultados em função do nível percentual de ruído para os diversos métodos, para um total de 6 volumes de 181 fatias de 3 bandas, totalizando 1086 imagens coloridas, ou 3258 imagens em níveis de cinza.

As tabelas 6.2 e 6.4 mostram o resultado das comparações entre os pares de média e desvio médio das medidas de  $\epsilon_{ME}$ ,  $\epsilon_{MAE}$ ,  $\epsilon_{MSE}$ ,  $\epsilon_{RMSE}$  e  $\epsilon_{PSNR}$ , para os métodos KO, CM, KM, ODC-CAN e ODC-PME, de acordo com teste F de hipótese nula, enquanto as tabelas 6.3 e 6.5 mostram comparações entre os métodos ODC-PME e ODC-CAN, e os métodos KO, CM e KM, respectivamente, usando o teste de  $\chi^2$  para medir o grau de dependência (considerado aqui como grau de similaridade) entre os resultados dos métodos ODC-PME e ODC-CAN e os resultados dos outros métodos; as sequências de valores observados foram construídas usando os pares de média e desvio médio dos índices de fidelidade  $\epsilon_{ME}$ ,  $\epsilon_{MAE}$ ,  $\epsilon_{RMSE}$  e  $\epsilon_{PSNR}$ . Todos os testes foram executados considerando um intervalo de confiança de 95%. As figuras 6.11 e 6.12 mostram o resultado da similaridade medida entre os métodos ODC-PME e ODC-CAN e os outros métodos, respectivamente, usando o teste de  $\chi^2$ , em função do ruído percentual.

**Tabela 6.1** Resultados de quantização expressos pelos índices de fidelidade, para 0% de ruído

	KO	CM	KM	ODC-PME	ODC-CAN
$\epsilon_{ME}$	50 ± 5	106 ± 15	84 ± 22	50 ± 3	46 ± 6
$\epsilon_{MAE}$	10 ± 3	10 ± 2	13 ± 4	12 ± 2	10 ± 2
$\epsilon_{MSE}$	186 ± 52	258 ± 81	347 ± 261	249 ± 60	169 ± 46
$\epsilon_{RMSE}$	13 ± 2	16 ± 31	18 ± 5	15 ± 2	13 ± 2
$\epsilon_{PSNR}$	26 ± 2	24 ± 2	24 ± 3	25 ± 1	26 ± 2

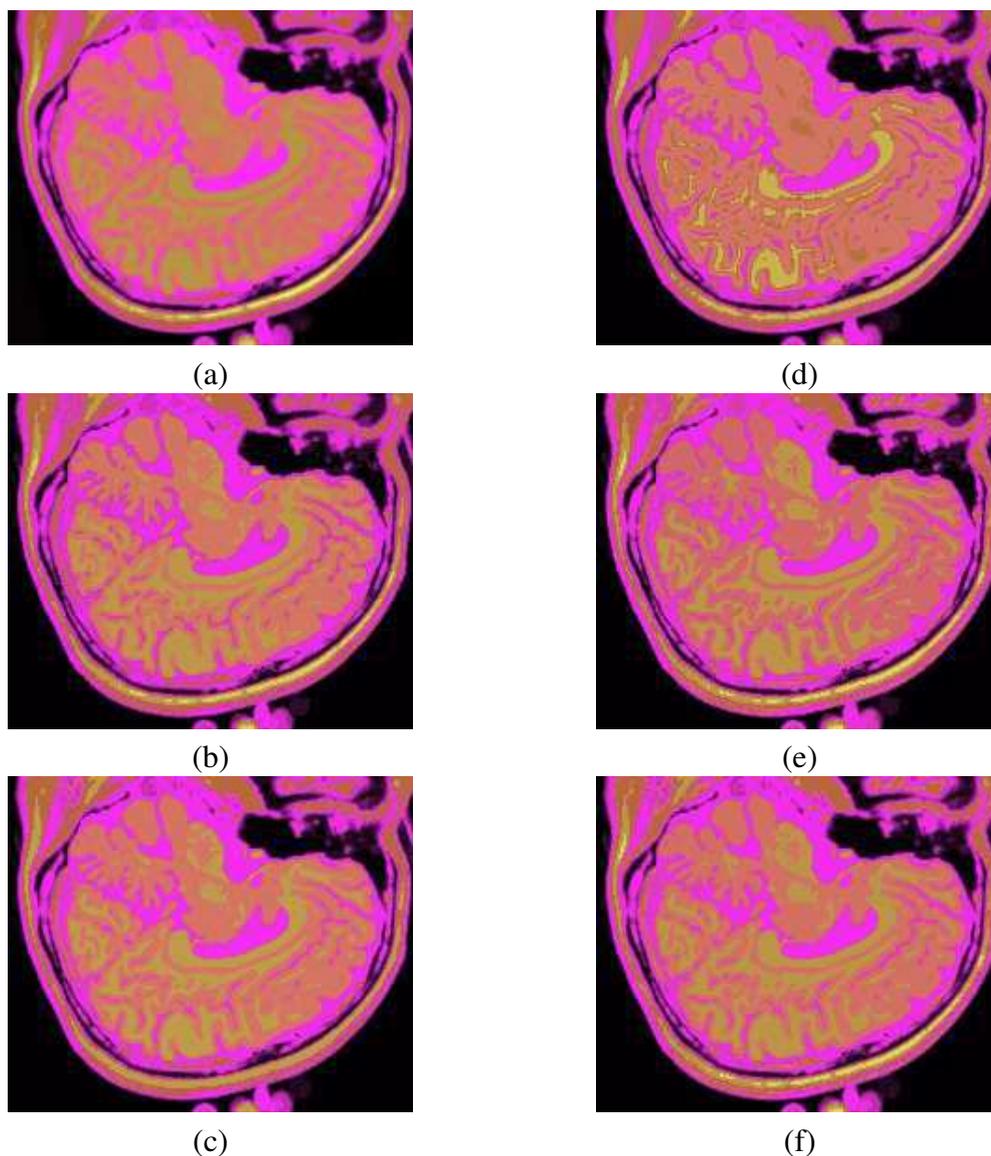
Observando o gráfico da figura 6.7 pode-se notar que os métodos ODC-CAN, ODC-PME e KO são os que atingem os melhores resultados para  $\epsilon_{ME}$ , para 0% de ruído. No entanto, os resultados vão se aproximando à medida em que o nível de ruído aumenta, e acabam por se tornar totalmente coincidentes a partir de 5% de ruído. Já os gráficos das figuras 6.8, 6.9 e 6.10 mostram que os resultados de  $\epsilon_{MAE}$ ,  $\epsilon_{RMSE}$  e  $\epsilon_{PSNR}$  obtidos pelos métodos de classificação utilizados não são distinguíveis, não sendo possível comparar os métodos entre si usando os



**Figura 6.5** Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD,  $T_1$  e  $T_2$  (a) e resultados de classificação usando os métodos KO (b), CM (c), KM (d), ODC-PME (e) e ODC-CAN (f)

**Tabela 6.2** Graus de similaridade entre o método ODC-PME e os métodos KO, CM e KM, de acordo com o teste de hipótese nula F, para 0% de ruído

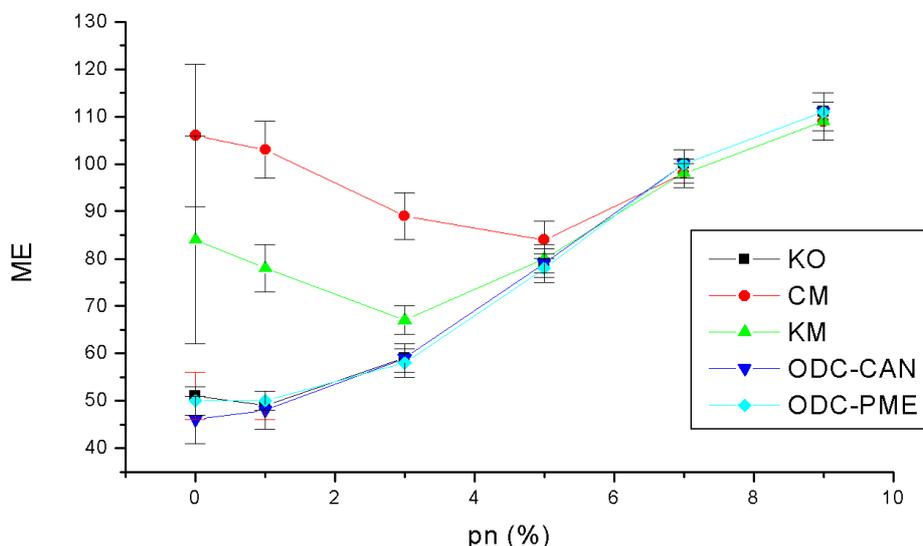
	$\mu(\epsilon_{ME})$	$\mu(\epsilon_{MAE})$	$\mu(\epsilon_{RMSE})$	$\mu(\epsilon_{PSNR})$
ODC-PME-KO	0,99	0,86	0,89	1,00
ODC-PME-CM	0,61	0,86	0,89	1,00
ODC-PME-KM	0,83	0,86	0,89	1,00



**Figura 6.6** Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD,  $T_1$  e  $T_2$  (a) e resultados de quantização usando os métodos KO (b), CM (c), KM (d), ODC-PME (e) e ODC-CAN (f)

**Tabela 6.3** Graus de similaridade entre o método ODC-PME e os métodos KO, CM e KM, de acordo com o teste de  $\chi^2$ , em função do ruído

pn (%)	ODC-PME-KO	ODC-PME-CM	ODC-PME-KM
0	0,956111	0,000000	0,000012
1	0,970435	0,000060	0,079578
3	0,999864	0,112552	0,986766
5	0,999971	0,994726	0,998794
7	1,000000	0,995641	0,992783
9	0,994829	0,998837	0,998837



**Figura 6.7** Resultados de  $\epsilon_{ME}$  em função do ruído percentual (ODC)

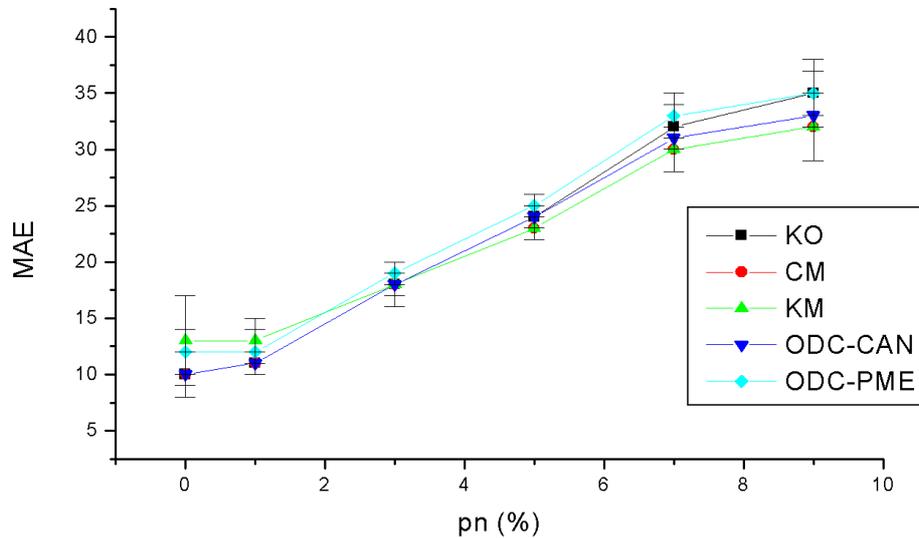
**Tabela 6.4** Graus de similaridade entre o método ODC-CAN e os métodos KO, CM e KM, de acordo com o teste de hipótese nula F, para 0% de ruído

	$\mu(\epsilon_{ME})$	$\mu(\epsilon_{MAE})$	$\mu(\epsilon_{RMSE})$	$\mu(\epsilon_{PSNR})$
ODC-CAN-KO	0,93	1,00	1,00	0,97
ODC-CAN-CM	0,54	1,00	1,00	0,97
ODC-CAN-KM	0,74	1,00	1,00	0,97

índices de fidelidade citados. Conseqüentemente, a comparação dos métodos índice a índice não é muito conclusiva quanto à classificação dos métodos de segmentação, como mostram os gráficos das figuras 6.7, 6.8, 6.9 e 6.10, para todos os níveis de ruído, e os testes F exibidos nas tabelas 6.2 e 6.4, que comparam os métodos ODC-PME e ODC-CAN com os métodos KO, CM e KM, para 0% de ruído.

Uma vez que as comparações índice a índice não são muito conclusivas, foi utilizado o teste de  $\chi^2$  para fazer comparações globais entre os métodos ODC-PME e ODC-CAN e os métodos KO, CM e KM. A tabela 6.3 mostra os resultados do teste de  $\chi^2$  entre o método ODC-PME e os outros métodos, mostrando que, para 0% de ruído, os métodos ODC-PME e KO são idênticos. O gráfico da figura 6.11 ilustra os resultados para os níveis de ruído de 0% a 9%, mostrando que os métodos ODC-PME e KO são praticamente iguais quanto aos resultados dos índices de fidelidade, apesar de todos os métodos atingirem resultados idênticos a partir de 5% de ruído.

A tabela 6.5 mostra os resultados do teste de  $\chi^2$  entre o método ODC-CAN e os outros métodos, mostrando que, para 0% de ruído, os métodos ODC-CAN e KO são similares. O gráfico da figura 6.12 ilustra os resultados para os níveis de ruído de 0% a 9%, mostrando



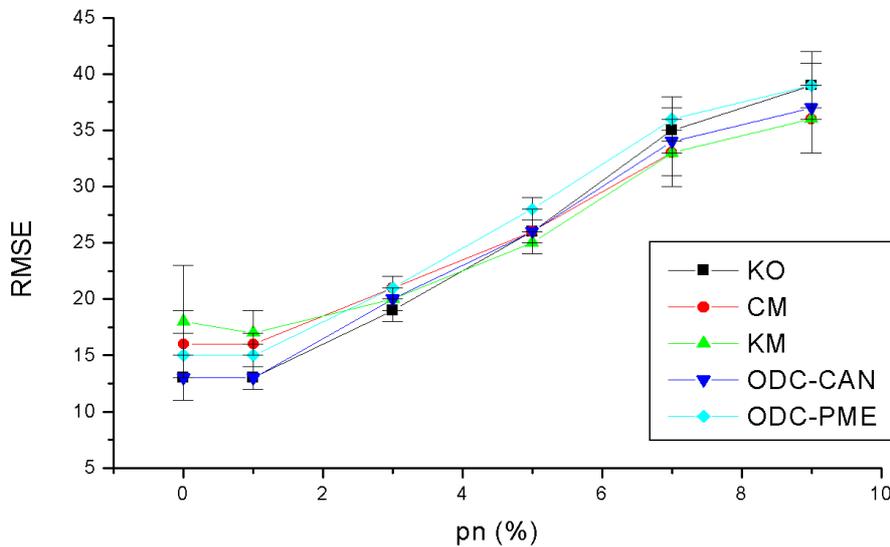
**Figura 6.8** Resultados de  $\varepsilon_{MAE}$  em função do ruído percentual (ODC)

**Tabela 6.5** Graus de similaridade entre o método ODC-CAN e os métodos KO, CM e KM, de acordo com o teste de  $\chi^2$ , em função do ruído

pn (%)	ODC-CAN-KO	ODC-CAN-CM	ODC-CAN-KM
0	0,994985	0,000000	0,000005
1	0,999750	0,000069	0,044842
3	0,993934	0,100309	0,962277
5	0,956870	0,915891	0,954338
7	0,992681	0,998602	0,976113
9	0,747098	0,997809	0,997809

que os métodos ODC-CAN e KO são muito similares quanto aos resultados dos índices de fidelidade até 7% de ruído, apesar de todos os métodos atingirem resultados semelhantes a partir de 5% de ruído, enquanto a partir de 7% de ruído os métodos ODC-CAN e KO passam a se diferenciar, comportamento bem diferente daquele observado no método ODC-PME quando comparado com o método KO.

Esses resultados mostram os classificadores dialéticos objetivos em sua versão canônica e na versão com entropia maximizada podem atingir resultados de segmentação e quantização similares àqueles obtidos usando mapas auto-organizados de Kohonen, principalmente quando se trata da versão com entropia maximizada, o que é uma característica muito importante, uma vez que os mapas auto-organizados de Kohonen são quantizadores vetoriais ótimos [33], com a vantagem de os classificadores dialéticos objetivos terem número de classes adaptável, não sendo necessário conhecer o número de classes presente na imagem, bastando inicializar a classificação com um número de classes maior do que o esperado.



**Figura 6.9** Resultados de  $\epsilon_{RMSE}$  em função do ruído percentual (ODC)

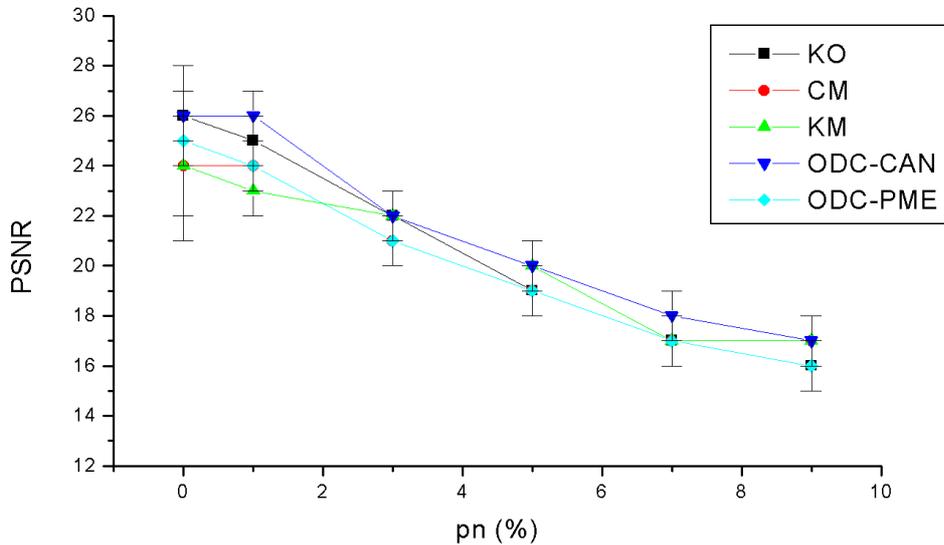
### 6.3.2 Classificadores K-Médias Otimizados pelo Método Dialético

Os gráficos das figuras 6.13, 6.14, 6.15 e 6.16 mostram os resultados das medidas dos índices de fidelidade  $\epsilon_{ME}$ ,  $\epsilon_{MAE}$ ,  $\epsilon_{RMSE}$  e  $\epsilon_{PSNR}$ , enquanto os gráficos das figuras 6.17 e 6.18 exibem os resultados do erro de quantização,  $J_e$ , e do índice de Omran,  $J_o$ , para os métodos KO, CM, KM, IC-CAN, EQ-CAN, IC-PME e EQ-PME, em função dos níveis percentuais de ruído, para os 6 volumes de 181 fatias de 3 bandas. A figura 6.19 mostra os resultados de classificação, enquanto a figura 6.20 exhibe os resultados de quantização para a imagem sem ruído da fatia 97, figura 6.4, usando os métodos IC-CAN-KM, IC-PME-KM, EQ-CAN-KM e EQ-PME-KM. É possível perceber a partir das imagens que a introdução da otimização dialética permitiu ao k-médias identificar diferentes classes dentro do fundo da imagem.

O gráfico da figura 6.13 mostra que o método EQ-PME, ou seja, o mapa de k-médias otimizado pelo método dialético de entropia maximizada em função do erro de quantização  $J_e$ , foi o que apresentou melhor resultado quanto ao índice de fidelidade  $\epsilon_{ME}$ .

Já o gráfico da figura 6.14 mostra que os melhores resultados do índice de fidelidade  $\epsilon_{MAE}$  são aqueles obtidos pelos métodos KO e CM, de 0% a 3% de ruído percentual, mas a partir de 3% o método EQ-PME passa a apresentar os melhores resultados, mostrando que o mapa auto-organizado de Kohonen e o mapa *fuzzy* c-médias são pouco robustos em relação ao ruído, enquanto o mapa de k-médias otimizado pelo método dialético com máxima entropia em função do erro de quantização tem desempenho menos sensível aos níveis de ruído; esse resultado, contudo, não é muito diferente do resultado obtido pelo uso do mapa de k-médias.

O gráfico da figura 6.15 mostra que o mapa auto-organizado de Kohonen apresentou os



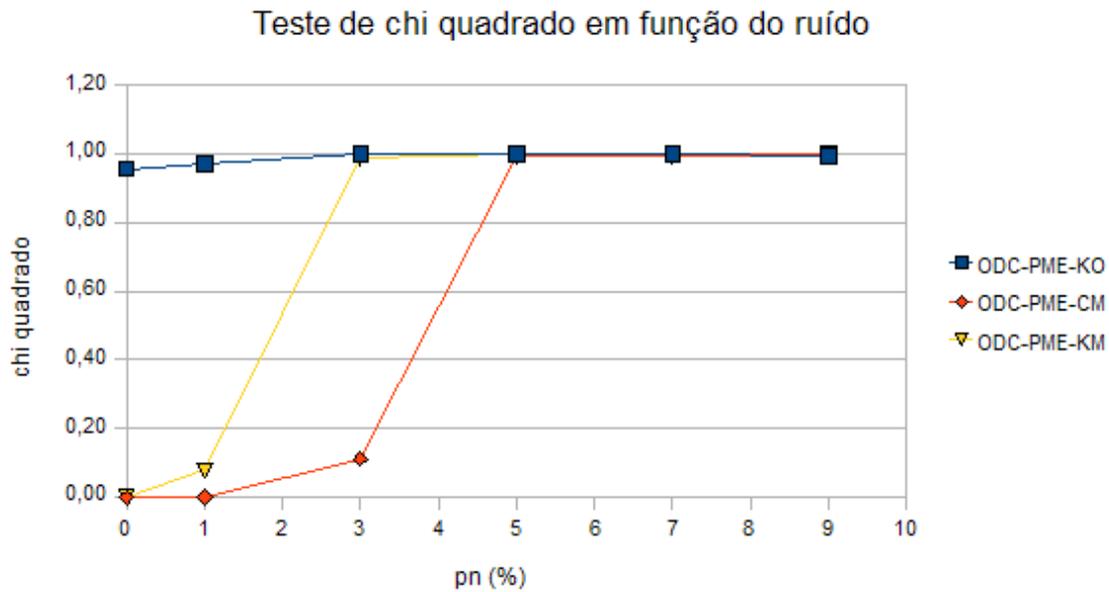
**Figura 6.10** Resultados de  $\epsilon_{\text{PSNR}}$  em função do ruído percentual (ODC)

melhores resultados do índice de fidelidade  $\epsilon_{\text{RMSE}}$  para níveis de ruído percentual de 0% a 3%. Contudo, a partir de 3% os resultados obtidos pelo método KO simplesmente não são mais distinguíveis dos resultados para os outros métodos. Já gráfico da figura 6.16 ilustra que os resultados da aplicação de todos os métodos não são estatisticamente distintos entre si de acordo com o índice de fidelidade  $\epsilon_{\text{RMSE}}$ , e isso levando em conta todos os níveis de ruído percentual, mas há uma leve indicação de superioridade dos resultados obtidos com o método de k-médias otimizado usando o método dialético com entropia otimizada em função do erro de quantização, EQ-PME.

Entretanto, o gráfico da figura 6.17 ilustra que, embora os resultados de 0% a 1% de ruído mostrem que os resultados não são distinguíveis entre si quanto ao erro de quantização, a partir de 1% de ruído os valores do erro de quantização para o método EQ-PME são bem menores do que as medidas de erro de quantização para os outros métodos, mostrando que a otimização do mapa de k-médias usando o método dialético com entropia maximizada em função do erro de quantização resulta resultados qualitativamente superiores àqueles obtidos com os outros métodos, conforme também a figura 6.20.

As diferenças entre os métodos KO, CM e KM e os métodos EQ-CAN, EQ-PME, IC-CAN e IC-PME são mais gritantes quanto a avaliação é feita usando o índice combinado de Omran, como ilustra a figura 6.17: todos os métodos baseados na otimização do mapa de k-médias por versões do método dialético retornam medidas de  $J_o$  que são pelo menos metade das mesmas medidas obtidas pelos outros métodos.

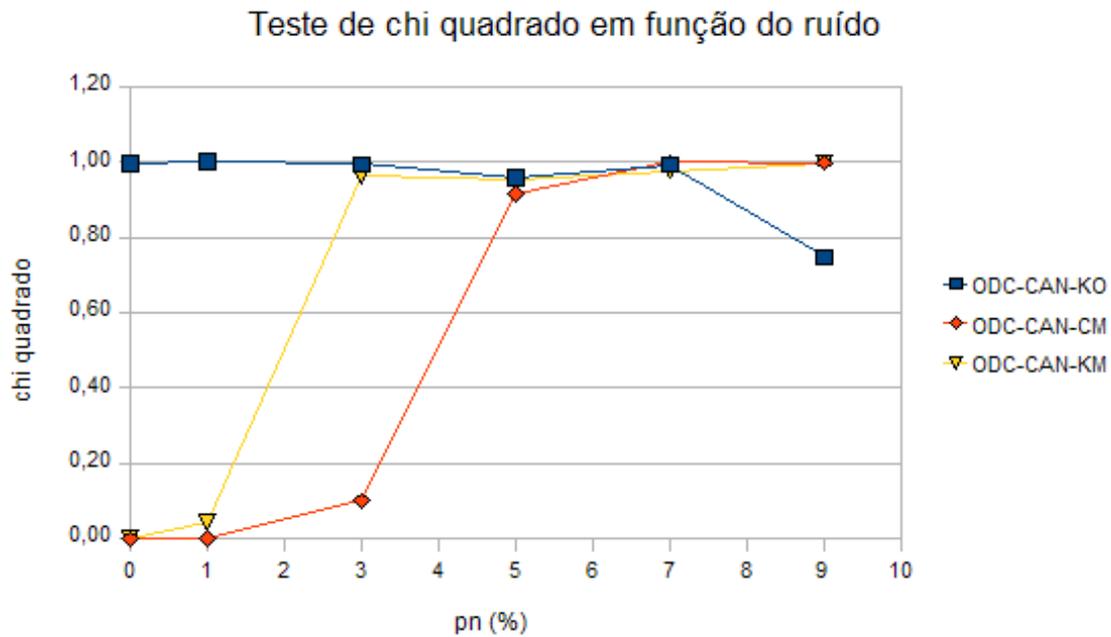
No entanto, da figura 6.17 também fica claro que, embora todos os métodos baseados na otimização pelo método dialético retornem medidas de  $J_o$  muito melhores que aquelas obtidas



**Figura 6.11** Graus de similaridade do método ODC-PME em relação aos métodos KO, CM e KM, usando o teste de  $\chi^2$ , em função do ruído percentual

usando KO, CM e KM, o método EQ-PME é o que apresenta o melhor resultado dentre todos, destacando-se dos outros métodos a partir de 1% de ruído percentual, ou seja: o método baseado no mapa de k-médias otimizado usando o método dialético de entropia maximizada em função do erro de quantização  $J_e$ . Esse resultado não deixa de ser interessante e curioso, pois mostra que o índice de Omran  $J_o$ , no caso das imagens utilizadas, é dominado pelo erro de quantização  $J_e$ , dado que a otimização em função do índice de Omran usando os métodos dialéticos canônico (IC-CAN) e com entropia maximizada (IC-PME) são idênticos e qualitativamente inferiores, assim como os resultados obtidos usando o método dialético canônico em função de  $J_e$  (EQ-CAN). Isso também indica que a aplicação do Princípio da Máxima Entropia diferenciou sensivelmente o algoritmo de sua versão canônica, tendo acelerado a convergência para a otimização em função do erro de quantização, como bem atestam os resultados.

Esses resultados, combinados àqueles apresentados anteriormente, mostram que, para a aplicação ilustrada neste capítulo, o método EQ-CAN foi superior aos outros métodos de classificação não supervisionada e quantização apresentados, tanto quando são utilizados índices de fidelidade quanto índices de validade do agrupamento, o que mostra que, para as imagens utilizadas, tanto existe um predomínio do erro de quantização  $J_e$  sobre o índice combinado de Omran  $J_o$  (onde este último, por sua vez, contém o primeiro conforme expressões 6.17 e 6.11), quanto existe uma forte influência da aplicação do Princípio da Máxima Entropia ao método dialético de busca e otimização, que nesta aplicação ajudou o método a ter sua convergência acelerada ou até mesmo a vencer um mínimo local.

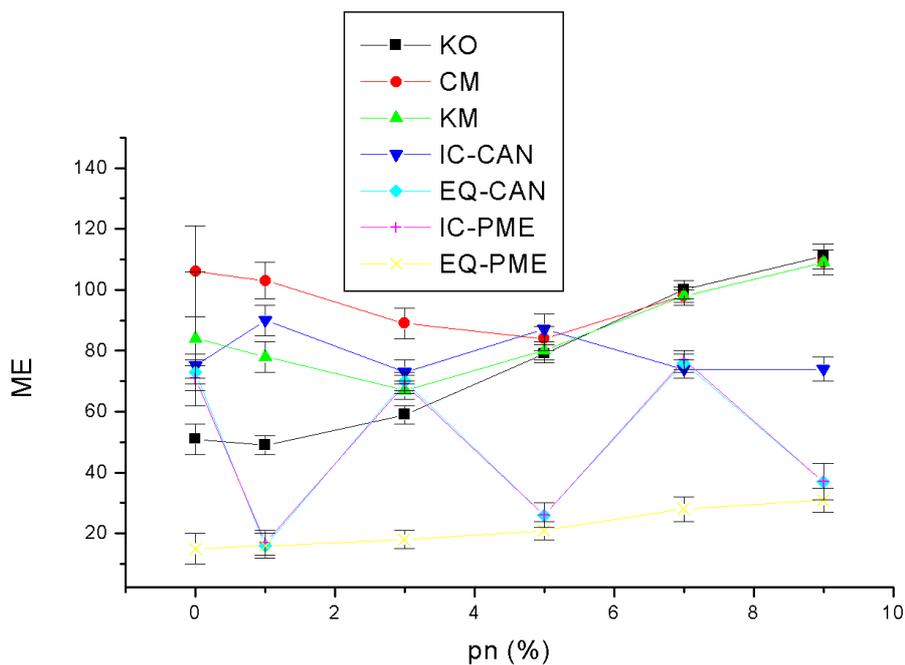


**Figura 6.12** Graus de similaridade do método ODC-CAN em relação aos métodos KO, CM e KM, usando o teste de  $\chi^2$ , em função do ruído percentual

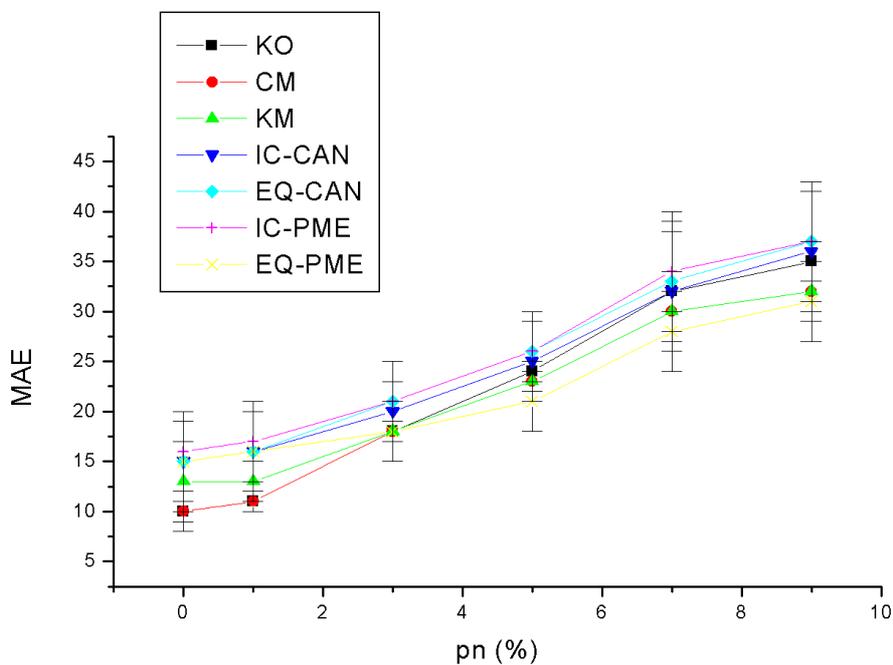
## 6.4 Conclusões

Este capítulo mostrou duas aplicações do método dialético: classificadores dialéticos objetivos e classificadores baseados na otimização dialética. Os classificadores dialéticos objetivos são uma implementação direta do método dialético como definido neste trabalho para problemas de classificação e agrupamento. Foi visto que esses classificadores podem chegar a atingir resultados tão bons quanto os resultados atingidos usando o mapa auto-organizado de Kohonen, considerado um método ótimo para quantização vetorial, com a vantagem de não ser necessário especificar o número exato de classes no início do processo de treinamento, fazendo com que o método seja executado até que se atinja um bom número de classes, necessariamente menor ou igual ao número inicial de classes definidas no início do treinamento.

Já os classificadores baseados na otimização dialética mostraram que é possível obter classificadores otimizados segundo algum critério usando as diversas versões do método dialético e, com eles, conseguir resultados de quantização e de fidelidade superiores àqueles obtidos utilizando o método com parâmetros gerados pela utilização de seu algoritmo próprio de treinamento. Esses classificadores otimizados, quando comparados aos que não fazem uso de otimização e sim utilizam seus próprios algoritmos de treinamento, podem ser uma importante alternativa para evitar com mais facilidade que o processo fique preso em mínimos locais, dada a boa capacidade do método dialético de evitar mínimos locais, como se percebe das situações onde a função objetivo é multimodal, com destaque para a versão com entropia maximizada.



**Figura 6.13** Resultados de  $\epsilon_{ME}$  em função do ruído percentual (ODM)



**Figura 6.14** Resultados de  $\epsilon_{MAE}$  em função do ruído percentual (ODM)

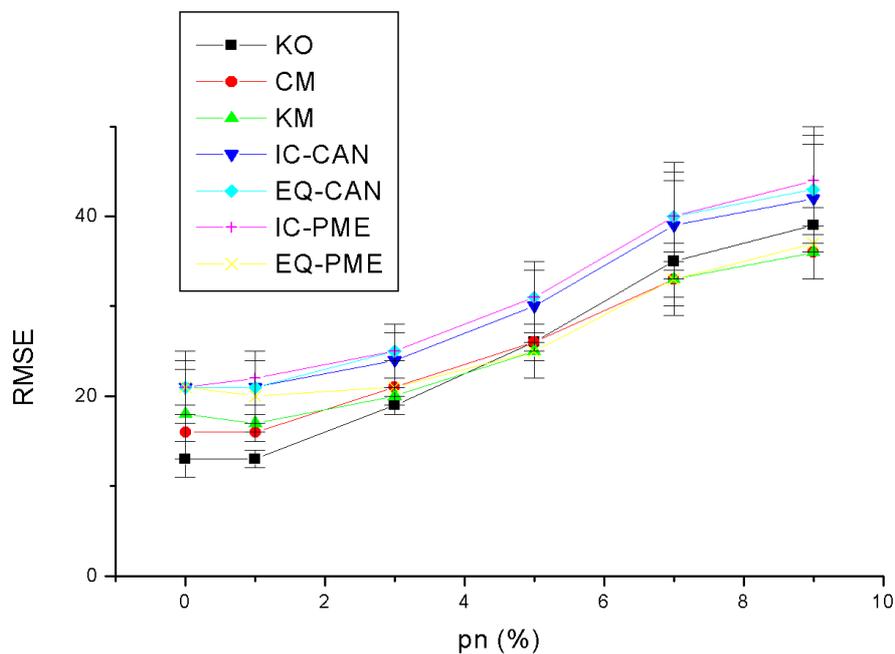


Figura 6.15 Resultados de  $\epsilon_{\text{RMSE}}$  em função do ruído percentual (ODM)

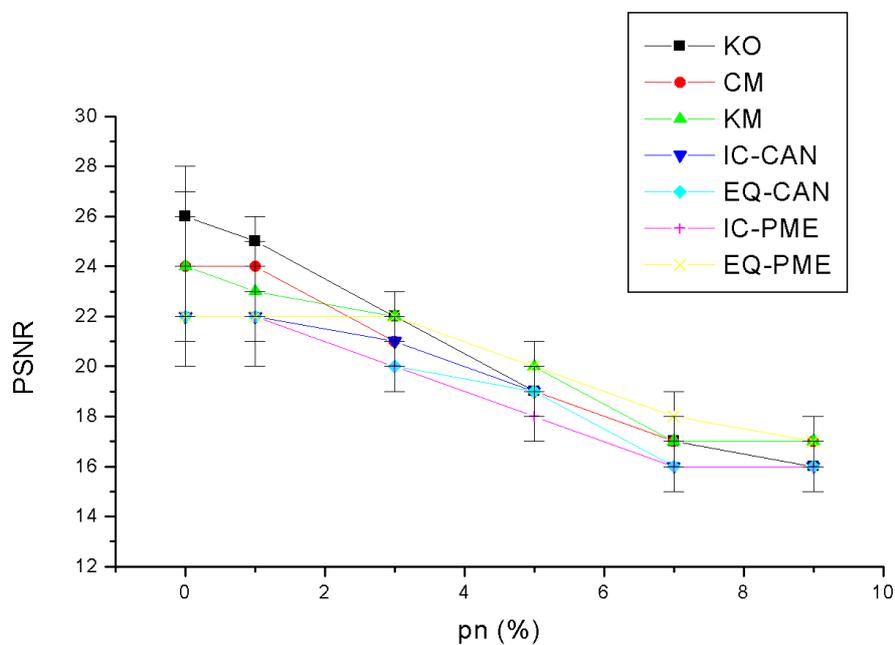
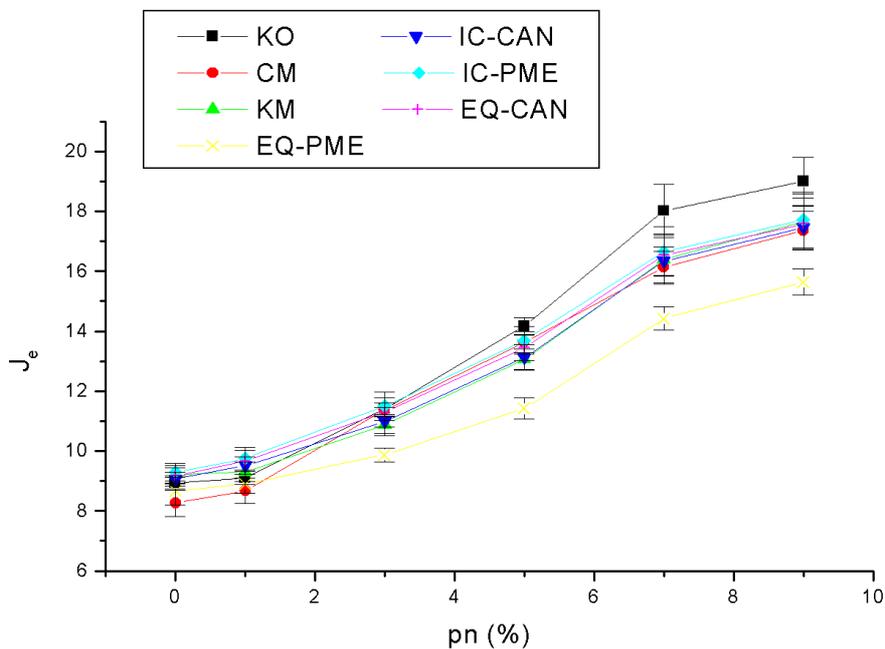
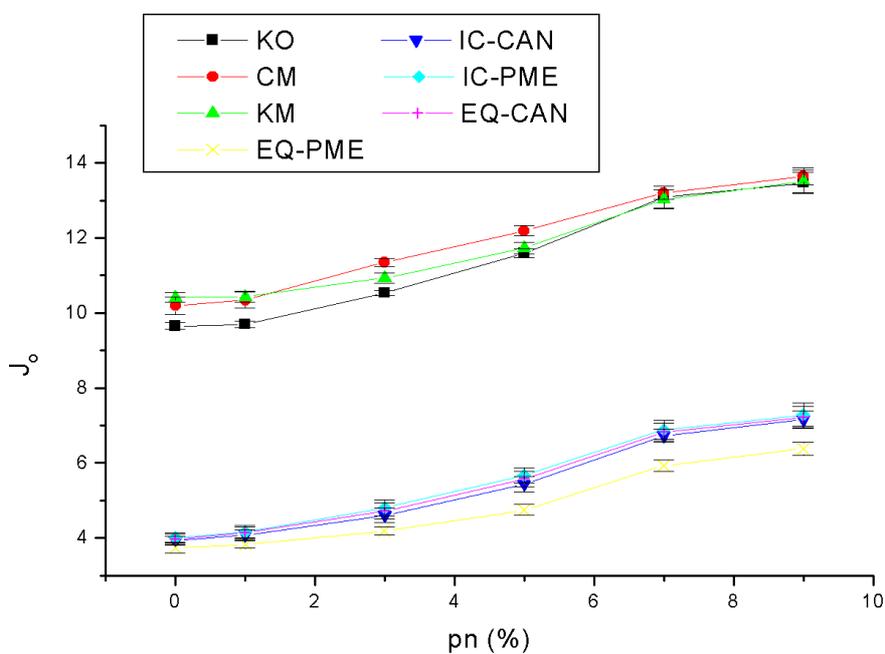


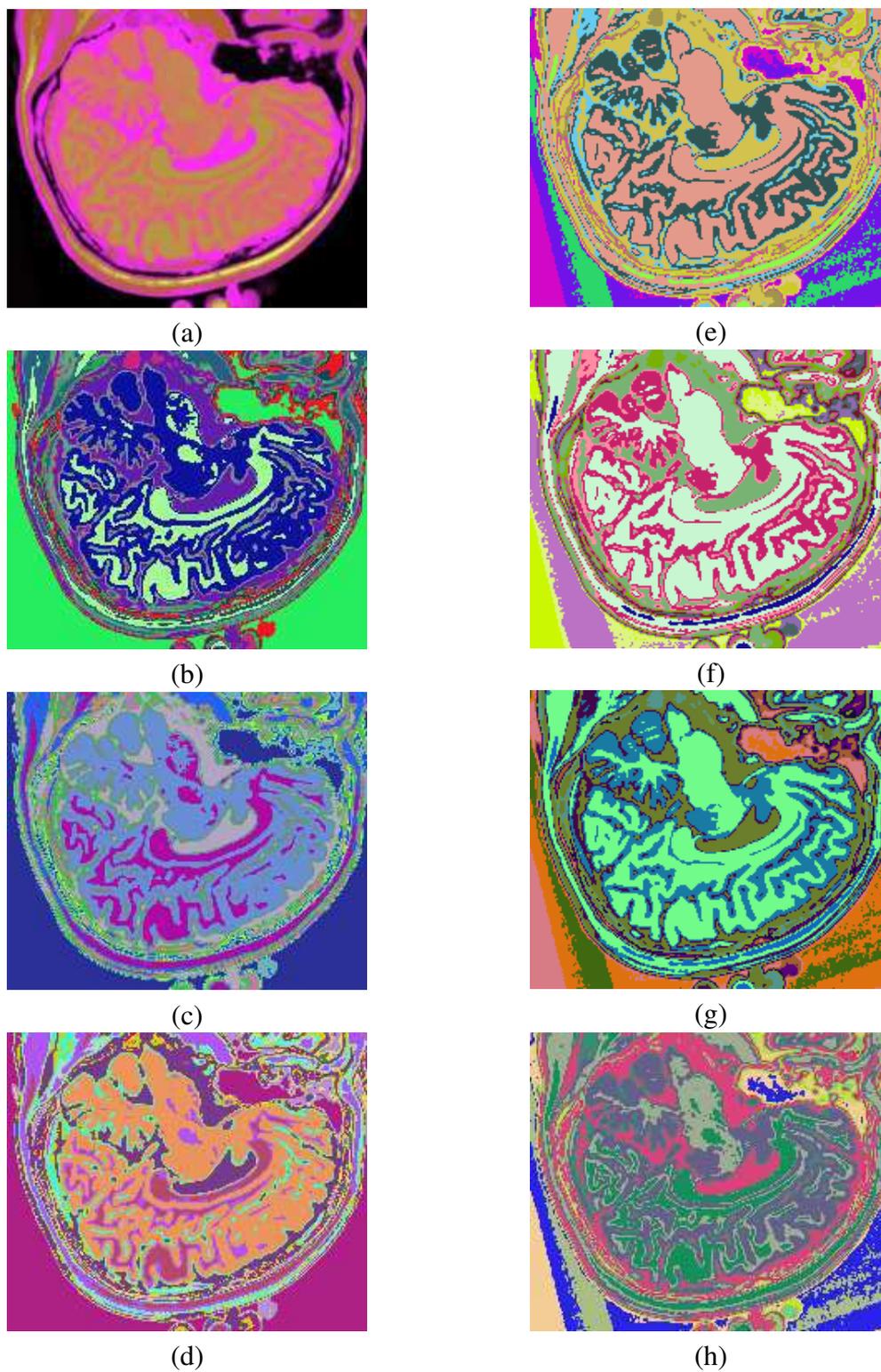
Figura 6.16 Resultados de  $\epsilon_{\text{PSNR}}$  em função do ruído percentual (ODM)



