

**RODRIGO JANASIEVICZ GOMES PINHEIRO**

**DEFINIÇÃO DE UMA METODOLOGIA  
DE AVALIAÇÃO DA QUALIDADE DE  
SEGMENTAÇÃO DE IMAGENS**

**Curitiba  
Julho de 2005**

RODRIGO JANASIEVICZ GOMES PINHEIRO

DEFINIÇÃO DE UMA METODOLOGIA  
DE AVALIAÇÃO DA QUALIDADE DE  
SEGMENTAÇÃO DE IMAGENS

Dissertação de Mestrado, apresentada ao  
Curso de Mestrado em Informática Apli-  
cada da Pontifícia Universidade Católica  
do Paraná .

Área de Atuação: Processamento e  
Análise de Imagens

Orientador: Prof. Dr. Jacques Facon.

Curitiba  
Julho de 2005

Pinheiro, Rodrigo Janasiewicz Gomes

Definição de uma Metodologia de Avaliação da Qualidade de Segmentação de Imagens, Curitiba 2005. 101p

Dissertação (Mestrado) – Pontifícia Universidade Católica do Paraná.  
Programa de Pós-Graduação em Informática Aplicada.

1. Processamento e Análise de imagens 2. Processamento de Documentos.

I. Pontifícia Universidade Católica do Paraná II. Centro de Ciências Exatas e de Tecnologia III. Curso de Mestrado em Informática Aplicada

Dedico aos meus pais, irmãos e amigos,  
por toda a paciência e compreensão

# AGRADECIMENTOS

Ao Professor, orientador e amigo Prof. Dr. Jacques Facon, pelo incentivo, apoio e sobretudo, pela compreensão em alguns momentos de dificuldade.

Aos colegas e amigos David Menoti e Luiz Antonio Pereira Neves, pela "força" concedida.

Aos professores e colegas do PPGIA.

A Pontifícia Universidade Católica do Paraná e ao Programa de Pós Graduação em Informática Aplicada, pela oportunidade concedida.

A todos que contribuíram de forma direta ou indireta para este trabalho.

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
1.1	Desafio . . . . .	2
1.2	Motivação . . . . .	2
1.3	Proposta . . . . .	2
1.4	Contribuição . . . . .	3
1.5	Organização . . . . .	3
<b>2</b>	<b>DESCRIÇÃO DA ÁREA DE ESTUDO</b>	<b>5</b>
2.1	Introdução . . . . .	5
2.2	Segmentação por região . . . . .	6
2.2.1	Introdução . . . . .	6
2.2.2	Divisão e fusão ou <i>split and merge</i> . . . . .	6
2.2.3	Limiarização, binarização. . . . .	7
2.2.4	Metodologias de segmentação por região de envelopes postais .	12
2.2.5	Segmentação <i>Wavelet</i> de envelopes postais de [MFBJ03] . . . .	12
2.2.5.1	Decomposição <i>Wavelet</i> . . . . .	13
2.2.5.2	Identificação dos pontos salientes . . . . .	14
2.2.5.3	Rotulação controlada das janelas de pontos salientes	15
2.2.5.4	Projeção reversa dos pontos salientes selecionados . .	16
2.2.5.5	Perseguição de contorno através dos pontos salientes	16
2.2.6	Segmentação de envelopes postais de Eiterer et al [EFM04] utilizando uma abordagem fractal . . . . .	19
2.2.6.1	Dimensão fractal para imagens em níveis de cinza . .	19
2.2.6.2	<i>Clustering</i> . . . . .	20
2.2.6.3	Identificação das classes . . . . .	22

<b>3</b>	<b>METODOLOGIAS DE AVALIAÇÃO DA QUALIDADE DA SEG- MENTAÇÃO</b>	<b>25</b>
3.1	Introdução . . . . .	25
3.2	Imagens <i>ground truth</i> . . . . .	26
3.3	Discrepância baseada no número de pixels mal-segmentados de Zhan [Zha96] . . . . .	28
3.4	Discrepância baseada na posição dos pixels mal-segmentados de Zhan [Zha96] . . . . .	29
3.5	Nova metodologia para avaliar a qualidade de segmentação de en- velopes postais . . . . .	29
3.5.1	Métrica de classificação . . . . .	30
<b>4</b>	<b>EXPERIMENTOS E RESULTADOS</b>	<b>33</b>
4.1	Introdução . . . . .	33
4.2	Avaliação do desempenho de métodos de avaliação de qualidade da segmentação. . . . .	34
4.2.1	Introdução . . . . .	34
4.2.2	Avaliação utilizando imagens artificialmente modificadas . . .	38
4.2.2.1	Erosão . . . . .	39
4.2.2.2	Dilatação . . . . .	46
4.2.2.3	Ruído pimenta . . . . .	53
4.2.2.4	Ruído sal . . . . .	57
4.2.2.5	Ruído sal-pimenta . . . . .	61
4.2.2.6	Erosão + ruído sal-pimenta . . . . .	65
4.2.2.7	Dilatação + ruído sal-pimenta . . . . .	69
4.2.3	Análise da eficiência e da sensibilidade da métrica proposta . .	73
4.2.4	Avaliação de qualidade sobre resultados obtidos de processos reais de segmentação . . . . .	75
<b>5</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>81</b>

# Lista de Figuras

2.1	(a) Imagem de células em níveis de cinza, (b) Imagem segmentada por limiarização global contendo duas classes o fundo (branco) e as células (preto). . . . .	5
2.2	Exemplo da utilização de um <i>quadtree</i> na segmentação por divisão e fusão. . . . .	7
2.3	(a) Imagem original em níveis de cinza, (b) Imagem limiarizada utilizando o algoritmo de limiarização global de [Ots79] . . . . .	8
2.4	(a) Imagem original em níveis de cinza, (b) Imagem limiarizada com o algoritmo global de Otsu, (c) Imagem limiarizada utilizando o algoritmo adaptativo de Wellner . . . . .	9
2.5	(a) Imagem original em níveis de cinza, (b) Resultado da multilimiarização de Yen et al, (c) Classe 1, (d) Classe 2, (e) Classe 3, (f) Classe 4, (g) Classe 5 . . . . .	11
2.6	Diagrama de blocos do método de segmentação de [MFBJ03], indicando os 5 passos principais. . . . .	13
2.7	Diagrama de blocos ilustrando o algoritmo de decomposição <i>Wavelet</i> de Mallat . . . . .	14
2.8	Resultado do algoritmo de segmentação proposto. (a) Imagem original (b) $I_{CONTORNO}$ em tons de cinza; (c) Imagem $I_{FINAL}$ binária . . . . .	18
2.9	Esquema da abordagem de segmentação fractal de [EFM04]. . . . .	19
2.10	(a) Imagem carta040 da base real a ser segmentada, (b) Imagem da característica Dimensão Fractal, (c) Imagem gerada pelo K-means, (d) Imagem Final Segmentada. . . . .	24



