

Universidade Federal do Espírito Santo
Centro Tecnológico
Programa de Pós-Graduação em Engenharia Elétrica

Klaus Fabian Côco

**Modelos Fractais para a Função de Vizinhança
na Análise Topográfica de Componentes
Independentes**

Vitória
2007

Klaus Fabian Côco

Modelos Fractais para a Função de Vizinhança na Análise Topográfica de Componentes Independentes

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do grau de Doutor em Engenharia Elétrica, na área de concentração em Automação.

Orientadores: Prof. Dr. Mário Sarcinelli Filho e Prof. Dr. Evandro Ottoni Teatini Salles.

Vitória
2007

Dados Internacionais de Catalogação-na-publicação(CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

Côco, Klaus Fabian, 1971-
C667m Modelos fractais para a função de vizinhança na análise topográfica
de componentes independentes / Klaus Fabian Côco.-2007.
180 f.:il.

Orientador: Mário Sarcinelli Filho.

Co-Orientador: Evandro Ottoni Teatini Salles.

Tese (doutorado): Universidade Federal de Espírito Santo, Centro
Tecnológico.

1. Fractais. 2. Processamento de imagens. 3. Variáveis latentes. 4.
Estatística matemática. I. Sarcinelli Filho, Mário. II. Salles, Evandro
Ottoni Teatini. III. Universidade Federal de Espírito Santo. Centro
Tecnológico. IV. Título.

CDU: 621.3

Modelos Fractais para a Função de Vizinhança na Análise Topográfica de Componentes Independentes

Klaus Fabian Côco

Tese submetida ao Programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do grau de Doutor em Engenharia Elétrica - Automação.

Aprovada em 01 de outubro de 2007.

Prof. Dr. Mário Sarcinelli Filho
Universidade Federal do Espírito Santo
Orientador

Prof. Dr. Evandro Ottoni Teatini Salles
Universidade Federal do Espírito Santo
Orientador

Prof. Dr. Nelson Delfino d'Ávila Mascarenhas
Universidade Federal de São Carlos

Prof. Dr. Adrião Duarte Doria Neto
Universidade Federal do Rio Grande do Norte

Prof. Dr. Hans-Jorg Andreas Schneebeli
Universidade Federal do Espírito Santo

Prof. Dr. Thomas Walter Rauber
Universidade Federal do Espírito Santo

Vitória
2007

*“Aos meus pais, meu filho Eduardo e à minha querida esposa
Ana Paula”.*

Agradecimentos

Durante todo o processo de estudo em minha vida tive ao meu lado minha família, meus amigos, colegas e professores que contribuíram pacientemente na preparação do conhecimento que vim acumulando ao longo de todos esses anos.

Nos últimos anos tenho contado com o apoio muito forte da minha esposa, Ana Paula, e em especial dos amigos Evandro, Durval e José Mário, sempre incentivando meu trabalho e procurando entender os desdobramentos das pesquisas realizadas.

Ao professor orientador Mário, sempre persuasivo quanto aos resultados obtidos na pesquisa em busca de publicações que venham validar o trabalho elaborado, meus sinceros agradecimentos à sua dedicação, paciência e otimismo fortalecendo meu ânimo nos momentos menos favoráveis.

Ao amigo e professor orientador Evandro fico extremamente grato pela disposição, paciência e tolerância durante todos esses anos de estudos. Não esqueço os momentos filosóficos de debate das várias nuances que orbitam a pesquisa desenvolvida. Durante muitas vezes me vi com pensamentos fixos por dias e até semanas buscando a compreensão e a solução para as questões levantadas nas reuniões, que por vezes tomavam horas sem que pudséssemos notar o passar do tempo.

Sou grato aos professores, colegas e amigos estudantes da pós-graduação em Engenharia Elétrica pelo ambiente agradável de estudo e trabalho, e em especial aos colegas de laboratório, Alexandre e Jorge, pela convívio aprazível e estimulante durante esses anos.

Resumo

Este trabalho propõe o uso de descritores fractais no critério topográfico na Análise Topográfica de Componentes Independentes (TICA - *Topographic Independent Component Analysis*) como uma implementação melhorada do modelo existente.

A implementação proposta visa contribuir para a melhoria da representação de imagens pelo modelo TICA, em especial, aquelas que possuem características estatísticas fractais, como, por exemplo, as imagens naturais.

A Análise de Componentes Independentes (ICA - *Independent Component Analysis*) é um método não-supervisionado de separação cega de fontes estatisticamente independentes, utilizado para representação de sinais e imagens através de um modelo linear. Estudos mostram que esse método é o que melhor representa o comportamento das células simples do córtex visual primário do sistema de visão humano.

Uma extensão desse método, denominada TICA, mostra-se capaz de modelar o comportamento das células complexas dessa mesma região do córtex visual, responsável pela organização espacial das células simples através de interações laterais entre os neurônios e da resposta à orientação de fase.

Nesse sentido, a proposta de uso de funções fractais no modelo topográfico é capaz de adequá-lo à modelagem do comportamento das células do córtex visual secundário, sem a perda da capacidade de modelar o comportamento das células do córtex visual primário.

Para os testes aqui apresentados foram selecionadas imagens artificiais, naturais e sintéticas, para avaliar a adequação da metodologia proposta. Os resultados obtidos são bastante encorajadores, e mostram que a TICA baseada em modelos fractais é adequada para tratar imagens com características fractais estatísticas de alta-ordem.

Abstract

This work discusses the use of fractal descriptors in the topographic criterion in Topographic Independent Component Analysis (TICA), as an improvement in the implementation of the existing model.

The proposed implementation aims at to contribute with the improvement of the representation of images by the TICA model, in special to the TICA-based representation of images characterized by fractal statistics as, for example, the natural images.

The Independent Component Analysis (ICA) is a non-supervised method for the blind separation of statistically independent sources, used for representing signals and images through a linear model. Studies have shown that such a method is the one that better represents the behavior of the simple cells of the primary visual cortex of the human visual system.

An extension of this method, the TICA, has shown to be able to model the behavior of the complex cells of the same region of the visual cortex, responsible for the spatial organization of the simple cells through the side iterations among the neurons and the response to the phase orientation.

In this sense, the proposal of using fractal functions in the topographic model is able to make it suitable for modeling the behavior of the cells of the secondary visual cortex, without losing the capability of modeling the behavior of the cells of the primary visual cortex.

For the tests presented here, artificial, natural and synthetic images, were selected in order to evaluate the suitability of the proposed methodology. The results obtained so far, are quite encouraging and have shown that TICA based on fractal models is suitable to deal with images having characteristics of high-order fractal statistics.

Sumário

I	Introdução	1
1	O Problema	2
II	Fundamentação Teórica	15
2	Texturas e Imagens Naturais	16
2.1	Texturas Naturais	18
2.2	Imagens Naturais e Informação Espectral	19
2.3	Texturas e a Percepção Humana	26
3	Análise de Componentes Independentes	32
3.1	Estatísticas de alta-ordem	32
3.2	A Análise de Componentes Independentes	35
3.2.1	Maximização da Não-Gaussianidade	38
3.2.2	Técnicas de Estimação na ICA	40
3.2.3	ICA e o Modelo de Visão	41
3.3	Análise Topográfica de Componentes Independentes	44
3.3.1	Modelo de Estimação na TICA	46
3.3.2	TICA e o Modelo de Visão	49
4	Fractais	51
4.1	Fractais Estatísticos	53
4.1.1	Fractais Estatísticos e Texturas Naturais	53
4.1.2	Coeficiente de Variação Espacial Estatístico	55
4.1.3	Distribuição α -estável	56
4.1.4	Fractais e o Modelo de Visão	57
III	Modelagem	60
5	Elaboração dos Modelos	61

5.1	Espectro Fractal Baseado em Morfologia	65
5.2	Modelagem Baseada em Fractal	69
5.2.1	Modelo Baseado na Energia	72
5.2.2	Modelo Baseado no Descritor LMME	74
5.2.3	Modelo Baseado na Energia e no Descritor LMME . . .	75
5.2.4	Modelo Baseado no Descritor LMME Modificado . . .	77
5.2.5	Modelo Baseado no Descritor LMME obtido em Múltiplas Escalas	79
5.2.6	Modelo Baseado no Descritor LMME obtido em Múltiplas Cenas	82
5.2.7	Modelo Baseado na Lacunaridade Fractal	84
5.2.8	Modelo Baseado no Coeficiente de Variação Espacial Estatístico	85
IV	Experimentos	87
6	Aplicação dos Modelos em Segmentação	88
7	Preparação	93
8	Resultados e Análises	101
9	Conclusões e Projetos Futuros	141
A	Processamento de Imagens	144
A.1	Morfologia	144
A.2	Segmentação, Representação, Descrição e Classificação	148

Lista de Figuras

1.1	Regiões geometricamente auto-similares em diferentes escalas em uma imagem natural (observe a marcação branca em “V” nos caules da planta).	5
1.2	Regiões estatisticamente auto-similares em diferentes escalas em uma imagem natural.	6
1.3	Exemplo de espectro $1/f$	6
1.4	Regiões visualmente identificadas na imagem fractal.	8
1.5	Esquema do olho humano onde pode-se observar a localização da retina e do nervo óptico [74]	10
1.6	Células da retina [15]	11
1.7	Trajetória do estímulo visual entre a retina e o córtex visual primário V1.	11
1.8	Exemplo de percepção dos estímulos visuais entre a retina e o córtex visual [102]	12
1.9	Nesta imagem o cérebro supõe que exista um quadrado que sobrepõe a figura interrompendo os círculos, porém, a imagem apenas possui círculos cortados para induzir a essa percepção.	13
2.1	Exemplo de espectro $1/f$ para cenas naturais e cenas artificiais [103].	20
2.2	Relação entre a assinatura espectral e a distância de observação para imagens artificiais (ambientes humanos). (a) Exemplos de cenas artificiais em diferentes distâncias (b) Espectros de potências das imagens de exemplo (c) Assinatura espectral de 300 a 400 imagens semelhantes às dos exemplos [103].	21
2.3	Relação entre a assinatura espectral e a distância de observação para imagens naturais. (a) Exemplos de cenas naturais em diferentes distâncias (b) Espectros de potências das imagens de exemplo (c) Assinatura espectral de 300 a 400 imagens semelhantes às dos exemplos [103].	22
2.4	Efeito da distância da imagem no julgamento da cena. A imagem à esquerda é a imagem à direita invertida (espelhamento horizontal) [103]. Observe que a complexidade das cenas dificulta, em um primeiro momento, a percepção da orientação das imagens.	23

2.5	Efeito da distância da imagem no julgamento da cena. Observe que é facilmente percebida, pelo observador, a inversão da imagem devido à distância de observação.	24
2.6	Exemplo de módulo e fase de Fourier. (a) Imagem de teste com figuras geométricas. (b) Informação do módulo de Fourier da imagem de teste. (c) Informação da fase de Fourier da imagem de teste.	24
2.7	Reconstrução de imagem a partir apenas da informação de fase do espectro de Fourier. (a) Imagem original, (b) Imagem reconstruída, (c) Imagem original, (d) Imagem reconstruída.	25
2.8	Texturas estatísticas. (a) Mosaico das texturas com médias $\mu_1 = \mu_2 = 0.1$, variâncias $\sigma_1^2 = \sigma_2^2 = 1$, skewness $skew_1 = 0.1$ e $skew_2 = -0.1$ e kurtosis $kurt_1 = 2$ e $kurt_2 = 4$. (b) Mosaico das texturas com médias $\mu_1 = \mu_2 = 0.1$, variâncias $\sigma_1^2 = 1$ e $\sigma_2^2 = 1.6$, skewness $skew_1 = skew_2 = 0.1$ e kurtosis $kurt_1 = kurt_2 = 2$. (c) Borda de separação entre texturas nos mosaicos.	26
2.9	Campos receptivos 2-D das células em V1 do córtex cerebral de um gato são funções de Gabor [22].	28
2.10	As linhas horizontais da figura são paralelas, porém a disposição espacial das cores causam a impressão de que as linhas não são paralelas.	29
2.11	Os quadrados marcados com A e B possuem a mesma cor, porém, devido à disposição espacial das cores e à sombra do cilindro, eles aparentam possuir cores diferentes [79].	29
3.1	Modelo ICA para representação de imagens. A imagem é obtida pela soma do produto das componentes independentes s_i pelas componentes da base, obtida da matriz de mistura.	37
3.2	Exemplo de uma base TICA de textura natural [49].	47
4.1	Fractais geométricos [33].	52
4.2	Fractais estatísticos [99].	52
5.1	Dilatações do elemento estruturante Y_ϵ	67
5.2	Imagem composta pelos descritores fractais estatísticos (LMME) da Figura 5.4.	68
5.3	Imagem resultante do processo de detecção de borda aplicado à Figura 5.4 utilizando o método Sobel.	68
5.4	Imagem Natural.	69
5.5	Imagem de teste para a função de vizinhança (20x20).	72
5.6	Função de Vizinhança Energia. Elemento estruturante “disco” de raio 1.	73
5.7	Função de Vizinhança Energia. Elemento estruturante “quadrado” de lado 3.	74

5.8	Função de Vizinhança: Descritor LMME.	75
5.9	Função de Vizinhança: Fractal-Energia.	77
5.10	Função de Vizinhança: Descritor LMME Modificado.	78
5.11	Função de Vizinhança: Fractal Multi-Escala. Elemento estruturante com duas decimações.	80
5.12	Função de Vizinhança: Fractal Multi-Escala sem decimação do elemento estruturante.	81
5.13	Função de Vizinhança: Fractal Multi-Escala com uma decimação do elemento estruturante.	81
5.14	Função de Vizinhança: Fractal Multi-Cena sem dilatações do elemento estruturante.	83
5.15	Função de Vizinhança: Fractal Multi-Cena com uma dilatação do elemento estruturante.	83
5.16	Função de Vizinhança: Lacunaridade Fractal obtida pela segunda ordem estatística.	84
5.17	Função de Vizinhança: Lacunaridade Fractal obtida pela quarta ordem estatística.	85
5.18	Função de Vizinhança: Coeficiente de Variação Espacial Estatística. . . .	86
7.1	Imagens artificiais. (a) Totens (b) Mulher.	94
7.2	Imagens naturais. (a) Corujas (b) Praia (c) Montanhas (d) Peixe. . . .	95
7.3	Imagens com texturas estatísticas sintéticas. (a) Imagem estocástica com cumulantes (a notação é $C_{ordem}^{textura}$) $C_1^1 = C_1^2 = 0$, $C_2^1 = C_2^2 = 1$, $C_3^1 = 0,01$ e $C_3^2 = -0,01$, e $C_4^1 = 2$ e $C_4^2 = 3$; e (b) Imagem não estocástica com médias $\mu_1 = \mu_2 = 0$, variâncias $\sigma_1^2 = \sigma_2^2 = 2$, skewness $skew_1 = 0$, $skew_2 = 0,1$ e kurtosis $kurt_1 = 2$ e $kurt_2 = 3$, onde a textura um é a externa e a textura dois é a interna do mosaico (256×256)	95
7.4	Imagens de Brodatz (256×256) com as texturas: (a) D17 (interna) e D16 (externa), (b) D52 (interna) e D20 (externa), (c) D16 (interna) e D103 (externa) e (d) D55 (interna) e D22 (externa)	96
7.5	Espetros $1/f$ das imagens. (a) Totem (b) Loira (c) Corujas (d) Praia (e) Montanhas (f) Peixe.	97
7.6	Espetros $1/f$ das texturas estatísticas sintéticas. (a) Figura 7.3a Externa (b) Figura 7.3a Interna (c) Figura 7.3b Externa (d) Figura 7.3b Interna.	98
7.7	Espetros $1/f$ das texturas de Brodatz. (a) Figuras 7.4a e 7.4c Interna (D17) (b) Figura 7.4a Externa (D16) (c) Figura 7.4b Interna (D52) (d) Figura 7.4b Externa (D20) (e) Figura 7.4c Externa (D103) (f) Figura 7.4d Interna (D55) (g) Figura 7.4d Externa (D22).	99

8.1	Regiões segmentadas, da imagem da Figura 7.1a (pág. 94), de forma não supervisionada para dois grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	102
8.2	Segmentação não-supervisionada da imagem da Figura 7.1a (pág. 94) para dois grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	103
8.3	Diferenças da segmentação não-supervisionada entre a TICAIE e as demais técnicas aplicadas à imagem da Figura 7.1a (pág. 94) para dois grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	104
8.4	Regiões segmentadas, da imagem da Figura 7.1b (pág. 95), de forma não supervisionada para dois grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	106
8.5	Segmentação não-supervisionada da imagem da Figura 7.1b (pág. 95) para dois grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	107
8.6	Diferenças da segmentação não-supervisionada entre a TICAIE e as demais técnicas aplicadas à imagem da Figura 7.1b (pág. 95) para dois grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	108
8.7	Regiões segmentadas, da imagem da Figura 7.1b (pág. 94), de forma não supervisionada para três grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	110
8.8	Segmentação não-supervisionada da imagem da Figura 7.1b (pág. 94) para três grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	111

8.9	Diferenças da segmentação não-supervisionada entre a TICA E e as demais técnicas aplicadas à imagem da Figura 7.1b (pág. 94) para três grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	112
8.10	Regiões segmentadas, da imagem da Figura 7.2a (pág. 95), de forma não supervisionada para dois grupamentos usando o <i>k-means</i> : (a) TICA E, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	113
8.11	Segmentação não-supervisionada da imagem da Figura 7.2a (pág. 95) para dois grupamentos usando o <i>k-means</i> : (a) TICA E, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	114
8.12	Diferenças da segmentação não-supervisionada entre a TICA E e as demais técnicas aplicadas à imagem da Figura 7.2a (pág. 95) para dois grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	114
8.13	Regiões segmentadas, da imagem da Figura 7.2b (pág. 95), de forma não supervisionada para três grupamentos usando o <i>k-means</i> : (a) TICA E, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	115
8.14	Segmentação não-supervisionada da imagem da Figura 7.2b (pág. 95) para três grupamentos usando o <i>k-means</i> : (a) TICA E, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	116
8.15	Diferenças da segmentação não-supervisionada entre a TICA E e as demais técnicas aplicadas à imagem da Figura 7.2b (pág. 95) para três grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	117
8.16	Regiões segmentadas, da imagem da Figura 7.2c (pág. 95), de forma não supervisionada para dois grupamentos usando o <i>k-means</i> : (a) TICA E, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	119

8.17	Segmentação não-supervisionada da imagem da Figura 7.2c (pág. 95) para dois grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	120
8.18	Diferenças da segmentação não-supervisionada entre a TICAIE e as demais técnicas aplicadas à imagem da Figura 7.2c (pág. 95) para dois grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	120
8.19	Regiões segmentadas, da imagem da Figura 7.2c (pág. 95), de forma não supervisionada para três grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	121
8.20	Segmentação não-supervisionada da imagem da Figura 7.2c (pág. 95) para três grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	122
8.21	Diferenças da segmentação não-supervisionada entre a TICAIE e as demais técnicas aplicadas à imagem da Figura 7.2c (pág. 95) para três grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	122
8.22	Regiões segmentadas, da imagem da Figura 7.2d (pág. 95), de forma não supervisionada para três grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	123
8.23	Segmentação não-supervisionada da imagem da Figura 7.2d (pág. 95) para três grupamentos usando o <i>k-means</i> : (a) TICAIE, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenar de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	124
8.24	Diferenças da segmentação não-supervisionada entre a TICAIE e as demais técnicas aplicadas à imagem da Figura 7.2d (pág. 95) para três grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	124

8.25	Regiões segmentadas, da imagem da Figura 7.2d (pág. 95), de forma não supervisionada para quatro grupamentos usando o <i>k-means</i> : (a) TICAÉ, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenças de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	125
8.26	Segmentação não-supervisionada da imagem da Figura 7.2d (pág. 95) para quatro grupamentos usando o <i>k-means</i> : (a) TICAÉ, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (cenças de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).	126
8.27	Diferenças da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.2d (pág. 95) para quatro grupamentos usando o <i>k-means</i> : (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICAVEE, (g) TICA FMC (cenças de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).	126
8.28	Erro da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.3a (pág. 95). (a) TICAÉ, TICA FM e TICA FMM, (b) TICAÉ, TICALF2 e TICALF4, (c) TICAÉ, TICA FMC e TICA FME, (d) TICAÉ, TICAVEE e TICA FE.	127
8.29	Erro da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.3b (pág. 95). (a) TICAÉ, TICA FM e TICA FMM, (b) TICAÉ, TICALF2 e TICALF4, (c) TICAÉ, TICA FMC e TICA FME, (d) TICAÉ, TICAVEE e TICA FE.	129
8.30	Erro da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.4a (pág. 96). (a) TICAÉ, TICA FM e TICA FMM, (b) TICAÉ, TICALF2 e TICALF4, (c) TICAÉ, TICA FMC e TICA FME, (d) TICAÉ, TICAVEE e TICA FE.	130
8.31	Erro da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.4b (pág. 96). (a) TICAÉ, TICA FM e TICA FMM, (b) TICAÉ, TICALF2 e TICALF4, (c) TICAÉ, TICA FMC e TICA FME, (d) TICAÉ, TICAVEE e TICA FE.	131
8.32	Diferenças da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.4c (pág. 96). (a) TICAÉ, TICA FM e TICA FMM, (b) TICAÉ, TICALF2 e TICALF4, (c) TICAÉ, TICA FMC e TICA FME, (d) TICAÉ, TICAVEE e TICA FE.	133
8.33	Diferenças da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.4d (pág. 96). (a) TICAÉ, TICA FM e TICA FMM, (b) TICAÉ, TICALF2 e TICALF4, (c) TICAÉ, TICA FMC e TICA FME, (d) TICAÉ, TICAVEE e TICA FE.	135

8.34	Bancos de filtros da Figura 7.4d (pág. 96) obtidos pela: (a) TICAE, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (cenos de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).	137
8.35	Espectro de frequências e assinatura espectral para um filtro de cada banco obtido pelas técnicas: (a) TICAE, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (cenos de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).	139
A.1	Exemplo da dilatação morfológica.	145
A.2	Exemplo da erosão morfológica.	145
A.3	Exemplo de adjacência [36]. (a) Arranjo dos pixels em uma parte da imagem; (b) Conjunto de pixels 4-adjacentes em relação ao pixel central; (c) Conjunto de pixels 8-adjacentes e (d) Conjunto de pixels m-adjacentes.	147

Lista de Tabelas

6.1	Desempenho médio de processamento para 2000 amostras 8×8 da mistura em 100 iterações.	91
8.1	Melhor caso aproximado para a segmentação não-supervisionada da Figura 7.3a.	132

Prefácio

As pesquisas em processamento de imagens e sinais desenvolvidas no laboratório Prof. Ailson Rosetti de Almeida - CISNE - da Universidade Federal do Espírito Santo têm enfatizado, já a algum tempo, os campos da Engenharia Elétrica relacionadas à biometria, redes neurais e biomedicina, com contribuições nas áreas de reconhecimento de faces, forma de caminhar, impressões digitais, sinais biomédicos e modelagem baseada em sistemas naturais.

O estudo apresentado neste trabalho vem contribuir com as pesquisas de processamento de imagens realizadas no laboratório, especificamente na área de modelagem baseada em sistemas naturais, com a inclusão de modelos fractais na análise topográfica de componentes independentes.

O presente trabalho teve como inspiração o funcionamento de um conjunto de células do sistema de visão humano, mais especificamente as células do córtex visual V2, que apresentam comportamento muito semelhante aos modelos fractais estatísticos utilizados neste trabalho.

O problema de tratar a representação de imagens naturais se deve à enorme gama de variáveis e relações presentes entre os componentes dessas imagens e a grande disparidade das imagens naturais existentes à nossa volta, todavia, sabe-se que essas imagens possuem uma boa parcela de características estatísticas fractais, o que permite explorar esse modelo na representação proposta nesta pesquisa.

Assim, o texto elaborado na tese explica a modelagem do sistema proposto, organizando o trabalho em três partes fundamentais.

Na primeira, o problema a ser abordado é introduzido e caracterizado, seguido pela motivação e importância que levou à realização do estudo apresentado e pela definição dos objetivos a serem alcançados. Nesta primeira parte, elementos de inspiração neurocientíficos, baseados no modelo humano de visão, já são introduzidos como conceitos passíveis de comparação com a abordagem de tese proposta neste trabalho.

Na segunda parte, é montado um referencial teórico acerca do conhecimento existente em torno dos elementos utilizados para a solução do problema. Esses elementos são as texturas e as imagens naturais, além das téc-

nicas de modelagem ICA (*Independent Component Analysis*), sua variação TICA (*Topographic Independent Component Analysis*) e os modelos Fractais Estatísticos.

Em cada etapa da modelagem teórica, da segunda parte, foi elaborada uma seção correlacionando as características das técnicas apresentadas com as respectivas características observadas e conhecidas do sistema neurobiológico de visão. Novamente, a comparação compões apenas um maneira inspiradora de tecer, a modelagem apresentada nesta tese, com o sistema humano de visão.

O sistema humano de visão, por sua vez, é a referência natural para comparação da qualidade da representação de imagens proposta neste trabalho. Dessa forma, todas as abordagens neurofisiológicas usadas nesta pesquisa são pautadas em experimentos científicos amplamente conhecidos e devidamente documentados em publicações de notória relevância pela comunidade científica.

Na terceira parte é elaborado o modelo, utilizando fractais no critério topográfico do modelo ICA, bem como sua avaliação através de experimentos de segmentação não-supervisionada em algumas imagens artificiais, sintéticas e naturais envolvendo texturas ou imagens da natureza. Na terceira parte também são criticados os resultados e elaborada a conclusão da tese.

Um apêndice sobre processamento de imagens trata de noções básicas acerca das teorias sobre morfologia matemática e conceitos de segmentação, classificação, representação e descrição.

Considerações finais são expostas no último momento, retratando os trabalhos futuros e sugestões para a complementação desta pesquisa.

Parte I

Introdução

Capítulo 1

O Problema

Caracterização do Problema e Motivações para a Abordagem

Diariamente nos deparamos, de maneira quase sempre inconsciente, com inúmeras situações das quais demandamos de complexos circuitos neuronais para aquisição, segmentação, classificação e avaliação das imagens percebidas no nosso sistema de visão.

A grande diversidade de estruturas que compõem o mundo à nossa volta formam um aparato de imagens que vão desde estruturas muito simples e regulares até modelos altamente complexos e que são interpretados de maneira satisfatória pelo mecanismo humano de visão.

Imagens como paisagens vegetais, texturas de mármore, granitos, fotos de satélite e texturas de madeira, naturalmente reconhecidos e interpretados pelo mecanismo de visão humana, são exemplos onde o indeterminismo e a alta variabilidade dos elementos descritores dessas imagens torna difícil a sua caracterização por sistemas automatizados.

Imagens como as descritas no parágrafo anterior são conhecidas como *imagens naturais* e compõem uma parcela considerável do universo de imagens percebidas ao longo da existência humana.

Embora o sistema humano seja bastante eficaz na identificação dessas imagens, alguns problemas são associados ao elemento humano quando lhe é exigida uma atividade um pouco automatizada de classificação. Por exemplo, no controle de qualidade que envolva inspeção visual.

É comum observar uma perda da qualidade de classificação do avaliador ao longo do tempo nas imagens que possuem características visuais muito complexas.

Provavelmente, isso ocorre devido à capacidade de aprendizado sempre

estar ativa no indivíduo, permitindo, por exemplo, que uma seqüência de imagens ligeiramente semelhantes, quando observadas de maneira seqüencial, porém com um pequeno grau de distinção relevante entre as extremidades da seqüência (o que permitiria a observação de diferenças entre as imagens), não seja percebido pelo avaliador humano. Outra hipótese seria a perda da eficiência classificatória por cansaço visual e mental.

Nesse sentido, a busca por entender e inspirar-se no sistema de visão humano se deve, em grande parte, à qualidade e eficiência da percepção e interpretação das imagens observadas por esse sistema, bem como também à necessidade de se criar mecanismos automáticos de visão que possam funcionar adequadamente sem, contudo, possuir os problemas como os descritos no parágrafo anterior.

O sistema de visão humano pode detectar, segmentar e classificar uma diversidade muito grande de estímulos visuais: imagens em movimento, padronagens desconhecidas, em duas ou três dimensões, com variação de luminosidade e oclusão.

O interesse em desenvolver métodos de processamento de imagens digitais, segundo Gonzalez e Woods [36], se deve principalmente a duas áreas de aplicação: a melhoria da informação visual para interpretação humana e o processamento de dados de cenas para a percepção automática através de máquinas.

Ao longo de muitos anos isso motivou os estudiosos e cientistas a desenvolverem uma variedade de técnicas para processamento de imagens que, atualmente, são empregadas nos mais diversos campos de atividades da sociedade moderna como: na indústria (reconhecimento de padrões de texturas, controle de processos), na segurança pública (reconhecimento de face, impressão digital, biometria), no lazer (câmeras fotográficas digitais, jogos eletrônicos), na área médica (diagnóstico de doenças, cirurgias), entre outras.

Apesar do considerável avanço das técnicas de processamento de imagens, ainda hoje não existe um modelo artificial de visão que permita o reconhecimento de uma ampla variedade de imagens de maneira tão precisa como o elaborado pelo sistema de visão humano.

Os recentes avanços da engenharia eletrônica e de computação, com o surgimento de equipamentos biomédicos como a ressonância magnética e a tomografia computadorizada, propiciaram uma evolução nas teorias médicas acerca do funcionamento de alguns órgãos, em especial, o cérebro.

Os cientistas têm buscado entender o cérebro humano e criar modelos artificiais que possam imitar, com algum grau de qualidade, o mecanismo de visão natural.

Sistemas baseados em redes neurais [54, 87, 93], mapas auto-organizados [10, 72], modelos estatísticos [48, 50, 88], wavelets [57, 66], dentre outros,

vêm sendo estudados buscando aproximar os modelos artificiais ao sistema de visão humano.

A abordagem baseada em modelagem estatística é utilizada nesta pesquisa, pois, é a que apresenta melhores resultados dentre as técnicas citadas para a modelagem do sistema de visão [48]. De fato, descobertas no campo da neuro-ciência têm mostrado que as células do sistema de visão realizam filtragem estatística das imagens observadas [3, 5, 8, 13, 23, 24, 32, 42, 48, 50, 55, 76, 81, 82, 107].

As cenas naturais, geralmente, não são homogêneas e são razoavelmente ergódicas¹ (pelo menos para medidas estatísticas de baixa-ordem).

Sabe-se ainda que as frequências espaciais em texturas naturais não são idênticas, porém, a teoria da informação mostra que os métodos de filtragem mais eficientes são aqueles que melhor capturam as regularidades estatísticas dos dados.

As propriedades estatísticas de cenas naturais não são inteiramente estacionárias, por exemplo, em paisagens, o brilho médio da parte superior da imagem geralmente é maior que o brilho médio da parte inferior da imagem. A parte superior também possui menor variância.

É importante, aqui, efetuar uma distinção entre texturas naturais e imagens naturais. As *texturas naturais*, pertencem ao conjunto das imagens naturais, porém, são definidas assim devido à distância do observador à imagem (muito próximo ou muito distante), e à organização espacial entre os pixels em suas respectivas vizinhanças. A definição de textura será abordada mais à frente de maneira mais completa.

Todavia, embora exista uma grande variedade de sistemas para processamento de imagens voltados especificamente para texturas ou para imagens naturais, eles normalmente são baseados em bancos de filtros construídos a partir de funções de Gabor [43].

De fato, as funções de Gabor estão presentes no modelo de processamento de uma classe de células do primeiro estágio de processamento do sistema de visão humano, como demonstrado por Daugman [22].

Todavia, o modelo de visão humano é composto por vários estágios de processamento, contendo, cada qual, conjuntos de células com características próprias, inclusive com percepção de caracteres estatísticos em alta-ordem [48, 50, 65, 95] ou até mesmo com atributos fractais [5, 95].

É conhecido que o espectro de potências de imagens naturais possui uma distribuição sempre na forma $1/f$ (onde f é a frequência espacial) o que caracteriza uma distribuição estatística fractal [108].

¹Ergódica no sentido de que as estatísticas de uma imagem são uma boa estimativa para as estatísticas de um conjunto inteiro de imagens.



Figura 1.1: Regiões geometricamente auto-similares em diferentes escalas em uma imagem natural (observe a marcação branca em “V” nos caules da planta).

Intuitivamente, isso mostra que há objetos em diferentes escalas que podem ser vistos de uma certa distância em relação ao observador (Figura 1.1), ou ainda, que existem estruturas com comportamentos estatísticos similares em diferentes escalas na imagem (Figura 1.2). A Figura 1.3 apresenta o aspecto de uma distribuição $1/f$.

A existência de características fractais no sistema de visão humano permite especular o porque da facilidade humana em interpretar imagens da natureza, uma vez que estas são muito bem representadas por modelos fractais [78, 104]. Por outro lado, é conhecido o fato de que bancos de filtros baseados em funções de Gabor não são adequados para segmentação de imagens fractais [90].

De fato, a elaboração de modelos que possam representar os primeiros estágios de visão já é uma realidade. Sabe-se hoje que as células simples do córtex visual primário (V1) correspondem a um banco de filtros estatísticos de alta-ordem e independentes entre si [47].

Estudos de modelagem desse sistema como o desenvolvido por Hyvärinen et al. [47] e conhecido como ICA (*Independent Component Analysis*) quando submetido à representação de imagens naturais têm permitido a geração de filtros com funções próximas às de Gabor, correspondendo aos estudos de Daugman [22] para as células simples do córtex em V1. Embora estudos mais recentes mostrem que os filtros ICA possuem melhor sintonia que os filtros de Gabor [53].



Figura 1.2: Regiões estatisticamente auto-similares em diferentes escalas em uma imagem natural.

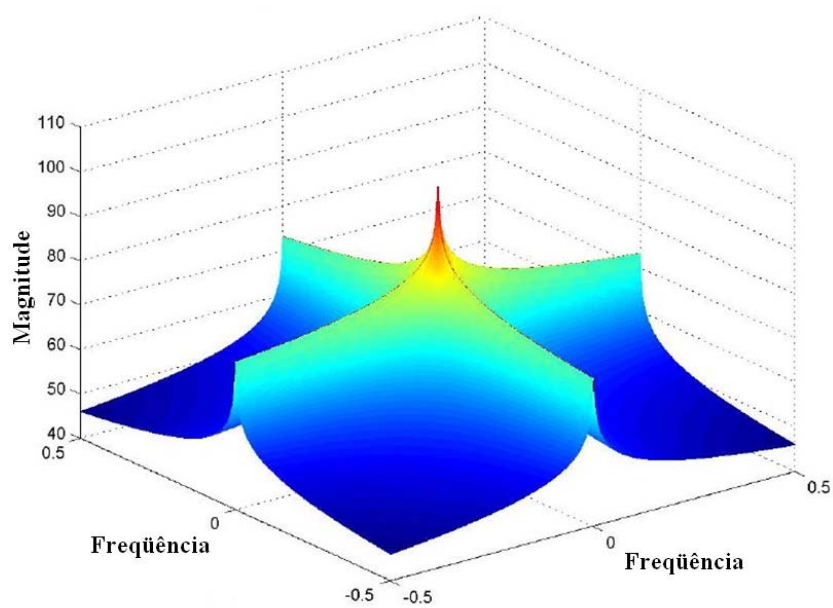


Figura 1.3: Exemplo de espectro $1/f$.

Todavia, o modelo ICA, na forma proposta na literatura, não leva em consideração as características fractais existentes nas imagens naturais. Isso porque essas características são predominantes nas células posteriores à primeira camada de células visuais, ou seja, elas existem em uma região do córtex visual denominada de *córtex visual secundário*, ou V2, localizado imediatamente depois da primeira camada de células do córtex visual, V1. Sabe-se, ainda, que a ICA atualmente é o modelo que melhor representa o comportamento das células simples do córtex V1 [47].

A construção de um modelo que possa abarcar grande parte dessas características poderia sinalizar uma aproximação do comportamento conhecido das células que compõem as primeiras camadas do córtex visual.

Esse, portanto, é um estímulo para este trabalho, ou seja, como propor uma conjectura para o modelo de representação estatística das células neuronais nas regiões V1 e V2 do córtex visual?

Sabe-se, todavia, que a separação de fractais estatísticos diretamente pelo uso da ICA mostrou-se inadequada, com o estudo elaborado por Potter e Kinsner [83]. Porém, o uso de fractais no critério topográfico mostrou-se satisfatório para realçar as características fractais nas componentes estatisticamente independentes do modelo ICA [18].

A explicação para a ineficácia da separação de fractais diretamente nas componentes independentes da ICA pode ser entendida no seguinte exemplo: a imagem da Figura 1.4 foi gerada a partir de um único fractal estatístico com distribuição $1/f^3$, porém, é possível efetuar visualmente uma separação grosseira entre as regiões claras e escuras, como marcado na imagem pelo retângulo tracejado.

Isso mostra que a segmentação visual pode dividir a imagem em regiões, o que permite afirmar que nem sempre um único fractal pode representar uma única região visual na imagem segmentada.

A questão é que tonalidades diferentes remetem a diferenças estatísticas de baixa-ordem (fora da região de transição). Porém, os modelos fractais estatísticos normalmente são definidos por estatísticas de ordens mais elevadas. Então, como o sistema de visão humano reconhece as mais variadas imagens existentes na natureza, uma vez que essas imagens são bem modeladas por fractais de alta-ordem estatística?

O comportamento do modelo biológico da visão humana motivou o seu estudo e inspirou a elaboração deste trabalho através da aproximação do comportamento das células no córtex visual secundário (V2), do cérebro humano, pelo uso de modelos fractais no critério topográfico da TICA.

O uso da TICA como modelo de base se deve à análise do comportamento dos neurônios em V2 e aos estudos que demonstram ser a ICA o melhor modo de implementar o modelo de redução de redundâncias da camada V1

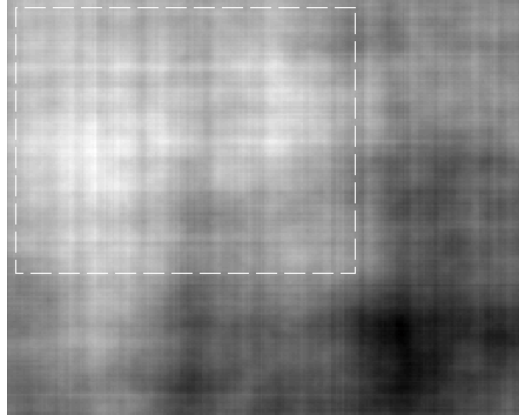


Figura 1.4: Regiões visualmente identificadas na imagem fractal.

do córtex visual humano [1].

Baseado nas características de modelagem das células em V1 do modelo ICA, na predominância de estruturas fractais das imagens naturais e no comportamento fractal das células em V2, surgiu, então, a motivação de adequar o modelo ICA para representar também o comportamento fractal conhecido das células em V2, com o intuito de permitir uma melhor adequação das imagens naturais quando representadas como bancos de filtros estatísticos e independentes em alta-ordem pelo modelo ICA.

A motivação em desenvolver técnicas de processamento de imagens, em especial, para texturas naturais se deve ao fato dessas imagens serem altamente variáveis e o nosso sistema de visão possuir a habilidade suficiente para reconhecer essas imagens.

Outra motivação para a escolha de um método estatístico como a ICA é que uma grande parte das texturas naturais são reconhecidas pelo sistema de visão humano como instâncias de um mesmo tipo de textura [51].

Isso permite dizer que uma importante parte da visão envolve a captura das variações estatísticas dessas texturas naturais. Como a ICA efetua a separação estatística de fontes e representa bem as células simples e complexas (critério topográfico) em V1, torna-se uma técnica adequada ao modelo de representação de imagens proposto nesta pesquisa.

Objetivo

De acordo com as motivações citadas anteriormente, o objetivo desta pesquisa é a criação de uma representação de imagens em tons de cinza, baseada no modelo TICA, mas que seja mais adequada à representação de imagens

naturais que a TICA.

Para isso, se propõe a inclusão de funções fractais estatísticas no critério topográfico desse modelo. A inclusão de fractais do modelo TICA, conseqüentemente, permite simular alguns dos comportamentos conhecidos das células do córtex visual secundário, V2, que possuem características fractais, como será abordado mais adiante.

A proposta de representação elaborada aqui é validada por experimentos de segmentação não-supervisionada em imagens, e os resultados obtidos são discutidos ao final deste trabalho.

A aplicação da segmentação permite visualizar as diferentes regiões caracterizadas pelos filtros estatísticos obtidos na modelagem proposta e comparar com as regiões obtidas pelos filtros do modelo existente. Uma vez que o resultado da segmentação permite comprovar a modelagem proposta, o resultado também valida o modelo de representação proposto.

O Sistema de Visão Humano

Nesta seção serão abordados alguns dos conceitos básicos do mecanismo de visão humano que inspiraram o trabalho desenvolvido e que serão necessários para a compreensão da modelagem proposta nesta pesquisa.

A recepção do estímulo visual nos seres humanos ocorre em uma região interna do olho denominada retina. A retina é um tecido sensível à luz capaz de convertê-la em impulsos elétricos que são enviados ao cérebro através do nervo óptico (veja Figura 1.5).

O estímulo elétrico é produzido nas células foto receptoras da retina, denominadas de cones e bastonetes (veja Figura 1.6), e se propaga para as células horizontais responsáveis por calcular a média dos estímulos gerados pelos cones e bastonetes [15].

Cada célula bipolar recebe entradas de uma célula fotorreceptora (cones e bastonetes) e de uma célula horizontal e produz um sinal que é proporcional à diferença logarítmica dos sinais produzidos pelas duas células, o que equivale à razão entre a intensidade local (proveniente da célula fotorreceptora) e a iluminação de fundo (proveniente da célula horizontal) [15].

Por isso, áreas grandes da retina, com iluminação uniforme, produzem sinais fracos de estímulo, enquanto áreas de maior variação, como o contorno de objetos, produzem sinais mais fortes, ou seja, a retina é sensível, basicamente, à variações de luminosidade da imagem [15].

Por outro lado, as células horizontais têm uma resposta mais lenta que as células fotorreceptoras de forma que, enquanto o estímulo visual varia rapidamente nos cones e bastonetes, as células horizontais ainda possuem

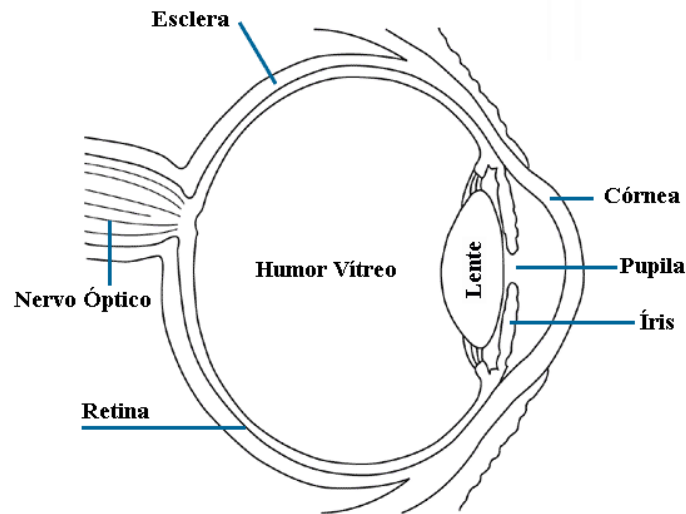


Figura 1.5: Esquema do olho humano onde pode-se observar a localização da retina e do nervo óptico [74]

informações de instantes anteriores, fazendo com que as células amácrinas, que recebem o estímulo das células bipolares, possam detectar o movimento da imagem [15].

Das células amácrinas, o estímulo visual para a ser conduzido pelas células ganglionares da retina (RGC - *Retinal Ganglion Cells*), que possui como uma de suas atribuições, a função de atuar como um relé controlando a abertura e fechamento da pupila.

Das células RGC, o estímulo passa pelas células do núcleo geniculado lateral (LGN - *Lateral Geniculate Nucleus*) que antecedem o córtex visual primário V1 (veja Figura 1.7).

Observa-se, então, que o processamento da visão inicia-se na retina, porém, a maior parte da análise ocorre no córtex visual, sendo que apenas algumas ações são executadas nas células que antecedem o córtex primário (veja Figura 1.8).

No córtex visual, os estímulos trafegam pelas diferentes camadas e em cada camada existe um conjunto de células especializadas em extrair diferentes características usadas para compor o processo de interpretação das imagens recebidas pela retina.

No córtex o processamento é dividido entre os diferentes grupamentos de células dispostas em cinco camadas bem definidas (V1, V2, V3, V4 e V5). É através do córtex primário, V1, que chegam os estímulos visuais ao cérebro

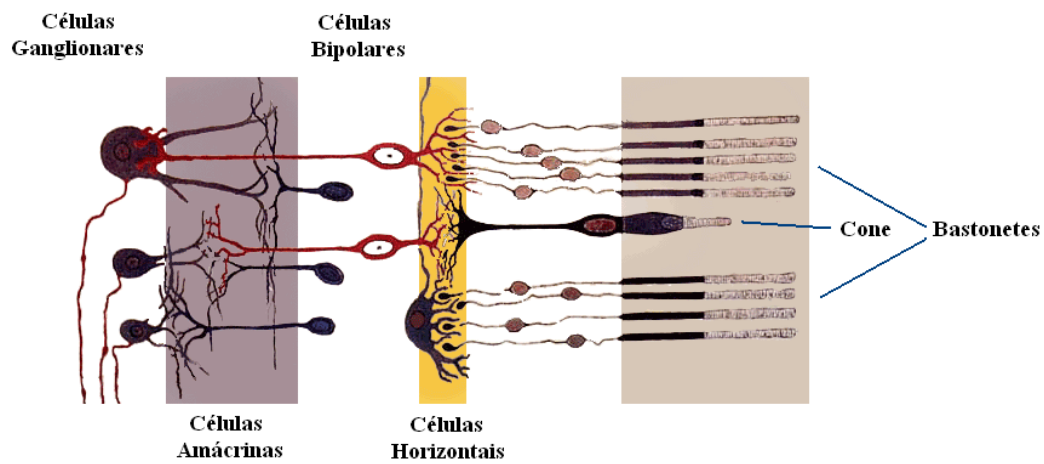


Figura 1.6: Células da retina [15]

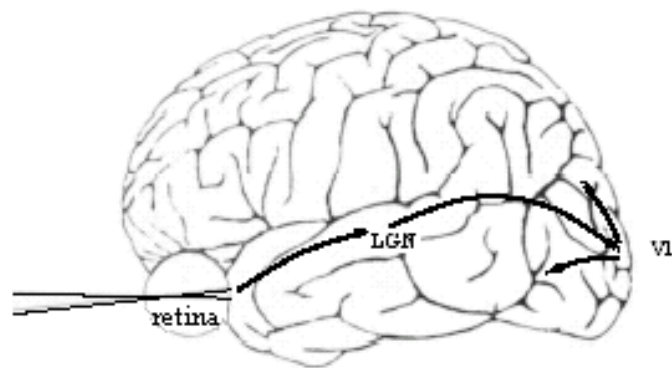


Figura 1.7: Trajeto do estímulo visual entre a retina e o córtex visual primário V1.

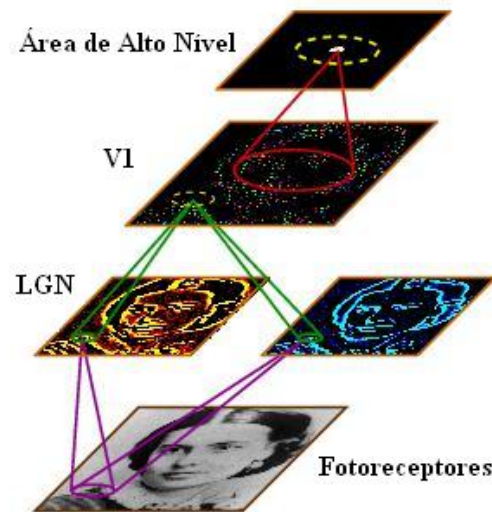


Figura 1.8: Exemplo de percepção dos estímulos visuais entre a retina e o córtex visual [102]

provenientes das células do núcleo geniculado lateral (veja Figura 1.7). É a região da visão mais amplamente estudada no cérebro.

O córtex estriado, V2, recebe forte estímulo de V1 possuindo muitas propriedades em comum com o córtex primário como sensibilidade à orientação e à frequência espacial.

Porém, em V2 ocorrem interpretações mais complexas que em V1, como a orientação de contornos ilusórios (veja Figura 1.9) onde o estímulo é proveniente apenas de parte da figura [29].

Pesquisas recentes mostram que as células em V2 respondem mais facilmente a padrões mais complexos que devem ser dirigidos por múltiplas orientações em diferentes sub-regiões com um simples campo receptivo [29].

Apesar dos esforços dos pesquisadores pouco se sabe acerca do comportamento das células em V2² e menos ainda acerca do comportamento das regiões superiores do córtex (V3, V4 e V5). Não há consenso, mas acredita-se existir uma sexta camada denominada V6 [34].

As células no primeiro estágio do processamento cortical (V1 e V2) são capazes de responder a um único estímulo de borda de uma maneira complexa, mas que permite a detecção do objeto mesmo enquanto suas definições mudam, ou estão parcialmente ocluídas (exemplo da Figura 1.9) [67].

²Todavia, as descobertas neurocientíficas são amplamente publicadas, sendo utilizados nesta pesquisa apenas os resultados tidos como consensuais entre os pesquisadores, conforme abordado nas diversas publicações referenciadas neste trabalho.

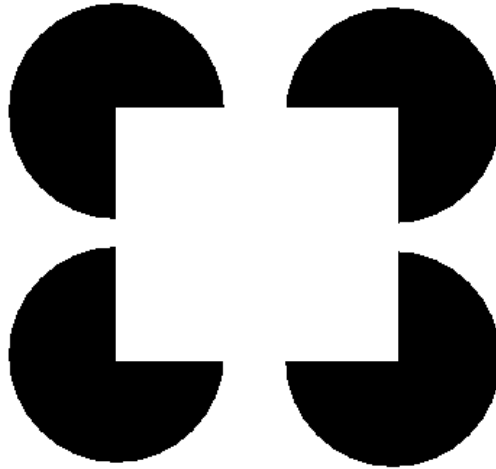


Figura 1.9: Nesta imagem o cérebro supõe que exista um quadrado que sobrepõe a figura interrompendo os círculos, porém, a imagem apenas possui círculos cortados para induzir a essa percepção.

Muitos pesquisadores vêm contribuindo para o entendimento do mecanismo de visão humano e cada vez mais descobertas e modelos são produzidos para buscar o aperfeiçoamento dos mecanismos artificiais de visão. Nos parágrafos seguintes estão relacionados alguns dos resultados de pesquisas e seus respectivos autores.

Leventhal et al. [67] demonstrou que uma sub-população de células nas regiões V1 e V2 podem responder a estímulos de fronteira (semelhantes a uma barra orientada) na forma de sinal invariante e que, embora essa propriedade seja rara em V1, ela é predominante em V2.

Hubel and Wiesel [46] demonstraram a presença de células no córtex V1 de primatas que respondem, preferencialmente, ao estímulo visual em ambos os olhos em oposição ao estímulo em um olho apenas e que células no LGN são exclusivamente monoculares.

Por outro lado, demonstraram que o córtex V1 é o primeiro local onde observa-se forte seletividade relacionada à orientação e à direção [46].