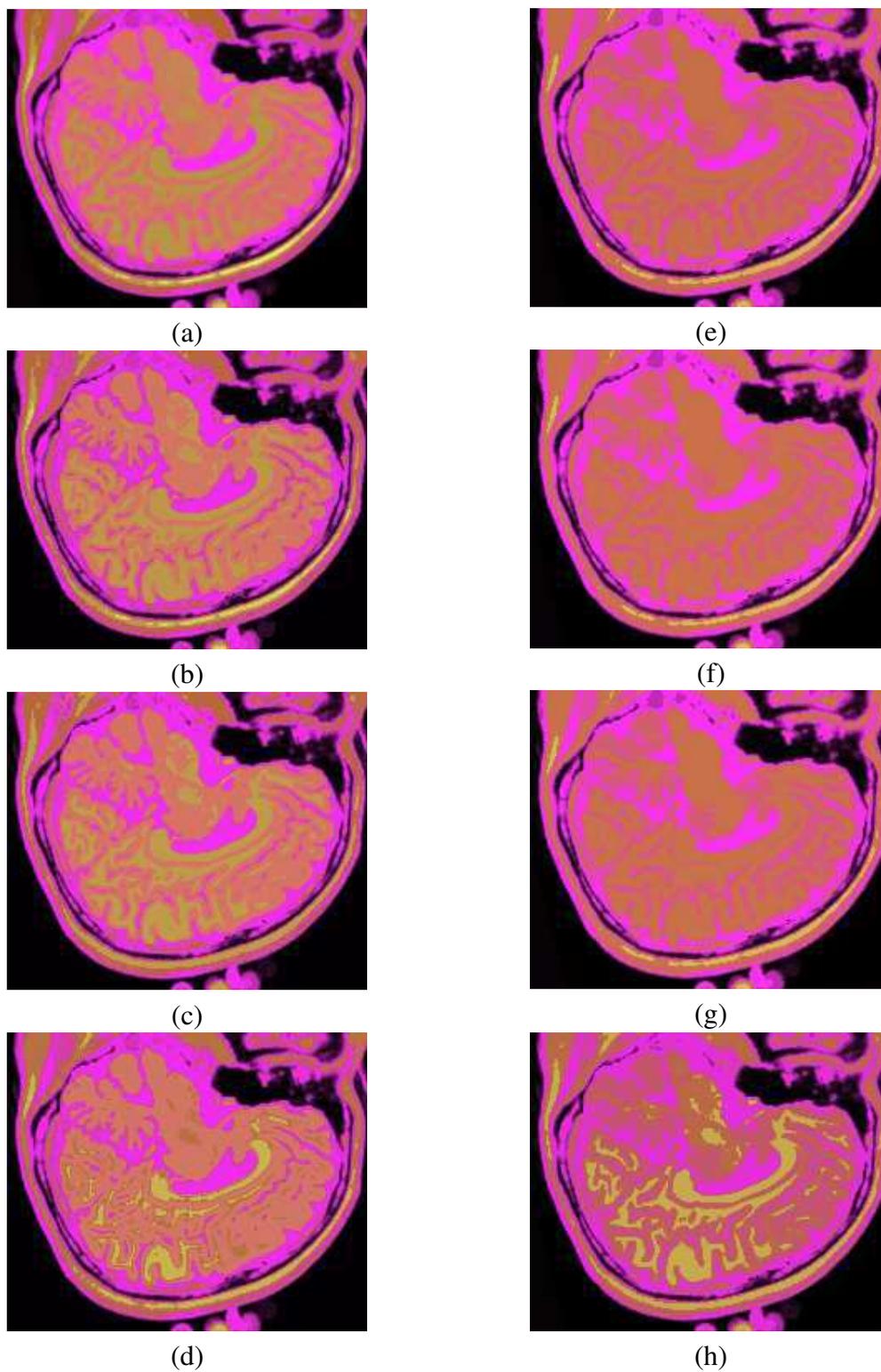
**Figura 6.17** Resultados de  $J_e$  em função do ruído percentual



**Figura 6.18** Resultados de  $J_o$  em função do ruído percentual



**Figura 6.19** Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD,  $T_1$  e  $T_2$  (a) e resultados de classificação usando os métodos KO (b), CM (c), KM (d), IC-CAN-KM (e), IC-PME-KM (f), EQ-CAN-KM (g) e EQ-PME-KM (h)



**Figura 6.20** Composição colorida R0-G1-B2 das imagens da fatia 97 ponderadas em PD,  $T_1$  e  $T_2$  (a) e resultados de quantização usando os métodos KO (b), CM (c), KM (d), IC-CAN-KM (e), IC-PME-KM (f), EQ-CAN-KM (g) e EQ-PME-KM (h)



# Classificação de Imagens de Difusão de Alzheimer

## 7.1 Introdução

A doença de Alzheimer é a maior causa de demência, tanto no grupo etário senil quanto no pré-senil, observando-se seu aumento gradual à medida em que o indivíduo envelhece [144, 145]. As principais manifestações da doença compreendem o comprometimento cognitivo com gradual perda de memória, além de sintomas psicológicos, neurológicos e comportamentais que indicam declínio nas atividades da vida diária como um todo [15].

A doença de Alzheimer é caracterizada pela redução da substância cinzenta e pelo aumento dos sulcos. A substância cinzenta é responsável pela memória e sua redução explica a perda gradual de memória no indivíduo senil afetado por esta doença. Entretanto, também a substância branca é afetada, apesar de ser desconhecida a relação entre a doença e a substância branca [146, 147, 148, 149].

A aquisição de imagens de ressonância magnética ponderadas em difusão possibilita a visualização da dilatação dos cornos temporais dos ventrículos laterais, além de realçar o aumento dos sulcos, que estão relacionados com o avanço da doença de Alzheimer [1].

Assim, a medição volumétrica de estruturas cerebrais é de grande importância para o diagnóstico e a avaliação do progresso de doenças como a doença de Alzheimer [144, 150, 151, 152], em especial a medição das áreas ocupadas pelos sulcos e pelos ventrículos laterais, pois estas medidas permitem a adição de informação quantitativa à informação qualitativa expressa pelas imagens de ressonância magnética ponderadas em difusão [153].

A avaliação do progresso da doença de Alzheimer por meio da análise de imagens de ressonância magnética ponderadas em difusão é feita a partir da aquisição de no mínimo três imagens de cada fatia de interesse, onde cada uma dessas imagens é adquirida usando a sequência eco de spin Stejskal-Tanner, com expoentes de difusão diferentes, sendo um deles igual a  $0 \text{ s/mm}^2$ , ou seja, uma das três imagens é uma imagem ponderada em  $T_2$  [147, 1].

Utilizando essas três imagens assim adquiridas, é montada uma quarta imagem: o mapa de coeficientes de difusão aparente (*Apparent Diffusion Coefficient Map*, ADC), onde cada *pixel* corresponde ao coeficiente de difusão aparente do *voxel* correspondente. Quanto maior o brilho do *pixel* no mapa ADC, maior o coeficiente de difusão aparente [1].

Este trabalho propõe uma nova abordagem para avaliação do progresso da doença de Alzheimer: uma vez que o mapa ADC usualmente apresenta *pixels* com intensidades não nulas em regiões não ocupadas pela amostra, uma certa incerteza também pode ser levantada quanto aos coeficientes de difusão do interior da amostra, sendo sensível à presença de ruídos nas imagens [1].

Neste estudo de caso, as imagens são utilizadas na composição de uma única imagem mul-

tiespectral, onde cada uma das imagens ponderadas em difusão se comporta como uma banda espectral de uma imagem multiespectral sintética. A imagem multiespectral sintética assim montada é então classificada utilizando os seguintes métodos: as redes neurais perceptron multicamadas (MLP), as redes *fuzzy c*-médias, os mapas auto-organizados de Kohonen, os classificadores LVQ de Kohonen, as redes de função de base radial e o classificador dialético objetivo.

Segundo diversos trabalhos, é comum assumir que não é possível distinguir entre substância cinzenta e substância branca em imagens ponderadas em difusão [154, 155, 156, 157, 158, 159, 160]. Este trabalho mostra que essa distinção entre substância cinzenta e substância branca é possível utilizando-se classificadores não supervisionados adaptativos, em especial o classificador dialético objetivo, proposto neste trabalho, para classificação da imagem multiespectral sintética construída, permitindo observar a redução da substância cinzenta no lóbulo frontal, onde se localiza a memória recente, o que é fundamental para avaliação do progresso da doença de Alzheimer.

## 7.2 Materiais e Métodos

### 7.2.1 Imagens Ponderadas em Difusão e Mapas ADC

As imagens de ressonância magnética (RM) ponderadas em difusão foram obtidas a partir do banco de imagens clínicas do Laboratório de Imagens de Ressonância Magnética do Departamento de Física da Universidade Federal de Pernambuco. O banco é composto de imagens clínicas reais obtidas por um tomógrafo clínico de RM de 1,5 T.

Neste trabalho foram utilizadas 80 imagens de ressonância magnética axiais ponderadas em difusão, ou seja, 4 volumes cerebrais de 20 fatias axiais cada, sendo um desses volumes composto por mapas ADC, correspondentes a um paciente do sexo masculino, de 70 anos de idade, portador da doença de Alzheimer.

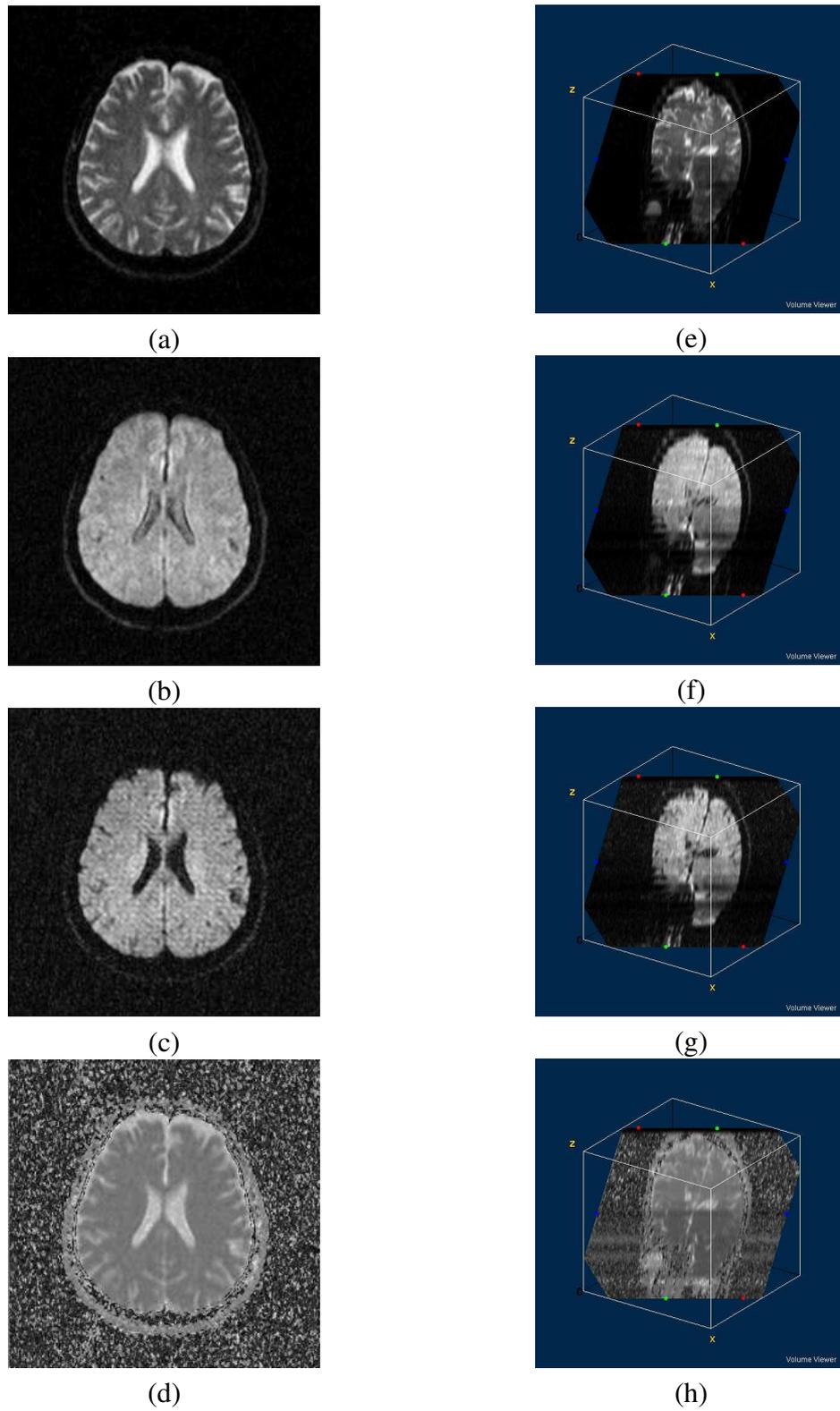
Para realizar o treinamento foi escolhida a fatia 13 de cada amostra volumétrica de 20 fatias (figuras 7.1 (a), (b) e (c)), dado que essa fatia mostra os cornos temporais dos ventrículos laterais, o que permite uma avaliação mais clara por parte do especialista e facilita a correlação entre os dados gerados pela ferramenta computacional e o conhecimento *a priori* do especialista. Além disso, a fatia 13 apresenta artefatos fora da região cranial em quantidade considerável.

Pode-se considerar imagens como funções matemáticas, onde seu domínio é uma região do plano dos inteiros, a *grade*, e seu contradomínio é o conjunto dos valores possíveis de serem assumidos pelos *pixels* correspondentes a cada posição da grade.

Seja  $f_i : S \rightarrow W$  o conjunto das imagens de RM ponderadas em difusão, onde  $1 \leq i \leq 3$ ,  $S \subseteq \mathbb{Z}^2$  é a grade da imagem  $f_i$ , onde  $W \subseteq \mathbb{R}$  é seu contradomínio. Cada imagem multiespectral sintética  $f : S \rightarrow W^3$  integrante do volume multiespectral a ser classificado é composta pelas 3 imagens de RM ponderadas em difusão obtidas dos volumes mostrados nas figuras 7.1 (e), (f) e (g), sendo dada por:

$$f(\mathbf{u}) = (f_1(\mathbf{u}), f_2(\mathbf{u}), f_3(\mathbf{u}))^T, \quad (7.1)$$

onde  $\mathbf{u} \in S$  é a posição do *pixel* na imagem  $f$ , e  $f_1$ ,  $f_2$  e  $f_3$  são as imagens de RM de difusão. Considerando que cada *pixel*  $f_i(\mathbf{u})$  é aproximadamente proporcional ao sinal do voxel corres-



**Figura 7.1** Imagens da fatia 13 (esquerda) e volumes (direita) ponderados em difusão, com expoentes de 0 s/mm<sup>2</sup> (a, e), 500 s/mm<sup>2</sup> (b, f) e 1000 s/mm<sup>2</sup> (c, g), e mapa ADC (d, h)

pondente como segue [161, 162, 163, 164, 165, 166, 167]:

$$f_i(\mathbf{u}) = K\rho(\mathbf{u})e^{-T_E/T_2(\mathbf{u})}e^{-b_iD_i(\mathbf{u})}, \quad (7.2)$$

onde  $D_i(\mathbf{u})$  é o coeficiente de difusão de spin associado ao *voxel* da posição  $\mathbf{u}$ , medido após o  $i$ -ésimo experimento;  $\rho(\mathbf{u})$  é a densidade de spins nucleares no referido *voxel*;  $K$  é uma constante de proporcionalidade;  $T_2(\mathbf{u})$  é o tempo de relaxação transversal no *voxel*;  $T_E$  é o tempo de eco e  $b_i$  é o expoente de difusão, dado por [1, 22, 168]:

$$b_i = \gamma^2 G_i^2 T_E^3 / 3, \quad (7.3)$$

onde  $\gamma$  é a constante giromagnética e  $G_i$  é o gradiente aplicado no  $i$ -ésimo experimento. As figuras 7.1 (e), (f) e (g) mostram os volumes de imagens ponderadas em difusão com expoentes de difusão 0 s/mm<sup>2</sup>, 500 s/mm<sup>2</sup> e 1000 s/mm<sup>2</sup>, respectivamente.

A análise de imagens de RM ponderadas em difusão é usualmente realizada através dos mapas ADC,  $f_{\text{ADC}} : S \rightarrow W$ , calculados como segue [169]:

$$f_{\text{ADC}}(\mathbf{u}) = \frac{C}{b_2} \ln \left( \frac{f_1(\mathbf{u})}{f_2(\mathbf{u})} \right) + \frac{C}{b_3} \ln \left( \frac{f_1(\mathbf{u})}{f_3(\mathbf{u})} \right), \quad (7.4)$$

onde  $C$  é uma constante de proporcionalidade.

Considerando  $n$  experimentos, pode-se generalizar a equação 7.4 como segue:

$$f_{\text{ADC}}(\mathbf{u}) = \sum_{i=2}^n \frac{C}{b_i} \ln \left( \frac{f_1(\mathbf{u})}{f_i(\mathbf{u})} \right). \quad (7.5)$$

Portanto, o mapa ADC é dado por:

$$f_{\text{ADC}}(\mathbf{u}) = C\bar{D}(\mathbf{u}), \quad (7.6)$$

onde  $\bar{D}(\mathbf{u})$  é a média amostral das medidas do coeficiente de difusão  $D(\mathbf{u})$  [163, 170, 171].

Portanto, os *pixels* do mapa ADC são proporcionais aos coeficientes de difusão nos *voxels* correspondentes. Entretanto, como as imagens são adquiridas em diferentes momentos, deve-se considerar a ocorrência de ruído em todos os experimentos. Nas figuras 7.1 (d) e (h) podem-se notar diversos artefatos devidos à presença de ruído. Nas regiões da imagem onde a relação sinal-ruído é pobre (por exemplo,  $s/n \approx 1$ ), o mapa ADC produz artefatos como consequência do cálculo do logaritmo (ver equações 7.4 e 7.5).

Esses fatores levam à seguinte conclusão: os *pixels* do mapa ADC não necessariamente correspondem aos coeficientes de difusão, uma vez que diversos *pixels* indicam altas taxas de difusão em *voxels* onde simplesmente não há amostra ou em áreas muito sólidas, como a caixa cranial, por exemplo, como pode ser notado nas figuras 7.1 (d) e (h). Essa é a razão pela qual esses mapas indicam coeficientes de difusão *aparentes*, e não *reais*.

Neste trabalho propõe-se uma alternativa à análise do mapa ADC: a análise multiespectral do volume composto pelas imagens  $f : S \rightarrow W^3$  utilizando métodos baseados em redes neurais e no classificador dialético objetivo.

### 7.2.2 Análise do Mapa ADC usando Fuzzy C-Médias

Para comparar os métodos propostos e o mapa ADC é preciso extrair informação quantitativa e qualitativa do mapa ADC. Isso é possível aplicando uma classificação mono-espectral não supervisionada ao mapa ADC utilizando um método baseado em agrupamento [172, 173]. Foi escolhido então o *fuzzy c*-médias (ADC-CM) com 3 entradas, 3 saídas, máximo de 200 iterações e taxa de aprendizado inicial  $\eta_0 = 0,1$ .

As 3 entradas são associadas ao vetor  $\mathbf{x} = (x_1, x_2, x_3)^T$ , onde  $x_i = f_{\text{ADC}}(\mathbf{u})$ , para  $1 \leq i \leq 3$ . As saídas da rede representam as classes de interesse e estão associadas ao vetor  $\mathbf{y} = (y_1, y_2, y_3)^T$ , onde cada saída corresponde à classe de mesmo índice. O critério de decisão empregado é o mesmo da subseção 7.2.3: a maior saída indica a classe mais provável [133, 174]. O conjunto de treinamento é composto por *pixels* da imagem  $f_{\text{ADC}}$  e construído usando conhecimento especialista na seleção das regiões de interesse [7].

### 7.2.3 Análise Multiespectral utilizando Redes Neurais

Seja o universo de classes de interesse  $\Omega = \{C_1, C_2, C_3\}$ ,  $C_1$  representa os coeficientes de difusão associados ao fluido cérebro-espinal (também chamado de líquido);  $C_2$ , substância branca e cinzenta, uma vez que, a partir dos mapas ADC, elas não podem ser diferenciadas por meio de imagens de difusão, pois seus coeficientes de difusão são muito próximos;  $C_3$  corresponde ao fundo da imagem.

Para a análise multiespectral utilizando redes neurais, as entradas são associadas ao vetor  $\mathbf{x} = (x_1, x_2, x_3)^T$ , onde  $x_i = f_i(\mathbf{u})$ , para  $1 \leq i \leq 3$ . As saídas da rede representam as classes de interesse e estão associadas ao vetor  $\mathbf{y} = (y_1, y_2, y_3)^T$ , onde cada saída corresponde à classe de mesmo índice. O critério de decisão é baseado no critério de Bayes: a maior saída indica a classe mais provável [133, 174]. O conjunto de treinamento é construído utilizando conhecimento especialista durante a seleção das regiões de interesse [7].

O volume de imagens multiespectrais sintéticas foi classificado usando os seguintes métodos:

1. *Perceptron multicamadas* (MLP): Taxa de aprendizado inicial  $\eta_0 = 0,2$ , erro de treinamento  $\varepsilon = 0,05$ , máximo de 1000 iterações, 3 entradas, 3 saídas, 2 camadas, 60 neurônios na camada 1 [7];
2. *Rede de função de base radial* (RBF): 3 entradas, 2 camadas; camada 1: mapa de *k*-médias com 18 neurônios na camada 1, taxa de aprendizado inicial  $\eta_0 = 0,1$ , máximo de 200 iterações; camada 2: 3 saídas, máximo de 200 iterações, taxa de aprendizado inicial  $\eta_0 = 0,1$  [7].
3. *Classificador LVQ de Kohonen* (LVQ): 3 entradas, 3 saídas, máximo de 200 iterações, taxa de aprendizado inicial  $\eta_0 = 0,1$  [7].
4. *Mapa auto-organizado de Kohonen* (KO): 3 entradas, 3 saídas, máximo de 200 iterações, taxa de aprendizado inicial  $\eta_0 = 0,1$  [7].

O número de entradas corresponde ao número de bandas das imagens de entrada, dado que a classificação é feita *pixel a pixel*. Já o número de saída corresponde ao número de classes. Os demais parâmetros foram definidos empiricamente.

O perceptron multicamadas foi escolhido para se avaliar o desempenho da classificação multiespectral baseada em redes neurais clássicas de duas camadas. O número de entradas e saídas corresponde ao número de bandas e de classes de interesse, respectivamente. O erro de treinamento foi escolhido considerando o máximo ruído estimado em imagens ponderadas em difusão. O número de neurônios na camada 1 e a taxa de aprendizado foram determinados empiricamente.

A rede de função de base radial foi escolhida para avaliar o desempenho da classificação multiespectral baseada em uma estratégia orientada a problemas locais. O número de entradas e saídas corresponde ao número de bandas e de classes de interesse, respectivamente. A taxa de aprendizado foi determinada empiricamente.

O mapa auto-organizado de Kohonen foi escolhido para avaliar o desempenho de um método de classificação não supervisionada na tarefa de classificação multiespectral baseada em agrupamentos. Já o classificador LVQ de Kohonen foi escolhido como alternativa baseada em classificação supervisionada ao perceptron multicamadas e à rede de função de base radial.

#### 7.2.4 Análise Multiespectral pelo Mapa Fuzzy C-Médias

O *mapa fuzzy c-médias* (CM) também foi utilizado para realizar a classificação das imagens multiespectrais sintéticas geradas a partir das imagens de difusão, de forma semelhante às redes neurais. Foi treinado um mapa de 3 entradas e 3 saídas, durante um máximo de 200 iterações, com taxa de aprendizado inicial  $\eta_0 = 0,1$ .

#### 7.2.5 Análise Multiespectral pelo Classificador Dialético

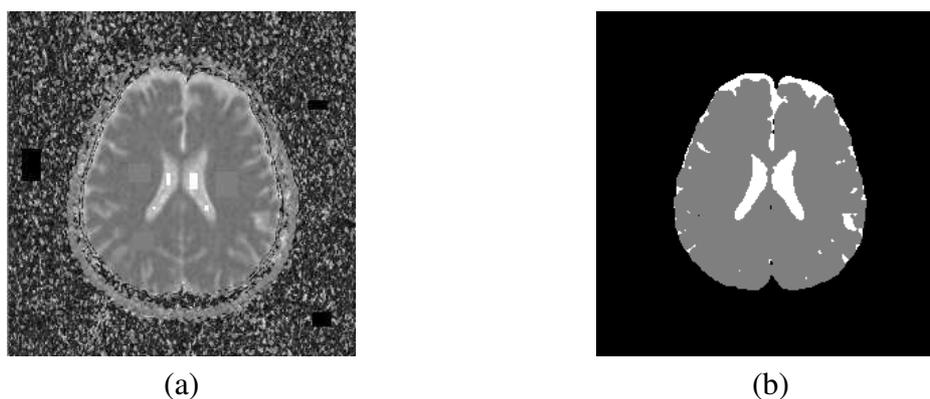
O *classificador dialético objetivo* (ODC) foi treinado a partir de um sistema inicial com 10 classes integrantes, condicionado por 3 condições de entrada, estudado durante 5 fases históricas, com duração de 100 cada e passo histórico inicial  $\eta_0 = 0,1$ . Nas etapas de crise revolucionária, considera-se uma medida de força mínima de 1%, contradição mínima de 25% e máxima crise de 25%. O critério de parada utilizado foi o número de classes final, no caso, 4 classes. As condições de entrada são os valores dos *pixels* em cada uma das 3 bandas.

Para acelerar a convergência, não foi utilizado o operador de geração de novas classes na etapa de crise revolucionária, durante o treinamento do classificador. Como funções de anticontração foram utilizadas as seguintes funções:

$$g_i(\mathbf{x}) = \left( \sum_{k=1}^{n_C} \frac{\|\mathbf{x} - \mathbf{w}_i\|^2}{\|\mathbf{x} - \mathbf{w}_k\|^2} \right)^{-1}, \quad (7.7)$$

onde  $1 \leq i \leq n_C$ . Os pesos foram atualizados na etapa de evolução, em cada fase histórica, pela seguinte expressão:

$$w_{i,j}(t+1) = \begin{cases} w_{i,j}(t) + \eta(t)g_i^2(\mathbf{x}(t))(x_j(t) - w_{i,j}(t)), & i = k \\ w_{i,j}(t), & i \neq k \end{cases}, \quad (7.8)$$



**Figura 7.2** Conjunto de treinamento (a) e imagem verdade da fatia 13 gerada pela classificação usando a rede polinomial (b)

onde  $\eta(t)$  é o passo da fase histórica e pode ser reduzido a cada iteração de forma semelhante à utilizada na atualização da taxa de aprendizado no treinamento das redes auto-organizáveis de Kohonen [7].

### 7.2.6 Ferramentas Computacionais

Para implementar os métodos utilizados e reconstruir os volumes, foi desenvolvida uma ferramenta computacional que recebeu o nome de *AnImed*, construída utilizando a linguagem de programação orientada a objeto *Object Pascal* no ambiente de desenvolvimento *Delphi 5*. A visualização dos volumes foi efetuada por meio da ferramenta baseada em *Java*, *ImageJ*, desenvolvida pelo NIH (*National Institute of Health*, EUA) e do *plugin VolumeViewer*.

## 7.3 Resultados

Para avaliar objetivamente os resultados da classificação volumétrica, foram utilizados três métodos: o índice  $\kappa$ , a *acurácia global* e a *matriz de confusão*. A avaliação subjetiva foi realizada utilizando conhecimento especialista. As classes fundo da imagem ( $C_3$ ), substância branca e cinzenta ( $C_2$ ) e líquido cérebro-espinhal ( $C_1$ ) foram associadas às seguintes cores: preto, cinza e branco, respectivamente. A figura 7.2 (a) mostra o conjunto de treinamento montado sobre a fatia 13 do volume de mapas ADC, enquanto a figura 7.2 (b) mostra a imagem verdade para a fatia 13.

Para a tarefa de classificação, assume-se que as classes de interesse são separáveis por hiperquádricas. Portanto, uma rede polinomial de segundo grau foi escolhida para classificar o volume multiespectral original e gerar um volume verdade.

O grau do polinômio foi empiricamente determinado pelo seu incremento gradual até que não houvesse diferenças significativas entre a classificação presente e a classificação imediatamente anterior. Essas diferenças foram avaliadas quantitativamente usando o índice de fidelidade de Wang [142], um índice estatístico de similaridade. O índice de Wang é dado por

[142]:

$$Q_w(f, g) = \frac{4\mu_f\mu_g\sigma_{f,g}}{(\mu_f^2 + \mu_g^2)(\sigma_f^2 + \sigma_g^2)}, \quad (7.9)$$

onde  $f$  e  $g$  são as imagens monoespectrais a serem comparadas, enquanto  $\mu_f$ ,  $\mu_g$ ,  $\sigma_f^2$ ,  $\sigma_g^2$  e  $\sigma_{f,g}$  são a média de  $f$ , a média de  $g$ , a variância de  $f$ , a variância de  $g$ , e a correlação entre  $f$  e  $g$ , respectivamente.

Assim, as similaridades entre os resultados de classificação obtidos por polinômios de quarto grau e de terceiro grau em relação ao resultado do polinômio de grau dois foram 0,9816 e 0,9904, respectivamente. Optou-se então por escolher o resultado da classificação pelo polinômio de segundo grau.

A rede polinomial é uma rede de duas camadas: a primeira é uma rede multiplicativa que gera todos os termos do polinômio de grau 2 a partir das 3 entradas; a segunda camada consiste em um perceptron de camada única com taxa de aprendizado inicial  $\eta_0 = 0,1$  e erro de treinamento  $\varepsilon = 0,05$ , máximo de 200 iterações de treinamento, responsável pelo cálculo dos coeficientes do polinômio que modela as funções discriminantes de cada classe [133, 174]. A rede polinomial é um aproximador polinomial. A taxa de aprendizado e o erro de treinamento foram determinados empiricamente.

A *matriz de confusão* para o universo de classes de interesse  $\Omega$  é uma matriz  $m \times m$ ,  $\mathbf{T} = [t_{i,j}]_{m \times m}$ , onde cada elemento  $t_{i,j}$  representa o número de objetos pertencentes à classe  $C_j$  classificados como  $C_i$ , onde  $\Omega = \{C_1, C_2, \dots, C_m\}$  [133, 175].

A *acurácia global*  $\phi$  é a razão entre o número de objetos corretamente classificados e o total de objetos, definida como segue [133, 175]:

$$\phi = \rho_v = \frac{\sum_{i=1}^m t_{i,i}}{\sum_{i=1}^m \sum_{j=1}^m t_{i,j}}. \quad (7.10)$$

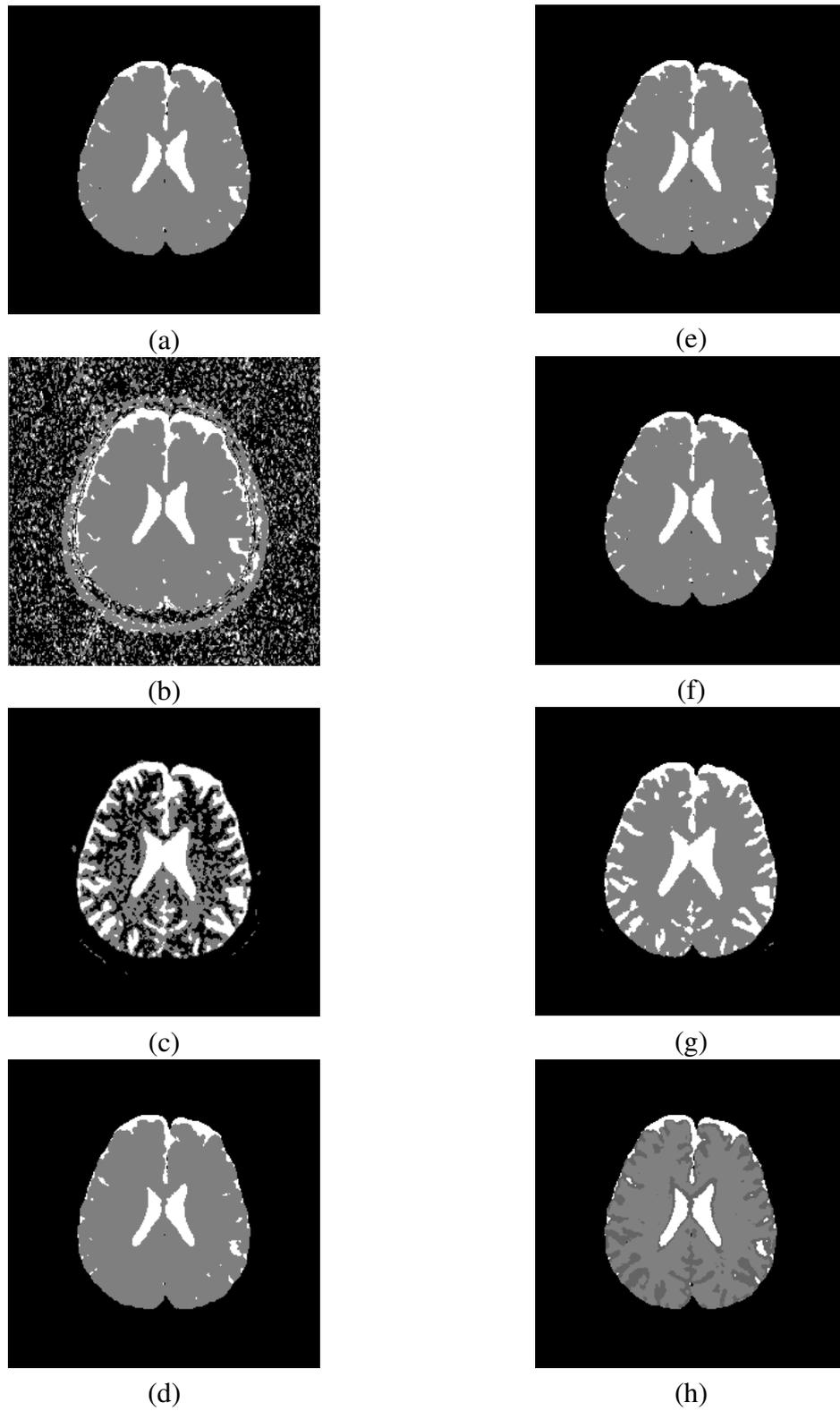
O *índice*  $\kappa$  é um índice de correlação estatística definido como segue [133]:

$$\kappa = \frac{\rho_v - \rho_z}{1 - \rho_z}, \quad (7.11)$$

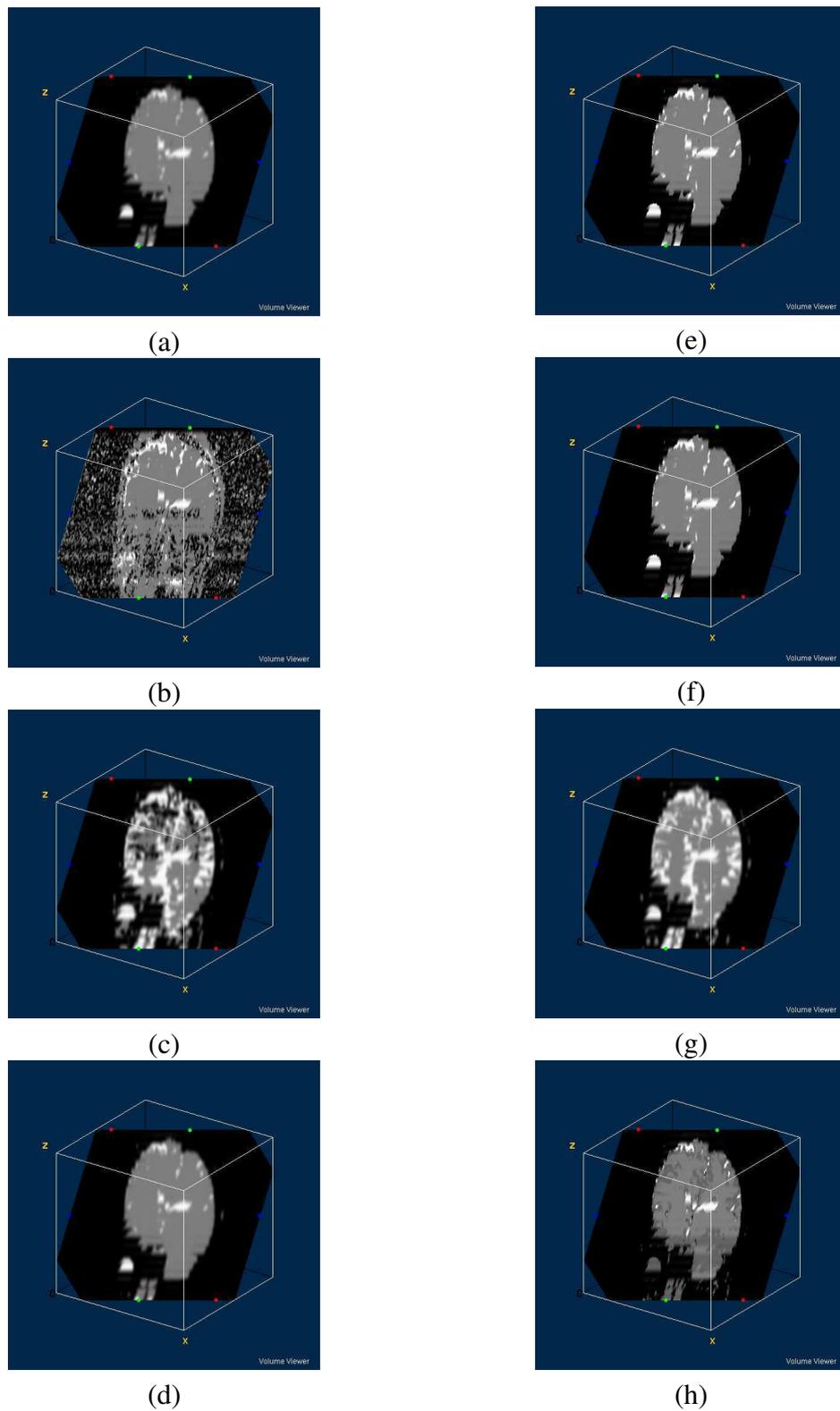
onde

$$\rho_z = \frac{\sum_{i=1}^m (\sum_{j=1}^m t_{i,j})(\sum_{j=1}^m t_{j,i})}{(\sum_{i=1}^m \sum_{j=1}^m t_{i,j})^2}. \quad (7.12)$$

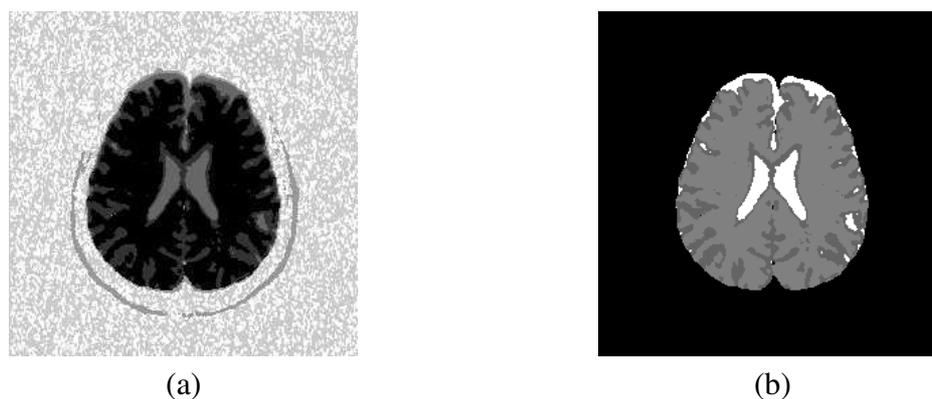
A figura 7.4 (b) mostra o resultado da classificação do mapa ADC usando o método *fuzzy* c-médias. As figuras 7.4 (c), (d), (e), (f) e (g) mostram os resultados da classificação do volume de imagens multiespectrais sintéticas composto pelos volumes das imagens das figuras 7.1 (e), (f) e (g) usando os métodos MLP, RBF, KO, LVQ e CM, respectivamente. A tabela 7.1 apresenta o índice  $\kappa$  e a acurácia global  $\phi$ , enquanto a tabela 7.2 mostra os volumes percentuais  $V_1$ ,  $V_2$  e  $V_3$  ocupados pelas classes de interesse  $C_1$ ,  $C_2$  e  $C_3$ , nesta ordem, bem como a razão entre o volume de líquido cérebro-espinal e o volume total de substância branca e cinzenta, denominada simplesmente de razão fluido-substância, expressa por  $V_1/V_2$ . A figura 7.3 (b) mostra a fatia 13 do volume da classificação do mapa ADC pelo método CM, enquanto as



**Figura 7.3** Fatia 13 da classificação usando PO (a), ADC-CM (b), MLP (c), RBF (d), KO (e), LVQ (f), CM (g), e ODC (h)



**Figura 7.4** Volumes de classificação usando PO (a), ADC-CM (b), MLP (c), RBF (d), KO (e), LVQ (f), CM (g), e ODC (h)



**Figura 7.5** Fatia 13 da classificação pelo classificador dialético objetivo antes (a) e depois (b) da pós-rotulação

**Tabela 7.1** Acurácia global  $\phi$  (%) e índice  $\kappa$  para os métodos de classificação

	MLP	RBF	KO	LVQ	CM	ADC-CM
$\phi$ (%)	88,5420	99,3587	99,3647	99,3683	93,9364	58,1154
$\kappa$	0,6081	0,9681	0,9688	0,9689	0,7755	0,2495

figuras 7.3 (c), (d), (e), (f) e (g) mostram as fatias 13 dos volumes de 20 fatias gerados pela classificação segundo os métodos MLP, RBF, KO, LVQ e CM, nesta ordem.

A figura 7.5 (a) mostra a fatia 13 do resultado da classificação utilizando o classificador dialético objetivo. A figura 7.5 (b) mostra a fatia 13 do resultado após efetuada a pós-rotulação. A figura 7.4 (h) mostra todo o volume gerado pela classificação pelo ODC.