3.1	(a)Imagem real de objetos poliédricos, (b)Imagem <i>ground truth</i> gerada artificialmente para a imagem de objetos poliédricos, (c)Imagem real de um envelope postal, (d)Imagem <i>ground truth</i> gerada artificialmente a partir do envelope postal . . . . .	27
4.1	(a) Imagem <i>ground truth</i> , (b) Pimenta(15%), (c) Sal(20%), (d) Sal-Pimenta(50%), (e) Dilatação( $EE_{cruz}$ 1 iteração), (f) Dilatação( $EE_{quadrado}$ 3 iterações), (g) Dilatação( $EE_{cruz}$ 2 iterações) + Sal-Pimenta(20%), (h) Erosão( $EE_{cruz}$ 2 iterações), (i) Erosão( $EE_{quadrado}$ 5 iterações), (j) Erosão( $EE_{cruz}$ 2 iterações) + Sal-Pimenta(35%). . . . .	37
4.2	Gráfico referente a normalização dos valores do <i>foreground</i> para Erosão utilizando EE cruz . . . . .	40
4.3	Gráfico referente a normalização dos valores do <i>background</i> para erosão utilizando EE cruz . . . . .	41
4.4	Gráfico referente a normalização dos valores gerais para erosão utilizando EE cruz . . . . .	42
4.5	Gráfico referente a normalização dos valores do <i>foreground</i> para erosão utilizando EE quadrado . . . . .	43
4.6	Gráfico referente a normalização dos valores do <i>background</i> para erosão utilizando EE quadrado . . . . .	44
4.7	Gráfico referente a normalização dos valores gerais para erosão utilizando EE quadrado . . . . .	45
4.8	Gráfico referente a normalização dos valores do <i>foreground</i> para dilatação utilizando EE cruz . . . . .	47
4.9	Gráfico referente a normalização dos valores do <i>background</i> para dilatação utilizando EE cruz . . . . .	48
4.10	Gráfico referente a normalização dos valores gerais para dilatação utilizando EE cruz . . . . .	49
4.11	Gráfico referente a normalização dos valores do <i>foreground</i> para dilatação utilizando EE quadrado . . . . .	50
4.12	Gráfico referente a normalização dos valores do <i>background</i> para dilatação utilizando EE quadrado . . . . .	51
4.13	Gráfico referente a normalização dos valores gerais para dilatação utilizando EE quadrado . . . . .	52

4.14	Gráfico referente a normalização dos valores do <i>foreground</i> para o ruído pimenta. . . . .	54
4.15	Gráfico referente a normalização dos valores do <i>background</i> para o ruído pimenta. . . . .	55
4.16	Gráfico referente a normalização dos valores gerais para o ruído pimenta.	56
4.17	Gráfico referente a normalização dos valores do <i>foreground</i> para ruído sal . . . . .	58
4.18	Gráfico referente a normalização dos valores do <i>background</i> para ruído sal . . . . .	59
4.19	Gráfico referente a normalização dos valores gerais para ruído sal . . .	60
4.20	Gráfico referente a normalização dos valores do <i>foreground</i> para ruído sal-pimenta . . . . .	62
4.21	Gráfico referente a normalização dos valores do <i>background</i> para ruído sal-pimenta . . . . .	63
4.22	Gráfico referente a normalização dos valores gerais para ruído sal-pimenta . . . . .	64
4.23	Gráfico referente a normalização dos valores do <i>foreground</i> para erosão 1 iteração + ruído sal-pimenta . . . . .	66
4.24	Gráfico referente a normalização dos valores do <i>background</i> para erosão 1 iteração + ruído sal-pimenta . . . . .	67
4.25	Gráfico referente a normalização dos valores gerais para erosão 1 iteração + ruído sal-pimenta . . . . .	68
4.26	Gráfico referente a normalização dos valores do <i>foreground</i> para dilatação 1 iteração + ruído sal-pimenta . . . . .	70
4.27	Gráfico referente a normalização dos valores do <i>background</i> para dilatação 1 iteração + ruído sal-pimenta . . . . .	71
4.28	Gráfico referente a normalização dos valores gerais para dilatação 1 iteração + ruído sal-pimenta . . . . .	72
4.29	Exemplos da segmentação do bloco de endereços: (a) Imagem original No 1 contendo selo, (b) Imagem <i>ground truth</i> , (c) Resultado da segmentação, (d) Imagem original No 2 que não contém selo, (e) Imagem <i>ground truth</i> , (f) Resultado da segmentação. . . . .	77

4.30 Resultado da segmentação de envelopes postais. (a) Imagem *ground truth* do envelope 028, (b) Imagem *ground truth* do envelope 075, (c) resultado da segmentação do envelope 028 pelo método de [MFBJ03], (d) resultado da segmentação do envelope 075 pelo método de [MFBJ03], (e) resultado da segmentação do envelope 028 pelo método de [EFM04], (f) resultado da segmentação do envelope 075 pelo método de [EFM04]. 79

# Lista de Tabelas

4.1	Resultados iniciais para ruído pimenta. . . . .	35
4.2	Resultados iniciais para ruído sal. . . . .	36
4.3	DBNPMS Multi-classe I e II - erosão ( $EE_{cruz}$ uma iterações) com % de ruído sal-pimenta . . .	38
4.4	Resultados médios para erosão utilizando EE cruz . . . . .	39
4.5	Resultados para erosão utilizando EE quadrado . . . . .	39
4.6	Resultados médios para dilatação utilizando EE cruz . . . . .	46
4.7	Resultados médios para dilatação utilizando EE quadrado . . . . .	46
4.8	Resultados médios para ruído pimenta . . . . .	53
4.9	Resultados médios para ruído sal . . . . .	57
4.10	Resultados médios para ruído sal-pimenta . . . . .	61
4.11	Resultados médios para erosão 1 iteração + ruído sal-pimenta . . . . .	65
4.12	Resultados médios para dilatação 1 iteração + ruído sal-pimenta . . . . .	69
4.13	Tabela de sensibilidade entre metodologias de avaliação e ruídos . . . . .	74
4.14	Média das medidas de avaliação de segmentação para a segmentação <i>Wavelet</i> de envelopes postais de [MFBJ03] . . . . .	75
4.15	Média das medidas de avaliação de segmentação para a segmentação de envelopes postais de [EFM04] utilizando uma abordagem fractal . . . . .	75
4.16	Média das metodologias de avaliação de segmentação para a seg- mentação <i>Wavelet</i> de envelopes postais de [MFBJ03] . . . . .	76
4.17	Média das metodologias de avaliação de segmentação para a seg- mentação de envelopes postais de [EFM04] utilizando uma abordagem fractal . . . . .	76

# Resumo

A área de Processamento e Análise de Imagens, objetiva reproduzir computacionalmente as decisões visuais tomadas pelo ser humano. O processo de segmentação, consiste em subdividir uma imagem em conjuntos semanticamente relevantes e é provavelmente uma das etapas mais críticas na análise e processamento de imagens. Na literatura existem muitos métodos de segmentação de imagens. Entretanto, nenhum algoritmo de segmentação é aplicável a todos os tipos de imagens, diversos algoritmos têm atuações diferentes em um mesmo grupo de imagens e não há uma imagem que seja bem segmentada por qualquer algoritmo de segmentação. Desta forma faz-se necessário avaliar a qualidade da segmentação utilizando conceitos objetivos de qualidade.

A seguir, uma abordagem de avaliação de discrepância de métodos de segmentação por região é proposta. Nesse trabalho, a abordagem empregada usou critérios objetivos e qualitativos para gerar avaliações quantitativas.

A abordagem proposta foi testada sobre duas bases de imagens, a primeira formada de imagens artificiais onde foram gerados vários cenários de má segmentação. E a segunda base formada por imagens obtidas por dois mecanismos reais de segmentação de envelopes postais.

Os resultados obtidos mostraram a importância do uso da base de imagens artificiais para quantificar a influência de cada método de avaliação para vários cenários de má segmentação. Conclusões e futuras pesquisas são apresentadas ao final deste trabalho.

**Palavras-chaves:** 1. Segmentação. 2. Qualidade 3. Mecanismos de avaliação. 4. Envelopes postais.

# Abstract

Image analysis usually refers to processing images by computer with the goal of finding what objects are in the image. Image segmentation is one of the most critical tasks in automatic image analysis. It consists of subdividing an image into its constituent parts and extracting the parts of interest (objects). A great variety of segmentation algorithms have been developed in the last few decades. Since none of the proposed segmentation algorithms are generally applicable to all images and different algorithms are not equally suitable for a particular application, the performance of segmentation algorithms is indispensable. It is the reason why establish certain criteria, other than human subjective ones, to evaluate the performance evaluation of segmentation algorithms is needed.