Enquanto a grande maioria das células em V1 mostram algum grau de seletividade em relação à orientação, apenas aproximadamente 25% a 30% das células em V1 são fortemente seletivas em relação à direção [24, 92].

Alguns trabalhos buscam modelar células do sistema de visão como: células simples em V1 [22, 42, 51], células complexas em V1 [49] e células de grade em V1 e V2 (*grating cells*) [45, 64].

Estudos acerca das áreas de ativação neuronal a diferentes estímulos, permitiram uma série de conjecturas, pautadas em experimentos, que buscam

determinar um modelo artificial do mecanismo de percepção humano das imagens visuais.

Várias pesquisas foram desenvolvidas com intuito de entender o comportamento do sistema de reconhecimento neural de imagens formadas na retina [2, 3, 4, 17, 23, 31, 32, 48, 50, 55, 81, 82, 95].

Parte II

Fundamentação Teórica

Capítulo 2

Texturas e Imagens Naturais

A caracterização de texturas é uma atividade de extrema importância na análise de imagens. As texturas estão presentes em nosso cotidiano através das imagens de satélites, imagens microscópicas, biomédicas e imagens da natureza.

Apesar de sua grande importância, não existe ainda um método geral para sua discriminação. Isso se deve em grande parte pela complexidade na caracterização do que seja uma textura.

Do ponto de vista etimológico, o termo tem sua origem do latim *textura* que quer dizer tecido, trama ou contextura [30]. Estes, por sua vez, levam a diferentes significados que reunidos em um mesmo contexto representam: um preparado, uma teia, uma estrutura, aquilo que prende, que enreda, que emaranha, uma organização, uma ligação entre as partes de um todo [30]. Portanto, do ponto de vista etimológico, pode-se inferir, de forma generalista, que:

Definição 1 *Uma textura é uma estrutura organizada que permite a construção de um todo a partir da ligação entre suas partes.*

Todavia, no contexto de processamento de imagens, é necessário ser mais preciso nessa definição. Do ponto de vista da modelagem matemática, a textura é uma expressão relacionada com as propriedades que representam a superfície de objetos. Para Sonka, Hlavac e Boyle [98],

Definição 2 *Textura pode ser definida como algo que consiste de elementos mutuamente relacionados. Um elemento refere-se a uma primitiva¹ de textura.*

¹Para Sonka *et al.* [98], uma primitiva de textura pode ser um pixel ou um conjunto de pixels. Para Jain [52], os níveis de cinza, a forma e a homogeneidade também são primitivas da textura.

Para Haralick *et al.* [39],

Definição 3 *Uma textura é definida através de coeficientes de uniformidade, densidade, aspereza, regularidade, intensidade, entre outras características da imagem.*

Para Sklansky [96], uma região em uma imagem apresenta uma textura única se um conjunto de características estatísticas ou outras propriedades locais da imagem forem constantes, com pouca variação ou aproximadamente periódicas.

Gonzalez e Woods [36] conceituam textura através de um descritor que fornece medidas de propriedades como suavidade, rugosidade e regularidade.

Alguns autores, como Jain [52], definem texturas como propriedades de uma região da imagem que descrevem o comportamento da variação tonal dos pixels em uma determinada área.

Nesse caso, a textura é caracterizada pela repetição de um modelo (texels) sobre uma região. Esse modelo pode ser repetido de forma exata ou com pequenas variações como: tamanho, formato, cor e orientação. A variação encontrada na maneira como os “texels” se relacionam permite diferenciar duas texturas.

Observa-se, portanto, que cada autor apresenta a sua definição. Isso ocorre porque não existe uma formalização única precisa, nem uma aproximação matemática formal para a caracterização de texturas.

Entretanto, esses autores caracterizaram a textura de maneira bidimensional, em que uma dimensão contém propriedades primitivas de tonalidade e a outra os relacionamentos espaciais entre essas primitivas, e ainda, verifica-se uma certa precedência que corrobora para a definição generalista formulada a partir do sentido etimológico do termo.

As texturas estão presentes em quase todos os objetos², o que torna imprescindível o desenvolvimento de mecanismos que permitam efetuar o seu reconhecimento e tratamento de forma automática nos sistemas de visão artificial.

A grande variedade de texturas com os mais variados comportamentos organizacionais (morfológicos, estruturais e estatísticos) dificultam muito o desenvolvimento de uma única técnica de processamento, bem como proporciona a criação de modelos de análise com alta complexidade e alto custo computacional.

²Não há textura no vidro transparente.

2.1 Texturas Naturais

Há uma grande variação das propriedades que caracterizam as texturas em suas diferentes realizações. Na natureza, por exemplo, existem vários tipos de materiais com uma diversidade muito grande de texturas, como as superfícies de mármore, granito, madeira, pele de animais, ou até mesmo nas imagens de cenas e paisagens, com folhagens, grama, florestas e águas.

Essa disparidade de características torna difícil definir, de forma estrita, o que são texturas naturais. Assim, costuma-se definir *texturas naturais* àquelas texturas que ocorrem na natureza e que não são sintetizadas pelo homem. Por exemplo, as texturas acima citadas e ainda, imagens biomédicas e astronômicas.

Entretanto, alguns autores buscam definir essas texturas através de características mais específicas. Jain [52], por exemplo, define *texturas naturais* como formas de construção, geralmente, aleatórias, enquanto *texturas artificiais* são freqüentemente determinísticas ou periódicas.

Apesar das diferentes abordagens, cada autor busca caracterizar texturas da forma que melhor convém ao entendimento da abordagem analítica usada em sua análise.

Um descritor de texturas adequado para essa pesquisa pode ser feito pelo número de propriedades primitivas que compõem a textura e pela organização espacial da mesma.

Essa organização espacial pode ser aleatória ou possuir uma dependência relacionada a uma vizinhança ou a todas as demais primitivas. Essa dependência pode ser estrutural, funcional ou probabilística [12].

Uma textura pode ser descrita qualitativamente tendo uma ou mais propriedades de suavidade, granulidade, rugosidade, aleatoriedade, linearidade e irregularidade, por exemplo [52].

Cada uma dessas propriedades se traduz em alguma propriedade das primitivas tonais e de interação espacial entre estas primitivas [12]. A tonalidade dos pixel de uma trama e a identificação dessa trama como textura são dois conceitos interdependentes, possuindo uma relação análoga à relação entre onda e partícula no modelo atômico [12].

Essa relação pode ser entendida da seguinte forma: quando uma área pequena de uma imagem possui uma variação pequena das primitivas tonais, a propriedade predominante desta área é o tom dos pixels.

Quando uma área pequena de uma imagem possui uma variação grande das primitivas tonais, a propriedade predominante desta área é a de trama.

Observe que o tamanho da área é crucial para essa classificação. Se a área se reduz a apenas um pixel, então há somente uma característica discreta presente, o tom do pixel. Assim, se o número de primitivas distinguíveis

diminui, predominam as propriedades tonais e se o número de primitivas distinguíveis aumenta, predominam as propriedades texturais.

O estudo de texturas pode envolver análises estatísticas e/ou estruturais, como já mencionado. A primeira é aplicada de forma genérica, enquanto a segunda é utilizada apenas se for possível extrair primitivas da imagem [12].

Para Jain [52] a análise estatística corresponde ao uso de transformadas, matriz de co-ocorrência, modelos de campos aleatórios e função de autocorrelação.

Para análise estrutural, Jain divide a abordagem em duas partes: a primeira é a análise estrutural periódica e consiste na determinação de primitivas e de regras de colocação, como período, adjacências e distâncias próximas; e a segunda é a análise estrutural aleatória e consiste no uso de densidades e comprimentos de corrida (ou deslocamento).

Essas análises não são mutuamente excludentes e em texturas naturais, por exemplo, a análise estatística é a que proporciona melhores resultados, enquanto que para a análise estrutural algumas propriedades estruturais podem necessitar de uma descrição estatística [12, 52].

Texturas puramente estruturais possuem “texels” determinísticos que se repetem de acordo com alguma colocação de regras, sejam elas determinísticas³ ou aleatórias⁴ [52].

2.2 Imagens Naturais e Informação Espectral

Texturas naturais compõem apenas um subconjunto das imagens existentes na natureza. Normalmente a denominação de textura natural obriga posicionar o observador a uma distância pequena da imagem, ou seja, em uma pequena escala da cena.

Contudo, a relação de aproximação é apenas um pré-requisito para a caracterização de uma textura, sendo necessário utilizar uma das abordagens definidas no início deste capítulo para reconhecer uma imagem como uma textura.

A relação entre a aproximação da cena em imagens naturais e artificiais e seus respectivos espectros de potências foi bem estudado por Torralba e Oliva [103]. Nesse trabalho, os autores observaram a existência de uma forte correlação entre diversas categorias de objetos em cenas naturais e a imagem em seu entorno e mostraram que a organização de partes e regiões de cenas naturais é muito menos restrita que em cenas de objetos artificiais, resultando em imagens que não são espacialmente estacionárias [103].

³Neste caso a textura é rotulada como *forte*.

⁴Neste caso a textura é rotulada como *fraca*.

Torralba e Oliva mostraram ainda, que tanto nas imagens naturais quanto nas elaboradas pelo homem, há uma predominância nas orientações verticais e horizontais das bordas dos elementos das imagens em relação às orientações oblíquas (Figura 2.1).

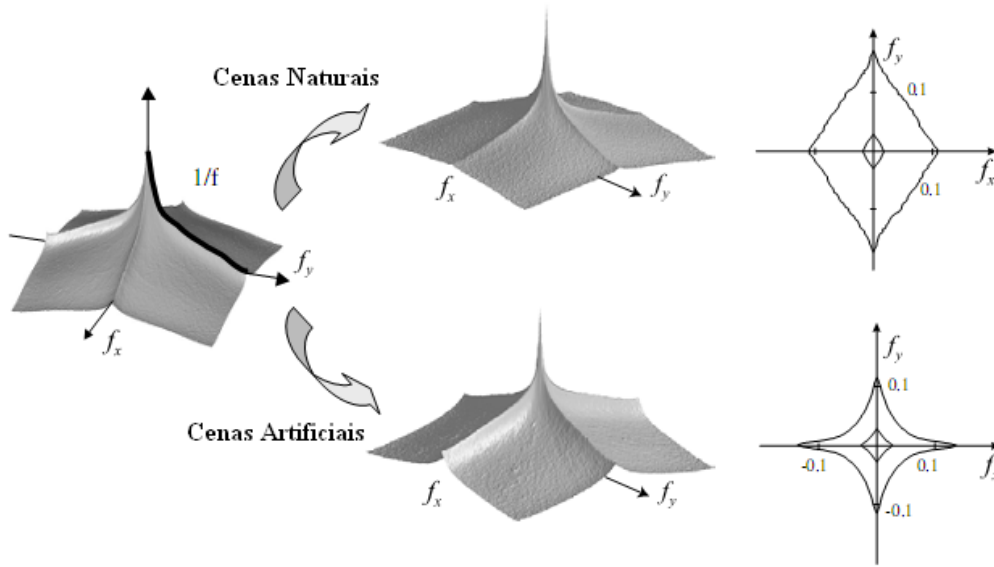


Figura 2.1: Exemplo de espectro $1/f$ para cenas naturais e cenas artificiais [103].

Apesar de imagens naturais serem estatisticamente redundantes [95], elas possuem uma certa regularidade estatística retratada pela forma do espectro médio de potências da imagem, que segue a distribuição $1/f^\lambda$, (onde f é a frequência espacial⁵) com $\lambda \approx 2$ (realce de bordas [95]) ou $\lambda \approx 1$ (superfície com mesma luminosidade, ou seja, todas as escalas⁶ possuem a mesma amplitude).

A forma do espectro médio de potências é uma função de orientação que, representada na forma polar, pode ser escrita como,

$$P[|I(f, \theta)|^2] \cong \frac{H(\theta)}{f^{\lambda(\theta)}}, \quad (2.1)$$

onde θ é a orientação da imagem I em relação ao observador, $H(\theta)$ é um fator de escala em amplitude para cada orientação, $\lambda(\theta)$ é um expoente de frequência (ou parâmetro espectral) como uma função da orientação e f é a frequência espacial da imagem I .

⁵A frequência espacial refere-se à informação em frequência resultante da Transformada de Fourier 2-D da imagem.

⁶Neste contexto, escalas determinam o tamanho da região observada.

O fator de escala e o expoente de frequência variam de acordo com a orientação das cenas e com o tipo de ambiente retratado na imagem, contribuindo com a forma do espectro de potências [7, 103].

A relação 2.1 determina a classe de funções $1/f$ que dão origem às distribuições α -estáveis⁷ e ao movimento Browniano fracional (*fractional Brownian motion* - fBm ⁸), ambas com auto-similaridades estatísticas (ou seja, representam fractais estatísticos [58]) [110]. O parâmetro espectral contém a dimensão fractal⁹ das distribuições fractais.

As Figuras 2.2 e 2.3 exemplificam o comportamento do espectro de potências e da assinatura espectral¹⁰ para imagens naturais e imagens de estruturas feitas pelo homem (artificiais), onde é possível observar a forma do espectro de potências como uma função da orientação e do ambiente (natural ou artificial).

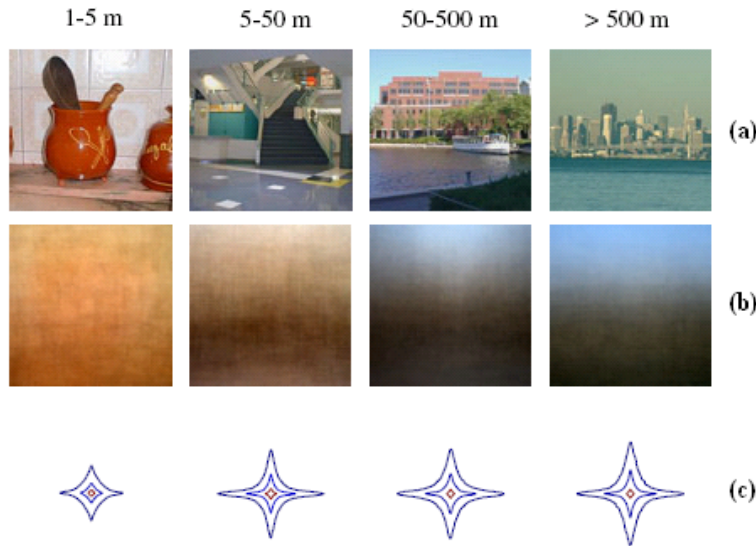


Figura 2.2: Relação entre a assinatura espectral e a distância de observação para imagens artificiais (ambientes humanos). (a) Exemplos de cenas artificiais em diferentes distâncias (b) Espectros de potências das imagens de exemplo (c) Assinatura espectral de 300 a 400 imagens semelhantes às dos exemplos [103].

⁷As distribuições α -estáveis, que serão comentadas adiante, dão origem a uma série de funções densidade de probabilidade, entre elas a distribuição gaussiana.

⁸O fBm é um fractal estatístico cujo parâmetro de auto-similaridade é restrito à segunda ordem estatística [26].

⁹A dimensão fractal será abordada mais adiante.

¹⁰A assinatura espectral corresponde à média do espectro de potências de um conjunto de imagens [103].

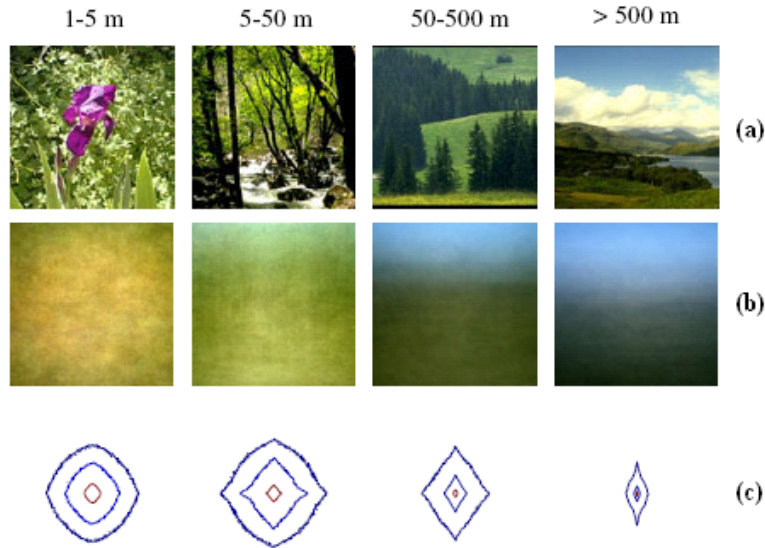


Figura 2.3: Relação entre a assinatura espectral e a distância de observação para imagens naturais. (a) Exemplos de cenas naturais em diferentes distâncias (b) Espectros de potências das imagens de exemplo (c) Assinatura espectral de 300 a 400 imagens semelhantes às dos exemplos [103].

As figuras enfatizam as diferenças entre imagens naturais e artificiais em diferentes faixas de escala. Visões próximas de objetos construídos pelo homem tendem a produzir imagens que são compostas por superfícies lisas e suaves e, conseqüentemente, a energia do espectro de potências fica concentrada, principalmente, em baixa frequência espacial.

Quando a distância entre o observador e a cena aumenta, o campo visual passa a compreender um largo espaço compondo vários objetos, então, a imagem percebida aparece como uma coleção de fragmentos ou de pequenas peças. Nesse caso, a energia das componentes espectrais de alta frequência torna-se maior devido à maior desordem produzida pelos vários objetos da imagem.

Nas imagens naturais, quando a distância entre o observador e a cena aumenta, a estrutura natural torna-se maior e mais suave (os pequenos “grãos” desaparecem). Portanto, na média, com o aumento da distância o nível de desordem decresce diminuindo a energia das componentes de alta frequência.

Entretanto, nas duas figuras, os padrões de orientação variam com a escala. Visões próximas nas estruturas naturais tendem a ser isotrópicas¹¹ em relação à orientação e o aumento da distância promove uma polarização nas orientações vertical e horizontal (principalmente devido às imagens de

¹¹Propriedades idênticas em todas as direções.

árvores e do horizonte) com uma concentração da energia principalmente nas frequências espaciais verticais [25, 103].

Todavia, em escalas muito altas a energia passa a se concentrar, principalmente, nas frequências espaciais horizontais nas imagens naturais [103].

A distribuição anisotrópica¹² das orientações é compatível com o modelo neurobiológico, no qual o número de células nos estágios iniciais do córtex (células em V1) varia em relação ao ajuste da orientação e da escala espacial (ex. a fovea possui mais células orientadas horizontalmente e verticalmente que de maneira oblíqua) [103]

Outra característica de extrema importância em imagens naturais é a forma com que as estatísticas da imagem mudam com a localização espacial (Veja Figuras 2.4 e 2.5). Isso ocorre porque, em imagens naturais, as localizações espaciais em uma vizinhança são fortemente correlacionadas pela intensidade dos pixels [95].



Figura 2.4: Efeito da distância da imagem no julgamento da cena. A imagem à esquerda é a imagem à direita invertida (espelhamento horizontal) [103]. Observe que a complexidade das cenas dificulta, em um primeiro momento, a percepção da orientação das imagens.

Assim, quando são consideradas todas as possíveis direções de observação, é de se esperar que as imagens naturais sejam invariantes à escala e estacionárias¹³ [32].

Esse é o caso de observações próximas à imagem, ou seja, em baixa escala. Em observações à distância o ponto de vista do observador exerce uma influência considerável na determinação das características estatísticas da imagem.

Estatísticas de segunda ordem são completamente descritas pelo espectro de potências de imagens naturais [25]. Todavia, estatísticas de alta-ordem

¹²Diferentes propriedades em diferentes direções.

¹³As características estatísticas da imagem são igualmente distribuídas em relação a localização.



Figura 2.5: Efeito da distância da imagem no julgamento da cena. Observe que é facilmente percebida, pelo observador, a inversão da imagem devido à distância de observação.

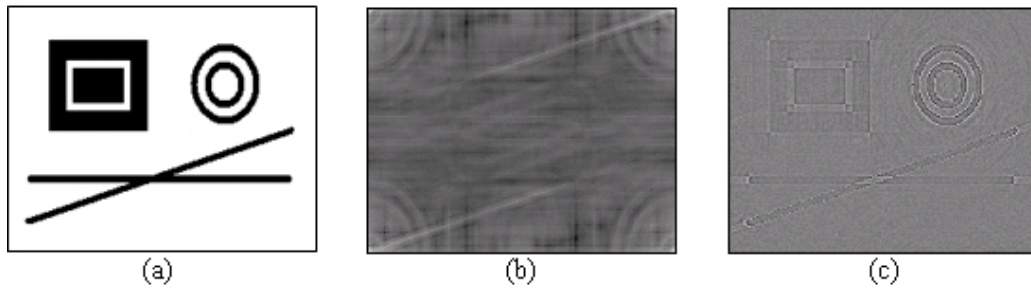


Figura 2.6: Exemplo de módulo e fase de Fourier. (a) Imagem de teste com figuras geométricas. (b) Informação do módulo de Fourier da imagem de teste. (c) Informação da fase de Fourier da imagem de teste.

residem no espectro de fases da transformada de Fourier e complementam o espectro de potências [25].

O fato é que o espectro de fases de Fourier carrega muito mais informações perceptivas relevantes da imagem devido à sua representação das informações de bordas, isso porque, é na localidade das bordas que as fases de diferentes frequências são alinhada [25].

A Figura 2.6 exemplifica o problema mostrando as informações de módulo e de fase de Fourier de uma imagem com figuras geométricas. Observe que utilizando apenas a informação de fase é possível reconhecer a imagem de teste.

Na Figura 2.7, é mostrada a reconstrução de imagens naturais a partir apenas das informações de Fase do espectro de Fourier. Observe que é possível reconhecer muitas das características da imagem original na imagem

reconstruída.

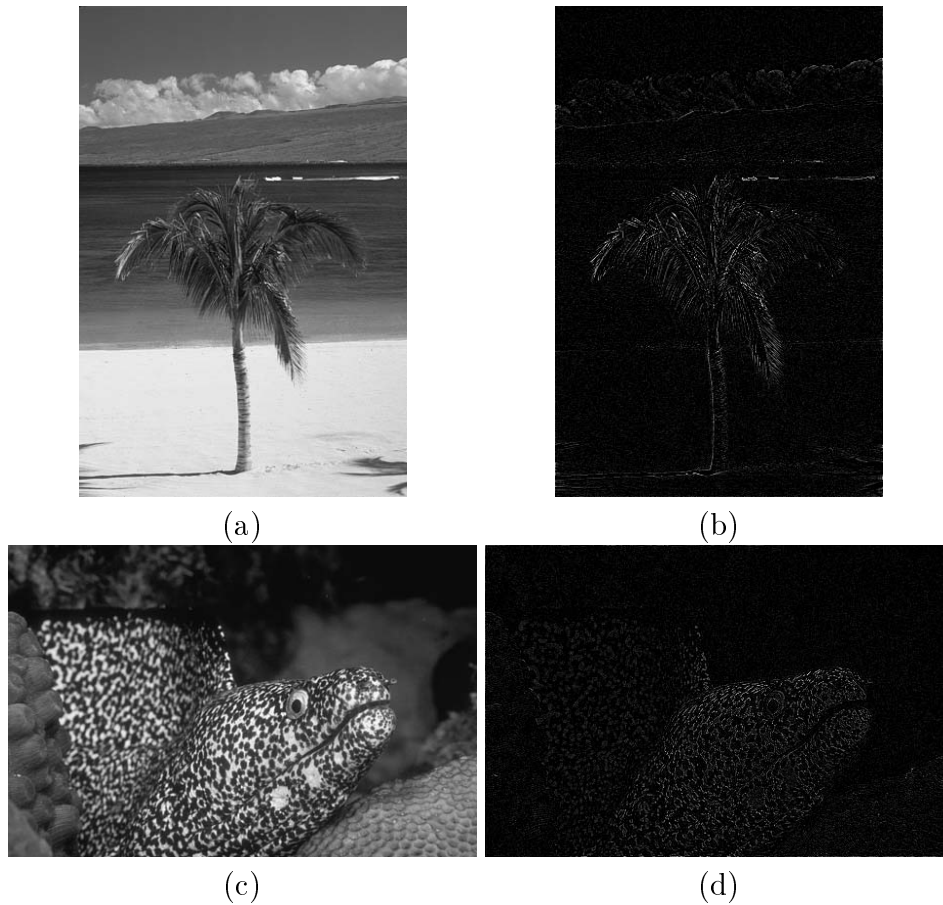


Figura 2.7: Reconstrução de imagem a partir apenas da informação de fase do espectro de Fourier. (a) Imagem original, (b) Imagem reconstruída, (c) Imagem original, (d) Imagem reconstruída.

Uma abordagem mais detalhada pode ser vista no trabalho de Doi e Lewicki [25]. Os autores mostram que a síntese de imagens a partir de um mesmo espectro de fases de uma imagem natural e diferentes espectros de potências permite a identificação visual da imagem sintetizada com a imagem original.

Quando o espectro de potências da amplitude segue a lei $1/f$ e a fase da imagem original é preservada, a imagem sintetizada é muito próxima da textura natural original, sugerindo que a informação exata de amplitude possui um papel menor na percepção da textura original a partir da textura sintetizada [25].

Assim, como apresentado nos estudos em [25, 47], o espectro de fases de texturas naturais contém muito mais informação relevante que o espectro

de potências, como as informações de borda, sua localização e as fases das diferentes frequências que compõem a imagem.

2.3 Texturas e a Percepção Humana

Em nível estatístico é possível definir uma textura através de medidas estatísticas extraídas de um conjunto de propriedades locais da imagem [12]. Julesz, conforme citado em [84], na década de sessenta, procurou determinar os parâmetros de texturas com importância perceptiva para o ser humano.

Sua pesquisa mostrou que o sistema de visão humano somente consegue distinguir texturas que possuam pequenas diferenças pelo menos na primeira e na segunda ordem estatística.

A Figura 2.8a mostra um mosaico entre duas texturas sintéticas diferentes apenas em altas ordens estatísticas, onde não é possível perceber a distinção entre as duas texturas, e um mosaico de texturas com diferentes variâncias (Figura 2.8b), onde se observa a região de cada textura.

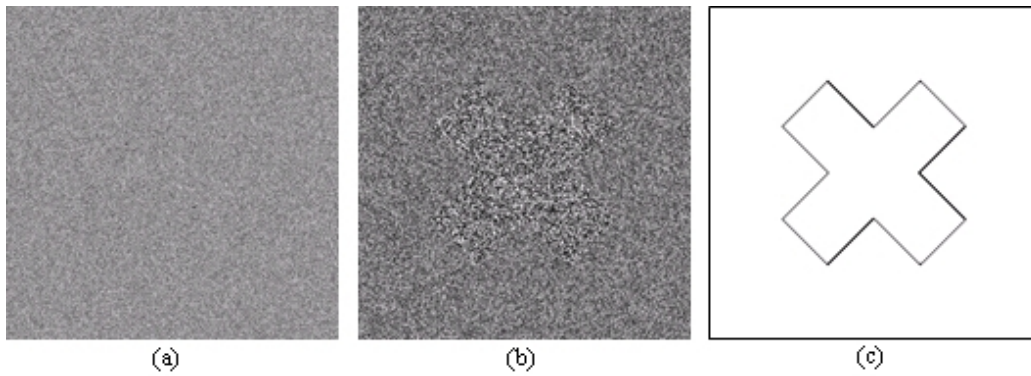


Figura 2.8: Texturas estatísticas. (a) Mosaico das texturas com médias $\mu_1 = \mu_2 = 0.1$, variâncias $\sigma_1^2 = \sigma_2^2 = 1$, skewness $skew_1 = 0.1$ e $skew_2 = -0.1$ e kurtosis $kurt_1 = 2$ e $kurt_2 = 4$. (b) Mosaico das texturas com médias $\mu_1 = \mu_2 = 0.1$, variâncias $\sigma_1^2 = 1$ e $\sigma_2^2 = 1.6$, skewness $skew_1 = skew_2 = 0.1$ e kurtosis $kurt_1 = kurt_2 = 2$. (c) Borda de separação entre texturas nos mosaicos.

Estudos elaborados mais tarde por Pratt e Faugeras, conforme citação em [84], vieram enriquecer o trabalho de Julesz mostrando que o sistema de visão humano é sensível não somente às diferenças na média e variância, mas também na autocorrelação entre os pares de texturas.

Descobertas mais recentes sugerem que o cérebro humano efetua análise frequencial multicanal e análise de orientação das imagens visuais formadas na retina [89].

Assume-se, de forma geral, que os filtros fisiológicos são funções lineares com respostas estatisticamente independentes e espacialmente balanceadas, aproximando-se das células simples do córtex e com forma espacial circular ou ligeiramente alongada [47].

Isso também foi estudado por Daugman [22] que, em seus experimentos, mostrou que os campos receptivos 2-D medidos para células simples do córtex de um gato, células em V1, são aproximados por funções de Gabor, representados em múltiplas frequências com informação de fase e de localização (Figura 2.9).

Durante algum tempo esses resultados motivaram a modelagem de texturas pelas funções de Gabor, o que, a princípio, permitiria a construção de classificadores capazes de esgotar os limites da percepção humana, modelando quase que perfeitamente o seu sistema de visão. Entretanto, naquela época, muito pouco se conhecia acerca das funções das outras camadas do córtex visual (V2 a V5).

Contudo, Pratt, Faugeras e Gagalowicz [85] já haviam mostrado que o modelo de visão humano não é sempre adequado à classificação de texturas, uma vez que não permite a distinção entre texturas que possuem pequenas diferenças apenas em altas ordens estatísticas (Figura 2.8a).

Imagens com diferenças pronunciadas em alta-ordem estatística podem ser percebidas pelo sistema de visão humano em um mecanismo de esforço atento (se a diferença entre as estatísticas das texturas for pequena, mas suficiente para a distinção), ou pré-atentivo, quando as diferenças são maiores. Entretanto, o mecanismo atento depende de informação temporal, ou seja, a imagem precisa ser observada ao longo do tempo [94].

Estudos mostram que no esforço atento há um aumento das atividades de algumas células em V1, relacionadas à atenção, porém, em V4 ocorre um aumento muito grande das atividades celulares [86]. Em alguns experimentos, não houve variações nas atividades em V1, durante o esforço atento, mas em V4 as mudanças ocorreram em mais de 50% das células [86].

Acredita-se que o fato das células em V1 possuir uma forte realimentação das células em V2 contribua com o mecanismo de interpretação dos estímulos no processo de esforço atento [86].

Retomando a questão, qual seria a finalidade de mecanismos que detectassem pequenas diferenças em alta-ordem estatística? A resposta está relacionada a uma série de motivos para a existência desses sistemas:

- A detecção de falhas na trama da textura em objetos cuja percepção visual não identifique, mas que possam comprometer a resistência ou a durabilidade do material. Ex: trincas podem não ser visíveis em uma textura, mas são perceptíveis aos sistemas estatísticos de alta-

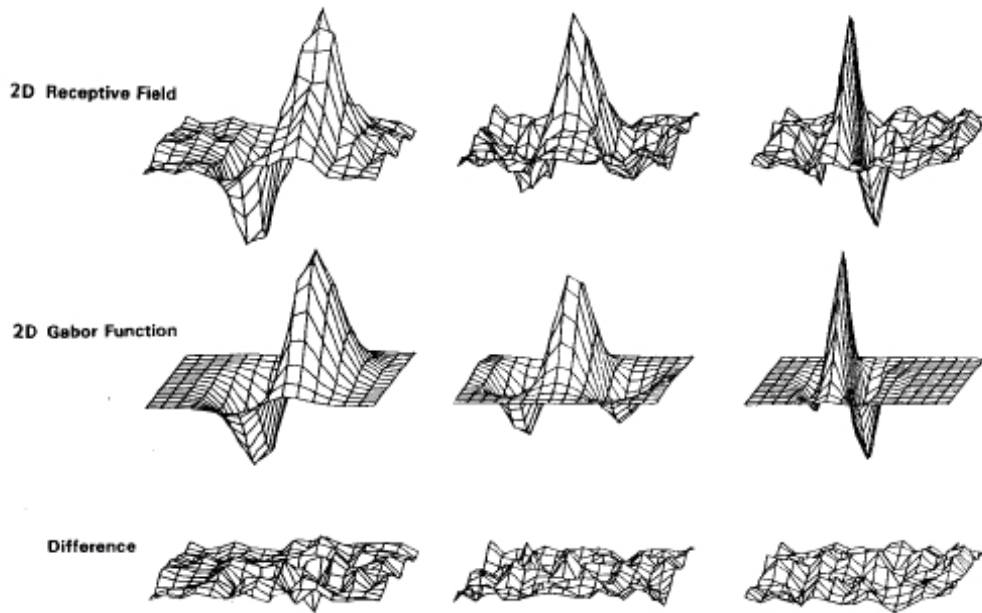


Figura 2.9: Campos receptivos 2-D das células em V1 do córtex cerebral de um gato são funções de Gabor [22].

ordem, que percebem as trincas como bordas na textura diferenciando um objeto sem trincas com outro defeituoso, mas de mesma aparência visual.

- Entender a gênese de estruturas naturais através da observação da textura de vegetais ou tecidos animais em diferentes estágios de desenvolvimento. Visivelmente semelhantes as texturas podem conter diferenças estatísticas de alta-ordem ao longo da sua formação que possam contribuir para a compreensão da estrutura em desenvolvimento ou da evolução de uma patologia.
- Ilusões de ótica podem modificar a percepção de cores e de espaço tornando a percepção de uma mesma imagem diferente em diferentes condições de baixa-ordem estatística como iluminação ambiente e disposição espacial de cores (Figura 2.10 e Figura 2.11) [37, 54]. O mecanismo cognitivo que torna possível o surgimento de ilusões vem sendo amplamente estudado [37, 54]. Sistemas baseados em estatísticas de altas ordens são menos influenciados por essas condições, desde que não haja inclusão de falsas bordas na imagem devido à iluminação [6, 81].

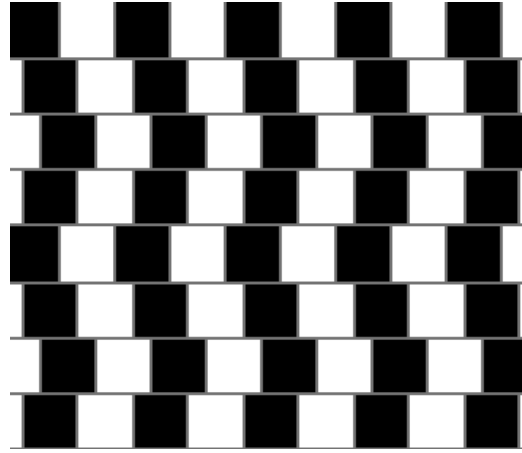


Figura 2.10: As linhas horizontais da figura são paralelas, porém a disposição espacial das cores causam a impressão de que as linhas não são paralelas.

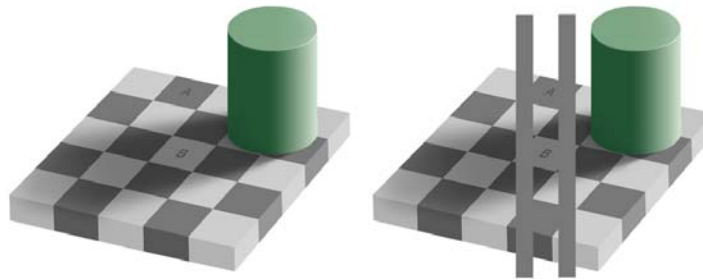


Figura 2.11: Os quadrados marcados com A e B possuem a mesma cor, porém, devido à disposição espacial das cores e à sombra do cilindro, eles aparentam possuir cores diferentes [79].

- Um sistema de alta-ordem estatística que modela o comportamento das células V2 do córtex (que são de alta-ordem estatística) permite estudar e entender melhor o comportamento do mecanismo de visão humano.

Uma analogia ao mecanismo de distinção de Pratt, Faugeras e Gagalowicz pode ser feita com o surgimento dos equipamentos óticos. Até a invenção do microscópio ou do telescópio, pouco ou quase nada se sabia acerca da existência dos microorganismos e do cosmos como é conhecido hoje. Os ambientes visualizados com esses equipamentos não são perceptíveis ao sistema de visão humano naturalmente, mas eles existem e possuem uma gama enorme de informações que ainda estão sendo descobertas.

Retomando a questão da modelagem de texturas, não somente o córtex primário (V1) é o responsável pelo reconhecimento de padrões visuais pro-

cessados nas diferentes camadas do sistema de visão. As células simples e complexas de V1 efetuam apenas uma etapa do processamento, sendo as células das camadas V2 a V5 responsáveis por completar o complexo mecanismo de interpretação das imagens.

Em relação à questão de estatísticas de baixa-ordem, a distribuição gaussiana poderia caracterizar adequadamente as informações de segunda ordem estatística determinadas pelo espectro de amplitude das texturas naturais, porém, sabe-se que essas imagens, em geral, não apresentam a função de distribuição de probabilidade gaussiana [31, 32, 95].

Assim, torna-se necessário encontrar uma função não-gaussiana que permita a representação das estatísticas de alta-ordem. Porém, a explicação para o uso de funções de Gabor por Daugman e Jones [23, 55], para a captura de regularidades estatísticas de alta-ordem nas células em V1, decorre de uma aproximação da representação de imagens naturais pela composição de funções que descrevem essas estatísticas através das informações de fase das funções de Gabor.

Logo, estatísticas de ordem superior, como a distância dos deslocamentos de níveis de cinza ou o espectro de fase de Fourier, também podem ser medidos e usados para a classificação de texturas naturais [12]. Entretanto, o uso de funções não-gaussianas para a captura de estatísticas de alta-ordem é reforçado com o estudo apresentado por Jenssen [53] que compara a segmentação de texturas entre ICA e funções de Gabor, e mostra que a ICA permite melhores resultados, ou seja, melhor caracterização das informações estatísticas de alta-ordem.

Sabe-se ainda que as funções de Gabor não são adequadas para imagens com características fractais [90]. Como será visto mais adiante, as imagens naturais são ricas em estatísticas de alta-ordem e são muito bem representadas por modelos fractais estatísticos.

A ICA tem mostrado que a modelagem de texturas naturais com estatísticas de alta-ordem não gaussiana é possível e que as funções dos filtros no modelo ICA correspondem aos campos receptivos das células simples, em V1, de estrutura localizada e orientada [11, 51, 70].

Apesar da ICA buscar pela representação estatisticamente independente, os resultados obtidos em texturas naturais exibem uma dependência residual, que significa uma necessidade de organização espacial entre as componentes [49, 59]. Sabe-se ainda que o modelo ICA permite a modelagem apenas das células simples (lineares) da classe de células em V1, não sendo possível modelar, diretamente, as células complexas que possuem as propriedades de invariância à fase e ao deslocamento [47].

Nesse sentido, uma variante da técnica ICA conhecida como Análise Topográfica de Componentes Independentes (TICA - *Topographic Independent*

Component Analysis), permite a aproximação do modelo ICA ao sistema de visão do córtex primário de maneira mais precisa [49].

Uma vez que a TICA insere uma organização espacial no modelo ICA, o que pode ser vista como uma organização topográfica das células simples em V1 [49], ela estabelece um modelo de relação de dependência residual¹⁴ na estrutura das células do córtex [49], buscando encontrar subespaços com características estatísticas invariantes em alta-ordem estatística [47].

O critério topográfico, que será melhor abordado mais à frente, é visto como uma extensão do modelo ICA, incluindo outras variáveis de alto nível que servem como parâmetros de organização espacial capturando as dependências de variâncias entre os coeficientes da ICA [59, 60]. Esse modelo representa o conjunto de células complexas em V1 responsáveis pelo controle do ganho e invariância à fase [49].

¹⁴Na proposta desta pesquisa é mostrado que essa relação de dependência não é meramente residual, quando se utiliza modelos fractais para a função de vizinhança.

Capítulo 3

Análise de Componentes Independentes

3.1 Estatísticas de alta-ordem

Antes de abordar as características da Análise de Componentes Independentes, é importante entender o significado do termo *estatísticas de alta-ordem*, que é utilizado nesta análise.

Segundo Therrien [101] *estatísticas de alta-ordem* (HOS - *Higher Order Statistics*) ou *espectros de alta-ordem* (HOS - *Higher Order Spectra*) são as medidas estatísticas que possuem momentos de ordem superior a dois. Análises em que se utiliza até o segundo momento estatístico são normalmente suficientes para extrair as características mais importantes da maioria dos processos aleatórios.

Todavia, se o processo não é gaussiano, (por exemplo, quando a fase de um processo aleatório é importante) medidas do segundo momento estatístico (função de autocorrelação, densidade espectral de potências e função densidade espectral complexa) não carregam informações de fase do processo aleatório.

Nesse caso, o uso de estatísticas de alta-ordem se faz necessário como uma boa aproximação do processo aleatório, uma vez que a obtenção da função densidade de probabilidade nem sempre é possível, seja pela complexidade do processo, seja por ser sua obtenção impraticável.

Os momentos estatísticos permitem, em muitos casos, descrever completamente o comportamento de variáveis aleatórias, uma vez que a partir de todos os momentos é possível reconstruir a função densidade de probabilidade do processo [101]. Como nem sempre é possível obter todos os momentos (infinitos momentos), uma aproximação da função densidade não apenas

permite um método de análise tratável no domínio do sinal, como também permite métodos tratáveis no domínio da frequência, através de extensões da Análise de Fourier para processos aleatórios estacionários [101].

A equação geral dos momentos para um processo aleatório estacionário no sentido amplo é definida por Stark [100] como segue.

Definição 4 *O momento de ordem r de uma variável aleatória discreta V é definido como:*

$$M_V^r \triangleq E[V^r] = \sum_i v_i^r P_V(v_i). \quad (3.1)$$

onde, E é o valor esperado, i refere-se a uma amostra da variável aleatória, r é a ordem do momento e P_V é a função probabilidade de massa (Probability Mass Function - PMF).

Uma vez que o processo aleatório é estacionário a função densidade de probabilidade é uma função apenas de espaçamento entre as amostras e não de localizações absolutas [101].

Portanto, $M_V^0 = 1$, $M_V^1 = \mu$ (a média constante) e $M_V^2[n_1 - n_0] = E[v(n_1)v^*(n_0)] = E[v^*(n)v(n+l)] = R_v[l] = M_V^2[l]$ (a função de correlação), a qual depende apenas de um único argumento l , denominado *lag* (atraso), que é a diferença entre n_1 e n_0 (localizações das amostras).

Embora os momentos representem satisfatoriamente as características do processo aleatório, geralmente, momentos de altas ordens correspondem a correlações usadas em estatísticas de baixa-ordem. Ou seja, momentos de altas ordens possuem medidas de baixa ordem estatística (veja Equação 3.3).

Por isso normalmente é preferível trabalhar com quantidades denominadas de *cumulantes*, por serem partes contáveis de alta-ordem das covariâncias e que exibem, cumulativamente, informações adicionais trazidas dentro das estatísticas de altas ordens [51, 101].

Em outras palavras, os momentos carregam características estatísticas das ordens inferiores a eles, por exemplo, um momento de 3ª ordem pode possuir estatísticas de 1ª e 2ª ordem, além da de 3ª ordem. Já os cumulantes possuem apenas as características estatísticas de sua ordem, ou seja, um cumulante de 3ª ordem possui apenas estatísticas de 3ª ordem.

Os cumulantes são obtidos a partir da expansão em série de Taylor da função característica de uma variável aleatória V . A função característica, por sua vez, é obtida através da Transformada de Fourier da função densidade de probabilidade de V [51, 80, 101]. Assim, o cumulante de ordem r é obtido por:

$$C_r = (-j)^r \left. \frac{d^r \ln(E[e^{j\omega V}])}{d\omega^r} \right|_{\omega=0}, \quad (3.2)$$

onde, ω é a frequência no domínio da transformada.

As quantidades espectrais relacionadas aos cumulantes são denominadas de “espectros cumulantes” (*cumulant spectra*) ou “espectro de alta-ordem” (*higher order spectra*) ou “poli-espectro” (*polyspectra*). Os espectros de segunda ordem são denominados de “espectros de potência ordinários” (*ordinary power spectrum*), os de terceira e quarta ordem são chamados de bi-espectros (*bispectrum*) e tri-espectro (*trispectrum*), respectivamente.

O espectro de alta-ordem pode ser definido como a transformada de Fourier de estatísticas de alta-ordem e que retém informações de fase do sinal [51].

A relação entre cumulantes e momentos, até a quarta ordem estatística, é dada por:

$$\begin{cases} M_V^1 = C_1 \\ M_V^2 = C_2 + C_1^2 \\ M_V^3 = C_3 + 3C_2C_1 + C_1^3 \\ M_V^4 = C_4 + 4C_3C_1 + 3C_2^2 + 6C_2C_1^2 + C_1^4. \end{cases} \quad (3.3)$$

Quando a variável aleatória possui média zero, os três primeiros cumulantes são iguais aos respectivos momentos [51]. Essa semelhança entre momentos e cumulantes quando a média é zero é motivo de uma série de confusões quando afirmam alguns autores que a distribuição gaussiana possui apenas momentos estatísticos até a segunda ordem, o que não é verdade.

A confusão se faz porque a distribuição gaussiana possui apenas cumulantes diferentes de zero nas duas primeiras ordens estatísticas e zeros nas demais ordens, enquanto existem momentos pares de ordem elevada que são diferentes de zero [77, 80, 101].

Esse esclarecimento permite entender melhor porque os cumulantes são desejáveis como medidas estatísticas de altas ordens, quando se quer trabalhar com processos aleatórios de altas ordens estatísticas, ou seja, processos aleatórios não-gaussianos.

Partindo dessa premissa, Therrien [101] afirma que a medida dos cumulantes de um processo aleatório é feita a partir do processo aleatório gaussiano, ou seja, para estudar um processo aleatório de alta-ordem estatística basta excluir da análise as componentes da distribuição gaussiana. Esse princípio é a base da Análise de Componentes Independentes que será vista na próxima seção.

3.2 A Análise de Componentes Independentes

Quando se aplica a Análise de Componentes Principais (PCA - *Principal Component Analysis*) em imagens, normalmente assume-se que as propriedades das imagens são invariantes à translação, ou seja, que as imagens são estacionárias [44]. Dessa forma, a correlação das intensidades em duas localizações de uma imagem dependem apenas do caminho entre essas localizações e não de cada localização absoluta [95]. Por esse motivo, a transformada de Fourier sempre garante um conjunto válido de eixos componentes principais em que a variância, ao longo de cada um dos eixos, é o espectro de potências de Fourier [95].

A existência de um conjunto válido de eixos permite salientar que sempre será possível efetuar o *branqueamento* em imagens, ou seja, estimação das características estatísticas de baixa-ordem.

Isso porque o *branqueamento* pode ser efetuado pela transformada de Fourier. O procedimento consiste em dividir cada componente em frequência da transformada pela raiz quadrada de sua variância seguida da transformada inversa de Fourier [95].

Outra forma de efetuar o *branqueamento* é através da diagonalização de matriz, como definido do teorema abaixo [100]:

Teorema 1 (*Branqueamento*) *Seja $P_{n \times n}$ uma matriz simétrica real, definida positiva, onde ϕ_i e γ_i (para $i = 1, \dots, n$) são os autovetores e autovalores de $P_{n \times n}$, respectivamente. Seja $U = (\phi_1, \dots, \phi_n)$ a matriz de autovetores de P e $M \triangleq U^T P U = \text{diag}(\gamma_1, \dots, \gamma_n)$, a matriz de autovalores de P .*

Sendo $P_{n \times n}$ definida positiva, todos os autovalores são positivos e a matriz diagonal $Z \triangleq \text{diag}(\gamma_1^{-1/2}, \dots, \gamma_n^{-1/2})$ existe e é real.

*Assim, a transformação de similaridade $(UZ)^T P (UZ)$ não apenas diagonaliza P , mas também reduz P a uma matriz identidade. Este processo é chamado de *branqueamento*, e a matriz $W = UZ$ é a matriz de *branqueamento* de P .*

Apesar da PCA ser usada para recuperar um conjunto de eixos ortogonais estatisticamente independentes que representem dados gaussianos¹, essa técnica falha quando os dados possuem estatísticas de alta-ordem, ou seja, não-gaussianas [95].

¹A PCA estima a média e a variância de maneira não-supervisionada da mistura, mesmo esta não possuindo distribuições gaussianas. Entretanto, devido à estimativa das grandezas de baixa ordem estatística serem suficientes para definição de uma aproximação por distribuição gaussiana, utiliza-se dizer que a PCA representa dados gaussianos.

A explicação é simples, quando os dados são uma mistura linear de fontes não-gaussianas, uma rotação adicional no sistema de coordenadas se faz necessária para recuperar os eixos originalmente independentes. Todavia, a rotação apropriada somente pode ser estimada observando-se as propriedades estatísticas dos dados além da covariância (i.e. ordem estatística maior que dois) [95].

Os algoritmos que buscam determinar essa orientação procuram maximizar os momentos de alta-ordem estatística, decompondo as amostras em componentes tão estatisticamente independentes quanto possível. Esses algoritmos compõem o que é denominado de Análise de Componentes Independentes.

O problema da separação estatística de alta-ordem é que não há garantias de que as componentes resultantes sejam efetivamente independentes, a menos que se garanta uma mistura linear de fontes de alta-ordem estatística (ou seja, fontes que sejam funções de distribuição de cauda-longa) [95].

Portanto, podemos formalizar o conceito da ICA como uma técnica de separação cega de fontes que utiliza estatística de alta-ordem para obtenção de um sistema linear de mistura de componentes tão estatisticamente independentes quanto possível.

Para imagens representadas no modelo ICA, a mistura X é a informação à priori e é composta por N janelas aleatórias da imagem representadas como colunas $\vec{x} = [x_1, x_2, \dots, x_M]^T$ da matriz X . As matrizes A e S devem ser estimadas maximizando a independência estatística de S . Matematicamente, o modelo é expresso como:

$$X_{N \times M} = A_{N \times P} S_{P \times M}, \quad (3.4)$$

onde, $A_{N \times P}$ é a matriz de mistura, inversível² ($W = A^{-1}$), $S_{P \times M}$ é a matriz de componentes tão estatisticamente independentes quanto possível em alta-ordem estatística e N , M e P são as dimensões das matrizes.

A matriz de mistura $A_{N \times P}$ fornece a maneira como as componentes independentes se misturam, linearmente, para formar a matriz $X_{N \times M}$. Essa combinação linear pode ser representada através de banco de filtros. Nesse caso, uma imagem monocromática é representada no modelo como:

$$I(x, y) = \sum_{i=1}^P a_i(x, y) s_i, \quad (3.5)$$

em que I é uma imagem cinza, a_i são os vetores que compõem a matriz de mistura, s_i são as componentes independentes da imagem (ou coeficientes

²Note que quando $N \neq P$, W é obtida pela pseudo-inversa da matriz A .

estocásticos) e x e y são as coordenadas dos pixels da imagem.

A Figura 3.1 ilustra o modelo ICA representando as funções da base dos filtros através de imagens.

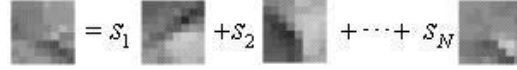


Figura 3.1: Modelo ICA para representação de imagens. A imagem é obtida pela soma do produto das componentes independentes s_i pelas componentes da base, obtida da matriz de mistura.