O treinamento do ODC resultou em 6 classes, reduzidas a 4 após a pós-rotulação manual que agrupou as 3 classes presentes fora do crânio, fundo da imagem, ruído e caixa craniana, numa só classe, fundo da imagem. A pós-rotulação é feita manualmente porque as duas populações são estatisticamente diferentes e só conceitualmente fazem parte de uma única classe. Da figura 7.5 (b) pode-se perceber que o ODC conseguiu fazer distinção entre substância branca e substância cinzenta, essa última presente na interface entre o líquido e a substância branca.

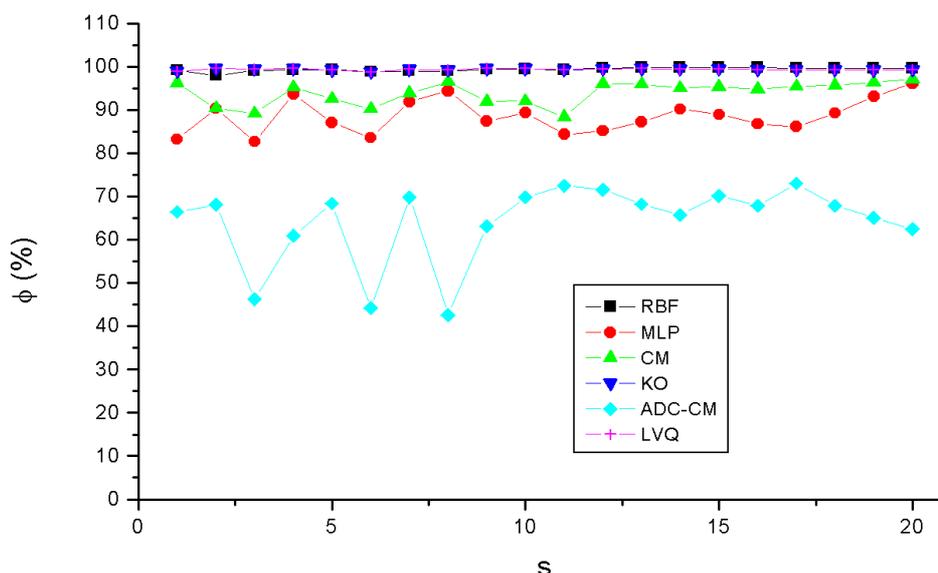
O treinamento dos métodos foi realizado a partir da fatia 13 e generalizado para as outras

**Tabela 7.2** Volumes percentuais e razão fluido-substância para os métodos de classificação

	MLP	RBF	KO	LVQ	CM	ADC-CM	PO
$V_1$ (%)	7,607	1,612	2,186	2,185	6,262	18,743	1,697
$V_2$ (%)	11,546	17,187	16,398	16,394	13,922	34,354	17,010
$V_3$ (%)	80,847	81,201	81,416	81,421	79,816	46,903	81,293
$V_1/V_2$	0,659	0,094	0,133	0,133	0,450	0,546	0,100

**Tabela 7.3** Índice de generalização dos métodos para as demais bandas

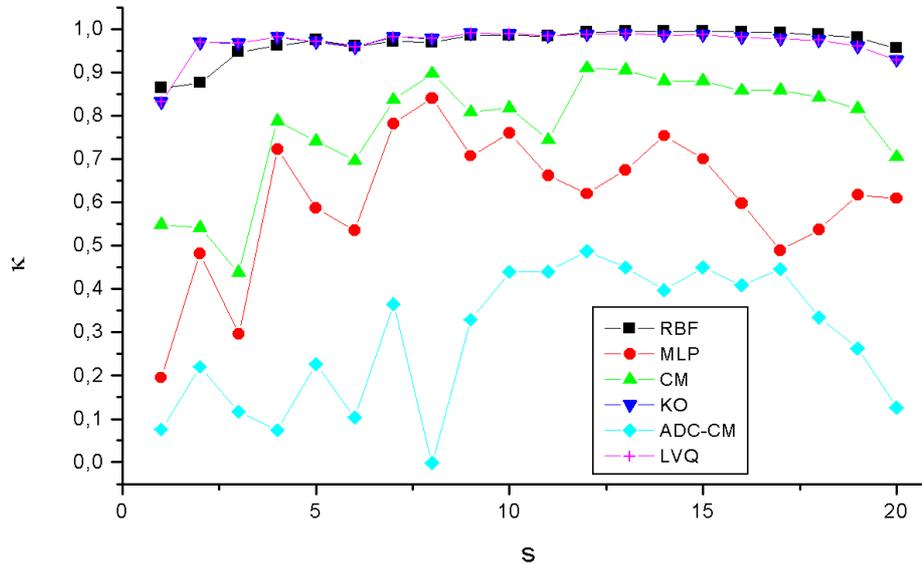
	MLP	RBF	KO	LVQ	CM	ADC-CM
$\Delta\kappa/\bar{\kappa}$	0,7398	0,9623	0,9634	0,9635	0,8295	0,4519
$\Delta\phi/\bar{\phi}$	0,9560	0,9956	0,9979	0,9979	0,9715	0,8577

**Figura 7.6** Comportamento da acurácia global  $\phi$  em função da fatia  $s$ 

19 fatias. As figuras 7.6 e 7.7 ilustram o comportamento dos índices  $\phi$  e  $\kappa$  em função da fatia  $s$ , onde  $1 \leq s \leq 20$ . A partir desses gráficos é possível estudar o *índice de generalização* de cada método, definido aqui como  $\Delta\kappa/\bar{\kappa}$ , resultando nos valores da tabela 7.3, onde se nota o relativamente baixo índice de generalização do método MLP, de 0,7398, frente aos outros métodos, de 0,8295 a 0,9635. Já o método ADC-CM possui o pior índice de generalização de todos, 0,4519.

Da tabela 7.3 também é possível concluir que não é possível definir um índice de generalização  $\Delta\phi/\bar{\phi}$ , uma vez que o menor valor, embora tenha sido para o método ADC-CM, foi de 0,8577, uma ordem de grandeza que simplesmente não reflete a verdade, se esse valor é comparado ao obtido com o índice de generalização baseado na estatística  $\kappa$ , de 0,4519.

As figuras 7.8 (a), (b) e (c) mostram imagens binárias da fatia 13 da classificação pelo ODC, mostrando o líquido (*cerebrospinal fluid*, CSF), a substância cinzenta (*gray matter*, GM) e a substância branca (*white matter*, WM), respectivamente, enquanto as figuras 7.9 e 7.10, mostram os respectivos espectros morfológicos e as distribuições morfológicas acumuladas.



**Figura 7.7** Comportamento do índice  $\kappa$  em função da variação da fatia  $s$

Tanto os espectros morfológicos quanto as distribuições morfológicas acumuladas foram obtidos usando como elemento estruturante o quadrado  $3 \times 3$  centrado em  $(2, 2)$ .

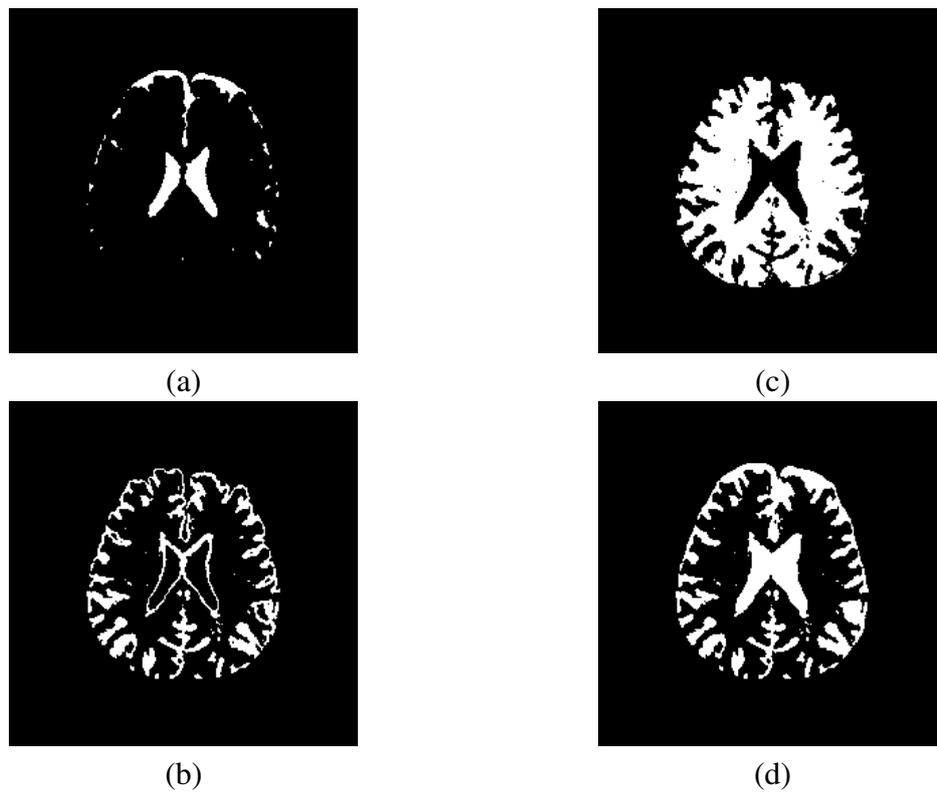
A figura 7.8 (d) ilustra o líquido mais a substância cinzenta (CSF+GM) da fatia 13 detectados pelo ODC, enquanto as imagens 7.9 e 7.10 mostram os gráficos do espectro morfológico e da distribuição morfológica acumulada associados.

As figuras 7.11 (a) e (b) mostram a fatia 13 do volume de líquido detectado pelos métodos CM e MLP, enquanto as figuras 7.12 e 7.13 mostram os respectivos espectros morfológicos e distribuições morfológicas acumuladas.

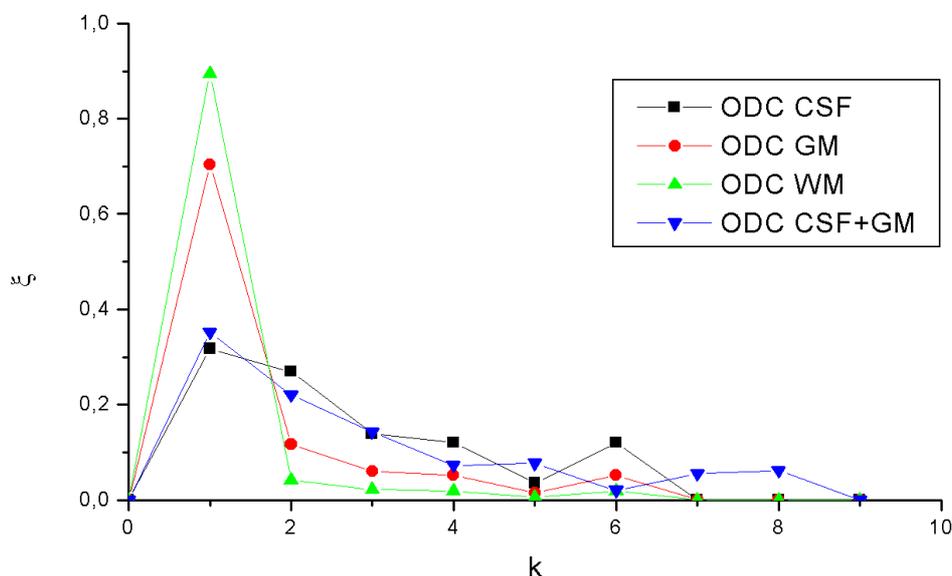
Comparando-se esses resultados com aqueles obtidos a partir do volume de líquido detectado pelo ODC, pode-se perceber que, na verdade, os métodos CM e MLP associaram líquido e substância cinzenta à mesma classe líquido. Isso se comprova tanto pelos espectros morfológicos quanto pelos valores do índice de similaridade morfológica entre os resultados do CM e do MLP em relação ao ODC, ou seja, 0,7274 e 0,8450, respectivamente.

As figuras 7.11 (c), (d), (e) e (f) mostram a fatia 13 do volume de líquido detectado pelos métodos KO, LVQ, RBF e PO, enquanto as figuras 7.14 e 7.15 mostram os respectivos espectros morfológicos e as distribuições morfológicas acumuladas.

Comparando-se esses resultados com aqueles obtidos a partir do volume de líquido detectado pelo ODC, pode-se perceber que, os métodos KO, LVQ, RBF e ODC detectaram corretamente a classe líquido, o que é atestado tanto pelos espectros morfológicos quanto pelos valores do índice de similaridade morfológica entre os resultados dos métodos KO, LVQ, RBF e ODC ao PO, ou seja, 0,8554, 0,8554, 0,9619 e 0,9006, respectivamente.



**Figura 7.8** Fatia 13 dos volumes de líquido (a), de substância cinzenta (b), de substância branca (c), e de líquido e substância cinzenta (d) detectados pelo ODC



**Figura 7.9** Espectros morfológicos da fatia 13 do volume de líquido, de substância branca e de substância cinzenta detectados pelo ODC

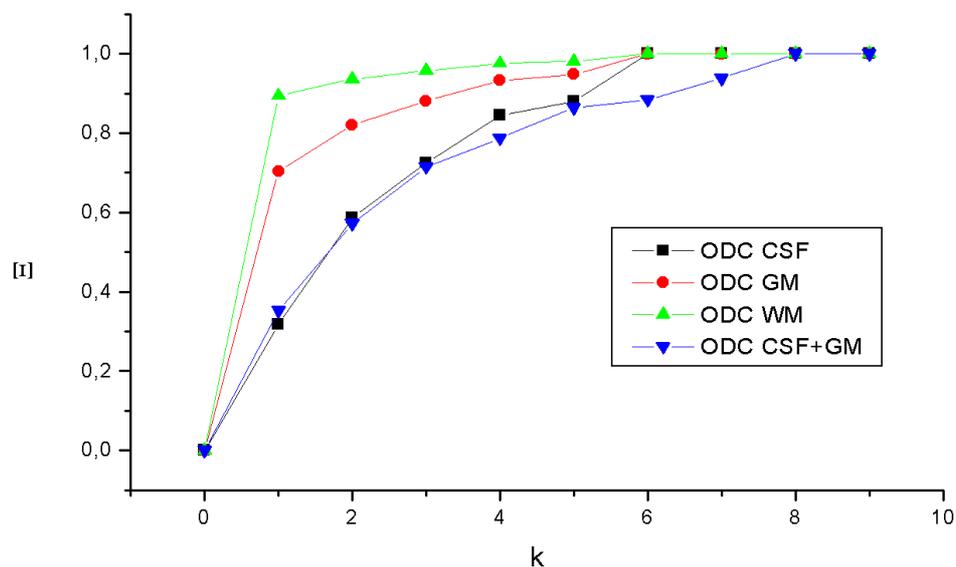
A partir do resultado obtido pelo ODC pode-se montar um novo conjunto de treinamento (ver figura 7.16 (b)) e, assim, treinar os classificadores supervisionados novamente. A tabela 7.4 mostra os valores de  $\kappa$  e  $\phi$  para a classificação da fatia 13 pelos métodos supervisionados MLP, LVQ e RBF em função do resultado obtido com o método PO para polinômios de grau 2. A tabela 7.5 mostra os valores das áreas e a razão líquido-substância para a fatia 13.

As figuras 7.16 (c), (d), (e) e (f) mostram os resultados da classificação pelos métodos MLP, RBF, LVQ e PO, respectivamente. Nota-se que a rede MLP retornou uma classificação onde as classes substância branca e substância cinzenta foram todas classificadas como substância cinzenta, o que não impediu que a razão líquido-substância tivesse um valor próximo do correto, apesar dos baixos valores de  $\kappa$  e  $\phi$  e da confusão entre substância cinzenta e fundo da imagem.

A fidelidade dos resultados obtidos pelos métodos MLP, RBF e LVQ ao resultado verdade obtido da classificação pela rede PO foi estimada usando o índice de similaridade morfológica, resultando 0,1533, 0,9877 e 0,9877, respectivamente.

## 7.4 Discussão e Conclusões

Os resultados da classificação multiespectral forneceram uma melhor avaliação, tanto quantitativa quanto qualitativa, dos volumes adquiridos, tornando possível a medição dos volumes de interesse, em especial a razão entre o volume ocupado pelo líquido cérebro-espinal e o volume



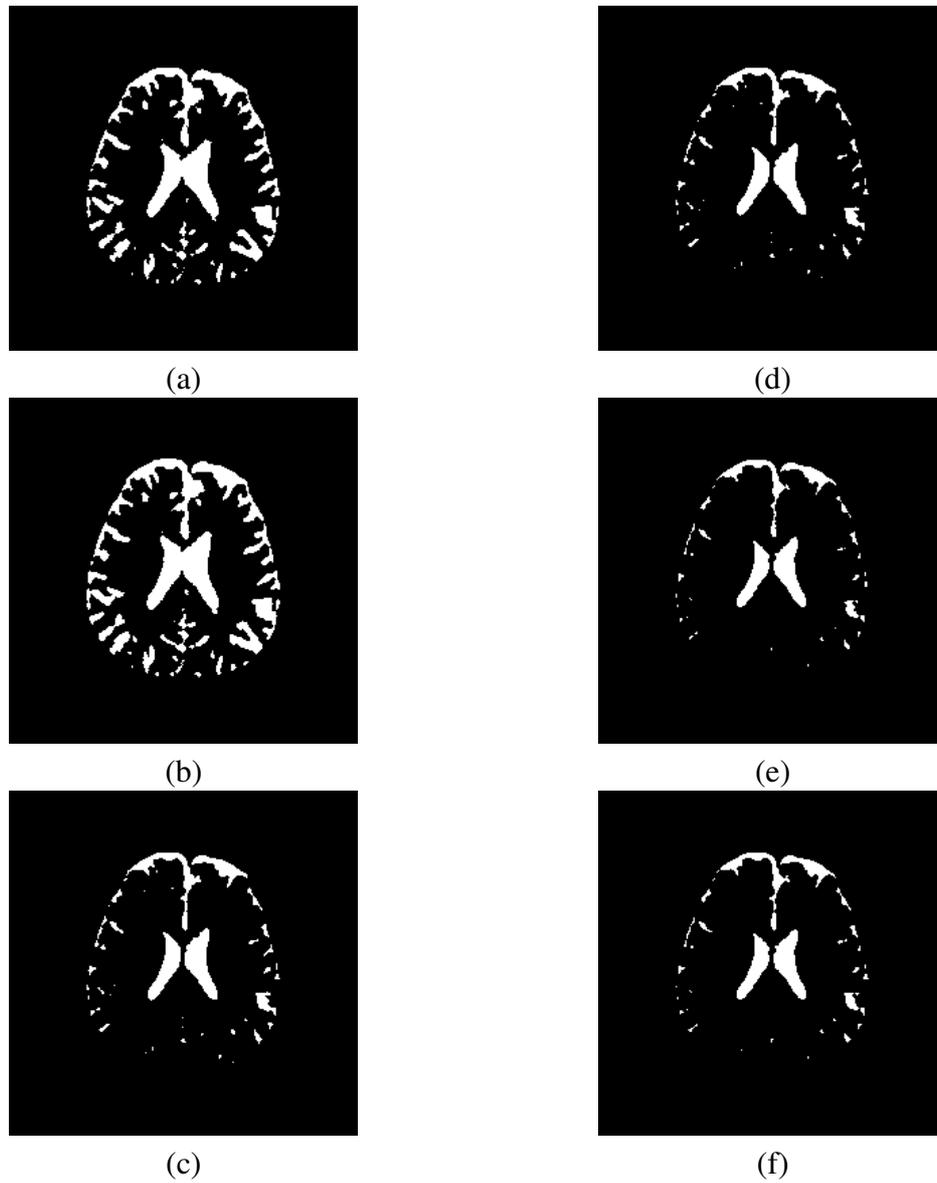
**Figura 7.10** Distribuições morfológicas acumuladas da fatia 13 do volume de líquido, de substância branca e de substância cinzenta detectados pelo ODC

**Tabela 7.4** Acurácia global  $\phi$  (%) e índice  $\kappa$  para os métodos de classificação treinados usando o resultado do ODC

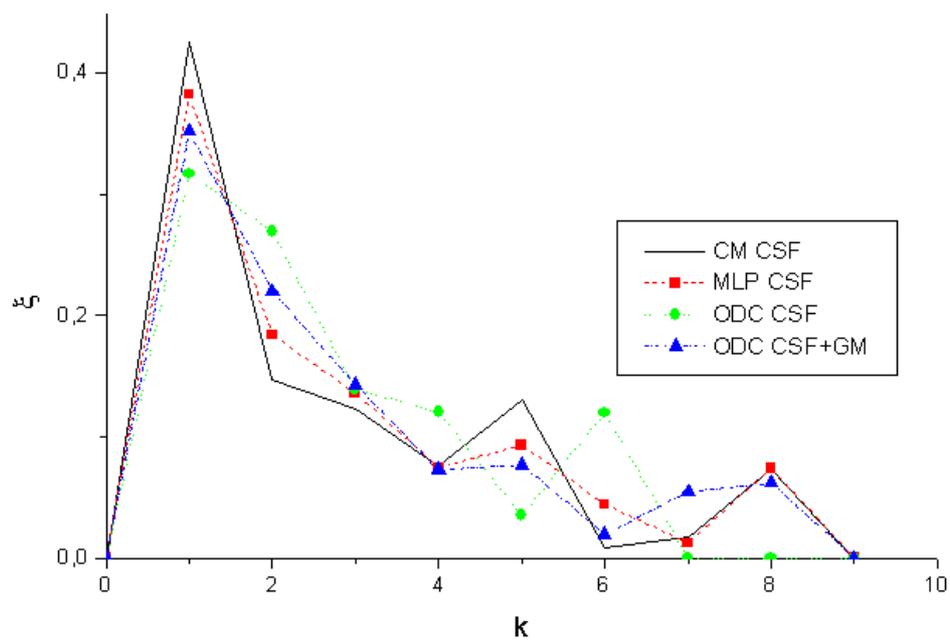
	MLP	RBF	LVQ
$\phi$ (%)	79,7699	98,6923	98,6923
$\kappa$	0,5370	0,9701	0,9701

**Tabela 7.5** Áreas percentuais e razão fluido-substância para os métodos de classificação treinados usando o resultado do ODC

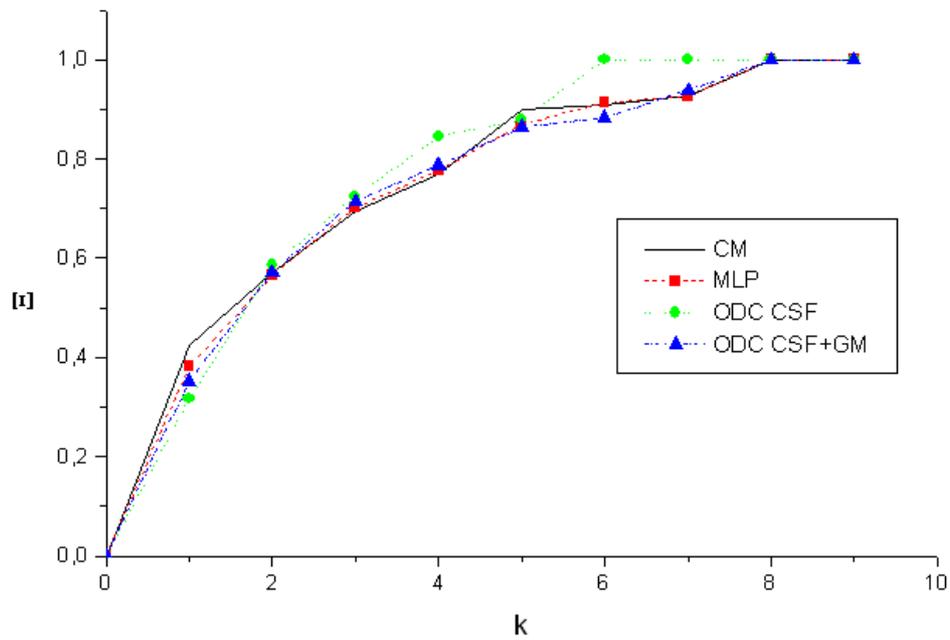
	MLP	RBF	LVQ	PO
$A_{CSF}$ (%)	2,681	2,820	2,820	2,458
$A_{GM}$ (%)	21,603	6,496	6,496	6,682
$A_{WM}$ (%)	0,000	18,256	18,256	18,481
$A_{CSF}/A_{GM+WM}$	0,124	0,114	0,114	0,098



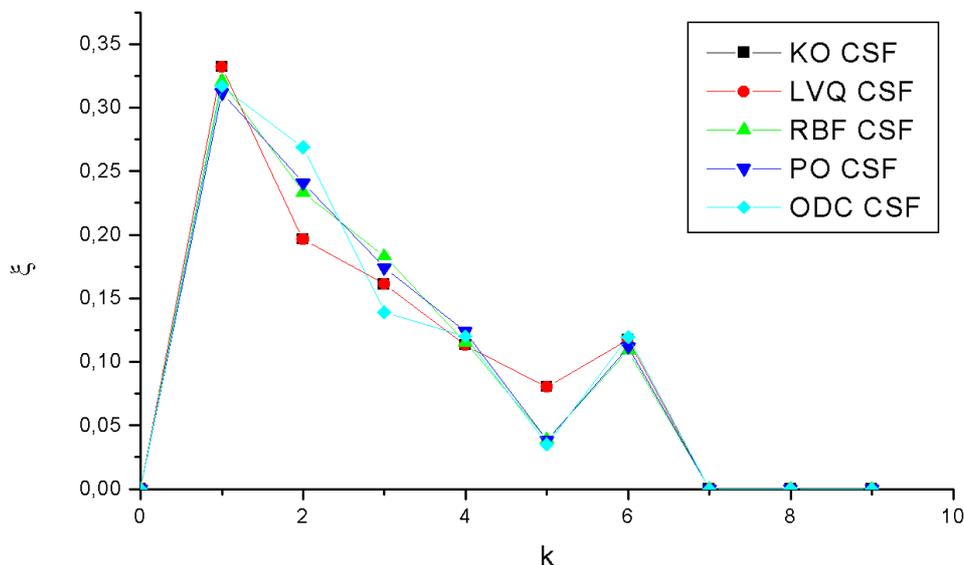
**Figura 7.11** Volumes da classe líquido detectados pelo CM (a), MLP (b), KO (c), LVQ (d), RBF (e) e PO (f)



**Figura 7.12** Espectros morfológicos da fatia 13 do volume da classe líquido detectados pelos métodos CM, MLP e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC



**Figura 7.13** Distribuições morfológicas acumuladas da fatia 13 do volume da classe líquido detectados pelos métodos CM, MLP e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC



**Figura 7.14** Espectros morfológicos da fatia 13 do volume da classe líquido detectados pelos métodos KO, LVQ, RBF, PO e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC

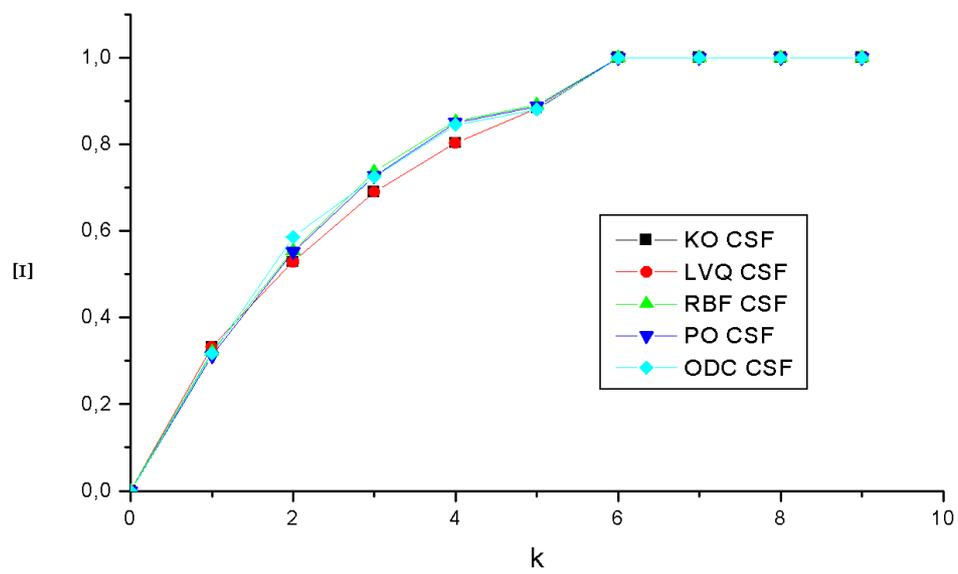
ocupado pelas substâncias branca e cinzenta, o que pode possibilitar o estabelecimento de uma correlação entre essa razão, juntamente com outras variáveis clínicas a serem consideradas, e o avanço de doenças neurodegenerativas, como a doença de Alzheimer.

Da tabela 7.1 pode-se notar que a abordagem multiespectral, com índice  $\kappa$  de 0,6081, 0,9681, 0,9688, 0,9689 e 0,7755, para o perceptron multicamadas, a rede de função radial, o mapa auto-organizado de Kohonen, o classificador LVQ de Kohonen e o mapa *fuzzy* c-médias, respectivamente, é superior à análise do mapa ADC pelo método *fuzzy* c-médias, com índice  $\kappa$  de 0,2495.

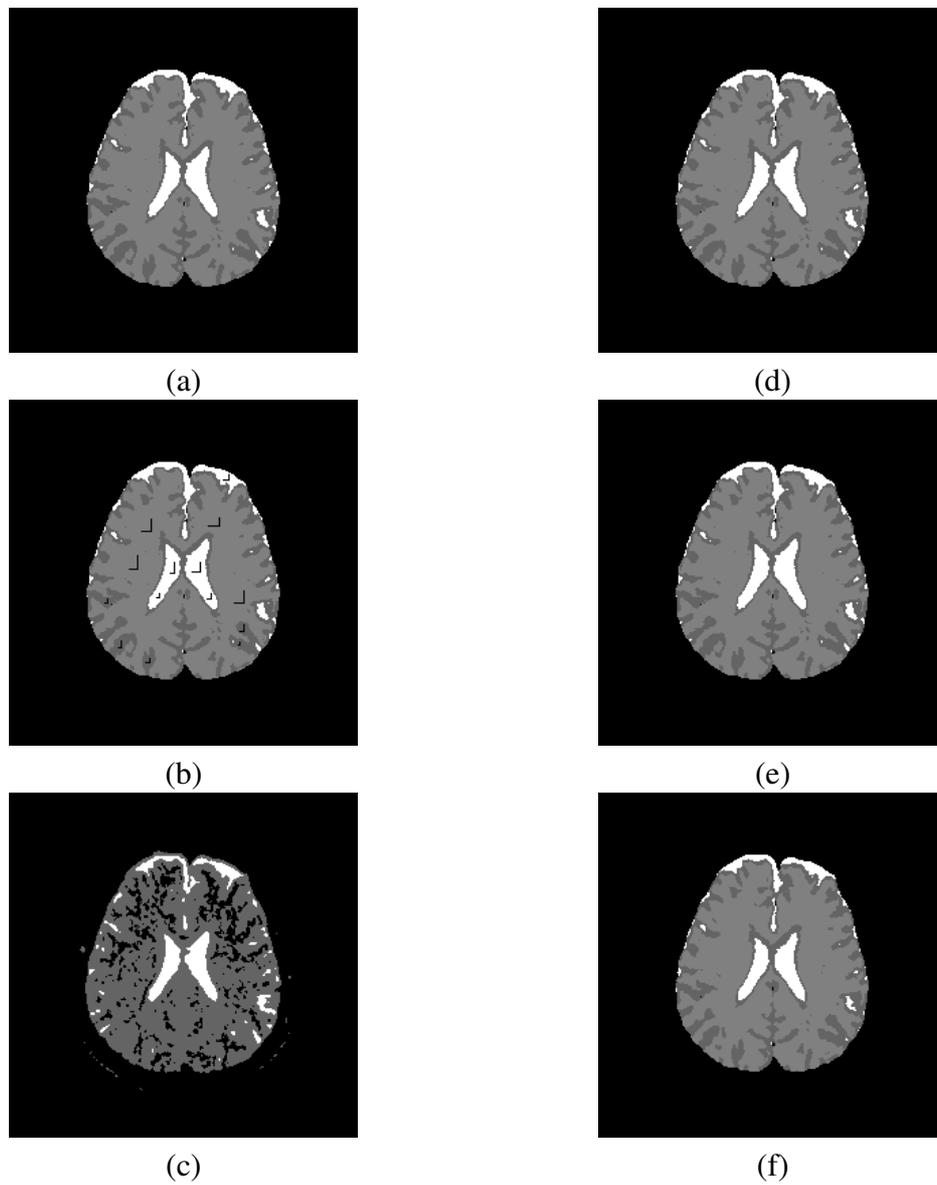
Esses resultados são confirmados quando se observam os volumes classificados das figuras 7.4 (c), (d), (e), (f) e (g), e se comparam com o volume classificado ADC da figura 7.4 (b), onde se podem notar diversas áreas fora da amostra e na caixa craniana marcadas como líquido cérebro-espinal ou substância branca ou cinzenta.

A partir dos resultados obtidos fica evidente que o perceptron multicamadas e o mapa *fuzzy* c-médias superestimaram o volume ocupado pelo líquido cérebro-espinal (ver figuras 7.4 (c) e (g)). Quando esse resultado é comparado com o volume com expoente de difusão 0 (ver figura 7.1 (e)), pode-se perceber que os ventrículos esquerdo e direito são separados. Além disso, os sulcos também foram superestimados, o que poderia levar um especialista a talvez considerar esse caso de Alzheimer como mais avançado do que realmente é.

A superestimação dos volumes também pode ser percebida nos resultados na tabela 7.2, onde a razão fluido-substância ( $V_1/V_2$ ) é 0,659 para o perceptron multicamadas e 0,450 para



**Figura 7.15** Distribuições morfológicas acumuladas da fatia 13 do volume da classe líquido detectados pelos métodos KO, LVQ, RBF, PO e ODC, e do volume de líquido e substância cinzenta detectado pelo método ODC



**Figura 7.16** Conjunto de treino (b) sobre o resultado do ODC (a) e novas classificações usando MLP (c), RBF (d), LVQ (e) e PO (f)

o mapa *fuzzy* c-médias, da mesma ordem, portanto, do resultado obtido pela classificação do mapa ADC pelo método *fuzzy* c-médias, com  $V_1/V_2 = 0,546$ , quase 6 vezes a taxa obtida pela rede de função de base radial, com  $V_1/V_2 = 0,094$ , pelo mapa auto-organizado de Kohonen e pelo classificador LVQ, com  $V_1/V_2 = 0,133$ , muito próximas, portanto, da taxa obtida pela rede polinomial,  $V_1/V_2 = 0,100$ . Consequentemente, as redes RBF, KO e LVQ podem ser consideradas boas estimadoras da razão fluido-substância, enquanto a rede MLP e o método CM poderiam ser descartados nesta aplicação, uma vez que o interesse principal aqui é a razão fluido-substância.

Contudo, a classificação utilizando o classificador dialético objetivo mostrou que, na verdade, a superestimação dos volumes resultante da classificação pelo perceptron multicamadas e pelo mapa *fuzzy* c-médias tem origem na associação das classes líquido e substância cinzenta na classe líquido. Obviamente, isso é muito dependente do conjunto de treinamento, no caso do perceptron multicamadas, já que este executa uma classificação supervisionada. Portanto, para um diferente conjunto de treinamento, e mesmo para uma quantidade diferente de neurônios na camada intermediária e diferentes parâmetros de treinamento da rede, o perceptron multicamadas resultante poderia ter associado a substância branca e a substância cinzenta a uma única classe, de forma semelhante à que se deu com o classificador LVQ e com a rede de função de base radial. Nesse caso, uma possibilidade para melhorar o resultado obtido pelo perceptron multicamadas seria utilizar algum método de otimização global para encontrar a melhor arquitetura para a rede neural.

No entanto, o resultado da classificação pela rede MLP e pelo mapa *fuzzy* c-médias, este, por sua vez, um método de classificação não supervisionada, ao confundir substância cinzenta com líquido, mostrou que, na verdade, há valores de coeficiente de difusão intermediários entre os coeficientes de difusão da substância cinzenta e de líquido, mostrando que é possível diferenciá-los. Isso é muito interessante, pois permite que se extraia muito mais informação de imagens de difusão do que os resultados provenientes da análise dos mapas ADC prevêem. Isso pode vir a otimizar a extração de informação de imagens de difusão, além de reduzir a relação custo benefício, pois novos exames para extração de informação anatômica, como a aquisição de imagens de ressonância magnética ponderadas em  $T_1$ ,  $T_2$  e densidade de prótons, podem se tornar desnecessários, sem contar o fato de que, muitas vezes, novos exames são um transtorno para o paciente, mesmo que usem tecnologias não invasivas, como a RM. Além disso, quanto mais tempo o paciente passa dentro do tomógrafo de RM, mais erros podem ocorrer, dado que é impossível que ele fique sem se movimentar dentro do aparelho, e sedar um paciente não é algo simples, podendo ter sérias consequências.

Observando a imagem da figura 7.8 (b), que indica a detecção de substância cinzenta pelo classificador dialético objetivo, pode-se notar a significativa redução de substância cinzenta no lóbulo frontal do cérebro, na parte superior da imagem, o que está diretamente relacionado com a redução da memória recente.

Em síntese, o classificador dialético objetivo se mostrou uma boa ferramenta para a extração de informação anatômica de imagens de difusão, permitindo fazer distinção entre as substâncias branca e cinzenta, o que não seria possível através da simples análise do mapa ADC.

O uso do classificador dialético objetivo permite realizar classificações que identifiquem classes estatisticamente significativas em situações onde o número inicial de classes é des-

conhecido. Isso permite a detecção de classes relevantes ou mesmo de singularidades não previstas inicialmente, o que pode auxiliar o especialista tanto na medição de volumes de interesse na tentativa de estabelecer uma correlação com o avanço de doenças neurodegenerativas, como a doença de Alzheimer, quanto na detecção de alterações significativas nos valores dos coeficientes de difusão medidos, por exemplo, potencialmente enriquecendo qualitativa e quantitativamente a análise do especialista.

# Detecção de Imagens Funcionais

## 8.1 Introdução

Um dos maiores desafios no campo das ciências cognitivas é mapear as regiões do cérebro responsáveis pelas funções motoras e comportamentais. A aquisição de Imagens de Ressonância Magnética Funcional (*functional Magnetic Resonance Imaging*, fMRI) é uma técnica fundamental para o estudo da atividade neural do cérebro humano, não sendo considerada invasiva quando comparada com a Tomografia por Emissão de Pósitrons (*Positron Emission Tomography*, PET) e com a Tomografia Computadorizada por Emissão de Fóton Único (*Single Photon Emission Computerized Tomography*, SPECT).

As imagens fMRI são adquiridas utilizando a sequência Eco-Planar de Aquisição de Imagens (*Echo Planar Imaging sequence*, EPI) e a técnica BOLD, que exploram tanto a propriedade diamagnética da oxi-hemoglobina quanto a propriedade paramagnética da desoxi-hemoglobina como um mecanismo de contraste natural [5], elevando o sinal nos *voxels* tipicamente de 1% a 5% em tomógrafos de 1,5 T [3, 4, 1, 6], produzindo artefatos na imagem resultante [1].

Um exame de fMRI consiste em três etapas de aquisição: aquisição *anatômica*, *sob estímulo* e *sem estímulo* [176]. Durante a primeira etapa são adquiridas imagens anatômicas de alta resolução sobre as quais é montado o *mapa de ativação*. Nas etapas de aquisição sob estímulo e sem estímulo são adquiridas todas as imagens para reconstrução do volume cerebral, sendo que, na aquisição sob estímulo, o voluntário está sob o experimento de interesse, enquanto na aquisição sem estímulo as imagens são adquiridas quando o voluntário se encontra em repouso. As imagens são adquiridas em quantidade suficiente para melhorar a relação sinal-ruído e portanto facilitar a detecção estatística das regiões com atividade neural relevante, chamadas aqui simplesmente de *regiões ativadas* [14, 5].

Devido ao baixo contraste das imagens adquiridas, para possibilitar a identificação das regiões ativadas é necessário realizar um pós-processamento sobre as imagens sob estímulo e sem estímulo. Os métodos de detecção podem ser estatísticos paramétricos ou não paramétricos.

Este capítulo apresenta uma nova abordagem para a detecção de regiões cerebrais ativadas: os métodos estatísticos clássicos de detecção de regiões ativadas baseados em testes de hipóteses são apresentados como métodos de análise de imagens multiespectrais sintéticas, sendo apresentada uma alternativa a esses métodos baseada na análise de imagens multiespectrais e mono-espectrais sintéticas utilizando métodos não paramétricos baseados em redes neurais e no classificador dialético objetivo.

Para validar esta proposta, foi realizado um estudo de caso onde foram adquiridas imagens funcionais eco-planares reais, sob estímulo e sem estímulo. Essas imagens são provenientes

de um experimento padrão onde se deseja detectar a área cerebral associada à atenção espacial visual. Para avaliar os métodos propostos foram calculadas as taxas de falsos positivos, falsos negativos, verdadeiros positivos e verdadeiros negativos, obtidas a partir das matrizes de confusão calculadas durante a etapa de pós-classificação.

## 8.2 Materiais e Métodos

As imagens de fMRI foram fornecidas pelo Departamento de Física e pelo Laboratório de Percepção Visual do Departamento de Psicologia da Universidade Federal de Pernambuco, Recife, Brasil. Essas imagens foram adquiridas a partir de um tomógrafo clínico de 1,5 T durante o ano de 2004. Foram adquiridas 60 imagens: 24 sob estímulo, 24 sem estímulo e 12 imagens de referência para corrigir os efeitos de movimento [177].

### 8.2.1 Métodos de Detecção Estatísticos Multiespectrais

Uma sequência temporal de imagens eco-planares pode ser modelada como uma única imagem multiespectral de  $n$  bandas,  $f : S \rightarrow [0, 1]^n$ , com vetor de ativação

$$\mathbf{v} = (v_1, v_2, \dots, v_n)^T,$$

onde, se  $v_k = 1$ , a banda  $f_k : S \rightarrow [0, 1]$  corresponde a uma imagem sob estímulo; caso  $v_k = 0$ , a banda  $f_k : S \rightarrow [0, 1]$  está associada a uma imagem sem estímulo;  $S$  é a grade da imagem  $f$ , onde  $1 \leq k \leq n$ .

Seja  $n_0$  o número de bandas associadas às imagens sem estímulo e  $n_1$  o número de bandas relacionadas às imagens sob estímulo, se  $n = n_0 + n_1$ , considerando as seguintes populações por *voxel*:

$$F_0(\mathbf{u}) = \{f_k(\mathbf{u}) | v_k = 0\}, \quad (8.1)$$

$$F_1(\mathbf{u}) = \{f_k(\mathbf{u}) | v_k = 1\}, \quad (8.2)$$

têm-se as seguintes estatísticas de interesse:

$$\bar{F}_0(\mathbf{u}) = \frac{1}{n_0} \sum_{v_k=0} f_k(\mathbf{u}), \quad (8.3)$$

$$S_0^2(\mathbf{u}) = \frac{1}{n_0 - 1} \sum_{v_k=0} (f_k(\mathbf{u}) - \bar{F}_0(\mathbf{u}))^2, \quad (8.4)$$

$$\bar{F}_1(\mathbf{u}) = \frac{1}{n_1} \sum_{v_k=1} f_k(\mathbf{u}), \quad (8.5)$$

$$S_1^2(\mathbf{u}) = \frac{1}{n_1 - 1} \sum_{v_k=1} (f_k(\mathbf{u}) - \bar{F}_1(\mathbf{u}))^2, \quad (8.6)$$

onde  $\bar{F}_0(\mathbf{u})$ ,  $S_0^2(\mathbf{u})$ ,  $\bar{F}_1(\mathbf{u})$  e  $S_1^2(\mathbf{u})$  são os respectivos pares média e desvio padrão que estimam os pares esperança e variância  $\mu_0(\mathbf{u})$ ,  $\sigma_0^2(\mathbf{u})$ ,  $\mu_1(\mathbf{u})$  e  $\sigma_1^2(\mathbf{u})$  para as populações adquiridas sem estímulo e sob estímulo, respectivamente, para cada *pixel* em  $f$  com posição  $\mathbf{u} \in S$ .