The focus of this work is the evaluation methods of region-based segmentation. It consists in comparing the results obtained by a segmentation algorithm with a reference (ground-truth) and measuring the differences (or discrepancy). Degradation mechanisms are used to compare segmentation evaluation methods onto deteriorated ground-truth segmentation images. In order to test the accuracy of quality evaluation in real segmentation, discrepancy criteria were applied on real approaches for postal envelopes.

Experiments showed the importance of using degradation mechanisms to compare segmentation evaluation methods.

**Key words:** 1. Segmentation, 2. Quality 3. Evaluation mechanisms 4. Postal envelopes.

# Capítulo 1

## INTRODUÇÃO

A área de Processamento e Análise de Imagens consiste em reproduzir computacionalmente as decisões visuais tomadas pelo ser humano. A segmentação é provavelmente uma das etapas críticas deste processo, pois "dela depende o eventual sucesso ou fracasso da análise" [GW92]. Segmentar uma imagem consiste em subdividi-la em conjuntos semanticamente relevantes e em escolher os conjuntos apropriados para computacionalmente automatizar o processo. Segundo [Fac01] "a grande dificuldade da segmentação reside no fato de não se conhecer de antemão o número e tipos de estrutura que se encontrarão na imagem. Essas estruturas são identificadas através da sua geometria, forma, topologia, textura, cor ou brilho sendo escolhidas aquelas que possibilitem a melhor identificação".

Na literatura existem muitos métodos de segmentação de imagens, envolvendo mais de 30 anos de pesquisa, entretanto pesquisas envolvendo a avaliação objetiva destes métodos de segmentação são muito recentes e carecem de mais pesquisas. Considerando que nenhum algoritmo de segmentação é aplicável a todos os tipos de imagens, que diversos algoritmos tem atuações diferentes em um mesmo grupo de imagens e que não há uma imagem que seja segmentada por qualquer algoritmo de segmentação, faz-se necessário avaliar a qualidade da segmentação utilizando conceitos objetivos de qualidade. As abordagens a serem desenvolvidas precisam empregar critérios objetivos e qualitativos para gerar

avaliações quantitativas.

## 1.1 Desafio

Os desafios para a presente pesquisa são desenvolver uma metodologia de avaliação de discrepância para metodologias de segmentação por região, no caso específico de segmentação binária, capaz de avaliar vários cenários de má segmentação, propor um mecanismo de geração destes cenários de má segmentação através da geração de ruídos e deformações em uma imagem que representa o resultado ideal de uma segmentação real e testar e comparar o novo mecanismo de avaliação de discrepância com metodologias de segmentação por região existentes na literatura, utilizando uma base de imagens obtidas através de mecanismos reais de segmentação de envelopes postais e através da base de imagens que representam os cenários de má segmentação.

## 1.2 Motivação

A segmentação de imagens é uma etapa primordial na área de processamento e análise de imagens; desta forma se faz necessário a avaliação objetiva de métodos de segmentação tendo em vista que trabalhos neste sentido são muito recentes e ainda carecem de mais pesquisas se comparado com sua importância, como pode-se observar nos trabalhos de [HBJJ<sup>+</sup>96], [Zha96] e [RLA<sup>+</sup>01].

## 1.3 Proposta

A proposta da presente pesquisa é desenvolver um mecanismo de avaliação de discrepância de metodologias de segmentação por região, testá-las e compará-las numa base de imagens sintetizadas e duas bases de imagens obtidas de mecanismos reais de segmentação de envelopes postais.



## 1.4 Contribuição

A principal contribuição deste trabalho é provar que o novo mecanismo de avaliação proposto permite uma avaliação mais apurada, de diversos cenários de má segmentação, que os mecanismos existentes na literatura. Provar que a utilização de imagens artificiais de cenários de má segmentação auxiliam na avaliação objetiva pois pode-se saber quais são os efeitos de cada perturbação e deformação presentes na segmentação na avaliação.

## 1.5 Organização

Essa dissertação está dividida em 5 Capítulos descritos da seguinte maneira:

O Capítulo 1 apresentou uma descrição sucinta dos desafios, motivação, proposta, contribuição e organização bem como da importância do estudo da segmentação de imagens de documentos, particularmente de envelopes postais brasileiros.

O Capítulo 2 trata das técnicas de segmentação por região de imagens e mais especificamente, a segmentação de imagens de envelopes postais.

O Capítulo 3 faz uma revisão dos mecanismos de avaliação da qualidade de segmentação, bem como a proposta do novo mecanismo de avaliação da qualidade de segmentação.

O Capítulo 4 descreve a metodologia e análise de resultados empregados na avaliação da qualidade de segmentação de imagens.

O Capítulo 5, por fim, apresenta as conclusões e propostas de trabalhos futuros na segmentação de envelopes postais brasileiros.



## Capítulo 2

# DESCRIÇÃO DA ÁREA DE ESTUDO

### 2.1 Introdução

Segmentar uma imagem consiste em subdividi-la em conjuntos semanticamente relevantes e em escolher os conjuntos apropriados para computacionalmente automatizar o processo. A figura 2.1 é um exemplo de segmentação por limiarização global, onde uma imagem em níveis de cinza contendo células é segmentada resultando em uma imagem binária (contendo apenas duas classes branco e preto) onde o fundo torna-se branco e as células tornam-se pretas.

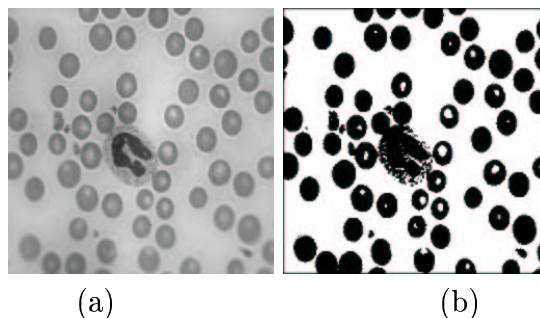


Figura 2.1: (a) Imagem de células em níveis de cinza, (b) Imagem segmentada por limiarização global contendo duas classes o fundo (branco) e as células (preto).

A segmentação é uma etapa de grande importância para a análise de imagens pois dela depende o sucesso ou fracasso da análise.

Entretanto esta não é uma tarefa fácil, pois segundo [Fac01] *a grande dificuldade da segmentação reside no fato de não se conhecer de antemão o número e tipos de estrutura que se encontram na imagem. Essas estruturas são identificadas através da sua geometria, forma, topologia, textura, cor ou brilho sendo escolhidas aquelas que possibilitem a melhor identificação.*

Na literatura existem várias metodologias diferentes para segmentar imagens. Classificam-se os diversos tipos de segmentação através do tipo de primitiva empregada. Assim sendo, as metodologias de segmentação são subdivididas em segmentação por região, segmentação por contorno, segmentação por textura. Neste trabalho apresentaremos apenas a segmentação por região.

## 2.2 Segmentação por região

### 2.2.1 Introdução

Pode definir uma região em uma imagem como sendo um conjunto de pontos ligados entre si que contenham características comuns. A principal característica para se caracterizar uma região de uma imagem é a homogeneidade. Pois segundo [Fac01] *uma região pode ser definida como sendo um conjunto de pontos que respeitam um mesmo predicado de homogeneidade. A escolha da propriedade de homogeneidade é com certeza o parâmetro decisivo para extrair "boas regiões".*

Os métodos utilizados na segmentação por região são:

### 2.2.2 Divisão e fusão ou *split and merge*

O processo de divisão e fusão consiste basicamente em agrupar os pixels em regiões, onde o critério de homogeneidade é verificado. Esse processo é feito através de inúmeras divisões e fusões, manipulando a

imagem até que estas regiões obedecem o critério de homogeneidade como pode ser encontrado em [CMVM86]. Uma forma de decompor e agrupar partes da imagem é utilizar uma estrutura quadtree como pode ser encontrado em [CA84]. Um exemplo de como a segmentação por divisão e fusão é feita, utilizando um quadtree, pode ser visto na figura 2.2.