Cada imagem da Figura 3.1 representa um filtro da base (não ortogonal) da imagem (mistura) representada pela ICA e as componentes independentes representam o conjunto de funções densidade de probabilidade marginais. Assim, para uma imagem I e um conjunto de N filtros espaciais $\{h_i\}$ centrados em (x, y) , diz-se que a matriz

$$F(x, y) = [f_1(x, y) = I * h_1(x, y), \dots, f_N(x, y) = I * h_N(x, y)], \quad (3.6)$$

formada pela convolução da imagem com os diferentes filtros, é a representação de característica da imagem no ponto (x, y) [70].

Assume-se que uma textura é completamente representada pela função densidade de probabilidade conjunta da matriz de características $F(x, y)$ (que por estacionariedade, é independente de x, y) [70].

Outra maneira de abordar o modelo de Análise de Componentes Independentes é visualizá-lo como o problema de encontrar a função $g(\vec{x})$, em que $\vec{x}$ é uma variável aleatória m -dimensional, para que o espaço transformado n -dimensional $\vec{s} = [s_1, s_2, \dots, s_N]^T$, definido por $\vec{s} = g(\vec{x})$ tenha algumas propriedades desejadas, como a independência estatística em altas ordens e variância unitária [51].

A separação cega de fontes pode ser efetuada pela otimização da *função contraste*, i.e. uma medida escalar de alguma propriedade da distribuição da variável $\vec{s}$ [16]. As funções de contraste são baseadas na entropia, na independência mútua, na decorrelação de alta-ordem e na divergência entre a distribuição conjunta de $\vec{s}$ e outras distribuições [16].

O enfoque da separação de fontes é essencialmente espacial (“olha-se” para a estrutura através dos sensores, não através do tempo). A consequência de ignorar qualquer estrutura de tempo é que as informações contidas nos dados são exaustivamente representadas pelas amostras da distribuição da matriz de mistura [16].

Todavia, como visto no início da seção, antes de efetuar a separação estatística de alta-ordem é necessário estimar as características estatísticas

de baixa-ordem das variáveis aleatórias efetuando-se um *branqueamento* das amostras. Essa etapa do processo é denominada de *estimação gaussiana* da ICA e sempre é possível de ser realizada [51].

Amostras aleatórias branqueadas satisfazem a condição de média zero e a variância unitária que, por sua vez, são preservadas mediante transformações ortogonais [51]. O branqueamento das amostras na ICA torna impossível determinar a variância das componentes independentes, bem como a ordem (posição) dessas componentes na matriz S .

As funções de distribuição de probabilidade das fontes independentes são desconhecidas e denotadas por $p_{si}(\cdot)$, embora se assume que a média é zero e a variância é unitária em cada componente. A distribuição conjunta, $p_s(\vec{s})$ é dada pelo produto das distribuições marginais de cada componente independente:

$$p_s(\vec{s}) = p_{s1}(s_1) \cdot p_{s2}(s_2) \cdot \dots \cdot p_{sn}(s_n) = \prod_{i=1}^n p_{si}(s_i). \quad (3.7)$$

No caso da ICA, a função densidade de probabilidade conjunta é obtida pelo produto das funções de densidade marginais, uma vez que as componentes marginais são estatisticamente independentes entre si.

As bordas da distribuição conjunta das variáveis aleatórias estão na direção dos vetores formados pelas colunas da matriz de mistura A , assim, a estimação do modelo ICA é feita estimando primeiramente a densidade conjunta das variáveis aleatórias e, em seguida, localizando as suas bordas [51].

3.2.1 Maximização da Não-Gaussianidade

Conforme visto na seção 3.1, estatísticas de alta-ordem remetem a cumulantes de ordem maior que dois. Como a distribuição gaussiana é definida apenas pelos cumulantes de primeira e segunda ordem (todos os demais são zero), estimar distribuições estatisticamente independentes em altas ordens estatísticas é estimar distribuições não-gaussianas. Assim, a não-gaussianidade é o ponto de partida para elaborar algoritmos para a ICA.

A *kurtosis* e a *negentropy* são duas medidas de não-gaussianidade, de particular interesse para elaboração do algoritmo ICA, devido basicamente à sua simplicidade.

A kurtosis é o quarto momento estatístico que, para distribuições gaussianas com média zero, apresenta também valor zero. Maximizar a kurtosis equivale a buscar por distribuições não-gaussianas no algoritmo da ICA.

A negentropy, por sua vez, é uma medida de entropia diferencial baseada na teoria da informação e que se fundamenta no fato de que variáveis

gaussianas possuem entropia muito grande quando comparadas com variáveis de outras distribuições com a mesma variância. Logo, medidas de entropia podem ser usadas para medir não-gaussianidades [51].

A Equação 3.8 apresenta a negentropia para uma distribuição $\vec{y}$.

$$J(\vec{y}) = H(\vec{v}) - H(\vec{y}) \quad (3.8)$$

onde, $J(\vec{y})$ é a negentropia de $\vec{y}$, $\vec{y}$ é uma variável aleatória de média zero e variância unitária, $\vec{v}$ é um vetor gaussiano aleatório com a mesma matriz de covariância de $\vec{y}$ e $H(\cdot)$ é a entropia diferencial.

Para a distribuição gaussiana, a entropia diferencial é dada pela equação 3.9.

$$H(\vec{v}) = \frac{1}{2} \log |\det \Psi| + \frac{n}{2} (1 + \log 2\pi) \quad (3.9)$$

onde, n é a dimensão de $\vec{v}$ e Ψ é a covariância de $\vec{v}$ que é a mesma de $\vec{y}$.

O problema do uso da kurtosis é que existem variáveis não-gaussianas cuja kurtosis é zero, mesmo que muito raras [51]. Ainda, a kurtosis é muito sensível a poucas observações podendo acarretar resultados errôneos ou irrelevantes [51].

A vantagem de se utilizar a negentropia como medida de não-gaussianidade é porque ela é bem justificada pela teoria estatística, bem como possui invariabilidade às transformações lineares inversíveis [51].

O problema da negentropia é sua difícil computabilidade, o que força o uso de uma aproximação do seu valor dado pela equação 3.10:

$$J(\vec{y}) \cong k_1 (E[G_1(\vec{y})])^2 + k_2 (E[G_2(\vec{y})] - E[G_2(\vec{v})])^2, \quad (3.10)$$

onde, J é a negentropia, k_1 e k_2 são constantes positivas; $\vec{v}$ é uma variável gaussiana de média zero e variância unitária; $\vec{y}$ é uma variável aleatória de média zero e variância unitária; G_1 e G_2 são funções gerais não quadráticas; e $E[G_i(\cdot)]$ é o valor esperado da função não quadrática e representa momentos estatísticos não-polinomiais [51].

Pode-se simplificar a expressão acima fazendo $G_1 = G_2 = G$, logo

$$J(\vec{y}) \propto [E[G(\vec{y})] - E[G(\vec{v})]]^2. \quad (3.11)$$

Escolhendo G que não cresce tão rápido é possível obter um estimador mais robusto [51], por exemplo:

$$\begin{aligned} G(\vec{y}) &= \frac{1}{a_1} \log(\cosh(a_1 \vec{y})); & 1 \leq a_1 \leq 2 \\ G(\vec{y}) &= -e^{-\frac{\vec{y}^2}{2}} \end{aligned}$$

3.2.2 Técnicas de Estimação na ICA

Algoritmos para ICA usando diferentes abordagens para independência estatística são encontrados nas referências [51] e [49]. Neste trabalho, será abordado apenas o modelo de estimação pela máxima verossimilhança uma vez que a comparação entre o modelo ICA e o modelo topográfico, que será apresentado na seção 3.3, é didaticamente melhor caracterizado nessa abordagem.

O desenvolvimento que se segue mostra como foi obtida a equação do modelo ICA, usando a máxima verossimilhança, para a estimação das componentes estatisticamente independentes. Tal desenvolvimento é necessário para o entendimento de como a proposta de modificação do modelo TICA foi elaborada na seção 5.2 (pág. 69), que será abordada adiante.

Na seção 3.2 vimos que o modelo ICA é linear e não singular, de forma que $X = A \times S$ e $S = A^{-1} \times X$. Então, densidade³ dessa transformação linear pode ser formulada como [51]

$$p_x(\vec{x}) = |\det W| p_s(\vec{s}) = |\det W| \prod_i p_i(s_i), \quad (3.12)$$

onde $W = A^{-1}$, p_x é a densidade da mistura, p_s é dada na equação 3.7 (pág. 38) e p_i denota a densidade das componentes independentes.

Da representação vetorial das componentes que formam as matrizes da equação 3.4 (pág. 36), é possível expressar s_i em função de $\vec{w}_i$ e $\vec{x}$ para n componentes independentes, como:

$$p_x(\vec{x}) = |\det W| \prod_{i=1}^n p_i(\vec{w}_i^T \vec{x}). \quad (3.13)$$

Assumindo que existem T observações de $\vec{x}$, representadas por $\vec{x}(1), \vec{x}(2), \dots, \vec{x}(T)$, então a verossimilhança entre essas observações, L , pode ser obtida como o produto da densidade na equação 3.20 avaliada nos T pontos de observações. Logo,

$$L(W) = \prod_{t=1}^T \prod_{i=1}^n p_i(\vec{w}_i^T \vec{x}(t)) |\det W|. \quad (3.14)$$

Para simplificar algebricamente a expressão da equação 3.26 e torná-la prática, utiliza-se o logaritmo da verossimilhança, desde que o máximo do logaritmo seja obtido no mesmo ponto do máximo da verossimilhança [51]. O logaritmo da verossimilhança é dado por

³A densidade de uma transformação linear $y = g(x)$ mostra como obter $p_y(y)$ a partir de $p_x(x)$ [51]

$$\log L(W) = \sum_{t=1}^T \sum_{i=1}^n \log p_i(\vec{w}_i^T \vec{x}(t)) + T \log |\det W|. \quad (3.15)$$

Portanto essa é a equação utilizada para a obtenção do modelo ICA. Hyvärinen, Karhunen e Oja [51], sugerem o uso do algoritmo *FastICA* descrito a seguir pela estimação da máxima verossimilhança para obtenção da matriz de mistura.

Algoritmo 1 *Algoritmo FastICA pela estimação da máxima verossimilhança.*

1. Centralize os dados, $\vec{x}$, obtendo média zero e variância unitária (branqueamento).
2. Calcule a matriz de correlação $C = E[\vec{x}\vec{x}^T]$.
3. Inicialize aleatoriamente a matriz W .
4. Compute

$$\vec{s} = W\vec{x}$$

$$\beta_i = -E[s_i g(s_i)], \text{ para } i = 1, \dots, n$$

$$\alpha_i = -1/(\beta_i + E[\dot{g}(s_i)]), \text{ para } i = 1, \dots, n$$
5. Atualize a matriz W por:

$$W \leftarrow W + \text{diag}(\alpha_i)[\text{diag}(\beta_i) + E[g(\vec{s})\vec{s}^T]]W$$
6. Descorrelate e normalize por:

$$W \leftarrow (WCW^T)^{-1/2}W$$
7. Se não convergiu, retorne ao passo 4.

No algoritmo acima, $g(\cdot)$ é a derivada de $G(\cdot)$ e $\dot{g}$ é a derivada de $g(\cdot)$. A convergência no algoritmo consiste em não haver mais mudanças na matriz W .

3.2.3 ICA e o Modelo de Visão

A representação de texturas por estatísticas marginais é uma aproximação simples e atraente em que um dado conjunto de funções densidade de probabilidade marginais representam a descrição estatística conjunta de uma característica da textura [70].

Vários experimentos envolvendo a ICA têm sido usados visando à geração de modelos genéricos para representação de texturas [70, 73]. Muitos desses experimentos permitem a obtenção de bases em que as componentes da base são funções de Gabor semelhantes aos obtidos por Daugman [22] para o modelo de visão do córtex V1.

Essas bases podem ser comparadas às células simples em V1, ou seja, neurônios independentes que efetuam a separação estatística de alta-ordem das componentes visuais da imagem focada na retina [51].

Cada neurônio efetua a identificação de estatísticas marginais que irão compor a descrição estatística conjunta das características da imagem.

Em uma interpretação neurocientífica, as variáveis latentes⁴ s_i modelam as respostas das células simples em V1 e a matriz de mistura A_i é uma aproximação de seus campos receptivos [48].

A relação das propriedades de células simples com estatísticas de imagens naturais foi explorada extensivamente por Field [31, 32]. Field propôs que as células simples são otimizadas especificamente para codificar imagens naturais. Isso porque elas produzem representações de estímulo que ajudam na detecção de similaridades [9].

Isso está em consonância com Barlow [8] que afirma que os nossos detectores corticais de características visuais devem ser o resultado final de um processo de redução de redundâncias no qual a ativação de cada detector de características é suposta ser tão estatisticamente independentes quanto possível entre os detectores.

Por outro lado, Field e Olshausen [76] mostraram que impondo um espalhamento na saída dos campos receptivos treinados sobre imagens naturais houve a produção de campos receptivos similares aos das células simples do córtex. Esse resultado foi baseado no experimento de Vinje e Gallant [107], onde mostraram, fisiologicamente, que as respostas dos neurônios no córtex V1 são mais esparsas durante a apresentação de estímulos de cenas naturais.

Nesse sentido, a ICA, em imagens naturais, produz campos receptivos como aqueles das células simples do córtex visual primário [76]. Isso porque o modelo ICA é estritamente baseado no estímulo estatístico, que por sua vez é abundante nas cenas naturais e, portanto, as células simples de V1 podem ser modeladas como filtros estatísticos de alta-ordem.

Não linearidades e adaptações podem ser adicionados em um processo de melhoria da extração de características do modelo ICA, que parece ser uma estratégia de sucesso em estágios anteriores da visão como nas células ganglionares da retina (*Retinal Ganglion Cells - RGCs*) e nos Núcleos Geni-

⁴Significa que não podem ser observadas diretamente.

culados Laterais⁵ (*Lateral Geniculate Nucleous* - *LGN*), pois, sabe-se que o sistema de visão, nos primeiros estágios, responde a estímulos relacionados a não-linearidades [42].

Estudos mais recentes mostram que os campos receptivos localizados nas células RGC⁶ e LGN do córtex visual removem as redundâncias de segunda ordem estatística (bem caracterizadas por funções gaussianas), efetuando um branqueamento nas imagens visualizadas, enquanto em V1 são removidas apenas as redundâncias de alta-ordem estatística (não caracterizadas por funções gaussianas) [2, 3, 4, 11, 95].

Entretanto, os neurônios do LGN não branqueiam genericamente qualquer estímulo, mas somente aqueles que exibem a mesma estrutura correlacional de imagens naturais [95]. É por isso que uma das componentes independentes do modelo ICA pode ser uma gaussiana.

O modelo ICA do córtex visual não é um modelo completo das células simples. Por ser linear⁷ e não adaptativo, muitos aspectos das células simples são ignorados, como a adaptação ao contraste, normalização do contraste, não-linearidades envolvidas no processo de ajuste de orientação e adaptação a vários estímulos estatísticos.

As células simples respondem, preferencialmente, à linhas e bordas orientadas, podendo serem vistas como detectores de linhas ou de bordas [11, 42]. As componentes independentes da ICA podem ser consideradas como elementos constituintes das imagens que se assemelham a bordas e os filtros da matriz de mistura são similares aos campos receptivos das células simples do córtex [42]. Os filtros são usualmente passa-baixa em uma direção e passa-banda na direção ortogonal [42].

Exceção à correspondência entre os filtros ICA e as células simples do córtex está na distribuição do pico de resposta da frequência espacial. Enquanto a ICA leva o pico da frequência espacial próximo à máxima frequência espacial seguido por amostras oscilatórias, medidas nas células simples mostram uma distribuição bem mais larga [42]. E ainda, enquanto as células simples têm campos receptivos atuando em diferentes escalas espaciais, os filtros da ICA mostram muito menos variabilidade [42].

Por outro lado, existe uma boa correspondência entre as distribuições para a largura de banda da frequência espacial, ajuste da orientação da largura de banda, relação de forma e tamanho do campo receptivo [42].

⁵Uma das atribuições do LGN é funcionar como um relé controlando a abertura e fechamento da pupila.

⁶Os axônios destas células vão dar origem às fibras nervosas do nervo óptico.

⁷Numerosos estudos têm confirmado que as células simples em V1 podem ser consideradas basicamente como operadores lineares no tempo e no espaço (veja Kagan et al. [56]).

O modelo apresentado por Balboa e Grzywacz [6] mostra que variações no tamanho dos campos receptivos são ótimas para representação da intensidade das bordas em imagens.

3.3 Análise Topográfica de Componentes Independentes

Embora a analogia do modelo ICA ao sistema de visão humano seja satisfatório em grande parte dos experimentos, acredita-se que o modelo possa ser aperfeiçoado para corresponder melhor ao mecanismo de visão.

O estudo apresentado por Jenssen e Eltoft [53] reforça essa hipótese quando compara a segmentação de texturas usando ICA com funções de Gabor, e mostra que a ICA possui melhor qualidade de segmentação em relação às funções de Gabor.

Por outro lado, o modelo ICA representado na equação (3.15) (pág. 41) é inapropriado para descrever a resposta das células complexas do córtex visual primário [82]. Isso porque no modelo ICA as variáveis s_i não possuem uma ordem estatística particular e nem uma relação é assumida entre elas, a menos da completa independência estatística.

Entretanto, existem aplicações na qual o modelo ICA não remove completamente a dependência entre as componentes. Todavia, as dependências de alta-ordem que a ICA não remove são usadas para definir uma ordem topográfica, de forma que as células (filtros ICA) próximas tendem a serem ativas⁸ (ou inativas) ao mesmo tempo [73].

Assim, a organização topográfica dos filtros (células simples) elimina a dependência residual de alta-ordem das componentes do modelo ICA, alterando e organizando o banco de filtros do modelo.

O desenvolvimento que segue mostra como o critério topográfico foi idealizado e modelado por Hyvärinen, Hoyer e Inki [49] para compor o modelo ICA apresentado anteriormente.

Hyvärinen, Hoyer e Inki [49] propuseram um modelo de correlação de alta-ordem baseado na energia das componentes independentes, que pode ser interpretado intuitivamente como ativações simultâneas de células topograficamente próximas em uma imagem.

A TICA leva em consideração que componentes próximos entre si, em uma representação topográfica, são fortemente dependentes em relação à

⁸Do ponto de vista neurobiológico uma célula é ativa quando está transmitindo impulso elétrico às demais células de suas ligações.

correlação, ou à informação mútua, promovendo um relaxamento da relação de independência do modelo ICA [49].

Assim, o modelo $X = A \times S$ da Análise Topográfica de Componentes Independentes consiste em definir uma função densidade de probabilidade conjunta de S baseada no modelo de estímulos laterais exercidos pelas células em uma vizinhança.

O modelo proposto por Hyvärinen, Hoyer e Inki [49] parte da premissa de que dois conjuntos de variáveis s_i e s_j topograficamente próximas e linearmente descorrelacionadas possam ser definidas como produtos de componentes independentes de alta-ordem, z_i e z_j , de média zero e variância unitária por uma grandeza estatística comum, κ , como mostrado nas equações (3.16).

$$\begin{aligned} s_i &= \kappa z_i \\ s_j &= \kappa z_j. \end{aligned} \tag{3.16}$$

Nesse caso, z_i é um conjunto de variáveis aleatórias estatisticamente independentes e que possuem a mesma função densidade de probabilidade das variáveis s_i .

A proximidade espacial entre as variáveis permite que a representação mostrada na equação 3.16 tenha a mesma grandeza estatística κ .

O cálculo da covariância das componentes estatisticamente independentes mostra a interdependência entre essas variáveis:

$$\begin{aligned} \text{cov}(s_i, s_j) &= E[s_i s_j] - E[s_i] E[s_j], \\ \text{cov}(s_i, s_j) &= E[\kappa z_i \kappa z_j] - E[\kappa z_i] E[\kappa z_j], \end{aligned}$$

como z_i e z_j são estatisticamente independentes de média zero e variância unitária, temos

$$\text{cov}(s_i, s_j) = E[\kappa^2] E[z_i] E[z_j] - E[\kappa]^2 E[z_i] E[z_j] = 0. \tag{3.17}$$

O resultado da equação (3.17) mostra que, quando z_i e z_j são estatisticamente independentes com média zero e variância unitária, não há como determinar uma dependência entre as variáveis topograficamente próximas.

Hyvärinen, Hoyer e Inki [49] propuseram o cálculo da covariância do quadrado das componentes estatisticamente independentes da equação (3.16) para mostrar que existe uma correlação entre essas componentes próximas e que é dada pela medida estatística comum, κ . Assim,

$$\text{cov}(s_i^2, s_j^2) = E[s_i^2 s_j^2] - E[s_i^2] E[s_j^2]$$

que resulta em:

$$\text{cov}(s_i^2, s_j^2) = E[\kappa^4] - E[\kappa^2]^2 \neq 0, \quad \kappa \neq 0, \tag{3.18}$$

o resultado da equação (3.18) mostra que existe uma interdependência entre as variáveis vizinhas, dada pela energia das variáveis, devido à proximidade topográfica entre elas. O uso da energia mostra que a medida estatística comum, κ , entre as variáveis s_i e s_j é a variância [49].

Nesse sentido, o critério topográfico⁹ foi definido por Hyvärinen, Hoyer e Inki [49], especificando uma função de vizinhança, $h(i, j)$ que expressa a proximidade entre as componentes (i e j) da imagem de forma monotonicamente decrescente em função de alguma medida de distância.

O decréscimo em função da distância é para modelar as iterações laterais curtas das células complexas de V1, como será abordado na seção 3.3.2 (pág. 49).

Hyvärinen, Hoyer e Inki [49] definiram u_k como a energia das componentes independentes em alta-ordem estatística e a aplicação da função de vizinhança nas componentes do modelo foi efetuada como mostra a equação (3.19).

$$\kappa = \phi \left\{ \sum_k h(i, j) u_k \right\}, \quad (3.19)$$

onde ϕ é uma não-linearidade escalar e κ é a variância, devido à energia, logo, a dependência entre s_i e s_j é descrita pela dependência de suas variâncias. As variáveis u_k e κ são mutuamente independentes.

A Figura 3.2 mostra uma base TICA de textura natural em que todos os filtros (campos receptivos das células em V1) estão localizados e orientados [49].

As funções da base correspondem aos padrões que as variáveis da matriz de mistura codificam e que parecem padrões de bordas. De fato, o estudo apresentado em [11] mostra que as componentes independentes de imagens naturais são aproximações das bordas da imagem.

3.3.1 Modelo de Estimação na TICA

Baseado na abordagem da ICA pela estimação da máxima verossimilhança, a adaptação do método de TICA para essa mesma abordagem pode ser verificada como mostram os parágrafos seguintes.

De maneira análoga ao desenvolvimento da seção 3.2.2 (pág. 40), a densidade da transformação linear não singular $X = W^{-1} \times S$ pode ser formulada

⁹O modelo topográfico aqui não segue a regra de atualização sináptica proposta por Kohonen [63].

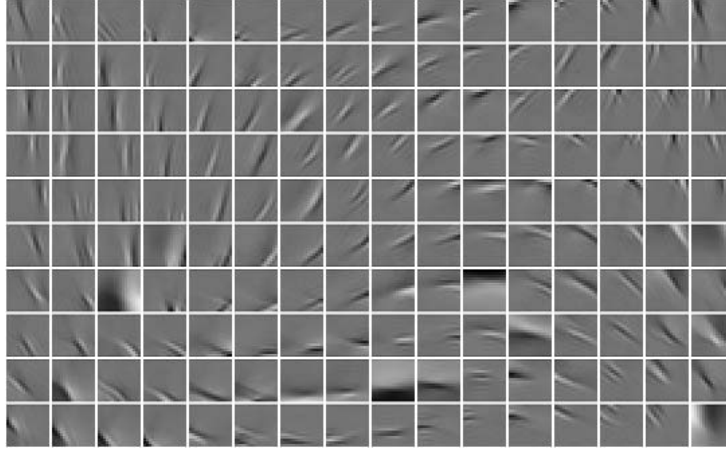


Figura 3.2: Exemplo de uma base TICA de textura natural [49].

como [51]

$$p_x(\vec{x}) = |\det W| p_s(\vec{s}) = |\det W| \prod_j p_j(s_j), \quad (3.20)$$

onde p_x é a densidade da mistura, p_s é dada na equação 3.7 (pág. 38) e p_j denota a densidade das componentes independentes.

Do resultado obtido na equação 3.18 (pág. 45), pode-se definir o parâmetro u_i como a relação estatística das componentes independentes em alta-ordem no critério topográfico como sendo

$$u_j = s_j^2 = (\vec{w}_j^T \vec{x}(t))^2. \quad (3.21)$$

Assim, pode-se adequar a equação 3.20, substituindo s_j por u_j

$$p_x(\vec{x}) = |\det W| \prod_j p_j(u_j), \quad (3.22)$$

A densidade das componentes independentes, nesse caso, é a densidade de uma região e pode ser definida como

$$p_j(u_j) = \left\{ \sum_i h(i, j) u_i \right\} \quad (3.23)$$

O critério topográfico demanda a existência de uma função de vizinhança da qual os neurônios, nesta vizinhança, se correlacionam de uma maneira não linear através de interações laterais, ou seja,

$$p_j(u_j) = \left\{ \sum_i \phi(h(i, j) u_i) \right\} = \phi \left(\sum_{i=1}^n h(i, j) u_i \right) \quad (3.24)$$

Substituindo a equação 3.24 na equação 3.22, temos

$$p_x(\vec{x}) = |\det W| \prod_{j=1}^n \phi \left(\sum_{i=1}^n h(i, j) u_i \right) \quad (3.25)$$

Assumindo que existem T observações de $\vec{x}$, representadas por $\vec{x}(1), \vec{x}(2), \dots, \vec{x}(T)$, então a verossimilhança entre essas observações, L , pode ser obtida como o produto da densidade na equação 3.25 avaliada nos T pontos de observações. Logo,

$$L(W) = \prod_{t=1}^T \left\{ |\det W| \prod_{j=1}^n \phi \left(\sum_{i=1}^n h(i, j) u_i \right) \right\} \quad (3.26)$$

Para simplificar algebricamente a expressão da equação 3.26 e torná-la prática, utiliza-se o logaritmo da verossimilhança, desde que o máximo do logaritmo seja obtido no mesmo ponto do máximo da verossimilhança [51]. O logaritmo da verossimilhança é dado por

$$\log L(W) = \log \prod_{t=1}^T \left\{ |\det W| \prod_{j=1}^n \phi \left(\sum_{i=1}^n h(i, j) u_i \right) \right\}. \quad (3.27)$$

Simplificando,

$$\log L(W) = \sum_{t=1}^T \log |\det W| + \sum_{t=1}^T \sum_{j=1}^n \log \phi \left(\sum_{i=1}^n h(i, j) u_i \right). \quad (3.28)$$

Assim, o modelo TICA é obtido à partir da organização da equação acima, resultando em [49]:

$$\log L(W) = \sum_{t=1}^T \sum_{j=1}^n G \left(\sum_{i=1}^n h(i, j) (\vec{w}_i^T \vec{x}(t))^2 \right) + T \log |\det W|, \quad (3.29)$$

em que G é uma transformação não linear monotônica de números reais positivos, $h(i, j)$ é a função de vizinhança, $(\vec{w}_i^T \vec{x}(t))^2$ é a energia das componentes independentes não negativas de alta-ordem estatística, n é o tamanho da janela da vizinhança, T é o número de observações de $\vec{x}$ e W é a inversa da matriz de mistura. As propriedades do modelo TICA são descritas em [49].

3.3.2 TICA e o Modelo de Visão

Sabe-se que os neurônios em V1 são sensíveis a uma série de estímulos, incluindo, orientação, direção de movimento e coloração [17]. O efeito de seletividade neuronal em V1 depende dos padrões de conectividade do circuito. Porém, neurônios das células complexas em V1 exibem pequena sensibilidade para a localização espacial do estímulo visual em seus campos receptivos [17].

Células simples são seletivas à orientação, frequência espacial e fase espacial enquanto as células complexas são seletivas a orientação e a frequência espacial, mas não são seletivas à fase espacial [17], portanto, a distinção entre células simples e complexas em V1 é baseada na separação espacial de suas respostas aos vários estímulos visuais [17].

A atividade das células simples é modulada na mesma frequência do estímulo, enquanto a resposta da célula complexa varia em até duas vezes essa frequência [17]. Estas respostas temporais diferentes são uma consequência direta da falta de seletividade das células complexas em relação à fase [17].

Todavia, as células simples e complexas em V1 efetuam interações laterais que compõem um mecanismo de dependência residual na estrutura celular do córtex [81], que é modelada pelo critério topográfico no modelo TICA [49].

Uma vez que a TICA insere uma organização espacial no modelo da ICA, o que pode ser vista como uma organização topográfica das células V1 do córtex visual [49], ela estabelece uma relação de dependência residual na estrutura das células do córtex, buscando encontrar sub-espacos com características invariantes em alta-ordem estatística [47].

O uso do critério topográfico na ICA pode ser visto como uma particularização da Análise de Sub-espacos Independentes (ISA - *Independent Subspace Analysis*) [47] e permite, além da aproximação do modelo de células simples de V1 do córtex visual, uma aproximação do modelo cortical das células complexas do sistema de visão [47].

Evidências sugerem que a área V2 do córtex replica a estrutura da área V1, mas em uma escala espacial muito grande [38]. Isto é explicado, em parte, pela existência de influências modulatórias através de conexões monossinápticas horizontais curtas entre as células vizinhas simples e complexas em V1 com mecanismo de realimentação em V2 e células com conexões monossinápticas horizontais longas entre si em V2 [38, 81].

As conexões horizontais efetuam o agrupamento perceptivo, pela excitação/inibição célula a célula, formando os grupos sensíveis a determinadas características estatísticas da imagem [38]. Essa propriedade dos neurônios é conhecida como *propriedade bipolar de ativação dos neurônios*.

A propriedade bipolar de ativação dos neurônios no córtex cerebral tem sido usada para explicar e prever muitos dos agrupamentos perceptivos

dos dados visuais, entretanto, há muito mais evidências neurofisiológicas da propriedade bipolar para ativação dos neurônios em V2 que em V1, havendo apenas umas poucas células que efetuam o suporte para ativação/desativação da propriedade bipolar em V1 [38]. Isso sugere que o critério topográfico da ICA deva ser mais eficiente quando efetuado pelas células em V2 que em V1.

Capítulo 4

Fractais

O matemático Mandelbrot [68] despertou na comunidade científica o interesse na modelagem de objetos e superfícies com formato altamente irregular quando introduziu o conceito de fractal (do latim *fractio* que quer dizer fração ou parte [30]) baseando-se nas idéias do matemático Euclides, que afirmou ser possível reduzir todas as formas da natureza em formas geométricas simples.

Mandelbrot deu início a uma série de trabalhos que culminou no surgimento do que é hoje denominado de *geometria fractal*. Imagens como galáxias, flocos de neve, contorno da costa marinha e silhuetas de montanhas são modeladas por equações baseadas na geometria fractal [41]. Essas imagens costumam ser muito irregulares de forma que mesmo ajustando a escala para aproximar a imagem do observador, as irregularidades continuam a existir em um grau elevado de complexidade [41].

A base da teoria fractal está no conceito de auto-similaridade que, no espaço Euclidiano, pode ser formulada da seguinte maneira [105]:

Conceito 1 *Considere um conjunto limitado F em um n -espaço Euclidiano. O conjunto F é dito ser auto-similar se F está na união de K distintas (não-sobrepostas) cópias de si mesmo, em que cada cópia está escalada para baixo por uma razão menor que um em todas as coordenadas.*

O conceito de auto-similaridade no espaço Euclidiano foi estendido, mais tarde, para modelos estatísticos permitindo classificar um objeto fractal em duas sub-áreas da teoria fractal [105]:

1. *Objeto Auto-Similar Determinístico*: no qual o fractal (então denominado fractal euclidiano) é composto de características distintas que remontam cada outro objeto idêntico em diferentes escalas (característica invariante a escala).

2. *Objeto Auto-Similar Estatístico*: no qual as características do fractal (então denominado fractal estatístico ou, ainda, *random fractal*) podem mudar em diferentes escalas, mas suas propriedades estatísticas em todas as escalas são as mesmas (invariância estatística à escala).

As Figuras 4.1 e 4.2 apresentam exemplos de fractais geométricos e estatísticos, respectivamente.

Fractais, portanto, são definidos pela propriedade de auto-similaridade ao longo de diferentes escalas de uma imagem. Essa característica é que permite a distinção de um objeto euclidiano clássico de um objeto fractal [105].

No mundo real, a descrição fractal é estatística (*random fractals*), e é a mais útil na modelagem de determinadas texturas [105].



Figura 4.1: Fractais geométricos [33].



Figura 4.2: Fractais estatísticos [99].

4.1 Fractais Estatísticos

Os conjuntos fractais geométricos são definidos pela seguinte equação de auto-similaridade:

$$f(x, y) = \rho^{-d} f(\rho x, \rho y), \quad (4.1)$$

onde ρ é o fator de escalamento e d é o expoente de escalamento, ou *parâmetro de auto-similaridade determinístico*, e $f(x, y)$ é a função de uma imagem qualquer em tons de cinza.

A equação 4.1 representa um processo auto-similar determinístico, uma vez que $\rho^{-d} f(\rho x, \rho y)$ e $f(x, y)$ possuem uma semelhança geométrica em diferentes escalas.

A definição de fractais estatísticos segue a formulação da equação:

$$\text{fdp}(f(x, y)) = \rho^{-r} \text{fdp}(f(\rho x, \rho y)), \quad (4.2)$$

onde ρ é o fator de escalamento e r é o expoente de escalamento, ou *parâmetro de auto-similaridade estatística*, $f(x, y)$ é a função de uma imagem qualquer e fdp é a função densidade de probabilidade de $f(x, y)$.

A equação 4.2 representa um processo auto-similar estatístico, uma vez que $\rho^{-r} \text{fdp}(f(\rho x, \rho y))$ e $\text{fdp}(f(x, y))$ possuem a mesma função densidade de probabilidade estatística [28].

4.1.1 Fractais Estatísticos e Texturas Naturais

Existe uma relação entre fractais estatísticos e imagens naturais dada pela distribuição $1/f$, ou ainda, uma relação de fractais estatísticos com o espectro de potência de Fourier, mostrando que as imagens naturais são fractais estatísticos possuindo auto-similaridade inerente [104].

De fato Pentland [78], em seus trabalhos, já havia mostrado que muitas das superfícies de texturas naturais são fractais espacialmente isotrópicos e que a intensidade das imagens dessas superfícies são também fractais, compondo assim uma terceira dimensão nas imagens supostamente planas.

Sabe-se, também, que fractais estatísticos são excelentes descritores de bordas, ou seja, os fractais estatísticos se relacionam com o espectro de fase de Fourier, o que permite afirmar que também efetuam análise estatística de alta-ordem [58, 61, 78, 109, 111].

Os fractais estatísticos interpretam uma superfície textural 2-D como um modelo 3-D não se restringindo somente à representação de texturas pela segunda ordem estatística de Julesz, mas também incorporando à representação, através da frequência espacial inserida indiretamente na dimensão

fractal, informações estatísticas de alta-ordem, permitindo explorar as características de texturas em concordância com Pratt, Faugeras e Gagalowicz [78].

A quantificação de texturas a partir de fractais estatísticos é baseada principalmente nas características de auto-similaridade, lacunaridade e dimensão fractal. Fractais, por outro lado, são também excelentes descritores de bordas, baseando os algoritmos nas mesmas características já mencionadas.

Como visto no início do capítulo, a *auto-similaridade estatística* dos fractais refere-se ao fato de a medida estatística de um sinal ser invariante em relação à transformação de escala [78]. Nesse sentido, uma superfície fractal é uma superfície que pode ser precisamente aproximada por uma função fractal simples sobre uma faixa de escalas [78].

Métodos multi-escalares, como a transformada wavelet, são usados para a determinação de medidas estatísticas invariantes em diferentes escalas no domínio do problema [26]. Todavia, os métodos baseados em wavelets possuem problemas na estimação precisa do parâmetro de auto-similaridade quando seu valor é pequeno ($< 1/2$) [26].

A *lacunaridade fractal* é uma característica fractal de alta-ordem, distinta e independente da dimensão fractal, geometricamente invariável, que expressa uma medida da densidade da textura [111]. Se a textura é densa, a lacunaridade é pequena. Em texturas naturais a medida de lacunaridade varia rapidamente na borda de diferentes regiões, permitindo extrair as bordas da imagem [111].

A fórmula de lacunaridade, definida por Mandelbrot, e apresentada na equação 4.3, usa momentos estatísticos de ordem um e dois [69],

$$\gamma(x) = \frac{\sigma^2(x) - \mu^2(x)}{\mu^2(x)}, \quad (4.3)$$

onde, $\gamma(\cdot)$ denota a lacunaridade, σ^2 é a variância e μ é a média da variável aleatória x que é, por exemplo, o nível de cinza da imagem.

Turner [105], estende a definição da lacunaridade para momentos estatísticos de ordem mais elevada, expressando a lacunaridade como:

$$\gamma(x) = E \left[\left| \frac{x}{E[x]} - 1 \right|^n \right]^{\frac{1}{n}}, \quad (4.4)$$

onde, $E[\cdot]$ é o valor esperado e n é a ordem estatística.

A *dimensão fractal*, por sua vez, é uma medida matemática baseada em topologia e conhecida como dimensão de Hausdorff-Besicovitch [105]. Ela pode ser calculada tomando o limite do quociente das mudanças logarítmicas em magnitude e em escala, à medida que a escala se aproxima de zero.

Se uma curva pode ser subdividida em N partes, sendo o todo M vezes maior que cada uma delas, a dimensão de Hausdorff-Besicovitch será [105]:

$$D = \frac{\log N}{\log M}. \quad (4.5)$$

A dimensão fractal é muito utilizada para a caracterização de bordas e superfícies e é bem explorada na literatura científica [58, 78, 111]. Basicamente, o princípio se baseia na confrontação da dimensão fractal da imagem com a dimensão fractal de cada pixel da imagem (as dimensões dos pixels são calculadas levando em conta sua vizinhança). Os pixels cujo valor da dimensão fractal é menor que a dimensão fractal da imagem são localizados nas regiões de bordas [78].

A caracterização de superfícies fractais pode ser feita pelo uso da dimensão fractal. Uma superfície é fractal se a dimensão fractal é estável¹ sobre transformações lineares da intensidade dos pixels em uma larga faixa de transformações de escala [78].

4.1.2 Coeficiente de Variação Espacial Estatístico

Conforme visto nas seções anteriores, a observação da invariância estatística depende da medição do descritor fractal. Este, por sua vez, corresponde a um parâmetro que busca representar o comportamento estatístico da imagem em diferentes escalas da mesma região analisada.

Nesse sentido, pode-se afirmar que o descritor fractal comporta-se como um coeficiente que retrata a dependência estatística de uma região da imagem mediante uma variação espacial (escala) dessa mesma região, porém, observada a partir de altas ordens estatísticas.

Embora o *Coeficiente de Variação Espacial Estatístico*, CVEE, não tenha sido criado por Nunes e Conci [75] com intuito claro de modelagem fractal, o CVEE possui um comportamento descritivo parecido com fractais, porém, sob a ótica de baixa-ordem estatística. Isso porque o CVEE procura representar o comportamento estatístico de baixa-ordem de uma região sob variações de escala dessa mesma região.

A concepção do CVEE surgiu da necessidade de quantificar características de texturas através de medidas estatísticas que pudessem descrever as variações espaciais de intensidade da imagem. Para isso, foram empregadas duas medidas estatísticas no modelo: uma de posição (média) e uma de dispersão (coeficiente de variação) do conjunto de pixels pertencentes a uma

¹O conceito de estabilidade é: se um número de variáveis aleatórias independentes e identicamente distribuídas possuem uma distribuição estável, então uma combinação linear dessas variáveis terá a mesma distribuição.

região da imagem. A equação 4.6 exhibe o modelo.

$$\text{CVEE} = \frac{\pi}{180} \left(\sqrt{\bar{c}^2 + \text{CV}^2} \right) \tan^{-1} \left(\frac{\bar{c}}{\text{CV}} \right), \quad (4.6)$$

onde, $\bar{c}$ é a média de intensidade dos pixels e CV é o coeficiente de variação definido como $\text{CV} = (\sigma/\bar{c}) \times 100$ em que σ é o desvio padrão da intensidade dos pixels na região considerada.

Da mesma forma que o descritor fractal, o CVEE também é invariante à rotação [75].

4.1.3 Distribuição α -estável

O interesse na distribuição α -estável está na representação de funções de cauda-longa, ou seja, funções que possuem dependência estatística ao longo de uma distância ou de um período de tempo muito grande [26].

De maneira mais específica, a *dependência de cauda-longa* ocorre quando existe um decaimento muito lento da correlação entre as amostras e permite que o passado distante influencie o futuro, mesmo que este seja independente daquele para uma certa distância ao longo de um período de tempo [26].

Como visto anteriormente, a dependência estatística ao longo de uma distância (escalas²) é bem modelada por fractais estatísticos, portanto, as distribuições α -estáveis, que modelam essa dependência, são também funções que representam fractais estatísticos.

A formalização desse conceito é feita pelo *Teorema de Zolotarev* (veja Kuruoglu [65]) que mostra que variáveis aleatórias α -estáveis não-gaussianas possuem parâmetro estatístico constante em diferentes escalas, ou seja, possuem auto-similaridade estatística. De fato, funções de cauda-longa são bem modeladas por funções $1/f$ que originam as distribuições α -estáveis e as funções fractais estatísticas [108, 110].

No sistema de visão humano, a extensão espacial de integração de um neurônio no córtex V1 (células complexas) inclui não somente os campos receptivos “clássicos”, mas também campos receptivos que respondem a estímulos de cauda-longa [35, 97].

Estes campos possuem longas conexões intracorticais com 80% de suas sinapses em neurônios excitatórios (ativadores) [97]. Apesar das células em V1 processarem estímulos de cauda-longa, os estímulos são muito mais claros quando processados pelas células em V2 [81].

A função característica das distribuições α -estáveis é definida como [65]

²Que podem ser de espaço ou de tempo.

$$\varphi(t) = \begin{cases} e^{j\mu t - \gamma|t|^\alpha [1 + j\beta \text{sign}(t) \text{tg}(\frac{\alpha\pi}{2})]} & \text{se } \alpha \neq 1 \\ e^{j\mu t - \gamma|t|^\alpha [1 + j\beta \text{sign}(t) \frac{2}{\pi} \log|t|]} & \text{se } \alpha = 1, \end{cases} \quad (4.7)$$

onde, $-\infty < \mu < \infty$, $\gamma > 0$, $0 < \alpha \leq 2$ e $-1 < \beta < 1$. Os quatro parâmetros α (expoente característico), β (parâmetro de simetria), γ (parâmetro de escala) e μ (parâmetro de localização) determinam uma distribuição α -estável de maneira única e completa. Quando $\alpha = 2$ a distribuição é gaussiana, quando $\alpha = 1$ e $\beta = 0$ a distribuição é a de *Cauchy* e quando $\gamma = 1$ e $\mu = 0$ a distribuição α -estável é dita ser “padrão”.

4.1.4 Fractais e o Modelo de Visão

O sistema de visão humano é ainda bastante complexo para os modelos desenvolvidos. A percepção de texturas naturais é uma das mais complexas abstrações do sistema visual humano. A rugosidade, por exemplo, é uma característica originalmente criada para texturas táteis, não visuais, porém, sua determinação é precedida de um julgamento subjetivo, quando medido pela variação da energia total nos níveis de cinza, abstraindo-se daí um parâmetro tátil.

No mecanismo de visão natural, a rugosidade é obtida através da percepção das irregularidade da superfície do objeto observado de maneira rápida e eficiente. A observação do comportamento irregular de determinadas regiões estimulou Mandelbrot [68] a construir o modelo fractal apresentado na seção 4.1.

Sabe-se hoje que, no sistema de visão humano, os neurônios do córtex visual primário, V1, respondem aos estímulos de imagens naturais com trens de pulsos distribuídos exponencialmente [95], auto-similares [5] e dependentes estatisticamente entre si ao longo do tempo [5].

Além disso, o espectro de Fourier da função de distribuição de probabilidade dos intervalos entre disparos segue a lei de potências $1/f$ [5]. Essas pesquisas mostram, portanto, que o sistema de visão humano possui características que podem ser comparadas com os modelos fractais vistos na seção 4.1.

Os atributos de alta-ordem estatística das imagens naturais são amplamente seletivos para uma orientação simples dos campos receptivos das células em V1 [50], porém, as bordas típicas de imagens naturais (alta-ordem estatística) não são muito similares à banda passante típica das funções de Gabor (ou campos receptivos em V1), devido à grande gama de componentes frequenciais dessas bordas [50].

Uma representação adequada de cada largura de banda de borda deve requerer um agregado com diferentes frequências. É fato que imagens naturais possuem bordas semelhantes a degraus e que não estão contidas em uma única faixa de frequência, sendo necessária a combinação de várias componentes de frequência para representar uma única borda [50].

Essa característica justifica o por que da assinatura espectral dos filtros dos modelos TICA baseados em fractal possuem várias componentes frequenciais em oposição ao modelo TICA baseado na energia, com apenas poucas componentes frequenciais, como será visto na Figura 8.35 (pág. 139).

Observa-se, ainda, que esse preceito mostra que os modelos TICA baseados em fractais possuem assinatura espectral semelhantes àquela predominantes nas imagens naturais, como apresentado na seção 2.2 (pág. 19).

Sabe-se, todavia, que fractais são excelentes descritores de bordas e que existem células em V2 em que o estímulo é ótimo em bordas relativamente finas e semelhantes a degraus [38, 48, 50]. Como citado nos parágrafos anteriores, a representação das bordas em degrau por células em V2 deve possuir um agregado de diferentes frequências capazes de representar adequadamente a transição abrupta da borda.

A existência de múltiplas componentes frequências nos modelos TICA baseados em fractais, associado ao comportamento semelhante entre as células em V2 e modelos fractais, permite fazer uma conjectura, para os modelos elaborados nesta pesquisa, acerca da aproximação de grande parte do comportamento conhecido das células em V2 do córtex visual.

A existência de comportamentos semelhantes entre as células de V1 e de V2 permitem estender a modelagem das células em V1 no modelo ICA para as células em V2 de forma a permitir a melhoria da qualidade de representação do modelo, como a que se propõe nesta pesquisa. Por exemplo, Leventhal et al. [67] demonstrou que uma sub-população de células nas regiões V1 e V2 podem responder a estímulos de borda (semelhantes a uma barra orientada) na forma de sinal invariante. Embora essa propriedade seja rara em V1 ela é predominante em V2.

As células em V1 também são alongadas, porém muito menos que as células em V2 [81], e podem representar bordas com grande largura de banda [50]. Evidências sugerem que a área V2 do córtex replica a estrutura da área V1, mas em uma escala espacial muito grande [38]. O ótimo estímulo em V2 também é mais longo que o observado usualmente em V1, assim, as células em V2 também apresentam alta seletividade em relação à orientação, sendo essa característica, mais pronunciada em V2 que em V1 [48, 50].

Estímulos longos, como os que ocorrem em V2, são descritos por funções de cauda-longa, que representam estatísticas de altas-ordens [95], modeladas por distribuições α -estáveis que, por sua vez, são modelos matemáticos

estatisticamente alto-similares, ou seja, fractais [65].

O córtex secundário, V2, gera forte realimentação para V1 possuindo muitas propriedades em comum com o córtex primário como sensibilidade à orientação e à frequência espacial, porém, em V2 ocorrem interpretações mais complexas que em V1, como a orientação de contornos ilusórios onde o estímulo é proveniente apenas de parte da figura [29].

Pesquisas recentes mostram que as células em V2 têm estímulo ajustado para padrões mais complexos que devem ser dirigidos por múltiplas orientações em diferentes sub-regiões com um simples campo receptivo [29].

A realimentação de V2 para V1 é que torna possível a modificação dos filtros do modelo ICA para adequar o banco de filtros à representação das características percebidas por V2 na imagem analisada. Isso explica por que o banco de filtros TICA baseado em fractais modifica bastante o banco de filtros ICA quando comparado com o banco TICA baseado na energia, pois o modelo fractal busca representar, principalmente, as características das células em V2.

As células no primeiro estágio do processamento cortical (V1 e V2) são capazes de responder a um único estímulo de fronteira de uma maneira complexa, mas que permite a detecção do objeto mesmo enquanto suas definições mudam, ou estão parcialmente oclusas [67].

Leventhal et. al [67], mostrou que, através de uma série de estímulos visuais variando padrões de texturas, luminância e contraste, uma classe de células em V2 possui forte ativação para bordas mais salientes, independente dos sinais que definem os limites de fronteira da imagem. Essas células invariantes ao sinal (*cue invariant* - CI) são raras em V1, mas são facilmente encontradas em V2.

Em seus experimentos, Leventhal et. al. [67], mostrou que muitas células CI em V2 responderam a contornos ilusórios e exibiram o mesmo grau de orientação e direção quando testadas com fronteiras definidas por diferentes sinais. Sua pesquisa sugere que a resposta das células invariantes a sinais (CI) perfazem o primeiro estágio de processamento das células corticais (V2).

Assim, baseado nos trabalhos e relações descritos acima, pode-se afirmar que o sistema de visão humano possui modelos fractais específicos para descrição de características estatísticas de alta-ordem oriundas das informações de bordas da imagem processada no córtex em V1 e, principalmente, em V2. Esta conjectura será usada como um dos suportes aos modelos desenvolvidos nesta pesquisa.

Parte III

Modelagem

Capítulo 5

Elaboração dos Modelos

A Análise de Componentes Independentes, como vista na seção 3.2 (pág. 35), é um modelo que permite a representação de imagens através de um sistema linear de componentes estatisticamente independentes em alta-ordem estatística. A ICA, e a sua variante a TICA, segundo Hyvärinen, Hoyer e Inki [49], modelam adequadamente as células simples e complexas no córtex visual primário, V1.

Todavia, estudos mais elaborados sugerem a existência de características fractais no sistema de visão, como:

- Disparos neuronais por trens de pulsos auto-similares e dependentes estatisticamente entre si ao longo do tempo nas células em V1 [38];
- Estímulo maximizado nas células em V2 para bordas em degrau nas imagens [38, 48, 50, 67]. Sabe-se, todavia, que fractais estatísticos são bons descritores de bordas, podendo modelar esse comportamento de V2;
- Células em V1 alongadas, porém muito menos que as células em V2 [81], mostrando que as células em V2 extraem melhor as estatísticas de alta-ordem de cauda-longa inerente aos modelos fractais estatísticos;
- Há muito mais evidências neurofisiológicas da propriedade bipolar para ativação dos neurônios em V2 que em V1, havendo apenas umas poucas células que efetuam o suporte para ativação/desativação da propriedade bipolar em V1 [38] sugerindo que o critério topográfico da ICA deva ser mais eficiente quando efetuado pelas células em V2 que em V1. A propriedade bipolar¹ corresponde a um sistema de memória que

¹Veja capítulo 1 (pág. 9) e seção 3.3.2 (pág. 49).

permite comparar o estímulo visual no presente com o ocorrido no passado recente perfazendo uma dependência entre os estímulos ao longo do tempo. Esse comportamento pode ser modelado por funções de dependência de cauda-longa, ou seja, funções α -estáveis, que são fractais estatísticos.

- As componentes independentes da ICA aplicada em imagens naturais são aproximações de bordas [11];
- Imagens naturais são bem representadas por modelos fractais. As imagens naturais possuem assinatura espectral com múltiplas componentes freqüenciais, distribuídas quase que uniformemente em todas as direções. Os estímulos de bordas em degrau, disparados por V2, são representados por múltiplas componentes freqüenciais, logo, células em V2 podem ser bem representadas por modelos fractais.