A maioria dos métodos estatísticos utilizados em soluções comerciais para detecção de níveis de ativação em imagens de ressonância funcional são testes de hipótese nula [178], onde se deseja medir o desvio por *voxel* entre as populações sob estímulo e sem estímulo [179]. Se esse desvio estiver fora de um determinado intervalo de confiança de probabilidade  $P_c$ , o *pixel* correspondente à posição  $\mathbf{u}$  no *mapa binário de ativação*  $f_b : S \rightarrow \{0, 1\}$  assume  $f_b(\mathbf{u}) = 1$  e é considerado *ativado*; caso contrário,  $f_b(\mathbf{u}) = 0$  e o *pixel* correspondente é considerado *não ativado* [5].

Os métodos de detecção baseados em testes de hipótese nula variam de acordo com a natureza do teste, como segue:

1. *Teste t de Student* (TTS): As variâncias  $\sigma_0^2(\mathbf{u})$  e  $\sigma_1^2(\mathbf{u})$  são diferentes e desconhecidas.

$$T(\mathbf{u}) = \frac{\bar{F}_0(\mathbf{u}) - \bar{F}_1(\mathbf{u})}{\sqrt{\frac{S_0^2(\mathbf{u})}{n_0} + \frac{S_1^2(\mathbf{u})}{n_1}}}, \quad (8.7)$$

com

$$\lambda = \frac{\left(\frac{S_0^2(\mathbf{u})}{n_0} + \frac{S_1^2(\mathbf{u})}{n_1}\right)^2}{\frac{(S_0^2(\mathbf{u})/n_0)^2}{n_0+1} + \frac{(S_1^2(\mathbf{u})/n_1)^2}{n_1+1}} - 2 \quad (8.8)$$

graus de liberdade, onde  $\mathbf{u} \in S$ ;

2. *Teste t Combinado* (TTC): As variâncias  $\sigma_0^2(\mathbf{u})$  e  $\sigma_1^2(\mathbf{u})$  são iguais e desconhecidas.

$$T(\mathbf{u}) = \frac{\bar{F}_0(\mathbf{u}) - \bar{F}_1(\mathbf{u})}{S_p(\mathbf{u})\sqrt{\frac{1}{n_0} + \frac{1}{n_1}}}, \quad (8.9)$$

onde

$$S_p^2(\mathbf{u}) = \frac{(n_0 - 1)S_0^2(\mathbf{u}) + (n_1 - 1)S_1^2(\mathbf{u})}{n_0 + n_1 - 2} \quad (8.10)$$

é o *estimador combinado* de  $\sigma^2(\mathbf{u}) = \sigma_0^2(\mathbf{u}) = \sigma_1^2(\mathbf{u})$  com  $\lambda = n_0 + n_1 - 2$  graus de liberdade, onde  $\mathbf{u} \in S$ ;

3. *Teste t Emparelhado* (TTP): Considera as estatísticas das diferentes populações

$$D(\mathbf{u}) = \{f_i(\mathbf{u}) - f_j(\mathbf{u}) | v_i = 0, v_j = 1\}, \quad (8.11)$$

onde  $n_0 = n_1 = n/2$ ; se  $\bar{D}(\mathbf{u})$  e  $S_D^2(\mathbf{u})$  são a média e o desvio padrão da população  $D(\mathbf{u})$ , então o testes estatístico é dado por:

$$T(\mathbf{u}) = \frac{\bar{D}(\mathbf{u})}{S_D(\mathbf{u})/\sqrt{n/2}}. \quad (8.12)$$

Utilizando as estatísticas  $T$  obtidas a partir dos cálculos anteriores e distribuídas segundo uma distribuição  $t$  com  $\lambda$  graus de liberdade, pode-se calcular as probabilidades  $P_{\mathbf{u}} = P(t \leq T(\mathbf{u}))$ ,  $\forall \mathbf{u} \in S$ .

O mapa de ativação binário é composto como segue:

$$f_b(\mathbf{u}) = \begin{cases} 1, & P_{\mathbf{u}} < \frac{1-P_c}{2} \vee P_{\mathbf{u}} > \frac{1+P_c}{2} \\ 0, & \frac{1-P_c}{2} \leq P_{\mathbf{u}} \leq \frac{1+P_c}{2} \end{cases}, \quad (8.13)$$

onde  $\mathbf{u} \in S$ . A presença de ruído é caracterizada pela existência de *pixels* isolados marcados como ativados em regiões onde não há amostra cranial. Nestes casos, pode-se aplicar o filtro da moda sobre o mapa  $f_b$  para eliminar as falsas detecções, gerando o mapa de ativação filtrado  $f'_b : S \rightarrow \{0, 1\}$ .

Para avaliar qualitativamente as áreas detectadas é comum compor uma imagem temática a partir da superposição do mapa binário de ativação sobre a imagem anatômica correspondente, rotulando as áreas ativadas com uma cor pré-definida, que recebe o nome de *marcador* [3, 5].

A imagem temática de ativação  $f_a : S \rightarrow [0, 1]^3$ , com mapa de ativação binário filtrado  $f'_b : S \rightarrow \{0, 1\}$  e marcador  $\mathbf{r} = (r_1, r_2, r_3)^T$ , pode ser obtida pela expressão:

$$f_a(\mathbf{u}) = f'_b(\mathbf{u})\bar{F}_1(\mathbf{u})\mathbf{r} + (1 - f'_b(\mathbf{u}))\bar{F}_1(\mathbf{u})(1, 1, 1)^T, \quad (8.14)$$

onde  $\mathbf{u} \in S$ .

### 8.2.2 Métodos Multiespectrais Não Paramétricos

Os métodos não paramétricos são baseados em geral na classificação *voxel a voxel* das séries temporais normalizadas utilizando métodos não supervisionados, tais como os mapas de *k*-médias, os mapas auto-organizados de Kohonen e os mapas *fuzzy c*-médias [75, 78, 79, 16].

Contudo, pode-se considerar a tarefa de detectar regiões como na verdade a classificação da imagem multiespectral normalizada de  $n$  bandas  $f'$  considerando o conjunto de classes  $\Omega = \{C_1, C_2, \dots, C_m\}$ , onde as primeiras  $m - 1$  classes representam os diversos níveis de ativação de interesse detectados e a  $m$ -ésima classe representa o fundo da imagem.

Consequentemente, a imagem multiespectral normalizada  $f' : S \rightarrow [-1, 1]^n$  é definida a partir de  $f$  como segue [78]:

$$f'_k(\mathbf{u}) = \begin{cases} f_k(\mathbf{u}) - \bar{F}_0, & v_k = 0 \\ f_k(\mathbf{u}) - \bar{F}_1, & v_k = 1 \end{cases}, \quad (8.15)$$

onde  $1 \leq k \leq n$  e  $\mathbf{u} \in S$ .

A classificação automática de  $f' : S \rightarrow [-1, 1]^n$  resulta no mapa de classificação  $g : S \rightarrow \{1, 2, \dots, m\}$ . O mapa binário de ativação é gerado a partir da seleção da  $k$ -ésima classe associada a uma ativação, como segue:

$$f_b(\mathbf{u}) = \begin{cases} 1, & g(\mathbf{u}) = k \\ 0, & g(\mathbf{u}) \neq k \end{cases}, \quad (8.16)$$

onde  $\mathbf{u} \in S$ .

Para reduzir a influência das falsas ativações originadas pelo ruído, o mapa binário de ativação  $f_b$  é filtrado pelo filtro da moda, resultando  $f'_b$ . A imagem temática de ativação  $f_a$  é gerada pela superposição de  $f'_b$  sobre o mapa anatômico sob estímulo  $\bar{F}_1$  usando o marcador  $\mathbf{r}$ , como descrito na equação 8.14.

### 8.2.3 Método Não Paramétrico Monoespectral Proposto

A maior desvantagem do emprego de métodos não paramétricos é seu alto custo computacional associado, apesar de seus bons resultados na detecção e classificação de diferentes níveis de ativação. Esses métodos lidam com imagens multiespectrais sintéticas com no mínimo 60 bandas durante a fase de treinamento e são responsáveis pela detecção de diferentes níveis de ativação.

Como uma boa solução pode-se gerar uma imagem monoespectral com as diferenças de níveis. Uma imagem candidata pode ser  $f_d : S \rightarrow [0, 1]$ , definida como as diferenças entre os níveis médios dos sinais sob estímulo e sem estímulo por *voxel*:

$$f_d(\mathbf{u}) = \bar{F}_1(\mathbf{u}) - \bar{F}_0(\mathbf{u}), \quad (8.17)$$

onde  $\mathbf{u} \in S$ . Contudo, as regiões ativadas com baixo nível médio de sinal podem não ser detectadas durante a etapa de classificação. Portanto, pode-se modificar o histograma da imagem diferença, gerando a imagem diferença modificada  $f'_d : S \rightarrow [0, 1]$  definida por:

$$f'_d(\mathbf{u}) = \left( \frac{\bar{F}_1(\mathbf{u}) - \bar{F}_0(\mathbf{u}) - \lambda_1}{\lambda_2 - \lambda_1} \right)^{1/2}, \quad (8.18)$$

onde  $\mathbf{u} \in S$ ,

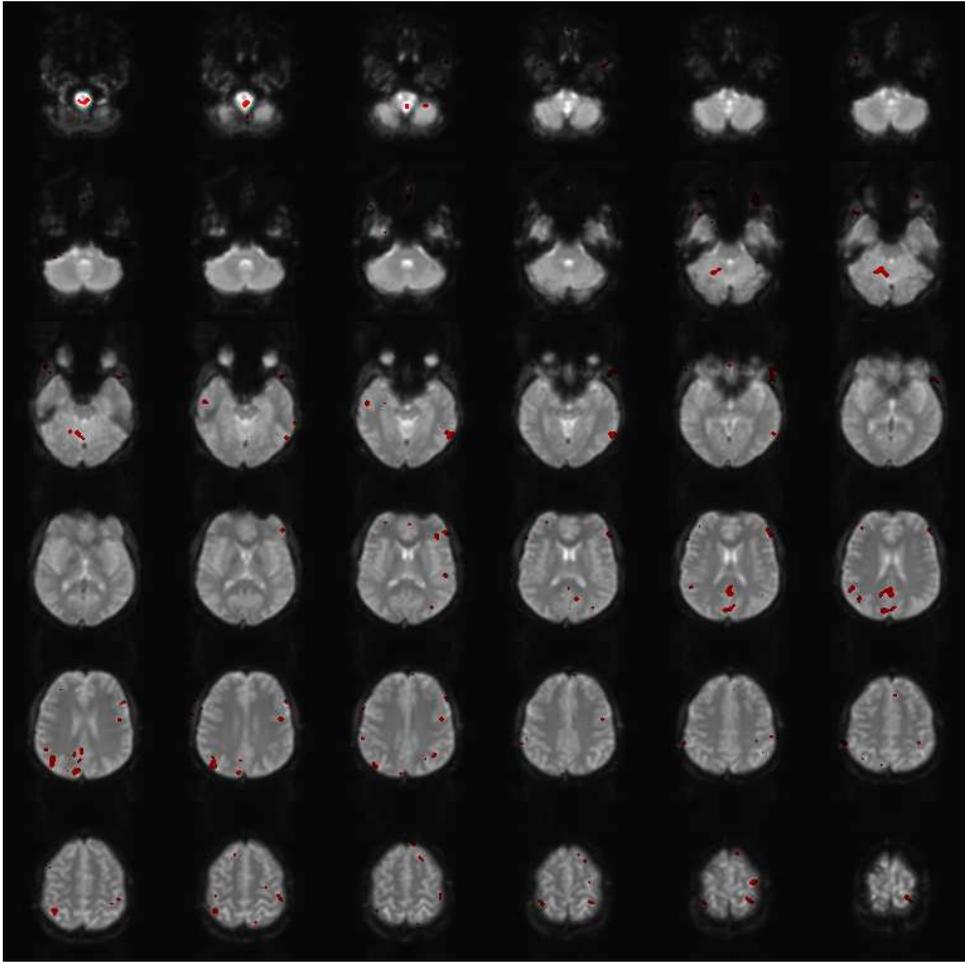
$$\lambda_1 = \bigwedge_{\mathbf{v} \in S} (\bar{F}_1(\mathbf{v}) - \bar{F}_0(\mathbf{v})), \quad (8.19)$$

e

$$\lambda_2 = \bigvee_{\mathbf{v} \in S} (\bar{F}_1(\mathbf{v}) - \bar{F}_0(\mathbf{v})). \quad (8.20)$$

A classificação automática de  $f'_d : S \rightarrow [0, 1]$  resulta no mapa de classificação  $g : S \rightarrow \{1, 2, \dots, m\}$ . O mapa de ativação binário  $f_b : S \rightarrow \{0, 1\}$  é gerado a partir da seleção da  $k$ -ésima classe de interesse associada à ativação, como descrito na equação 8.16. Os passos seguintes são idênticos aos passos seguidos pelos métodos não paramétricos citados anteriormente. Na etapa de classificação podem-se utilizar diversos métodos não paramétricos, tais como o mapa de  $k$ -médias, os mapas auto-organizados de Kohonen e os mapas *fuzzy*  $c$ -médias [18, 17].

A avaliação do resultado é feita a partir de  $g$ , utilizando um conjunto de regiões de interesse gerado a partir da seleção da região ativada correspondente, utilizando como referência um atlas funcional que mostra a região do cérebro responsável pela atenção espacial visual [6]. Para avaliar quantitativamente os resultados utiliza-se o índice  $\kappa$ , a acurácia global e a matriz de confusão. Da matriz de confusão calculada foram obtidas as taxas de falsos positivos, falsos negativos, verdadeiros positivos e verdadeiros negativos para cada um dos métodos de detecção apresentados.



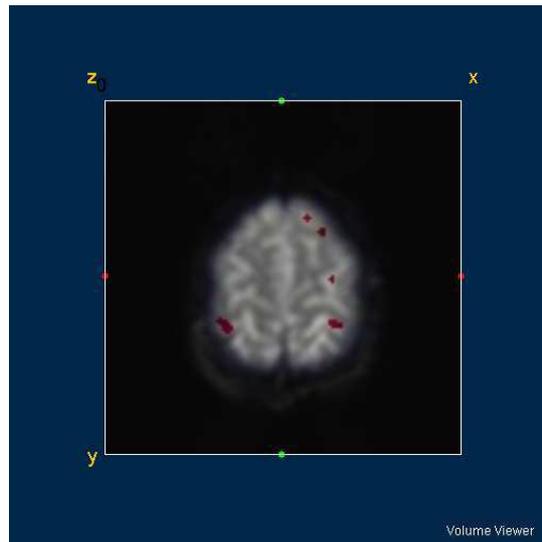
**Figura 8.1** Mapa temático de ativação do teste  $t$  de Student

### 8.3 Resultados

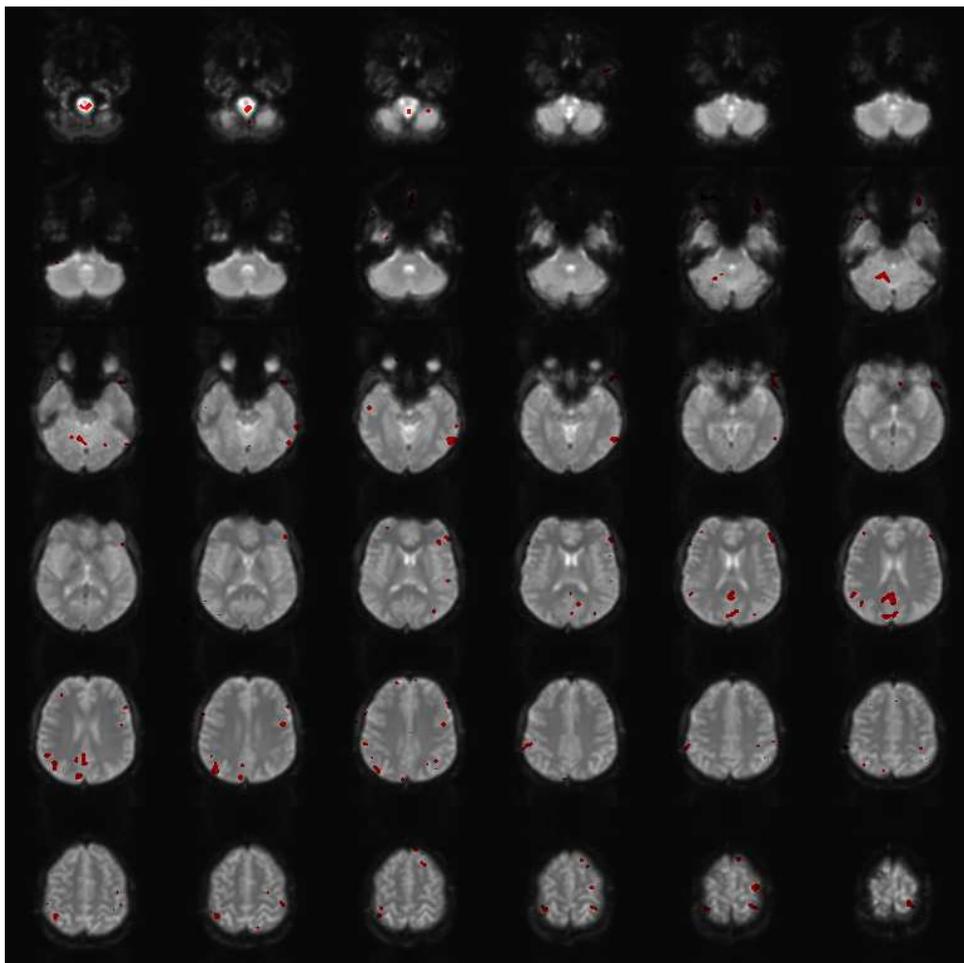
As figuras 8.1, 8.3 e 8.5 mostram os mapas temáticos de ativação  $f_a : S \rightarrow [0, 1]^3$  resultantes da detecção usando o teste  $t$  de Student, o teste  $t$  emparelhado e o teste  $t$  combinado, com probabilidade de confiança  $P_c = 95\%$  e o marcador  $\mathbf{r}$  (vermelho), considerando a notação RGB,  $\mathbf{r} = (1, 0, 0)^T$ , enquanto as figuras 8.2, 8.4 e 8.6 mostram suas respectivas composições volumétricas.

As figuras 8.7 e 8.8 mostram a imagem diferença modificada  $f'_d : S \rightarrow [0, 1]$  na forma de representação plana e de composição volumétrica, respectivamente. Nelas é possível notar tanto as regiões ativadas quanto as falsas ativações devidas à presença de ruído durante todos os experimentos.

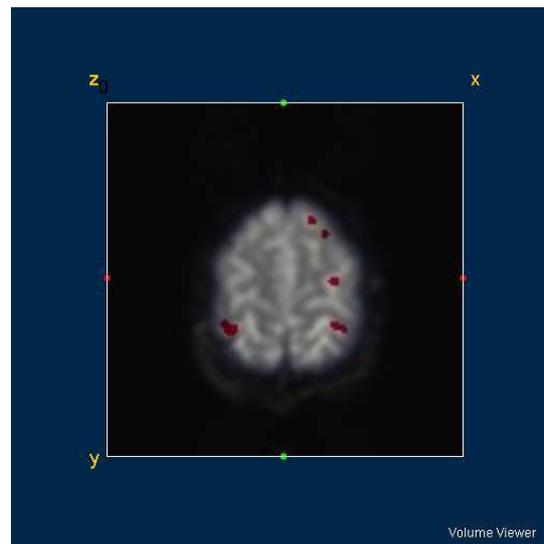
As figuras 8.9 e 8.10 mostram o mapa temático de ativação  $f_a : S \rightarrow [0, 1]^3$  na forma de representação plana e de composição volumétrica, respectivamente, obtida do mapa anatômico de ativação  $\bar{F}_1 : S \rightarrow [0, 1]$  e do mapa binário de ativação filtrado  $f'_b : S \rightarrow [0, 1]$  resultante da classificação da imagem normalizada segundo o mapa auto-organizado de Kohonen (KON),



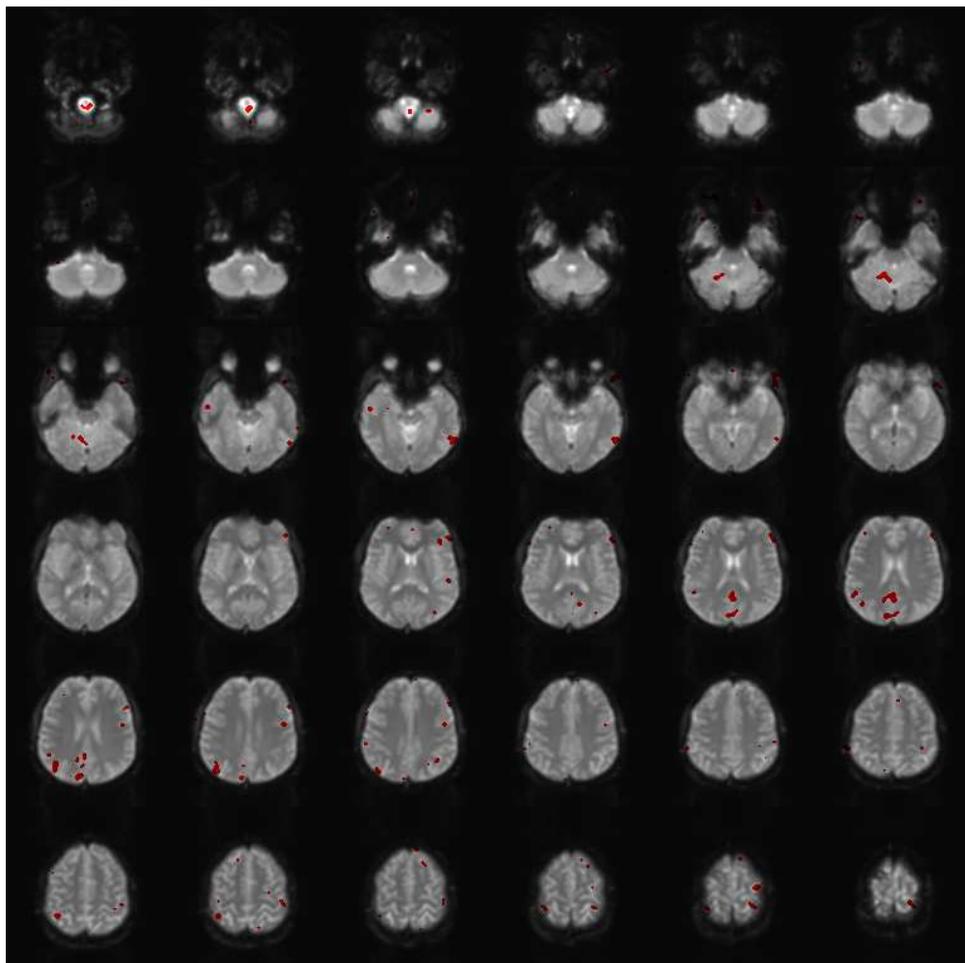
**Figura 8.2** Composição volumétrica do mapa temático de ativação do teste  $t$  de Student



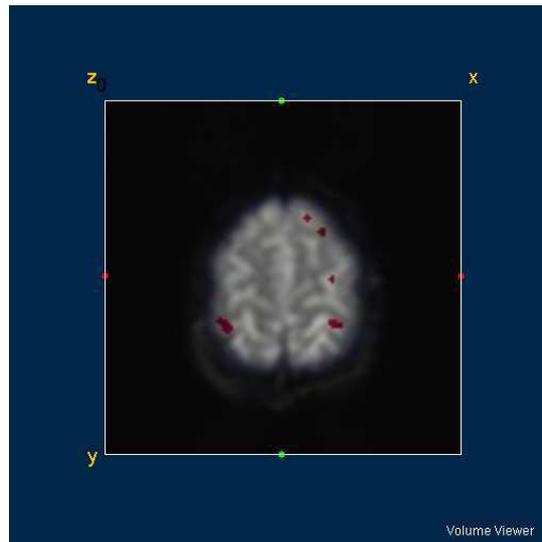
**Figura 8.3** Mapa temático de ativação do teste  $t$  emparelhado



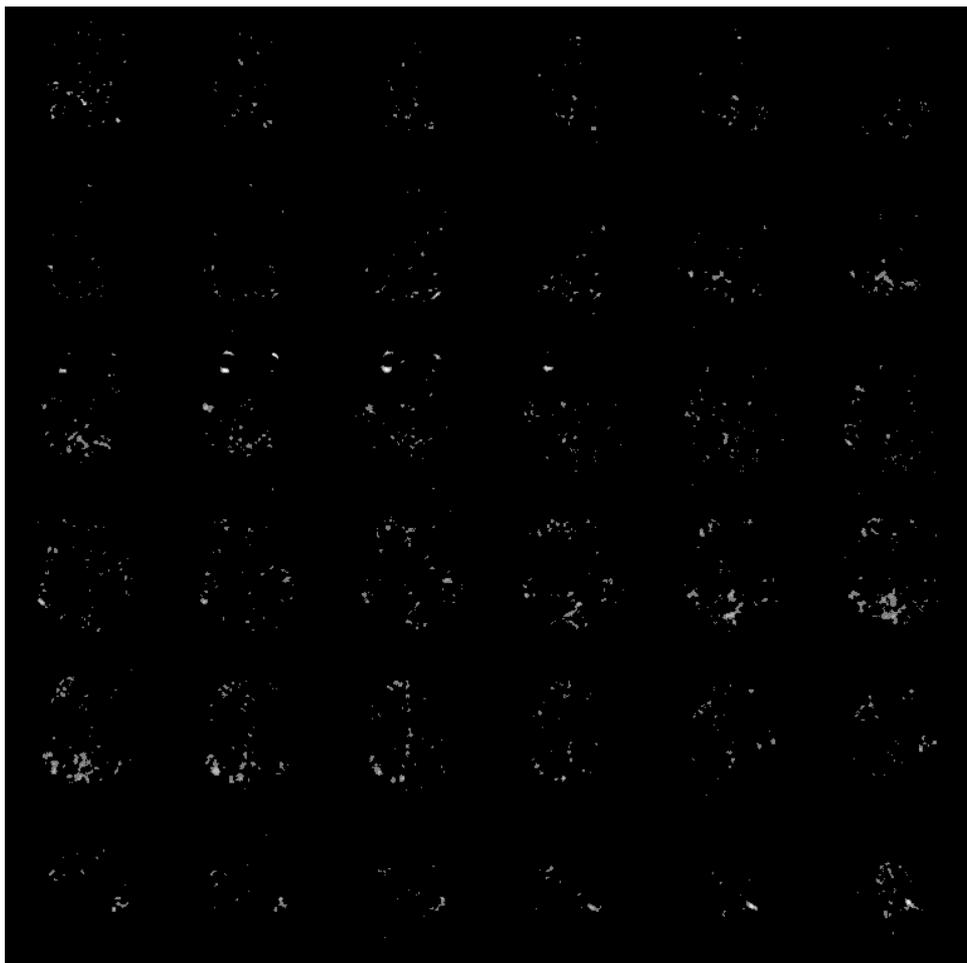
**Figura 8.4** Composição volumétrica do mapa temático de ativação do teste  $t$  emparelhado



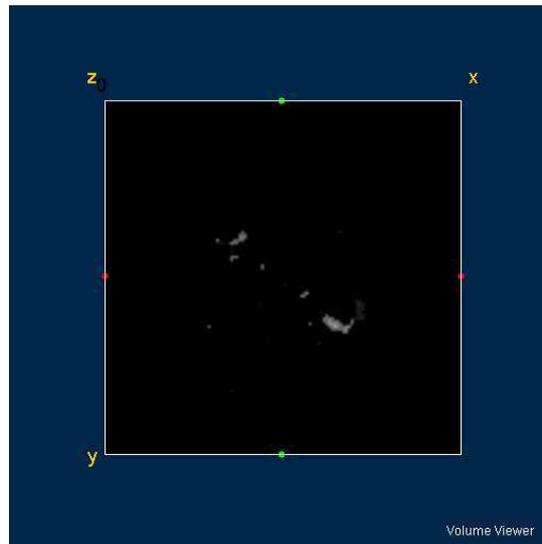
**Figura 8.5** Mapa temático de ativação do teste  $t$  combinado



**Figura 8.6** Composição volumétrica do mapa temático de ativação do teste  $t$  combinado



**Figura 8.7** Imagem diferença modificada



**Figura 8.8** Composição volumétrica da imagem diferença modificada

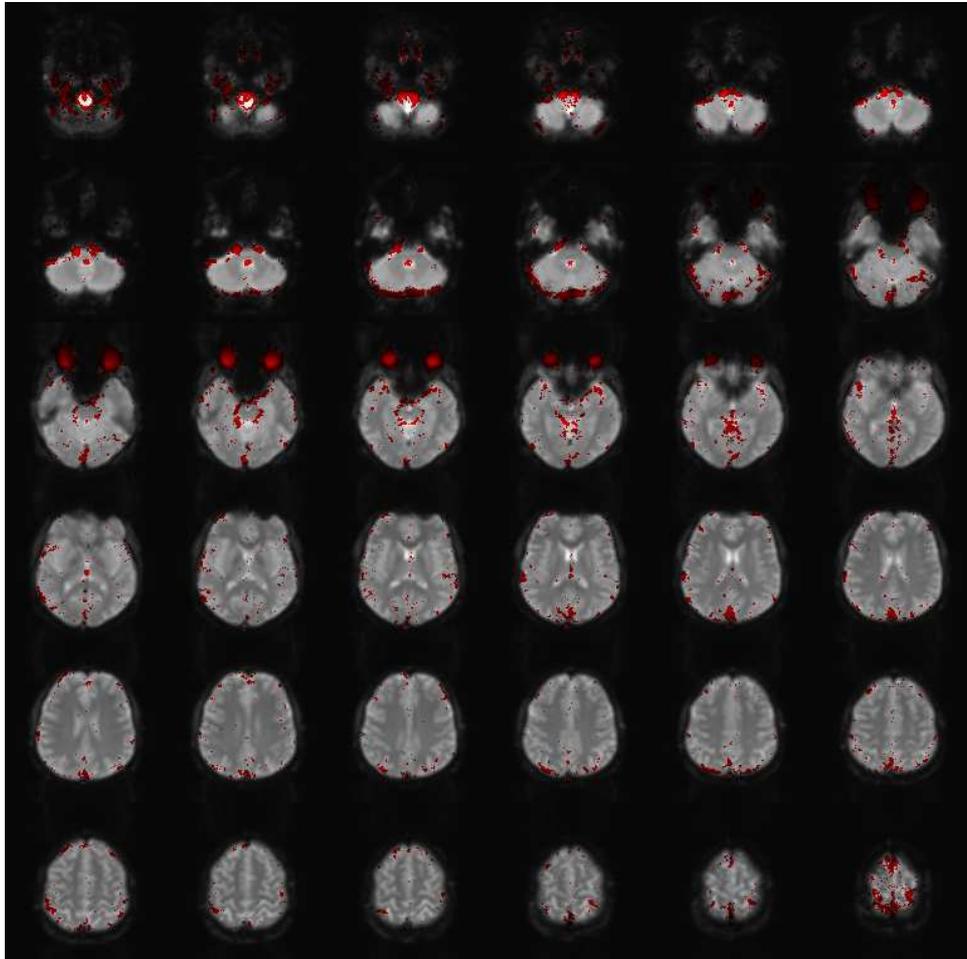
	TTS	TTC	TTP	KON	KO	ODC
$\kappa$	0,59	0,59	0,66	0,26	0,71	0,71
$\phi$ (%)	78,53	78,53	82,61	57,88	85,33	85,33

**Tabela 8.1** Avaliação de desempenho usando a taxa de acerto global  $\phi$  e o índice  $\kappa$

com 48 entradas, 3 saídas, 3 neurônios linearmente distribuídos, 200 iterações durante a etapa de treinamento e taxa de aprendizado inicial de  $\eta_0 = 0,1$ .

As figuras 8.11 e 8.12 mostram o mapa temático de ativação  $f_a : S \rightarrow [0, 1]^3$  na representação plana e de composição volumétrica, respectivamente, obtido do mapa anatômico de ativação  $\bar{F}_1 : S \rightarrow [0, 1]$  e do mapa binário de ativação filtrado  $f'_b : S \rightarrow [0, 1]$  resultante da classificação da imagem diferença modificada usando o mapa auto-organizado de Kohonen (KO), com uma entrada, 3 saídas, 3 neurônios linearmente distribuídos, 50 iterações durante a etapa de treinamento e taxa de aprendizado inicial  $\eta_0 = 0,1$ . As figuras 8.13 e 8.14 mostram a representação plana e de composição volumétrica, respectivamente, do mapa temático de ativação adquirido usando o classificador dialético objetivo (ODC) com uma entrada, 4 pólos iniciais, 4 fases históricas com duração de 100 iterações cada, passo histórico inicial  $\eta_0 = 0,1$ , limiares de força e de contradição de 0,1 e 0,75, respectivamente, e sem geração de novos pólos.

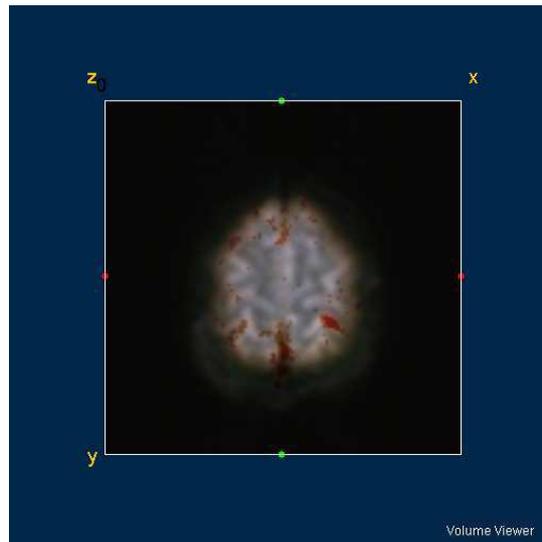
O desempenho de classificação foi avaliado usando regiões de interesse adquiridas de acordo com um atlas de fMRI baseado em experimentos de tomógrafos de ressonância de 4 T [6]. A tabela 8.1 mostra os resultados da avaliação dos métodos de detecção apresentados usando o índice  $\kappa$  e a taxa de acerto global  $\phi$  (%), enquanto a tabela 8.2 mostra as taxas de falsos positivos (FP), falsos negativos (FN), verdadeiros positivos (TP) e verdadeiros negativos (TN).



**Figura 8.9** Mapa temático de ativação resultante da classificação da imagem normalizada pelo mapa auto-organizado de Kohonen

	TTS	TTC	TTP	KON	KO	ODC
TN (%)	100,00	100,00	100,00	100,00	100,00	100,00
FN (%)	0,00	0,00	0,00	0,00	0,00	0,00
FP (%)	34,80	34,80	28,19	68,28	23,79	23,79
TP (%)	65,20	65,20	71,81	31,72	76,21	76,21

**Tabela 8.2** Avaliação de desempenho usando as taxas de detecção



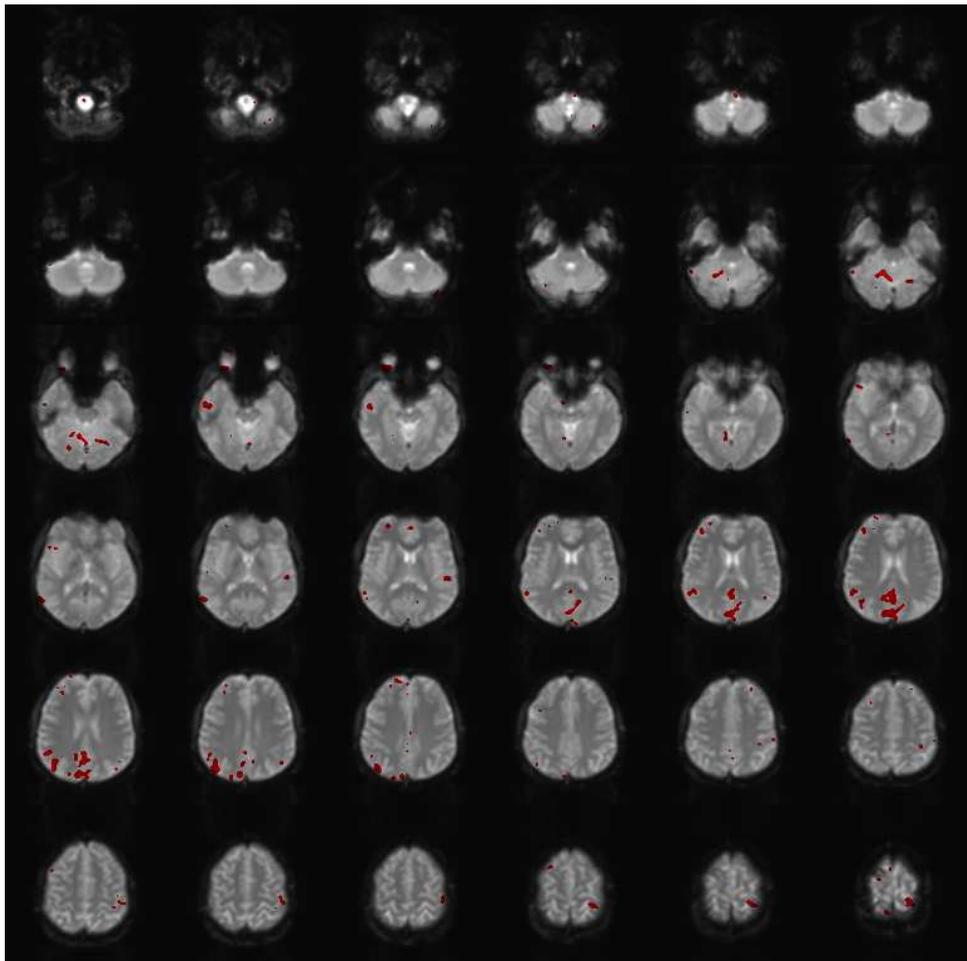
**Figura 8.10** Composição volumétrica do mapa temático de ativação resultante da classificação da imagem normalizada pelo mapa auto-organizado de Kohonen

## 8.4 Discussão e Conclusões

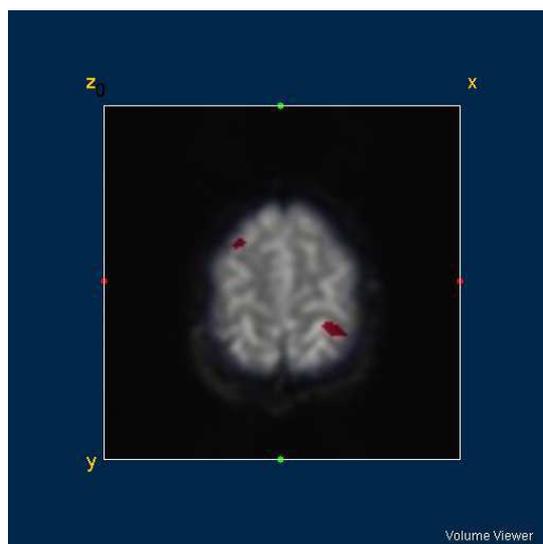
Analisando os índices de pós-classificação  $\kappa$  da tabela 8.1, pode-se perceber que tanto o mapa auto-organizado de Kohonen quanto o classificador dialético objetivo associados à classificação da imagem diferença modificada resultaram em maiores correlações estatísticas ( $\kappa = 0,71$ ) com o conjunto de teste composto pelas regiões de interesse previamente selecionadas que os métodos estatísticos paramétricos ( $\kappa = 0,59$  para o teste  $t$  de Student e para o teste  $t$  combinado, e  $\kappa = 0,66$  para o teste  $t$  emparelhado) e para a classificação da imagem normalizada ( $\kappa = 0,26$ ). Esses resultados são confirmados pela taxa de acerto global:  $\phi = 85,33\%$  para os métodos KO e ODC contra  $\phi = 78,53\%$  para os métodos TTS e TTC, e  $\phi = 82,61\%$  para o método TTP.

A tabela 8.2 mostra as taxas TN, FN, FP e TP para os métodos de detecção apresentados. Essas taxas foram calculadas a partir das matrizes de confusão obtidas no processo de pós-classificação usando um conjunto de teste composto pelas referidas regiões de interesse e considerando apenas duas classes: sob estímulo e sem estímulo. Consequentemente, TN e TP são expressos na diagonal principal das matrizes de confusão.

Analisando a tabela 8.2 é possível notar que todos os métodos de detecção resultaram na ausência de falsos negativos nos *voxels* ativados. Entretanto, as taxas de falsos positivos foram consideráveis, devido à possível falta de precisão na seleção das regiões de interesse pelo especialista, uma vez que essa seleção é realizada através da análise visual de um atlas de ativação fMRI [6]. Esse erro de seleção pode inclusive aumentar devido à baixa resolução anatômica das imagens EPI [1]. A resolução anatômica também diminui com o aumento do campo magnético estático, como pode ser demonstrado pelos experimentos de fMRI com tomógrafos de ressonância magnética de 4 T [6]. Além do mais, para demonstrar os métodos propostos foram utilizadas neste trabalho imagens funcionais reais, o que significa dizer que existem mais



**Figura 8.11** Mapa temático de ativação resultante da classificação da imagem diferença modificada pelo mapa auto-organizado de Kohonen

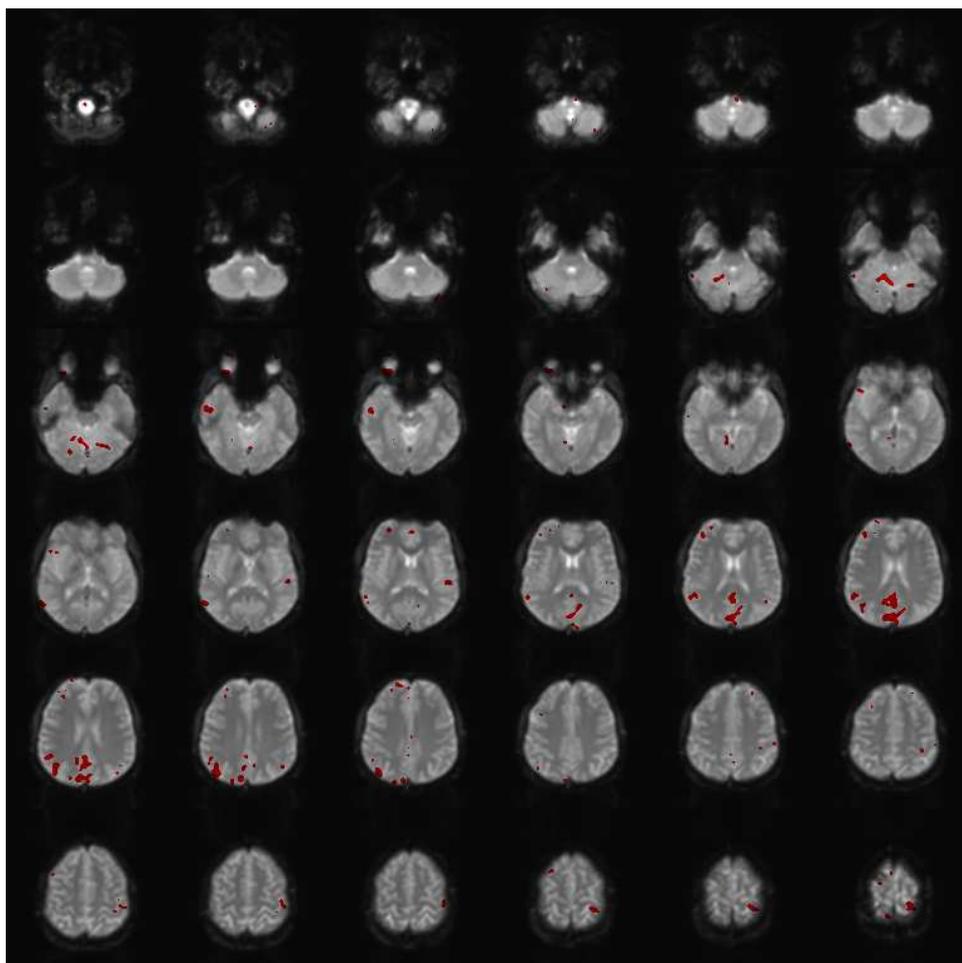


**Figura 8.12** Composição volumétrica do mapa temático de ativação resultante da classificação da imagem diferença modificada pelo mapa auto-organizado de Kohonen

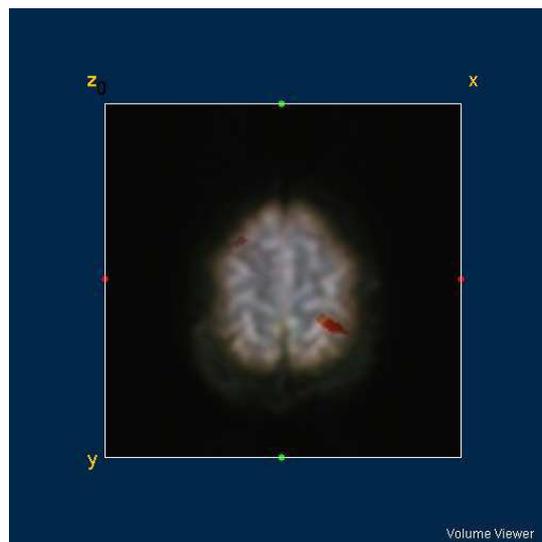
áreas ativadas do que as esperadas inicialmente, uma vez que o cérebro do voluntário está ocupado com a realização de outras tarefas que não o movimento dos dedos da mão direita, como exigia o experimento de controle, resultando que há ativações mais fortes nas áreas do córtex dedicadas à atenção visual direita do que na área dedicada ao movimento dos dedos da mão direita.