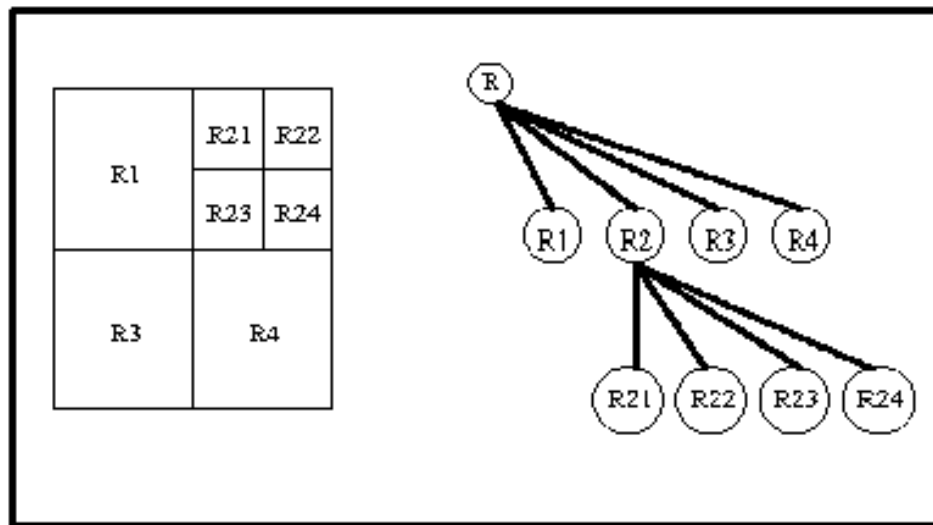


Figura 2.2: Exemplo da utilização de um *quadtree* na segmentação por divisão e fusão.

### 2.2.3 Limiarização, binarização.

O processo de limiarização consiste em separar os objetos de interesse, ou objetos do primeiro plano (*foreground*), do fundo (*background*). O objetivo é marcar os objetos do primeiro plano como pixels pretos e o fundo como pixels branco ou vice-versa, gerando assim apenas duas classes de elementos distintos, separando assim o fundo dos objetos de interesse. A limiarização é um dos processos de segmentação por região mais utilizados por ser um princípio muito simples. A seguir apresentaremos as três abordagens de limiarização que são: global, adaptativa ou

local e multilimiarização.

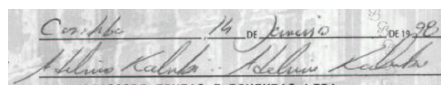
- Limiarização global.

A limiarização utiliza o histograma da imagem como base do processo. Uma imagem com o primeiro plano e o fundo bem definidos gera um histograma com dois picos e um vale entre eles, onde cada pico representa uma classe diferente. Desta forma a separação entre as classe será feita através do limiar  $T$  que encontra-se em alguma região do vale. No caso de um primeiro plano mais claro que o fundo, escolher no histograma um valor de limiar  $T$ , tal que para cada pixel  $P(x, y)$ :

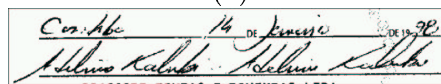
$$P(x, y) \in \begin{cases} 0 & \text{se o valor do nível de cinza do pixel } P(x, y) > T \\ 255 & \text{se o valor do nível de cinza do pixel } P(x, y) \leq T \end{cases} \quad (2.1)$$

Pode-se encontrar, devido ao grande número de pesquisas realizadas, inúmeras publicações sobre técnicas de limiarização global. As técnicas que se destacam por apresentarem eficiência em imagens variadas e complexas são as de [Ots79] e [LL93].

A figura 2.3 é um exemplo de limiarização global utilizando o algoritmo de [Ots79].



(a)



(b)

Figura 2.3: (a) Imagem original em níveis de cinza, (b) Imagem limiarizada utilizando o algoritmo de limiarização global de [Ots79]

- Limiarização adaptativa ou local.

Nem sempre serão trabalhadas imagens com o primeiro plano e o fundo bem definidos, devido a problema de iluminação e falta de qualidade da imagem, gerando um histograma com vários picos e vales, necessitando assim de diferentes valores de limiar para diferentes regiões da imagem. Uma forma de determinar estes vários limiares para diferentes regiões seria utilizar as técnicas de limiarização global em uma janela para percorrer toda a imagem, gerando assim valores de limiares diferentes de acordo com a região da imagem. Entretanto é preciso descobrir um tamanho de janela que possa ser grande o suficiente para garantir que o número de pixels do fundo seja capaz de obter uma boa estimativa do valor médio, mas não tão grande como a média sobre as intensidades de segundo plano não uniformes. Uma forma de contornar este problema é ter o tamanho da janela variável, sendo automaticamente ajustado dependendo da aplicação.

A literatura fornece várias abordagens como os métodos de [Ber86], o método de [ETM91] e o método de [Wel93], que consistem nos métodos mais interessantes atualmente..

A figura 2.4 demonstra a diferença de resultado obtido entre um algoritmo global de Otsu [Ots79] e um algoritmo local ou adaptativo de Wellner [Wel93] para uma imagem complexa.

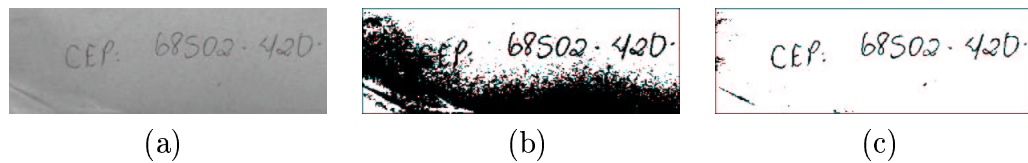


Figura 2.4: (a)Imagem original em níveis de cinza, (b)Imagem limiarizada com o algoritmo global de Otsu, (c)Imagem limiarizada utilizando o algoritmo adaptativo de Wellner

- Multilimiarização.

Dependendo da complexidade de uma imagem, a limiarização desta imagem resultará em apenas duas classes, o *foreground* e o *background*, podendo desta forma perder informações relevantes à análise desta imagem. Uma alternativa para este problema é a multilimiarização, que consiste em separar a imagem em várias classes distintas. O problema desta abordagem consiste em se saber qual é o número de classes necessárias para particionar a imagem. Existem vários métodos de multilimiarização, porém muitos destes métodos não prevêm automaticamente o número necessário de classes, dificultando a utilização dos mesmos. Entretanto existem algumas abordagens que processam automaticamente o número de classes necessárias para a multilimiarização, como é o caso do método proposto por Yen et al [YCC95].

A figura 2.5 demonstra a utilização do algoritmo de multilimiarização de Yen et al [YCC95] onde submetemos a imagem Lenna ao processo obtendo cinco classes distintas, onde estas classes representam vários níveis intermediários de limiar.