Nesse sentido, o presente trabalho busca adequar o modelo ICA à representação fractal existente no comportamento das células do córtex primário (V1) e secundário (V2) do sistema de visão.

Não devia ser surpresa que o sistema de visão humano tivesse mecanismos descritores fractais para interpretar as imagens da natureza, pois, essas mesmas imagens são melhor caracterizadas por modelos fractais estatísticos, como denota a citação traduzida do Pentland [78]:

“...todavia, a aparência natural dos fractais é uma forte evidência de que eles capturam todas as formas estruturais perceptivas relevantes das superfícies naturais.”

As *formas estruturais perceptivas relevantes*, citadas por Pentland, correspondem ao aspecto (forma estrutural) dos objetos visíveis (perceptíveis) que, por sua vez, são importantes (relevantes) para a interpretação e entendimento das imagens naturais.

A representação da imagem através de filtros estatísticos obtidos do modelo ICA mostra-se bastante adequada à modelagem do sistema de visão, uma vez que é observado que o modelo de visão natural efetua a filtragem estatística das imagens percebidas na retina [22].

Os filtros mapeiam os valores da imagem em cada vizinhança em um *subespaço de perceptividade relevante*², reduzindo o tamanho da representação enquanto preserva a informação estrutural [53]. Por isso, os filtros são adequados para capturar o comportamento estatístico das imagens naturais,

²Ou seja, os filtros capturam apenas as informações que podem ser processadas nas regiões do cérebro preparadas para interpretar essas informações.

uma vez que essas imagens podem ser caracterizadas pelas propriedades de homogeneidade e localidade da representação, ou seja, pelo comportamento estatístico e em uma janela de tamanho limitado [53].

A construção do modelo aqui proposto partiu da formulação de duas hipóteses básicas. A primeira é fundamentada em duas características fractais observadas no sistema de visão: a interação lateral entre as células em V1 e, de forma mais acentuada, em V2 [81], correlacionando as diversas regiões de ativação na vizinhança (essa característica é bem representada pela distribuição de cauda-longa das funções fractais [65, 95]); e evidências neurofisiológicas da propriedade bipolar em V2, mostrando que o critério topográfico pode ser mais eficiente nas células dessa região [38], como visto na seção 4.1.4 (pág. 57), essas células são bem caracterizadas pelo descritor fractal.

Hipótese 1 *Descritores fractais são adequados como critério topográfico no modelo de Análise Topográfica de Componentes Independentes.*

O termo *adequação* refere-se à finalidade do modelo TICA, ou seja, os descritores fractais correspondem melhor às características que inspiraram a elaboração do critério topográfico.

A segunda hipótese é uma consequência da primeira, ou seja, se o modelo TICA baseado em descritores fractais é mais adequado à modelagem do sistema de visão que o modelo TICA baseado na energia, então espera-se uma melhor representação das imagens no modelo proposto.

Hipótese 2 *O uso de descritores fractais no critério topográfico da Análise Topográfica de Componentes Independentes pode facilitar a obtenção dessas componentes, bem como melhorar sua caracterização no modelo, quando as imagens analisadas possuírem características estatísticas fractais.*

Todavia, a formulação da segunda hipótese se baseia principalmente nas seguintes informações: as componentes independentes de imagens naturais são aproximações de bordas [11]; fractais são excelentes descritores de bordas [58, 61, 78, 109, 111]; as células em V2 possuem estímulo maximizado para bordas em degrau das imagens [38, 48, 50].

A TICA, como visto, é um modelo de aproximação do comportamento das células simples e complexas em V1, porém, com a modificação introduzida nessa pesquisa, estende-se a representação do modelo ao comportamento das células em V2.

Tal extensão não aufere modificações drásticas no modelo, uma vez que, as células em V1 e em V2 são conectadas e apresentam comportamentos semelhantes [38], porém as células em V2 apresentam características topográficas mais acentuadas [35, 38, 81, 97] e de estatísticas de mais alta-ordem

que as células complexas em V1 [38, 48, 50]. Entretanto, cada região V1 e V2 possui especializações que se destacam e que são mantidas na modelagem desenvolvida.

Cabe ressaltar que o que se modela na ICA são estágios do processo de visão em que já foram efetuados processamentos anteriores, como por exemplo, o branqueamento realizado pelas células LGN e RGC que conectam a retina ao córtex visual primário.

Ou seja, se uma imagem pode ser bem caracterizada através da modelagem por PCA, ela será resolvida pelas células LGN, e em V1 são removidas apenas as redundâncias de alta-ordem estatística [2, 3, 4, 11, 42, 95].

No mais, o modelo ICA é estritamente baseado no estímulo estatístico e não linearidades e adaptações podem ser adicionados em um processo de refinamento. Todavia, o modelo ICA do córtex visual não é um modelo completo das células simples.

Por ser linear e não adaptativo, muitos aspectos das células simples são ignorados, como a adaptação ao contraste, normalização do contraste, não-linearidades envolvidas no processo de ajuste de orientação e a adaptação a vários estímulos estatísticos.

Por outro lado, o critério topográfico pode contribuir, não de forma residual, mas de maneira expressiva na segmentação não-supervisionada de algumas texturas. Essa afirmação é contrária àquela estabelecida em alguns trabalhos que defendem que o critério topográfico contribui apenas de maneira residual na segmentação de texturas [51, 49, 73].

Ora, uma vez que o critério topográfico modelado aqui não se restringe apenas à organização das células complexas em V1, limitadas apenas ao segundo momento estatístico (como no modelo usando a energia), mas também a altas ordens proporcionadas pelo descritor fractal, que modela o comportamento das células em V2, espera-se que as componentes estatisticamente independentes do modelo possam ser influenciadas pela medida estatística de alta-ordem, incorporada ao critério topográfico, não apenas de forma residual.

Uma melhor caracterização das componentes estatisticamente independentes do modelo deve levar a uma melhor representação das componentes estatísticas formadoras da imagem em análise. Mesmo que o cerne do processamento ICA para a obtenção das componentes estatisticamente independentes, não seja modificado na proposta desta tese, mas somente uma adequação do critério topográfico é efetuada para realçar as características de funcionamento das células em V2, através do uso dos modelos fractais.

Cada hipótese aqui formulada será retomada no momento oportuno, todavia, observa-se aqui que a proposta desse trabalho não é criar uma técnica de segmentação ou de classificação de imagens, mas contribuir com o mo-

delo ICA, de representação de imagens, na modelagem do sistema de visão humano pela inclusão de mecanismos que aproximem a modelagem do comportamento conhecido das células do córtex visual secundário, V2.

5.1 Espectro Fractal Baseado em Morfologia

Uma das formas de se representar uma textura é através da obtenção de características da distribuição de seus pixels. A obtenção das características dos pixels pode ser realizada através do uso de descritores locais. Todavia, a escolha dos descritores depende do tipo de técnica de reconhecimento que é usada no modelo em análise. Sabe-se, ainda, que diferentes tipos de distribuições requerem diferentes técnicas de reconhecimento [12].

De maneira geral, os descritores são sintonizados para extrair relações entre os pixels de acordo com o foco da análise que se queira aplicar na textura. Por exemplo, descritores que acentuam as transições dos valores dos pixels são importantes para o reconhecimento de bordas na imagem, pois nos dão pistas sobre as possíveis fronteiras entre os objetos da cena.

A escolha do descritor adequado torna-se mais fácil quando são envolvidos modelos matemáticos das transformadas bidimensionais aplicadas à imagem, pois, o resultado de uma transformada bidimensional, aplicada em uma zona pequena ao redor de um ponto, pode ser considerado um descritor local associado a esse ponto de forma análoga ao processo de convolução, com a diferença de que, para os resultados serem tratáveis, é preciso diminuir o volume de dados resultantes, tornando-os mais significativos [12].

Por exemplo, no caso de coeficientes de uma transformada, pode-se tomar aqueles de maior valor, como na transformada de Karhunen-Loève, ou ajustar a superfície a uma outra predeterminada cujos parâmetros seriam os descritores locais [12].

A transformada wavelet é um exemplo do ajuste de superfície cujos coeficientes são os descritores, a wavelet mãe define a superfície predeterminada e a escala da wavelet determina a região local. Outros descritores podem ser obtidos de qualquer conjunto de parâmetros de um ajuste local a um determinado tipo de função [12].

Um descritor de texturas adequado para essa pesquisa pode ser feito pelo número de propriedades primitivas que compõem a textura e pela organização espacial da mesma.

Essa organização espacial pode ser aleatória ou possuir uma dependência relacionada a uma vizinhança ou a todas as demais primitivas. Essa dependência pode ser estrutural, funcional ou probabilística [12].

Um descritor que possui essas características, e ainda, possui uma cor-

respondência com o comportamento das células em V2 para a descrição de imagens, baseada em critérios estatísticos da distribuição dos pixels em uma vizinhança da textura, é o descritor fractal.

Um dos primeiros descritores estatísticos fractal foi implementado por Sarkar e Chaudhuri [91] e denominado de *Differential Box-Counting*.

Recentemente, uma adaptação do modelo de Sarkar e Chaudhuri, denominada de Expoente Morfológico Local Multifractal (*Local Morphologic Multifractal Exponent* - LMME) [109] foi implementada com o uso de elementos estruturantes morfológicos como delimitadores da região local.

A determinação do descritor LMME parte do princípio de que uma imagem $N \times N$ é considerada uma superfície 3-D definida como um conjunto triplo $\{i, j, f(i, j); i = 1, \dots, N, j = 1, \dots, N\}$, no caso, uma imagem em escala de cinza. Para uma dada escala, ϵ , e um elemento estruturante, Y_ϵ , é definido um conjunto triplo $\{i_{\epsilon k}, j_{\epsilon k}, \beta\epsilon\}$, onde $k = 1, \dots, P_\epsilon$, sendo P_ϵ o número de elementos em Y_ϵ e β o fator de forma, definindo a "altura" do elemento estruturante, o que permite maior sensibilidade de descrição do modelo.

A dilatação da imagem com Y_ϵ no pixel (i, j) (Figura 5.1) é calculada como:

$$f_\epsilon(i, j) = \max_{k=1,2,\dots,P_\epsilon} \{f(i + i_{\epsilon k}, j + j_{\epsilon k})\} + \beta\epsilon. \quad (5.1)$$

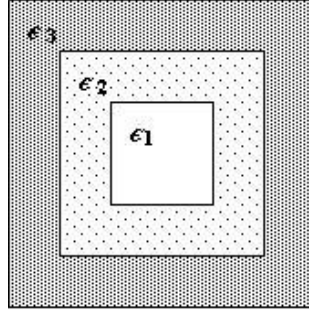
A *medida natural local*, $\mu_\epsilon(i, j)$, que corresponde à dimensão fractal de um pixel, representa a diferença relativa entre o maior tom de cinza dos pixels dentro de um elemento estruturante, centrado na posição (i, j) , após sucessivas dilatações deste, em relação à uma janela da textura de tamanho $W \times W$. A equação 5.2 exibe o modelo matemático dessa medida [109].

$$\mu_\epsilon(i, j) = \frac{|f_\epsilon(i, j) - f(i, j)|}{\sum_{i,j}^W |f_\epsilon(i, j) - f(i, j)|}. \quad (5.2)$$

A *medida de intensidade de ordem q* na escala ϵ , $I(q, \epsilon)$, é calculada conforme a equação 5.3 [109] e representa a intensidade dos tons de cinza da janela, devendo satisfazer a lei de potência das distribuições $1/f$.

$$I(q, \epsilon) = \sum_{i,j}^W \mu_\epsilon(i, j)^q. \quad (5.3)$$

O conjunto de descritores de texturas multifractal, denominado LMME da janela, L_q , é dado pela equação 5.4 [109], e define uma relação entre a *medida de intensidade de ordem q da janela* e a escala do elemento estruturante da

Figura 5.1: Dilatações do elemento estruturante Y_ϵ .

grade da janela. Essa relação é a medida de auto-similaridade estatística da janela.

$$L_q = \frac{1}{1-q} \lim_{\epsilon \rightarrow 0} \frac{\ln I(q, \epsilon)}{\ln \frac{N}{\epsilon}}. \quad (5.4)$$

O LMME da grade, $L_q(i, j)$, também denominado *espectro fractal*, define uma relação entre a *medida natural local* do pixel e a escala do elemento estruturante da grade da janela. Essa relação representa a medida da auto-similaridade estatística na grade. A equação 5.5 exibe a relação [109].

$$L_q(i, j) = \frac{1}{1-q} \lim_{\epsilon \rightarrow 0} \frac{\ln \mu_\epsilon(i, j)}{\ln \frac{N}{\epsilon}}. \quad (5.5)$$

O expoente morfológico multifractal local (LMME) é o descritor da textura definido em duas instâncias, na janela e na grade da janela. O LMME carrega a descrição da região da textura permitindo caracterizá-la através das informações extraídas a partir das medidas de intensidade dos pixels ou através da *medida natural local*.

Cada qual representa um conjunto de primitivas texturais com maior ou menor aproximação, respectivamente, entre o observador e a imagem, realçando características tonais e texturais da superfície analisada. A Figura 5.2 exibe o valor do descritor LMME para cada pixel da imagem da Figura 5.4 e a Figura 5.3 exibe as bordas detectadas na imagem da Figura 5.4 utilizando o método Sobel [36]. Observe que o descritor LMME não se resume apenas a um simples detector de bordas, como pode ser visto comparando-se as imagens das Figuras 5.2 e 5.3.

O descritor LMME é o parâmetro de auto-similaridade estatística pois representa a relação linear entre a *medida natural local* ou a *medida de intensidade de ordem q* e as múltiplas escalas de visualização nas regiões da imagem (grades ou janelas) no espaço logarítmico.



Figura 5.2: Imagem composta pelos descritores fractais estatísticos (LMME) da Figura 5.4.



Figura 5.3: Imagem resultante do processo de detecção de borda aplicado à Figura 5.4 utilizando o método Sobel.



Figura 5.4: Imagem Natural.

O uso de um elemento estruturante morfológico no modelo LMME permite focalizar a busca por estruturas fractais na imagem adequando a forma da região de cálculo do descritor (grade) às formas existentes na imagem de teste de maneira supervisionada.

5.2 Modelagem Baseada em Fractal

A modelagem baseada em fractal, que é abordada nesta seção, corresponde à formulação do uso do modelo fractal como critério topográfico na ICA. O processo de elaboração do modelo parte das mesmas premissas que os autores usaram no desenvolvimento do modelo TICA energia, proposto por Hyvärinen, Hoyer e Inki [49]. Analogias entre a energia e fractais, respeitando as diferenças e semelhanças de cada abordagem, são desenvolvidas nos parágrafos que seguem.

Sabe-se que o uso do critério topográfico no modelo ICA contribui na organização espacial do banco de filtros ICA e que pode ser visto como uma organização topográfica das células em V1 do córtex visual [49]. Isso se deve à caracterização, inserida pelo critério topográfico, de uma *forte correlação entre os pixels vizinhos que é determinada pela função de vizinhança*.

Como apresentado na seção 3.3 (pág. 44), a TICA pressupõe uma correlação de características entre pixels vizinhos de uma imagem, em que *a distância das componentes em uma representação topográfica é uma função da dependência das componentes* [49]. Embora se justifique o uso da correlação linear como a mais básica relação de dependência, para vários métodos de estimação ICA essa medida não é muito útil [51, 49].

Intuitivamente, a correlação mostra que as componentes tendem a serem ativas³, ou seja, conectadas (ver Hyvärinen, Hoyer e Inki [49]), mesmo que seus atuais valores não sejam facilmente preditos a partir das componentes vizinhas.

De maneira análoga à correlação da energia no modelo TICA, um fractal estatístico pressupõe uma correlação entre os pixels vizinhos de uma imagem, dada pela medida estatística constante do fractal, ou seja, pelo parâmetro de auto-similaridade estatística.

A distância das componentes integrantes do fractal, na representação topográfica, é uma função da dependência de cauda-longa inerente ao modelo fractal.

Segundo Hyvärinen, Hoyer e Inki [49], existem várias componentes neurais ativas ao mesmo tempo em uma vizinhança. Isso significa dizer que existem vários pixels da vizinhança contribuindo efetivamente com a energia, ou seja, representando ligações entre si.

De forma semelhante, pode-se considerar como células ativas em um modelo baseado em fractais, os pixels que contribuem efetivamente com o valor de descritor fractal, ou seja, pixels que apresentam ligações entre si compondo um mesmo fractal na região da vizinhança.

Observe que enquanto a energia apenas relaciona pixels que se ligam com mais ou menos intensidade em baixa-ordem estatística, o descritor relaciona pixels que se comunicam segundo uma organização auto-similar estatística de alta-ordem.

A predição dos valores atuais das componentes ativas, tal como na energia, também não são facilmente preditos a partir das componentes vizinhas. Isso porque a estimação de componentes fractais demanda a observação de um número grande de amostras devido à correlação de cauda-longa entre estas para a estimação precisa do parâmetro de auto-similaridade.

Todavia, a predição dos valores de energia, do modelo TICA energia, ou do descritor dos modelos TICA fractais, não é necessária para os modelos, uma vez que tanto a energia quanto o descritor fractal são medidos e não estimados.

A informação relativa à predição dos respectivos valores apenas mostra que o arranjo dos campos receptivos no critério topográfico não é preparado para prever qual o valor de uma componente ativa, baseando-se apenas nas demais da vizinhança, mas sim quais componentes estarão ativas.

De acordo com o que foi apresentado na seção 3.3 (pág. 44), relativo à modelagem do modelo TICA energia, e do desenvolvimento exposto neste ca-

³Do ponto de vista neurobiológico uma célula é ativa quando está transmitindo impulso elétrico às demais células de suas ligações.

pítulo, os parágrafos seguintes apresentam o mecanismo lógico desenvolvido para obtenção do modelo TICA baseado em fractal.

Na seção 3.3 (pág. 44), foi visto que Hyvärinen, Hoyer e Inki [49] propuseram a representação de duas variáveis s_i e s_j topograficamente próximas e linearmente descorrelacionadas como produtos de componentes independentes, z_i e z_j , de média zero e variância unitária, por uma grandeza estatística comum. Como o caso propunha a energia das componentes independentes, a grandeza estatística comum era a variância.

De maneira análoga, propõe-se o caso de duas variáveis s_i e s_j topograficamente próximas, linearmente descorrelacionadas e estatisticamente similares em altas ordens. Pode-se representar s_i e s_j como pertencentes a um mesmo processo estocástico auto-similar, ou seja, um mesmo fractal estatístico.

Dessa forma, análogo à proposição feita por Hyvärinen, Hoyer e Inki [49] ao modelo de energia nas equações 3.16 (pág. 45), pressupõe-se, aqui, que exista uma correlação entre as componentes s_i e s_j dada pelo mesmo parâmetro de auto-similaridade estatística r (veja equação 4.2 pág. 53).

Como o parâmetro de auto-similaridade define o comportamento estatístico de alta-ordem do escalamento [28, 58, 78, 105], então, pode-se retomar ao escopo da ICA para definir s_i e s_j como produtos de componentes independentes, z_i e z_j , de média zero e variância unitária por uma variável estatística comum, r_s , ou seja

$$\begin{aligned} s_i &= z_i r_s \\ s_j &= z_j r_s. \end{aligned} \quad (5.6)$$

Calculando a covariância, de maneira análoga a feita por Hyvärinen, Hoyer e Inki [49], porém, utilizando agora as equações em 5.6, tem-se:

$$\text{cov}(s_i^2, s_j^2) = E[r_s^4] - E[r_s^2]^2 \neq 0, \quad r_s \neq 0. \quad (5.7)$$

Esse resultado mostra que é possível utilizar o parâmetro de auto-similaridade estatística no critério topográfico da TICA para efetuar a organização espacial do banco de filtros considerando as características auto-similares da disposição das componentes estatisticamente independentes do modelo, uma vez que é possível correlacionar as componentes independentes através de sua auto-similaridade. Resultado análogo foi proposto em [18].

Assim, tal como a função u_k no modelo da TICA (equação 3.19 pág. 46), o uso do parâmetro de auto-similaridade, como representação das componentes independentes não negativas de alta-ordem, usadas para gerar a independência estatística, foi proposto no modelo em estudo, resultando na equação 5.8

$$r_s = \phi \left(\sum_{k=1}^n h(i, j) R(s_k) \right), \quad (5.8)$$

onde, r_s corresponde à função de vizinhança de alta-ordem estatística do modelo, ϕ é uma não linearidade escalar, $h(i, j)$ é uma função de vizinhança monotonicamente decrescente em função de alguma medida de distância e $R(s_k)$ é a função baseada no parâmetro de auto-similaridade para uma região da imagem.

Dessa maneira, o modelo TICA usando o parâmetro de auto-similaridade é representado como

$$\log L(W) = \sum_{t=1}^T \sum_{j=1}^n G \left(\sum_{i=1}^n h(i, j) R(\vec{w}_i^T \vec{x}(t)) \right) + T \log |\det W|, \quad (5.9)$$

onde G é a transformação não linear monotônica de números reais positivos, T é o número de observações da variável aleatória $\vec{x}$ e W é a inversa da matriz de mistura.

Nas seções seguintes foram elaborados diferentes algoritmos para a função de vizinhança envolvendo diferentes abordagens baseadas nas teorias de fractais estatísticos. Em cada abordagem é escolhida uma função fractal que será representada pelo termo $R(s_k)$ que será utilizada na equação 5.9.

Como exemplo da caracterização da função de vizinhança no modelo TICA, foi usada a imagem de teste (Figura 5.5), representando uma região de vizinhança simples, para visualização das características realçadas em cada modelo utilizado.

A imagem de teste não tem a função de esgotar o comportamento dos modelos nas imagens naturais, todavia, possui a vantagem de permitir a análise simplificada entre patamares e bordas (elementos básicos constituintes de várias imagens) além de possibilitar algumas comparações simples entre os métodos desenvolvidos.

5.2.1 Modelo Baseado na Energia

A concepção do modelo topográfico na ICA, por Hyvärinen, Hoyer e Inki [49], foi motivada na modelagem das interações laterais entre os neurônios

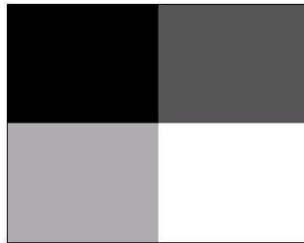


Figura 5.5: Imagem de teste para a função de vizinhança (20x20).

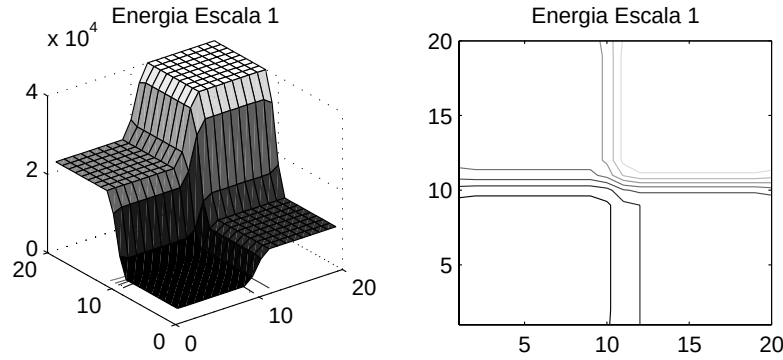


Figura 5.6: Função de Vizinhança Energia. Elemento estruturante “disco” de raio 1.

complexos do córtex V1, ou seja, no mecanismo de dependência residual na estrutura celular do córtex que permite organizar as células simples em V1 e promover um relaxamento da relação de independência do modelo ICA [81].

O modelo de uma função de vizinhança usando a energia foi proposto por Hyvärinen, Hoyer e Inki [49] com intuito de modelar o comportamento das células complexas. Nessa pesquisa, o modelo será denominado de TICAÉ (ou TICA Energia) para diferenciá-lo das demais implementações das funções de vizinhança que são propostas.

A aplicação da função de vizinhança TICAÉ no modelo de região da Figura 5.5 é apresentado nas Figuras 5.6 e 5.7. Para cada pixel da imagem de teste da Figura 5.5, foi calculada a energia da região envolvida pelo elemento morfológico estruturante “disco” de raio 1 (na Figura 5.6) e “quadrado” de lado 3 (na Figura 5.7).

O uso do elemento estruturante na dimensão adotada na Figura 5.6 se deve à escolha do elemento que permita a menor largura de borda possível para essa função de vizinhança.

O motivo é devido a duas características: a primeira vem da necessidade de comparação desse modelo com os demais modelos que serão propostos em que as características de bordas serão mais exploradas. A segunda, remete à inspiração no modelo biológico como visto na seção 3.3.2 (pág 49) que mostra que as células complexas em V1 possuem estímulo maximizado quando a imagem visualizada possui bordas largas, logo, é necessária a observação desse comportamento no modelo.

A imagem à esquerda na Figura (5.6) apresenta os diferentes níveis de energia que representam o mosaico e a imagem da direita mostra as curvas de nível referentes à imagem da esquerda.

Verifica-se que, apesar da dimensão de cada elemento estruturante ser o menor possível para o modelo, a transição de um nível de energia para

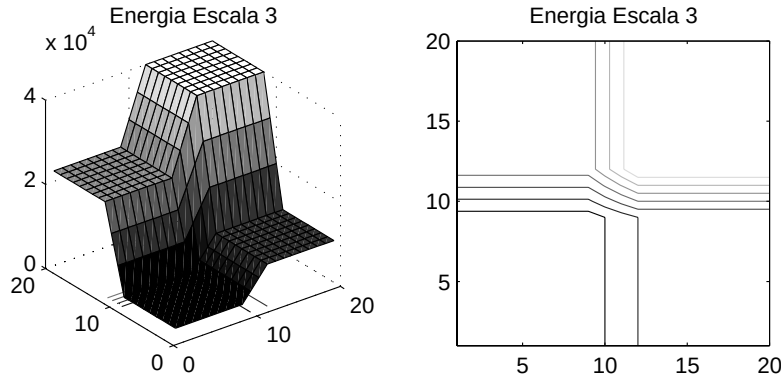


Figura 5.7: Função de Vizinhança Energia. Elemento estruturante “quadrado” de lado 3.

outro envolve até três pixels na região de borda. Ou seja, o modelo exibe um comportamento adequado ao das células complexas em V1 que modelam o estímulo nos neurônios com bordas largas, tal como ocorre no modelo biológico que inspirou a TICA (veja em [38, 48, 50]).

Outra característica da função de vizinhança é a representação das quatro regiões da imagem de teste na forma de patamares de energia (segunda ordem estatística). Isso porque o modelo TICAIE efetua uma organização das células simples através da variância residual obtida da correlação entre as componentes independentes das células simples (componentes da ICA).

5.2.2 Modelo Baseado no Descritor LMME

O uso do descritor LMME no modelo topográfico da ICA é motivado por duas características. A primeira é que fractais estatísticos são excelentes descritores de bordas [58, 78, 111] e a segunda é que as componentes independentes da ICA de texturas naturais são aproximações de bordas [11]. Isso contribuiu para formular as duas hipóteses desta pesquisa.

O modelo da função de vizinhança proposto aqui, utilizando o descritor LMME da seção 5.1 (pág. 65), foi denominado de TICA FM (TICA Fractal Morfológico). O modelo consiste na implementação da função $R(s_k)$ da equação 5.8 (pág. 71) como sendo o espectro fractal definido pela equação 5.5 (pág. 67).

Assim, para cada pixel da imagem de teste da Figura 5.5, o valor do descritor⁴ da região envolvida pelo elemento estruturante “disco” de raio 1 foi calculado e a representação do resultado está na imagem da Figura 5.8.

⁴Nos modelos baseados no descritor fractal deste capítulo foram utilizados os parâme-

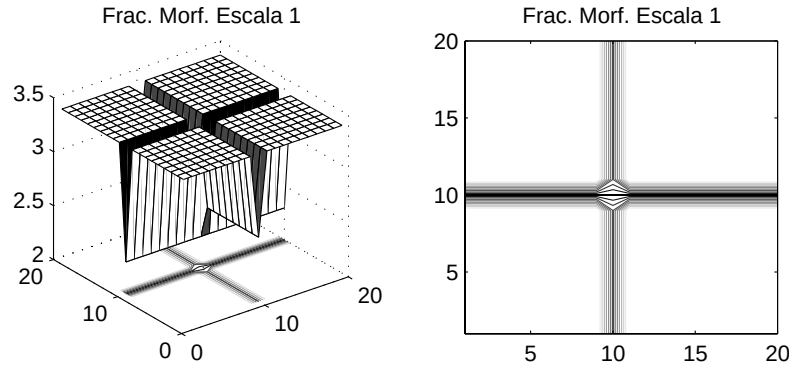


Figura 5.8: Função de Vizinhança: Descritor LMME.

A imagem à esquerda na Figura 5.8 apresenta os descritores LMME que representam o mosaico e a da direita as curvas de nível da imagem à esquerda. Observe que a região de bordas é bem caracterizada pelos descritores fractais, como já era esperado, com apenas um pixel de borda na transição entre as regiões. Porém, a informação relativa aos patamares de energia (baixa-ordem estatística) existentes no modelo TICAÉ não aparecem neste modelo.

Observe, ainda, que a região de bordas aproxima-se do comportamento característico de detecção de bordas, como é conhecido, das células em V2. A função de vizinhança implementada e proposta neste tópico prioriza apenas as informações de fase da textura, uma vez que realça apenas as bordas e ignora a informação de amplitude local dos pixels na vizinhança [19, 21].

Com esse modelo, o critério topográfico realça as características de alta-ordem estatística (que influenciam na organização espacial das células simples em V1) e efetua a inserção de informações estatísticas de alta-ordem, capturadas pelo descritor, nas componentes independentes estatísticas obtidas por essas células.

5.2.3 Modelo Baseado na Energia e no Descritor LMME

O modelo baseado na energia e no descritor LMME é proposto nesta seção como um modelo híbrido oriundo dos modelos TICAÉ e TICAFM apresentados nas seções anteriores. O modelo, então denominado de TICAFE (TICA Fractal-Energia), é construído incorporando os aspectos positivos de cada técnica, quais sejam:

Nas regiões de bordas das imagens, o valor do descritor LMME (dimensão fractal) é sempre menor que nas demais regiões da figura [78]. Todavia,

tros $\beta = 0.1$ e $q = 2$.

o descritor LMME não dispõe de um mecanismo que permita manter as informações dos diferentes patamares de energia relativos aos níveis de cinza das diferentes regiões da imagem.

Por outro lado, o modelo TICAE incorpora informações dos níveis de cinza nas várias regiões da imagem, porém com uma pobre caracterização de bordas.

O modelo fractal-energia para a função de vizinhança foi concebido da junção das características de seletividade de bordas, propiciada pelo descritor LMME, e informação de energia das diferentes regiões da imagem, obtida do modelo TICAE.

O procedimento para realização da extração de características da vizinhança obedece a seguinte heurística [20]:

1. Efetua-se o cálculo do LMME da grade (equação 5.5 pág. 67) para cada pixel de uma vizinhança da imagem, formando uma nova imagem da vizinhança composta apenas dos valores das dimensões fractais das grades;
2. Nessa nova imagem da vizinhança, os valores das dimensões das grades que são inferiores ao valor da dimensão LMME da janela (equação 5.4 pág. 67) são substituídos por zeros. Esses valores determinam as bordas das regiões da imagem;
3. Aplica-se o algoritmo de conectividade-4 na janela, mapeando as regiões separadas pelas bordas. O valor zero é tido como referência de cor de fundo para o algoritmo de conectividade da imagem;
4. Efetua-se uma dilatação por um elemento estruturante com uma máscara 2×2 para eliminar regiões muito pequenas (de 2 ou 3 pixels) envoltas por regiões maiores;
5. Calcula-se a energia das regiões mapeadas atribuindo aos pixels dessas regiões o seu respectivo valor de energia;
6. A vizinhança, agora, está mapeada por regiões e identificada pela energia de seus pixels.

A heurística acima representa a função $R(s_k)$ da equação (5.8) (pág. 71) que constitui a implementação da TICAFE.

Na heurística usada no modelo, os pixels de borda são identificados no algoritmo LMME no momento da transição de nível da imagem de maneira que o pixel identificado como borda irá pertencer a uma das duas regiões da imagem separadas pela borda.

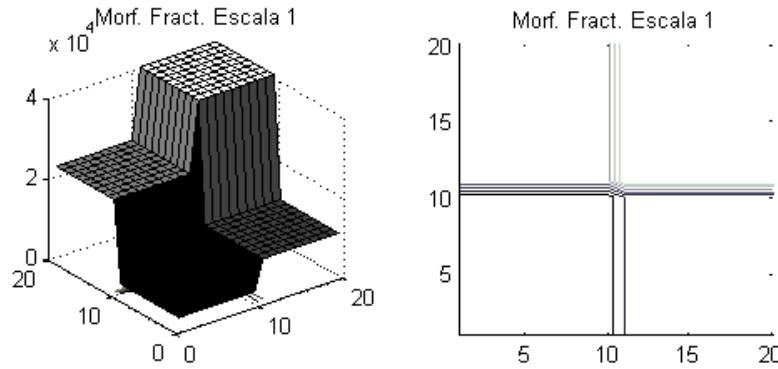


Figura 5.9: Função de Vizinhaça: Fractal-Energia.

No caso do algoritmo de conectividade-4, a varredura é feita da esquerda para a direita e de cima para baixo, logo, os pixels marcados como bordas devem pertencer à textura mapeada nos pixels mais à esquerda e mais acima da borda.

A máscara usada na dilatação e o processo de dilatação, além de eliminar micro-regiões, também contribui para a correção de localização das bordas identificando esses pixels nas texturas imediatamente à esquerda e acima.

A Figura 5.9 mostra o resultado da aplicação da heurística acima na imagem da Figura 5.5. A imagem à esquerda refere-se ao resultado da aplicação da heurística na imagem de teste e a da direita as curvas de nível da imagem à esquerda.

Observe na Figura 5.9 que não há pixel de borda, uma vez que a fronteira entre as diferentes regiões efetuam um salto, transitando de uma região em um pixel para outra região no pixel seguinte.

Essa característica permite modelar a função de vizinhança inspirada no modelo de excitação máxima dos neurônios em V2 [38, 50], e mantém as características de baixa-ordem da imagem, através dos patamares de energia, mantendo a atividade dos neurônios complexos em V1 (tal como inspirado para o modelo TICA Energia).

5.2.4 Modelo Baseado no Descritor LMME Modificado

A proposta de modificação do descritor LMME apresentado nesta seção se deve a duas observações adquiridas durante os experimentos da TICAFE.

No primeiro caso, observou-se um tempo exagerado de processamento na aplicação do modelo TICAFE devido ao número de operações realizadas dentro da heurística proposta. No segundo, a TICAFE insere um erro de um pixel na localização de bordas devido a aplicação do algoritmo de

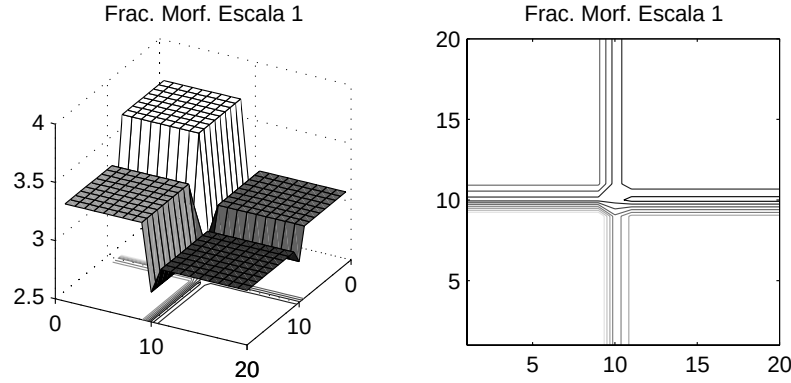


Figura 5.10: Função de Vizinhança: Descritor LMME Modificado.

conectividade-4.

A modificação consiste em não apenas considerar a diferença entre os níveis de cinza da imagem, mas também manter o patamar de referência do maior nível adicionando essa informação através da soma de $f_{\epsilon}(i, j)$ em cada cálculo da diferença entre os tons de cinza da equação da medida natural local (equação 5.2 pág. 66):

$$\mu_{\epsilon}(i, j) = \frac{|2f_{\epsilon}(i, j) - f(i, j)|}{\sum_{i,j}^W |2f_{\epsilon}(i, j) - f(i, j)|}. \quad (5.10)$$

Essa modificação permitiu adequar o erro de localização das bordas, pois não mais é necessário executar o algoritmo de conectividade-4, bem como diminuir substancialmente o tempo de processamento, pois não mais é necessário executar a heurística do modelo TICAFE.

Entretanto, a adequação resultou em uma aproximação da técnica TICAFE e não na sua substituição, como pode ser visto no teste da Figura 5.10.

O modelo da função de vizinhança usando o descritor LMME modificado foi proposto aqui e denominado de TICAFMM (TICA Fractal Morfológico Modificado) para diferenciar das demais implementações das funções de vizinhança. Para cada pixel da imagem de teste da Figura 5.5, o valor do descritor da região envolvida pelo elemento estruturante “disco” de raio 1 foi calculado e a representação do resultado está na imagem da Figura 5.10.

A imagem à esquerda na Figura 5.10 apresenta os descritores LMME, obtidos pela modificação da *medida natural local*, que representam o mosaico. A imagem da direita mostra as curvas de nível referentes à imagem da esquerda.

Observa-se que a transição de um nível de energia para outro envolve dois

pixels nas regiões de borda e que nas bordas o valor da dimensão fractal do descritor modificado é sempre menor que as dimensões fractais nas demais regiões.

Portanto, observa-se que, em relação ao modelo TICAFE, a modificação do descritor LMME permitiu realçar as informações de bordas mantendo a informação dos patamares de energia, todavia, diferentemente do modelo TICAFE, o modelo TICAFMM insere um *offset* nas características extraídas pelo descritor.

O *offset* representa a informação da magnitude dos níveis de cinza de toda a vizinhança, enquanto os patamares carregam as informações de magnitudes relativas entre as diferentes regiões dentro da vizinhança. Portanto, enquanto a TICAFE possui apenas a informação relativa das magnitudes das regiões dentro da vizinhança, a TICAFMM possui ainda uma informação absoluta dos níveis de cinza das diferentes vizinhanças da imagem.

Por outro lado, as regiões de bordas são destacadas no modelo LMME modificado, sendo sempre marcadas com dimensão inferior às dimensões das outras regiões da imagem.

5.2.5 Modelo Baseado no Descritor LMME obtido em Múltiplas Escalas

A modelagem em múltiplas escalas foi idealizada a partir da característica de auto-similaridade dos modelos fractais e da melhor representatividade da intensidade de bordas quando são efetuadas mudanças no tamanho dos campos receptivos do modelo de visão [6, 42, 95].

A primeira característica está vinculada ao comportamento estatístico similar, inerente aos fractais estatísticos, quando observados em diferentes escalas, portanto, a aplicação dos descritores fractais na vizinhança é efetuada a partir de um elemento estruturante que é escalado (decimado) algumas vezes a critério do tamanho da região de aplicação do descritor e do número de decimações definidos *a priori* no modelo.

De acordo com H  yvarinen et al. [47], o tamanho dos campos receptivos das c  lulas de vis  o corresponde ao tamanho dos filtros do modelo ICA. A implementa  o de modelos ICA com bancos de filtros de diferentes dimens  es n  o    efetuada neste trabalho, por  m, a obten  o dos descritores fractais a partir da varia  o do tamanho do elemento estruturante permite uma melhor caracteriza  o de bordas sem, contudo, modificar o tamanho dos campos receptivos do modelo.

A melhor caracteriza  o de bordas, com a varia  o da escala do descritor, ocorre porque as bordas permanecem nos diferentes escalamentos do descritor

fractal.

Neste modelo, para cada escala do elemento estruturante é calculada a *medida de intensidade de ordem q* (equação 5.3, pág. 66) que representa a intensidade dos tons de cinza da janela, devendo satisfazer a lei de potência das distribuições $1/f$.

A função $R(s_k)$ da equação (5.8) (pág. 71) é dada pela equação 5.11.

$$R(s_k) = \tan^{-1} \left\{ \text{fit} \left(\frac{\log I(q, \varepsilon)}{\log \varepsilon} \right) \right\}, \quad (5.11)$$

onde $I(q, \varepsilon)$ é a *medida de intensidade de ordem q* na escala ε . O termo “fit” significa que os valores de $\log I(q, \varepsilon) \times \log \varepsilon$ foram interpoladas por uma reta e o valor de $R(s_k)$ é o ângulo de inclinação dessa reta, que representa um parâmetro único de auto-similaridade para as diferentes escalas da janela dentro da região de vizinhança do modelo TICA.

O modelo da função de vizinhança usando a Equação (5.11) foi denominada de TICAFME (TICA Fractal Multi-Escala) para diferenciar das demais implementações das funções de vizinhança.

Para cada pixel da imagem de teste da Figura 5.5, os valores de $R(s_k)$ para o elemento estruturante “quadrado” de raio 5 com duas decimações foi calculado e a representação do resultado está na imagem da Figura 5.11.

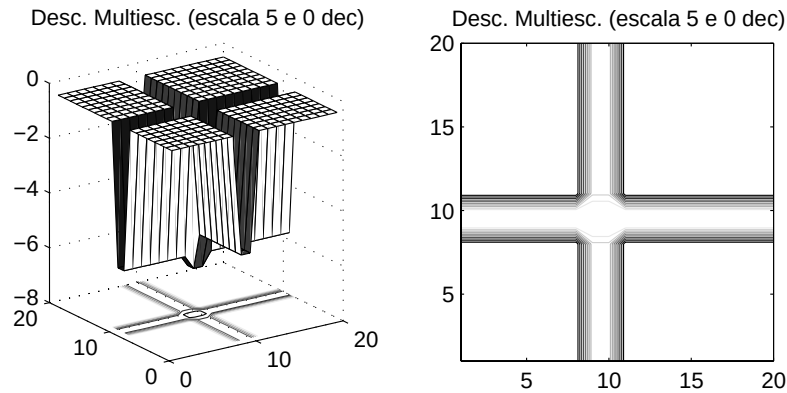


Figura 5.11: Função de Vizinhança: Fractal Multi-Escala. Elemento estruturante com duas decimações.

A imagem à esquerda na Figura 5.11 apresenta os valores de $R(s_k)$, nas diferentes escalas sobre o mosaico de teste. A imagem da direita mostra as curvas de nível referentes às imagens da esquerda.

A Figura 5.12 mostra que o modelo TICAFME é semelhante ao modelo TICAFM quando não são efetuadas decimações no elemento estruturante do descritor fractal.

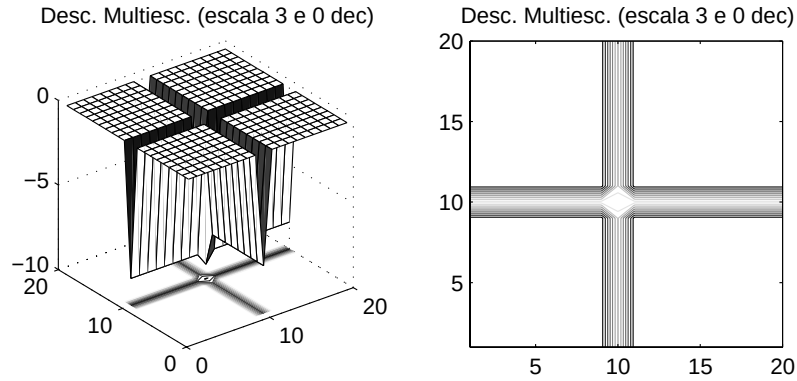


Figura 5.12: Função de Vizinhaça: Fractal Multi-Escala sem decimação do elemento estruturante.

A Figura 5.13 exibe o resultado para uma decimação no elemento estruturante usado na Figura 5.12.

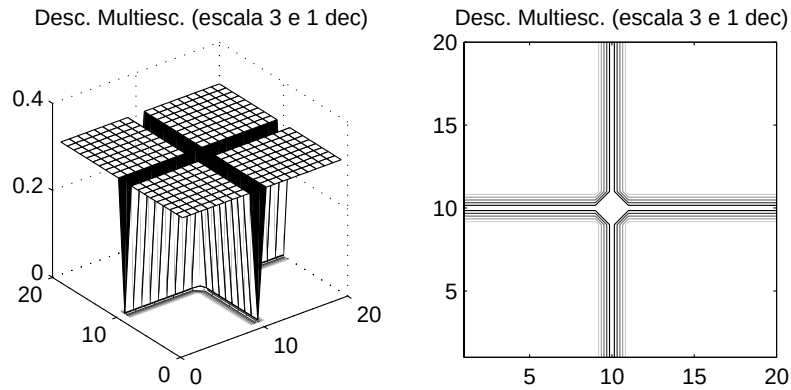


Figura 5.13: Função de Vizinhaça: Fractal Multi-Escala com uma decimação do elemento estruturante.

O resultado da aplicação desse modelo é muito semelhante ao resultado do modelo TICA FM, contudo há diferenças quanto à forma de obtenção das bordas da imagem. Nesta técnica utiliza-se a variação da escala dos campos receptivos (elementos estruturantes do descrito fractal) para detecção das bordas (informações de fase da imagem).

Entretanto, é possível notar que os pixels das bordas (base das figuras) são mais irregulares no modelo TICA FM, porém, a borda é mais fina (menor largura de borda).

Observa-se, ainda, que à medida que são efetuadas mais decimações, a região de borda torna-se mais larga, exibindo mais pixels nessa região, o que é esperado devido à mudança de escala. Esse comportamento define o

modelo como de transição entre o comportamento das células complexas em V1 (detecção de bordas largas) e as células em V2 (detecção de bordas finas).

Por outro lado, a abordagem em múltiplas escalas permite a captura do comportamento dinâmico das texturas ao longo do espaço, ou seja, torna possível obter características de formação da textura através do comportamento similar ao longo de uma trajetória ou de uma sequência de campos de Markov.

5.2.6 Modelo Baseado no Descritor LMME obtido em Múltiplas Cenas

Motivado pelo estudo apresentado por Torralba e Oliva [103] e pelo trabalho elaborado por Xia et al. [109] o modelo multicena utiliza de dilatações do elemento estruturante para cálculo do descritor fractal de cada pixel em uma vizinhança da imagem.

Segundo Torralba e Oliva [103], uma imagem multi-cena corresponde a diferentes distâncias do observador à imagem, proporcionando diferentes visões da mesma cena, com maior ou menor abertura do campo de visão em função da maior ou menor distância, respectivamente, entre o observador e a imagem.

Por outro lado os autores definem uma imagem em múltiplas escalas quando o observador permanece à mesma distância da imagem, porém a mesma é decimada sucessivamente, permitindo múltiplos escalamentos da imagem que está sendo observada.

Nesse sentido, o modelo baseado em múltiplas cenas é elaborado através de variações do tamanho do elemento estruturante na região de vizinhança (sem decimações).

Assim, para cada pixel da imagem de teste da Figura 5.5 os valores de $R(s_k)$, para o elemento estruturante “quadrado” de raio 3 sem dilatação foram calculados e a representação do resultado está na imagem da Figura 5.14.

As sucessivas dilatações do elemento estruturante permitem, ao modelo do descritor fractal, capturar o comportamento estatístico ao longo do espaço na região da vizinhança que está sendo analisada. Neste caso, as dilatações não levam a grandes modificações morfológicas, uma vez que elas se processam em uma região pequena da imagem analisada correspondente à grade de aplicação do descritor fractal.

A Figura 5.15 apresenta o cálculo do descritor fractal utilizando apenas uma dilatação do elemento estruturante.

Todavia, a região envolvida pelo elemento estruturante é muito pequena, mesmo com as sucessivas dilatações, porém suficiente para capturar a “di-

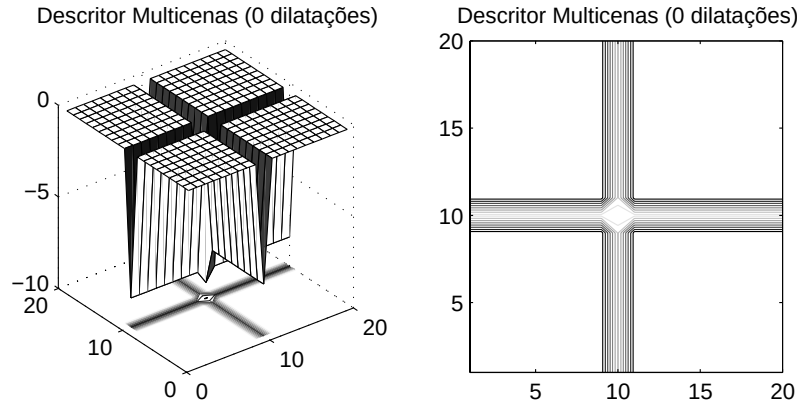


Figura 5.14: Função de Vizinhança: Fractal Multi-Cena sem dilatações do elemento estruturante.

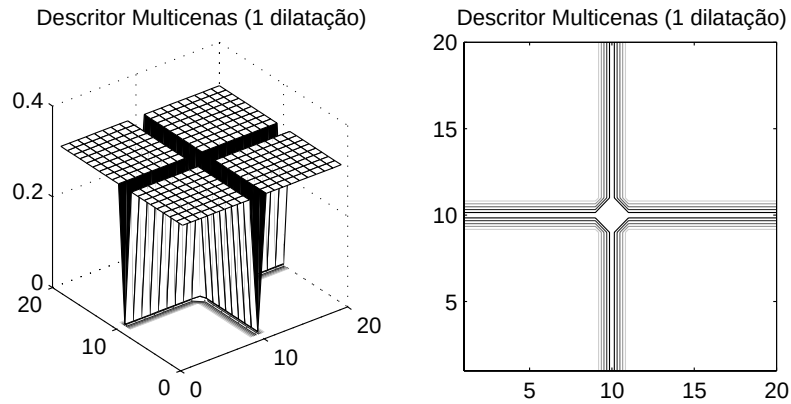


Figura 5.15: Função de Vizinhança: Fractal Multi-Cena com uma dilatação do elemento estruturante.

nâmica” da estrutura estatística da imagem. Devido a essa proximidade do observador⁵ das cenas era de se esperar que o comportamento do modelo, aqui denominado de TICA FMC (TICA Fractal Multi-Cena), fosse muito próximo ao comportamento do modelo TICA FME.

Observe que a semelhança não descaracteriza um dos modelos, pois, de um lado aborda-se a relação entre os tamanhos das cenas e do outro entre as escalas. A diferença entre as caracterizações de cena e escala em imagens é

⁵O termo “observador” é uma abstração usada para entender qual o tamanho da região “visualizada”, ou seja, se o elemento estruturante possui dimensões 7×7 é como se o “observador” pudesse visualizar apenas o conteúdo envolvido pelo elemento estruturante, pois seu campo de “visão” estaria restrito apenas à essa região devido a proximidade com a imagem.

bem representada por Torralba e Oliva [103].

Por outro lado, a simplicidade da imagem de teste não permite explorar profundamente as diferentes nuances de cada abordagem proposta, mas a escolha de um modelo simples permite explicitar as características principais de cada técnica (caracterização de bordas e patamares). Apenas os testes em imagens naturais poderão ressaltar comportamentos característicos de cada proposta elaborada nesta pesquisa.

Observando as imagens das Figuras 5.14 e 5.15 e comparando com a Figura 5.8 (pág. 75) observa-se uma forte semelhança na forma da vizinhança, porém, com diferenças na magnitude dos descritores.