A taxa de falsos positivos foi de 34,80% para os métodos TTS e TTC, 28,19% para o método TTP, 23,79% para os métodos não paramétricos KO e ODC, e 68,28% para o método KON. Esses resultados mostram que a classificação mono-espectral baseada no mapa auto-organizado de Kohonen e a abordagem baseada no classificador dialético objetivo foram superiores aos testes estatísticos de hipótese nula  $t$  e à versão multiespectral da classificação com mapa auto-organizado de Kohonen. Os métodos KO e ODC também são computacionalmente superiores se comparados à maioria dos métodos clássicos de detecção apresentados devido à sua natureza mono-espectral, em contraste com esses métodos, uma vez que foi demonstrado neste trabalho que a análise de sequências temporais de imagens EPI é na verdade a análise *pixel a pixel* da imagem multiespectral sintética composta pelas imagens EPI sob estímulo e sem estímulo (KON). Portanto, tanto o método de detecção baseado no mapa de Kohonen quanto o método dialético objetivo propostos neste trabalho têm a vantagem adicional de reduzir um problema de análise multiespectral a uma abordagem mono-espectral, eliminando o custo computacional associado às avaliações das funções de distribuição de probabilidade dos testes de hipótese.

Nesta aplicação o classificador dialético objetivo teve exatamente o mesmo desempenho do mapa auto-organizado de Kohonen. Os bons resultados obtidos com o método ODC se devem ao fato de o classificador dialético objetivo ter um bom desempenho na detecção de singularidades, como é o caso da detecção de regiões ativadas em imagens funcionais. Contudo, como a maioria dos métodos evolucionários, o método de detecção proposto baseado no classificador dialético objetivo tem a desvantagem de ter o desempenho de classificação altamente depen-



**Figura 8.13** Mapa temático de ativação resultante da classificação da imagem diferença modificada pelo classificador dialético objetivo



**Figura 8.14** Composição volumétrica do mapa temático de ativação resultante da classificação da imagem diferença modificada pelo classificador dialético objetivo

dente dos parâmetros iniciais, em particular do número inicial de pólos, do número de fases históricas, da duração de cada fase histórica, dos limiares de força e contradição e dos níveis de máxima crise, além do passo de treinamento. A determinação desses parâmetros é empírica e dependente do conhecimento e da experiência do especialista no treinamento. Contudo, este é um ponto onde se poderia considerar o uso de um método de otimização global para auxiliar na determinação desses parâmetros.

## Conclusão e Trabalhos Futuros

### 9.1 Comentários e Considerações

Um dos principais objetivos deste trabalho foi mostrar que é possível construir métodos de Inteligência Computacional buscando inspiração na Filosofia, em especial nos métodos filosóficos de investigação e análise da realidade, como é o caso da família de métodos filosóficos que recebe o nome de método dialético.

A dialética foi escolhida como inspiração filosófica deste trabalho devido ao seu caráter ao mesmo tempo sistemático e não linear. No entanto, a partir dos trabalhos de diversos filósofos e pensadores dialéticos dos séculos XIX e XX, fica patente a dificuldade de expressar, ou mesmo de traduzir, os conceitos contidos nas categorias da dialética, a saber, contradição, totalidade, movimento perpétuo e práxis, em expressões matemáticas ou mesmo na forma algorítmica. Esse fato é evidente ao se perceber que mesmo Marx, considerado o primeiro economista preocupado com rigor matemático em suas formulações, não está preocupado em expressar o uso das categorias dialéticas em expressões matemáticas, o que pode ser explicado pelo fato de que, como linguagem para expressão de análise da realidade ou de seus aspectos particulares, a filosofia alemã vivia seu auge, com forte predominância do pensamento dialético de cunho hegeliano.

Rosser Jr. possui um trabalho dos mais interessantes no que tange à tentativa de modelagem matemática das categorias dialéticas, em especial da dialética materialista, onde se busca aproximar essas categorias por meio de abordagens baseadas em sistemas dinâmicos não lineares, tais como aqueles presentes na Teoria da Catástrofe e na Teoria do Caos, visando expressar matematicamente os conceitos de contradição, totalidade, movimento perpétuo e práxis [35]. Este trabalho buscou mostrar as limitações da abordagem baseada em sistemas dinâmicos não lineares, em especial a impossibilidade de modelar a contento a categoria da práxis [35].

Dada a natureza imprecisa e nebulosa dos conceitos da dialética materialista, foi bastante natural buscar na Lógica *Fuzzy* o formalismo matemático para propor o Método Dialético Objetivo (*Objective Dialectical Method*, ODM), um método evolutivo de Inteligência Computacional baseado no método dialético materialista, mas obviamente construído a partir de simplificações e aproximações das categorias dialéticas, para facilitar e mesmo possibilitar sua implementação algorítmica.

O ODM foi proposto na forma de um método evolutivo dividido em fases históricas onde ocorre a dinâmica entre os pólos constituintes do sistema. Essa dinâmica se dá em duas etapas: evolução e revolução. Na etapa de evolução, a dinâmica entre os pólos se dá na forma de competição ou, para usar um jargão caro ao pensamento dialético materialista, luta de pólos, onde as interações entre os pólos afetam a natureza dos próprios pólos, modelada pelos

vetores de pesos que os representam e, portanto, instanciando assim as categorias de contradição, totalidade e movimento perpétuo, como desejado. Já na etapa de revolução acontecem de fato as mudanças qualitativas do sistema, onde pólos podem ser eliminados ou absorvidos por outros, enquanto novos pólos são gerados como resultado da síntese dos pólos envolvidos na contradição principal do sistema.

Assim, na proposta deste trabalho, na etapa de revolução está realmente implementada a categoria da práxis. Até o momento não havia modelo matemático, que se propusesse dialético, capaz de implementar a práxis com mudança do número de pólos integrantes do sistema, o que torna essa uma importante contribuição deste trabalho, dado que nem mesmo os pólos são explicitamente modelados por meio de sistemas dinâmicos. É importante destacar que, quando o ODM, aplicado à tarefa de classificação, quando passa a receber a denominação de Classificador Dialético Objetivo (*Objective Dialectical Classifier*, ODC), tem sua dinâmica reduzida a apenas uma única fase histórica, conforme o algoritmo proposto, a etapa de revolução não acontece para essa única fase histórica, e o método se reduz ao método de agrupamento pelo mapa *fuzzy* c-médias. Isso foi proposital, para evitar problemas com a convergência do modelo dialético proposto, no caso do classificador dialético, e utilizar como base um algoritmo de Inteligência Computacional bem conhecido e que fizesse uso do formalismo *fuzzy*, como planejado.

Contudo, ainda há limitações no modelo que são difíceis de serem contornadas, e que são ao mesmo tempo características do modelo dialético proposto. A principal limitação permeia praticamente todos os métodos evolutivos recentes: a quantidade razoável de parâmetros iniciais a serem definidos, tais como o número de fases históricas, o número inicial de pólos, a contradição mínima, a força mínima, a duração de cada fase histórica, e os passos histórico e contemporâneo. A segunda limitação do ODC está no esforço computacional do algoritmo, que ainda não é baixo, embora seja praticamente o mesmo esforço computacional do treinamento de tantos métodos *fuzzy* c-médias quantas sejam as fases históricas estudadas. Entretanto, não fez parte do escopo deste trabalho a medição do esforço computacional do ODC em uma determinada tarefa e sua comparação com outros métodos, tendo-se realizado comparações apenas quanto ao desempenho de classificação, segmentação e quantização, usando índices de fidelidade da imagem quantizada à imagem original, e medidas de validade de agrupamento. Caberia aqui como trabalho futuro pesquisar heurísticas que possam ajudar na determinação dos parâmetros iniciais de acordo com o problema abordado. Essas heurísticas, combinadas com métodos de otimização global, caso possam ser construídas, poderiam ajudar a reduzir o esforço na experimentação de parâmetros iniciais adequados, além de poder reduzir o custo computacional do treinamento.

Este trabalho também apresentou o método dialético objetivo adaptado a problemas de busca de otimização. Foi elaborada uma revisão bibliográfica a respeito das principais abordagens para busca e otimização no campo da Computação Evolucionária, a saber: algoritmos genéticos, programação evolucionária e otimização por enxame de partículas. Essa revisão bibliográfica foi importante para contextualizar a proposta: um método de otimização para resolução de problemas unimodais, multimodais e multiobjetivos, embora estes últimos não tenham sido muito explorados na avaliação.

Assim, foi proposta uma adaptação do método dialético a problemas de otimização e apre-

sentada de forma algorítmica, onde se percebe, da mesma forma que nas outras abordagens em Computação Evolucionária, muitas semelhanças entre conceitos importantes de algoritmos genéticos e programação evolucionária, tais como os conceitos de seleção natural, mutação e cruzamento, que mantém paralelos com os respectivos conceitos de luta de pólos, crise revolucionária, e pólos tese e antítese como geradores de pólos síntese no método dialético.

Este trabalho apresentou ainda uma nova versão do método dialético de busca e otimização baseada no Princípio da Máxima Entropia. Essa versão modificada do método foi usada para estudar a dinâmica da versão dita canônica. Contudo a entropia teve que ser redefinida para lidar com funções de pertinência *fuzzy* ao invés de probabilidades ou de funções densidade de probabilidade, dadas as diferenças conceituais entre probabilidades e pertinências, sendo estas últimas consideradas uma generalização das primeiras, dado que expressam a ideia de incerteza, que pode ser probabilística (probabilidades) ou possibilística.

Como resultado, obteve-se um método dialético com características de busca global reduzidas quando comparado com a versão canônica do algoritmo dialético de busca e otimização, mas otimizado quanto à convergência, o que resulta no uso de mais iterações para resolver alguns problemas que eram resolvidos com poucas iterações na versão canônica, embora os resultados sejam já muito próximos dos valores mínimos absolutos; já quanto à entropia, ou seja, a capacidade de exploração de novos candidatos à solução, houve um sensível incremento que significou um incremento na busca local e uma maior robustez a mínimos locais fortemente atrativos.

Ambas as versões do método dialético de busca e otimização foram utilizadas na geração de novos métodos de segmentação de imagens baseados no mapa de k-médias, que teve seu algoritmo de treinamento substituído por um processo de otimização de índices de validade de agrupamento usando o método dialético de otimização. Os resultados da segmentação de imagens multiespectrais sintéticas de três bandas compostas por imagens de ressonância magnética foram interessantes quanto à qualidade da segmentação, mostrando que é possível obter resultados ótimos segundo certas características desejáveis expressas pelos respectivos índices de avaliação da classificação e da quantização. Contudo, embora o desempenho tenha sido muito bom e os algoritmos obtidos sejam considerados rápidos, não houve neste trabalho preocupação em medir objetivamente o custo computacional das estratégias propostas, dado que o desempenho aparente, quando comparado com o desempenho de outros métodos de segmentação bem conhecidos, como o *fuzzy* c-médias e o k-médias canônico, não foi muito maior do que os deles.

Este trabalho buscou ainda apresentar, por meio de imagens reais de ressonância magnética, a análise de imagens de ressonância magnética sob uma perspectiva de análise multiespectral através de dois estudos de caso: a avaliação do progresso da doença de Alzheimer utilizando imagens de ressonância ponderadas em difusão e a detecção de regiões cerebrais de atividade motora em imagens funcionais de ressonância magnética utilizando métodos estatísticos paramétricos e não paramétricos.

No primeiro estudo de caso, a avaliação do progresso da doença de Alzheimer utilizando imagens de ressonância ponderadas em difusão, a análise multiespectral por redes neurais e arquiteturas híbridas é apresentada como uma alternativa ao exame convencional utilizando o mapa de coeficientes de difusão aparentes, dadas suas limitações quanto à classificação de

regiões de baixa mobilidade de *spins*.

No segundo estudo de caso, a detecção de regiões cerebrais de atividade motora em imagens funcionais de ressonância magnética utilizando métodos estatísticos e não estatísticos, a motivação é justamente a oposta: buscou-se mostrar que os métodos estatísticos paramétricos e não paramétricos usualmente empregados na detecção de regiões ativadas em imagens funcionais de ressonância são abordagens multiespectrais, e que é possível propor métodos não paramétricos mais simples e de menor custo computacional aparente que reduzam o problema à classificação mono-espectral de imagens, uma vez que a análise multiespectral neste caso está associada a um custo computacional relativamente alto, pois se analisam na verdade imagens de 48 bandas, com dimensão de  $768 \times 768$  *pixels*.

## 9.2 Contribuições

A principal contribuição deste trabalho foi a implementação e avaliação do classificador dialético objetivo, para problemas de agrupamento, segmentação e classificação não supervisionada, e do método dialético objetivo para busca e otimização, que por sua vez resultou em uma versão dita canônica e outra gerada a partir do uso do Princípio da Máxima Entropia adaptado a sistemas *fuzzy*. Essa contribuição extrapolou a motivação inicial de desenvolver os ditos métodos e aplicá-los ao estudo de imagens multiespectrais de ressonância magnética, mostrando que é possível e viável construir métodos de Computação Evolucionária com heurísticas inspiradas em métodos filosóficos investigativos e com desempenho melhorado em alguns aspectos pelo uso de elementos de Teoria da Informação, indicando à comunidade científica outras fontes de inspiração dentro do conhecimento humano para além das analogias biológicas e sociais e, inclusive, levando ao reencontro da ciência com a Filosofia, de onde se origina e com quem compartilha a intenção de melhor entender e explicar o ser humano e suas relações com tudo o que existe.

Este trabalho resultou nas seguintes publicações científicas:

1. *A Dialectical Approach for Classification of DW-MR Alzheimer's Images*, publicado nos anais do evento *IEEE Congress on Evolutionary Computation (CEC 2008)*, Qualis A2, acontecido como parte do *IEEE World Congress on Computational Intelligence (WCCI 2008)*, em Hong Kong, China, de 1o. a 6 de junho de 2008 [82];
2. *Evaluation of Alzheimer's Disease by Analysis of MR Images using Objective Dialectical Classifiers as an Alternative to ADC Maps*, publicado nos anais do evento *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2008)*, Qualis A1, acontecido de 20 a 24 de agosto de 2008 em Vancouver, Canadá [83];
3. *Dialectical Non-Supervised Image Classification*, publicado nos anais do evento *IEEE Congress on Evolutionary Computation (CEC 2009)*, Qualis A2, em Trondheim, Noruega, de 18 a 21 de maio de 2009 [180];
4. *Optimization based on Dialectics*, publicado nos anais do evento *IEEE International*

- Joint Conference on Neural Networks (IJCNN 2009)*, Qualis A1, em Atlanta, EUA, de 14 a 19 de junho de 2009 [181];
5. *Dialectical Classification of MR Images for the Evaluation of Alzheimer's Disease*, aceito para publicação como capítulo do livro *Recent Advances in Biomedical Engineering*, pela editora InTech, Viena, Áustria, 2009 [182];
  6. *A Dialectical Method to Classify Alzheimer's Magnetic Resonance Images*, aceito para publicação como capítulo do livro *Evolutionary Computation*, pela editora InTech, Viena, Áustria, 2009 [183];
  7. *A Monospectral Approach for fMRI Analysis using Kohonen Self-Organized Networks and Objective Dialectical Classifiers*, publicado no periódico internacional *International Journal of Innovative Computing and Applications*, Qualis B5, 2008 [184];
  8. *Dialectical Multispectral Classification of Diffusion-Weighted Magnetic Resonance Images as an Alternative to Apparent Diffusion Coefficients Maps to Perform Anatomical Analysis*, publicado no periódico internacional *Computerized Medical Imaging and Graphics*, Qualis B2, 2009 [185];
  9. *An Object-Oriented Approach to Design Dialectical Systems*, aceito para publicação como capítulo do livro *Object-Oriented Analysis and Design*, pela editora Nova Science, New York, 2009 [186];
  10. *MRI Segmentation using Dialectical Optimization*, publicado nos anais do evento *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2009)*, Qualis A1, acontecido de 2 a 6 de setembro de 2009 em Minneapolis, EUA [187].

### 9.3 Trabalhos Futuros

Levando em conta que o método dialético para busca e otimização acabou se tornando a principal contribuição deste trabalho, como trabalhos futuros sugere-se a aplicação do método à otimização de outros algoritmos de segmentação de imagens *pixel a pixel*, tais como o *fuzzy c-médias* e o mapa auto-organizado de Kohonen, segundo critérios diversos, desde índices de validade de agrupamento e índices de fidelidade até estimativas de erro. Também seria interessante incluir nesses trabalhos futuros análises de curvas ROC, para permitir uma melhor comparação entre as técnicas de classificação.

Uma outra proposta de trabalho futuro pode ser a aplicação do método dialético de busca e otimização para substituir o algoritmo de treinamento de redes neurais diversas, tendo em vista a obtenção de redes otimizadas, minimizando o efeito dos mínimos locais presente nos algoritmos de treinamento usuais. Poder-se-ia ainda estudar a aplicação do método dialético de otimização usando pólos iniciais gerados segundo a proposta de Saha e Bandyopadhyay [143], empregando os algoritmos de treinamento usuais na geração de metade dos pólos iniciais (uma vez que, de acordo com o algoritmo dialético, a outra metade é composta de pólos opostos)

sendo executados durante uma quantidade muito pequena de iterações (inicialmente 5 iterações [143]), comparando o desempenho dessa estratégia com os algoritmos de treinamento padrão quanto à robustez em relação aos mínimos locais, o desempenho de classificação, o erro final ou até mesmo o custo computacional.

Outra perspectiva de trabalho futuro pode ser ainda a análise dinâmica dos algoritmos propostos analisando critérios de convergência, utilizando o formalismo de sistemas dinâmicos, uma vez que praticamente todos os algoritmos evolucionários, e algoritmos heurísticos em geral, carecem desse tipo de estudo, dada a complexidade de se modelar tais algoritmos como sistemas dinâmicos, o que também está relacionado à grande quantidade de parâmetros iniciais necessários. Entretanto, um dos focos deste trabalho esteve nas aplicações do método dialético em análise de imagens multiespectrais, o que fez com que essa análise não fizesse parte deste trabalho.

Outra sugestão ainda seria o estudo do método dialético de busca e otimização na resolução de problemas de otimização multiobjetivo, dado que adaptações e até mesmo o uso de diversos otimizadores em conjunto são necessários para resolver muitos problemas dessa natureza.

# Resultados de Minimização de Funções de Teste

## A.1 Introdução

Este apêndice apresenta os resultados completos de minimização das 24 funções de teste padrão apresentadas no capítulo 5 e utilizadas para validação do método dialético de busca e otimização proposto. Os resultados estão organizados por método de otimização, considerando os métodos baseados em programação evolucionária; o método de otimização por enxame de partículas baseado em consciência individual e coletiva; e os métodos dialéticos canônico e baseado no princípio da máxima entropia, com passos histórico e contemporâneo fixos e variáveis.

## A.2 Resultados por Método

### A.2.1 Algoritmo CEP

As tabelas A.1, A.2, A.3 e A.4 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.5, A.6, A.7 e A.8 mostram as faixas dos resultados obtidos; as tabelas A.13, A.14, A.15 e A.16 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.9, A.10, A.11 e A.12 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o algoritmo CEP.

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	7,9205	0,6611	8,1114
	50	7,3138	0,6006	7,3982
	100	6,8131	0,5885	6,9375
$f_2$	20	55,3494	41,5114	27,3422
	50	25,8869	12,9367	18,5830
	100	17,6750	2,5090	17,0870
$f_3$	20	55568,4125	9130,2526	56923,8228
	50	46704,6927	6955,8246	46794,7088
	100	41266,5859	5738,8370	41968,3494
$f_4$	20	18,5143	2,6264	18,4533
	50	16,3907	1,6928	16,5043
	100	14,6756	1,7444	14,6589
$f_5$	20	1099,3313	121,9763	1103,3853
	50	1015,6713	92,9273	1013,3577
	100	911,4575	92,6984	922,4433
$f_6$	20	24,4083	0,5814	24,1922
	50	23,9138	0,3986	24,0871
	100	23,4679	0,4923	23,2232
$f_7$	20	61,6515	10,6503	61,5048
	50	53,5161	8,5535	53,9736
	100	47,0174	7,6480	48,1186

Tabela A.1 Média, desvio médio e mediana dos valores mínimos (CEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-9126,5932	320,7103	-9024,4796
	50	-9353,5791	231,1685	-9261,6300
	100	-9564,0389	220,9461	-9488,5419
$f_9$	20	261,8747	14,8856	264,2213
	50	248,7359	11,7464	250,4365
	100	245,5979	9,0282	248,2076
$f_{10}$	20	16,5149	4,8544	19,6893
	50	14,8999	6,0317	19,0344
	100	12,6505	6,7332	17,5719
$f_{11}$	20	0,5493	0,03445	0,5616
	50	0,5176	0,0318	0,5202
	100	0,4689	0,0361	0,4717
$f_{12}$	20	23,6710	1,1338	24,1481
	50	22,7151	1,2920	23,0378
	100	21,3860	1,3744	21,8736
$f_{13}$	20	1,3259	0,1239	1,3339
	50	1,2093	0,1017	1,2236
	100	1,1682	0,0785	1,1698

Tabela A.2 Média, desvio médio e mediana dos valores mínimos (CEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	1,7230	0,8523	1,0051
	50	1,0605	0,1117	1,0000
	100	1,0110	0,0196	0,9993
$f_{15}$	20	0,0009	4,44E-5	0,0009
	50	0,0009	4,89E-5	0,0009
	100	0,0009	5,37E-5	0,0009
$f_{16}$	20	-1,0258	0,0026	-1,0262
	50	-1,0261	0,0026	-1,0266
	100	-1,0253	0,0027	-1,0253
$f_{17}$	20	0,3988	0,0005	0,3988
	50	0,3989	0,0005	0,3989
	100	0,3989	0,0004	0,3989
$f_{18}$	20	3,0050	0,0025	3,0051
	50	3,0052	0,0023	3,0055
	100	3,0046	0,0026	3,0042
$f_{19}$	20	-3,8258	0,0149	-3,8239
	50	-3,8229	0,0123	-3,8210
	100	-3,8238	0,0140	-3,8194
$f_{20}$	20	-3,0034	0,0656	-2,9966
	50	-3,0630	0,0522	-3,0651
	100	-3,0873	0,0394	-3,0774
$f_{21}$	20	-9,5763	0,2470	-9,6217
	50	-9,7182	0,1344	-9,6794
	100	-9,7232	0,1372	-9,6995
$f_{22}$	20	-9,7826	0,1749	-9,7508
	50	-9,8228	0,1903	-9,8089
	100	-9,7805	0,1828	-9,7130
$f_{23}$	20	-9,8110	0,1965	-9,7441
	50	-9,8100	0,1963	-9,7851
	100	-9,8496	0,1933	-9,8333

**Tabela A.3** Média, desvio médio e mediana dos valores mínimos (CEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	679,1279	23,9403	687,0408
	50	680,3059	24,7969	691,8136
	100	671,0780	25,5134	676,9189

**Tabela A.4** Média, desvio médio e mediana dos valores mínimos (CEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	7,9205	5,3175	9,5257
	50	7,3138	4,8857	8,5908
	100	6,8131	3,8570	8,2006
$f_2$	20	55,3494	13,1071	147,9932
	50	25,8869	13,4404	116,4675
	100	17,6750	12,4682	30,6042
$f_3$	20	55568,4125	27038,2632	87665,8467
	50	46704,6927	27107,4890	68964,8935
	100	41266,5859	23036,5801	55019,1686
$f_4$	20	18,5143	8,5151	27,0460
	50	16,3907	10,4962	23,2764
	100	14,6756	7,8932	19,7368
$f_5$	20	1099,3313	799,0836	1360,1038
	50	1015,6713	679,3607	1186,7178
	100	911,4575	544,8669	1095,4566
$f_6$	20	24,4083	22,3161	25,0780
	50	23,9138	22,1623	25,0557
	100	23,4679	22,2143	24,2421
$f_7$	20	61,6515	31,4855	92,1680
	50	53,5161	24,3164	82,7943
	100	47,0174	30,1313	71,8722

Tabela A.5 Faixa dos valores mínimos encontrados (CEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-9126,5932	-10850,3771	-8699,0275
	50	-9353,5791	-10120,4616	-8817,4141
	100	-9564,0389	-10218,7027	-9074,0554
$f_9$	20	261,8747	221,3032	296,4740
	50	248,7359	219,9871	284,3681
	100	245,5979	220,5347	265,8310
$f_{10}$	20	16,5149	3,9727	20,2263
	50	14,8999	3,7417	20,0472
	100	12,6505	3,8743	19,6809
$f_{11}$	20	0,5493	0,4652	0,6543
	50	0,5176	0,3969	0,6026
	100	0,4689	0,3760	0,5379
$f_{12}$	20	23,6710	20,3588	25,6954
	50	22,7151	18,6523	25,1785
	100	21,3860	16,5431	23,8684
$f_{13}$	20	1,3259	0,8772	1,5357
	50	1,2093	0,9258	1,4038
	100	1,1682	0,9830	1,3216

Tabela A.6 Faixa dos valores mínimos encontrados (CEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	1,7230	0,9980	6,9033
	50	1,0605	0,9980	1,9920
	100	1,0110	0,9980	1,9920
$f_{15}$	20	0,0009	0,0006	0,0009
	50	0,0009	0,0005	0,0009
	100	0,0009	0,0006	0,0009
$f_{16}$	20	-1,0258	-1,0314	-1,0202
	50	-1,0261	-1,0315	-1,0201
	100	-1,0253	-1,0313	-1,0201
$f_{17}$	20	0,3988	0,3978	0,3999
	50	0,3989	0,3978	0,3999
	100	0,3989	0,3979	0,3999
$f_{18}$	20	3,0050	3,0001	3,0099
	50	3,0052	3,0000	3,0099
	100	3,0046	3,0000	3,0097
$f_{19}$	20	-3,8258	-3,8576	-3,8008
	50	-3,8229	-3,8626	-3,8012
	100	-3,8238	-3,8590	-3,8003
$f_{20}$	20	-3,0034	-3,1761	-2,7952
	50	-3,0630	-3,2089	-2,8479
	100	-3,0873	-3,2446	-2,9940
$f_{21}$	20	-9,5763	-10,0693	-4,9874
	50	-9,7182	-10,1039	-9,5028
	100	-9,7232	-10,1193	-9,5010
$f_{22}$	20	-9,7826	-10,2926	-9,5007
	50	-9,8228	-10,3096	-9,5055
	100	-9,7805	-10,2939	-9,5040
$f_{23}$	20	-9,8110	-10,4448	-9,5002
	50	-9,8100	-10,4682	-9,5032
	100	-9,8496	-10,4687	-9,5003

Tabela A.7 Faixa dos valores mínimos encontrados (CEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	679,1279	598,8475	714,8695
	50	680,3059	581,5333	715,7118
	100	671,0780	584,1038	715,5134

Tabela A.8 Faixa dos valores mínimos encontrados (CEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_2$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_3$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_4$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_5$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_6$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_7$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.9 Número de iterações para minimização (CEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_9$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{10}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{11}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{12}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{13}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.10 Número de iterações para minimização (CEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	20522,69	59	50000	59
	50	3028,59	29	50000	94
	100	523,63	23	19	99
$f_{15}$	20	3887,52	2479	1191	100
	50	1583,7	1248	1021	100
	100	1003,88	825	608	100
$f_{16}$	20	49,74	39	30	100
	50	30,85	29	23	100
	100	19,62	18	13	100
$f_{17}$	20	103,33	86	28	100
	50	48,96	43	38	100
	100	36,45	33	46	100
$f_{18}$	20	2102,38	1776,5	3638	100
	50	651,66	488,5	114	100
	100	522,36	376,5	833	100
$f_{19}$	20	429,33	279	1	100
	50	203,78	129,5	1	100
	100	84,77	60,5	1	100
$f_{20}$	20	50000	50000	50000	0
	50	49827,68	50000	50000	1
	100	48799,4	50000	50000	3,33
$f_{21}$	20	19436,52	16081,5	50000	90
	50	9740,7	6231,5	7198	100
	100	5453,5	4136,5	5556	100
$f_{22}$	20	6846,29	5136,5	8428	100
	50	3034,16	2240,5	402	100
	100	1549,25	1352	582	100
$f_{23}$	20	7800,64	5138	1770	100
	50	3441,38	2543	1391	100
	100	1747,15	1644	2391	100

**Tabela A.11** Número de iterações para minimização (CEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	925,67	126	89	100
	50	88,12	71	75	100
	100	63,12	58,5	77	100

**Tabela A.12** Número de iterações para minimização (CEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_2$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_3$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_4$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_5$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_6$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_7$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0

**Tabela A.13** Número de avaliações da função (CEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_9$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{10}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{11}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{12}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{13}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0

**Tabela A.14** Número de avaliações da função (CEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	820907,6	2360	2000000	59
	50	302859	2900	5000000	94
	100	104726	4600	3800	99
$f_{15}$	20	155500,8	99160	47640	100
	50	158370	124800	102100	100
	100	200776	165000	121600	100
$f_{16}$	20	1989,6	1560	1200	100
	50	3085	2900	2300	100
	100	3924	3600	2600	100
$f_{17}$	20	4133,2	3440	1120	100
	50	4896	4300	3800	100
	100	7290	6600	9200	100
$f_{18}$	20	84095,2	71060	145520	100
	50	65166	48850	11400	100
	100	104472	75300	166600	100
$f_{19}$	20	17173,2	11160	40	100
	50	20378	12950	100	100
	100	16954	12100	200	100
$f_{20}$	20	2000000	2000000	2000000	0
	50	4982768	5000000	5000000	1
	100	9759880	10000000	10000000	3,33
$f_{21}$	20	777460,8	643260	2000000	90
	50	974070	623150	719800	100
	100	1090700	827300	1111200	100
$f_{22}$	20	273851,6	205460	337120	100
	50	303416	224050	40200	100
	100	309850	270400	116400	100
$f_{23}$	20	312025,6	205520	70800	100
	50	344138	254300	139100	100
	100	349430	328800	478200	100

**Tabela A.15** Número de avaliações da função (CEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	37026,8	5040	3560	100
	50	8812	7100	7500	100
	100	12624	11700	15400	100

**Tabela A.16** Número de avaliações da função (CEP)

### A.2.2 Algoritmo FEP

As tabelas A.17, A.18, A.19 e A.20 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.21, A.22, A.23 e A.24 mostram as faixas dos resultados obtidos; as tabelas A.29, A.30, A.31 e A.32 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.25, A.26, A.27 e A.28 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o algoritmo FEP.

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	43,0987	3,1157	43,6922
	50	38,9664	3,0060	39,4102
	100	37,2969	3,5241	38,5579
$f_2$	20	145,2394	51,4863	135,8213
	50	107,1062	42,4184	102,7488
	100	71,1598	27,6138	61,6515
$f_3$	20	10919,2869	1047,9031	10913,6143
	50	10251,0863	711,7154	10157,6017
	100	9458,0318	893,0507	9725,3881
$f_4$	20	3,1603	0,2118	3,1524
	50	3,0150	0,1502	3,0237
	100	2,9426	0,1227	2,9534
$f_5$	20	7348,2309	934,5728	7459,8067
	50	6436,7052	846,6251	6355,1246
	100	5508,4471	733,7421	5589,6204
$f_6$	20	22,5082	0,3651	22,4813
	50	22,1868	0,2157	22,2849
	100	22,0220	0,3254	22,1952
$f_7$	20	137,6215	18,4529	136,8839
	50	113,0896	17,9912	112,1131
	100	104,9818	16,7573	109,6798

**Tabela A.17** Média, desvio médio e mediana dos valores mínimos (FEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-30853,3734	25167,4221	-13611,9112
	50	-22462,9419	10066,0629	-15820,8000
	100	-43468,0658	34655,6345	-23184,2698
$f_9$	20	309,0819	12,2012	310,8605
	50	298,0218	8,2268	299,5134
	100	283,9660	9,6695	286,5927
$f_{10}$	20	6,1909	0,1764	6,2100
	50	5,9118	0,1410	5,9151
	100	5,8816	0,1599	5,9453
$f_{11}$	20	0,8801	0,0333	0,8878
	50	0,8541	0,0338	0,8696
	100	0,8335	0,0352	0,8377
$f_{12}$	20	2,5054	0,3285	2,4011
	50	2,2063	0,3031	2,2207
	100	1,9328	0,1980	1,9225
$f_{13}$	20	7,4625	0,7172	7,5695
	50	6,9648	0,5478	6,8866
	100	6,3003	0,6454	6,2743

**Tabela A.18** Média, desvio médio e mediana dos valores mínimos (FEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	0,9996	0,0021	0,9982
	50	1,0014	0,0030	1,0002
	100	1,0011	0,0024	1,0002
$f_{15}$	20	0,0009	2,45E-5	0,0009
	50	0,0009	4,75E-5	0,0009
	100	0,0009	6,25E-5	0,0009
$f_{16}$	20	-1,0264	0,0028	-1,0262
	50	-1,0253	0,0032	-1,0256
	100	-1,0255	0,0034	-1,0238
$f_{17}$	20	0,3990	0,0005	0,3990
	50	0,3989	0,0005	0,3988
	100	0,3988	0,0005	0,3987
$f_{18}$	20	3,0042	0,0024	3,0041
	50	3,0048	0,0022	3,0051
	100	3,0041	0,0023	3,0036
$f_{19}$	20	-3,8252	0,0139	-3,8198
	50	-3,8223	0,0116	-3,8211
	100	-3,8194	0,0118	-3,8179
$f_{20}$	20	-2,8723	0,0977	-2,8783
	50	-2,9744	0,0719	-2,9656
	100	-3,0270	0,0600	-3,0295
$f_{21}$	20	-9,7245	0,1292	-9,7463
	50	-9,7271	0,1479	-9,6657
	100	-9,7035	0,1193	-9,6972
$f_{22}$	20	-9,8054	0,1895	-9,7692
	50	-9,8298	0,1810	-9,7533
	100	-9,8201	0,1650	-9,7861
$f_{23}$	20	-9,7874	0,1192	-9,7843
	50	-9,8297	0,2208	-9,7801
	100	-9,8495	0,2293	-9,7605

**Tabela A.19** Média, desvio médio e mediana dos valores mínimos (FEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	681,2091	21,2530	688,3708
	50	680,5946	25,4597	690,5921
	100	678,7354	27,4442	690,0542

**Tabela A.20** Média, desvio médio e mediana dos valores mínimos (FEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	43,0987	32,8119	52,5883
	50	38,9664	31,4161	44,4322
	100	37,2969	24,8811	43,6569
$f_2$	20	145,2394	34,9895	603,0189
	50	107,1062	33,9537	500,8914
	100	71,1598	32,3753	169,3894
$f_3$	20	10919,2869	7422,8460	13214,2522
	50	10251,0863	8013,1296	12133,4027
	100	9458,0318	6385,3050	11006,2905
$f_4$	20	3,1603	2,6254	3,8170
	50	3,0150	2,4963	3,4851
	100	2,9426	2,5543	3,2471
$f_5$	20	7348,2309	4724,5573	9148,2613
	50	6436,7052	4339,1239	8157,5859
	100	5508,4471	3125,9671	7316,6933
$f_6$	20	22,5082	20,6851	23,2726
	50	22,1868	21,4424	22,5258
	100	22,0220	20,6826	22,6032
$f_7$	20	137,6215	80,9238	191,1756
	50	113,0896	72,0246	150,4733
	100	104,9818	44,8517	131,3946

**Tabela A.21** Faixa dos valores mínimos encontrados (FEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-30853,3734	-253813,1137	-12504,6057
	50	-22462,9419	-55351,9695	-12565,6797
	100	-43468,0658	-272161,8007	-12617,3199
$f_9$	20	309,0819	274,4357	332,9920
	50	298,0218	280,0231	314,6610
	100	283,9660	254,4656	301,5856
$f_{10}$	20	6,1909	5,7757	6,5679
	50	5,9118	5,4050	6,2574
	100	5,8816	5,4796	6,1516
$f_{11}$	20	0,8801	0,7717	0,9413
	50	0,8541	0,7101	0,9153
	100	0,8335	0,7223	0,9119
$f_{12}$	20	2,5054	1,8695	3,6146
	50	2,2063	1,6236	2,9451
	100	1,9328	1,3314	2,4023
$f_{13}$	20	7,4625	6,1152	9,0985
	50	6,9648	4,8275	8,4308
	100	6,3003	4,9110	8,1055

**Tabela A.22** Faixa dos valores mínimos encontrados (FEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	0,9996	0,9980	1,0094
	50	1,0014	0,9980	1,0097
	100	1,0011	0,9980	1,0073
$f_{15}$	20	0,0009	0,0008	0,0010
	50	0,0009	0,0008	0,0009
	100	0,0009	0,0007	0,0009
$f_{16}$	20	-1,0264	-1,0316	-1,0202
	50	-1,0253	-1,0315	-1,0200
	100	-1,0255	-1,0314	-1,0200
$f_{17}$	20	0,3990	0,3979	0,3999
	50	0,3989	0,3980	0,3999
	100	0,3988	0,3979	0,3999
$f_{18}$	20	3,0042	3,0001	3,0097
	50	3,0048	3,0001	3,0094
	100	3,0041	3,0001	3,0097
$f_{19}$	20	-3,8252	-3,8568	-3,8007
	50	-3,8223	-3,8452	-3,8012
	100	-3,8194	-3,8570	-3,8002
$f_{20}$	20	-2,8723	-3,0678	-2,6385
	50	-2,9744	-3,1256	-2,7869
	100	-3,0270	-3,2451	-2,8808
$f_{21}$	20	-9,7245	-10,0113	-9,4968
	50	-9,7271	-10,0573	-9,5052
	100	-9,7035	-10,0501	-9,5136
$f_{22}$	20	-9,8054	-10,2139	-9,5101
	50	-9,8298	-10,2898	-9,5446
	100	-9,8201	-10,2895	-9,5325
$f_{23}$	20	-9,7874	-10,2246	-9,5005
	50	-9,8297	-10,3699	-9,5174
	100	-9,8495	-10,4552	-9,5144

**Tabela A.23** Faixa dos valores mínimos encontrados (FEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	681,2091	602,8248	715,9980
	50	680,5946	572,7786	715,6341
	100	678,7354	577,1295	715,4078

**Tabela A.24** Faixa dos valores mínimos encontrados (FEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_2$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_3$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_4$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_5$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_6$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_7$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.25 Número de iterações para minimização (FEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	41,46	32,5	28	100
	50	26,73	20	15	100
	100	15,03	14,5	2	100
$f_9$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{10}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{11}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{12}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{13}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.26 Número de iterações para minimização (FEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	68,63	39	44	100
	50	46,83	29,5	31	100
	100	23,56	21,5	19	100
$f_{15}$	20	11496,7	6587	3534	96,66
	50	4114,86	3842,5	3587	100
	100	2708,7	2321	1905	100
$f_{16}$	20	50,96	46	93	100
	50	36,53	32	26	100
	100	22,6	21	21	100
$f_{17}$	20	118,46	109,5	99	100
	50	82,1	74	48	100
	100	51,86	50	19	100
$f_{18}$	20	2038,03	1778	1498	100
	50	877,46	614	181	100
	100	485,7	428,5	425	100
$f_{19}$	20	1038,7	849	1967	100
	50	241,43	140,5	1	100
	100	178,76	110	1	100
$f_{20}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	49793,36	50000	50000	3,33
$f_{21}$	20	16967,46	15570,5	11836	96,66
	50	9585,1	6298	1089	100
	100	3166,13	2698,5	2109	100
$f_{22}$	20	6512,1	4511,5	18661	100
	50	3150,1	2925,5	3678	100
	100	2106,2	1643,5	755	100
$f_{23}$	20	5909,43	5017,5	2366	100
	50	2567,46	2188	1420	100
	100	2042,26	2091,5	687	100

**Tabela A.27** Número de iterações para minimização (FEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	284,39	266	268	100
	50	163,9	157	158	100
	100	138,06	143,5	67	100

**Tabela A.28** Número de iterações para minimização (FEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_2$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_3$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_4$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_5$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_6$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_7$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0

**Tabela A.29** Número de avaliações da função (FEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	1658,66	1300	1120	100
	50	2673,33	2000	1500	100
	100	3006,66	2900	400	100
$f_9$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{10}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{11}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{12}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{13}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0

**Tabela A.30** Número de avaliações da função (FEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	2745,33	1560	1760	100
	50	4683,33	2950	3100	100
	100	4713,33	4300	3800	100
$f_{15}$	20	459868	263480	141360	96,66
	50	411486,66	384250	358700	100
	100	541740	464200	381000	100
$f_{16}$	20	2038,66	1840	3720	100
	50	3653,33	3200	2600	100
	100	4520	4200	4200	100
$f_{17}$	20	4738,66	4380	3960	100
	50	8210	7400	4800	100
	100	10373,33	10000	3800	100
$f_{18}$	20	81521,33	71120	59920	100
	50	87746,66	61400	18100	100
	100	97140	85700	85000	100
$f_{19}$	20	41548	33960	78680	100
	50	24143,33	14050	100	100
	100	35753,33	22000	200	100
$f_{20}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	9958673,33	10000000	10000000	3,33
$f_{21}$	20	678698,66	622820	473440	96,66
	50	958510	629800	108900	100
	100	633226,66	539700	421800	100
$f_{22}$	20	260484	180460	746440	100
	50	315010	292550	367800	100
	100	421240	328700	151000	100
$f_{23}$	20	236377,33	200700	94640	100
	50	256746,66	218800	142000	100
	100	408453,33	418300	137400	100

**Tabela A.31** Número de avaliações da função (FEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	11375,6	10640	10720	100
	50	16390	15700	15800	100
	100	27612	28700	13400	100

**Tabela A.32** Número de avaliações da função (FEP)

### A.2.3 Algoritmo LEP

As tabelas A.33, A.34, A.35 e A.36 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.37, A.38, A.39 e A.40 mostram as faixas dos resultados obtidos; as tabelas A.45, A.46, A.47 e A.48 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.41, A.42, A.43 e A.44 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o algoritmo LEP, com  $\beta = 0,8$ .