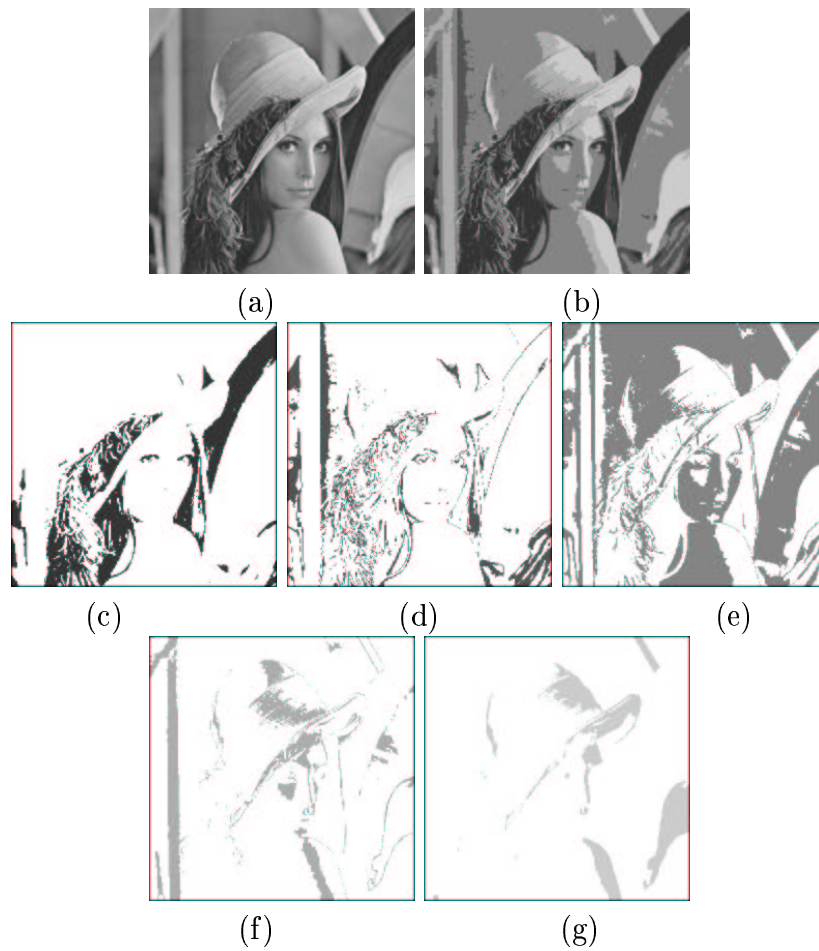


Figura 2.5: (a) Imagem original em níveis de cinza, (b) Resultado da multi-limiarização de Yen at al, (c) Classe 1, (d) Classe 2, (e) Classe 3, (f) Classe 4, (g) Classe 5

## 2.2.4 Metodologias de segmentação por região de envelopes postais

Serão apresentadas duas metodologias de segmentação por região, para envelopes postais. O resultado da segmentação destas metodologias será avaliado posteriormente pelas metodologias de avaliação de qualidade da segmentação implementadas.

A primeira metodologia é a segmentação *Wavelet* de envelopes postais de [MFBJ03] e a segunda metodologia é a segmentação de envelopes postais de Eiterer et al [EFM04] utilizando uma abordagem fractal.

## 2.2.5 Segmentação *Wavelet* de envelopes postais de [MFBJ03]

Existem várias abordagens de segmentação de texturas usando o espaço *Wavelet*. Aqui será descrita uma abordagem direcionada à segmentação de envelopes postais.

O trabalho de segmentação de [MFBJ03] busca selecionar características no espaço *Wavelet* para segmentar envelopes postais e localizar automaticamente o bloco de endereços. O método de é dividido em cinco passos principais:

1. Decomposição da imagem de envelope postal no espaço *Wavelet* utilizando o algoritmo de decomposição de Mallat [Mal89];
2. Identificação dos pontos salientes, baseada na intersecção dos coeficientes mais significativos dos canais de alta frequência (horizontal e vertical);
3. Controle da rotulação das janelas (de tamanho  $k$  e  $k^2$  elementos) de pontos salientes, para eliminar o ruído e separar o fundo do envelope, a partir de um teste de hipótese estatístico para um conjunto de janelas;
4. Projeção reversa dos pontos salientes das janelas selecionados na imagem de níveis de cinza original;
5. Perseguição de contorno através dos pixels referenciados pelos pontos salientes, baseada em teste de hipótese estatístico.

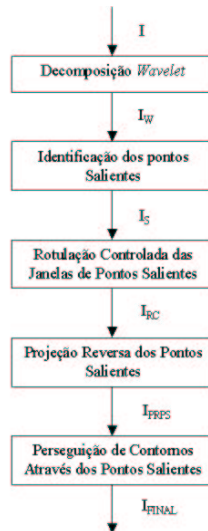


Figura 2.6: Diagrama de blocos do método de segmentação de [MFBJ03], indicando os 5 passos principais.

### 2.2.5.1 Decomposição *Wavelet*

A transformada *Wavelet* decompõe dados em blocos de construção fundamentais. Visto que é possível desenvolver decomposições *Wavelet* com uma grande variedade de funções básicas, e também enfatizar a redundância ou eliminá-la através dos níveis de decomposição. Existe como pode ser visto em [SN96] uma farta literatura de diferentes técnicas úteis para decomposição *Wavelet*. Como existe a necessidade de separar automaticamente o bloco de endereços, selos, carimbo e o fundo, a decomposição deve localizar descontinuidades na imagem que provavelmente sejam estes elementos. Para isso a decomposição de Mallat [Mal89] é usada. A decomposição de Mallat [Mal89] produz como saída quatro conjuntos relacionados à imagem original, onde uma saída é relacionada à banda de baixa frequência (*BB*) ou coeficientes de aproximação e as outras três saídas, na horizontal (*BA*), vertical (*AB*) e diagonal (*AA*), são relacionadas às bandas de alta frequência ou coeficientes de detalhe. A função de base Haar é usada neste caso por possibilitar a localização

de descontinuidades em imagens e por ser o filtro de menor comprimento conhecido, diminuindo a complexidade algorítmica da decomposição. A figura 2.7 representa o diagrama de blocos do algoritmo em questão.

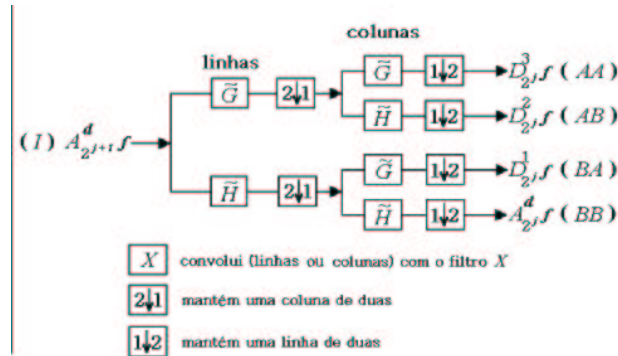


Figura 2.7: Diagrama de blocos ilustrando o algoritmo de decomposição *Wavelet* de Mallat

### 2.2.5.2 Identificação dos pontos salientes

Para identificar os pontos salientes, é necessário encontrar evidências, através das bordas para os objetos que são o foco da segmentação (bloco endereço, carimbo e selo) e também para a presença de ruídos. Uma imagem típica de envelope contém vários componentes que são de interesse deste algoritmo de segmentação. Todos estes componentes - bloco endereço, selos e carimbos - estão envolvidos pelo fundo do envelope, ou seja, as regiões do fundo do envelope e dos objetos de segmentação fazem fronteira. Estas regiões de transição são fortes evidências da existência de objetos de segmentação. A transformada *Wavelet*, através de suas bandas direcionais obtidas pela decomposição de Mallat [Mal89], pode representar estas zonas de transição. As zonas de transição com forte representação são denominadas de pontos salientes.

O objetivo neste passo é identificar evidências através de bordas de regiões mais consistentes que são parecidas com o fundo do envelope

ou as regiões de interesse (bloco de endereço, selos e carimbos). Então, define-se o conjunto de pontos salientes  $I_s$  como sendo o conjunto de pontos que têm forte evidência de ser um detalhe. Mesmo que estes pontos possam ser somente simples manchas, haverá outros passos, mais adiante, para verificar a consistência (eliminar ruídos) e incluir mais evidências.