Isso se deve, apesar de cada técnica possuir mecanismos distintos para o cálculo dos descritores fractais, à simplicidade da imagem de teste utilizada, que proporciona a obtenção dos mesmos descritores ou de uma aproximação entre seus valores em cada abordagem.

5.2.7 Modelo Baseado na Lacunaridade Fractal

Como visto na seção 4.1.1 (pág. 53), a lacunaridade é uma medida estatística de alta-ordem que permite identificar “lacunas” nas imagens testadas.

As bordas, são exemplos de “lacunas”, todavia, é necessário ressaltar que a referência usada para definir uma “lacuna” está na componente estatística da região e não no decaimento aparente dos níveis de cinza dessa região.

A Figura 5.16 exemplifica essa conotação. Observe que os patamares de menor nível de cinza das quatro regiões da imagem de teste da Figura 5.5 não são mapeados como “lacunas”, mas somente as bordas, pois estas apresentam características de maior magnitude nas componentes estatísticas da figura.

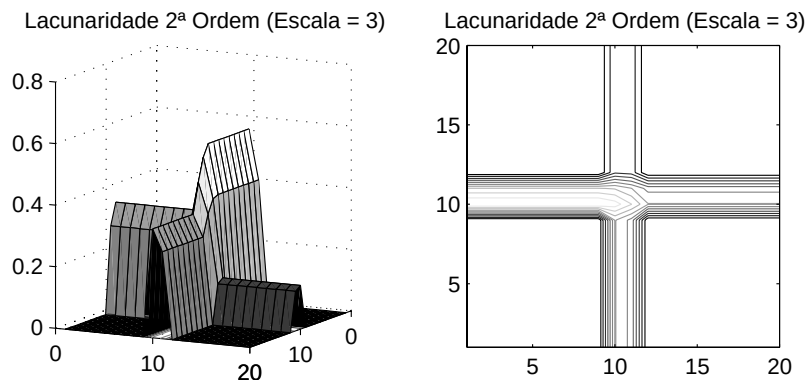


Figura 5.16: Função de Vizinhança: Lacunaridade Fractal obtida pela segunda ordem estatística.

Para cada pixel da imagem de teste da Figura 5.5, os valores da lacuna-

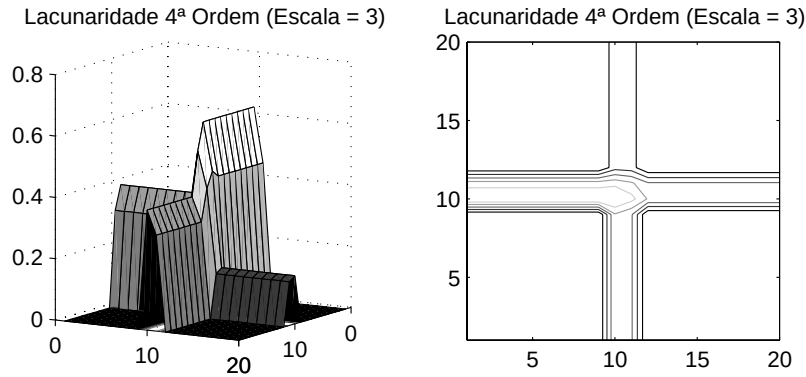


Figura 5.17: Função de Vizinhança: Lacunaridade Fractal obtida pela quarta ordem estatística.

ridade, $R(s_k)$, foram calculados usando informações estatísticas de 2ª ordem percorrendo a vizinhança com um elemento estruturante “quadrado” de lado 3 e a representação do resultado está na imagem da Figura 5.16.

A Figura 5.17 mostra os valores das lacunaridades fractais calculados na imagem de teste utilizando estatísticas de alta-ordem. Observe que tanto a Figura 5.16, que apresenta a lacunaridade calculada com medidas estatísticas de segunda ordem, quanto a Figura 5.17, obtida através de medidas de quarta ordem estatística, apresentam valores de lacunaridade muito próximos. Isso se deve, basicamente, ao fato de as bordas possuírem informações estatísticas em todas as ordens [25].

Nota-se que a medida de lacunaridade fractal caracteriza bem as informações de bordas na imagem testada. Comparando os valores da lacunaridade com os do descritor fractal da Figura 5.8 (pág. 75) observa-se uma certa correspondência nos valores das regiões de bordas [21].

Neste trabalho, o uso da lacunaridade no critério topográfico é simbolizado pelo termo TICALF (TICA Lacunaridade Fractal), sendo a forma TICALF2 correspondente ao uso da segunda ordem estatística e a forma TICALF4 à quarta ordem.

5.2.8 Modelo Baseado no Coeficiente de Variação Espacial Estatístico

A função $R(s_k)$ no modelo topográfico é substituída aqui neste modelo pelos coeficientes de variação espacial estatísticos, CVEE, definidos na equação 4.6 (pág. 56).

Para cada pixel da imagem de teste da Figura 5.5, os valores dos coeficientes de variação espacial estatístico, $R(s_k)$, foram calculados percorrendo

a vizinhança com um elemento estruturante “quadrado” de lado 3 (escala do elemento estruturante) e a representação do resultado está na imagem da Figura 5.18.

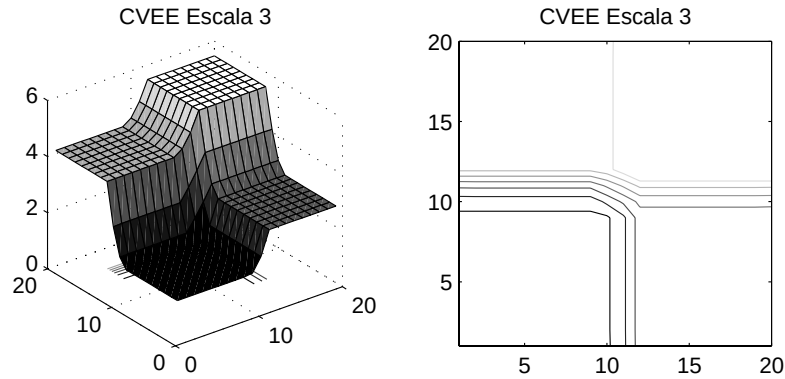


Figura 5.18: Função de Vizinhança: Coeficiente de Variação Espacial Estatística.

Observa-se, nesta representação da vizinhança pelos coeficientes de variação espacial estatísticos, uma semelhança muito grande na forma de caracterização dos patamares e bordas da imagem com o modelo de energia proposto por Hyvärinen, Hoyer e Inki [49] nas Figuras 5.6 e 5.7 (pág. 73).

Esse resultado era esperado uma vez que o CVEE é baseado em estatísticas de 1ª e 2ª ordem e trata a variação espacial da intensidade da imagem, o que também é bem delineada pela energia.

Neste trabalho, o uso do CVEE como critério topográfico no modelo TICA é representado por TICAVEE (TICA Variação Espacial Estatística) para diferenciar dos demais modelos apresentados.

Parte IV

Experimentos

Capítulo 6

Aplicação dos Modelos em Segmentação

Como visto na seção 3.2 (pág. 35) a ICA é uma técnica de separação cega de fontes (ou seja, é um método não-supervisionado) que permite representar sinais e imagens como uma combinação linear matricial de componentes estatisticamente independentes em altas-ordens estatísticas. Essa representação é feita através do uso de um banco de filtros linear denominado de matriz de mistura.

O uso do banco de filtros para a segmentação estatística não-supervisionada de imagens é um exemplo de aplicação do modelo TICA no reconhecimento de padrões [53].

Outros usos são atribuídos à representação de imagens pela TICA como, por exemplo, uso em pré-processamento para extração de características de interesse na análise de imagens [62] e classificação de imagens [13].

Nesta pesquisa, a segmentação não-supervisionada é usada para verificar as características dos modelos topográficos propostos frente ao modelo topográfico original da TICA baseada na energia. Esta nova abordagem traz qualidades ainda não exploradas na ICA, permitindo melhor caracterização das imagens representadas pelo modelo através da melhor sintonia dos filtros obtidos pela TICA, como será visto adiante.

A comparação com a TICA-E é feita porque a modelagem proposta visa uma melhoria na abordagem topográfica do modelo ICA, e não uma solução de segmentação de imagens que modifique completamente o algoritmo da análise de componentes independentes, uma vez que apenas o critério topográfico é modificado.

A segmentação não-supervisionada é aqui utilizada apenas como um exemplo de aplicação do modelo proposto para representação de imagens. A comparação de técnicas de segmentação com a técnica proposta não faz sentido

uma vez que a comparação estaria fortemente vinculada ao comportamento do algoritmo ICA, ou seja, seria comparar a ICA com as demais técnicas de segmentação. O critério topográfico melhora a qualidade de representação do modelo ICA, porém não causa modificações no cerne do modelo a ponto de descaracterizar a separação cega de fontes, que é o que a ICA elabora.

Por outro lado a ICA é uma representação de imagens em uma estrutura linear. Essa estrutura pode ser utilizada em uma série de aplicações como: extração de características de imagens [62], classificação de imagens [51] e segmentação estatística não-supervisionada [53].

Porém, existem situações em que a ICA não provê melhores resultados que outras técnicas de segmentação, por exemplo, nas imagens que possuem muita informação de baixa ordem estatística, técnicas como PCA apresentam melhores resultados [44].

Imagens contendo uma parcela de informações de baixa ordem estatística serão testadas nos experimentos que serão realizados mais à frente neste trabalho. Porém, o propósito aqui não é testar o funcionamento da ICA em baixa ordem estatística. Comentários acerca desses experimentos podem ser encontrados em [51].

A comparação da TICA energia com os modelos TICA fractais, aqui propostos, busca verificar se nas imagens fractais estatísticas onde a TICA energia possui bons resultados, as técnicas aqui propostas serão melhores. E, onde a TICA energia não consegue um bom modelo de representação, é porque a imagem não é adequada para ser representada por um modelo que utiliza estatísticas de alta-ordem.

A escolha da segmentação não-supervisionada para validar os modelos se deve, basicamente, à complexidade do processo de separação e identificação de regiões em imagens naturais.

A partir da obtenção do modelo TICA, que representa o mosaico de imagens de teste, temos a matriz de mistura do modelo como um banco de filtros estatísticos obtido de maneira não-supervisionada.

No modelo de banco de filtros, cada filtro responde a uma propriedade específica da textura, possuindo alta energia na região em que está sintonizado para filtrar e baixa energia nas demais regiões. Todavia, sobreposições de bandas entre filtros podem ocorrer, tornando o banco de filtros com várias componentes redundantes.

Visando minimizar a sobreposição de bandas, cada filtro do banco foi normalizado em energia e, para cada par de filtros da base, foi computada a norma entre eles. Os filtros com menor sobreposição de frequências foram selecionados.

A imagem a ser segmentada é submetida ao banco de filtros gerando uma componente da imagem para cada filtro (imagens características). Em

seguida, é aplicada a normalização logarítmica não linear, proposta por Unser e Edem [106], para cada componente da imagem resultante de cada filtro. A equação 6.1 representa esta operação.

$$e_k(i, j) = \log \left\{ \frac{1}{N^2} \sum_{a, b \in J} I_k^2(i - a, j - b) J(a, b) \right\}, \quad (6.1)$$

onde, I_k é a k -ésima imagem filtrada, J é a janela gaussiana ($N \times N$), em que $N = 2\sigma$, $\sigma = \frac{2}{\sqrt{3}}f_0$ e f_0 é a frequência espacial dominante estimada.

Cada imagem característica, e_k , resultante da normalização logarítmica, é submetida ao F -teste proposto por Kasparis et al. [61], que avalia se a imagem característica possui ou não variabilidade sobre a textura. O F -teste é efetuado pela equação 6.2.

$$F_k = \frac{(M - R)(\text{RSS}_k - \text{RSS}_{kT})}{(R - 1)\text{RSS}_{kT}}, \quad (6.2)$$

onde, M é o número total de pixels em cada região, k denota a k -ésima imagem característica e RSS é a soma residual dos quadrados. Os demais valores da equação 6.2 são obtidos da seguinte forma: cada imagem característica é dividida em R regiões igualmente espaçadas compondo Ω_ρ , $\rho = 1, 2, \dots, R$. Para cada região é calculada a *soma residual dos quadrados* ($\text{RSS}_{k\rho}$) (eq. 6.3),

$$\text{RSS}_{k\rho} = \sum_{i, j \in \Omega_\rho} [e_k(i, j) - m_{k\rho}]^2, \quad (6.3)$$

onde, ρ é a região, k denota a k -ésima imagem característica e $m_{k\rho}$ é a média das características na região Ω_ρ .

O valor de RSS_{kT} é obtido pela soma das regiões de $\text{RSS}_{k\rho}$ para cada k

$$\text{RSS}_{kT} = \sum_{\rho} [\text{RSS}_{k\rho}]. \quad (6.4)$$

O valor de RSS_k é obtido para a imagem característica toda (sem regiões) conforme a equação 6.5,

$$\text{RSS}_k = \sum [e_k - m_{ek\rho}]^2, \quad (6.5)$$

onde $m_{ek\rho}$ é o valor médio da imagem e_k .

A equação 6.2 mostra ainda que se RSS_{kT} é próximo de RSS_k , então a imagem característica k não possui informação local.

Após efetuado o F -Test as imagens características com pouca variabilidade sobre a textura são eliminadas do espaço característico. As imagens selecionadas para compor as características para segmentação não-supervisionada

são submetidas ao algoritmo *k-means* [27]. Outros algoritmos poderiam ser utilizados, todavia, na literatura consultada para segmentação fractal e segmentação usando ICA, predomina o uso desse algoritmo.

Alguns gráficos de erros de segmentação foram construídos variando o número de filtros utilizados no banco. A poda dos filtros ocorreu de acordo com os critérios definidos acima (sobreposições de frequências e *F-Test*).

O erro de segmentação das imagens sintéticas e texturas de Brodatz, Figuras 7.3 (pág. 95) e 7.4 (pág. 96), foi calculado contando o número de pixels mapeados fora das regiões em que deveriam estar localizados. O número de pixels contados é comparado com o número total de pixels da imagem e exibido graficamente em uma escala percentual.

Os bancos de filtros foram obtidos pela aplicação dos modelos TICAE (energia), TICAFM (descriptor LMME), TICAFE (energia + descriptor LMME), TICAFMM (descriptor LMME Modificado), TICAFME (descriptor LMME Multi-Escalas), TICAFMC (descriptor Multi-Cenas), TICALF (Lacunaridade Fractal) e TICAVEE (coeficiente de Variação Espacial Estatístico), como afirmado anteriormente.

Análises de desempenho elaboradas ao longo dos testes resultaram na Tabela 6.1 com a relação de desempenho para cada um dos algoritmos implementados.

Técnica	Tempo [s]	Relação	Otimização
TICAE	26,79	1,05%	não otimizado
TICAFM	60,87	2,40%	otimizado
TICAFMM	60,04	2,36%	otimizado
TICALF2	128,44	5,06%	otimizado
TICALF4	131,08	5,16%	otimizado
TICAFMC	2.386,30	93,98%	não otimizado
TICAFME	2.539,10	100,00%	não otimizado
TICAFE	895,28	35,26%	otimizado
TICAVEE	2.375,40	93,55%	não otimizado

Tabela 6.1: Desempenho médio de processamento para 2000 amostras 8×8 da mistura em 100 iterações.

As simulações foram efetuadas no MatlabTM [71] em um computador Pentium D 3.0 GHz com 1.0 GB de memória RAM DDR2. O algoritmo mais lento foi o do modelo TICAFME referenciado com 100% do tempo de processamento. Nas demais implementações o tempo de processamento foi comparado ao tempo da TICAFME e consta a relação na coluna da tabela.

Os modelos marcados como otimizados foram parcialmente implementa-

dos em C dentro do ambiente do MatlabTM. Os demais algoritmos foram elaborados exclusivamente com a linguagem do MatlabTM. A utilização dos métodos propostos, ou mesmo, do algoritmo ICA em sistemas que precisam de desempenho elevado obriga a implementação dos algoritmos em hardware. Isso, porque, não existe algoritmo rápido para ICA, uma vez que a base gerada pelo modelo não é ortogonal.

Capítulo 7

Preparação

Para exemplificar a aplicação dos métodos nesta pesquisa, foram utilizadas imagens artificiais (Figura 7.1), imagens estatísticas sintéticas (Figura 7.3), imagens naturais (Figura 7.2) e imagens de Brodatz [14] (Figura 7.4) com texturas naturais e artificiais. As imagens sintéticas (Figura 7.3) foram criadas utilizando o algoritmo de Pearson [71].

A imagem sintética da Figura 7.3a foi construída com uma organização de campos markovianos da forma

$$c(i, j) = q(i, j) + \frac{1}{2}c(i-1, j-1) + \text{pearsrnd}(m1, m2, m3, m4), \quad (7.1)$$

onde, $c(i, j)$ é o tom de cinza do pixel na posição (i, j) , pearsrnd é a função do MatlabTM [71] geradora de amostras aleatórias que segue a distribuição definida pelos quatro momentos estatísticos $m1$, $m2$, $m3$ e $m4$.

A imagem sintética da Figura 7.3b foi construída sem qualquer dependência espacial por campos markovianos. Cada pixel recebeu o tom de cinza baseado apenas na distribuição de valores do algoritmo de Pearson [71], ou seja, cada pixel é completamente independente espacialmente dos seus vizinhos.

A escolha das diferentes imagens propostas nos testes desta pesquisa se deve à medida de regularidade estatística contida em cada imagem. As imagens com características estatísticas fractais devem prover melhores resultados do modelo TICA baseado em fractais.

A mistura do modelo TICA foi montada com quinhentos a quatro mil janelas 8×8 aleatoriamente selecionadas para cada imagem, perfazendo bancos de filtros da matriz de mistura com sessenta e quatro filtros. A seleção aleatória das janelas é feita para evitar qualquer correlação temporal.

Os parâmetros ajustáveis em cada abordagem foram mantidos constantes na maioria dos experimentos e fixados em: $\beta = 0.1$ (eq. (5.1), pág. 66),

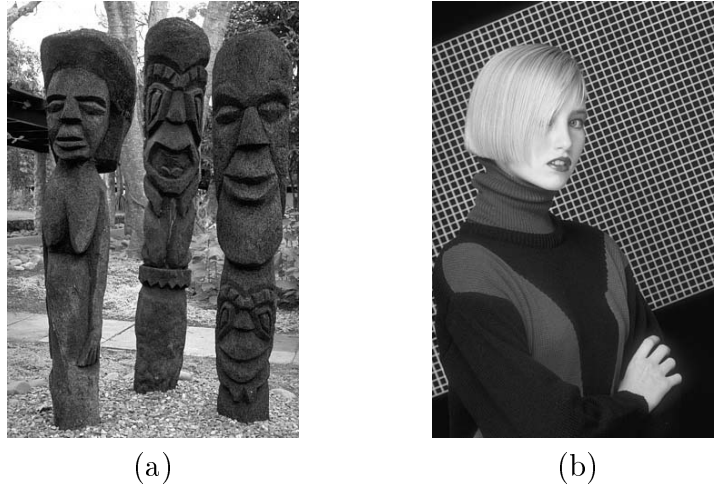


Figura 7.1: Imagens artificiais. (a) Totens (b) Mulher.

outros valores poderiam ser escolhidos interferindo pouco na sensibilidade do descritor; $q = 2$ (eq. 5.3, pág. 66), a escolha foi baseada nos experimentos em [109]. Variações nesses parâmetros, quando ocorrem, são informados nas simulações das imagens de teste.

Em todas as abordagens os elementos estruturantes utilizados foram o “disco” e o “quadrado”. o tamanho do elemento estruturante foi mantido em três e, nas técnicas TICAFME e TICAFMC o tamanho foi cinco, para permitir o escalamento.

Novamente, alterações nesses parâmetros são informados na apresentação dos resultados dos testes. As iterações do algoritmo ICA eram interrompidas quando a matriz de mistura se mantinha invariante ao longo de dez iterações consecutivas, sendo esse o critério de parada do algoritmo.

Experimentos anteriores mostraram que a estabilização das mudanças da matriz de mistura do modelo ICA ocorriam na grande maioria das vezes quando não mais haviam mudanças em dez iterações consecutivas.

Para cada uma das doze imagens de teste foi gerado um banco de filtros, correspondente à matriz de mistura, para cada uma das sete técnicas propostas, além da técnica existente baseada na energia.

Para a lacunaridade foram gerados dois bancos de filtros, sendo um para o modelo da lacunaridade de segunda ordem e outro para quarta ordem estatística. Dessa forma, foram nove bancos gerados para cada imagem, perfazendo um total de cento e oito bancos de filtros.

A Figura 7.5 mostra o espectro de aproximação $1/f$ para cada uma das imagens das Figuras 7.1 e 7.2.

A Figura 7.6 mostra o espectro de aproximação $1/f$ para cada uma das

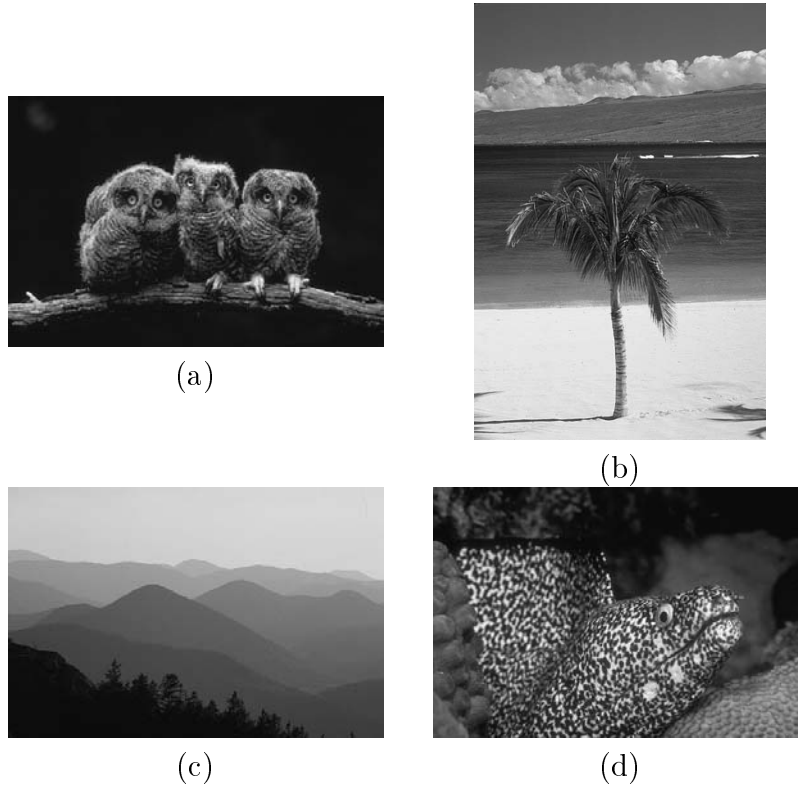


Figura 7.2: Imagens naturais. (a) Corujas (b) Praia (c) Montanhas (d) Peixe.

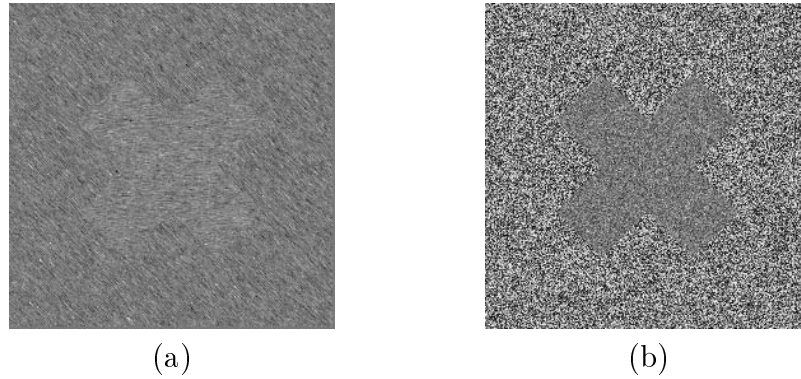


Figura 7.3: Imagens com texturas estatísticas sintéticas. (a) Imagem estocástica com cumulantes (a notação é $C_{\text{ordem}}^{\text{textura}}$) $C_1^1 = C_1^2 = 0$, $C_2^1 = C_2^2 = 1$, $C_3^1 = 0,01$ e $C_3^2 = -0,01$, e $C_4^1 = 2$ e $C_4^2 = 3$; e (b) Imagem não estocástica com médias $\mu_1 = \mu_2 = 0$, variâncias $\sigma_1^2 = \sigma_2^2 = 2$, skewness $skew_1 = 0$, $skew_2 = 0,1$ e kurtosis $kurt_1 = 2$ e $kurt_2 = 3$, onde a textura um é a externa e a textura dois é a interna do mosaico (256×256)

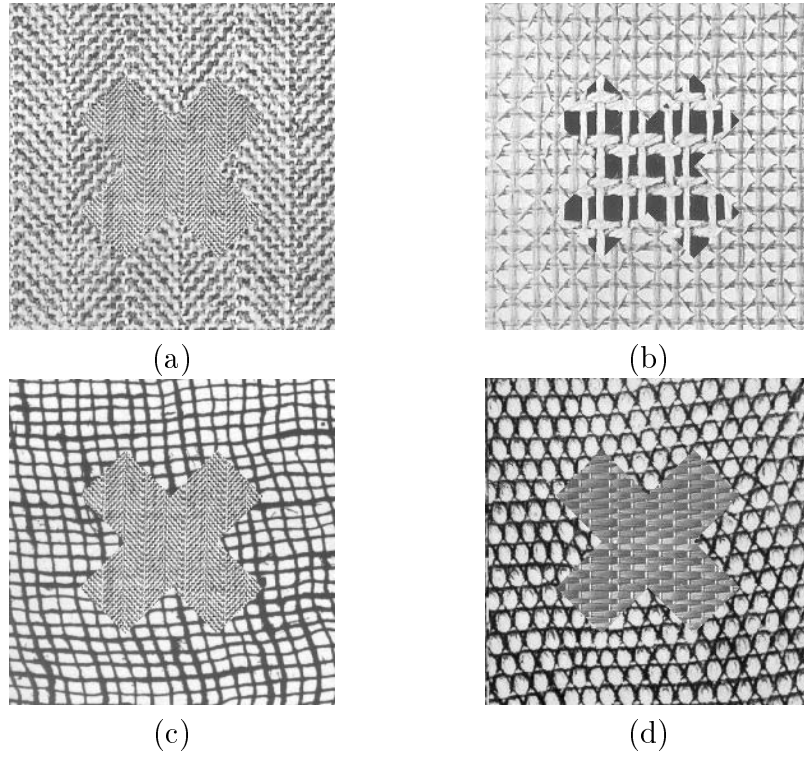


Figura 7.4: Imagens de Brodatz (256×256) com as texturas: (a) D17 (interna) e D16 (externa), (b) D52 (interna) e D20 (externa), (c) D16 (interna) e D103 (externa) e (d) D55 (interna) e D22 (externa)

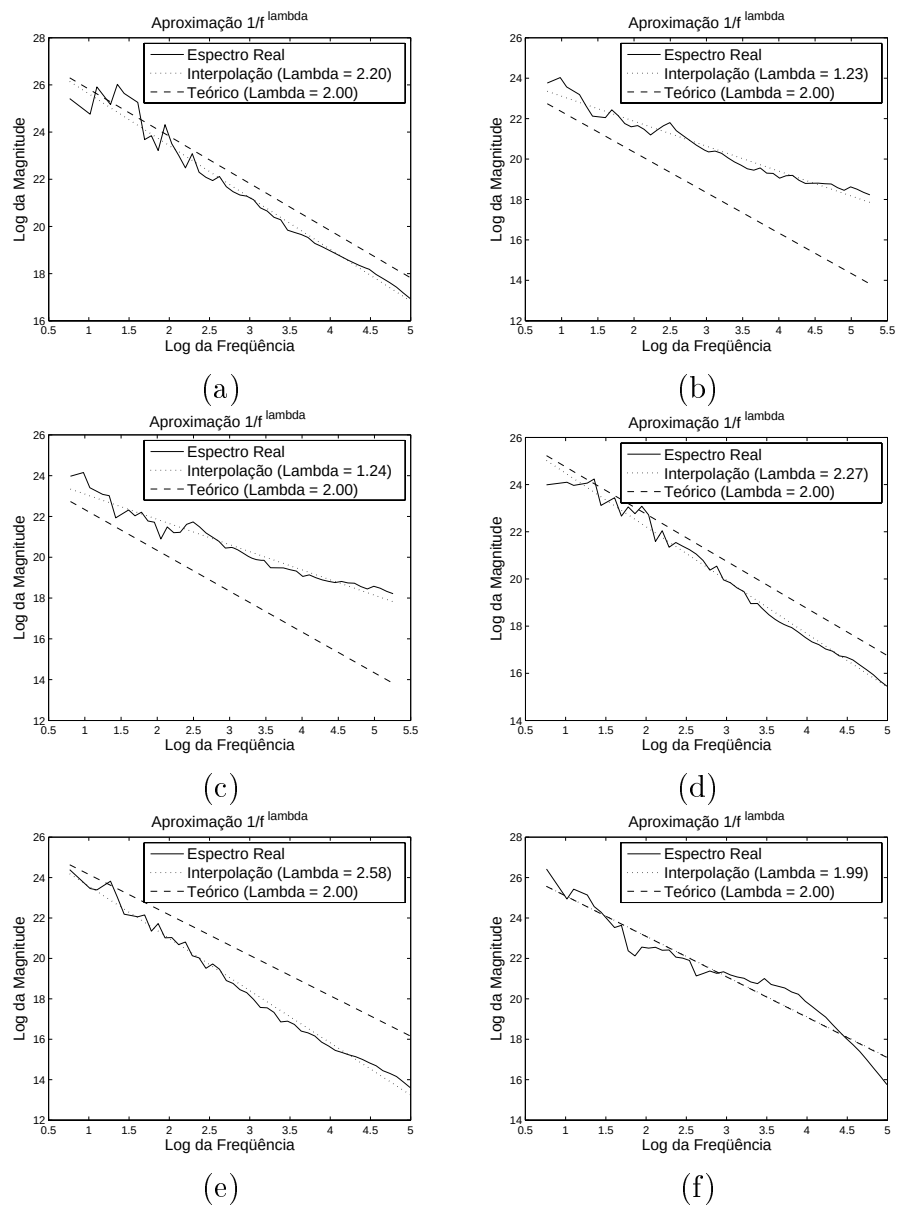


Figura 7.5: Espectros $1/f$ das imagens. (a) Toten (b) Loira (c) Corujas (d) Praia (e) Montanhas (f) Peixe.

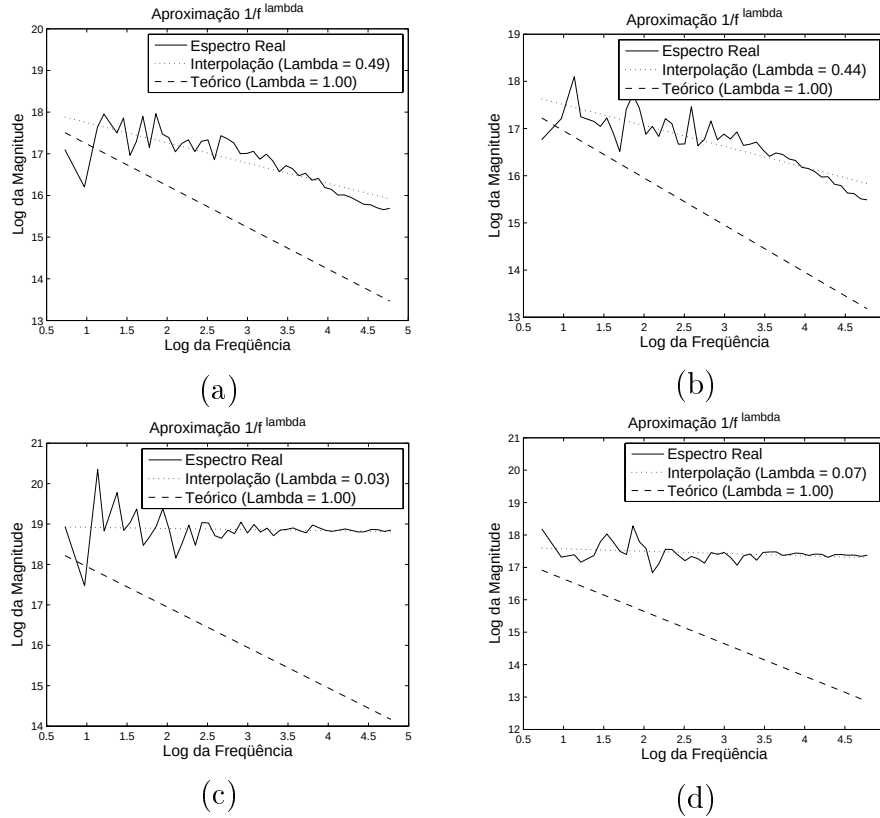


Figura 7.6: Espectros $1/f$ das texturas estatísticas sintéticas. (a) Figura 7.3a Externa (b) Figura 7.3a Interna (c) Figura 7.3b Externa (d) Figura 7.3b Interna.

texturas da Figura 7.3.

A Figura 7.7 mostra o espectro de aproximação $1/f$ para cada uma das texturas da Figura 7.4.

Como visto na seção 1 (pág. 2), o espectro $1/f^\lambda$ permite caracterizar uma imagem como sendo fractal estatística [108]. No mesmo sentido, na seção 2.2 (pág. 19) é apresentada a relação entre λ e a regularidade da superfície das imagens [95].

Nota-se, então, que de cordo com os espectros $1/f$ da Figura 7.5 (pág. 97), nas imagens das Figuras 7.1 (pág. 94) e 7.2 (pág. 95) há uma predominância fractal com $\lambda \approx 2$, logo, espera-se uma boa qualidade de representação do modelo TICA baseado em fractais nestas imagens.

Nas imagens sintéticas da Figura 7.3 (pág. 95), as texturas construídas com campos de Markov possuem pouca (Figura 7.3a) ou nenhuma (Figura 7.3b) característica fractal, como observado nas imagens da Figura 7.6 (pág. 98) portanto, neste caso deve prevalecer o comportamento do algoritmo ICA frente ao critério topográfico baseado em funções fractais, com pouca ou

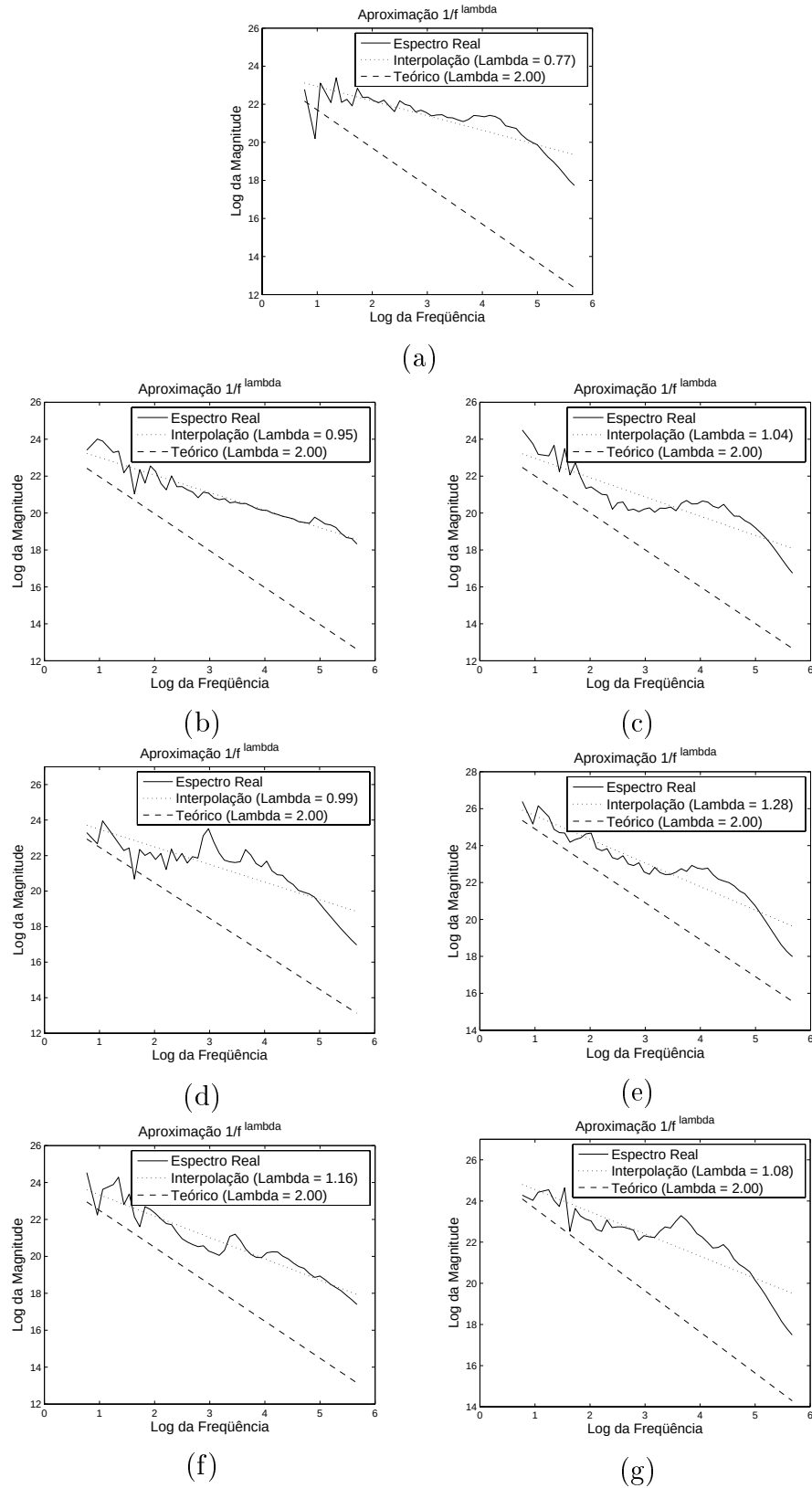


Figura 7.7: Espectros $1/f$ das texturas de Brodatz. (a) Figuras 7.4a e 7.4c Interna (D17) (b) Figura 7.4a Externa (D16) (c) Figura 7.4b Interna (D52) (d) Figura 7.4b Externa (D20) (e) Figura 7.4c Externa (D103) (f) Figura 7.4d Interna (D55) (g) Figura 7.4d Externa (D22).

quase nenhuma influência dos modelos fractais para a representação das texturas.

Os espectros de frequência das texturas de Brodatz [14] na Figura 7.7 (pág. 99) mostram que o critério topográfico fractal exercerá certa influência na segmentação não-supervisionada das texturas uma vez que os espectros de potências são em torno de $1/f^1$.

Imagens com forte contribuição fractal (e, portanto, do critério topográfico fractal) possuem $\lambda \cong 2$ [95].

Nas imagens artificiais e naturais das Figuras 7.1 e 7.2 foi usada a subtração das imagens segmentadas pelos métodos TICA baseados em fractais com a TICA Energia. Esse procedimento visa permitir uma melhor visualização das diferenças de segmentação proporcionadas por cada modelo.

As tonalidades mais escuras correspondem à regiões segmentadas pelos modelos TICA Fractal e que não foram segmentadas como a mesma região pela TICA Energia. Os pontos mais claros correspondem à regiões que foram segmentadas pela TICA Energia, mas que não foram segmentados como a mesma região pelos modelos TICA Fractal.

Capítulo 8

Resultados e Análises

As imagens da Figura 7.1 (pág. 94) foram incluídas nos testes com o propósito de verificar o comportamento das técnicas implementadas em relação à modelagem da detecção das bordas pela largura de banda (bordas largas para células complexas de V1 e em degrau para células em V2), como apresentado nas seções 4.1.4 (pág. 57) e 5 (pág. 61).

As Figuras 8.1, 8.2 e 8.3 apresentam o resultado da segmentação não-supervisionada sobre a imagem da Figura 7.1a para cada uma das técnicas abordadas nesta pesquisa.

Para a segmentação foram coletadas três mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo. Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “disco” de raio¹ três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Analisando os resultados da Figura 8.1 nota-se uma semelhança visual muito grande das técnicas propostas, Figuras 8.1(b)-(i), com a técnica existente baseada na energia, Figura 8.1a. Observa-se, ainda, que as bordas da figura são muito bem retratadas em todas as técnicas permitindo entender a imagem apenas pela observação das regiões segmentadas.

A Figura 8.2 apresenta a imagem original, com as marcações de bordas das regiões segmentadas (linha branca) obtidas de maneira não supervisionada pelo algoritmo TICA, com as diferentes técnicas implementadas, e as bordas reais da imagem.

A riqueza de detalhes torna difícil observar as diferenças entre cada abordagem, portanto, a Figura 8.3, obtida pela subtração dos resultados de cada técnica com a TICA, mostra melhor essas diferenças. Nota-se que os mo-

¹Exceto para as técnicas TICAFCM e TICAFME, ver Figuras 8.1(h)(i) e 8.1(h)(i).



Figura 8.1: Regiões segmentadas, da imagem da Figura 7.1a (pág. 94), de forma não supervisionada para dois grupamentos usando o k -means: (a) TICAЕ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (censo de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

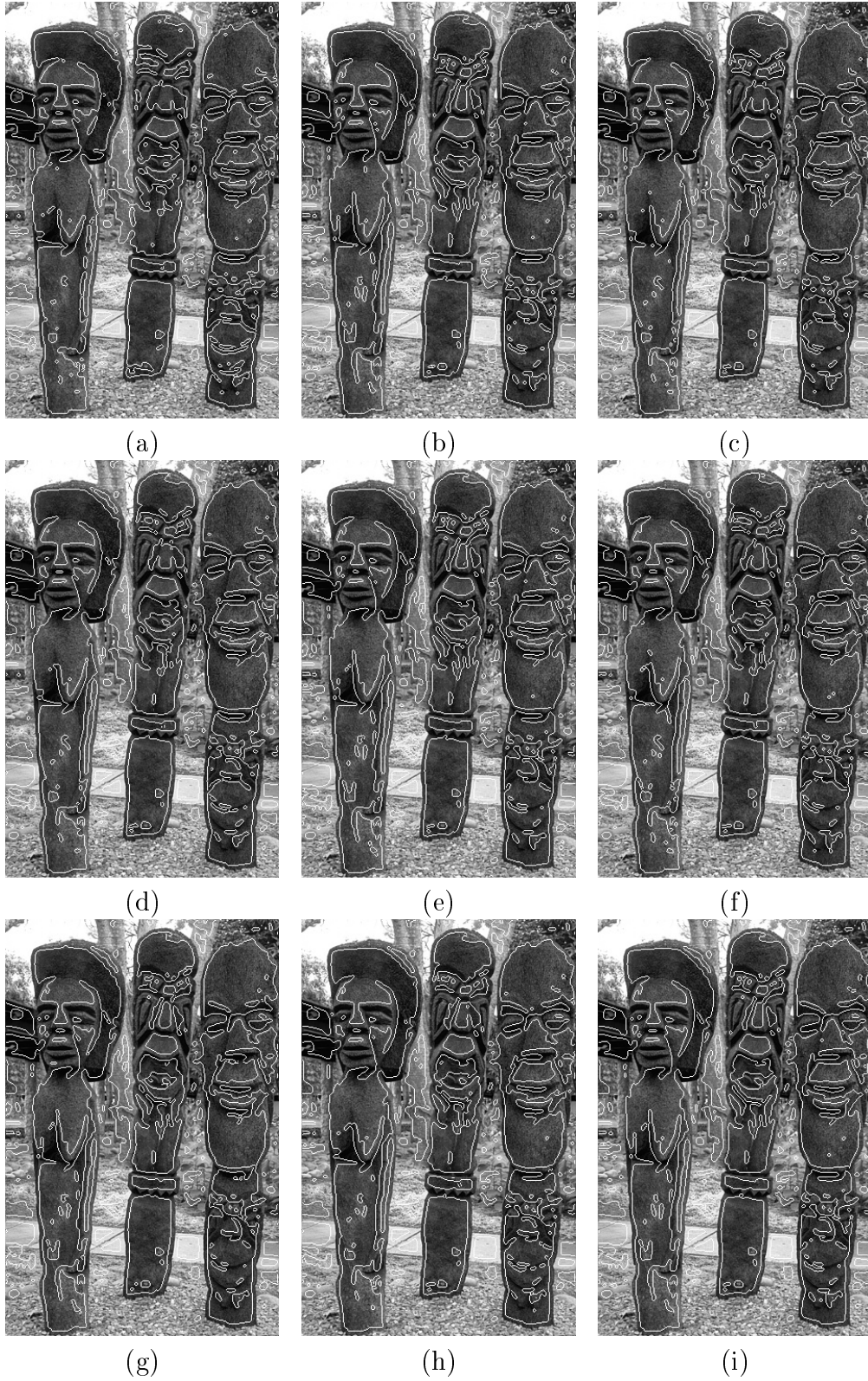


Figura 8.2: Segmentação não-supervisionada da imagem da Figura 7.1a (pág. 94) para dois grupamentos usando o *k-means*: (a) TICAЕ, (b) TICAФМ, (c) TICALF2, (d) TICALF4, (e) TICAФММ, (f) TICAФЕ, (g) TICAВЕЕ, (h) TICAФМC (cenas de tamanho 5, 3 e 1), (i) TICAФМЕ (escalas 5, 3 e 1).

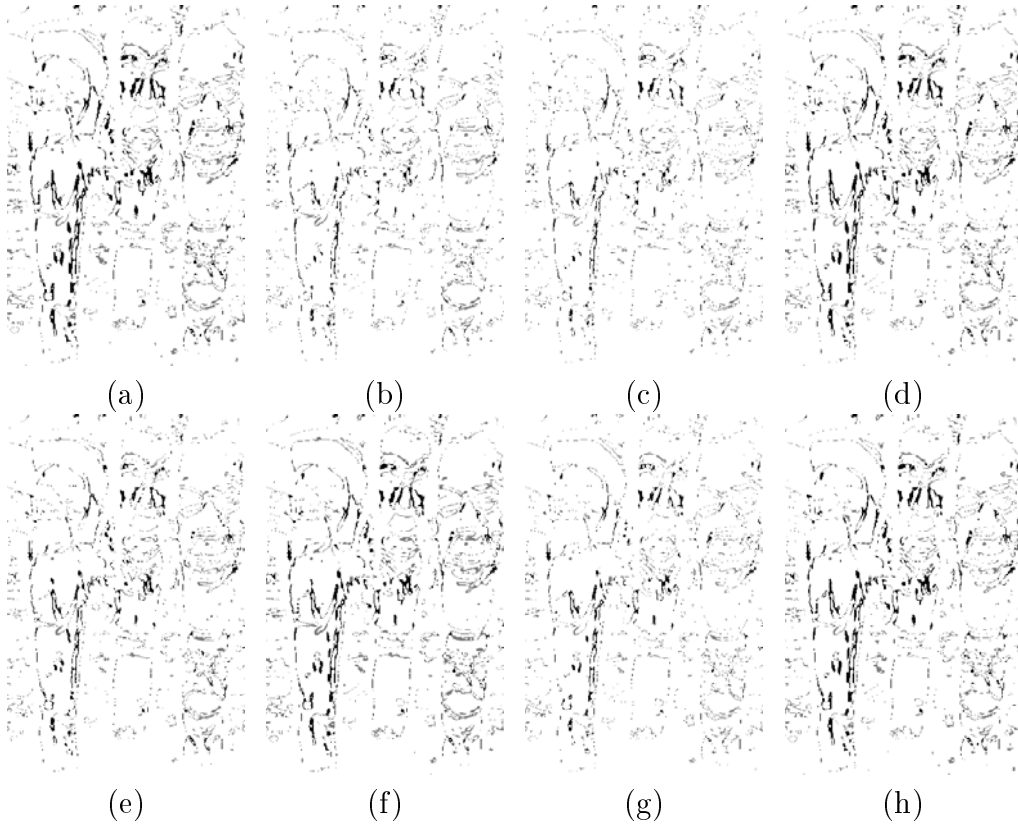


Figura 8.3: Diferenças da segmentação não-supervisionada entre a TICAE e as demais técnicas aplicadas à imagem da Figura 7.1a (pág. 94) para dois grupamentos usando o *k-means*: (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICA VEE, (g) TICA FMC (cenar de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).

delos TICALF2 e TICALF4 apresentam as menores diferenças.

O valor do fator $\lambda = 2, 2$, mostra que o critério fractal contribui efetivamente na obtenção dos filtros estatísticos do modelo.

As Figuras 8.4 a 8.9 apresentam o resultado da segmentação não-supervisionada sobre a imagem da Figura 7.1b para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas duas mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “disco” de raio² três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Nas imagens da Figura 8.4 é possível observar claramente a forma da pessoa na máscara segmentada. Novamente, as imagens são muito semelhantes, apesar das diferentes implementações do critério topográfico. Nota-se que a grade no plano de fundo da imagem é reconhecida como uma região da imagem, mostrando que o descritor fractal não é apenas um detector de bordas. A Figura 8.5 mostra a localização das bordas de segmentação em relação as bordas da imagem.

É possível observar as regiões estatisticamente semelhantes, embora o *k-means* agrupe algumas regiões distintas devido ao número de grupamentos definidos à priori no algoritmo.

O agrupamento, ou separação, é baseado na distância das variáveis envolvidas, portanto, as regiões com comportamento estatístico opostamente extremos são aquelas que definem os dois grupamentos na imagem da Figura 8.4.

A Figura 8.6 mostra as diferenças entre cada abordagem e a TICA E. Apesar das poucas diferenças observadas, cada imagem permite visualizar alguns contornos da figura original.

A explicação decorre do pequeno valor do fator λ ($\lambda = 1, 23$), o que mostra ser o critério fractal pouco expressivo na obtenção dos filtros estatístico, assim, as diferenças entre o modelo TICA Energia e os modelos TICA baseados em fractais é pequena.

Porém, isso mostra que as implementações baseadas em fractais também correspondem à modelagem do comportamento das células complexas em V1, uma vez que o resultado desses modelos se aproxima do resultado do modelo TICA Energia, inspirado no comportamento dessas células.

Com o objetivo de explorar melhor o banco de filtros, as imagens da Figura 7.1b foram segmentadas em três regiões e o resultado é apresentado

²Exceto para as técnicas TICA FMC e TICA FME, ver Figuras 8.4(h)(i) e 8.5(h)(i).