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	78,4822	7,7252	79,0620
	50	72,8104	6,8889	74,0547
	100	70,5875	5,2569	70,7249
$f_2$	20	13178,4324	23144,4183	599,2632
	50	349,8094	187,7557	254,3864
	100	190,8946	57,4286	172,0374
$f_3$	20	19250,0453	2092,7945	19374,5415
	50	17385,6093	1571,1859	17571,9795
	100	15413,9055	1418,3857	15873,3713
$f_4$	20	4,5793	0,2437	4,6659
	50	4,3427	0,2230	4,3944
	100	4,1455	0,2667	4,1940
$f_5$	20	7918,9705	1096,2267	7940,2351
	50	6839,4689	781,5602	6883,7885
	100	6217,1877	574,1143	6265,2128
$f_6$	20	22,3341	0,3691	22,3275
	50	22,1740	0,2821	22,2695
	100	21,8557	0,3758	22,0678
$f_7$	20	128,6865	22,7305	128,0524
	50	110,6662	15,2623	111,7489
	100	95,4334	14,9234	96,8876

**Tabela A.33** Média, desvio médio e mediana dos valores mínimos (LEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-71815,7862	68754,6341	-31351,5496
	50	-151471,5699	168510,3818	-32212,2313
	100	-60795,0904	48297,4677	-33337,3542
$f_9$	20	351,1672	17,3937	356,3454
	50	338,0304	14,2313	340,2547
	100	327,2967	14,1687	330,1137
$f_{10}$	20	7,4735	0,2658	7,4948
	50	7,2774	0,1930	7,2802
	100	7,0457	0,2012	7,0931
$f_{11}$	20	0,9870	0,0179	0,9916
	50	0,9729	0,0192	0,9770
	100	0,9587	0,0203	0,9665
$f_{12}$	20	4,8050	0,7897	4,9130
	50	4,1875	0,3932	4,2644
	100	3,7297	0,4596	3,8238
$f_{13}$	20	18,0250	2,4899	18,3030
	50	16,4714	1,8179	16,5427
	100	15,4826	1,5477	15,2346

**Tabela A.34** Média, desvio médio e mediana dos valores mínimos (LEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	1,0004	0,0023	0,9994
	50	1,0003	0,0017	0,9996
	100	1,0016	0,0029	1,0006
$f_{15}$	20	0,0009	2,21E-5	0,0009
	50	0,0009	3,80E-5	0,0009
	100	0,0009	3,76E-5	0,0009
$f_{16}$	20	-1,0259	0,0030	-1,0251
	50	-1,0242	0,0024	-1,0232
	100	-1,0252	0,0020	-1,0253
$f_{17}$	20	0,3990	0,0006	0,3990
	50	0,3989	0,0005	0,3989
	100	0,3987	0,0004	0,3987
$f_{18}$	20	3,0053	0,0024	3,0060
	50	3,0050	0,0020	3,0056
	100	3,0058	0,0025	3,0067
$f_{19}$	20	-3,8209	0,0136	-3,8131
	50	-3,8256	0,0132	-3,8258
	100	-3,8245	0,0148	-3,8214
$f_{20}$	20	-2,9094	0,0965	-2,9320
	50	-3,0029	0,0608	-2,9975
	100	-3,0322	0,0492	-3,0149
$f_{21}$	20	-9,6750	0,1193	-9,6253
	50	-9,7318	0,1367	-9,7041
	100	-9,7150	0,1456	-9,6423
$f_{22}$	20	-9,7398	0,1409	-9,6950
	50	-9,7822	0,2095	-9,6883
	100	-9,8287	0,1955	-9,7817
$f_{23}$	20	-9,8246	0,1924	-9,7359
	50	-9,8203	0,1896	-9,7515
	100	-9,7973	0,1987	-9,7680

**Tabela A.35** Média, desvio médio e mediana dos valores mínimos (LEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	683,2845	21,9196	689,9792
	50	677,2446	24,2793	682,6243
	100	683,7082	20,5954	691,0176

**Tabela A.36** Média, desvio médio e mediana dos valores mínimos (LEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	78,4822	48,1921	101,6835
	50	72,8104	34,1126	90,7492
	100	70,5875	47,4118	81,0999
$f_2$	20	13178,4324	108,9044	338956,4322
	50	349,8094	115,5038	1728,1688
	100	190,8946	86,7304	430,4621
$f_3$	20	19250,0453	11526,6208	24058,3596
	50	17385,6093	11605,8866	21699,3511
	100	15413,9055	10676,9177	18119,2662
$f_4$	20	4,5793	4,1268	5,2118
	50	4,3427	3,7224	4,8531
	100	4,1455	3,3690	4,6372
$f_5$	20	7918,9705	3744,8966	10848,8393
	50	6839,4689	3971,7227	9940,8713
	100	6217,1877	4480,1370	8087,6181
$f_6$	20	22,3341	21,4199	23,2343
	50	22,1740	20,6499	22,7209
	100	21,8557	20,4943	22,3541
$f_7$	20	128,6865	54,5832	198,7836
	50	110,6662	45,1841	154,2039
	100	95,4334	45,3614	133,1317

**Tabela A.37** Faixa dos valores mínimos encontrados (LEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-71815,7862	-485132,3496	-12824,0353
	50	-151471,5699	-1187267,6075	-13148,2291
	100	-60795,0904	-395285,1877	-12670,8168
$f_9$	20	351,1672	281,9848	388,8261
	50	338,0304	289,0484	379,5238
	100	327,2967	275,8434	356,8972
$f_{10}$	20	7,4735	6,7881	8,0857
	50	7,2774	6,7379	7,8393
	100	7,0457	6,2090	7,4388
$f_{11}$	20	0,9870	0,9052	1,0238
	50	0,9729	0,8749	1,0102
	100	0,9587	0,8642	0,9995
$f_{12}$	20	4,8050	3,0572	6,4811
	50	4,1875	2,9193	4,9992
	100	3,7297	1,9750	4,4548
$f_{13}$	20	18,0250	8,7317	25,2193
	50	16,4714	11,3870	21,6393
	100	15,4826	11,5912	19,5646

**Tabela A.38** Faixa dos valores mínimos encontrados (LEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	1,0004	0,9980	1,0094
	50	1,0003	0,9980	1,0071
	100	1,0016	0,9980	1,009
$f_{15}$	20	0,0009	0,0008	0,0010
	50	0,0009	0,0007	0,0009
	100	0,0009	0,0008	0,0009
$f_{16}$	20	-1,0259	-1,0314	-1,0202
	50	-1,0242	-1,0312	-1,0202
	100	-1,0252	-1,0302	-1,0200
$f_{17}$	20	0,3990	0,3979	0,3999
	50	0,3989	0,3979	0,3999
	100	0,3987	0,3978	0,3998
$f_{18}$	20	3,0053	3,0007	3,0094
	50	3,0050	3,0005	3,0093
	100	3,0058	3,0001	3,0097
$f_{19}$	20	-3,8209	-3,8617	-3,8018
	50	-3,8256	-3,8571	-3,8005
	100	-3,8245	-3,8625	-3,8000
$f_{20}$	20	-2,9094	-3,0727	-2,6538
	50	-3,0029	-3,2551	-2,8668
	100	-3,0322	-3,1511	-2,9328
$f_{21}$	20	-9,6750	-10,0563	-9,5093
	50	-9,7318	-10,1415	-9,5083
	100	-9,7150	-10,1117	-9,5017
$f_{22}$	20	-9,7398	-10,2000	-9,5133
	50	-9,7822	-10,2669	-9,5090
	100	-9,8287	-10,3396	-9,5038
$f_{23}$	20	-9,8246	-10,4509	-9,5483
	50	-9,8203	-10,4060	-9,5104
	100	-9,7973	-10,4589	-9,5182

**Tabela A.39** Faixa dos valores mínimos encontrados (LEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	683,2845	560,9580	715,8059
	50	677,2446	591,3238	715,8492
	100	683,7082	585,6153	715,2096

**Tabela A.40** Faixa dos valores mínimos encontrados (LEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_2$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_3$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_4$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_5$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_6$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_7$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.41 Número de iterações para minimização (LEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	12,6	11	3	100
	50	6,03	6	7	100
	100	4,2	3	3	100
$f_9$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{10}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{11}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{12}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{13}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.42 Número de iterações para minimização (LEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	87,23	52,5	55	100
	50	41,83	35,5	45	100
	100	28,23	28,5	16	100
$f_{15}$	20	18239,86	14293	50000	90
	50	6761,93	5324,5	5393	100
	100	3826,53	3264	291	100
$f_{16}$	20	62,23	58,5	23	100
	50	34,43	33,5	14	100
	100	25,73	23,5	25	100
$f_{17}$	20	146,73	140,5	67	100
	50	71,93	70	54	100
	100	46,63	48	13	100
$f_{18}$	20	1761,03	1754	1047	100
	50	654	492	852	100
	100	593,86	454,5	51	100
$f_{19}$	20	665,13	492	1797	100
	50	245,16	185	1	100
	100	176,23	106	82	100
$f_{20}$	20	50000	50000	50000	0
	50	49907,16	50000	50000	3,33
	100	50000	50000	50000	0
$f_{21}$	20	12419,43	10721,5	10910	100
	50	5668,48	4950,5	7298	100
	100	3381,63	3006	2967	100
$f_{22}$	20	6018,36	4646	11615	100
	50	2825,76	2786,5	3009	100
	100	1667,86	1400,5	1075	100
$f_{23}$	20	5206,16	4052,5	7874	100
	50	3029,19	2717,5	969	100
	100	1309,73	1168,5	3417	100

**Tabela A.43** Número de iterações para minimização (LEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	412,08	369	267	100
	50	240,79	236,5	182	100
	100	183,38	179,5	211	100

**Tabela A.44** Número de iterações para minimização (LEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_2$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_3$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_4$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_5$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_6$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_7$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0

**Tabela A.45** Número de avaliações da função (LEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	504	440	120	100
	50	603,33	600	700	100
	100	840	600	600	100
$f_9$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{10}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{11}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{12}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{13}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0

**Tabela A.46** Número de avaliações da função (LEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	3489,33	2100	2200	100
	50	4183,33	3550	4500	100
	100	5646,66	5700	3200	100
$f_{15}$	20	729594,66	571720	2000000	90
	50	676193	532450	539300	100
	100	765306,66	652800	58200	100
$f_{16}$	20	2489,33	2340	920	100
	50	3443,33	3350	1400	100
	100	5146,66	4700	5000	100
$f_{17}$	20	5869,33	5620	2680	100
	50	7193	7000	5400	100
	100	9326,66	9600	2600	100
$f_{18}$	20	70441,33	70160	41880	100
	50	65400	49200	85200	100
	100	118773,33	90900	10200	100
$f_{19}$	20	26605,33	19680	71880	100
	50	24516	18500	100	100
	100	35246,66	21200	16400	100
$f_{20}$	20	2000000	2000000	2000000	0
	50	4990716,66	5000000	5000000	3,33
	100	10000000	10000000	10000000	0
$f_{21}$	20	496777,33	428860	436400	100
	50	566848	495050	729800	100
	100	676326,66	601200	593400	100
$f_{22}$	20	240734,66	185840	464600	100
	50	282576,66	278650	300900	100
	100	333573,33	280100	215000	100
$f_{23}$	20	208246,66	162100	314960	100
	50	302919	271750	96900	100
	100	261946,66	233700	683400	100

**Tabela A.47** Número de avaliações da função (LEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	16483,2	14760	10680	100
	50	24079	23650	18200	100
	100	36676	35900	42200	100

**Tabela A.48** Número de avaliações da função (LEP)

### A.2.4 Algoritmo SPMEP

As tabelas A.49, A.50, A.51 e A.52 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.53, A.54, A.55 e A.56 mostram as faixas dos resultados obtidos; as tabelas A.61, A.62, A.63 e A.64 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.57, A.58, A.59 e A.60 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o algoritmo SPMEP, com  $\alpha = 3/2$ .

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	40,2018	41,9471	17,0400
	50	3,9365	4,4740	1,3486
	100	1,4205	1,4545	0,7918
$f_2$	20	0,4587	0,2734	0,3827
	50	0,1699	0,1071	0,1528
	100	0,1123	0,0795	0,0802
$f_3$	20	1583,1419	1784,8646	425,8885
	50	461,7035	544,7898	128,7501
	100	158,0057	173,6789	46,1962
$f_4$	20	4,7511	3,6420	2,8946
	50	1,8107	1,6114	0,8907
	100	1,2222	0,8083	1,0236
$f_5$	20	31,5558	3,3890	29,5623
	50	29,2511	0,3263	29,0429
	100	29,0241	0,1098	28,9928
$f_6$	20	0,9345	0,1221	1
	50	0,8468	0,2552	1
	100	0,6782	0,4289	1
$f_7$	20	0,0026	0,0039	1,50E-5
	50	7,34E-5	0,0001	1,39E-6
	100	7,65E-6	1,35E-5	1,20E-7

**Tabela A.49** Média, desvio médio e mediana dos valores mínimos (SPMEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-9220,9110	799,5028	-9089,4230
	50	-10200,6806	549,2788	-10185,0588
	100	-10449,4176	683,3381	-10128,5840
$f_9$	20	1,8655	1,4539	1,5251
	50	0,9626	0,6295	0,9987
	100	0,4429	0,5032	0,1409
$f_{10}$	20	1,1333	0,7891	0,9363
	50	0,5090	0,2489	0,5671
	100	0,2951	0,1883	0,2917
$f_{11}$	20	1,0382	0,7104	0,8267
	50	0,4441	0,2900	0,3086
	100	0,2890	0,2177	0,1867
$f_{12}$	20	1,9599	0,3074	1,8354
	50	1,6400	0,0615	1,6140
	100	1,6215	0,0551	1,6136
$f_{13}$	20	1128,2965	2123,3159	3,2101
	50	3,0248	0,1105	2,9724
	100	3,0219	0,0857	2,9922

**Tabela A.50** Média, desvio médio e mediana dos valores mínimos (SPMEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	1,0017	0,0035	1,0009
	50	1,0004	0,0027	0,9989
	100	1,0003	0,0022	0,9991
$f_{15}$	20	0,0009	4,77E-5	0,0009
	50	0,0008	7,99E-5	0,0009
	100	0,0008	9,46E-5	0,0008
$f_{16}$	20	-1,0259	0,0027	-1,0254
	50	-1,0263	0,0026	-1,0274
	100	-1,0262	0,0024	-1,0260
$f_{17}$	20	0,3990	0,0004	0,3989
	50	0,3989	0,0005	0,3987
	100	0,3989	0,0004	0,3989
$f_{18}$	20	3,0046	0,0030	3,0050
	50	3,0045	0,0022	3,0048
	100	3,0046	0,0026	3,0049
$f_{19}$	20	-3,8258	0,0156	-3,8239
	50	-3,8258	0,0148	-3,8255
	100	-3,8207	0,0139	-3,8183
$f_{20}$	20	-3,2321	0,0164	-3,2252
	50	-3,2249	0,0155	-3,2190
	100	-3,2277	0,0183	-3,2202
$f_{21}$	20	-9,5198	0,1928	-9,5391
	50	-9,7125	0,1244	-9,7112
	100	-9,6901	0,1375	-9,6847
$f_{22}$	20	-9,7355	0,1809	-9,6808
	50	-9,8411	0,1860	-9,8246
	100	-9,8206	0,1446	-9,8258
$f_{23}$	20	-9,8002	0,1858	-9,7461
	50	-9,8515	0,1907	-9,8586
	100	-9,8138	0,1707	-9,7495

**Tabela A.51** Média, desvio médio e mediana dos valores mínimos (SPMEP)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	685,5533	23,7413	700,1099
	50	686,7939	20,2551	696,2014
	100	685,5913	20,6976	694,0414

**Tabela A.52** Média, desvio médio e mediana dos valores mínimos (SPMEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	40,2018	0,0043	365,7526
	50	3,9365	0,0099	50,0118
	100	1,4205	0,0055	11,4338
$f_2$	20	0,4587	0,0344	1,4967
	50	0,1699	0,0012	0,5176
	100	0,1123	0,0035	0,3412
$f_3$	20	1583,1419	0,2645	16536,4479
	50	461,7035	0,3052	3567,8411
	100	158,0057	0,2505	913,7459
$f_4$	20	4,7511	0,0830	20,7253
	50	1,8107	0,0760	13,5780
	100	1,2222	0,0074	5,2570
$f_5$	20	31,5558	28,9901	76,8883
	50	29,2511	28,8921	30,7164
	100	29,0241	28,7897	29,5553
$f_6$	20	0,9345	0,0092	1
	50	0,8468	0,0017	1
	100	0,6782	0,0001	1
$f_7$	20	0,0026	5,72E-12	0,0293
	50	7,34E-5	3,98E-29	0,0014
	100	7,65E-6	9,54E-14	0,0001

**Tabela A.53** Faixa dos valores mínimos encontrados (SPMEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-9220,9110	-11962,7219	-7903,2903
	50	-10200,6806	-11901,2548	-9165,6408
	100	-10449,4176	-12227,3043	-8889,7099
$f_9$	20	1,8655	0,0002	8,2115
	50	0,9626	0,0001	4,0257
	100	0,4429	0,0001	1,6816
$f_{10}$	20	1,1333	0,0161	4,5029
	50	0,5090	0,0406	1,5023
	100	0,2951	0,0002	0,9437
$f_{11}$	20	1,0382	0,0260	3,1123
	50	0,4441	0,0014	1,2706
	100	0,2890	0,0004	1,2937
$f_{12}$	20	1,9599	1,5966	3,6293
	50	1,6400	1,3607	2,0155
	100	1,6215	1,0957	1,8624
$f_{13}$	20	1128,2965	2,9089	32978,0353
	50	3,0248	2,8923	4,0015
	100	3,0219	2,9001	3,2275

**Tabela A.54** Faixa dos valores mínimos encontrados (SPMEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	1,0017	0,9980	1,0093
	50	1,0004	0,9980	1,0098
	100	1,0003	0,9980	1,0097
$f_{15}$	20	0,0009	0,0007	0,0009
	50	0,0008	0,0006	0,0009
	100	0,0008	0,0004	0,0009
$f_{16}$	20	-1,0259	-1,0315	-1,0201
	50	-1,0263	-1,0315	-1,0200
	100	-1,0262	-1,0305	-1,0203
$f_{17}$	20	0,3990	0,3979	0,3999
	50	0,3989	0,3980	0,3999
	100	0,3989	0,3980	0,3999
$f_{18}$	20	3,0046	3,0002	3,0097
	50	3,0045	3,0002	3,0094
	100	3,0046	3,0000	3,0098
$f_{19}$	20	-3,8258	-3,8581	-3,8003
	50	-3,8258	-3,8612	-3,8006
	100	-3,8207	-3,8556	-3,8001
$f_{20}$	20	-3,2321	-3,2952	-3,2042
	50	-3,2249	-3,2660	-3,2029
	100	-3,2277	-3,2899	-3,2004
$f_{21}$	20	-9,5198	-10,0856	-8,7841
	50	-9,7125	-10,0025	-9,3412
	100	-9,6901	-10,0484	-9,4855
$f_{22}$	20	-9,7355	-10,2288	-9,3859
	50	-9,8411	-10,3183	-9,5411
	100	-9,8206	-10,1400	-9,5128
$f_{23}$	20	-9,8002	-10,1738	-9,2190
	50	-9,8515	-10,3332	-9,5140
	100	-9,8138	-10,2544	-9,5420

**Tabela A.55** Faixa dos valores mínimos encontrados (SPMEP)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	685,5533	579,7882	715,2444
	50	686,7939	605,6463	715,3412
	100	685,5913	580,8417	715,6905

**Tabela A.56** Faixa dos valores mínimos encontrados (SPMEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_2$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_3$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_4$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_5$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_6$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	46666,8	50000	50000	6,66
$f_7$	20	13334,8	2	2	73,33
	50	1668,6	2	2	96,66
	100	2	2	2	100

Tabela A.57 Número de iterações para minimização (SPMEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_9$	20	48333,4	50000	50000	3,33
	50	48333,4	50000	50000	3,33
	100	43333,6	50000	50000	13,33
$f_{10}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	48333,4	50000	50000	3,33
$f_{11}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	48333,4	50000	50000	3,33
$f_{12}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{13}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.58 Número de iterações para minimização (SPMEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	304,8	317	328	100
	50	207,9	187	38	100
	100	137,8	131,5	65	100
$f_{15}$	20	1303,73	935,5	1440	100
	50	594,63	465,5	395	100
	100	487,53	411,5	706	100
$f_{16}$	20	52,66	39	20	100
	50	27,06	26	23	100
	100	14,83	13,5	13	100
$f_{17}$	20	327,3	228	674	100
	50	144,06	97	145	100
	100	77,73	67,5	5	100
$f_{18}$	20	475,83	319	3	100
	50	370,43	309	425	100
	100	109,73	93	2	100
$f_{19}$	20	39,7	35,5	34	100
	50	21,6	19	13	100
	100	14,2	15	9	100
$f_{20}$	20	256,86	237	374	100
	50	215,03	225	225	100
	100	152,16	157,5	95	100
$f_{21}$	20	28049,76	26458,5	50000	63,33
	50	15888,2	17790,5	7513	96,66
	100	14325,36	12391,5	24957	96,66
$f_{22}$	20	19669,26	16622	50000	86,66
	50	13338,33	9327	1006	100
	100	5433,86	3850,5	1308	100
$f_{23}$	20	15162,73	10067,5	31555	96,66
	50	7603,33	5215,5	17757	100
	100	4726	3538,5	7381	100

**Tabela A.59** Número de iterações para minimização (SPMEP)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	211,26	192	263	100
	50	123,25	108,5	148	100
	100	78,47	75	54	100

**Tabela A.60** Número de iterações para minimização (SPMEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_2$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_3$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_4$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_5$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_6$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	9333360	10000000	10000000	6,66
$f_7$	20	533392	80	80	73,33
	50	166860	200	200	96,66
	100	400	400	400	100

**Tabela A.61** Número de avaliações da função (SPMEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_9$	20	1933336	2000000	2000000	3,33
	50	4833340	5000000	5000000	3,33
	100	8666720	10000000	10000000	13,33
$f_{10}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	9666680	10000000	10000000	3,33
$f_{11}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	9666680	10000000	10000000	3,33
$f_{12}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0
$f_{13}$	20	2000000	2000000	2000000	0
	50	5000000	5000000	5000000	0
	100	10000000	10000000	10000000	0

**Tabela A.62** Número de avaliações da função (SPMEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	12192	12680	13120	100
	50	20790	18700	3800	100
	100	27560	26300	13000	100
$f_{15}$	20	52149,33	37420	57600	100
	50	59463,33	46550	39500	100
	100	97506,66	82300	141200	100
$f_{16}$	20	2106,66	1560	800	100
	50	2706,66	2600	2300	100
	100	2966,66	2700	2600	100
$f_{17}$	20	13092	9120	26960	100
	50	14406,66	9700	14500	100
	100	15546,66	13500	1000	100
$f_{18}$	20	19033,33	12760	120	100
	50	37043,33	30900	42500	100
	100	21946,66	18600	400	100
$f_{19}$	20	1588	1420	1360	100
	50	2160	1900	1300	100
	100	2840	3000	1800	100
$f_{20}$	20	10274,66	9480	14960	100
	50	21503,33	22500	22500	100
	100	30433,33	31500	19000	100
$f_{21}$	20	1121990,66	1058340	2000000	63,33
	50	1588820	1779050	751300	96,66
	100	2865073,33	2478300	4991400	96,66
$f_{22}$	20	786770,66	664880	2000000	86,66
	50	1333833,33	932700	100600	100
	100	1086773,33	770100	261600	100
$f_{23}$	20	606509,33	402700	1262200	96,66
	50	760333,33	521550	1775700	100
	100	945200	707700	1476200	100

**Tabela A.63** Número de avaliações da função (SPMEP)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	8450,4	7680	10520	100
	50	12325	10850	14800	100
	100	15694	15000	10800	100

**Tabela A.64** Número de avaliações da função (SPMEP)

### A.2.5 Algoritmo PSO

As tabelas A.65, A.66, A.67 e A.68 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.69, A.70, A.71 e A.72 mostram as faixas dos resultados obtidos; as tabelas A.77, A.78, A.79 e A.80 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.73, A.74, A.75 e A.76 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o algoritmo PSO.

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	4800,5886	1159,2403	4884,2539
	50	2542,3720	926,2590	2217,9781
	100	1585,4484	571,9671	1428,3820
$f_2$	20	32,1989	5,9909	31,7894
	50	24,5347	5,1144	22,8713
	100	20,0254	3,8711	19,9038
$f_3$	20	1228339,8446	432918,2463	1092038,5264
	50	791364,4626	407829,1950	667350,0975
	100	360430,1678	183569,9165	273563,0512
$f_4$	20	23,0032	3,0581	23,0347
	50	20,0656	3,7964	18,9517
	100	15,8389	2,8472	15,6891
$f_5$	20	277,0581	72,2573	277,1713
	50	143,2292	33,5102	141,4566
	100	129,9445	50,2110	119,6252
$f_6$	20	27,5742	1,6751	29
	50	28,9001	0,5000	29
	100	28,9667	0,2577	29
$f_7$	20	0,8086	0,4597	0,6253
	50	0,2743	0,1979	0,1950
	100	0,1479	0,1394	0,0855

**Tabela A.65** Média, desvio médio e mediana dos valores mínimos (PSO)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-17008,5966	3097,8216	-15695,8164
	50	-16949,4957	3368,2168	-15068,7966
	100	-15931,7941	2398,2516	-15039,6060
$f_9$	20	167,4201	19,3534	161,2381
	50	130,1778	20,0383	131,6258
	100	116,3684	15,9510	115,5775
$f_{10}$	20	13,1784	1,1615	13,2858
	50	11,5211	1,6328	11,0666
	100	10,9572	2,5298	9,7627
$f_{11}$	20	57,4452	15,6720	55,3680
	50	23,1885	8,0612	22,3225
	100	18,1383	7,1415	14,9742
$f_{12}$	20	39203,2194	47975,7125	4117,4837
	50	2659,7743	4533,1874	26,1096
	100	3258,2922	6030,0574	12,9500
$f_{13}$	20	1978299,0819	1696442,5980	1201588,4180
	50	334034,2819	306451,5590	203747,8183
	100	146883,3263	184004,9518	25720,7428

**Tabela A.66** Média, desvio médio e mediana dos valores mínimos (PSO)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	1,0010	0,0026	0,9996
	50	1,0006	0,0028	0,9988
	100	0,9999	0,0021	0,9986
$f_{15}$	20	0,0021	0,0016	0,0010
	50	0,0056	0,0074	0,0017
	100	0,0024	0,0017	0,0017
$f_{16}$	20	-1,0251	0,0033	-1,0243
	50	-1,0255	0,0028	-1,0245
	100	-1,0256	0,0028	-1,0261
$f_{17}$	20	0,3988	0,0005	0,3988
	50	0,3989	0,0005	0,3990
	100	0,3989	0,0005	0,3989
$f_{18}$	20	3,0048	0,0022	3,0043
	50	3,0043	0,0025	3,0037
	100	3,0051	0,0024	3,0058
$f_{19}$	20	-3,8274	0,0149	-3,8276
	50	-3,8254	0,0158	-3,8222
	100	-3,8258	0,0136	-3,8302
$f_{20}$	20	-3,2122	0,0131	-3,2011
	50	-3,2169	0,0195	-3,2022
	100	-3,2151	0,0196	-3,2010
$f_{21}$	20	-8,6186	1,7391	-9,6303
	50	-8,6305	1,8283	-9,6709
	100	-8,4603	1,8967	-9,6373
$f_{22}$	20	-8,6020	1,8185	-9,6293
	50	-9,1287	1,2324	-9,7488
	100	-9,6583	0,3574	-9,7707
$f_{23}$	20	-8,2623	2,4865	-9,8484
	50	-9,2195	1,0361	-9,6722
	100	-9,3426	0,9870	-9,7957

**Tabela A.67** Média, desvio médio e mediana dos valores mínimos (PSO)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	3,33E49	4,44E49	715,6325
	50	1,33E49	2,31E49	691,0728
	100	682,0361	22,1945	684,7302

**Tabela A.68** Média, desvio médio e mediana dos valores mínimos (PSO)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	4800,5886	2554,5494	8086,6544
	50	2542,3720	991,4660	7795,8265
	100	1585,4484	383,6730	4899,9604
$f_2$	20	32,1989	17,2155	48,0208
	50	24,5347	13,1505	39,3872
	100	20,0254	8,2720	34,5259
$f_3$	20	1228339,8446	480995,8698	2540519,5944
	50	791364,4626	192591,5256	2157792,0020
	100	360430,1678	99540,9108	990570,8310
$f_4$	20	23,0032	12,7545	33,1182
	50	20,0656	11,3029	32,2729
	100	15,8389	9,8894	25,6987
$f_5$	20	277,0581	111,7196	530,3116
	50	143,2292	71,0037	243,7155
	100	129,9445	42,5821	298,1300
$f_6$	20	27,5742	23,0000	30
	50	28,9001	27	30
	100	28,9667	28	30
$f_7$	20	0,8086	0,1522	2,3882
	50	0,2743	0,0423	1,0415
	100	0,1479	0,0009	0,9264

**Tabela A.69** Faixa dos valores mínimos encontrados (PSO)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-17008,5966	-33176,7653	-12630,5209
	50	-16949,4957	-26896,0908	-12886,0035
	100	-15931,7941	-24937,7130	-12626,0250
$f_9$	20	167,4201	131,3831	217,1409
	50	130,1778	86,4192	187,7726
	100	116,3684	65,8134	174,0348
$f_{10}$	20	13,1784	10,8981	17,1883
	50	11,5211	8,2275	20,0000
	100	10,9572	5,9293	20,9917
$f_{11}$	20	57,4452	20,1093	101,3341
	50	23,1885	5,9935	48,4868
	100	18,1383	7,8028	74,7079
$f_{12}$	20	39203,2194	22,9478	201925,3503
	50	2659,7743	3,5878	53321,9066
	100	3258,2922	4,2748	71345,7953
$f_{13}$	20	1978299,0819	70540,5400	11928473,5709
	50	334034,2819	184,4659	1572972,9078
	100	146883,3263	14,3334	838219,4235

**Tabela A.70** Faixa dos valores mínimos encontrados (PSO)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	1,0010	0,9980	1,0085
	50	1,0006	0,9980	1,0099
	100	0,9999	0,9980	1,0095
$f_{15}$	20	0,0021	0,0009	0,0203
	50	0,0056	0,0007	0,1165
	100	0,0024	0,0006	0,0128
$f_{16}$	20	-1,0251	-1,0315	-1,0200
	50	-1,0255	-1,0311	-1,0203
	100	-1,0256	-1,0308	-1,0200
$f_{17}$	20	0,3988	0,3979	0,3999
	50	0,3989	0,3979	0,3999
	100	0,3989	0,3979	0,3999
$f_{18}$	20	3,0048	3,0001	3,0098
	50	3,0043	3,0000	3,0099
	100	3,0051	3,0002	3,0098
$f_{19}$	20	-3,8274	-3,8577	-3,8002
	50	-3,8254	-3,8501	-3,8003
	100	-3,8258	-3,8538	-3,8020
$f_{20}$	20	-3,2122	-3,2597	-3,1960
	50	-3,2169	-3,2909	-3,2000
	100	-3,2151	-3,2892	-3,2000
$f_{21}$	20	-8,6186	-10,0861	-2,6304
	50	-8,6305	-10,0554	-2,6304
	100	-8,4603	-9,9920	-2,6304
$f_{22}$	20	-8,6020	-10,0356	-2,7519
	50	-9,1287	-10,3147	-2,7658
	100	-9,6583	-10,2971	-5,1288
$f_{23}$	20	-8,2623	-10,3279	-1,8594
	50	-9,2195	-10,2973	-1,8594
	100	-9,3426	-10,4544	-2,8711

**Tabela A.71** Faixa dos valores mínimos encontrados (PSO)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	3,33E49	660,5888	1E50
	50	1,33E49	589,3605	1E50
	100	682,0361	623,5146	713,6958

**Tabela A.72** Faixa dos valores mínimos encontrados (PSO)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_2$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_3$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_4$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_5$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_6$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_7$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	48355,86	50000	50000	3,33

Tabela A.73 Número de iterações para minimização (PSO)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	10,23	10	10	100
	50	7,6	7,5	8	100
	100	6,93	6	6	100
$f_9$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{10}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{11}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{12}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0
$f_{13}$	20	50000	50000	50000	0
	50	50000	50000	50000	0
	100	50000	50000	50000	0

Tabela A.74 Número de iterações para minimização (PSO)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	1753,2	1506,5	2087	100
	50	898,9	895	3	100
	100	833,2	572	1959	100
$f_{15}$	20	36325,5	50000	50000	30
	50	40309,8	50000	50000	20
	100	40053,73	50000	50000	20
$f_{16}$	20	89,83	35,5	4	100
	50	44,7	23	3	100
	100	16,76	7,5	3	100
$f_{17}$	20	628,63	307,5	1003	100
	50	496,2	52	10	100
	100	457,8	32	12	100
$f_{18}$	20	217,8	129,5	68	100
	50	100,53	55	27	100
	100	129,9	60,5	8	100
$f_{19}$	20	22,66	15	1	100
	50	10	7	3	100
	100	8,8	4	2	100
$f_{20}$	20	2285,43	667	24	96,66
	50	317,46	197	392	100
	100	268	300	182	100
$f_{21}$	20	10667,8	810	50000	80
	50	9196,86	1076	50000	83,33
	100	12304,3	903	50000	76,66
$f_{22}$	20	8735,3	406	50000	83,33
	50	7386,3	964	50000	86,66
	100	2045,13	248,5	1198	96,66
$f_{23}$	20	13695,96	647	50000	73,33
	50	5741,36	852,5	50000	90
	100	5548,9	441	50000	90

**Tabela A.75** Número de iterações para minimização (PSO)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	17052,7	718,5	50000	66,66
	50	7162,4	243,5	50000	86,66
	100	287,53	207,5	208	100

**Tabela A.76** Número de iterações para minimização (PSO)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_2$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_3$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_4$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_5$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_6$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_7$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	4835586,66	5000000	5000000	3,33

**Tabela A.77** Número de avaliações da função (PSO)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	204,66	200	200	100
	50	380	375	400	100
	100	693,33	600	600	100
$f_9$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_{10}$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_{11}$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_{12}$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0
$f_{13}$	20	1000000	1000000	1000000	0
	50	2500000	2500000	2500000	0
	100	5000000	5000000	5000000	0

**Tabela A.78** Número de avaliações da função (PSO)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	35064	30130	41740	100
	50	44945	44750	150	100
	100	83320	57200	195900	100
$f_{15}$	20	726510	1000000	1000000	30
	50	2015490	2500000	2500000	20
	100	4005373,33	5000000	5000000	20
$f_{16}$	20	1796,66	710	80	100
	50	2235	1150	150	100
	100	1676,66	750	300	100
$f_{17}$	20	12572,66	6150	20060	100
	50	24810	2600	500	100
	100	45780	3200	1200	100
$f_{18}$	20	4356	2590	1360	100
	50	5026,66	2750	1350	100
	100	12990	6050	800	100
$f_{19}$	20	453,33	300	20	100
	50	500	350	150	100
	100	880	400	200	100
$f_{20}$	20	45708,66	13340	480	96,66
	50	15873,33	9850	19600	100
	100	26800	30000	18200	100
$f_{21}$	20	213356	16200	1000000	80
	50	459843,33	53800	2500000	83,33
	100	1230430	90300	5000000	76,66
$f_{22}$	20	174706	8120	1000000	83,33
	50	369315	48200	2500000	86,66
	100	204513,33	24850	119800	96,66
$f_{23}$	20	273919,33	12940	1000000	73,33
	50	287068,33	42625	2500000	90
	100	554890	44100	5000000	90

**Tabela A.79** Número de avaliações da função (PSO)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	341054	14370	1000000	66,66
	50	358120	12175	2500000	86,66
	100	28753,33	20750	20800	100

**Tabela A.80** Número de avaliações da função (PSO)

### A.2.6 Método Dialético Canônico com Passos Variáveis

As tabelas A.81, A.82, A.83 e A.84 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.85, A.86, A.87 e A.88 mostram as faixas dos resultados obtidos; as tabelas A.93, A.94, A.95 e A.96 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.89, A.90, A.91 e A.92 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o método dialético canônico com passos variáveis (ODM-CAN-V).

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	0,0008	8,90E-5	0,0009
	50	0,0008	9,91E-5	0,0009
	100	0,0009	7,66E-5	0,0009
$f_2$	20	0,0035	0,0045	0,0009
	50	0,0023	0,0027	0,0009
	100	0,0009	3,79E-5	0,0009
$f_3$	20	0,0019	0,0020	0,0009
	50	0,0008	7,37E-5	0,0009
	100	0,0009	6,87E-5	0,0009
$f_4$	20	0,0009	5,15E-5	0,0009
	50	0,0009	5,99E-5	0,0009
	100	0,0009	5,95E-5	0,0009
$f_5$	20	0,0008	0,0001	0,0009
	50	0,0008	0,0001	0,0009
	100	0,0008	0,0001	0,0009
$f_6$	20	0,0008	0,0001	0,0009
	50	0,0009	8,24E-5	0,0009
	100	0,0008	0,0001	0,0009
$f_7$	20	6,90E-5	0,0001	1,46E-6
	50	2,28E-5	4,19E-5	1,21E-6
	100	1,15E-6	1,64E-7	1,17E-6

**Tabela A.81** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-14022,7736	1609,6865	-13037,2672
	50	-14037,2307	1214,0145	-13493,6480
	100	-14109,6964	1329,1921	-13381,3191
$f_9$	20	0,0010	0,0002	0,0009
	50	0,0010	0,0004	0,0009
	100	0,0008	0,0001	0,0009
$f_{10}$	20	0,0014	0,0009	0,0009
	50	0,0010	0,0002	0,0009
	100	0,0009	5,81E-5	0,0009
$f_{11}$	20	0,0008	7,18E-5	0,0009
	50	0,0008	8,95E-5	0,0009
	100	0,0009	6,55E-5	0,0009
$f_{12}$	20	0,0008	0,0001	0,0009
	50	0,0008	0,0001	0,0009
	100	0,0009	5,10E-5	0,0009
$f_{13}$	20	0,0008	0,0001	0,0009
	50	0,0008	0,0001	0,0009
	100	0,0012	0,0007	0,0009

**Tabela A.82** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	2,9443	1,7192	1,9920
	50	1,6663	0,9787	1,0086
	100	1,0834	0,1532	1,0070
$f_{15}$	20	0,0022	0,0015	0,0011
	50	0,0013	0,0003	0,0011
	100	0,0011	0,0002	0,0010
$f_{16}$	20	-1,0234	0,0032	-1,0213
	50	-1,0247	0,0029	-1,0241
	100	-1,0255	0,0034	-1,0257
$f_{17}$	20	0,3995	0,0004	0,3998
	50	0,3995	0,0004	0,3998
	100	0,3995	0,0004	0,3998
$f_{18}$	20	3,0079	0,0021	3,0092
	50	3,0082	0,0017	3,0092
	100	3,0081	0,0017	3,0090
$f_{19}$	20	-3,8045	0,0220	-3,8067
	50	-3,8170	0,0139	-3,8128
	100	-3,8186	0,0154	-3,8117
$f_{20}$	20	-3,2093	0,0218	-3,2056
	50	-3,2099	0,0128	-3,2014
	100	-3,2101	0,0128	-3,2018
$f_{21}$	20	-8,0142	2,4765	-9,5289
	50	-8,7163	1,5823	-9,5485
	100	-8,9975	1,1460	-9,5369
$f_{22}$	20	-8,1916	2,3873	-9,5508
	50	-8,5634	1,8552	-9,5312
	100	-8,8957	1,3485	-9,5489
$f_{23}$	20	-9,7220	0,2248	-9,5892
	50	-9,6631	0,2742	-9,5866
	100	-9,6598	0,2675	-9,5930

**Tabela A.83** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	867,4929	83,7013	910,3974
	50	801,7279	111,7175	799,0266
	100	765,6315	106,3743	710,6659

**Tabela A.84** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	0,0008	1,68E-6	0,0009
	50	0,0008	1,84E-5	0,0009
	100	0,0009	0,0003	0,0009
$f_2$	20	0,0035	0,0003	0,0870
	50	0,0023	0,0007	0,0861
	100	0,0009	0,0003	0,0009
$f_3$	20	0,0019	0,0002	0,0736
	50	0,0008	5,04E-6	0,0009
	100	0,0009	0,0002	0,0009
$f_4$	20	0,0009	2,90E-5	0,0009
	50	0,0009	0,0002	0,0009
	100	0,0009	8,70E-5	0,0009
$f_5$	20	0,0008	0,0001	0,0009
	50	0,0008	7,56E-5	0,0009
	100	0,0008	8,45E-6	0,0009
$f_6$	20	0,0008	0,0002	0,0009
	50	0,0009	8,49E-5	0,0009
	100	0,0008	3,01E-6	0,0009
$f_7$	20	6,90E-5	4,62E-7	0,0009
	50	2,28E-5	5,43E-7	0,0009
	100	1,15E-6	6,93E-7	1,54E-6

**Tabela A.85** Faixa dos valores mínimos encontrados (ODM-CAN-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-14022,7736	-23862,3006	-11647,6325
	50	-14037,2307	-20631,0086	-12503,9056
	100	-14109,6964	-20942,7113	-12508,9776
$f_9$	20	0,0010	1,65E-5	0,0085
	50	0,0010	4,07E-7	0,0213
	100	0,0008	1,66E-6	0,0009
$f_{10}$	20	0,0014	0,0005	0,0156
	50	0,0010	9,68E-6	0,0095
	100	0,0009	4,33E-5	0,0009
$f_{11}$	20	0,0008	7,13E-5	0,0009
	50	0,0008	6,76E-6	0,0009
	100	0,0009	2,48E-5	0,0009
$f_{12}$	20	0,0008	8,78E-5	0,0009
	50	0,0008	0,0001	0,0009
	100	0,0009	0,0002	0,0009
$f_{13}$	20	0,0008	0,0001	0,0009
	50	0,0008	0,0001	0,0009
	100	0,0012	1,61E-5	0,0145

**Tabela A.86** Faixa dos valores mínimos encontrados (ODM-CAN-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	2,9443	0,9980	10,7631
	50	1,6663	0,9980	4,9504
	100	1,0834	0,9980	4,9504
$f_{15}$	20	0,0022	0,0006	0,0323
	50	0,0013	0,0004	0,0028
	100	0,0011	0,0006	0,0024
$f_{16}$	20	-1,0234	-1,0315	-1,0200
	50	-1,0247	-1,0314	-1,0200
	100	-1,0255	-1,0314	-1,0200
$f_{17}$	20	0,3995	0,3979	0,3999
	50	0,3995	0,3979	0,3999
	100	0,3995	0,3979	0,3999
$f_{18}$	20	3,0079	3,0002	3,0099
	50	3,0082	3,0006	3,0099
	100	3,0081	3,0000	3,0099
$f_{19}$	20	-3,8045	-3,8605	-3,2970
	50	-3,8170	-3,8598	-3,8000
	100	-3,8186	-3,8615	-3,8000
$f_{20}$	20	-3,2093	-3,3013	-3,0873
	50	-3,2099	-3,2743	-3,1839
	100	-3,2101	-3,2948	-3,1993
$f_{21}$	20	-8,0142	-10,1041	-2,6304
	50	-8,7163	-10,1497	-2,6304
	100	-8,9975	-10,1519	-2,6304
$f_{22}$	20	-8,1916	-10,4021	-2,7658
	50	-8,5634	-10,4011	-2,7658
	100	-8,8957	-10,2850	-2,7658
$f_{23}$	20	-9,7220	-10,5334	-9,5002
	50	-9,6631	-10,5336	-2,8066
	100	-9,6598	-10,5301	-2,8066