O procedimento de identificação de pontos salientes pode ser resumido segundo a Equação 2.2.

$$I_s \leftarrow I_{WCS(LH)} \cap I_{(HL)} \quad (2.2)$$

### 2.2.5.3 Rotulação controlada das janelas de pontos salientes

O conjunto de pontos salientes é evidência de textura esparsa de componentes conectados. Alguns pontos salientes devem aparecer em uma determinada região com uma certa distribuição diferente da distribuição de outras regiões, por exemplo, uma região com alta densidade de pontos salientes presentes (medidas em uma janela quadrada de tamanho  $k$ ), e uma outra região com baixa densidade. As regiões com alta densidade de pontos salientes são mais prováveis de serem regiões conectadas e portanto objetos de segmentação, enquanto que as outras regiões, de baixa concentração, passam a ser consideradas como ruído. A decisão de estabelecer um limite entre estas regiões de alta densidade e de baixa densidade, não pode ser feita de modo fixo, por um limiar, visto que para cada imagem de envelope os pixels da imagem e suas distribuições mudam muito. Para resolver este problema foi desenvolvido um algoritmo iterativo de controle para realizar testes de significância estatística [PTVF96] sobre as janelas de pontos salientes. O algoritmo de controle desenvolvido executa os testes de hipótese de forma global, isto é, o algoritmo não leva em consideração a posição das janelas dentro da imagem.

Estas duas regiões são determinadas de forma isolada, isto é, uma região não interfere na construção da outra. O teste de hipótese é utilizado para verificar quais janelas pertencem a uma dada classe/região

(de alta ou baixa densidade) em cada iteração. O algoritmo começa a execução em ambos os níveis, da base, pelas regiões/janelas de baixa densidade de pontos salientes e do topo, pela região/janela de maior densidade de pontos salientes, ao mesmo tempo em que vai decidindo os rótulos das janelas/regiões gerando assim o novo conjunto com as janelas rotuladas como de alta densidade e descartando as janelas/regiões rotuladas como sendo de baixa densidade.

#### 2.2.5.4 Projeção reversa dos pontos salientes selecionados

O conjunto  $I_{RC}$  é uma imagem de  $n/2k \times m/2k$  representando janelas que contém os pontos salientes que pertencem a uma região dita de alta densidade de pontos salientes. Cada ponto saliente, tem quatro filhos relacionados à imagem original. Uma janela de tamanho  $k$  possui  $k^2$  elementos relacionados à imagem de intersecção  $I_s$ . Cada um destes elementos está relacionado com quatro filhos da imagem original em tons de cinza. Todavia, somente os pontos salientes identificados no conjunto  $I_s$  e rotulados pelo conjunto  $I_{RC}$  serão utilizados para a projeção reversa. A projeção reversa destes quatro filhos de cada ponto saliente origina o conjunto  $I_{PRPS}$  com dimensão igual à imagem original. Este conjunto é composto pelos pixels com o nível de cinza da imagem original. Os pixels que não são originários de pontos salientes rotulados como de alta densidade são inicializados como não pertencentes ao conjunto  $I_{PRPS}$ .

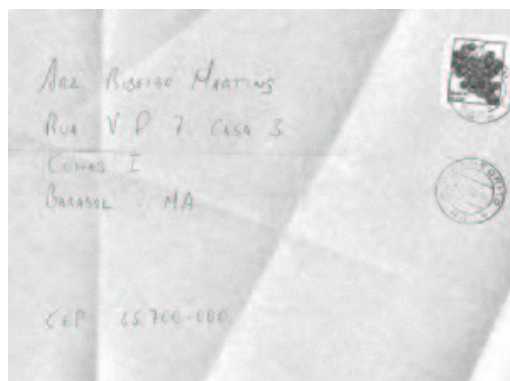
#### 2.2.5.5 Perseguição de contorno através dos pontos salientes

Neste passo a imagem  $I_{PRPS}$  deve conter os pixels que provavelmente pertençam a uma das classes – bloco de endereços, carimbo ou selos. A grande maioria do fundo já está selecionada pelo complemento de  $I_{PRPS}$ , tendo como universo a imagem original. Desta forma, esta evidência deve ser usada apropriadamente para encontrar no resto da imagem somente os pixels em tons de cinza que são coerentes à imagem  $I_{PRPS}$ , bem como a idéia de que regiões locais devem compartilhar atributos similares. Este é o estágio final da segmentação e duas hipóteses são estabelecidas para a tomada de decisão de quais pontos devem ser selecionados.

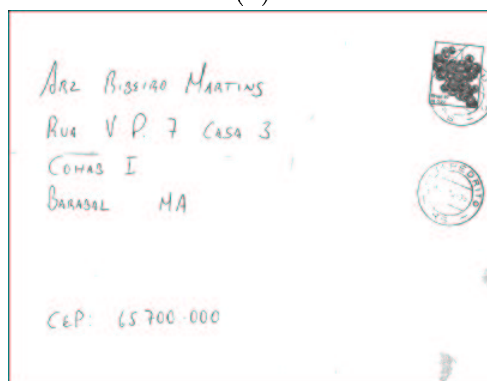
- Uma informação de contexto muito forte; considera-se que os objetos de segmentação (bloco endereço, carimbos e selos) têm uma intensidade maior que o fundo do envelope postal, ou seja, os tons de cinza dos objetos de segmentação estão mais próximos do preto que o fundo. Esta é uma hipótese que leva em conta a distribuição dos tons de cinza de uma forma global. Para isto, faz-se necessário supor uma distribuição para a imagem do envelope postal.
- Supõe-se que os pontos salientes representem regiões (pixels) de transição, isto é, possuem um alto grau de energia (valor absoluto) no sentido dos coeficientes *Wavelet* de alta frequência (*AB* e *BA*), e são a transição entre as regiões de objeto de segmentação e o fundo.

Baseando-se nestas duas hipóteses, foi desenvolvido o passo final de segmentação. Este tem início com a imagem  $I_{PRPS}$ , de cada um dos  $2 \times 2$  pixels ( $I_{SET}(i, j), I_{SET}(i + 1, j), I_{SET}(i, j + 1), I_{SET}(i + 1, j + 1)$ ) referenciados pelos pontos salientes. O pixel com o menor tom de cinza - mais próximo ao objeto de segmentação - deve ser selecionado, como pertencente ao contorno do objeto de segmentação, gerando assim a imagem  $I_{CONTORNO}$ .

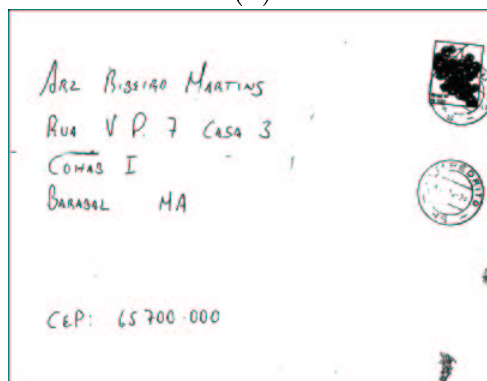
Após executado o algoritmo de perseguição de contorno, o conjunto  $I_{CONTORNO}$ , contém os tons de cinza de todos os pixels que foram segmentados 2.8. Entretanto, esta não é a representação recomendada para uma segmentação. Então, todos os pixels que foram rotulados/classificados/segmentados como bloco endereço, selo, ou carimbo são transformados em pixels pretos de uma imagem binária. Gerando, desta forma, a imagem de saída do algoritmo de segmentação proposto  $I_{FINAL}$  2.8.



(a)



(b)



(c)

Figura 2.8: Resultado do algoritmo de segmentação proposto. (a) Imagem original (b)  $I_{CONTORNO}$  em tons de cinza; (c) Imagem  $I_{FINAL}$  binária



## 2.2.6 Segmentação de envelopes postais de Eiterer et al [EFM04] utilizando uma abordagem fractal

Existem poucas abordagens de segmentação de texturas usando o paradigma de dimensão fractal. Aqui será descrita uma abordagem direcionada à segmentação de envelopes postais.