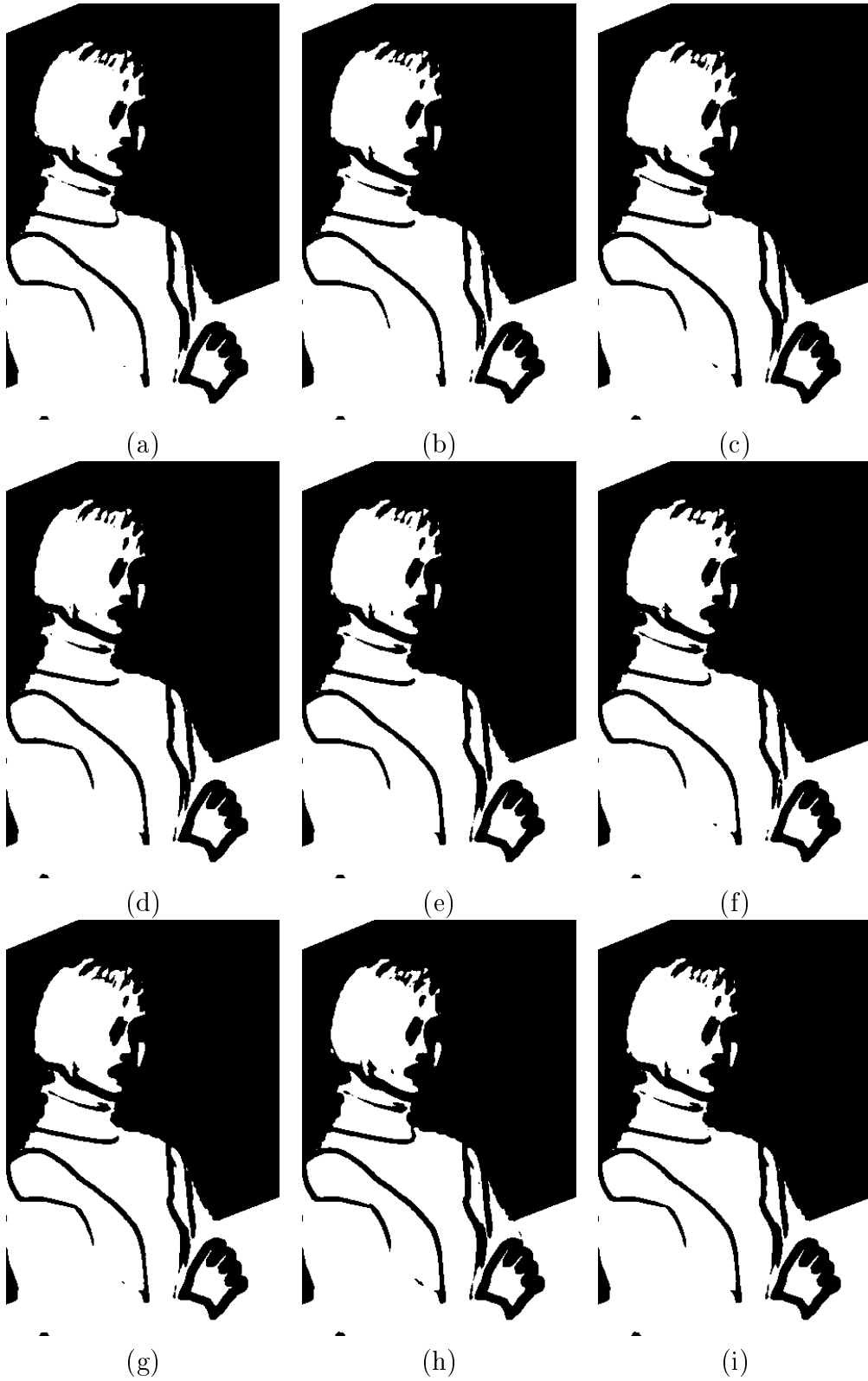


Figura 8.4: Regiões segmentadas, da imagem da Figura 7.1b (pág. 95), de forma não supervisionada para dois grupamentos usando o k -means: (a) TICAÉ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (censo de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

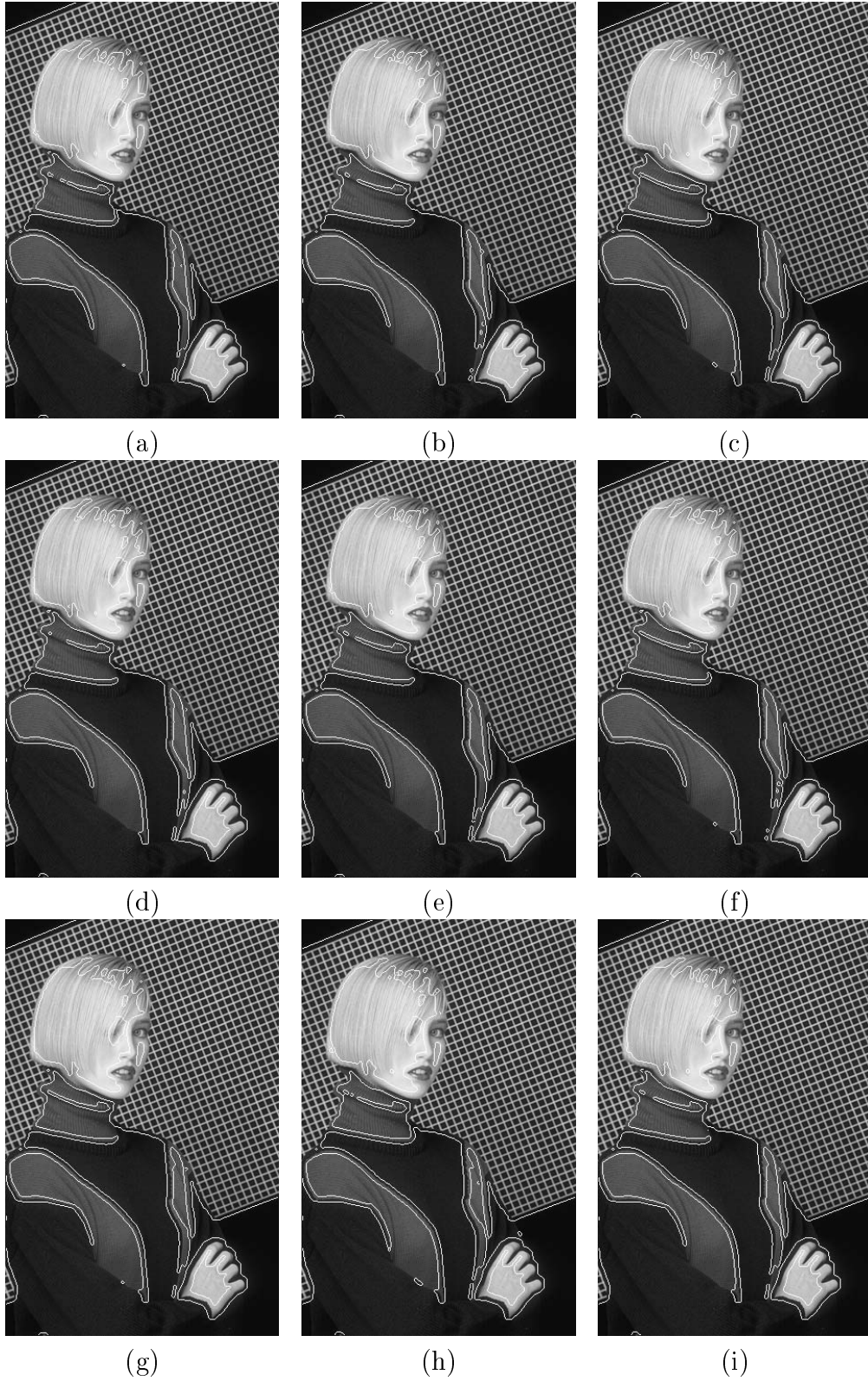


Figura 8.5: Segmentação não-supervisionada da imagem da Figura 7.1b (pág. 95) para dois grupamentos usando o *k-means*: (a) TICAЕ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (cenas de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

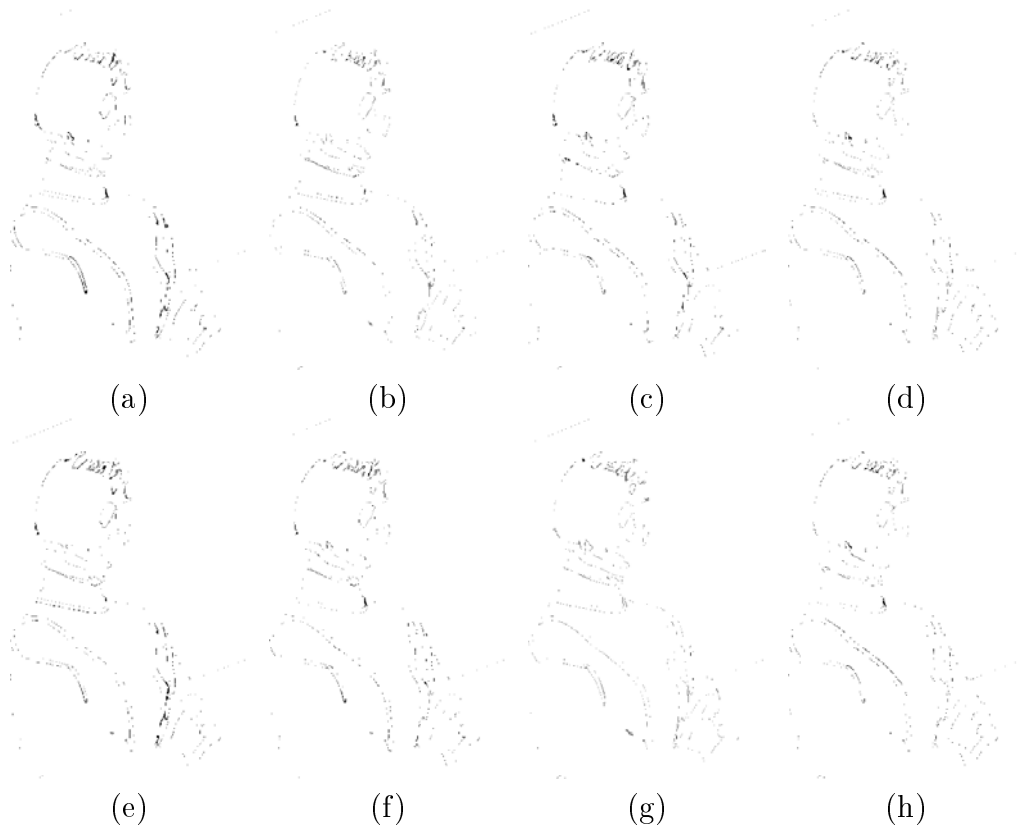


Figura 8.6: Diferenças da segmentação não-supervisionada entre a TICA E e as demais técnicas aplicadas à imagem da Figura 7.1b (pág. 95) para dois grupamentos usando o *k-means*: (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICA VEE, (g) TICA FMC (cenas de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).

nas Figuras 8.7 a 8.9.

Nota-se, contudo, um melhor detalhamento das regiões estatisticamente semelhantes, em particular nas regiões da linha do nariz, fundo quadriculado e mão direita mapeadas como a mesma região de segmentação. Todas as abordagens apresentaram pequenas diferenças em relação à TICA E, como mostra a Figura 8.9.

As Figuras 8.10 a 8.12 apresentam o resultado da segmentação não-supervisionada sobre a imagem da Figura 7.2a (pág. 95) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas três mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “disco” de raio³ três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Todas as abordagens apresentaram resultados semelhantes na segmentação não-supervisionada da Figura 7.2a, como pode ser observado na Figura 8.12.

Novamente, o fator $\lambda = 1,24$ mostra ser o critério fractal pouco expressivo na obtenção dos filtros estatístico para essa imagem, assim, as diferenças entre o modelo TICA Energia e os modelos TICA baseados em fractais é pequena.

Da mesma forma que nas Figuras 8.6 (pág. 108) e 8.9 (pág. 112), as implementações baseadas em fractais também correspondem à modelagem do comportamento das células complexas em V1, uma vez que o resultado desses modelos se aproxima do resultado do modelo TICA Energia, inspirado no comportamento dessas células.

As Figuras 8.13 a 8.15 apresentam o resultado da segmentação não-supervisionada sobre a imagem da Figura 7.2b (pág. 95) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas quatro mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “quadrado” de lado⁴ três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Observando os resultados da segmentação não-supervisionada da Figura 7.2b na Figura 8.13 é possível concluir o seguinte: os métodos TICA E e

³Exceto para as técnicas TICA FMC e TICA FME, ver Figuras 8.10(h)(i) e 8.11(h)(i).

⁴Exceto para as técnicas TICA FMC e TICA FME, ver Figuras 8.13(h)(i) e 8.14(h)(i).

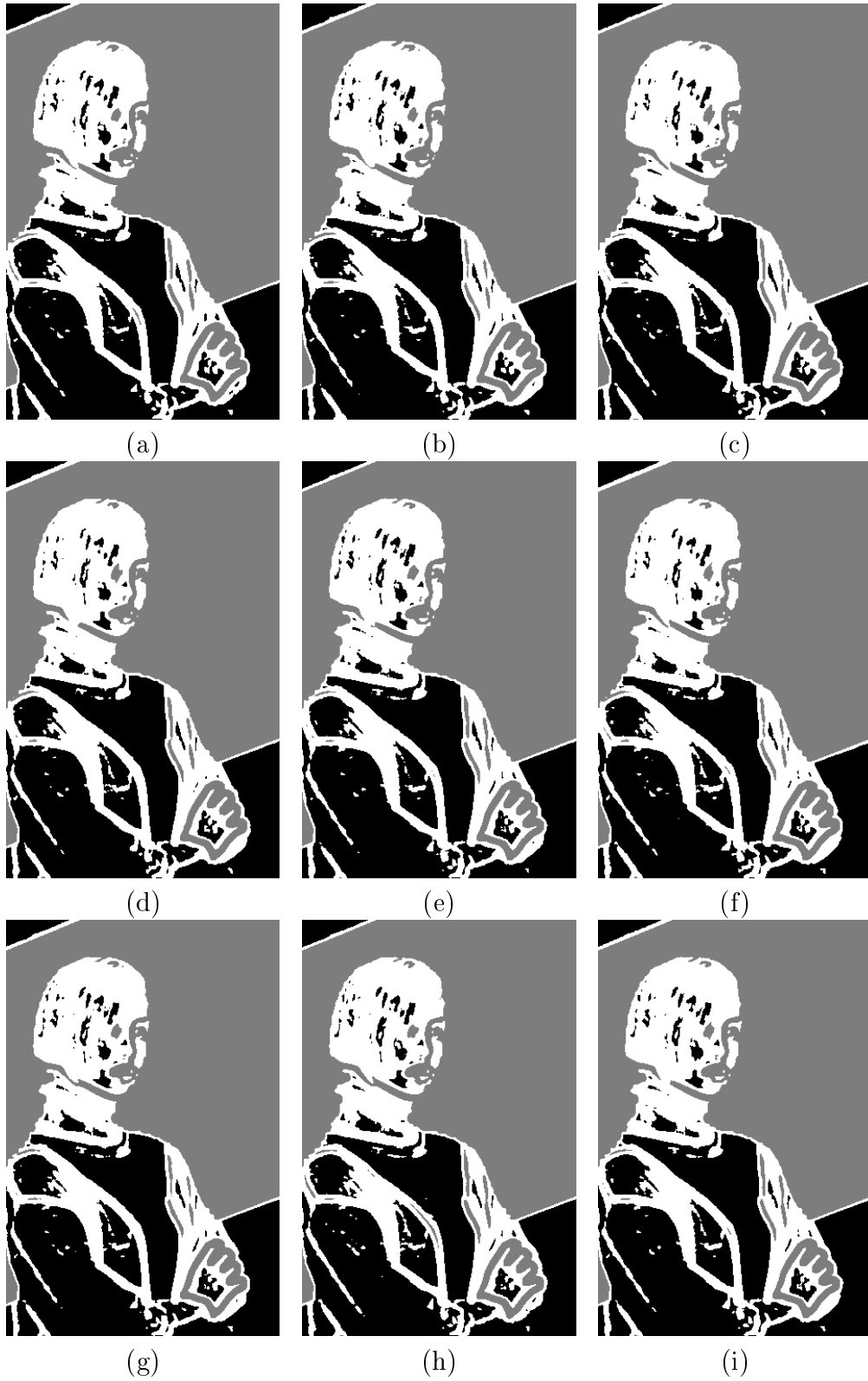


Figura 8.7: Regiões segmentadas, da imagem da Figura 7.1b (pág. 94), de forma não supervisionada para três grupamentos usando o *k-means*: (a) TICAÉ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (censo de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

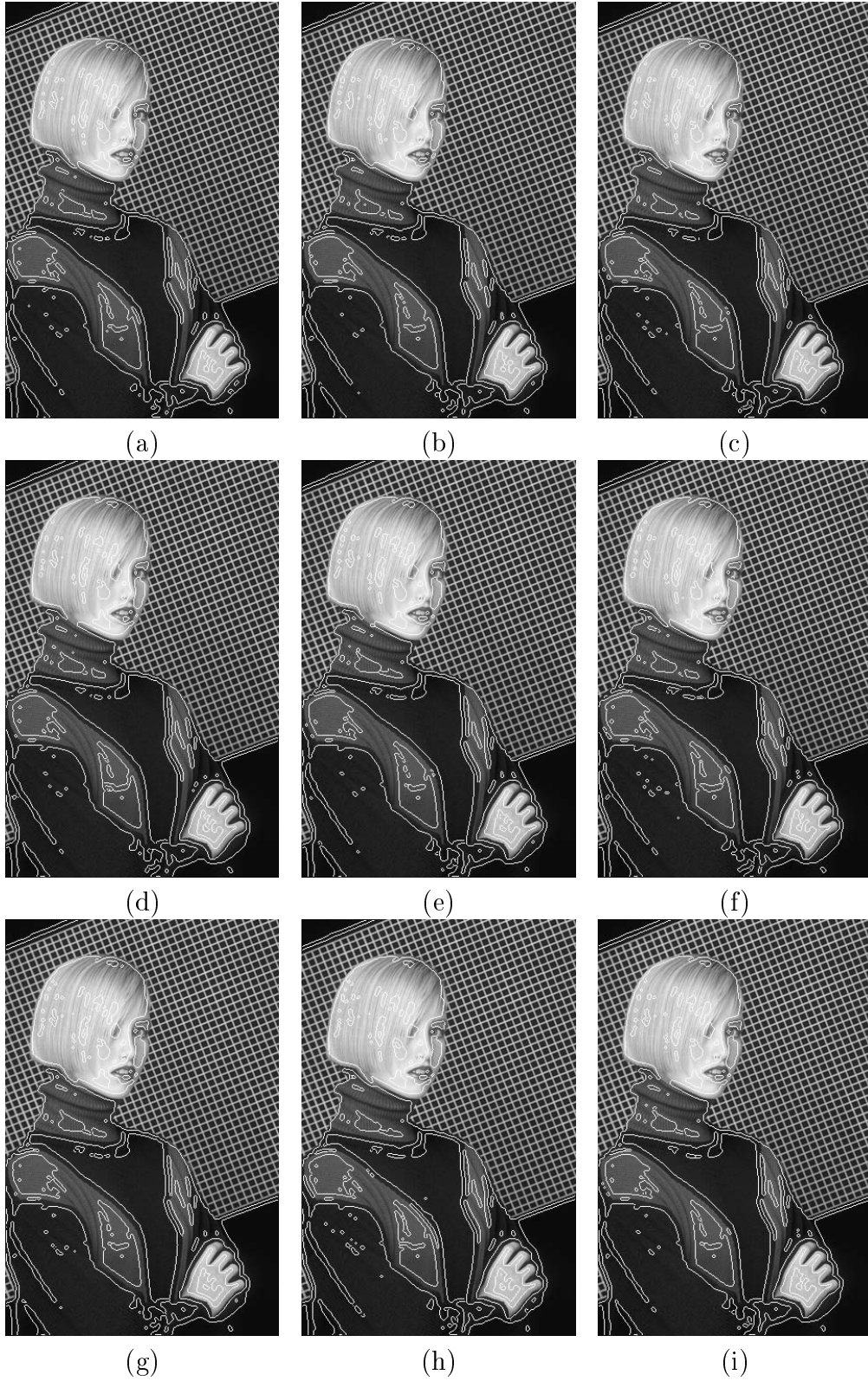


Figura 8.8: Segmentação não-supervisionada da imagem da Figura 7.1b (pág. 94) para três grupamentos usando o *k-means*: (a) TICAЕ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (escalas de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

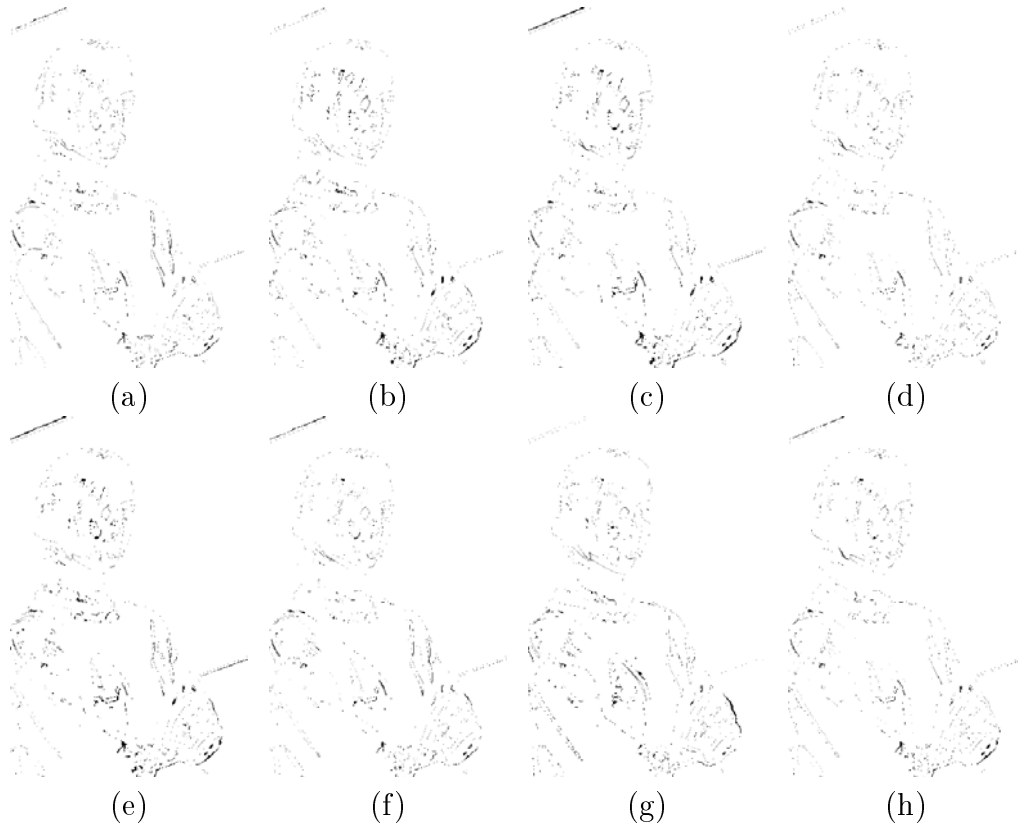


Figura 8.9: Diferenças da segmentação não-supervisionada entre a TICAe e as demais técnicas aplicadas à imagem da Figura 7.1b (pág. 94) para três grupamentos usando o *k-means*: (a) TICAfM, (b) TICALF2, (c) TICALF4, (d) TICAfMM, (e) TICAfE, (f) TICAfVEE, (g) TICAfMC (censo de tamanho 5, 3 e 1), (h) TICAfME (escalas 5, 3 e 1).

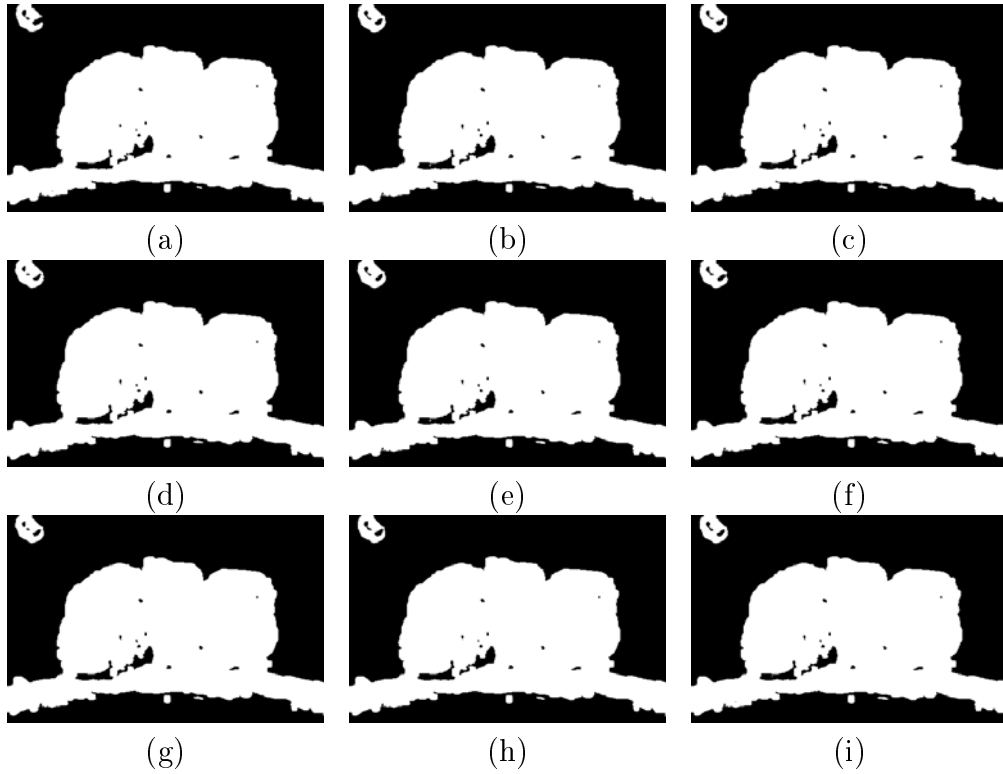


Figura 8.10: Regiões segmentadas, da imagem da Figura 7.2a (pág. 95), de forma não supervisionada para dois grupamentos usando o *k-means*: (a) TICA E, (b) TICA FM, (c) TICA LF2, (d) TICA LF4, (e) TICA FM, (f) TICA FE, (g) TICA VEE, (h) TICA FMC (cenos de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).

TICA FE apresentam resultados muito parecidos, porém a TICA FE separa melhor as regiões segmentadas, como pode ser melhor observado na areia ao pé do coqueiro. Por outro lado, os demais métodos apresentam maior seletividade das bordas da imagem, ou seja, maior realce das características estatísticas de alta-ordem.

As diferenças pronunciadas entre a TICA Energia e os métodos TICA Fractais na Figura 8.15 (pág. 117) se deve à efetiva contribuição dos modelos fractais no critério topográfico da ICA. Tal fato é notado com a observação do valor de $\lambda = 2,27$, o que mostra ter a imagem uma série de características fractais que foram percebidas pelos descritores do modelo topográfico.

As Figuras 8.16 a 8.21 apresentam o resultado da segmentação não-supervisionada sobre a imagem da Figura 7.2c para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas duas mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

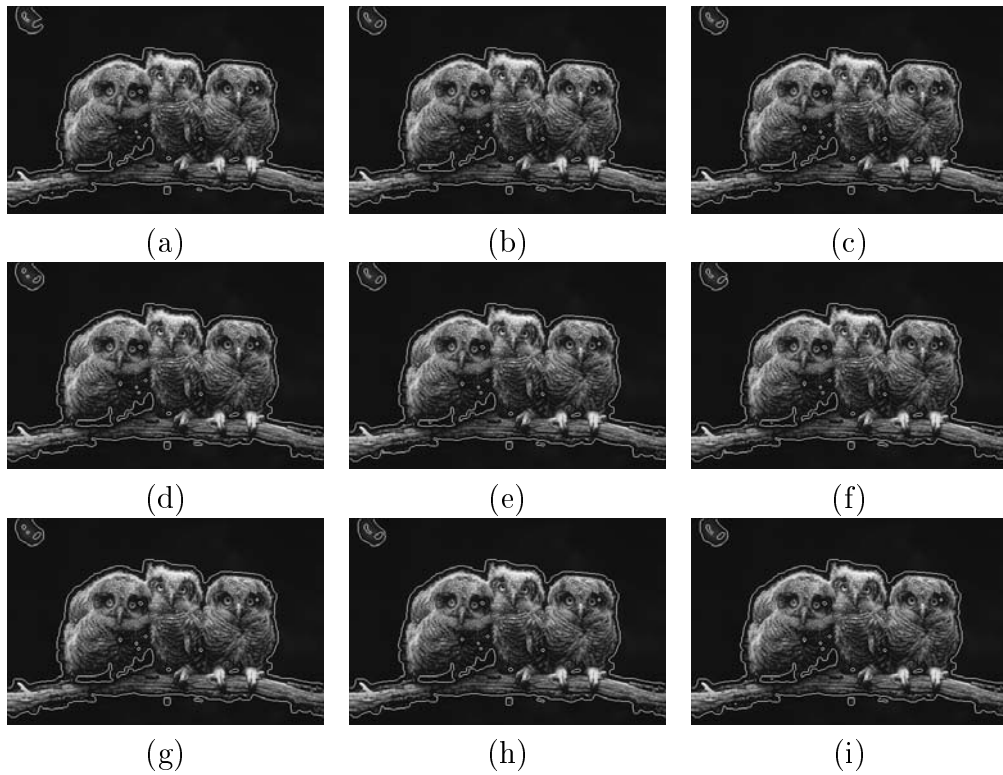


Figura 8.11: Segmentação não-supervisionada da imagem da Figura 7.2a (pág. 95) para dois grupamentos usando o *k-means*: (a) TICAЕ, (b) TICAФM, (c) TICALF2, (d) TICALF4, (e) TICAФMM, (f) TICAФE, (g) TICAVEE, (h) TICAФMC (cenais de tamanho 5, 3 e 1), (i) TICAФME (escalas 5, 3 e 1).

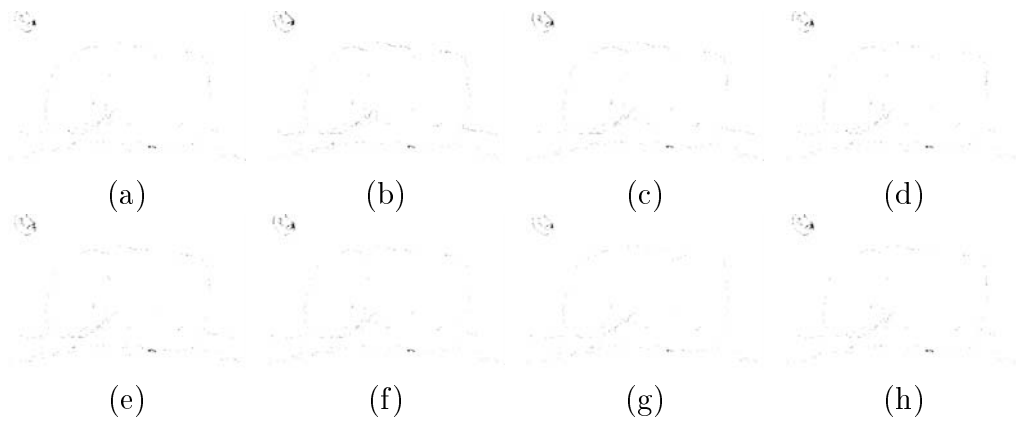


Figura 8.12: Diferenças da segmentação não-supervisionada entre a TICAЕ e as demais técnicas aplicadas à imagem da Figura 7.2a (pág. 95) para dois grupamentos usando o *k-means*: (a) TICAФM, (b) TICALF2, (c) TICALF4, (d) TICAФMM, (e) TICAФE, (f) TICAVEE, (g) TICAФMC (cenais de tamanho 5, 3 e 1), (h) TICAФME (escalas 5, 3 e 1).

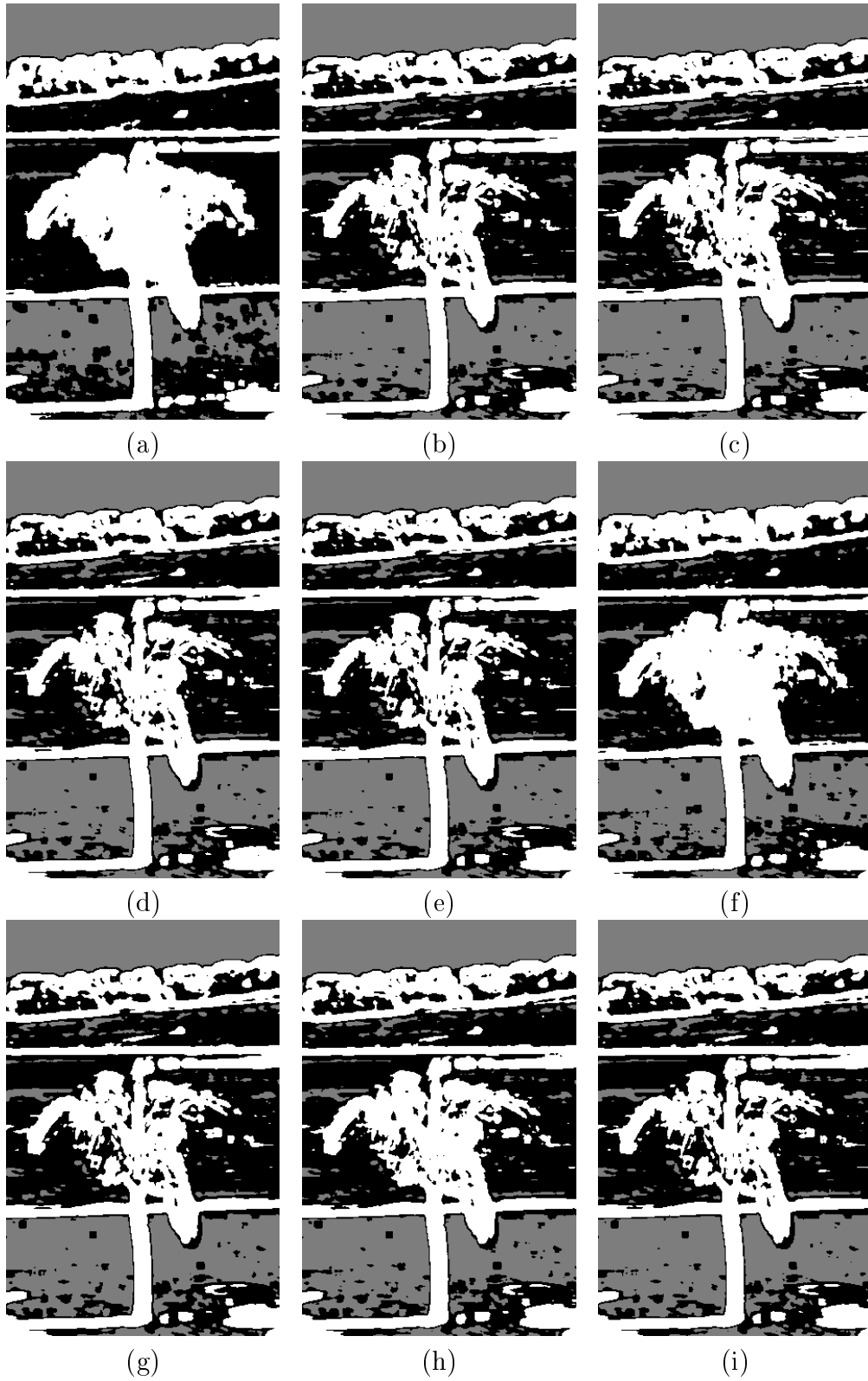


Figura 8.13: Regiões segmentadas, da imagem da Figura 7.2b (pág. 95), de forma não supervisionada para três grupamentos usando o *k-means*: (a) TICAÉ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (cenos de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

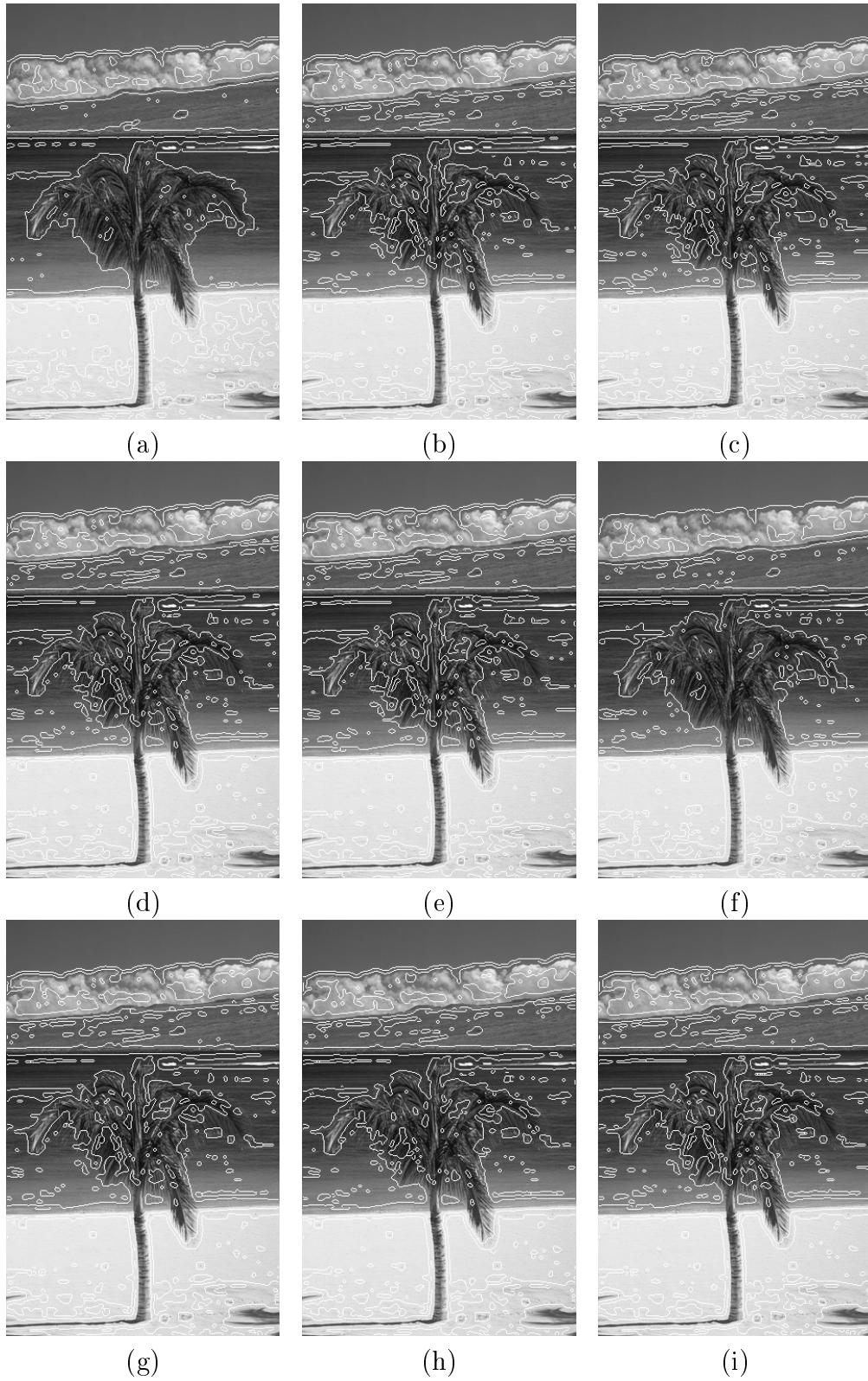


Figura 8.14: Segmentação não-supervisionada da imagem da Figura 7.2b (pág. 95) para três grupamentos usando o *k-means*: (a) TICAE, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (cenas de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

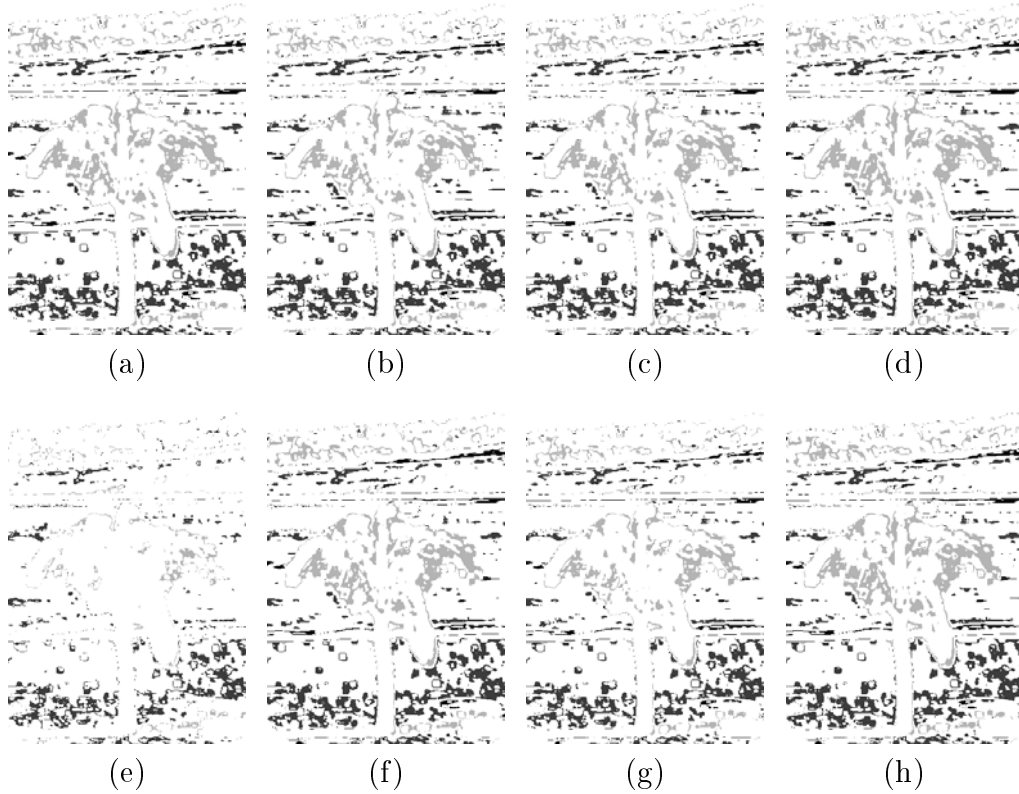


Figura 8.15: Diferenças da segmentação não-supervisionada entre a TICAE e as demais técnicas aplicadas à imagem da Figura 7.2b (pág. 95) para três grupamentos usando o *k-means*: (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICA VEE, (g) TICA FMC (cenos de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “disco” de raio⁵ três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

A segmentação não-supervisionada da imagem da Figura 7.2c (pág. 95) vista na Figura 8.16 mostra de forma bem clara as regiões estatisticamente semelhantes (bordas e patamares) descritas pela técnica de segmentação estatística usada neste trabalho.

Sabe-se que as componentes independentes da ICA são aproximações das bordas (estatísticas de alta-ordem) de imagens naturais [11], como pode ser visto no resultado da Figura 8.16.

Em todos os exemplos apresentados, os métodos baseados em fractais possuem maior sensibilidade na localização de bordas das imagens naturais, logo, contribuindo de forma melhor na obtenção das componentes estatisticamente independentes do modelo ICA.

As Figuras 8.19 e 8.20 apresentam muito bem a sensibilidade dos filtros das abordagens baseadas em fractal frente ao método baseado na energia. Os detalhes da textura, com linhas curvas nas encostas das montanhas não são destacadas na TICA E.

A diferença dos resultados entre os métodos propostos e a TICA E é muito grande (veja Figura 8.21), mostrando uma gama de informações extras que permitem melhorar a caracterização das imagens naturais na representação do modelo estatístico linear obtido pela TICA. Isso ocorre devido aos filtros estatísticos obtidos estarem bem sintonizados com as características estatísticas fractais da imagem, realçando os elementos fractais comuns da imagem, que são apresentados como regiões da segmentação.

Observa-se que as diferenças entre a TICA E e os modelos TICA Fractais é a mais pronunciada dentre todos os experimentos que estão representados nesta pesquisa. Isso se explica pela ocorrência do grande valor de λ quando comparado com os valores das demais imagens de exemplo utilizadas nesse trabalho ($\lambda = 2,58$).

O elevado valor de λ mostra que há uma quantidade grande de características fractais que são tratadas pelos modelos fractais do critério topográfico das técnicas propostas nesta pesquisa.

A percepção dos modelos baseados em fractais é tão grande que nas imagens da Figura 8.21 (pág. 122), em todas as abordagens, é possível observar nitidamente o relevo das montanhas da imagem de teste.

As Figuras 8.22 a 8.27 apresentam o resultado da segmentação não-supervisionada sobre a imagem da Figura 7.2d para cada uma das técni-

⁵Exceto para as técnicas TICA FMC e TICA FME, ver Figuras 8.16(h)(i) e 8.17(h)(i).

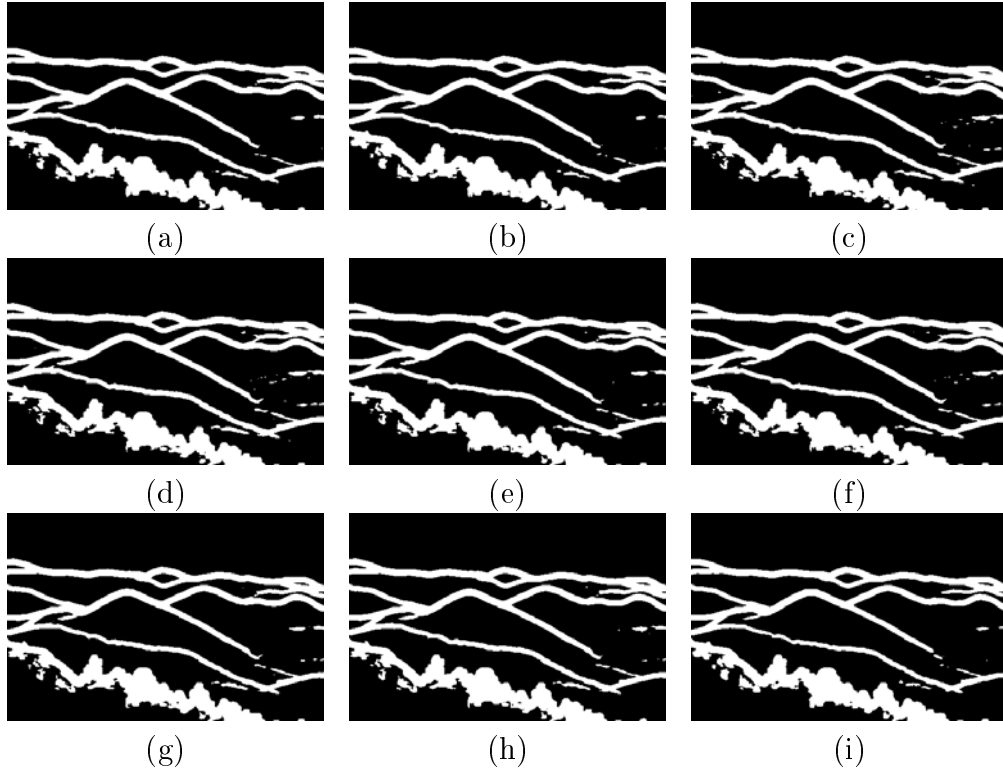


Figura 8.16: Regiões segmentadas, da imagem da Figura 7.2c (pág. 95), de forma não supervisionada para dois grupamentos usando o *k-means*: (a) TICAÉ, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICA VEE, (h) TICA FMC (cenos de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).

cas abordadas nesta pesquisa. Para a segmentação foram coletadas três mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “disco” de raio⁶ três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

As imagens segmentadas das Figuras 8.22 a 8.27 apresentam resultados coerentes com o esperado para a modelagem TICA baseada em fractais uma vez que $\lambda = 1,99$ mostra que a imagem possui características fractais suscetíveis de detecção pelos modelos fractais implementados nesta pesquisa.

A Figura 8.28 apresenta o erro da segmentação não-supervisionada sobre a imagem da Figura 7.3a (pág. 95) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas quinhentas amostras

⁶Exceto para as técnicas TICA FMC e TICA FME, ver Figuras 8.22(h)(i) e 8.23(h)(i).

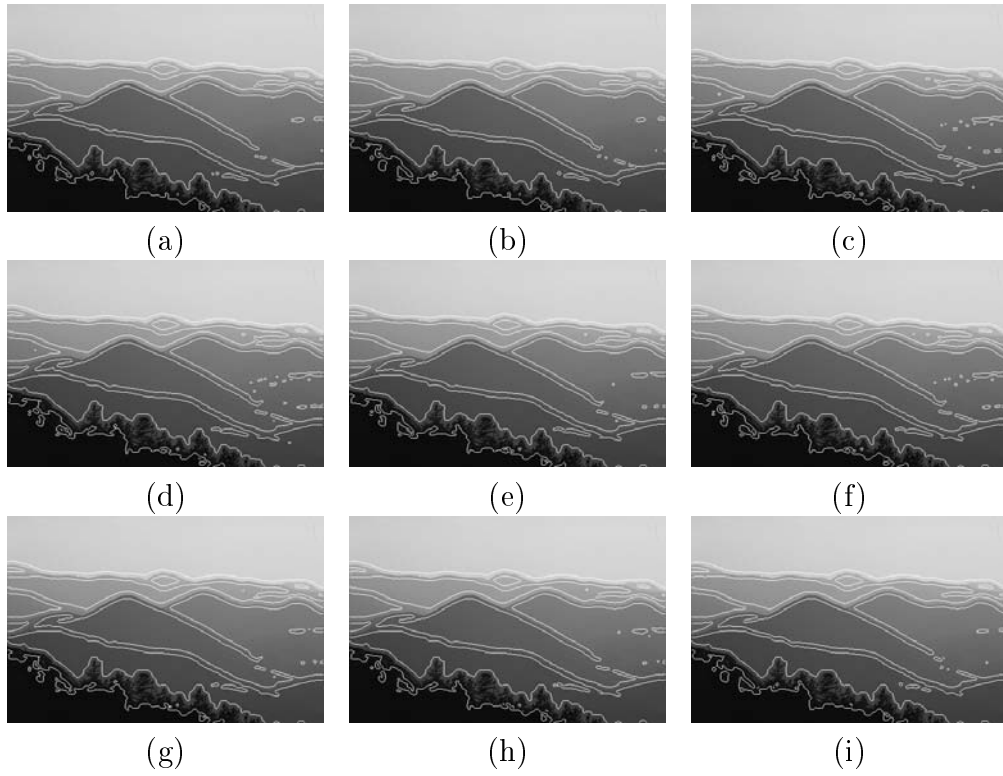


Figura 8.17: Segmentação não-supervisionada da imagem da Figura 7.2c (pág. 95) para dois grupamentos usando o *k-means*: (a) TICAЕ, (b) TICAФM, (c) TICALF2, (d) TICALF4, (e) TICAФMM, (f) TICAФE, (g) TICAФEE, (h) TICAФMC (cenais de tamanho 5, 3 e 1), (i) TICAФME (escalas 5, 3 e 1).

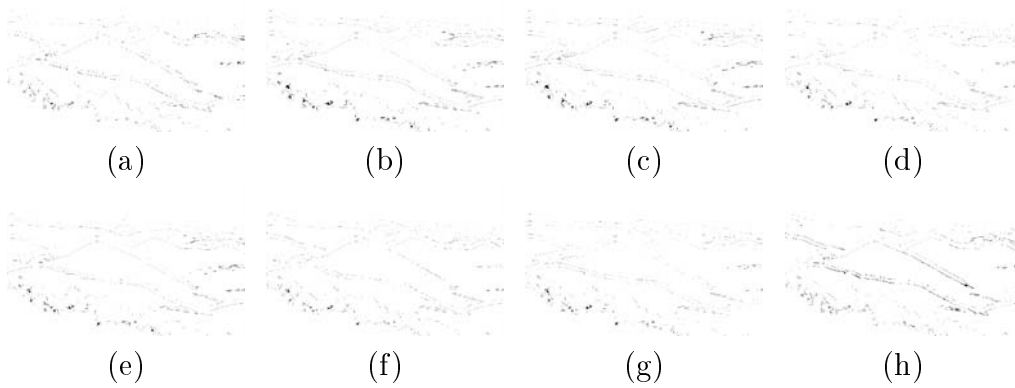


Figura 8.18: Diferenças da segmentação não-supervisionada entre a TICAЕ e as demais técnicas aplicadas à imagem da Figura 7.2c (pág. 95) para dois grupamentos usando o *k-means*: (a) TICAФM, (b) TICALF2, (c) TICALF4, (d) TICAФMM, (e) TICAФE, (f) TICAФEE, (g) TICAФMC (cenais de tamanho 5, 3 e 1), (h) TICAФME (escalas 5, 3 e 1).