**Tabela A.87** Faixa dos valores mínimos encontrados (ODM-CAN-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	867,4929	641,4760	962,3323
	50	801,7279	569,9940	962,3323
	100	765,6315	591,5393	962,3324

**Tabela A.88** Faixa dos valores mínimos encontrados (ODM-CAN-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	579,89	20	20	100
	50	686,76	9	9	100
	100	4,89	5	5	100
$f_2$	20	6022,9	26	26	88
	50	2010,55	11	11	96
	100	472,98	6	6	100
$f_3$	20	3867,67	27	27	94
	50	812,4	11	11	100
	100	15,95	6	6	100
$f_4$	20	794,18	25	25	100
	50	85,58	11	11	100
	100	121,54	6	6	100
$f_5$	20	8626,93	7863,5	8476	100
	50	9615,09	9046	10916	100
	100	10768,5	10163	2290	100
$f_6$	20	616,29	21	21	100
	50	56,92	9	9	100
	100	117,44	5	5	100
$f_7$	20	73,21	1	1	100
	50	47,29	1	1	100
	100	1	1	1	100

Tabela A.89 Número de iterações para minimização (ODM-CAN-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	992,97	2,5	1	99
	50	1,37	1	1	100
	100	1,04	1	1	100
$f_9$	20	3069,23	20	20	96
	50	736,74	8	8	99
	100	4,06	4	4	100
$f_{10}$	20	6487,92	25	25	88
	50	1380,71	10	10	98
	100	5,55	6	6	100
$f_{11}$	20	46,35	21	21	100
	50	15,69	9	9	100
	100	54,96	5	5	100
$f_{12}$	20	1151,43	1077,5	730	100
	50	1053,26	905	1402	100
	100	195,6	4	4	100
$f_{13}$	20	2907,68	2232	3804	100
	50	2973,46	2275,5	3333	100
	100	5829,11	2715,5	50000	97

Tabela A.90 Número de iterações para minimização (ODM-CAN-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	32502,17	50000	50000	35
	50	13001,97	3	3	74
	100	1501,51	2	2	97
$f_{15}$	20	33950,69	50000	50000	35
	50	32035,86	50000	50000	38
	100	29825,92	50000	50000	43
$f_{16}$	20	6,11	6	8	100
	50	2,03	2	2	100
	100	1,09	1	1	100
$f_{17}$	20	9,23	9	10	100
	50	3,54	4	4	100
	100	1,94	2	2	100
$f_{18}$	20	10,2	11	11	100
	50	3,93	4	4	100
	100	1,97	2	2	100
$f_{19}$	20	1004,54	4	4	98
	50	1,5	1	1	100
	100	1	1	1	100
$f_{20}$	20	9505,69	6	6	81
	50	3003,45	3	2	94
	100	501,82	1	1	99
$f_{21}$	20	12469,88	418	50000	77
	50	7946,05	133	4	87
	100	5729,55	2	2	91
$f_{22}$	20	13389,89	574,5	9	78
	50	9955,59	175,5	4	84
	100	6542,99	2	2	89
$f_{23}$	20	6662,72	2411	9	100
	50	4411,01	160	4	99
	100	3255,49	2	2	99

Tabela A.91 Número de iterações para minimização (ODM-CAN-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	42670,36	50000	50000	15
	50	27678,05	50000	50000	45
	100	19000,65	1	1	62

Tabela A.92 Número de iterações para minimização (ODM-CAN-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	1649,74	400	400	100
	50	2147,28	450	450	100
	100	489	500	500	100
$f_2$	20	12674	520	520	88
	50	4719,5	550	550	96
	100	1624,12	600	600	100
$f_3$	20	8373,26	540	540	94
	50	2368,28	550	550	100
	100	712,02	600	600	100
$f_4$	20	2132,14	500	500	100
	50	780,28	550	550	100
	100	1104,5	600	600	100
$f_5$	20	19053,86	17527	18752	100
	50	24030,18	22892	26632	100
	100	31337	30126	14380	100
$f_6$	20	2020,1	420	420	100
	50	683,82	450	450	100
	100	910,16	500	500	100
$f_7$	20	306,98	20	20	100
	50	285,14	50	50	100
	100	100	100	100	100

Tabela A.93 Número de avaliações da função (ODM-CAN-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	3617,12	50	20	99
	50	68,5	50	50	100
	100	104	100	100	100
$f_9$	20	6608,98	400	400	96
	50	1990,44	400	400	99
	100	406	400	400	100
$f_{10}$	20	13609,44	500	500	88
	50	3419,5	500	500	98
	100	555	600	600	100
$f_{11}$	20	637,52	420	420	100
	50	594,96	450	450	100
	100	786,12	500	500	100
$f_{12}$	20	4102,86	3955	3260	100
	50	6906,52	6610	7604	100
	100	2407,06	400	400	100
$f_{13}$	20	7615,36	6264	9408	100
	50	10746,92	9351	11466	100
	100	21458,22	15231	109800	97

Tabela A.94 Número de avaliações da função (ODM-CAN-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	71945,4	108200	160	35
	50	29710,5	150	150	74
	100	3685	200	200	97
$f_{15}$	20	69811,22	101800	101800	35
	50	68368,68	104800	104800	38
	100	66889,46	109800	109800	43
$f_{16}$	20	122,2	120	160	100
	50	101,5	100	100	100
	100	109	100	100	100
$f_{17}$	20	184,6	180	200	100
	50	177	200	200	100
	100	194	200	200	100
$f_{18}$	20	204	220	220	100
	50	196,5	200	200	100
	100	197	200	200	100
$f_{19}$	20	2130,8	80	80	98
	50	75	50	50	100
	100	100	100	100	100
$f_{20}$	20	19455,8	120	120	81
	50	6460,5	150	100	94
	100	1280	100	100	99
$f_{21}$	20	26170,78	2636	101800	77
	50	18570,98	5066	200	87
	100	16265,02	200	200	91
$f_{22}$	20	27875,62	2949	180	78
	50	22819,1	5151	200	84
	100	16547,34	200	200	89
$f_{23}$	20	14578,16	6622	180	100
	50	11412,98	5120	200	99
	100	10166,46	200	200	99

**Tabela A.95** Número de avaliações da função (ODM-CAN-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	86921,66	101800	101800	15
	50	58186,66	104800	104800	45
	100	41789	100	100	62

**Tabela A.96** Número de avaliações da função (ODM-CAN-V)

### A.2.7 Método Dialético com Máxima Entropia e Passos Variáveis

As tabelas A.97, A.98, A.99 e A.100 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.101, A.102, A.103 e A.104 mostram as faixas dos resultados obtidos; as tabelas A.109, A.110, A.111 e A.112 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.105, A.106, A.107 e A.108 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o método dialético baseado no Princípio da Máxima Entropia com passos variáveis (ODM-PME-V).

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	0,0006	0,0003	0,0007
	50	0,0004	0,0002	0,0004
	100	0,0004	0,0003	0,0004
$f_2$	20	0,0170	0,0157	0,0100
	50	0,0158	0,0172	0,0040
	100	0,0201	0,0205	0,0074
$f_3$	20	0,0008	0,0003	0,0009
	50	0,0008	0,0004	0,0009
	100	0,0007	0,0004	0,0009
$f_4$	20	0,0011	0,0006	0,0008
	50	0,0009	0,0002	0,0008
	100	0,0009	0,0002	0,0008
$f_5$	20	0,0333	0,0515	0,0008
	50	0,3531	0,5918	0,0049
	100	0,0891	0,1145	0,0068
$f_6$	20	0,0006	0,0002	0,0007
	50	0,0006	0,0002	0,0007
	100	0,0006	0,0002	0,0007
$f_7$	20	0,0003	0,0003	0,0001
	50	0,0004	0,0002	0,0005
	100	0,0007	0,0001	0,0007

**Tabela A.97** Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-14174,4209	1765,8502	-13212,0134
	50	-13787,2807	999,8902	-13340,1375
	100	-13765,5399	1139,1479	-13078,3738
$f_9$	20	0,0007	0,0003	0,0009
	50	0,0008	0,0002	0,0009
	100	0,0007	0,0002	0,0009
$f_{10}$	20	0,0011	0,0007	0,0008
	50	0,0009	0,0002	0,0008
	100	0,0010	0,0002	0,0009
$f_{11}$	20	0,0005	0,0002	0,0007
	50	0,0006	0,0002	0,0007
	100	0,0006	0,0002	0,0006
$f_{12}$	20	0,0006	0,0002	0,0007
	50	0,0006	0,0002	0,0007
	100	0,0006	0,0002	0,0007
$f_{13}$	20	0,0006	0,0002	0,0007
	50	0,0006	0,0002	0,0007
	100	0,0006	0,0002	0,0007

**Tabela A.98** Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	4,6508	2,8614	3,9682
	50	3,1008	1,6415	2,9821
	100	2,2481	0,9619	1,9920
$f_{15}$	20	0,0031	0,0023	0,0016
	50	0,0032	0,0031	0,0015
	100	0,0015	0,00078	0,0010
$f_{16}$	20	-1,0258	0,0032	-1,0250
	50	-1,0253	0,0028	-1,0253
	100	-1,0243	0,0028	-1,0233
$f_{17}$	20	0,3990	0,0005	0,3990
	50	0,3991	0,0004	0,3993
	100	0,3993	0,0004	0,3995
$f_{18}$	20	3,8154	1,5721	3,0051
	50	3,0047	0,0026	3,0043
	100	3,0050	0,0026	3,0053
$f_{19}$	20	-3,7800	0,0606	-3,8080
	50	-3,8191	0,0132	-3,8137
	100	-3,8225	0,0150	-3,8189
$f_{20}$	20	-3,1550	0,0928	-3,2051
	50	-3,2067	0,0331	-3,2079
	100	-3,2154	0,0178	-3,2067
$f_{21}$	20	-9,8180	0,1716	-9,7787
	50	-9,8155	0,1900	-9,7673
	100	-9,7879	0,2005	-9,7244
$f_{22}$	20	-9,9142	0,2559	-9,8433
	50	-9,8217	0,2066	-9,7846
	100	-9,7806	0,2091	-9,7026
$f_{23}$	20	-9,4362	0,9684	-9,9021
	50	-9,4342	0,8408	-9,7652
	100	-9,6863	0,3843	-9,6941

**Tabela A.99** Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	878,3669	91,7582	936,4728
	50	816,2446	116,2557	825,4084
	100	786,1422	114,5629	713,9279

**Tabela A.100** Média, desvio médio e mediana dos valores mínimos (ODM-PME-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	0,0006	1,06E-6	0,0009
	50	0,0004	9,14E-7	0,0009
	100	0,0004	3,36E-7	0,0009
$f_2$	20	0,0170	5,12E-5	0,0561
	50	0,0158	2,18E-5	0,0605
	100	0,0201	8,36E-5	0,0927
$f_3$	20	0,0008	2,64E-7	0,0016
	50	0,0008	1,44E-5	0,0019
	100	0,0007	7,17E-6	0,0018
$f_4$	20	0,0011	3,14E-5	0,0108
	50	0,0009	6,23E-5	0,0085
	100	0,0009	0,0001	0,0118
$f_5$	20	0,0333	5,26E-6	0,7197
	50	0,3531	2,66E-6	28,7036
	100	0,0891	2,03E-5	0,9832
$f_6$	20	0,0006	4,84E-6	0,0009
	50	0,0006	4,30E-6	0,0013
	100	0,0006	9,64E-7	0,0009
$f_7$	20	0,0003	1,67E-8	0,0009
	50	0,0004	1,40E-9	0,0009
	100	0,0007	1,09E-6	0,0009

**Tabela A.101** Faixa dos valores mínimos encontrados (ODM-PME-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-14174,4209	-65947,6391	-10673,6594
	50	-13787,2807	-17766,2372	-11787,9309
	100	-13765,5399	-18180,0563	-12506,2042
$f_9$	20	0,0007	3,53E-6	0,0016
	50	0,0008	9,97E-7	0,0014
	100	0,0007	9,68E-6	0,0014
$f_{10}$	20	0,0011	0,0003	0,0130
	50	0,0009	0,0001	0,0078
	100	0,0010	0,0001	0,0076
$f_{11}$	20	0,0005	1,73E-6	0,0009
	50	0,0006	6,73E-7	0,0009
	100	0,0006	1,28E-6	0,0009
$f_{12}$	20	0,0006	6,25E-6	0,0009
	50	0,0006	5,37E-6	0,0009
	100	0,0006	3,04E-6	0,0009
$f_{13}$	20	0,0006	1,65E-6	0,0009
	50	0,0006	3,63E-6	0,0009
	100	0,0006	1,95E-6	0,0009

**Tabela A.102** Faixa dos valores mínimos encontrados (ODM-PME-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	4,6508	0,9980	12,6705
	50	3,1008	0,9980	12,6705
	100	2,2481	0,9980	5,9288
$f_{15}$	20	0,0031	0,0004	0,0260
	50	0,0032	0,0005	0,1119
	100	0,0015	0,0005	0,0199
$f_{16}$	20	-1,0258	-1,0316	-1,0203
	50	-1,0253	-1,0315	-1,0201
	100	-1,0243	-1,0314	-1,0200
$f_{17}$	20	0,3990	0,3979	0,3999
	50	0,3991	0,3978	0,3999
	100	0,3993	0,3979	0,3999
$f_{18}$	20	3,8154	3,0001	30,0544
	50	3,0047	3,0000	3,0098
	100	3,0050	3,0000	3,0099
$f_{19}$	20	-3,7800	-3,8603	-3,2965
	50	-3,8191	-3,8552	-3,8003
	100	-3,8225	-3,8586	-3,8000
$f_{20}$	20	-3,1550	-3,2649	-1,4377
	50	-3,2067	-3,2899	-3,0194
	100	-3,2154	-3,3079	-3,1745
$f_{21}$	20	-9,8180	-10,1529	-9,5000
	50	-9,8155	-10,1531	-9,5003
	100	-9,7879	-10,1530	-9,5010
$f_{22}$	20	-9,9142	-10,4021	-9,5091
	50	-9,8217	-10,3881	-9,5015
	100	-9,7806	-10,4023	-9,5009
$f_{23}$	20	-9,4362	-10,5167	-2,4220
	50	-9,4342	-10,5354	-2,4273
	100	-9,6863	-10,5352	-2,4273

**Tabela A.103** Faixa dos valores mínimos encontrados (ODM-PME-V)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	878,3669	585,9763	1001,1928
	50	816,2446	620,7964	981,2394
	100	786,1422	630,4399	987,8290

**Tabela A.104** Faixa dos valores mínimos encontrados (ODM-PME-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	3864,36	2354	8	100
	50	2748,56	1402,5	3	100
	100	2422,17	4	2	100
$f_2$	20	31145,56	50000	50000	38
	50	27711,99	50000	50000	45
	100	36030,71	50000	50000	30
$f_3$	20	24644	22360,5	50000	58
	50	22902,54	21101,5	50000	66
	100	20783,1	5353,5	3	75
$f_4$	20	14175,89	9360,5	50000	89
	50	6822,45	3	3	95
	100	5200,19	2	2	98
$f_5$	20	31902,92	39584	50000	59
	50	38792,79	50000	50000	37
	100	40448,62	50000	50000	32
$f_6$	20	9134,19	6206	2	100
	50	9272,65	5307	2	98
	100	8296,22	6707	2	100
$f_7$	20	148,72	3	1	100
	50	6,03	1	1	100
	100	5,02	1	1	100

Tabela A.105 Número de iterações para minimização (ODM-PME-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	516,25	3	1	99
	50	501,57	1	1	99
	100	1,05	1	1	100
$f_9$	20	20533,91	11806	50000	79
	50	19988,28	8906	50000	88
	100	15427,92	4306	50000	94
$f_{10}$	20	4002,76	3	3	92
	50	2537,98	3	3	96
	100	2841,06	2	2	95
$f_{11}$	20	3202,75	2106	3	100
	50	4340,83	4054	201	100
	100	3427,85	2901,5	201	100
$f_{12}$	20	2687,43	2003	1603	100
	50	2895,67	2354	2304	100
	100	2563,49	2104,5	2203	100
$f_{13}$	20	4794,34	4402	2602	100
	50	3971,54	2508	1309	100
	100	3705,28	1707,5	506	100

Tabela A.106 Número de iterações para minimização (ODM-PME-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	38019,13	50000	50000	24
	50	36011,24	50000	50000	28
	100	34005,45	50000	50000	32
$f_{15}$	20	43249,71	50000	50000	18
	50	39931,71	50000	50000	24
	100	32561,83	50000	50000	39
$f_{16}$	20	491,41	1	1	100
	50	1	1	1	100
	100	1	1	1	100
$f_{17}$	20	338,77	1	1	100
	50	1	1	1	100
	100	1	1	1	100
$f_{18}$	20	1552,06	4	4	97
	50	1,08	1	1	100
	100	1	1	1	100
$f_{19}$	20	11500,8	1	1	77
	50	1	1	1	100
	100	1	1	1	100
$f_{20}$	20	19012,71	1	1	62
	50	18500,64	1	1	63
	100	7000,92	1	1	86
$f_{21}$	20	1735,69	101	1	100
	50	2656,8	201,5	1	100
	100	2726,55	1	1	100
$f_{22}$	20	2124,12	101,5	1	100
	50	1816,81	1	1	100
	100	1121,98	1	1	100
$f_{23}$	20	8564,95	751	1	93
	50	6122,87	1	1	94
	100	3838,96	1	1	98

**Tabela A.107** Número de iterações para minimização (ODM-PME-V)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	43140,58	50000	50000	14
	50	30000,71	50000	50000	40
	100	23228,29	1	1	54

**Tabela A.108** Número de iterações para minimização (ODM-PME-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	9186,54	6508	160	100
	50	8527,84	7605	150	100
	100	9278,84	400	200	100
$f_2$	20	63499,64	101800	101800	38
	50	58195,02	104800	104800	45
	100	79243,84	109800	109800	30
$f_3$	20	50761,52	46521	101800	58
	50	49289,88	47003	104800	66
	100	48328,2	20507	300	75
$f_4$	20	30099,22	20521	101800	89
	50	15744,42	150	150	95
	100	14341,94	200	200	98
$f_5$	20	65605,84	80968	101800	59
	50	82385,58	104800	104800	37
	100	90697,24	109800	109800	32
$f_6$	20	19980,36	14212	40	100
	50	23063,06	15414	100	98
	100	26104,32	23214	200	100
$f_7$	20	1071,62	60	20	100
	50	250,62	50	50	100
	100	399,1	100	100	100

Tabela A.109 Número de avaliações da função (ODM-PME-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	2185,56	60	20	99
	50	2122,5	50	50	99
	100	105	100	100	100
$f_9$	20	42717,88	25412	101800	79
	50	44729,52	22612	104800	88
	100	40463,76	18412	109800	94
$f_{10}$	20	8199,2	60	60	92
	50	5499,32	150	150	96
	100	6553,34	200	200	95
$f_{11}$	20	7838,66	6012	60	100
	50	13340,54	12908	5202	100
	100	16558,68	15603	10202	100
$f_{12}$	20	7174,86	5806	5006	100
	50	10591,34	9508	9408	100
	100	14926,98	14009	14206	100
$f_{13}$	20	11388,68	10604	7004	100
	50	12743,08	9816	7418	100
	100	17210,56	13215	10812	100

Tabela A.110 Número de avaliações da função (ODM-PME-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	78913,72	102800	103000	24
	50	76871,68	105900	50	28
	100	75873,28	110800	100	32
$f_{15}$	20	88283,96	101800	101800	18
	50	84220,7	104800	104800	24
	100	72718,2	109800	109800	39
$f_{16}$	20	1043,48	20	20	100
	50	50	50	50	100
	100	100	100	100	100
$f_{17}$	20	845,26	20	20	100
	50	50	50	50	100
	100	100	100	100	100
$f_{18}$	20	3249,02	80	80	97
	50	54	50	50	100
	100	100	100	100	100
$f_{19}$	20	23430	20	20	77
	50	48,5	50	50	100
	100	100	100	100	100
$f_{20}$	20	38738,58	20	20	62
	50	38808	50	50	63
	100	15464	100	100	86
$f_{21}$	20	4398,56	2002	20	100
	50	8070,24	5203	50	100
	100	9334,88	100	100	100
$f_{22}$	20	5192,88	2003	20	100
	50	5534,9	50	50	100
	100	5446,6	100	100	100
$f_{23}$	20	18092,36	3302	20	93
	50	14337,1	50	50	94
	100	11171,62	100	100	98

Tabela A.111 Número de avaliações da função (ODM-PME-V)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	87875,06	101800	101800	14
	50	62915,5	104800	104800	40
	100	51115,5	100	100	54

Tabela A.112 Número de avaliações da função (ODM-PME-V)

### A.2.8 Método Dialético Canônico com Passos Fixos

As tabelas A.113, A.114, A.115 e A.116 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.117, A.118, A.119 e A.120 mostram as faixas dos resultados obtidos; as tabelas A.125, A.126, A.127 e A.128 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.121, A.122, A.123 e A.124 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o método dialético canônico com passos fixos (ODM-CAN-F).

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	0,0005	0,0003	0,0006
	50	5,53E-5	9,15E-5	6,25E-6
	100	0,0007	5,40E-5	0,0007
$f_2$	20	0,0235	0,0146	0,0233
	50	0,0026	0,0047	0,0001
	100	0,0020	0,0029	0,0004
$f_3$	20	0,0064	0,0088	0,0008
	50	0,0032	0,0064	1,63E-5
	100	0,0002	3,41E-5	0,0002
$f_4$	20	0,0012	0,0010	0,0007
	50	0,0008	2,52E-5	0,0008
	100	0,0003	2,26E-5	0,0003
$f_5$	20	0,0961	0,1229	0,0177
	50	0,0580	0,0679	0,0182
	100	0,0525	0,0632	0,0105
$f_6$	20	0,0005	0,0002	0,0005
	50	5,71E-5	9,12E-5	7,10E-6
	100	0,0004	0,0003	0,0004
$f_7$	20	3,53E-5	6,24E-5	1,49E-6
	50	2,03E-5	3,70E-5	1,22E-6
	100	1,11E-6	1,64E-7	1,11E-6

**Tabela A.113** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-13648,0752	1485,0847	-13350,8757
	50	-14011,6573	1220,6374	-13475,2850
	100	-13722,7083	1061,2044	-13287,1638
$f_9$	20	0,0031	0,0030	0,0014
	50	0,0001	0,0002	3,53E-6
	100	0,0005	0,0003	0,0004
$f_{10}$	20	0,0039	0,0032	0,0021
	50	0,0005	0,0001	0,0004
	100	0,0002	0,0001	0,0001
$f_{11}$	20	0,0005	0,0002	0,0005
	50	4,16E-5	4,98E-5	1,56E-5
	100	0,0002	0,0003	3,01E-6
$f_{12}$	20	0,0006	0,0002	0,0007
	50	0,0006	0,0002	0,0006
	100	0,0006	0,0002	0,0007
$f_{13}$	20	0,0007	0,0002	0,0009
	50	0,0007	0,0002	0,0008
	100	0,0007	0,0002	0,0008

**Tabela A.114** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	5,1110	2,9368	3,9682
	50	2,1468	1,1887	1,9920
	100	1,0405	0,0775	0,9991
$f_{15}$	20	0,0032	0,0027	0,0016
	50	0,0032	0,0028	0,0014
	100	0,0017	0,0009	0,0011
$f_{16}$	20	-1,0274	0,0029	-1,0285
	50	-1,0271	0,0028	-1,0281
	100	-1,0254	0,0031	-1,0255
$f_{17}$	20	0,3988	0,0005	0,3987
	50	0,3987	0,0005	0,3986
	100	0,3986	0,0004	0,3986
$f_{18}$	20	5,7033	4,9674	3,0038
	50	3,0040	0,0025	3,0038
	100	3,0036	0,0023	3,0031
$f_{19}$	20	-3,7729	0,0684	-3,8102
	50	-3,8106	0,0203	-3,8098
	100	-3,8168	0,0145	-3,8091
$f_{20}$	20	-2,9227	0,2455	-3,0384
	50	-3,1725	0,0687	-3,2024
	100	-3,1882	0,0383	-3,2023
$f_{21}$	20	-8,3770	2,2233	-9,6364
	50	-8,7252	1,9503	-9,8670
	100	-9,0901	1,4211	-9,8781
$f_{22}$	20	-8,6483	2,0369	-9,7332
	50	-8,9892	1,7425	-9,9179
	100	-9,2788	1,3025	-9,9807
$f_{23}$	20	-9,9314	0,3882	-9,9976
	50	-9,9883	0,3533	-10,0666
	100	-10,1252	0,2295	-10,1152

**Tabela A.115** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	887,9872	68,0259	923,4461
	50	839,3678	103,3755	853,1816
	100	773,6154	114,7776	710,7234

**Tabela A.116** Média, desvio médio e mediana dos valores mínimos (ODM-CAN-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	0,0005	9,93E-6	0,0009
	50	5,53E-5	4,70E-6	0,0009
	100	0,0007	0,0004	0,0008
$f_2$	20	0,0235	0,0005	0,0703
	50	0,0026	9,89E-5	0,0844
	100	0,0020	0,0004	0,0835
$f_3$	20	0,0064	2,61E-6	0,1336
	50	0,0032	1,17E-5	0,3264
	100	0,0002	0,0001	0,0003
$f_4$	20	0,0012	7,11E-5	0,0193
	50	0,0008	0,0005	0,0008
	100	0,0003	0,0002	0,0009
$f_5$	20	0,0961	5,54E-5	1,7195
	50	0,0580	0,0001	0,9586
	100	0,0525	7,63E-5	0,8950
$f_6$	20	0,0005	2,56E-5	0,0012
	50	5,71E-5	1,07E-10	0,0009
	100	0,0004	2,28E-9	0,0009
$f_7$	20	3,53E-5	5,20E-7	0,0006
	50	2,03E-5	6,60E-7	0,0009
	100	1,11E-6	6,89E-7	1,70E-6

**Tabela A.117** Faixa dos valores mínimos encontrados (ODM-CAN-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-13648,0752	-19647,7974	-8282,0201
	50	-14011,6573	-21233,7101	-10530,7047
	100	-13722,7083	-19925,2094	-11298,7175
$f_9$	20	0,0031	2,41E-5	0,0203
	50	0,0001	2,27E-6	0,0119
	100	0,0005	0,0001	0,0137
$f_{10}$	20	0,0039	0,0003	0,0185
	50	0,0005	3,24E-5	0,0035
	100	0,0002	2,11E-6	0,0053
$f_{11}$	20	0,0005	1,98E-5	0,0009
	50	4,16E-5	8,48E-6	0,0008
	100	0,0002	1,53E-6	0,0009
$f_{12}$	20	0,0006	6,05E-5	0,0009
	50	0,0006	1,31E-6	0,0009
	100	0,0006	5,92E-6	0,0009
$f_{13}$	20	0,0007	8,97E-6	0,0009
	50	0,0007	1,84E-5	0,0009
	100	0,0007	4,58E-5	0,0009

**Tabela A.118** Faixa dos valores mínimos encontrados (ODM-CAN-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	5,1110	0,9980	12,6705
	50	2,1468	0,9980	9,8038
	100	1,0405	0,9980	3,9682
$f_{15}$	20	0,0032	0,0004	0,0359
	50	0,0032	0,0003	0,0341
	100	0,0017	0,0003	0,0275
$f_{16}$	20	-1,0274	-1,0316	-1,0207
	50	-1,0271	-1,0315	-1,0201
	100	-1,0254	-1,0315	-1,0200
$f_{17}$	20	0,3988	0,3978	0,3999
	50	0,3987	0,3978	0,3999
	100	0,3986	0,3978	0,3999
$f_{18}$	20	5,7033	3,0000	84,0000
	50	3,0040	3,0000	3,0097
	100	3,0036	3,0000	3,0097
$f_{19}$	20	-3,7729	-3,8576	-3,2955
	50	-3,8106	-3,8596	-3,6012
	100	-3,8168	-3,8609	-3,8000
$f_{20}$	20	-2,9227	-3,2629	-1,0871
	50	-3,1725	-3,2909	-2,7411
	100	-3,1882	-3,3061	-2,9362
$f_{21}$	20	-8,3770	-10,1466	-2,1351
	50	-8,7252	-10,1528	-2,6299
	100	-9,0901	-10,1531	-2,6304
$f_{22}$	20	-8,6483	-10,3980	-2,2504
	50	-8,9892	-10,4014	-2,7650
	100	-9,2788	-10,4016	-2,7658
$f_{23}$	20	-9,9314	-10,5350	-2,4129
	50	-9,9883	-10,5342	-2,4273
	100	-10,1252	-10,5362	-9,5000

**Tabela A.119** Faixa dos valores mínimos encontrados (ODM-CAN-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	887,9872	606,2376	968,8115
	50	839,3678	577,7530	1007,4225
	100	773,6154	583,1522	985,3457

**Tabela A.120** Faixa dos valores mínimos encontrados (ODM-CAN-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	10874,75	10552	13902	100
	50	784,1	2	2	100
	100	2	2	2	100
$f_2$	20	47441,53	50000	50000	8
	50	2501,9	2	2	95
	100	3829,94	2	2	93
$f_3$	20	29771,25	33752,5	50000	63
	50	839	2	2	99
	100	91	2	2	100
$f_4$	20	20425,89	16203	50000	80
	50	4,04	2	2	100
	100	12,1	2	2	100
$f_5$	20	41106,96	50000	50000	32
	50	43175,42	50000	50000	21
	100	42176,56	50000	50000	26
$f_6$	20	7096,94	3303	3303	99
	50	752,21	2	2	100
	100	3607,77	802,5	2	100
$f_7$	20	45,14	1	1	100
	50	23,08	1	1	100
	100	1	1	1	100

Tabela A.121 Número de iterações para minimização (ODM-CAN-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	8544,27	2	2	83
	50	1001,27	1	1	98
	100	1500,98	1	1	97
$f_9$	20	35598,1	50000	50000	47
	50	1660,11	2	2	99
	100	3068,14	2	2	98
$f_{10}$	20	41202,81	50000	50000	31
	50	1001,96	2	2	98
	100	1001,96	2	2	98
$f_{11}$	20	2777,47	1302,5	204	100
	50	80,01	2	2	100
	100	648,09	2	2	100
$f_{12}$	20	2772,61	2205,5	204	100
	50	3464,3	2903,5	2002	100
	100	3362,48	2554,5	704	100
$f_{13}$	20	6472,56	5904,5	404	100
	50	6416,53	5753,5	5505	100
	100	7094,57	6656,5	204	100

Tabela A.122 Número de iterações para minimização (ODM-CAN-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	42534,47	50000	50000	15
	50	25500,91	50000	50000	49
	100	1001,58	2	2	98
$f_{15}$	20	41062,87	50000	50000	25
	50	44195,55	50000	50000	16
	100	34483,28	50000	50000	40
$f_{16}$	20	425,07	2	2	100
	50	1,83	2	2	100
	100	1,07	1	1	100
$f_{17}$	20	1345,38	2	2	100
	50	1,97	2	2	100
	100	1,93	2	2	100
$f_{18}$	20	4210,87	2	2	92
	50	1,99	2	2	100
	100	1,96	2	2	100
$f_{19}$	20	20001,11	2	2	60
	50	4001,18	1	1	92
	100	1	1	1	100
$f_{20}$	20	44074,32	50000	50000	12
	50	22501,1	2	2	55
	100	17000,7	1	1	66
$f_{21}$	20	12971,72	4052,5	50000	81
	50	9872,8	103	2	84
	100	7790,86	104	2	89
$f_{22}$	20	13311,42	3553,5	50000	83
	50	9621,77	252	2	86
	100	5894,29	2	2	90
$f_{23}$	20	8257,87	6352,5	2	99
	50	7432,55	954	2	99
	100	3723,36	2	2	100

**Tabela A.123** Número de iterações para minimização (ODM-CAN-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	47159,18	50000	50000	6
	50	37346,51	50000	50000	26
	100	19128,71	1	1	62

**Tabela A.124** Número de iterações para minimização (ODM-CAN-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	23549,5	22904	29604	100
	50	1993,48	100	100	100
	100	200	200	200	100
$f_2$	20	96683,06	101800	101800	8
	50	5335	100	100	95
	100	8624,2	200	200	93
$f_3$	20	61342,5	69305	101800	63
	50	1915,12	100	100	99
	100	474,04	200	200	100
$f_4$	20	42651,78	34206	101800	80
	50	151,12	100	100	100
	100	508,32	200	200	100
$f_5$	20	84029,98	101800	101800	32
	50	91158,84	104800	104800	21
	100	94175,12	109800	109800	26
$f_6$	20	15993,88	8406	8406	99
	50	1976,74	100	100	100
	100	13269,98	11405	200	100
$f_7$	20	250,84	20	20	100
	50	236,72	50	50	100
	100	100	100	100	100

Tabela A.125 Número de avaliações da função (ODM-CAN-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	17488,14	40	40	83
	50	2159,5	50	50	98
	100	3392	100	100	97
$f_9$	20	72872,72	101800	101800	47
	50	3698,46	100	100	99
	100	7292,68	200	200	98
$f_{10}$	20	84205,62	101800	101800	31
	50	2194	100	100	98
	100	2392	200	200	98
$f_{11}$	20	7354,94	4405	2208	100
	50	491,22	100	100	100
	100	3797,14	200	200	100
$f_{12}$	20	7345,22	6211	2208	100
	50	11728,6	10607	8804	100
	100	16526,96	14909	11208	100
$f_{13}$	20	14745,12	13609	2608	100
	50	17633,06	16307	15810	100
	100	23989,14	23113	10208	100

Tabela A.126 Número de avaliações da função (ODM-CAN-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	87778	102200	101800	15
	50	53631,5	104800	100	49
	100	2354	200	200	98
$f_{15}$	20	84220,18	101800	101800	25
	50	94018,36	104800	104800	16
	100	77643,1	109800	109800	40
$f_{16}$	20	938,88	40	40	100
	50	91,5	100	100	100
	100	107	100	100	100
$f_{17}$	20	2867,7	40	40	100
	50	98,5	100	100	100
	100	193	200	200	100
$f_{18}$	20	8616,5	40	40	92
	50	99,5	100	100	100
	100	196	200	200	100
$f_{19}$	20	40742,2	40	40	60
	50	8443	50	50	92
	100	100	100	100	100
$f_{20}$	20	89772,24	101800	101800	12
	50	47215	100	100	55
	100	37402	100	100	66
$f_{21}$	20	27443,56	9905	101800	81
	50	22393,76	5006	100	84
	100	21068	10008	200	89
$f_{22}$	20	28162,46	8907	101800	83
	50	22024,82	5304	100	86
	100	14869,78	200	200	90
$f_{23}$	20	18196,7	14505	40	99
	50	17566,3	6708	100	99
	100	11688,4	200	200	100

**Tabela A.127** Número de avaliações da função (ODM-CAN-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	96065,26	101800	101800	6
	50	78361,66	104800	104800	26
	100	42240,14	100	100	62

**Tabela A.128** Número de avaliações da função (ODM-CAN-F)

### A.2.9 Método Dialético com Máxima Entropia e Passos Fixos

As tabelas A.129, A.130, A.131 e A.132 mostram os resultados obtidos quanto aos valores mínimos obtidos; as tabelas A.133, A.134, A.135 e A.136 mostram as faixas dos resultados obtidos; as tabelas A.141, A.142, A.143 e A.144 exibem as estatísticas do número de avaliações da função objetivo para minimização; enquanto as tabelas A.137, A.138, A.139 e A.140 exibem os resultados quanto ao número de iterações dispendidas até se alcançar os valores limite  $f_{lim}$ , definidos segundo a tabela 5.5, para funções unimodais de alta dimensionalidade, multimodais de alta dimensionalidade e muitos mínimos locais, multimodais de baixa dimensionalidade e poucos mínimos locais e para uma função que modela um problema multi-objetivo, respectivamente, usando o método dialético otimizado pelo Princípio da Máxima Entropia com passos fixos (ODM-PME-F).

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_1$	20	0,0006	0,0003	0,0006
	50	0,0006	0,0002	0,0007
	100	0,0006	0,0002	0,0007
$f_2$	20	0,0224	0,0115	0,0225
	50	0,0147	0,0145	0,0077
	100	0,0198	0,0171	0,0204
$f_3$	20	0,0009	0,0002	0,0009
	50	0,0009	0,0002	0,0009
	100	0,0009	0,0002	0,0009
$f_4$	20	0,0013	0,0010	0,0008
	50	0,0008	0,0002	0,0008
	100	0,0008	0,0001	0,0007
$f_5$	20	0,0449	0,0643	0,0023
	50	0,3313	0,6020	0,0051
	100	0,0615	0,0901	0,0026
$f_6$	20	0,0006	0,0002	0,0007
	50	0,0006	0,0002	0,0007
	100	0,0006	0,0002	0,0007
$f_7$	20	0,0002	0,0002	5,40E-5
	50	0,0004	0,0002	0,0005
	100	0,0007	0,0001	0,0007

**Tabela A.129** Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_8$	20	-13962,7099	1768,7182	-13145,2049
	50	-13900,4271	1048,9225	-13518,2708
	100	-13898,5910	1058,7716	-13653,0867
$f_9$	20	0,0007	0,0002	0,0008
	50	0,0008	0,0002	0,0009
	100	0,0008	0,0002	0,0009
$f_{10}$	20	0,0049	0,0037	0,0035
	50	0,0044	0,0033	0,0027
	100	0,0050	0,0040	0,0030
$f_{11}$	20	0,0007	0,0002	0,0008
	50	0,0006	0,0002	0,0007
	100	0,0006	0,0002	0,0007
$f_{12}$	20	0,0007	0,0002	0,0008
	50	0,0007	0,0001	0,0008
	100	0,0007	0,0002	0,0008
$f_{13}$	20	0,0007	0,0002	0,0008
	50	0,0007	0,0002	0,0008
	100	0,0006	0,0002	0,0006

**Tabela A.130** Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{14}$	20	5,0198	3,2088	3,9682
	50	2,8939	1,6874	1,9920
	100	2,3168	1,2212	1,9920
$f_{15}$	20	0,0029	0,0021	0,0016
	50	0,0024	0,0017	0,0015
	100	0,0013	0,0004	0,0010
$f_{16}$	20	-1,0258	0,0028	-1,0257
	50	-1,0253	0,0028	-1,0253
	100	-1,0249	0,0028	-1,0243
$f_{17}$	20	0,3990	0,0005	0,3988
	50	0,3991	0,0004	0,3991
	100	0,3993	0,0004	0,3995
$f_{18}$	20	5,4344	4,4217	3,0056
	50	3,0045	0,0027	3,0046
	100	3,0053	0,0026	3,0057
$f_{19}$	20	-3,8006	0,0277	-3,8084
	50	-3,8196	0,0129	-3,8139
	100	-3,8237	0,0151	-3,8192
$f_{20}$	20	-3,1591	0,0789	-3,2008
	50	-3,1944	0,0409	-3,2055
	100	-3,2115	0,0216	-3,2004
$f_{21}$	20	-9,8216	0,1844	-9,7841
	50	-9,8516	0,2070	-9,8493
	100	-9,7913	0,1890	-9,7811
$f_{22}$	20	-9,8942	0,2351	-9,8494
	50	-9,8448	0,2423	-9,7280
	100	-9,8466	0,2506	-9,7259
$f_{23}$	20	-9,6624	0,5987	-9,9458
	50	-9,8727	0,3602	-9,8525
	100	-9,8176	0,3486	-9,7546

**Tabela A.131** Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)

	$m(0)$	$\bar{y}$	$\sigma_y$	$m_d(y)$
$f_{26}$	20	884,8741	79,3431	928,3530
	50	802,8226	111,5446	770,0345
	100	762,5059	98,9598	708,7597

**Tabela A.132** Média, desvio médio e mediana dos valores mínimos (ODM-PME-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_1$	20	0,0006	2,35E-7	0,0009
	50	0,0006	9,55E-5	0,0009
	100	0,0006	3,07E-7	0,0009
$f_2$	20	0,0224	0,0004	0,0463
	50	0,0147	2,37E-5	0,0508
	100	0,0198	3,55E-5	0,0504
$f_3$	20	0,0009	0,0001	0,0019
	50	0,0009	1,76E-5	0,0015
	100	0,0009	3,21E-6	0,0015
$f_4$	20	0,0013	6,58E-17	0,0125
	50	0,0008	1,04E-18	0,0054
	100	0,0008	7,20E-5	0,0037
$f_5$	20	0,0449	0,0001	1,5360
	50	0,3313	4,00E-6	28,7070
	100	0,0615	5,79E-5	1,4224
$f_6$	20	0,0006	1,74E-6	0,0009
	50	0,0006	4,56E-6	0,0009
	100	0,0006	7,41E-7	0,0010
$f_7$	20	0,0002	6,46E-9	0,0009
	50	0,0004	2,77E-8	0,0009
	100	0,0007	3,00E-8	0,0009

**Tabela A.133** Faixa dos valores mínimos encontrados (ODM-PME-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_8$	20	-13962,7099	-25008,7535	-9007,2524
	50	-13900,4271	-17584,6602	-11629,4165
	100	-13898,5910	-22053,6486	-12503,8126
$f_9$	20	0,0007	3,04E-7	0,0012
	50	0,0008	8,34E-5	0,0014
	100	0,0008	0,0001	0,0014
$f_{10}$	20	0,0049	2,66E-5	0,0185
	50	0,0044	0,0002	0,0211
	100	0,0050	0,0001	0,0175
$f_{11}$	20	0,0007	6,12E-8	0,0009
	50	0,0006	1,00E-7	0,0009
	100	0,0006	2,39E-7	0,0009
$f_{12}$	20	0,0007	3,92E-5	0,0009
	50	0,0007	3,40E-5	0,0009
	100	0,0007	6,51E-5	0,0009
$f_{13}$	20	0,0007	2,54E-5	0,0009
	50	0,0007	4,64E-5	0,0009
	100	0,0006	5,34E-5	0,0009

**Tabela A.134** Faixa dos valores mínimos encontrados (ODM-PME-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{14}$	20	5,0198	0,9980	12,6705
	50	2,8939	0,9980	10,7631
	100	2,3168	0,9980	5,9288
$f_{15}$	20	0,0029	0,0004	0,0248
	50	0,0024	0,0004	0,0313
	100	0,0013	0,0005	0,0078
$f_{16}$	20	-1,0258	-1,0315	-1,0201
	50	-1,0253	-1,0314	-1,0201
	100	-1,0249	-1,0313	-1,0200
$f_{17}$	20	0,3990	0,3979	0,3999
	50	0,3991	0,3979	0,3999
	100	0,3993	0,3979	0,3999
$f_{18}$	20	5,4344	3,0000	30,0000
	50	3,0045	3,0000	3,0097
	100	3,0053	3,0002	3,0099
$f_{19}$	20	-3,8006	-3,8563	-3,6410
	50	-3,8196	-3,8620	-3,8004
	100	-3,8237	-3,8618	-3,8002
$f_{20}$	20	-3,1591	-3,2861	-2,3046
	50	-3,1944	-3,2971	-2,9556
	100	-3,2115	-3,2945	-3,1226
$f_{21}$	20	-9,8216	-10,1524	-9,5022
	50	-9,8516	-10,1526	-9,5003
	100	-9,7913	-10,1508	-9,5137
$f_{22}$	20	-9,8942	-10,3968	-9,5045
	50	-9,8448	-10,4006	-9,5013
	100	-9,8466	-10,4026	-9,5007
$f_{23}$	20	-9,6624	-10,5261	-2,4219
	50	-9,8727	-10,5362	-2,4273
	100	-9,8176	-10,5352	-2,4273