O trabalho de segmentação de Eiterer et al [EFM04] propõe segmentar envelopes postais baseado num mecanismo de classificação por *clustering* da dimensão fractal dos diversos conjuntos de informações semânticas relevantes de um envelope postal (bloco-endereço, carimbo(s) e selo(s)), sem conhecimento a priori destes componentes, como pode ser visto na figura 2.9.

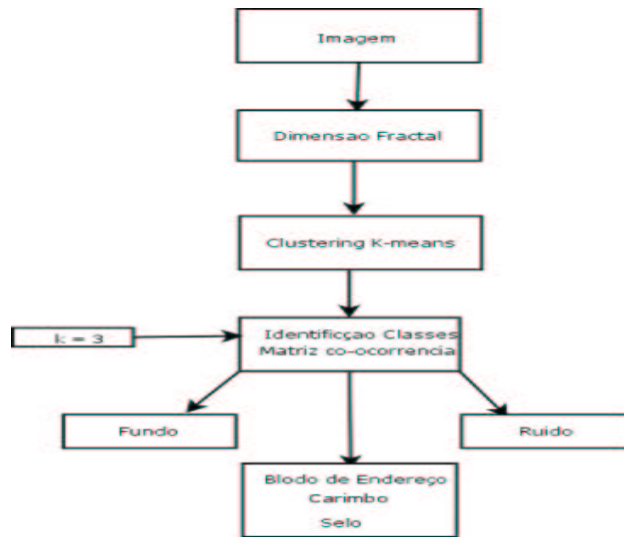


Figura 2.9: Esquema da abordagem de segmentação fractal de [EFM04].

### 2.2.6.1 Dimensão fractal para imagens em níveis de cinza

Na literatura, existem alguns métodos para se calcular o  $FD$  para imagens em níveis de cinza. Provavelmente o método mais popular é o *box-counting* de [CSK93].

O *2D variation procedure*, proposto por [DQRC<sup>+</sup>89] [Cla86], considerado um método do tipo *box-counting*, analisa o ambiente de pixels

para diferentes distâncias  $r$ . Os quadrados com diferentes tamanhos  $r$  são executados pixel a pixel da esquerda para a direita e de cima para baixo. O algoritmo determina o mínimo  $min$  e o máximo  $max$  valor de cinza dentro do quadrado de tamanho  $r$ . Desde o valor da intensidade  $v$  de um pixel  $p$  é relacionado com a altura dos objetos na imagem, pode-se de uma maneira significativa definir a  $p$  um volume  $r \times r \times v$ . Definindo o volume de diferença de  $p$  como  $r \times r \times (max - min)$ , pode-se denotar  $V(r)$  como a soma das diferenças entre o máximo e o mínimo para a escala  $r$  para toda a imagem, resultando em:  $V(r) = const r^s$ . No mapeamento de Richardson-Mandelbrot ( $\log(V(r))$  versus  $\log(r)$ ), a dependência deste volume  $V(r)$  deve ser linear com o quadrado de tamanho  $r$ . Usando o método dos mínimos quadrados nessa regressão linear, pode-se computar o declive  $s$  e deduzir  $FD$ :

$$FD = 3 - s/2 \quad (2.3)$$

Os autores constataram que a técnica de calculo da dimensão fractal *2D variation procedure* é menos complexa do que outras disponíveis (por exemplo, as coberturas morfológicas) e tem mostrado ser promissora em imagens complexas (ver [Rin03]) e por estes motivos foi a escolhida. Os autores aplicaram o método *2D variation procedure* em imagens de envelopes postais da base de imagens reais para várias faixas de tamanho de janela: ( $r = \{3, 5\}$ ,  $r = \{3, 5, 7, \}$ ,  $r = \{3, 5, 7, 9\}$ ,  $r = \{3, 5, 7, 9, 11\}$ ,  $r = \{3, 5, 7, 9, 11, 13\}$ ,  $r = \{3, 5, 7, 9, 11, 13, 15\}$ ) ; Os resultados da  $FD$  para cada pixel num intervalo de [2.0, 3.0] (ver equação 2.3).

### 2.2.6.2 Clustering

O mecanismo de *clustering* pode ser visto como um problema de reconhecimento de padrões não supervisionado. Para agrupar automaticamente as classes do bloco de endereços, selos e carimbos que foram extraídos através da dimensão fractal, foi utilizado o *K-means* que é um algoritmo não supervisionado robusto, utilizado em problemas de segmentação de imagens de texturas. O *K-means* é baseado na minimização do índice de performance que é definido como a soma das distâncias

quadráticas de todos os pontos do cluster em relação a seu centro. Este procedimento consiste dos seguintes passos:

Passo 1: escolher  $k$  centros iniciais dos clusters  $z_1(1), z_2(1), \dots, z_k(1)$ . Esta seleção é arbitrária e usualmente seleciona aleatoriamente os centros iniciais;

Passo 2: a  $k$  – ésima iteração distribui as amostras  $\{x\}$  entre os  $k$  clusters, usando a relação:

$$X \in S_j(k) \quad \text{se} \quad \|x - z_j(k)\| < \|x - z_i(k)\| \quad (2.4)$$

para  $i = 1, 2, \dots, k$  e  $i \neq j$ , onde  $S_j(k)$  denota o conjunto de amostras cujos centros de clusters é  $z_j(k)$ .

Passo 3: do resultado do passo 2, calcule-se os novos centros dos clusters  $z_j(k+1)$ ,  $j = 1, 2, \dots, k$  tal como a soma das distâncias quadráticas de todos os pontos em  $S_j(k)$  em relação ao novo centro do cluster seja minimizada. Dito de outra maneira, o novo centro do cluster  $z_j(k+1)$  é calculado de modo que o índice de performance seja minimizado.

$$J_j = \sum_{x \in S_j(k)} \|x - z_j(k+1)\|^2, j = 1, 2, \dots, k \quad (2.5)$$

O  $z_j(k+1)$  o qual minimiza este índice de performance é simplesmente a média das amostras  $S_j(k)$ . Conseqüentemente, o novo centro do cluster é dado por:

$$z_j(k+1) = \frac{1}{N_j} \sum_{x \in S_j(k)} x, j = 1, 2, \dots, k \quad (2.6)$$

onde  $N_j$  é o número de amostras em  $S_j(k)$ . O nome *K-means* é obviamente derivado da maneira com a qual os centros dos clusters são seqüencialmente atualizados.

Passo 4: se  $z_j(k+1) = z_j(k)$  para  $j = 1, 2, \dots, k$  o algoritmo convergiu e o procedimento é finalizado.

O comportamento do algoritmo é influenciado pelo número de centros de *clusters* especificados, pela escolha inicial dos centros, pela ordem de escolha das amostras, e naturalmente pelas propriedades geométricas dos dados.

Os parâmetros iniciais (padrões) usados no *K-means* pelos autores foram:

- número de características igual a 1, a própria *FD*;
- número de clusters igual a três, que são o fundo, o ruído e o bloco de endereço, selo e carimbo;
- número de iterações igual a 10;
- valor de melhoria mínima igual a 0,0001.

O algoritmo *k-means* varia de acordo com as condições iniciais de processamento, isto é, bem conhecido. Nesta abordagem os autores usaram *min*,  $(max \times min)/2$  e *max*, para inicializar o valor dos centróides dos clusters, onde *max* e *min* são respectivamente os valores da máxima e da mínima *FD* do conjunto. A convergência do algoritmo *K-means* em uma imagem  $I_{FD}$  resulta em uma imagem rotulada  $I_{k-means}$  onde cada pixel é rotulado em um dos três clusters.