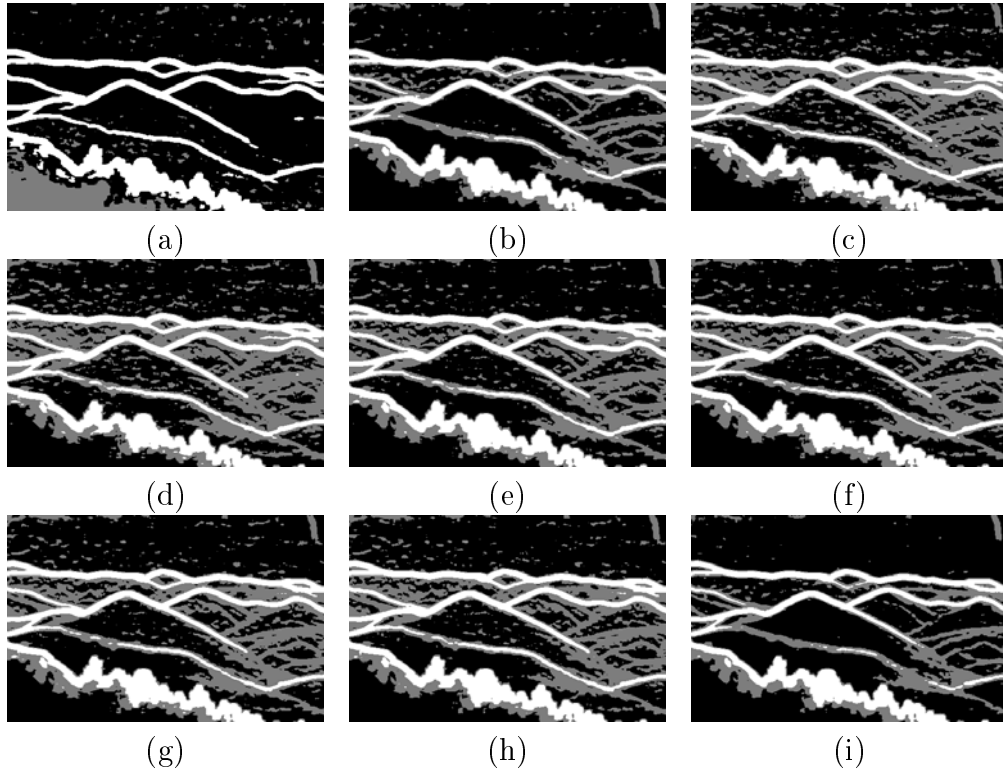


Figura 8.19: Regiões segmentadas, da imagem da Figura 7.2c (pág. 95), de forma não supervisionada para três grupamentos usando o *k-means*: (a) TICAÉ, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICAVEE, (h) TICA FMC (censo de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).

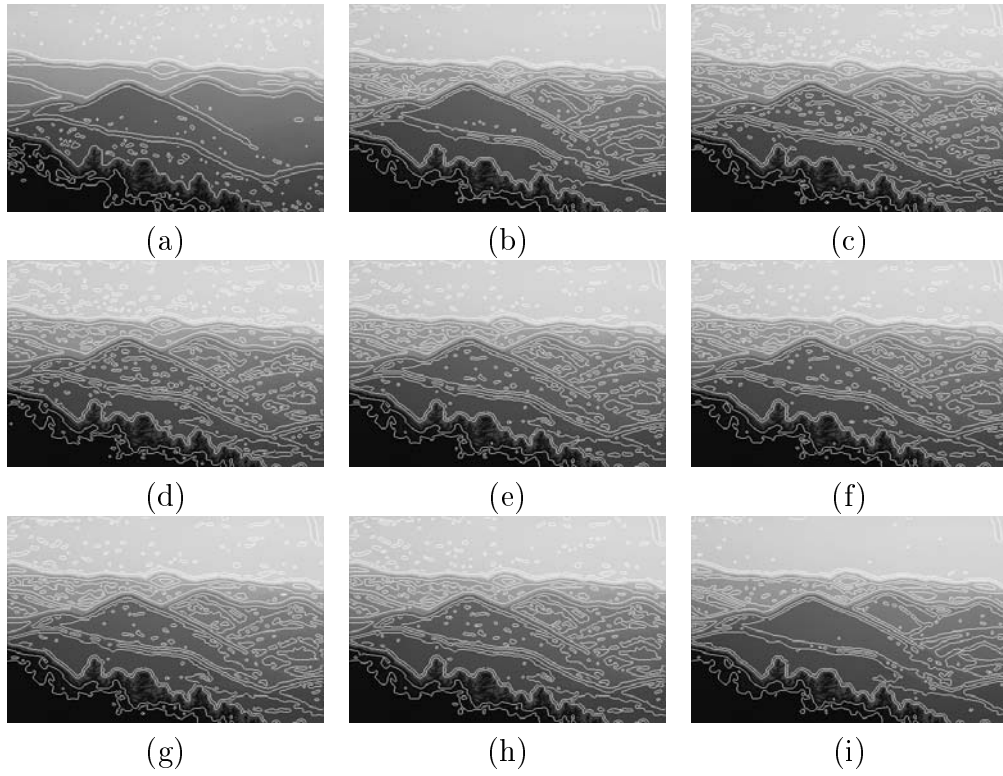


Figura 8.20: Segmentação não-supervisionada da imagem da Figura 7.2c (pág. 95) para três grupamentos usando o *k-means*: (a) TICAЕ, (b) TICAФМ, (c) TICALF2, (d) TICALF4, (e) TICAФММ, (f) TICAФЕ, (g) TICAФЕЕ, (h) TICAФМС (cenas de tamanho 5, 3 e 1), (i) TICAФМЕ (escalas 5, 3 e 1).

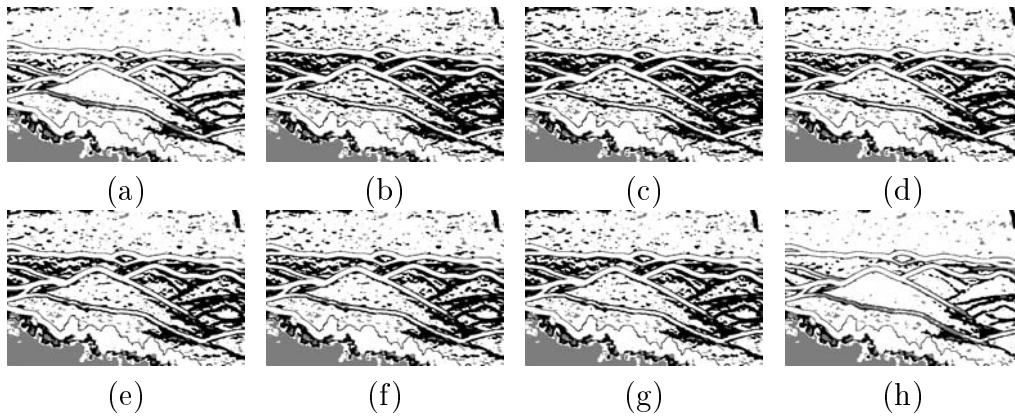


Figura 8.21: Diferenças da segmentação não-supervisionada entre a TICAЕ e as demais técnicas aplicadas à imagem da Figura 7.2c (pág. 95) para três grupamentos usando o *k-means*: (a) TICAФМ, (b) TICALF2, (c) TICALF4, (d) TICAФММ, (e) TICAФЕ, (f) TICAФЕЕ, (g) TICAФМС (cenas de tamanho 5, 3 e 1), (h) TICAФМЕ (escalas 5, 3 e 1).

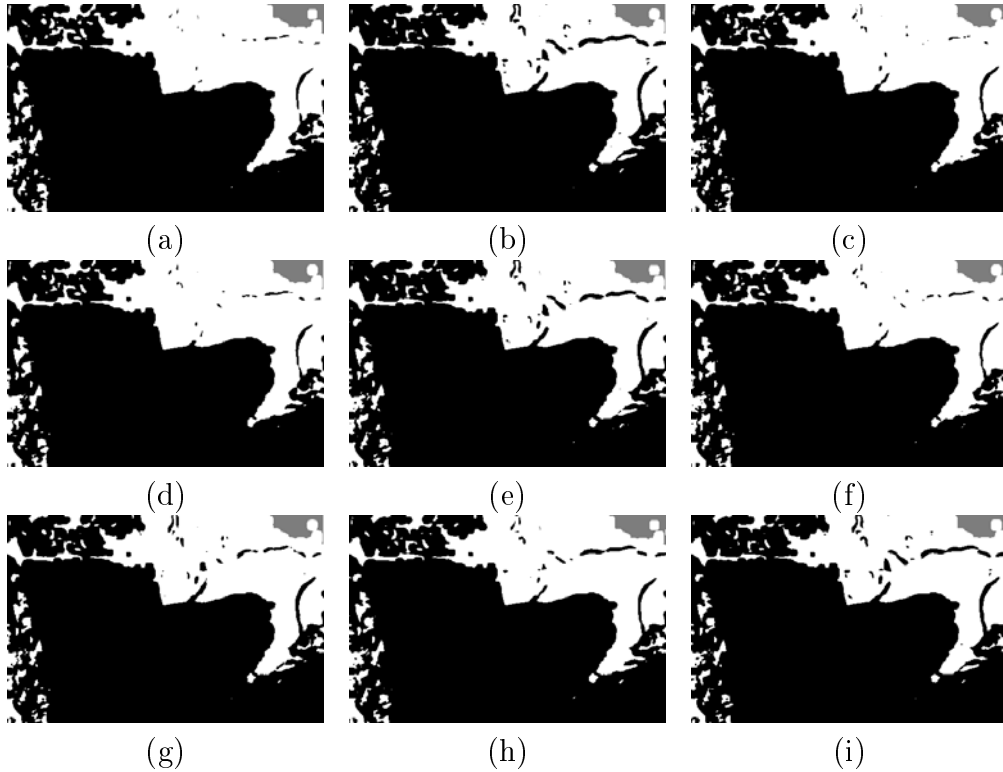


Figura 8.22: Regiões segmentadas, da imagem da Figura 7.2d (pág. 95), de forma não supervisionada para três grupamentos usando o *k-means*: (a) TICA_E, (b) TICA_{FM}, (c) TICA_{LF2}, (d) TICA_{LF4}, (e) TICA_{FM}_M, (f) TICA_{FE}, (g) TICA_{VEE}, (h) TICA_{FM}_C (cenas de tamanho 5, 3 e 1), (i) TICA_{FM}_E (escalas 5, 3 e 1).

de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “disco” de raio⁷ três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Na Figura 7.3a (pág. 95) as texturas foram montadas com duas seqüências de campos de Markov com estatísticas idênticas de baixa-ordem e diferenças na terceira e quarta ordem estatística. As técnicas TICA_{FM} e TICA_{FM}_M (Figura 8.28a) apresentaram melhores resultados de segmentação, com diferenças de até 7% em relação à TICA_E.

Como esses modelos se baseiam fortemente nos descritores fractais, o resultado mostra que a contribuição das estatísticas de alta-ordem de cada textura foram bem representadas por estes elementos.

⁷Exceto para as técnicas TICA_{FM}_C e TICA_{FM}_E, em que as escalas foram 5, 3 e 1.

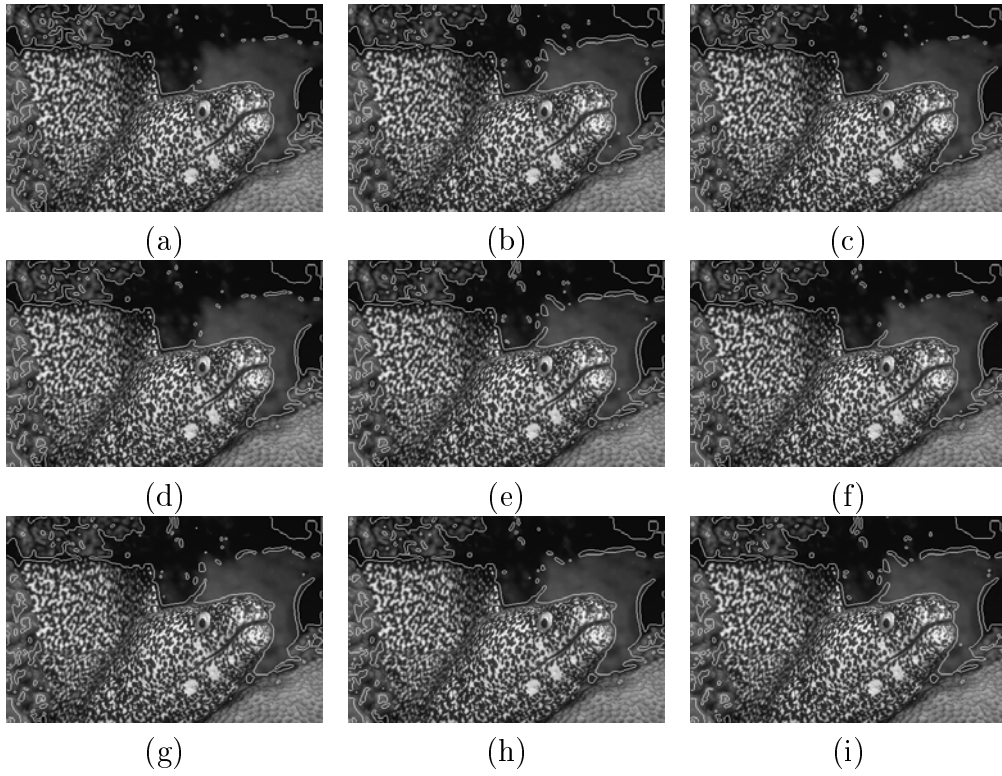


Figura 8.23: Segmentação não-supervisionada da imagem da Figura 7.2d (pág. 95) para três grupamentos usando o *k-means*: (a) TICAE, (b) TICA FM, (c) TICA LF2, (d) TICA LF4, (e) TICA FMM, (f) TICA FE, (g) TICA VEE, (h) TICA FMC (cenas de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).

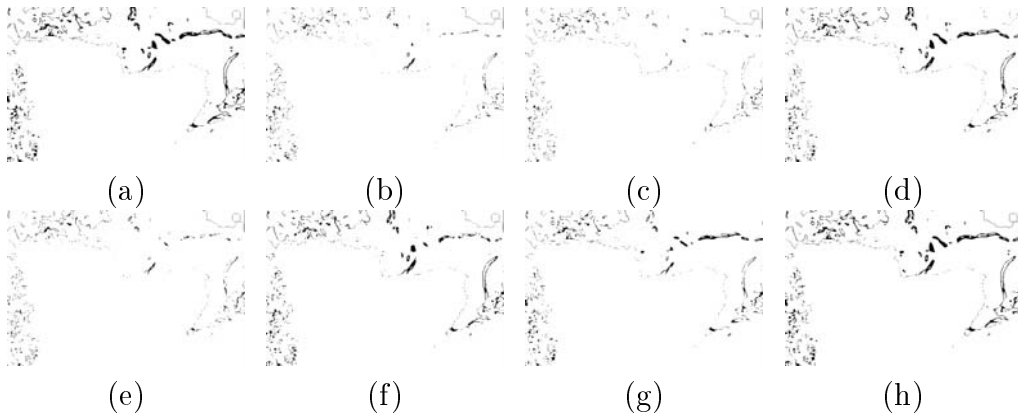


Figura 8.24: Diferenças da segmentação não-supervisionada entre a TICAE e as demais técnicas aplicadas à imagem da Figura 7.2d (pág. 95) para três grupamentos usando o *k-means*: (a) TICA FM, (b) TICA LF2, (c) TICA LF4, (d) TICA FMM, (e) TICA FE, (f) TICA VEE, (g) TICA FMC (cenas de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).

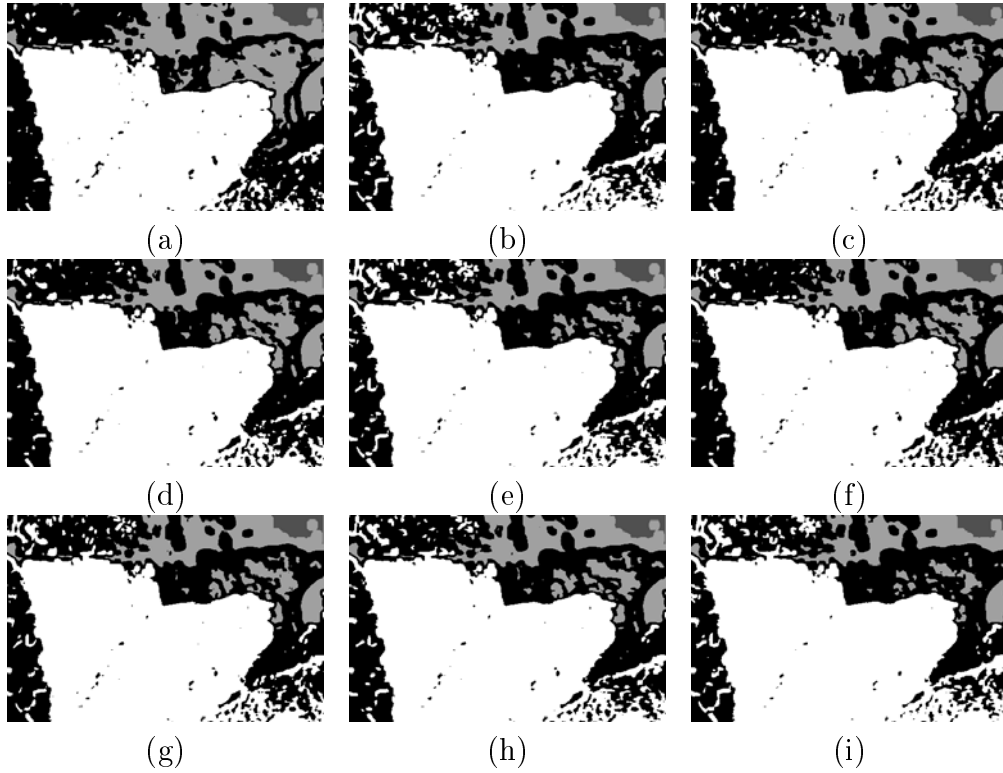


Figura 8.25: Regiões segmentadas, da imagem da Figura 7.2d (pág. 95), de forma não supervisionada para quatro grupamentos usando o k -means: (a) TICAÉ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (censo de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

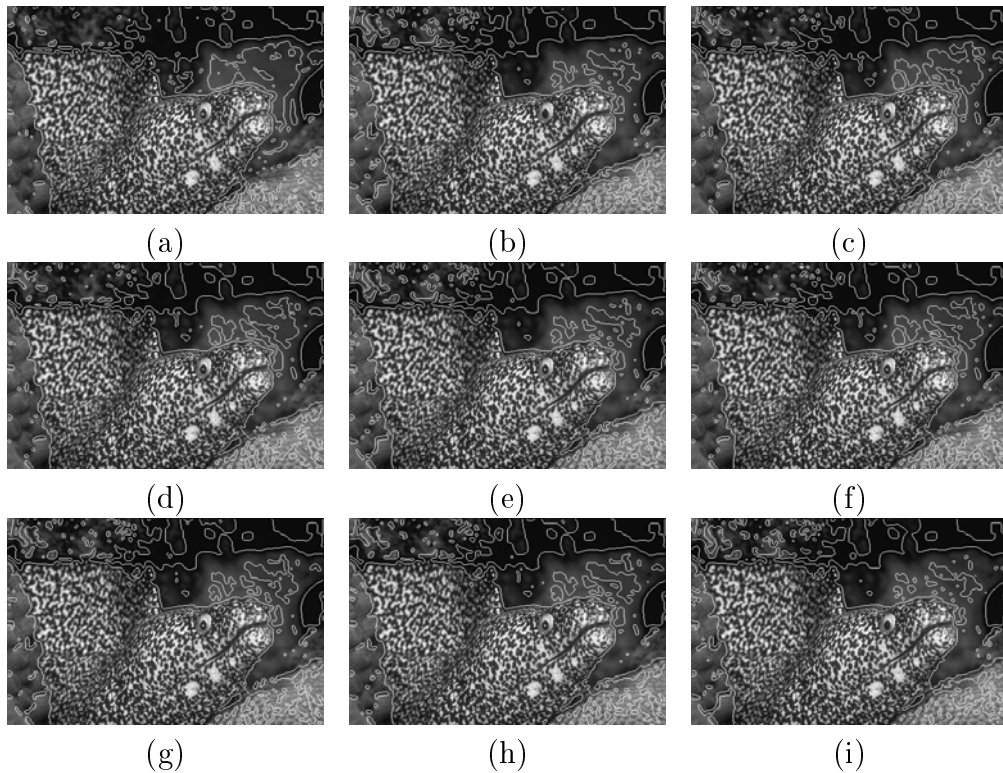


Figura 8.26: Segmentação não-supervisionada da imagem da Figura 7.2d (pág. 95) para quatro grupamentos usando o *k-means*: (a) TICAÉ, (b) TICA FM, (c) TICALF2, (d) TICALF4, (e) TICA FMM, (f) TICA FE, (g) TICA VEE, (h) TICA FMC (cenais de tamanho 5, 3 e 1), (i) TICA FME (escalas 5, 3 e 1).

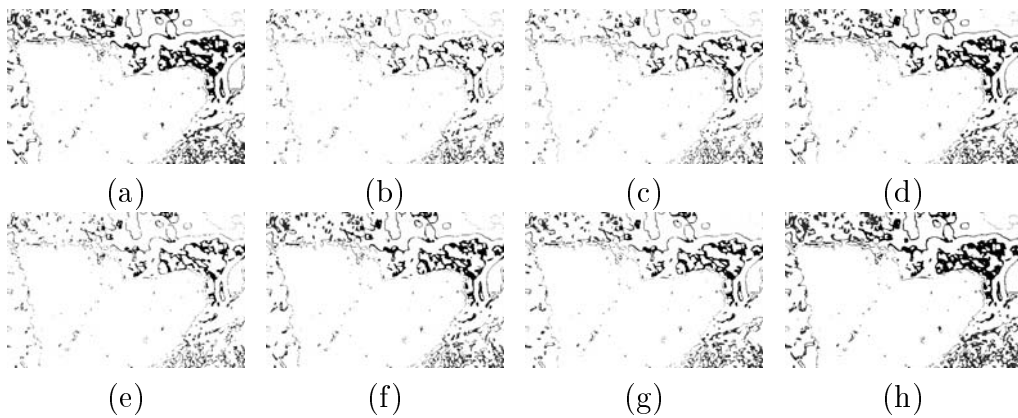


Figura 8.27: Diferenças da segmentação não-supervisionada entre a TICAÉ e as demais técnicas aplicadas à imagem da Figura 7.2d (pág. 95) para quatro grupamentos usando o *k-means*: (a) TICA FM, (b) TICALF2, (c) TICALF4, (d) TICA FMM, (e) TICA FE, (f) TICA VEE, (g) TICA FMC (cenais de tamanho 5, 3 e 1), (h) TICA FME (escalas 5, 3 e 1).

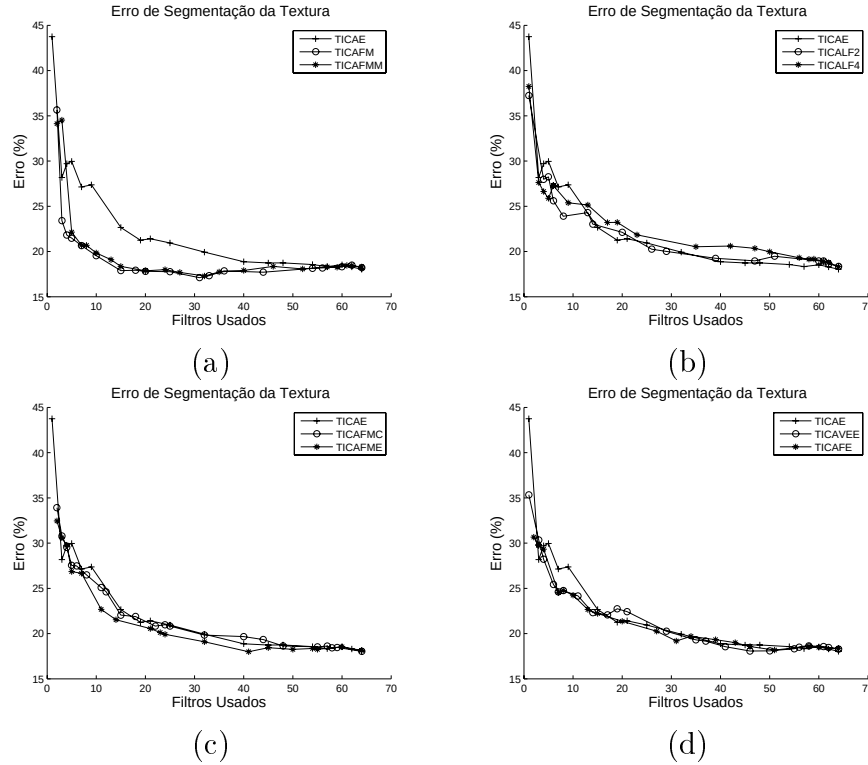


Figura 8.28: Erro da segmentação não-supervisionada entre a TICA e as demais técnicas aplicadas à imagem da Figura 7.3a (pág. 95). (a) TICA, TICA e TICA, (b) TICA, TICALF2 e TICALF4, (c) TICA, TICA e TICA, (d) TICA, TICA e TICA.

Já as abordagens baseadas na lacunaridade (Figura 8.28b) mantiveram o erro de segmentação próximo à TICA. Isso ocorre devido às texturas não apresentarem lacunas na representação estatística.

As abordagens multi-escala e multi-cena (Figura 8.28c), mantiveram o erro próximo ao obtido pela TICA, porém, a TICA ficou um pouco abaixo do gráfico do erro da TICA em praticamente toda a extensão de filtros usados. Neste caso, a abordagem multi-escalar parece ter capturado melhor a organização markoviana dos campos nas duas texturas em detrimento da abordagem multi-cenas.

Nestes dois métodos o descritor fractal é obtido da variação do tamanho da cena na vizinhança (TICA) e dos escalamentos de uma região da vizinhança (TICA), capturando a relação dinâmica de formação da textura através dos descritores em cada cena ou escala da imagem, porém representados em um único fator numérico que representa a região dentro da vizinhança.

Nas abordagens TICA e TICA (Figura 8.28d), os resultados se-

guiram a TICAÉ, como era de se esperar, pois, o comportamento do CVEE é parecido com o da energia e o modelo TICAFE usa uma heurística baseada em fractais usando a energia das regiões da vizinhança, ou seja, é o TICAÉ mais robusto à localização de bordas (que são poucas na textura analisada).

Observe que na textura em questão, de maneira geral, todas as técnicas tiveram um erro de segmentação acima de 15%, cabe lembrar que tal diferença se deve basicamente ao algoritmo ICA na determinação das componentes estatisticamente independentes do modelo. A TICA é uma extensão da ICA, logo, onde a ICA não apresenta bom resultado, a TICA não apresentará resultado muito melhor, no máximo, um pouco melhor.

A Figura 8.29 apresenta o erro da segmentação não-supervisionada sobre a imagem da Figura 7.3b (pág. 95) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas quinhentas amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “disco” de raio⁸ três resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Na Figura 7.3b (pág. 95) as texturas foram montadas sem campos de Markov e com estatísticas idênticas de baixa-ordem e diferenças apenas nos terceiro e quarto momentos estatísticos.

As técnicas TICAFM e TICAFMM (Figura 8.29a), TICALF4 (Figura 8.29b) e TICAFE (Figura 8.29d) apresentaram melhores resultados de segmentação em relação à TICAÉ. Nestes casos os modelos estatísticos de alta-ordem no critério topográfico contribuíram para ajustar melhor a independência das componentes estatisticamente independentes do modelo.

Por outro lado, os modelos TICALF2 (Figura 8.29b), TICAFMC, TICAFME (Figura 8.29c) e TICAVEE (Figura 8.29d) mantiveram o resultado em torno da TICAÉ. Isso porque os métodos TICALF2 e TICAVEE se baseiam em estatísticas de segunda ordem e os métodos TICAFMC e TICAFME se baseiam na variação da textura ao longo do espaço.

Como as texturas foram sinteticamente construídas com estatísticas de alta-ordem, sem correlação nas frequências espaciais, os modelos baseados em estatísticas de baixa-ordem não obtiveram melhores resultados que a TICAÉ e os modelos multi-escala e multi-cena tiveram maximizadas as representações dos descritores na maior escala/cena, aproximando do comportamento característico das células complexas em V1 com larga banda de borda, como na TICAÉ.

⁸Exceto para as técnicas TICAFMC e TICAFME, em que as escalas foram 5, 3 e 1.

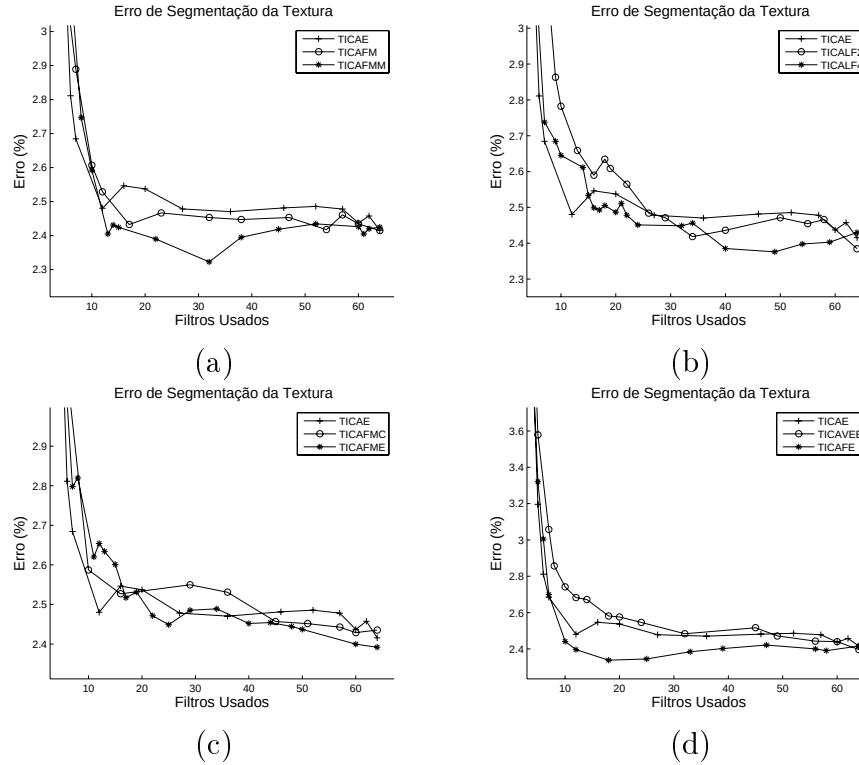


Figura 8.29: Erro da segmentação não-supervisionada entre a TICAe e as demais técnicas aplicadas à imagem da Figura 7.3b (pág. 95). (a) TICAe, TICAfM e TICAfMM, (b) TICAe, TICALF2 e TICALF4, (c) TICAe, TICAfMC e TICAfME, (d) TICAe, TICAVEE e TICAfE.

Cabe observar ainda que as texturas deste mosaico não possuem características fractais expressivas, como pode ser visto no espectro $1/f$ da Figura 7.5 (pág. 97). Logo, o que prevalece nos modelos implementados são as características do comportamento das células complexas de V1, ou seja, o mecanismo de interação lateral reduzido na organização das células simples.

A Figura 8.30 apresenta o erro da segmentação não-supervisionada sobre a imagem da Figura 7.4a (pág. 96) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas duas mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “quadrado” de lado⁹ três, resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

⁹Exceto para as técnicas TICAfMC e TICAfME, em que as escalas foram 5, 3 e 1.

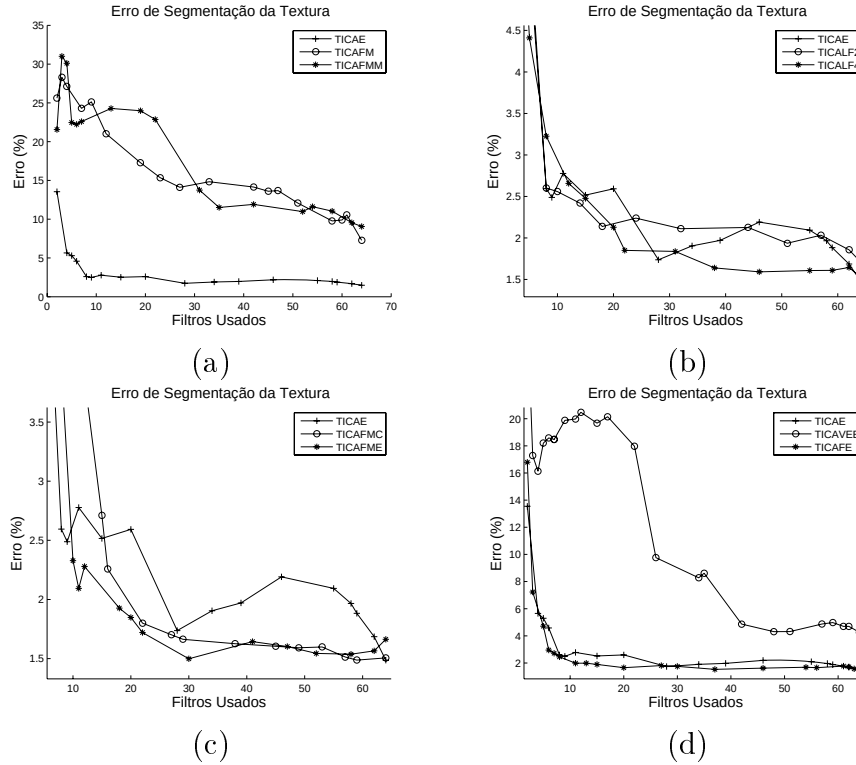


Figura 8.30: Erro da segmentação não-supervisionada entre a TICA e as demais técnicas aplicadas à imagem da Figura 7.4a (pág. 96). (a) TICA, TICAFM e TICAFMM, (b) TICA, TICALF2 e TICALF4, (c) TICA, TICAFMC e TICAFME, (d) TICA, TICAVEE e TICAFE.

Nas texturas de Brodatz da Figura 7.4a (pág. 96) as técnicas TICAFM, TICAFMM e TICAVEE (Figuras 8.30a e (d)) não obtiveram resultado satisfatório para a segmentação não-supervisionada na quantidade de amostras usadas para composição da mistura do modelo TICA.

As duas texturas parecem semelhantes, mas em diferentes escalas, e cada abordagem buscou extrair informações desconsiderando essas diferenças. Nos métodos TICAFMC e TICAFME (Figura 8.30c) essas diferenças foram detectadas e contribuíram para a classificação das duas texturas, proporcionando resultados melhores que os obtidos pela TICA.

As medidas baseadas na lacunaridade (Figura 8.30b) encontraram regras relacionadas, provavelmente, no tamanho das lacunas estatísticas de cada textura, que permitiram separá-las. O método TICAFE (Figura 8.30d), como dito anteriormente, é o método TICA com uma melhor definição de bordas, logo, o resultado deste modelo segue o resultado da TICA neste mosaico.

A explicação para esse comportamento entre a TICAFE e TICA nestas texturas pode ser constatado no espectro $1/f$ da Figura 7.7 (pág. 99), que

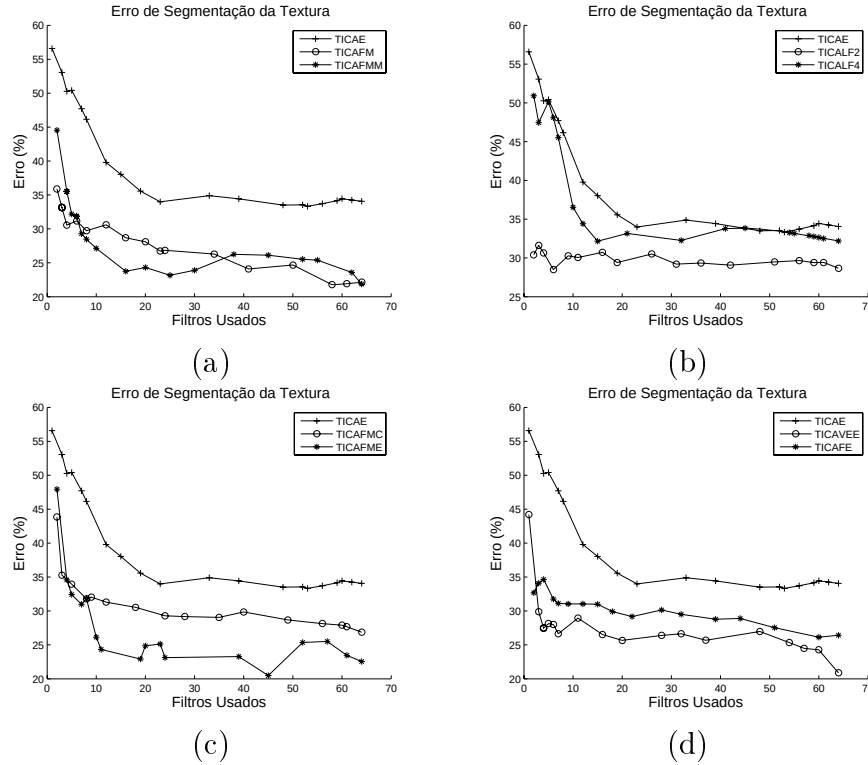


Figura 8.31: Erro da segmentação não-supervisionada entre a TICA e as demais técnicas aplicadas à imagem da Figura 7.4b (pág. 96). (a) TICA, TICA FM e TICA FMM, (b) TICA, TICA LF2 e TICA LF4, (c) TICA, TICA FMC e TICA FME, (d) TICA, TICA VEE e TICA FE.

mostra uma certa suavidade das imagens do ponto de vista fractal, pois $\lambda \approx 1$ nas duas texturas do mosaico, ou seja, a superfície aparenta ser bem homogênea [95].

A Figura 8.31 apresenta o erro da segmentação não-supervisionada sobre a imagem da Figura 7.4b (pág. 96) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “quadrado” de lado¹⁰ três, resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Para análise dos gráficos da Figura 8.31 a Tabela 8.1 mostra o resultado aproximado do erro da segmentação não-supervisionada no melhor caso para cada técnica. Pode-se observar uma disparidade no erro de segmentação de

¹⁰Exceto para as técnicas TICA FMC e TICA FME, em que as escalas foram 5, 3 e 1.

Método Utilizado	Erro Aproximado
TICAE	35%
TICAFM	22%
TICAFMM	22%
TICALF2	30%
TICALF4	34%
TICAFMC	28%
TICAFME	20%
TICAFE	27%
TICAVEE	21%

Tabela 8.1: Melhor caso aproximado para a segmentação não-supervisionada da Figura 7.3a.

até 15% entre as técnicas TICAE e TICAFME (Figura 8.31c).

Em todas as abordagens propostas o erro de segmentação foi menor que o erro apresentado pela TICAE. Este resultado mostra que o critério topográfico, inspirado no comportamento das células em V2, exerce um papel importante na obtenção das funções de base do modelo ICA nessas imagens.

Novamente aqui, o erro de segmentação é fortemente influenciado pelo algoritmo ICA, ou seja, pela forma com que a ICA obtém suas componentes. Todavia, mesmo com o alto erro de segmentação, o critério topográfico permite reduzir bastante esse erro em algumas das abordagens implementadas, como pode ser visto nos gráficos da Figura 8.31.

A Figura 8.32 apresenta o erro da segmentação não-supervisionada sobre a imagem da Figura 7.4c (pág. 96) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas duas mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “quadrado” de lado¹¹ três, resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Neste teste, a maioria das técnicas implementadas nesta pesquisa tiveram erros superiores ao proporcionado pela TICAE, porém com diferenças sempre inferiores a 2%. Exceção ocorreu com a TICAFME, perfazendo um resultado melhor que a TICAE com diferença de até 1%.

Este resultado mostra a dependência da ICA às amostras de análise da mistura, ou seja, cada imagem possui peculiaridades que podem ser melhor

¹¹Exceto para as técnicas TICAFMC e TICAFME, em que as escalas foram 5, 3 e 1.

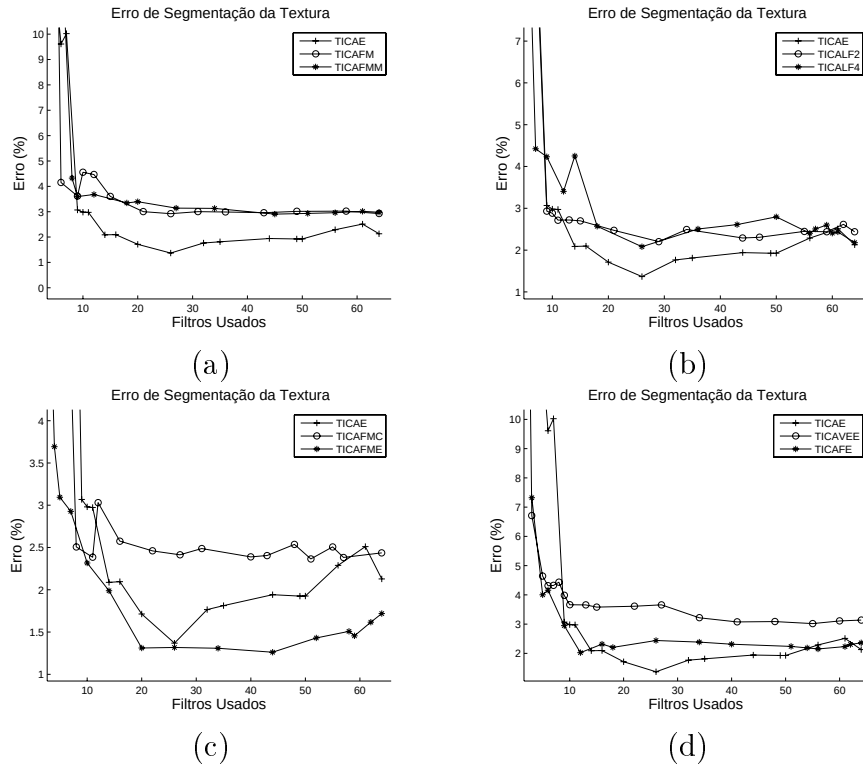


Figura 8.32: Diferenças da segmentação não-supervisionada entre a TICAE e as demais técnicas aplicadas à imagem da Figura 7.4c (pág. 96). (a) TICAE, TICAFM e TICAFMM, (b) TICAE, TICALF2 e TICALF4, (c) TICAE, TICAFMC e TICAFME, (d) TICAE, TICAVEE e TICAFE.

descritas quando analisadas de uma maneira que de outra.

Dependendo da forma como são medidas as características da mistura, mais fácil ou mais difícil é a obtenção das componentes estatisticamente independentes da imagem. No caso em questão, a abordagem multi-escala no critério topográfico extrai melhor as características que permitem segmentar as duas regiões da imagem.

A Figura 8.33 apresenta o erro da segmentação não-supervisionada sobre a imagem da Figura 7.4d (pág. 96) para cada uma das técnicas abordadas nesta pesquisa. Para a segmentação foram coletadas quatro mil amostras de tamanho 8×8 , aleatoriamente selecionadas na imagem, para compor a mistura do modelo.

Cada algoritmo TICA é aplicado à mistura utilizando um elemento morfológico estruturante “quadrado” de lado¹² três, resultando em um banco com sessenta e quatro filtros para cada modelo. Cada banco de filtros é, então, utilizado para a segmentação não-supervisionada da imagem de teste.

Na imagem da Figura 7.4d (pág. 96) todas as técnicas propostas apresentam resultado de segmentação melhor que o resultado obtido da técnica existente baseada na energia (TICAE), mostrando que a modelagem fractal do critério topográfico pode melhorar a obtenção das componentes estatisticamente independentes do modelo ICA, com a melhor sintonia dos filtros resultantes da representação da imagem pela matriz de mistura.

De maneira geral, os resultados dos testes permitem averiguar algumas informações. Na seção 3.2.3 (pág. 41), o modelo apresentado por Balboa e Grzywacz [6] mostra que variações no tamanho dos campos receptivos são ótimas para representação da intensidade¹³ das bordas em imagens.

Os modelos TICAFMC e TICAFME procuram modelar essa característica, proporcionando um diferencial na extração de informações dos campos receptivos (células simples) e na organização dos mesmos.

Observe ainda que, inspirado nos modelos biológicos, imagens simples são bem resolvidas pelas células que antecedem o córtex V1, como o LGN e RGC, que efetuam o branqueamento da imagem. Logo, informações que podem ser resolvidas por estatísticas de segunda ordem, através do PCA, não são processadas no córtex V1, o que explica por que imagens simples (baixa-ordem estatística) não são resolvidas adequadamente pela ICA (inspirada no comportamento das células simples em V1).

Cabe ressaltar que a sensibilidade da segmentação não-supervisionada depende do número de agrupamentos, definido como parâmetro do *k-means*.

¹²Exceto para as técnicas TICAFMC e TICAFME, em que as escalas foram 5, 3 e 1.

¹³A intensidade aqui está relacionada a diferença entre os níveis de cinza do interior da borda e da superfície próxima à borda em uma relação de profundidade.

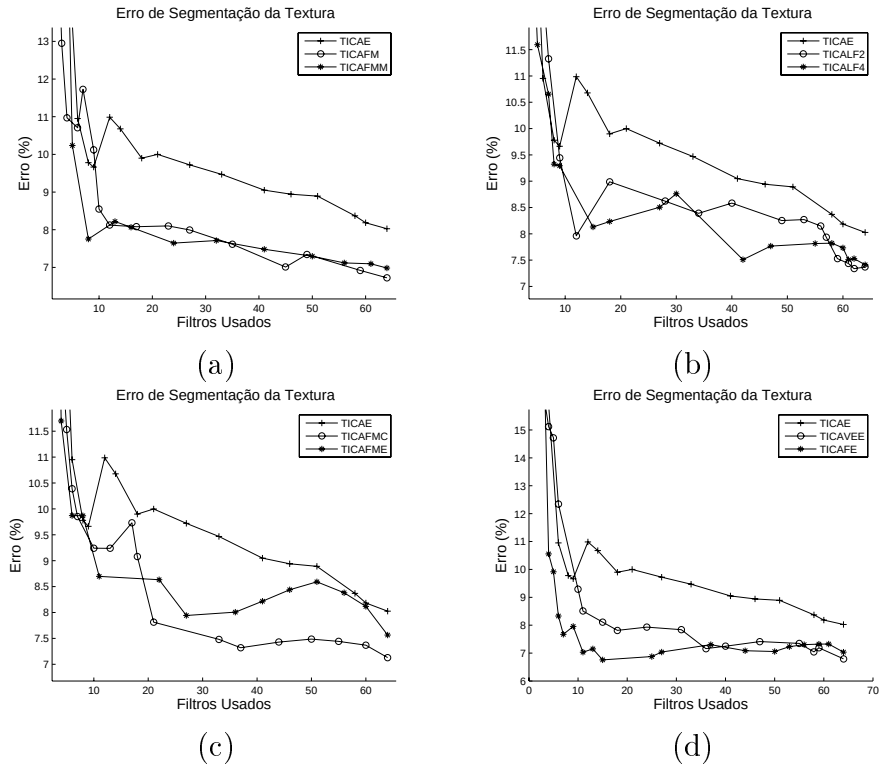


Figura 8.33: Diferenças da segmentação não-supervisionada entre a TICAE e as demais técnicas aplicadas à imagem da Figura 7.4d (pág. 96). (a) TICAE, TICAFCM e TICAFCMM, (b) TICAE, TICALF2 e TICALF4, (c) TICAE, TICAFCMC e TICAFCME, (d) TICAE, TICAVEE e TICAFE.

Assim, definindo poucos agrupamentos, as regiões segmentadas passam a comportar uma disparidade maior de características estatísticas dentro de uma mesma região.

Isso se deve ao algoritmo do *k-means* que sempre irá mapear as amostras da mistura em um dos agrupamentos (o número de grupos é uma informação disponível à priori nesse algoritmo e é definido de maneira aproximada pela inspeção visual da imagem). Todavia, o banco de filtros estatísticos obtido da TICA é o mesmo para diferentes escolhas do número de grupos do *k-means*.

Outra observação é que o modelo ICA possui alta sensibilidade para a detecção de regiões com o mesmo comportamento estatístico, uma vez que a ICA é um método de separação cega de fontes estatisticamente independentes em alta-ordem.

O uso de fractais, no critério topográfico, proporciona um aumento dessa sensibilidade do modelo, fazendo com que pequenas variações estatísticas sejam amplificadas e selecionadas para compor as componentes independentes da mistura.

O resultado pode ser observado no elevado nível de detalhamento das imagens naturais segmentadas neste trabalho, bem como na resposta em frequência dos filtros dos modelos baseados em fractais.

A Figura 8.34 apresenta as imagens dos bancos de filtros obtidos por cada técnica sobre a imagem da Figura 7.4d (pág. 96).

Observe que o critério topográfico modifica bastante os filtros obtidos pelas células simples influenciando efetivamente na caracterização da imagem visualizada. Isso pode ser notado observando-se os funções de Gabor na Figura 8.34(a) e a não visualização dessas funções na Figura 8.34b devido à existência de uma gama maior de componentes frequenciais selecionadas em cada filtro da base.

A presença de filtros orientados é nítida no banco obtido da TICAE (Figura 8.34a) e é visível em quase todas as células do banco.

Nas técnicas TICALF2 e TICALF4 (Figuras 8.34(c,d)) é possível observar uma pouco aparente orientação dos filtros e nas demais técnicas (Figuras 8.34(b,e,f,g,h,i)) não é possível verificar qualquer orientação dos filtros nos bancos.

Isto condiz com o estudo apresentado por Torralba e Oliva [103] acerca da disposição da energia do espectro de frequências de cenas com características fractais, ou seja, cenas com componentes em altas ordens estatísticas.

As bases da imagem obtidas pela TICAE (Figura 8.34a) apresentam claramente as orientações de fase das funções de Gabor, enquanto as bases obtidas pelas demais técnicas apresentam estruturas denominadas de *checkerboard-like* no domínio espacial, ou seja, as bases possuem várias componentes em frequência orientadas em diferentes direções.