**Tabela A.135** Faixa dos valores mínimos encontrados (ODM-PME-F)

	$m(0)$	$\bar{y}$	$y_{\min}$	$y_{\max}$
$f_{26}$	20	884,8741	674,1653	978,9077
	50	802,8226	614,7398	992,9648
	100	762,5059	609,2515	997,1779

**Tabela A.136** Faixa dos valores mínimos encontrados (ODM-PME-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_1$	20	5662,13	2406	2	100
	50	4995,07	3005,5	201	100
	100	5268,67	3452,5	201	100
$f_2$	20	48168,67	50000	50000	6
	50	29152,79	50000	50000	42
	100	32938,43	50000	50000	35
$f_3$	20	27536,57	36807	50000	60
	50	29369,92	40709,5	50000	62
	100	24667,78	25851	50000	65
$f_4$	20	14600,94	9712	50000	89
	50	11335,87	8805,5	50000	95
	100	12940,36	10359	50000	97
$f_5$	20	35350,7	50000	50000	40
	50	37335,95	50000	50000	35
	100	37625,51	50000	50000	38
$f_6$	20	9802,37	8103	12401	100
	50	11592,47	7153,5	3506	100
	100	10246,57	7153,5	2205	99
$f_7$	20	240,78	102	1	100
	50	4,03	1	1	100
	100	33,05	1	1	100

Tabela A.137 Número de iterações para minimização (ODM-PME-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_8$	20	6636,41	2	2	87
	50	1515,23	1	1	97
	100	1,03	1	1	100
$f_9$	20	18650,18	7757	50000	84
	50	20515,63	11604,5	50000	84
	100	21687,97	15709	50000	88
$f_{10}$	20	44895,6	50000	50000	20
	50	44013,87	50000	50000	21
	100	44079,24	50000	50000	25
$f_{11}$	20	3578,51	3103	2	100
	50	3496,44	2954	1	100
	100	3730,5	3053	201	100
$f_{12}$	20	2681,19	2005	4005	100
	50	2718,26	1904	2103	100
	100	3025,09	2105,5	704	100
$f_{13}$	20	6438,59	4806	305	100
	50	6638,34	4504,5	903	100
	100	4150,41	1505	206	100

Tabela A.138 Número de iterações para minimização (ODM-PME-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{14}$	20	40141,5	50000	50000	20
	50	32023,76	50000	50000	36
	100	29002,48	50000	50000	42
$f_{15}$	20	43432,67	50000	50000	18
	50	37235,86	50000	50000	29
	100	31587,74	50000	50000	40
$f_{16}$	20	578,48	1	1	100
	50	1	1	1	100
	100	1	1	1	100
$f_{17}$	20	316,48	1	1	100
	50	1	1	1	100
	100	1	1	1	100
$f_{18}$	20	5519,25	2	2	91
	50	1,07	1	1	100
	100	1	1	1	100
$f_{19}$	20	13000,74	1	1	74
	50	1	1	1	100
	100	1	1	1	100
$f_{20}$	20	23204,65	1,5	1	54
	50	21500,57	1	1	57
	100	15000,7	1	1	70
$f_{21}$	20	2460,91	103	1	100
	50	3376,01	952,5	1	100
	100	1649,81	1	1	100
$f_{22}$	20	1100,59	1	1	100
	50	1341,58	1	1	100
	100	1304,51	1	1	100
$f_{23}$	20	6731,73	202	1	96
	50	4551,6	1	1	99
	100	3294,64	1	1	99

**Tabela A.139** Número de iterações para minimização (ODM-PME-F)

	$m(0)$	$\bar{k}$	$m_d(k)$	$m_o(k)$	$\phi(\%)$
$f_{26}$	20	47921,13	50000	50000	5
	50	29983,73	50000	50000	41
	100	16742,71	1	1	67

**Tabela A.140** Número de iterações para minimização (ODM-PME-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_1$	20	12401,02	6612	40	100
	50	14601,98	10811	5202	100
	100	20337,34	16705	10202	100
$f_2$	20	98137,34	101800	101800	6
	50	61167,34	104800	104800	42
	100	72475,2	109800	109800	35
$f_3$	20	56873,14	75414	101800	60
	50	63210,56	86219	104800	62
	100	58559,32	61502	109800	65
$f_4$	20	30984,24	21224	101800	89
	50	27471,74	22411	104800	95
	100	35680,72	30518	109800	97
$f_5$	20	72515,42	101800	101800	40
	50	79497,9	104800	104800	35
	100	85071,14	109800	109800	38
$f_6$	20	21404,74	18006	26602	100
	50	27984,94	19107	11812	100
	100	30293,14	24107	14210	99
$f_7$	20	1552,92	2004	20	100
	50	151,1	50	50	100
	100	649,2	100	100	100

Tabela A.141 Número de avaliações da função (ODM-PME-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_8$	20	13814,98	40	40	87
	50	3326,62	50	50	97
	100	103	100	100	100
$f_9$	20	38976,88	17314	101800	84
	50	45824,22	28009	104800	84
	100	53146,92	41218	109800	88
$f_{10}$	20	91591,2	101800	101800	20
	50	92827,74	104800	104800	21
	100	97958,48	109800	109800	25
$f_{11}$	20	8674,42	8006	40	100
	50	11460,24	10708	50	100
	100	17066,96	15906	10202	100
$f_{12}$	20	7162,38	5810	9810	100
	50	10236,52	8608	9006	100
	100	15850,18	14011	11208	100
$f_{13}$	20	14677,18	11412	2410	100
	50	18076,68	13809	6606	100
	100	18100,82	12810	10212	100

Tabela A.142 Número de avaliações da função (ODM-PME-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{14}$	20	82756,5	101800	101800	20
	50	68222,72	104800	104800	36
	100	64588,16	109800	100	42
$f_{15}$	20	88633,7	101800	101800	18
	50	78634,44	104800	104800	29
	100	70161,9	109800	109800	40
$f_{16}$	20	1288,72	20	20	100
	50	50	50	50	100
	100	100	100	100	100
$f_{17}$	20	796,02	20	20	100
	50	50	50	50	100
	100	100	100	100	100
$f_{18}$	20	11550,42	40	40	91
	50	53,5	50	50	100
	100	100	100	100	100
$f_{19}$	20	26482,8	20	20	74
	50	50	50	50	100
	100	100	100	100	100
$f_{20}$	20	47282,84	30	20	54
	50	45092,5	50	50	57
	100	33010	100	100	70
$f_{21}$	20	5938,1	2006	20	100
	50	9793,78	6705	50	100
	100	7763,52	100	100	100
$f_{22}$	20	2967,8	20	20	100
	50	4727	50	50	100
	100	6005,7	100	100	100
$f_{23}$	20	14497,38	2204	20	96
	50	11432,16	50	50	99
	100	10374,04	100	100	99

**Tabela A.143** Número de avaliações da função (ODM-PME-F)

	$m(0)$	$\bar{q}$	$m_d(q)$	$m_o(q)$	$\phi(\%)$
$f_{26}$	20	97606,98	101800	101800	5
	50	62927,62	104800	104800	41
	100	36980,1	100	100	67

**Tabela A.144** Número de avaliações da função (ODM-PME-F)



## Referências Bibliográficas

- [1] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley-Liss, 1999.
- [2] P. M. Matthews. An introduction to functional magnetic resonance of the brain. In P. Jezzard, P. M. Matthews, and S. M. Smith, editors, *Functional MRI: an introduction to methods*, New York, 2004. Oxford University Press.
- [3] S. Malinen, Y. Hlushchuk, and R. Hari. Towards natural stimulation in fMRI: Issues of data analysis. *NeuroImage*, (35):131–139, 2007.
- [4] A. J. Bartsch, G. Homola, S. Thesen, P. Sahmer, R. Keim, C. F. Beckmann, A. Biller, C. Knaus, and M. Bendszus. Scanning for the scanner: fMRI of audition by read-out omissions from echo-planar imaging. *NeuroImage*, (35):234–243, 2007.
- [5] J. O. Giacomantone. Ressonância magnética funcional com filtragem pela difusão anisotrópica robusta. Master's thesis, Programa de Pós-Graduação em Engenharia Elétrica da Universidade de São Paulo, São Paulo, Brasil, 2005.
- [6] J. A. Maldjian, A. Gottschalk, R. S. Patel, J. A. Detre, and D. C. Alsop. The Sensory Somatotopic Map of the Human Hand Demonstrated at 4 Tesla. *NeuroImage*, (10):55–62, 1999.
- [7] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, New York, 1999.
- [8] A. P. Braga, A. C. P. L. F. Carvalho, and T. B. Ludermir. *Redes Neurais Artificiais: Teoria e Aplicações*. LTC, São Paulo, 2000.
- [9] Z. L. Kovács. *Redes Neurais Artificiais: Fundamentos e Aplicações*. Acadêmica São Paulo, São Paulo, 1996.
- [10] D. E. Cooper. *As Filosofias do Mundo: Uma Introdução Histórica*. Loyola, São Paulo, 2002.
- [11] A. S. Vázquez. *Filosofia da Práxis*. Expressão Popular, São Paulo, 2007.
- [12] A. Gramsci. Alguns Problemas para o Estudo da Filosofia da Práxis. In *Concepção Dialética da História*. Civilização Brasileira, 1989.

- [13] A. Gramsci. Introdução ao Estudo da Filosofia e do Materialismo Histórico. In *Concepção Dialética da História*. Civilização Brasileira, 1989.
- [14] D. A. Fair, B. L. Schlaggar, A. L. Cohen, F. M. Miezin, N. U. F. Dosenbach, K. K. Wenger, M. D. Fox, A. Z. Snyder, M. E. Raichle, and S. E. Petersen. A method for using blocked and event-related fMRI data to study ‘resting state’ functional connectivity. *NeuroImage*, (35):396–405, 2007.
- [15] S. Xie, J. X. Xiao, J. Bai, and X. X. Jiang. Patterns of brain activation in patients with mild Alzheimer’s disease during performance of subtraction: An fMRI study. *Clinical Imaging*, (29):94–97, 2005.
- [16] R. L. Somorjai, R. Vivanco, and N. Pizzi. A novel, direct spatio-temporal approach for analyzing fMRI experiments. *Artificial Intelligence in Medicine*, (25):5–17, 2002.
- [17] S. C. Ngan, E. S. Yacoub, W. F. Auffermann, and X. Hu. Node merging on Kohonen’s self-organizing of fMRI data. *Artificial Intelligence in Medicine*, (25):19–33, 2002.
- [18] M. Jarmasz and R. L. Somorjai. Exploring regions of interest with cluster analysis (EROICA) using a spectral peak statistic for selecting and testing the significance of fMRI activation time series. *Artificial Intelligence in Medicine*, (25):45–67, 2002.
- [19] R. K. S. Kwan, A. C. Evans, and G. B. Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging*, 18(11):1085–1097, 1999.
- [20] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and Construction of a Realistic Digital Brain Phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998.
- [21] C. P. Slichter. *Principles of Magnetic Resonance*. Springer-Verlag, New York, 1978.
- [22] Z. P. Liang and P. C. Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. IEEE Press, New York, 2000.
- [23] P. Jezzard and S. Clare. Principles of nuclear magnetic resonance and MRI. In P. Jezzard, P. M. Matthews, and S. M. Smith, editors, *Functional MRI: an introduction to methods*, New York, 2004. Oxford University Press.
- [24] R. N. Sener. Diffusion MRI findings in Wilson’s disease. *Computerized Medical Imaging and Graphics*, (27):17–21, 2003.
- [25] L. S. Raffin, L. A. Bacheschi, L. R. Machado, J. P. S. Nóbrega, C. Coelho, and C. C. Leite. Diffusion-weighted MR imaging of cystic lesions of neurocysticercosis: a preliminary study. *Arquivos de Neuropsiquiatria*, 4(59):839–842, 2001.

- [26] G. Jost, S. Hähnel, S. Heiland, C. Stippich, M. E. Bellemann, and K. Sartor. An automated method for volumetric quantification of magnetization transfer of the brain. *Magnetic Resonance Imaging*, (20):593–597, 2002.
- [27] K. Marx. Critique of Hegel’s dialectics and philosophy. In *Economic and Philosophic Manuscripts of 1844*. International Publishers, 1980.
- [28] F. Engels. The role played by labor in the transition from ape to man. In *Collected Works of Karl Marx and Frederik Engels*. International Publishers, 1975.
- [29] A. Gramsci. Introduction to the Study of Philosophy and Historical Materialism. In *Prison Notebooks*. Columbia University, 1992.
- [30] A. Gramsci. Some Problems in the Study of the Philosophy of Praxis. In *Prison Notebooks*. Columbia University, 1992.
- [31] N. Bobbio. *Saggi su Gramsci*. Feltrinelli, Milano, 1990.
- [32] L. Goldmann. *Dialética e cultura*. Paz e Terra, Rio de Janeiro, 1967.
- [33] S. Haykin. *Redes Neurais: Princípios e Prática*. Bookman, Porto Alegre, 2001.
- [34] C. Thornley and F. Gibb. A dialectical approach to information retrieval. *Journal of Documentation*, 63(5):755–764, 2007.
- [35] J. B. Rosser Jr. Aspects of dialectics and nonlinear dynamics. *Cambridge Journal of Economics*, 24(3):311–324, 2000.
- [36] L. Konder. *O futuro da Filosofia da Práxis: o pensamento de Marx no século XXI*. Paz e Terra, Rio de Janeiro, 1992.
- [37] G. A. Bornheim. *Dialética: teoria, práxis; Ensaio para uma crítica da fundamentação ontológica da Dialética*. Globo-EDUSP, São Paulo, 1977.
- [38] L. Konder. *O que é Dialética*. Brasiliense, São Paulo, 2004.
- [39] Aristóteles. *A Ética*. Ediouro, Rio de Janeiro, 1985.
- [40] L. Konder. *Marx: vida e obra*. Expressão Popular, São Paulo, 1998.
- [41] L. Konder. *Hegel: a razão quase enlouquecida*. Campus, Rio de Janeiro, 1991.
- [42] K. Marx. Crítica da dialética e da filosofia de Hegel. In *Manuscritos Econômico-Filosóficos*, São Paulo, 2002. Martin Claret.
- [43] V. Arcary. A concepção marxista da História e a centralidade do conceito de época revolucionária. In A. Boito Jr. and C. N. Toledo, editors, *Marxismo e ciências humanas*, São Paulo, 2003. Xamã.

- [44] P. P. A. Funari. História, contradições e conflitos. In A. Boito Jr. and C. N. Toledo, editors, *Marxismo e ciências humanas*, São Paulo, 2003. Xamã.
- [45] M. W. B. Almeida. Marxismo e antropologia. In A. Boito Jr. and C. N. Toledo, editors, *Marxismo e ciências humanas*, São Paulo, 2003. Xamã.
- [46] O. Ianni. *Dialética e Capitalismo: Ensaio sobre o Pensamento de Marx*. Vozes, São Paulo, 1988.
- [47] K. Marx. *Miséria da Filosofia*. Martin Claret, São Paulo, 2007.
- [48] R. Malagodi. *O que é Materialismo Dialético*. Brasiliense, São Paulo, 1988.
- [49] K. Marx and F. Engels. *A ideologia alemã*. Martins Fontes, São Paulo, 2002.
- [50] F. Engels. Sobre o papel do trabalho na transformação do macaco em homem. In R. Antunes, editor, *A Dialética do Trabalho*, São Paulo, 2004. Expressão Popular.
- [51] K. Marx. *O Capital*. Centauro, São Paulo, 2005.
- [52] K. Marx. *O Dezoito Brumário de Louis Bonaparte*. Centauro, São Paulo, 2003.
- [53] H. Lefebvre. *Lógica Formal Lógica Dialética*. Civilização Brasileira, Rio de Janeiro, 1969.
- [54] M. Löwy. *Ideologias e ciência social: elementos para uma análise marxista*. Cortez, São Paulo, 2003.
- [55] F. Carboni and M. Maestri. *A Linguagem Escravizada: Língua, História, Poder e Luta de Classes*. Expressão Popular, São Paulo, 2003.
- [56] M. Tsé-Tung. Sobre a prática. In *Sobre a Prática e sobre a Contradição*, São Paulo, 2001. Expressão Popular.
- [57] G. Plekhanov. Da filosofia da história. In *O Papel do Indivíduo na História*, São Paulo, 2003. Expressão Popular.
- [58] G. Plekhanov. Da concepção materialista da história. In *O Papel do Indivíduo na História*, São Paulo, 2003. Expressão Popular.
- [59] M. Harnecker. *Estratégia e Tática*. Expressão Popular, São Paulo, 2004.
- [60] M. Tsé-Tung. Sobre a contradição. In *Sobre a Prática e sobre a Contradição*, São Paulo, 2001. Expressão Popular.
- [61] G. Plekhanov. O papel do indivíduo na história. In *O Papel do Indivíduo na História*, São Paulo, 2003. Expressão Popular.
- [62] A. Singhal. Modern Information Retrieval: A Brief Overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 24(4):35–43, 2001.

- [63] B. Hjørland. *Information Seeking and Subject Representation: An Actively Theoretical Approach to Information Science*. Greenwood Press, Westport, 1997.
- [64] S. Jarupathirun and F. M. Zahedi. Dialectic decision support systems: System design and empirical evaluation. *Decision Support Systems*, 43:1553–1570, 2007.
- [65] J. B. Rosser Jr. The rise and fall of catastrophe theory applications in economics: Was the baby thrown out with the bathwater? *Journal of Economic Dynamics and Control*, 31:3255–3280, 2007.
- [66] W. E. Boyce and R. C. DiPrima. *Equações Diferenciais Elementares e Problemas de Valores de Contorno*. Livros Técnicos e Científicos, Rio de Janeiro, 1994.
- [67] K. T. W. Yiu and S. O. Cheung. A catastrophe model of construction conflict behavior. *Building and Environment*, 41:438–447, 2006.
- [68] I. G. Raftoyiannis, T. G. Constantakopoulos, G. T. Michaltsos, and A. N. Kounadis. Dynamic buckling of a simple geometrically imperfect frame using Catastrophe Theory. *International Journal of Mechanical Sciences*, 48:1021–1030, 2006.
- [69] E. J. Wagenmakers, P. C. M. Molenaar, R. P. P. P. Grasman, P. A. I. Hartelman, and H. L. J. van der Maas. Transformation invariant stochastic catastrophe theory. *Physica D*, 211:263–276, 2005.
- [70] S. R. Kulkarni, G. Lugosi, and S. S. Venkatesh. Learning Pattern Classification: A Survey. *IEEE Transactions on Information Theory*, 44(6):2178–2206, 1998.
- [71] F. M. Ham and I. Kostanic. *Principles of Neurocomputing for Science and Engineering*. McGraw-Hill, 2001.
- [72] F. A. Sala, F. M. Azevedo, and F. I. M. Argoud. Mapas de Kohonen na detecção de eventos epileptogênicos. In *III Congresso Latino-Americano de Engenharia Biomédica*, pages 533–536, João Pessoa, Brasil, 2004. Sociedade Brasileira de Engenharia Biomédica.
- [73] K. L. Wu and M. S. Yang. Alternative c-means clustering algorithms. *Pattern Recognition*, (35):2267–2278, 2002.
- [74] W. L. Hung, M. S. Yang, and D. H. Chen. Parameter selection for suppressed fuzzy c-means with an application to MRI segmentation. *Pattern Recognition Letters*, (27):424–438, 2006.
- [75] W. Chen, M. L. Giger, and U. Bick. A fuzzy c-means (fcm)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. *Academic Radiology*, 13(1):63–72, 2006.
- [76] C. Zhu and T. Jiang. Multicontext fuzzy clustering for separation of brain tissues in magnetic resonance images. *NeuroImage*, (18):685–696, 2003.

- [77] M. D. Alexiuk and N. J. Pizzi. Robust centroids using fuzzy clustering with feature partitions. *Pattern Recognition Letters*, (26):1039–1046, 2005.
- [78] E. Dimitriadou, M. Barth, C. Windischberger, K. Hornik, and E. Moser. A quantitative comparison of functional MRI cluster analysis. *Artificial Intelligence in Medicine*, (31):57–71, 2004.
- [79] C. Windischberger, M. Barth, C. Lamm, L. Schroeder, H. Bauer, R. C. Gur, and E. Moser. Fuzzy cluster analysis of high-field functional MRI data. *Artificial Intelligence in Medicine*, (29):203–223, 2003.
- [80] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, 1981.
- [81] R. Eberhart and Y. Shi. *Computational Intelligence: concepts to implementations*. Morgan Kaufmann, 2007.
- [82] W. P. Santos, R. E. Souza, P. B. Santos Filho, F. B. Lima Neto, and F. M. Assis. A Dialectical Approach for Classification of DW-MR Alzheimer’s Images. In *IEEE World Congress on Computational Intelligence (WCCI 2008)*, pages 1728–1735, Hong Kong, China, 2008. CIS-IEEE.
- [83] W. P. Santos, F. M. Assis, R. E. Souza, and P. B. Santos Filho. Evaluation of Alzheimer’s Disease by Analysis of MR Images using Objective Dialectical Classifiers as an Alternative to ADC Maps. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5506–5509, Vancouver, Canada, 2008. EMBS-IEEE.
- [84] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–256, 1948.
- [85] E. T. Jaynes. Information Theory and Statistical Mechanics I. *The Physical Review*, 106(4):620–630, 1957.
- [86] E. T. Jaynes. Information Theory and Statistical Mechanics II. *The Physical Review*, 108(2):171–190, 1957.
- [87] D. R. Farrier. Jaynes’ principle and maximum entropy spectral estimation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-32(6):1176–1183, 1984.
- [88] R. R. Yager. On the entropies of fuzzy measures. *IEEE Transactions on Fuzzy Systems*, 8(4):453–461, 2000.
- [89] J. L. Fan and Y. L. Ma. Some new fuzzy entropy formulas. *Fuzzy Sets and Systems*, 128:277–284, 2002.

- [90] H. M. Lee, C. M. Chen, J. M. Chen, and Y. L. Jou. An Efficient Fuzzy Classifier with Feature Selection Based on Fuzzy Entropy. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, 31(3):426–432, 2001.
- [91] C. Lucas and B. N. Araabi. Generalization of the Dempster-Shafer theory: a fuzzy-valued measure. *IEEE Transactions on Fuzzy Systems*, 7(3):255–270, 1999.
- [92] E. F. Combarro and P. Miranda. On the polytope of non-additive measures. *Fuzzy Sets and Systems*, 159:2145–2162, 2008.
- [93] R. Ferreira. *Bates, Darwin, Wallace e a Teoria da Evolução*. UnB-Edusp, 1990.
- [94] G. Rudolph. Convergence Analysis of Canonical Genetic Algorithms. *IEEE Transactions on Neural Networks*, 5(1):96–101, 1994.
- [95] Z. Bingul, A. Sekmen, and S. Zein-Sabatto. Evolutionary Approach to Multi-Objective Problems Using Adaptive Genetic Algorithms. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 1923–1927, Nashville, USA, 2000.
- [96] A. E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing*. Springer, 2007.
- [97] X. Yao, Y. Liu, and G. Lin. Evolutionary Programming Made Faster. *IEEE Transactions on Evolutionary Computation*, 3(2):82–102, 1999.
- [98] H. Dong, J. He, H. Huang, and W. Hou. Evolutionary Programming using a mixed mutation strategy. *Information Sciences*, 177:312–327, 2007.
- [99] M. Iwamatsu. Generalized evolutionary programming with Lévy-type mutation. *Computer Physics Communications*, 147:729–732, 2002.
- [100] M. Ji, H. Tang, and J. Guo. A single-point mutation evolutionary programming. *Information Processing Letters*, 90:293–299, 2004.
- [101] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82, 1997.
- [102] J. Kennedy and R. Eberhart. Particle swarm optimization. In *IEEE International Conference on Neural Networks*, pages 1942–1948, Perth, Australia, 1995. CIS-IEEE.
- [103] R. Eberhart and J. Kennedy. A new optimizer using particle swarm theory. In *IEEE Symposium on Micro Machine and Human Science*, pages 39–43, Nagoya, Japan, 1995. CIS-IEEE.
- [104] J. Chuanwena and E. Bompardb. A hybrid method of chaotic particle swarm optimization and linear interior for reactive power optimisation. *Mathematics and Computers in Simulation*, (68):57–65, 2005.

- [105] D. W. van der Merwe and A. P. Engelbrecht. Data clustering using particle swarm optimization. In *IEEE Congress on Evolutionary Computation*, pages 185–191, Canberra, Australia, 2003. CIS-IEEE.
- [106] X. Hu, R. C. Eberhart, and Y. Shi. Engineering optimization with particle swarm. In *IEEE Swarm Intelligence Symposium*, pages 53–57, Indiana, USA, 2003. CIS-IEEE.
- [107] I. C. Trelea. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters*, (85):317–325, 2003.
- [108] Y. Shi and R. A. Krohling. Co-evolutionary particle swarm optimization to solve min-max problems. In *IEEE Congress on Evolutionary Computation*, pages 1682–1687, Honolulu, USA, 2002. CIS-IEEE.
- [109] A. E. M. Zavala, A. H. Aguirre, and E. R. V. Diharce. Particle Evolutionary Swarm Optimization Algorithm (PESO). In *Proceedings of the Sixth Mexican International Conference on Computer Science (ENC'05)*, Mexico, 2005. CIS-IEEE.
- [110] M. Reyes-Sierra and C. A. C. Coello. Multi-Objective Particle Swarm Optimizers: A Survey of the State-of-the-Art. *International Journal of Computational Intelligence Research*, 2(3):287–308, 2007.
- [111] R. A. Krohling, F. Hoffmann, and L. S. Coelho. Co-evolutionary particle swarm optimization for min-max problems using Gaussian distribution. In *IEEE Congress on Evolutionary Computation*, pages 959–964, Honolulu, USA, 2004. CIS-IEEE.
- [112] M. J. Tahk and B. C. Sun. Coevolutionary Augmented Lagrangian Methods for Constrained Optimization. *IEEE Transactions on Evolutionary Computation*, 4(2):114–124, 2000.
- [113] X. Wang, Y. Wang, H. Zeng, and H. Zhou. Particle swarm optimization with escape velocity. In *IEEE International Conference on Computational Intelligence and Security*, pages 457–460, Guagzhou, China, 2006. CIS-IEEE.
- [114] H. Liu, A. Abraham, and W. Zhang. A fuzzy adaptive turbulent particle swarm optimisation. *International Journal of Innovative Computing and Applications*, 1(1):39–47, 2007.
- [115] J. Kennedy and R. C. Eberhart. A discrete binary version of the particle swarm algorithm. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 4104–4108, Orlando, USA, 1997.
- [116] E. C. Biscaia Jr., M. Schwaab, and J. C. Pinto. Um Novo Enfoque do Método do Enxame de Partículas. In *Workshop em Nanotecnologia e Computação Inspirada na Biologia (NanoBio2004)*, pages 1–7, Rio de Janeiro, 2004.

- [117] S. Rahnamayan, H. R. Tizhoosh, and M. M. A. Salama. A novel population initialization method for accelerating evolutionary algorithms. *Computers and Mathematics with Applications*, (53):1605–1614, 2007.
- [118] M. Jiang, Y. P. Luo, and S. Y. Yang. Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. *Information Processing Letters*, 102:8–16, 2007.
- [119] Y. L. Zheng, L. H. Ma, L. Y. Zhang, and J. X. Qian. On the convergence analysis and parameter selection in particle swarm optimization. In *Proceedings of the Second International Conference on Machine Learning and Cybernetics*, pages 1802–1807, Xian, China, 2003.
- [120] L. F. Q. Silveira, L. G. Q. Silveira-Júnior, F. M. Assis, and E. L. Pinto. Algoritmo genético aplicado na obtenção de constelações de sinais para transmissão digital com codificação wavelet em canais com desvanecimento Rayleigh. In *VIII Congresso Brasileiro de Redes Neurais*, Florianópolis, Brasil, 2007.
- [121] R. Poli and W. B. Langdon. Markov chain models of bare-bones particle swarm optimizers. In *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, pages 142–149, London, 2007.
- [122] E. F. Campana, G. Fasano, and A. Pinto. Dynamic system analysis and initial particles position in particle swarm optimization. In *IEEE Swarm Intelligence Symposium*, Indianapolis, USA, 2006.
- [123] C. J. F. T. Braak. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing*, 16(3):239–249, 2006.
- [124] E. Cantú-Paz. Markov chain models of parallel genetic algorithms. *IEEE Transactions on Evolutionary Computation*, 4(3):216–226, 2000.
- [125] A. H. Wright and Y. Zhao. Markov chain models of genetic algorithms. In *Proceedings of the Genetic and Evolutionary Computation (GECCO) Conference*, pages 734–741. Morgan Kaufmann, 1999.
- [126] J. Suzuki. A further result on the Markov chain model of genetic algorithms and its application to a simulated annealing-like strategy. *IEEE Transactions Systems, Man, Cybernetics*, 28:95–102, 1998.
- [127] D. Chen and C. Zhao. Data-driven fuzzy clustering based on maximum entropy principle and PSO. *Expert Systems with Applications*, 36:625–633, 2009.
- [128] K. Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE*, 86(11):2210–2239, 1998.

- [129] K. Rose, E. Gurewitz, and G. C. Fox. Statistical mechanics and phase transitions in clustering. *Physical Reviews Letters*, 65(8):945–948, 1990.
- [130] R. Zhang and J. Liu. Underwater image segmentation with maximum entropy based on particle swarm optimization (PSO). In *Proceedings of the First International Multi-Symposium on Computer and Computational Sciences (IMSCCS'06)*, 2006.
- [131] M. Omran, A. P. Engelbrecht, and A. Salman. Particle swarm optimization method for image clustering. *International Journal of Pattern Recognition and Artificial Intelligence*, 19(3):297–321, 2005.
- [132] H. Pedrini and W. R. Schwartz. *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. Thomson, São Paulo, 2008.
- [133] R. Duda, P. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, 2001.
- [134] R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1972.
- [135] F. van der Bergh and A. P. Engelbrecht. A new locally convergent particle swarm optimizer. In *IEEE International Conference on Systems, Man, and Cybernetics*, Hammamet, Tunisia, 2002. SMCS-IEEE.
- [136] C. D. Perttunen and B. E. Stuckman. The rank transformation applied to a multi-univariate method of global optimization. In *IEEE International Conference on Systems Engineering*, pages 217–220, 1989.
- [137] J. Grespan. A dialética do avesso. In A. Boito Jr. and C. N. Toledo, editors, *Marxismo e ciências humanas*, São Paulo, 2003. Xamã.
- [138] O. Tambosi. Marxismo e dialética: uma herança fatal. In A. Boito Jr. and C. N. Toledo, editors, *Marxismo e ciências humanas*, São Paulo, 2003. Xamã.
- [139] S. Das, A. Abraham, and A. Konar. Automatic clustering using an improved differential evolution algorithm. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 38(1):218–237, 2008.
- [140] S. Das, S. Sil, and U. K. Chakraborty. Kernel-Based Clustering of Image Pixels with Modified Differential Evolution. In *IEEE World Congress on Computational Intelligence (WCCI 2008)*, pages 3472–3479, Hong Kong, China, 2008. CIS-IEEE.
- [141] S. Das and A. Konar. Automatic image pixel clustering with an improved differential evolution. *Applied Soft Computing*, (9):226–236, 2009.
- [142] Z. Wang and A. C. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 9, 2002.

- [143] S. Saha and S. Bandyopadhyay. Application of a New Symmetry-Based Cluster Validity Index for Satellite Image Segmentation. *IEEE Geoscience and Remote Sensing Letters*, 5(2):166–170, 2008.
- [144] M. Ewers, S. J. Teipel, O. Dietrich, S. O. Schönberg, F. Jessen, R. Heun, P. Scheltens, L. van de Pol, N. R. Freymann, H. J. Moeller, and H. Hampela. Multicenter assessment of reliability of cranial MRI. *Neurobiology of Aging*, (27):1051–1059, 2006.
- [145] M. S. Mega, I. D. Dinov, J. C. Mazziotta, M. Manese, P. M. Thompson, C. Lindshield, J. Moussai, N. Tran, K. Olsen, C. I. Zoumalan, R. P. Woods, and A. W. Toga. Automated brain tissue assessment in the elderly and demented population: Construction and validation of a sub-volume probabilistic brain atlas. *NeuroImage*, (26):1009–1018, 2005.
- [146] O. Friman, G. Farneäck, and C. F. Westin. A bayesian approach for stochastic white matter tractography. *IEEE Transactions on Medical Imaging*, 25(8):965–978, 2006.
- [147] O. Naggara, C. Oppenheim, D. Rieu, N. Raoux, S. Rodrigo, G. D. Barba, and J. F. Meder. Diffusion tensor imaging in early Alzheimer’s disease. *Psychiatry Research Neuroimaging*, (146):243–249, 2006.
- [148] M. Bozzali, A. Falini, M. Franceschi, M. Cercignani, M. Zuffi, G. Scotti, G. Comi, and M. Filippi. White matter damage in alzheimer’s disease assessed in vivo using diffusion tensor magnetic resonance imaging. *Journal of Neurology, Neurosurgery and Psychiatry*, 72:742–746, 2002.
- [149] A. T. Du, N. Schuff, L. L. Chao, J. Kornak, F. Ezekiel, W. J. Jagust, J. H. Kramer, B. R. Reed, B. L. Miller, D. Norman, H. C. Chui, and M. W. Weiner. White matter lesions are associated with cortical atrophy more than entorhinal and hippocampal atrophy. *Neurobiology of Aging*, (26):553–559, 2005.
- [150] O. T. Carmichael, H. A. Aizenstein, S. W. Davis, J. T. Becker, P. M. Thompson, C. C. Meltzer, and Y. Liu. Atlas-based hippocampus segmentation in Alzheimer’s disease and mild cognitive impairment. *NeuroImage*, (27):979–990, 2005.
- [151] Y. Hirata, H. Matsuda, K. Nemoto, T. Ohnishi, K. Hirao, F. Yamashita, T. Asada, S. Iwabuchi, and H. Samejima. Voxel-based morphometry to discriminate early Alzheimer’s disease from controls. *Neuroscience Letters*, (382):269–274, 2005.
- [152] N. Pannacciulli, A. Del Parigi, K. Chen, D. S. N. T. Le, E. M. Reiman, and P. A. Tataranni. Brain abnormalities in human obesity: A voxel-based morphometric study. *NeuroImage*, (31):1419–1425, 2006.
- [153] S. Hayasaka, A. T. Du, A. Duarte, J. Kornak, G. H. Jahng, M. W. Weiner, and N. Schuff. A non-parametric approach for co-analysis of multi-modal brain imaging data: Application to Alzheimer’s disease. *NeuroImage*, (30):768–779, 2006.

- [154] W. P. Santos, R. E. Souza, A. F. D. Silva, and P. B. Santos Filho. Evaluation of Alzheimer's disease by analysis of MR images using multilayer perceptrons and committee machines. *Computerized Medical Imaging and Graphics*, 32(1):17–21, 2008.
- [155] W. P. Santos, R. E. Souza, A. F. D. Silva, and P. B. Santos Filho. Evaluation of Alzheimer's disease by analysis of MR images using multilayer perceptrons, polynomial nets and Kohonen LVQ classifiers. In *Lecture Notes in Computer Science: Computer Vision / Computer Graphics Collaboration Techniques and Applications (MIRAGE 2007)*, volume 2, pages 12–22, Rocquencourt, France, 2007. INRIA & CS-IEEE.
- [156] W. P. Santos, R. E. Souza, A. F. D. Silva, and P. B. Santos Filho. Avaliação da Doença de Alzheimer pela Análise Multiespectral de Imagens DW-MR por Mapas Auto-Organizados de Kohonen como Alternativa aos Mapas ADC. In *IV Congresso Latino-Americano de Engenharia Biomédica*, pages 399–402, Porlamar, Venezuela, 2007. IFMBE.
- [157] W. P. Santos, R. E. Souza, A. F. D. Silva, and P. B. Santos Filho. Avaliação da doença de Alzheimer pela análise multiespectral de imagens DW-MR por redes RBF como alternativa aos mapas ADC. In *VIII Congresso Brasileiro de Redes Neurais*, Florianópolis, Brasil, 2007. SBRN.
- [158] W. P. Santos, R. E. Souza, and P. B. Santos Filho. Evaluation of Alzheimer's disease by analysis of MR images using multilayer perceptrons and Kohonen SOM classifiers as an alternative to the ADC maps. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2118–2121, Lyon, France, 2007. EMBS-IEEE.
- [159] W. P. Santos, R. E. Souza, A. F. D. Silva, N. M. Portela, and P. B. Santos Filho. Análise multiespectral de imagens cerebrais de ressonância magnética ponderadas em difusão usando lógica nebulosa e redes neurais para avaliação de danos causados pela doença de Alzheimer. In *XI Congresso Brasileiro de Física Médica*, Ribeirão Preto, Brasil, 2006. Sociedade Brasileira de Física Médica.
- [160] W. P. Santos, R. E. Souza, A. F. D. Silva, N. M. Portela, and P. B. Santos Filho. Avaliação da doença de Alzheimer por análise de imagens de RMN utilizando redes MLP e máquinas de comitê. In *XX Congresso Brasileiro de Engenharia Biomédica*, São Pedro, Brasil, 2006. Sociedade Brasileira de Engenharia Biomédica.
- [161] W. C. Wu, K. H. Kao, P. H. Lai, and H. W. Chung. Multi-Component Decay Behavior on High-b-Value Diffusion Weighted MRI. In *Proceedings of the 23rd Annual EMBS International Conference*, pages 2286–2288, Istanbul, Turkey, 2001. EMBS-IEEE.
- [162] I. Elshafey. Diffusion tensor magnetic resonance imaging of lesions in multiple sclerosis patients. In *19th National Radio Science Conference*, pages 626–633, Alexandria, Egipto, 2002. EMBS-IEEE.

- [163] P. J. Basser and S. Pajevic. A Normal Distribution for Tensor-Valued Random Variables: Applications to Diffusion Tensor MRI. *IEEE Transactions on Medical Imaging*, 22(7):785–794, 2003.
- [164] B. Wang, P. K. Saha, J. K. Udupa, M. A. Ferrante, J. Baumgardner, D. A. Roberts, and R. R. Rizi. 3D airway segmentation method via hyperpolarized  $^3\text{He}$  gas MRI by using scale-based fuzzy connectedness. *Computerized Medical Imaging and Graphics*, (28):77–86, 2004.
- [165] Y. Chen, W. Guo, Q. Zeng, X. Yan, F. Huang, H. Zhang, G. He, B. C. Vemuri, and Y. Liu. Estimation, Smoothing, and Characterization of Apparent Diffusion Coefficient Profiles from High Angular Resolution DWI. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 588–593. CS-IEEE, 2004.
- [166] W. Guo, Q. Zeng, Y. Chen, and Y. Liu. Using multiple tensor detection to reconstruct white matter fiber traces with branching. In *Proceedings of the ISBI 2006*. CS-IEEE, 2006.
- [167] C.A. Castano-Moraga, C. Lenglet, R. Deriche, and J. Ruiz-Alzola. A Fast and Rigorous Anisotropic Smoothing Method for DT-MRI. In *Proceedings of the ISBI 2006*. CS-IEEE, 2006.
- [168] G. T. Gullberg, D. G. Roy, G. L. Zeng, A. L. Alexander, and D. L. Parker. Tensor Tomography. *IEEE Transactions on Nuclear Science*, 46(4):991–1000, 1999.
- [169] P. J. Basser. Diffusion-Tensor MRI: Theory, Experimental Design, and Data Analysis. In *Proceedings of the 2nd Joint EMBS BMES Conference*, pages 1165–1166, Houston, USA, 2002. EMBS-IEEE-BMES.
- [170] N. Kang, J. Zhang, E. S. Carlson, and D. Gembris. White matter fiber tractography via anisotropic diffusion simulation in the human brain. *IEEE Transactions on Medical Imaging*, 24(9):1127–1137, 2005.
- [171] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache. Clinical DT-MRI estimations, smoothing and fiber tracking with log-Euclidean metrics. In *Proceedings of the ISBI 2006*. CS-IEEE, 2006.
- [172] H. Li, T. Liu, G. Young, L. Guo, and S. T. C Wong. Brain Tissue Segmentation Based on DWI/DTI Data. In *Proceedings of the ISBI 2006*. CS-IEEE, 2006.
- [173] A. Bartesaghi and M. Nadar. Segmentation of Anatomical Structure from DT-MRI. In *Proceedings of the ISBI 2006*. CS-IEEE, 2006.
- [174] J. Sklansky and G. N. Wassel. *Pattern Classifiers and Trainable Machines*. Springer-Verlag, 1st edition, 1981.

- [175] D. Landgrebe. Hyperspectral image analysis. *IEEE Signal Processing Magazine*, Jan 2002.
- [176] P. Bandettini. Functional MRI today. *International Journal of Psychophysiology*, (63):138–145, 2007.
- [177] P. R. Bannister, J. M. Brady, and M. Jenkinson. Integrating temporal information with a non-rigid method of motion correction for functional magnetic resonance images. *Image and Vision Computing*, (25):311–320, 2007.
- [178] A. S. Lukic, M. N. Wernick, and S. C. Strother. An evaluation of methods for detecting brain activations from functional neuroimages. *Artificial Intelligence in Medicine*, (25):69–88, 2002.
- [179] B. Thirion, P. Pinel, S. Mériaux, A. Roche, S. Dehaene, and J. B. Poline. Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. *NeuroImage*, (35):105–120, 2007.
- [180] W. P. Santos, F. M. Assis, R. E. Souza, P. B. Mendes, H. S. S. Monteiro, and H. D. Alves. Dialectical Non-Supervised Image Classification. In *IEEE Congress on Evolutionary Computation (CEC 2009)*, pages 2480–2487, Trondheim, 2009. CIS-IEEE.
- [181] W. P. Santos and F. M. Assis. Optimization based on Dialectics. In *IEEE International Joint Conference on Neural Networks (IJCNN 2009)*, pages 2804–2811, Atlanta, 2009. CIS-IEEE.
- [182] W. P. Santos, F. M. Assis, R. E. Souza, and P. B. Santos Filho. Dialectical Classification of MR Images for the Evaluation of Alzheimer’s Disease. In A. Lazinica, editor, *Recent Advances in Biomedical Engineering*, Vienna, 2009. InTech.
- [183] W. P. Santos, F. M. Assis, R. E. Souza, P. B. Mendes, H. S. S. Monteiro, and H. D. Alves. A Dialectical Method to Classify Alzheimer’s Magnetic Resonance Images. In A. Lazinica, editor, *Evolutionary Computation*, Vienna, 2009. InTech.
- [184] W. P. Santos, F. M. Assis, R. E. Souza, A. C. T. C. Albuquerque, and M. L. B. Simas. A Monospectral Approach for fMRI Analysis using Kohonen Self-Organized Networks and Objective Dialectical Classifiers. *International Journal of Innovative Computing and Applications*, 1(4):260–273, 2008.
- [185] W. P. Santos, F. M. Assis, R. E. Souza P. B. Santos Filho, and F. B. Lima Neto. Dialectical Multispectral Classification of Diffusion-Weighted Magnetic Resonance Images as an Alternative to Apparent Diffusion Coefficients Maps to Perform Anatomical Analysis. *Computerized Medical Imaging and Graphics*, 33(6):442–460, 2009.
- [186] W. P. Santos, F. M. Assis, P. B. Mendes, H. S. S. Monteiro, and H. D. Alves. An Object-Oriented Approach to Design Dialectical Systems. In F. Columbus, editor, *Object-Oriented Analysis and Design*, New York, 2009. Nova Science.

- [187] W. P. Santos, F. M. Assis, and R. E. Souza. MRI Segmentation using Dialectical Optimization. In *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, USA, 2009. EMBS-IEEE.