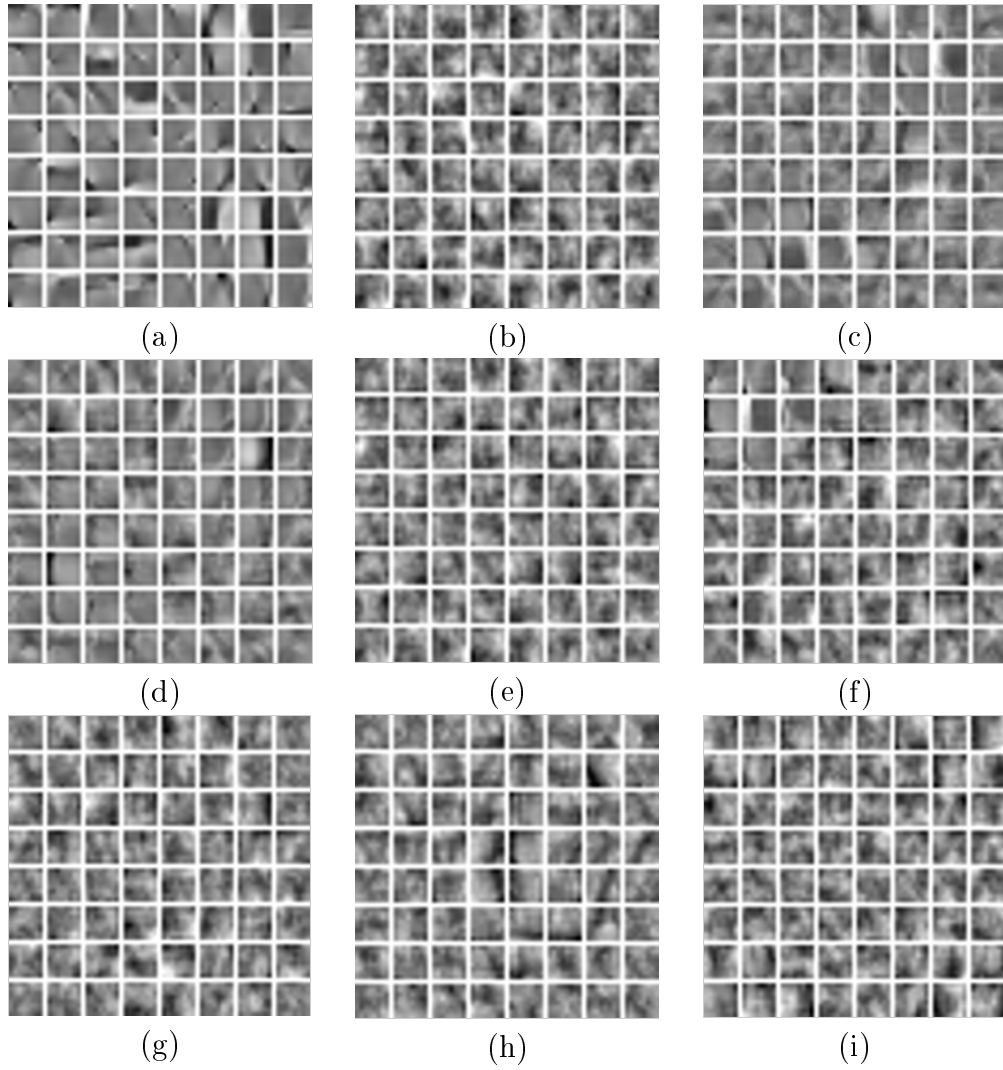


Figura 8.34: Bancos de filtros da Figura 7.4d (pág. 96) obtidos pela: (a) TICAЕ, (b) TICAФМ, (c) TICALF2, (d) TICALF4, (e) TICAФММ, (f) TICAФЕ, (g) TICAФЕЕ, (h) TICAФМC (cenais de tamanho 5, 3 e 1), (i) TICAФМЕ (escalas 5, 3 e 1).

A Figura 8.35 apresenta o espectro de energia e a assinatura espectral no domínio da frequência espacial para alguns filtros dos bancos de filtros de cada técnica.

Observa-se, em geral, nas abordagens baseadas em fractal, a ocorrência das componentes em frequência principais que correspondem à formação da estrutura de Gabor, todavia, em torno dessas há uma série de componentes em frequência espalhadas em várias direções.

Isso se deve basicamente a duas características: A primeira é baseada no estudo que afirma que as células em V1 eliminam as redundâncias de alta-ordem estatística [2, 3, 4, 11]. Essa teoria, em consonância com os descritores fractais no critério topográfico, promovem, no modelo teórico, a melhor caracterização das nuances fractais das texturas testadas, como pode ser observado no espalhamento das componentes frequenciais espaciais.

Isso mostra que as bases obtidas por componentes fractais representam melhor o espaço de características dessas texturas, inserindo outras componentes frequenciais que melhor caracterizam as imagens testadas.

A segunda característica é baseada no estudo, apresentado por Torralba e Oliva [103], e mostra que o espectro médio de potências de imagens naturais em uma escala muito próxima do observador tende a ser isotrópico em relação à orientação, como pode ser constatado na Figura 2.3 (pág. 22).

Isso reforça a proposta de uso de funções baseadas em fractais uma vez que, observando-se as Figuras 8.35(b,d,e,h) verifica-se uma certa isotropia em relação à orientação do espectro da base (observe que a textura D22 é natural), o que proporciona a formação das estruturas tipo *checkerboard-like* no domínio espacial.

Assim, pode-se justificar os resultados da segmentação não-supervisionada nos testes anteriores, por métodos fractais, devido à melhor extração de características quando comparado com a energia. Tal fato pode ser comprovado nas características de isotropia da potência em relação à orientação e pela existência de estatísticas de altas ordens nas texturas.

Observa-se, em alguns gráficos de erro da segmentação não-supervisionada das texturas analisadas, uma redução do número de filtros seguida de uma melhora discreta no resultado da segmentação. O comportamento do erro com a redução do número de filtros pode ser explicado por dois fatores.

Quando o erro de classificação é muito pequeno e a sua variação, com a retirada dos filtros, também é muito pequena, ocorre um fenômeno característico da existência vários filtros redundantes no banco gerado pelas técnicas empregadas.

A existência de filtros redundantes promove o surgimento de uma espécie de “ruído” na classificação das texturas, inserindo informações que podem atrapalhar na alocação dos pixels nas regiões de fronteira entre as texturas

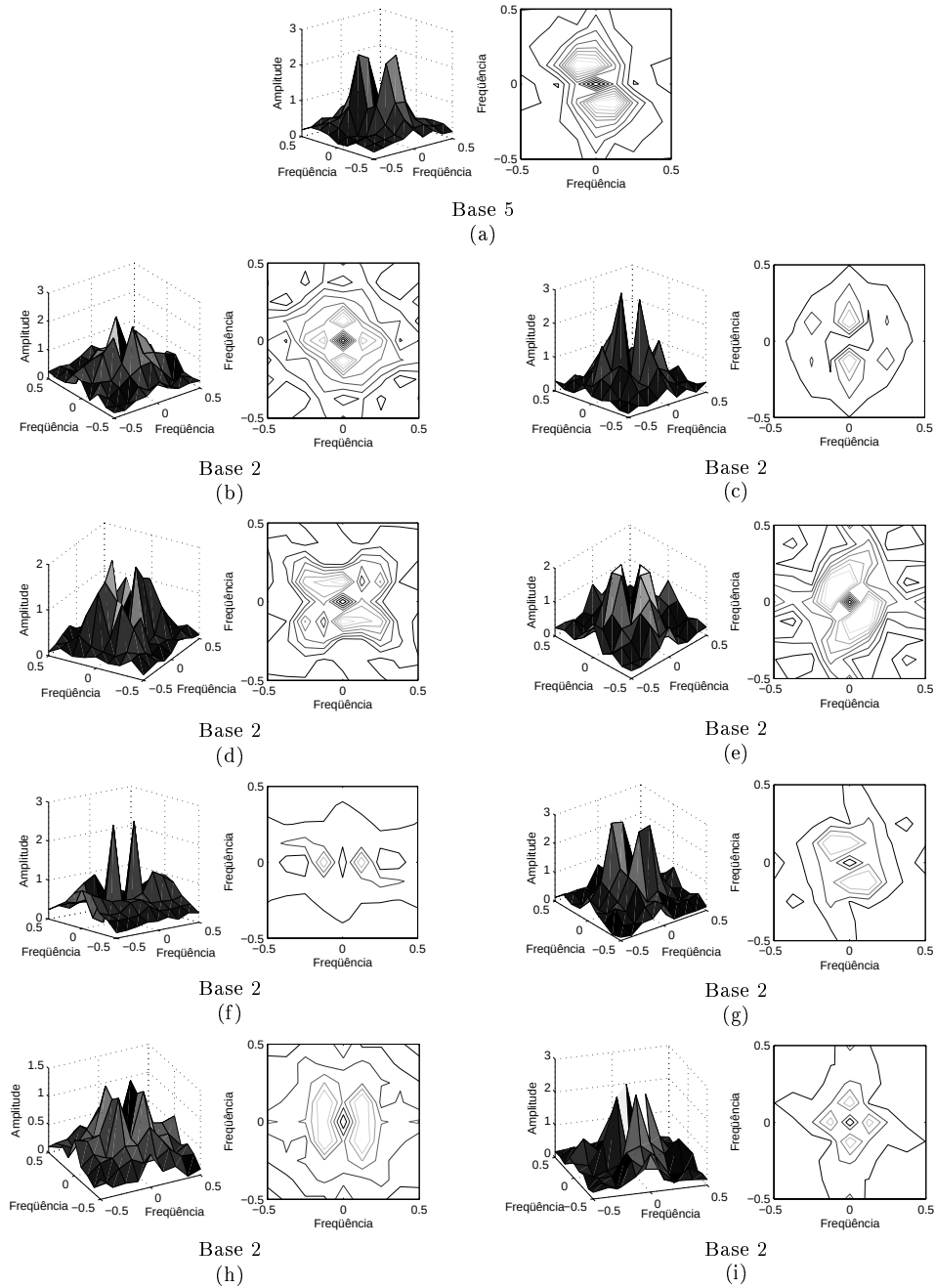


Figura 8.35: Espectro de frequências e assinatura espectral para um filtro de cada banco obtido pelas técnicas: (a) TICAЕ, (b) TICAFM, (c) TICALF2, (d) TICALF4, (e) TICAFMM, (f) TICAFE, (g) TICAVEE, (h) TICAFMC (cenas de tamanho 5, 3 e 1), (i) TICAFME (escalas 5, 3 e 1).

pelo método de classificação.

O segundo fator está relacionado com a técnica de aglomeração. O algoritmo *k-means* pode não ser robusto o suficiente para que se efetue a segmentação não-supervisionada, uma vez que o algoritmo procura agrupar as amostras das regiões informadas através da distância euclidiana entre os pontos médios e as amostras filtradas.

A redução de característica, nesse caso, favorece o mapeamento dos pixels nas regiões de fronteira pelo algoritmo de aglomeração. Todavia, o uso do *k-means* é motivado pela utilização desse método não-supervisionado nos trabalhos encontrados durante esta pesquisa.

Por outro lado, a segmentação não-supervisionada não é o objetivo principal deste trabalho e o uso de uma técnica pouco robusta, com os métodos propostos, podendo permitir a averiguação dos resultados de maneira correta, mostra que a representação proposta nesta pesquisa é bem ajustada permitindo que uma técnica de aglomeração pouco apurada funcione adequadamente com as imagens testadas.

De maneira geral, cada abordagem se adequa melhor a uma classe de texturas. As técnicas que usam o critério topográfico baseado em fractais parecem responder bem, também, quando as texturas apresentam $\lambda \approx 2$.

Esta última característica denota um comportamento de reforço ao modelo ICA, uma vez que as componentes independentes da ICA de texturas naturais são aproximações de bordas [11].

Capítulo 9

Conclusões e Projetos Futuros

Os resultados da segmentação não-supervisionada validam as duas hipóteses elaboradas nesta pesquisa. Ou seja, o uso de fractais é adequado no critério topográfico da ICA (hipótese 1), e esse uso permite melhor sintonia dos filtros na extração de características estatísticas fractais (hipótese 2), como pode ser observado nas imagens das assinaturas espectrais (Figura 8.35, pág. 139).

Os resultados mostram que imagens com características fractais são bem representadas pelos modelos propostos. A distribuição $1/f^\lambda$ é uma medida de quão estatisticamente fractal é uma imagem. Nos experimentos, quanto menor λ , melhor foram os resultados obtidos na segmentação não-supervisionada utilizada para validar os modelos propostos. Valores de $\lambda \approx 2$ nas imagens, têm mostrado resultados muito satisfatórios para a representação dos modelos propostos.

Como contribuições desta pesquisa, pode-se citar: uma melhor representação de imagens naturais no modelo TICA, através de filtros melhor sintonizados nas características fractais estatísticas dessas imagens; a elaboração de um modelo inspirado no comportamento das células V2 do sistema de visão humano, permitindo um melhor entendimento do comportamento desse sistema biológico; e uma extensão do modelo TICA, inspirado nas células em V1, como proposto por Hyvärinen et. al [49], para modelagem estatística fractal.

Esses atributos foram obtidos com a utilização dos descritores fractais no critério topográfico do modelo, permitindo uma melhor descrição das características estatísticas das imagens naturais e de imagens com características estatísticas fractais.

A contribuição de uma melhor aproximação do modelo TICA ao sistema de visão humano decorre da inclusão de algumas características das células do córtex estriado V2 ao modelo. Essas características são representadas pelo comportamento dos descritores fractais no critério topográfico.

Nesse sentido, os modelos TICA baseados em Fractais representam um avanço em relação ao modelo TICA original proposto por Hyvärinen, Hoyer e Inki [49], melhorando a representação de imagens naturais.

Essas duas contribuições determinam algumas vantagens em relação a representação de imagens naturais e a aproximação do comportamento do modelo ao mecanismo natural de visão. Dessa forma, possibilita entender melhor o comportamento visual humano, nas partes modeladas nesta pesquisa, bem como permite o uso em sistemas que necessitem interpretar imagens naturais. Sistemas de visão artificial para navegação de robôs em ambientes externos e segmentação de imagens são alguns dos exemplos.

Cada abordagem implementada possui características distintas de representação e desempenho. Porém, cada qual busca modelar características fractais no critério topográfico, representando as informações topográficas de forma bem coerente com os modelos fractais.

Todas as propostas baseadas em fractais, no geral, foram melhores que a abordagem TICA Energia. As abordagens TICAFE e TICAVEE, em alguns experimentos, tiveram resultados próximos à TICAE.

A primeira, provavelmente, devido à semelhança do método que utiliza a energia associada a heurística de localização de bordas baseada em fractais. A segunda, está relacionada às características estatísticas de segunda-ordem do CVEE.

Os modelos multi-escala (TICAFME) e multi-cena (TICAFMC) mostraram pior desempenho de processamento, demandando muito tempo do processador, porém, os resultados foram semelhantes às demais propostas baseadas em fractais.

As implementações baseadas na lacunaridade, TICALF2 e TICALF4, apresentaram desempenho bastante elevado e resultados satisfatórios, porém um pouco inferiores aos resultados dos modelos baseados somente no descritor LMME (TICAFM e TICAFMM).

Contudo, a pesquisa apresentada neste trabalho retrata apenas uma etapa da investigação inspirada na modelagem do sistema de visão humano visando possíveis aplicações em problemas de visão artificial. Estudos inspirados no mecanismo de realimentação entre os neurônios do córtex secundário e as células complexas do córtex primário podem ser implementados nos modelos propostos visando melhorar a representação dos modelos desenvolvidos aqui, especificamente, no aperfeiçoamento das funções de mapeamento da região de vizinhança [29].

A inserção de Cadeias de Markov no modelo fractal poderia permitir a modelagem da detecção de bordas oclusas, inspirado em outro comportamento das células de V2 [29], podendo proporcionar melhor representação dos objetos parcialmente oclusos nas imagens analisadas.

A implementação de funções de vizinhança multi-escalares para compor um banco com filtros de diferentes tamanhos poderá melhorar a seletividade do modelo ICA topográfico inserindo uma representação multi-escala para a base ICA [6, 42, 95].

A inclusão de informações temporais nos modelos TICA baseados em fractais, inspirado no mecanismo de ativação persistente das células em V2 [5], pode melhorar a representação de imagens, ou seja, as informações temporais de cauda-longa e alta-ordem estatística reforçam a ativação dos neurônios em V2 [5].

O estudo da implementação de funções fractais como critério topográfico da ICA não-linear [51] pode permitir a modelagem de outras características inspiradas no sistema visual humano relacionadas a não linearidades conhecidas do sistema de visão [42, 56].

O uso de outros classificadores, como o SOM (*Self-Organized Map*) ou ART (*Adaptive Resonance Theory*), no lugar do *k-means*, pode ser acrescentado como trabalho futuro no intuito de agrupar melhor as regiões segmentadas das imagens testadas. Provavelmente, o uso de classificadores com propriedades topológicas, como o SOM, melhore ainda mais os resultados da segmentação.

A análise dos modelos em texturas com características periódicas, também é uma das abordagens possíveis de serem desenvolvidos em estudos futuros nesta pesquisa.

Apêndice A

Processamento de Imagens

A.1 Morfologia

O termo *morfologia* normalmente remete o pensamento às diversas formas nas quais a matéria pode assumir, ou à lembrança dos modelos estruturas abstratos do mecanismo de linguagem, da conformação social nas cidades, dentre outras várias meações do termo.

Entretanto, a palavra *morfologia* tem aqui um contexto matemático que, conforme Gonzalez e Woods [36], representa um conjunto de técnicas para extração de componentes da imagem úteis na representação e descrição das formas de regiões, como fronteiras, esqueletos e regiões convexas.

A morfologia matemática é um modelo teórico baseado nas teorias de reticulados e de conjuntos topológicos que utiliza operações de conjunto e operações lógicas para extração de regiões da imagem.

Dentre as possíveis operações morfológicas a *dilatação* e a *erosão* são conhecidas como operações primitivas fundamentais. Assim, temos:

Definição 5 *Sejam A e B conjuntos no espaço inteiro 2-D, Z^2 . A dilatação de A por B , denotada de $A \oplus B$, é definida como:*

$$A \oplus B = \left\{ z \mid \left(\widehat{B} \right)_z \cap A \neq \emptyset \right\}, \quad (\text{A.1})$$

onde $\left(\widehat{B} \right)_z$ é a reflexão de B sob sua origem e deslocada por z .

Ou seja, a *dilatação* de A por B , então, é o conjunto de todos os deslocamentos, z , tal que B e A possuem, em comum, pelo menos um elemento. O conjunto B é normalmente referenciado como o *elemento estruturante* na *dilatação*. A Figura A.1 exemplifica a *dilatação* para dois conjuntos básicos A e B .

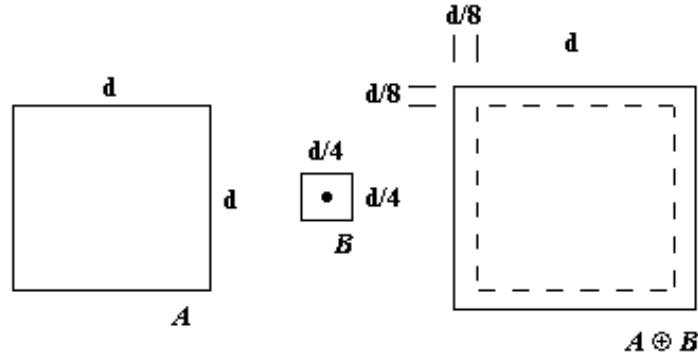


Figura A.1: Exemplo da dilatação morfológica.

Definição 6 Sejam A e B conjuntos no espaço inteiro 2-D, Z^2 . A erosão de A por B , denotada de $A \ominus B$, é definida como:

$$A \ominus B = \{z \mid (B)_z \subseteq A\}. \quad (\text{A.2})$$

Em outras palavras, a erosão de A por B é o conjunto de todos os pontos z tal que B , transladado por z , está contido em A . A Figura A.2 exemplifica a erosão para dois conjuntos A e B .

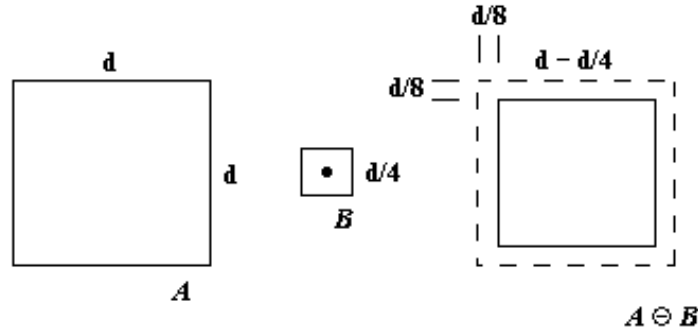


Figura A.2: Exemplo da erosão morfológica.

A extensão dos conceitos de *dilatação* e *erosão* em imagens binárias para imagens com níveis de cinza se dá através do uso de operadores de máximo e mínimo (supremo e ínfimo) de funções [40].

Assim, uma *superfície* de uma função é formada pelo valor máximo (ou supremo) da função para cada ponto pertencente a seu domínio e a *umbra* de uma função refere-se às coordenadas de todos os pontos cujos valores são menores (ínfimos) ou iguais à função.

Dessa forma Haralick [40] define a *dilatação* em níveis de cinza como a dilatação das *umbras* da imagem e do *elemento estruturante*. Ou ainda, como o máximo de uma série de somas correspondentes a translações da imagem pelo *elemento estruturante*. Matematicamente, pode-se expressar a *dilatação* em níveis de cinza da imagem f por b como,

$$(f \oplus b)(s, t) = \max \{f(s - x, t - y) + b(x, y)\}, \quad (\text{A.3})$$

onde, D_f e D_b são os domínios de f e b , respectivamente, de forma que $(s - x), (t - y) \in D_f$ e $(x, y) \in D_b$. Neste caso, b é o *elemento estruturante* do processo morfológico, sendo agora uma função e não um conjunto como na *dilatação* binária. As variáveis x, y, s e t são as coordenadas das funções (ou imagens).

Por outro lado, a *erosão* em níveis de cinza consiste na *superfície* das *erosões* binárias das *umbras* da imagem e do *elemento estruturante*. Ou ainda, no mínimo de uma série de diferenças correspondentes a translações da imagem pelo *elemento estruturante* refletido [40]. Matematicamente, pode-se expressar a *erosão* em níveis de cinza da imagem f por b como,

$$(f \ominus b)(s, t) = \min \{f(s + x, t + y) - b(x, y)\}, \quad (\text{A.4})$$

onde, D_f e D_b são os domínios de f e b , respectivamente, de forma que $(s + x), (t + y) \in D_f$ e $(x, y) \in D_b$. Todas as demais variáveis são definidas da mesma forma que na *dilatação*.

Outros dois conceitos de grande relevância na teoria morfológica matemática são o de *conectividade* e *vizinhança*. Um pixel p na coordenada (x, y) em uma imagem possui vizinhos horizontais e verticais dados pelas coordenadas $(x + 1, y)$, $(x - 1, y)$, $(x, y + 1)$ e $(x, y - 1)$. Este conjunto de pixels é chamado de 4-vizinhos de p e denotado por $N_4(p)$ e cada pixel é uma unidade de distância a partir de (x, y) .

Os quatro pixels das diagonais nas coordenadas $(x + 1, y + 1)$, $(x + 1, y - 1)$, $(x - 1, y + 1)$ e $(x - 1, y - 1)$ compõem o conjunto $N_D(p)$. Estes pontos juntamente com os 4-vizinhos de p são chamados de 8-vizinhos de p e denotados por $N_8(p)$.

A conectividade entre pixels, por sua vez, estabelece se dois pixels estão conectados, ou seja, se são adjacentes e seu nível de cinza satisfaz algum critério específico de similaridade [36]. Gonzalez e Woods [36] consideram três tipos de adjacências entre pixels:

1. A 4-adjacência. Dois pixels p e q com valores em uma imagem são 4-adjacentes se q está no conjunto $N_4(p)$.

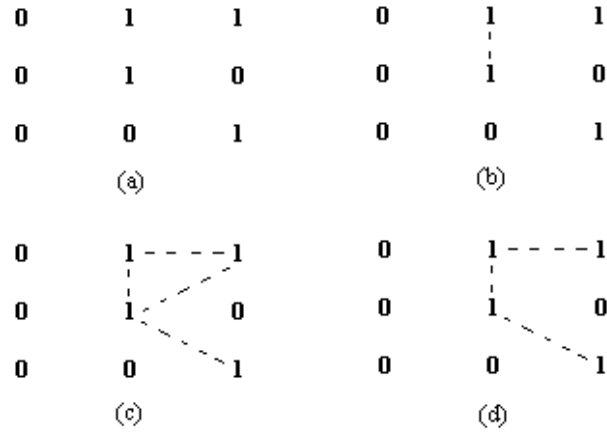


Figura A.3: Exemplo de adjacência [36]. (a) Arranjo dos pixels em uma parte da imagem; (b) Conjunto de pixels 4-adjacentes em relação ao pixel central; (c) Conjunto de pixels 8-adjacentes e (d) Conjunto de pixels m-adjacentes.

2. A 8-adjacência. Dois pixels p e q com valores em uma imagem são 8-adjacente se q está no conjunto $N_8(p)$.
3. A m-adjacência (adjacência mixada). Dois pixels p e q com valores em uma imagem são m-adjacente se:

- (a) q está em $N_4(p)$, ou
- (b) q está em $N_D(p)$ e o conjunto $N_4(p) \cap N_4(q)$ não possui pixels.

A Figura A.3 exemplifica os três tipos de adjacências para uma região de uma imagem binária. A similaridade é considerada para o valor um.

Para entender o conceito de conectividade, considere S um subconjunto de pixels de uma imagem. Dois pixels p e q são conectados em S se existe um caminho entre eles consistindo inteiramente por pixels em S . Assim, para qualquer pixel p em S , o conjunto de pixels que são conectados a p em S é chamado de *componente conectado* de S . Se existe apenas um componente conectado, então o conjunto S é chamado de *conjunto conectado* ou *região* da imagem. A partir do conceito de região define-se o conceito de *fronteira* (ou *contorno*) pelo conjunto de pixels na região que possui um ou mais vizinhos que não pertencem à região.

É importante diferenciar o conceito de *fronteira* do de *borda*. Enquanto a fronteira forma um caminho fechado em torno de uma região, bordas são formadas por pixels com um valor derivativo que excede um determinado limiar [36]. Assim, o conceito de *borda* é local e baseado, por exemplo, na descontinuidade dos níveis de cinza de uma imagem.

É possível ligar os pontos de bordas em segmentos que irão corresponder à fronteira de uma região, todavia, esse nem sempre é o caso [36]. A exceção em que bordas e fronteiras são iguais ocorre nas imagens binárias.

A.2 Segmentação, Representação, Descrição e Classificação

O processo de *segmentação*, em processamento de imagens, consiste em subdividir uma imagem em regiões ou objetos [36], ou ainda, na associação das sub-unidades da imagem a objetos da cena [105]. Porém, o nível de subdivisões realizadas no processo de segmentação depende do problema a ser resolvido. A segmentação deve parar quando os objetos de interesse da aplicação forem isolados.

Os algoritmos de segmentação, geralmente, são baseados em duas propriedades básicas dos valores de intensidade dos pixels: a *descontinuidade* e a *similaridade* [36]. Na primeira, mudanças abruptas na intensidade, como bordas, permitem particionar a imagem em regiões. Na segunda, procura-se isolar regiões que possuam semelhanças de acordo com um conjunto de critérios pré-definidos (limiarização e regiões de crescimento, redução e mistura são exemplos da segunda propriedade).

Já Turner [105] afirma que, na prática, existem três métodos principais de segmentação: métodos baseados em pixels, métodos baseados em regiões e métodos baseados em bordas. Nota-se que os dois primeiros métodos são casos relacionados à similaridade, enquanto o último à descontinuidade.

Apesar da aparente simplicidade, a segmentação de imagens pode ser uma das tarefas mais difíceis em processamento de imagens [36]. Texturas e imagens naturais são exemplos onde a segmentação é nada trivial. Isso se deve, em boa parte, à natureza estatística dessas imagens e à complexidade na identificação das similaridades entre suas diferentes regiões.

Algoritmos baseados em estatísticas de diferentes ordens têm mostrado bons resultados na segmentação dessas imagens [12, 48, 53, 58, 61]. Esses algoritmos usam o princípio da similaridade e descontinuidade nas propriedades estatísticas das regiões ou dos objetos da cena e, nesse caso, a segmentação é denominada de *segmentação estatística*.

Após a segmentação da imagem em regiões, os pixels agregados precisam ser representados e descritos em uma forma adequada para processamentos posteriores.

A *representação* de regiões envolve, basicamente, duas escolhas: A representação externa (através das fronteiras), ou interna (pelas características

dos pixels).

A *descrição* das regiões é baseada na escolha da representação, por exemplo, regiões representadas pela fronteira possuem descritores de fronteiras que fornecem características como o tamanho, a orientação, o número de concavidades, entre outras.

A imagem segmentada e devidamente representada e descrita permite que seja efetuada uma *classificação* de suas partes para orientar o processamento das informações de interesse na análise da imagem.

Enquanto a segmentação de imagens visa subdividir uma imagem em regiões, a classificação busca identificar essas regiões de acordo com características (ou atributos) que podem ser definidas *à priori* (*classificação supervisionada*) ou obtidas durante o processo de classificação (*classificação não-supervisionada*). O conceito, então, pode ser sintetizado como uma etapa do processo de reconhecimento de padrões onde são identificadas as estruturas segmentadas de uma imagem.

O método de classificação não-supervisionado usa somente os atributos estatísticos da imagem para estabelecer as fronteiras entre os conjuntos de dados que se agrupam naturalmente (mecanismo denominado de *clustering*). As classes espectrais resultantes são interpretadas de acordo com a correlação entre seus componentes.

Já o método de classificação supervisionado faz uso de amostras de treinamento obtidas de imagens semelhantes cujas características são conhecidas e irão permitir estabelecer as características espectrais das classes.

No caso de texturas, comumente são usadas três abordagens para a classificação: análise no espaço freqüencial, matrizes de co-ocorrência (dependência espacial em níveis de cinza) e autocorrelação direcional (para determinar periodicidade) [105].

O método ICA, usado nessa pesquisa, usa estatísticas de alta-ordem, de maneira não-supervisionada, para obtenção de filtros espaciais que separam amostras correlatadas em alta-ordem estatística.

Referências Bibliográficas

- [1] AN, K. H.; LEE, M.; SHIN, J. K., *Saliency map model based on the edge images of natural scenes*, IEEE Proceedings of the 2002 International Joint Conference on Neural Networks, 1, pp. 1023-1027, 2002.
- [2] ATICK, J. J.; REDLICH, A. N., *Towards a theory of early visual processing*, Neural Computation, 2, pp. 308-320, 1990.
- [3] ATICK, J. J.; REDLICH, A. N., *What does the retina know about natural scenes?*, Neural Computation, 4, pp. 196-210, 1992.
- [4] ATICK, J. J.; REDLICH, A. N., *Convergent algorithm for sensory receptive field development*, Neural Computation, 5, pp. 45-60, 1993.
- [5] BADDELEY, R.; ABBOTT, L. F.; BOOTH, M. C. A.; SENGPIEL, F.; FREEMAN, T.; WAKEMAN, E. A.; ROOLS, E. T., *Responses of neurons in primary and inferior temporal visual cortices to natural scenes*, Proc. R. Soc. Land, Royal Society, B(264), pp. 1175-1783, 1997.
- [6] BALBOA, R. M.; GRZYWACZ, N. M., *The role of early lateral inhibition: more than maximizing luminance information*, Vision Research, 17, pp. 77-89, 2000.
- [7] BALBOA, R. M.; GRZYWACZ, N. M., *Power spectra and distribution of contrast of natural images from different habitats*, Vision Research, 43, pp. 2527-2537, 2003.
- [8] BARLOW, H. B.; TOLHUST, D. J., *Why do you have edge detectors?*, Optical Society of America Technical Digest, 23(172), 1992.
- [9] BARLOW, H. B., *What is the computational goal of the neocortex?*, In Large scale neuronal theories of the brain (ed. C. Koch), Cambridge: MIT Press, pp. 1-22, 1994.

- [10] BEDNAR, J. A.; KELKAR, A.; MIIKKULAINEN, R., *Scaling self-organizing maps to model large cortical networks*, Neuroinformatics, 2, pp. 275-302, 2004.
- [11] BELL, A.; SEJNOWSKI, T., *The 'independent components' of natural scenes are edge filters*, Vision Research, 37, pp. 3327-3338, 1997.
- [12] BLANCO, José A. M., *Jerarquización de estructuras de nível bajo y médio para reconhecimento visual. Aplicações a Texturas y Formas.*, Tesis Doctoral, Departamento de Informática y Sistemas, Universidad Politécnica de Canarias, 1987.
- [13] BORGNE, H.; GUYADER, N.; GUÉRIN-DUGUÉ, A.; HERÁULT, J., *Classification of images: ica filters vs human perception*, Signal Processing and Its Applications, Proceedings IEEE, v.2, pp. 251-254, 2003.
- [14] BRODATZ, P., *Texture: a photograph album for artists and designs*, Dover, New York, 1956.
- [15] CAJAL, S. R. Y., *Histologie Du Système Nerveux de l'Homme et Des Vertébrés*, Maloine, Paris, 1991.
- [16] CARDOSO, J. F., *Blind signal separation: statistical principles*, Proceedings of the IEEE, 9(10), pp. 2009-2025, 1998.
- [17] CHANCE, F. S.; NELSON, S. B.; ABBOTT, L. F. *Complex cell as cortically amplified simple cells*, Nature Neuroscience, 3(2), pp. 277-282, 1999.
- [18] CÔCO, K. F.; SALLES, E. O. T.; SARCINELLI Filho, M., *Uso do parâmetro de hurst como critério topográfico na análise de componentes independentes*, VII SBAI / II IEEE LARS, São Luis - MA, v.1 pp. 33-37, 2005.
- [19] CÔCO, K. F.; SALLES, E. O. T.; SARCINELLI Filho, M., *Banco De Filtros Ica Topografico Fractal, Baseado em Morfologia, Aplicado na Segmentação de Texturas*. In: XVI Congresso Brasileiro de Automática (XVI CBA), Salvador - BA, v.1 pp. 1-6, 2006.
- [20] CÔCO, K. F.; SALLES, E. O. T.; SARCINELLI Filho, M., *Topographic Independent Component Analysis Based on Fractal Theory and Morphology Applied to Texture Segmentation*. Signal Processing (doi:10.1016/j.sigpro.2007.01.027), v.87 pp. 1966-1977, 2007.

- [21] CÔCO, K. F.; SALLES, E. O. T.; SARCINELLI Filho, M., *Comparando a Lacunaridade Fractal e o Descritor Morfológico Fractal como Critério Topográfico na Análise de Componentes Independentes*. Signal Processing VIII SBAI / VIII CBRN, Florianópolis - SC, 2007.
- [22] DAUGMAN, J. G., *Complete discrete 2-D gabor transforms by neural networks for image analysis and compression*, IEEE Transactions on Acoustics, Speech and Signal Processing. 36 (7) pp. 1169-1179, 1988.
- [23] DAUGMAN, J. G., *Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters*, J. Opt. Soc. Am. A, 2 pp.1160-1169, 1985.
- [24] DEVALOIS, R. L.; YUND, E. W.; HEPLER, N., *The orientation and direction selectivity of cells in the macaque visual cortex*, Vis Res, 22, pp. 531-544, 1982.
- [25] DOI, Eizaburo; LEWICKI, Michael S., *Relations between the statistical regularities of natural images and the response properties of the early visual system*, Japanese Cognitive Science Society, Sig P&P, Univ. Kyoto, July 28, 2005.
- [26] DOUKHAN, P.; OPPENHEIM, G.; TAQQU, M. S., *Theory and applications of long-range dependency*, Boston: Birkhäuser, 2003.
- [27] DUDA, R. O.; HART, P. E.; STORK, D. G., *Pattern classification*, New York: Willey, 2000.
- [28] FALCONER, Kenneth, *Fractal geometry: mathematical foundations and applications*, London: Wiley, 2003.
- [29] FANGTU, T. Q.; RÜDIGER, von der Heydt, *Figure and ground in the visual cortex: V2 combines stereoscopic cues with gestalt rules*, DOI 10.1016/j.neuron.2005.05.028, Elsevier, 2005.
- [30] FERREIRA, A. B. de H., *Novo dicionário aurélio - séc. XXI*, Editora Nova Fronteira, São Paulo, 1999.
- [31] FIELD, D. J., *Relations between the statistics of natural images and the response properties of cortical cells*, Opt. Soc. Am. A, 4, pp. 2379-2394, 1987.
- [32] FIELD, D. J., *What is the goal of sensory coding?*, Neural Computation, 6, pp. 559-601, 1994.

- [33] FRACTART, *Janelas Para o Infinito Exposição de Fractais*. [online] Disponível na Internet via WWW.URL: <http://www.fractarte.com.br/galeria2/galeria.php>. Imagens capturadas em 29 de abril de 2007.
- [34] GALLETTI, C.; GAMBERINI, M.; KUTZ, D. F.; BALDINOTTI, I.; FATTORI, P., *The relationship between V6 and PO in macaque extrastriate cortex*, European Journal of Neuroscience, doi:10.1111/j.1460-9568.2005.03911.x, 21, 2005.
- [35] GALUSKE, R. A. W.; SINGER, W., *The origin and topography of long-range intrinsic projections in cat visual cortex: a developmental study*, Cerebral Cortex, 6, pp. 417-430, 1996.
- [36] GONZALEZ, R. C.; WOODS, R. E., *Digital image processing*, New Jersey: Prentice-Hall, 2002.
- [37] GREGORY, R. L., *Knowledge in perception and illusion*, Phil. Trans. R. Soc. Lond. B, 352, pp. 1121-1125, 1997.
- [38] GROSSBERG, S.; WILLIAMSON, J. R., *A neural model of how horizontal and interlaminar connections of visual cortex develop into adult circuits that carry out perceptual grouping and learning*, Cerebral Cortex, Oxford University Press, 11, pp. 37-58, 2001.
- [39] HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I., *Textural features for image classification*, IEEE Trans. Syst., Man, Cybern., SMC-3, pp. 610-621, 1973.
- [40] HARALICK, R. M., *Mathematical morphology and computer vision*, Twenty-Second Asilomar Conference on Signals, Systems and Computers, vol 1, pp. 468-479, 1988.
- [41] HARTE, D., *Multifractals: theory and applications*, Florida: Chapman & Hall/CRC, 2001.
- [42] HATEREN, J. H.; SCHAAF, A., *Independent component filters of natural images compared with simple cells in primary visual cortex*, Proc. Royal Society of London, B265, pp. 359-366, 1998.
- [43] HAVLICEK, J. P.; HARDING, D. S.; BOVIC, A. C., *Multidimensional quasi-eigfunction approximations and multicomponent AM-FM models*, IEEE Trans. on Image Processing, 9(2), pp. 227-242, 2000.

- [44] HEIDEMANN, G., *The principal components of natural images revisited*, IEEE Trans. on Pattern Analysis and Machine Intelligence, 28(5), pp. 822-826, 2006.
- [45] HEYDT, R. V. D.; PETERHANS, E.; DÜRSTELER, M. R., *Periodic-pattern-selective cells in monkey visual cortex*, The journal of Neuroscience, 12(4), pp. 1416-1434, 1992.
- [46] HUBEL, D. H.; WIESEL, T. N., *Receptive fields and functional architecture of the monkey striate cortex*, Journal of Physiology, 195, pp. 215-243, 1968.
- [47] HYVÄRINEN, A.; HOYER, P., *Emergence of phase and shift invariant features by decomposition of natural images into independent feature subspaces*, Neural Computation, 12(7) pp. 1705-1720, 2000.
- [48] HYVÄRINEN, A.; HOYER, P. O.; HURRI, J.; GUTMANN, M., *Statistical models of images and early vision*, Proceedings of the Int. Symposium on Adaptive Knowledge Representation and Reasoning (AKRR2005), Espoo, Finland, 2005.
- [49] HYVÄRINEN, A.; HOYER, P. O.; INKI, M., *Topographic independent component analysis*, Neural Computation, 13(7), pp. 1527-1558, 2001.
- [50] HYVÄRINEN, A.; GUTMANN, M.; HOYER, P. O., *Statistical model of natural stimuli predicts edge-like pooling of spatial frequency channels in V2*, BMC Neuroscience, (6)12, 2005.
- [51] HYVÄRINEN, A.; KARHUNEN, J.; OJA, E., *Independent component analysis*, New York: John Wiley and Sons, 2001.
- [52] JAIN, A. K., *Fundamentals of digital image processing*, New Jersey: Prentice-Hall, 1989.
- [53] JENSSEN, R.; ELTOFT, T., *Ica filter bank for segmentation of textured images*, in Proc Int'l. Workshop on Independent Component Analysis and Blind Signal Separation (ICA2003), pp. 827-832, 2003.
- [54] JINHUI, c.; KISHIGAMI, T.; MINOWA, K.; TSUJII, S., *Artificial neural networks which can see geometric illusions in human vision*, Neural Network Proceedings IEEE, 3(25), pp. 2209-2212, 1993.
- [55] JONES, J. P.; PALMER, L. A., *An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex*, Journal of Neurophysiology, 58, pp. 1233-1258, 1987.

- [56] KAGAN, I.; GUR, M.; SNODDERLY, M., *Spatial organization of receptive fields of V1 neurons of alert monkeys: comparison with responses to gratings*, Journal of neurophysiology, The American Physiological Society, 88, pp. 2557-2574, 2002.
- [57] KAI, K.; NASELARIS, T.; GALLANT, J., *Estimation of voxel receptive fields in human visual cortex using natural images [Abstract]*, Journal of Vision, 7(9), pp. 79, 2007.
- [58] KAPLAN, Lance M., *Extended fractal analysis for texture classification and segmentation*, IEEE Trans. on Image Processing, 8(11) pp. 1572-1585, 1999.
- [59] KARKLIN, Y.; LEWICKI, M. S., *Learning higher-order structures in natural images*, Network: Comput. Neural Syst., 14, pp. 483-499, 2003.
- [60] KARKLIN, Y.; LEWICKI, M. S., *A hierarchical bayesian model for learning nonlinear statistical regularities in nonstationary natural signals*, Neural Computation, 17, pp. 397-423, 2005.
- [61] KASPARIS, T. et al., *Segmentation of textured images based on fractals and image filtering*, Pattern Recognition, 34 pp. 1963-1973, 2001.
- [62] KIVINIEMI, V.; KANTOLA, J.; JAUHIAINEN, J.; HIVÄRINEN, A.; TERVONEN, O., *Independent component analysis of nondeterministic fMRI signal sources*, NeuroImage, 19 pp. 253-260, 2003.
- [63] KOHONEN T., *Self-organizing maps*, Springer Series in Information Sciences, Vol. 30, Springer, Heidelberg, 1997.
- [64] KRUIZINGA, P.; PETKOV, N., *Nonlinear operator for oriented texture*, IEEE Transactions on Image Processing, 8(10), pp. 1395-1407, 1999.
- [65] KURUOGLU, E. E., *Signal processing in α -stable noise environments: a least l_p -norm approach*, PhD Thesis, University of Cambridge, 22 nov. 1998.
- [66] LEE, T. S., *Image Representation Using 2D Gabor Wavelets*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(10), pp. 959-971, 1996.
- [67] LEVENTHAL, A. G.; WANG, Y. C.; SCHMOLESKY, M. T.; ZHOU, Y., *Neural correlates of boundary perception*, Vis Neurosci, 15, pp. 1107-1118, 1998.

- [68] MANDELBROT B. B., *The fractal geometry of nature*, Freeman, New York, 1983.
- [69] MANDELBROT B. B., *A fractal's lacunarity, and how it can be tuned and measured*, Fractals in biology and medicine TF Nonnenmacher, GA Losa, ER Weibel (Basel and Boston: Birkhäuser Verlag), 1994.
- [70] MANDUCHI, R.; PORTILLA, J., *Independent component analysis of textures*, in proc. Int'l Conference on Computer Vision, Corfu, 9, pp. 1054-1060, 1999.
- [71] MATHWORKS, The, Inc. Matlab, *The language for technical computing*, R2006a, 2006.
- [72] MIIKKULAINEN, R.; BEDNAR, J. A.; CHOE, Y.; SIROSH, J., *Self-organization, plasticity, and low-level visual phenomena in a laterally connected map model of the primary visual cortex*, in Goldstone, R. L. and Schyns, P. G. and Medin, D. L., Eds. Psychology of Learning and Motivation, v36: Perceptual Learning, pp. 257-308. Academic Press, San Diego CA, 1997.
- [73] MUTIHAC, R.; VAN HULLE, M. M., *Statistics of feature extraction by topographic independent component analysis from natural images*, Proceedings 2nd Int. Conf. on Electronics, Control and Signal Processing, World Scientific and Engineering Academy and Society (WSEAS). (Singapore, 7-12 December, 2002), CD ROM, pp. 451-289, 2003.
- [74] NATIONAL INSTITUTE OF HEALTH, *Today Vision Tomorrow - National Eye Institute*. [online] Disponível na Internet via WWW.URL: <http://www.nei.nih.gov/health/eyediagram/eyeimages1.asp>. Imagens capturadas em 29 de agosto de 2007.
- [75] NUNES, E. O.; CONCI, A., *O coeficiente de hurst e o coeficiente de variação espacial na segmentação de texturas multiespectrais*, In: 4o Congresso Temático de Dinâmica, Controle e Aplicações, Bauru. anais do IV Congresso Temático de Dinâmica Controle e Aplicações, Petropolis: SBMAC, v.4. pp.1089-1096, 2005.
- [76] OLSHAUSEN, B. A.; FIELD, D. J., *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, Nature, 381, pp. 607-609, 1996.

- [77] PAPOULIS, A.; PILLAI, S. U., *Probability, random variables and stochastic processes*, 4^a ed. New York: McGraw-Hill, 2002.
- [78] PENTLAND, A. P., *Fractal based description of natural scenes*, IEEE Trans. Pattern. Anal. Mach. Intell., 6(6), pp. 661-674, 1984.
- [79] PERCEPTUAL SCIENCE GROUP AT MIT, *Checkerboard Illusion*. Massachusetts Institute of Technology, [online] Disponível na Internet via WWW.URL: <http://web.mit.edu/persci/>. Imagens capturadas em 29 de agosto de 2007.
- [80] PICINBONO, B., *Random signals and systems*, New Jersey: Prentice Hall International, 1993.
- [81] POLAT, U.; TYLER, C. W., *What pattern the eye sees best*, Vision Research, 39, pp. 887-895, 1999.
- [82] POLLEN, D.; RONNER, S., *Visual cortical neurons as localized spatial frequency filters*, IEEE Trans. on Systems, Man and Cybernetics, 13, pp. 907-916, 1983.
- [83] POTTER, M.; KINSNER, W., *Effects of ICA on the estimation of fractal sources*, Canadian Conference on Electrical and Computer Engineering, IEEE CCECE 2003, v.2, 4, pp. 1031-1034, 2003.
- [84] PRATT, W. K., *Digital image processing*, John Wiley and Sons, New York, 1991, Chapter 17, pp. 569-577.
- [85] PRATT, W. K.; FAUGERAS, O. D.; GAGALOWICZ, A., *Applications of stochastic texture field models to image processing*, Proceedings of IEEE, v.69, 5, pp. 542-551, 1981.
- [86] POSNER, M. I.; GILBERT, C. D., *Attention and primary visual cortex*, Proc. Natl. Acad. Sci. USA, v.96, pp. 2585-2587, 1999.
- [87] QIU, G., *Modelling the visual cortex using artificial neural networks for visual image reconstruction*, Fourth International Conference on Artificial Neural Networks, 26(28), pp. 127-132, 1995.
- [88] RAO, R. P. N., *Bayesian inference and attentional modulation in the visual cortex*, Cognitive Neuroscience and Neuropsychology, 16(16), pp. 1843-1848, 2005.
- [89] RUSS, C. J., *The image processing handbook*, Florida: CRC-Press, 1999.

- [90] SALLES, Evandro O. T., *Um Método de Classificação de Texturas com Rotação Baseado na Modelagem HMM de Características AM-FM.*, Tese de Doutorado, DECOM/FEEC, UNICAMP, Fevereiro, 2001.
- [91] SARKAR, N.; CHAUDHURI, B. B., *An efficient differential box-counting approach to compute fractal dimension of image*, IEEE Transactions on Systems, Man, and Cybernetics, 24(1), pp. 115-120, 1994.
- [92] SCHILLER, P. H.; FINLAY, B. H.; VOLMAN, S. F., *Quantitative studies of single-cell properties in monkey striate cortex*, I. Spatiotemporal organization of receptive fields, J Neurophys, 39, pp. 1288-1319, 1976.
- [93] SCHOFIELD, A. J.; FOSTER, D. H., *Artificial Neural Networks Simulating Visual Texture Segmentation and Target Detection in Line-Element Images*, Philosophical Transactions: Biological Sciences, 350(1334), pp. 401-412, 1995.
- [94] SERENCES, J. T.; SCHWARZBACH, J.; COURTNEY, S. M.; GOLAY, X.; YANTIS, S., *Control of object-based attention in human cortex*, Cerebral Cortex Advance Access, Oxford University Press, doi:10.1093/cercor/bhh095, pp. 1-12, 2004.
- [95] SIMONCELLI, E. P.; OLSHAUSEN, B. A., *Natural image statistics and neural representation*, Annu. Rev. Neurosci., 24, pp. 193-216, 2001.
- [96] SKLANSKY, J., *Image segmentation and feature extration*, IEEE Trans. Syst., Man, Cybern., 13(5), pp. 907-916, 1978.
- [97] SOMERS, D. C.; TODOROV, E. V.; SIAPAS, A. G.; TOTH, L. J.; KIM, D.; SUR, M., *A local circuit approach to understanding integration of long-range inputs in primary visual cortex*, Cerebral Cortex, 8, pp. 204-217, 1998.
- [98] SONKA, M.; HLAVAC, V.; BOYLE, R., *Image processing, analysis and machine vision*, London: Chapman & Hall, 1993.
- [99] SPROTT, *Sproot's fractal gallery*. [online] Disponível na Internet via WWW.URL: <http://sprott.physics.wisc.edu/fractals.htm>. Imagens capturadas em 29 de abril de 2007.
- [100] STARK, H.; WOODS, J. W., *Probability and random processes with applications to signal processing*, New Jersey: Prentice Hall, 2002.
- [101] THERRIEN, C. W., *Discrete random signals and statistical signal processing*, New Jersey: Prentice Hall, 1992.

- [102] TOPOGRAPHICA, *Neural Map Simulator*. [online] Disponível na Internet via WWW.URL: <http://topographica.org/Home/index.html>. Imagens capturadas em 29 de agosto de 2007.
- [103] TORRALBA, Antonio; OLIVA, Aude., *Statistics of natural image categories*, Institute of Physics Publishing: Computation in Neural Systems, 14, 2003.
- [104] TURIEL, Antonio Mato; GERMÁN, Néstor Parga, *Self-similarity properties of natural images resemble those of turbulent flows*, Physical Review Letters. 80(5) pp. 1098-1101, 1998.
- [105] TURNER, M. J.; BLACKLEDGE, Jonathan M.; ANDREWS, Patrick R., *Fractal geometry in digital imaging*, San Diego: Academic Press, 1998.
- [106] UNSER, M., EDEM, M., *Nonlinear operators for improving texture segmentation based on features extracted by spatial filtering*, IEEE Transactions on Systems, Man, Cybernetics, 20, pp. 804-815, 1990.
- [107] VINJE, W. E.; GALLANT, J. L., *Sparse coding and decorrelation in primary visual cortex during natural vision*, Science, 287, pp. 1273-1276, 2000.
- [108] WORNELL, G. W., *Signal processing with fractals. a wavelet-based approach*, Upper Saddle River: Prentice Hall, 1996.
- [109] XIA, Y.; FENG, D.; ZHAO, R., *Morphology-based multifractal estimation for texture segmentation*, IEEE Transaction on Image Processing, 15(3), pp. 614-623, 2006.
- [110] YIN, J.; PETROPULUS, A. P., *On the relationship between $1/f$ and α -stable processes*, Information Theory Proceedings IEEE, 16(21), pp. 1585-1588, 1998.
- [111] ZHANGHOUHONG, Du Gan, *Application of high-order fractal feature to edges detection*, Proceedings of ICSP 98 pp. 1006-1008, 1998.