### 2.2.6.3 Identificação das classes

Para segmentar o *cluster* do bloco de endereço, o carimbo e os selos é necessário sua identificação entre os três *clusters* de  $I_{K-means}$ . Primeiramente, a medida de distribuição espacial de cada classe na imagem rotulada é realizada por meio de uma matriz de co-ocorrência. Neste caso, a matriz de co-ocorrência de  $I_{K-means}$  é uma matriz  $C = [c_{ij}]$  de dimensão  $3 \times 3$  que contém a informação estatística a respeito do número de ocorrências do conjunto de dois pixels separados por um deslocamento  $(dx ; dy)$ . A matriz de co-ocorrência  $C = [c_{ij}(dx, dy)]$  de qualquer par de pixels  $(x, y)$  e  $(x + dx; y + dy)$  representando respectivamente os conjuntos  $(i, j)$  pode ser mostrado a seguir :

$$c_{ij}(dx, dy) = \#\{(x, y) / I_{Kmeans}(x, y) = i,$$

$$I_{Kmeans}(x + d_x, y + d_y) = j \} \quad (2.7)$$

onde  $\#$  é a cardinalidade e  $i, j \in \{1, 2, 3\}$ .

Para uma melhor medida de interação entre as três classes, a matriz de co-ocorrência é construída pela união das ocorrências dos pixels da vizinhança na direção horizontal, vertical e diagonal, respectivamente  $C_{hor}$ ,  $C_{ver}$  e  $C_{diag}$ . A matriz final de co-ocorrência  $C = [c_{ij}]$  usada em nossa abordagem é

$$C = C_{hor} + C_{ver} + C_{diag} = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{bmatrix} \quad (2.8)$$

Para cada cluster  $k$ , um acumulador  $Acc_k$  é definido como a seguir:

$$Acc_k = \sum_{i=1}^3 c_{i,k} + \sum_{j=1}^3 c_{k,j}, \text{ onde } 1 \leq k \leq 3 \quad (2.9)$$

O cluster do fundo do envelope é identificado pela seleção do cluster  $k$  o qual  $Acc_k$  é máximo:

$$k_{fundo} = Arg_k \max (Acc_k) \quad (2.10)$$

Depois desta identificação, pode se identificar o cluster do bloco de endereço como o *cluster* com a menor interação com o *cluster* do fundo e com mais do que 1,5% de pixels de toda a imagem  $I$  :

$$k_{blocoendereço} = Arg_k \min [c_{k,k_{fundo}} + c_{k_{fundo},k}] \quad (2.11)$$

onde  $1 \leq k \leq 3$  e  $k \neq k_{fundo}$  e  $N_l > 1.5\% \times N_I$  é o numero de pixels que pertence ao cluster  $l$ , e  $N_I$  é o total de numero de pixels na imagem original  $I$ . O fator 1,5% foi obtido depois de estudos empíricos dos autores das imagens dos envelopes. Quer dizer se o cluster é muito pequeno, provavelmente é um cluster de ruído.

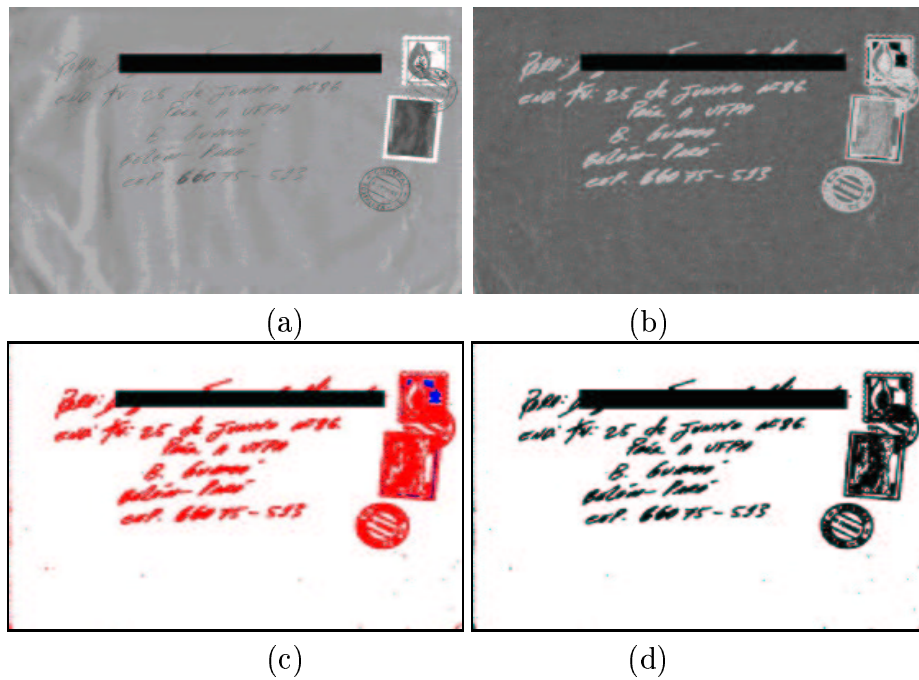


Figura 2.10: (a) Imagem carta040 da base real a ser segmentada, (b) Imagem da característica Dimensão Fractal, (c) Imagem gerada pelo K-means, (d) Imagem Final Segmentada.

## Capítulo 3

# METODOLOGIAS DE AVALIAÇÃO DA QUALIDADE DA SEGMENTAÇÃO

### 3.1 Introdução

Como visto no capítulo 2, existem diferentes métodos de segmentação de imagens, que utilizam as mais variadas técnicas o que demonstra sua importância e também sua complexidade. Da mesma forma, avaliar os resultados de segmentação de imagens tem sua importância e sua complexidade, pois não existe um método de segmentação simples que segmente bem todos os tipos de imagens, também não há um grupo de imagens que seja bem segmentada por qualquer método de segmentação e não existe uma única forma de avaliar a qualidade da segmentação.

De acordo com [Zha96], algoritmos de segmentação podem ser avaliados analiticamente ou empiricamente:

- avaliar analiticamente consiste em analisar as propriedades do algoritmo de segmentação;
- avaliar empiricamente consiste em aplicar o algoritmo de segmentação de imagens e medir a qualidade através de resultados de segmentação obtidos;

Em geral, avaliação empírica é mais satisfatória que a avaliação analítica.

[Zha96] classifica os métodos empíricos em:

- avaliação empírica de qualidade *goodness* (*AEQ*): métodos (*AEQ*) podem transformar propriedades desejáveis intuitivamente em "parâmetros de bondade";
- avaliação empírica de discrepância (*AED*): são comparados os resultados obtidos por um algoritmo de segmentação em comparação com uma imagem de referência *ground truth* e as diferenças (ou discrepância) são avaliadas e medidas.

Vários autores [HJJ+96] [SS01] [Zha96] [YALG95] [Voo95] trataram em seus trabalhos o problema de medir e avaliar a qualidade de segmentação de imagem propondo critérios para a avaliação da qualidade.

Entretanto não existe um padrão bem definido de avaliação de qualidade de segmentação. Geralmente esses critérios são para um tipo específico de aplicação e/ou imagens. Alguns destes trabalhos permitem a comparação de procedimentos.

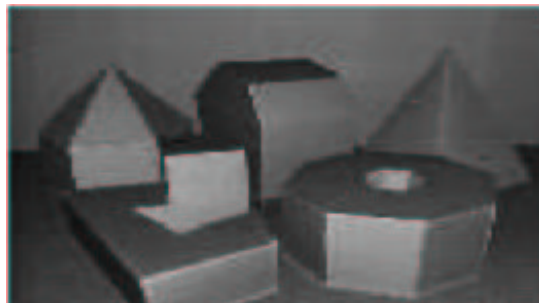
Serão apresentadas a seguir algumas abordagens de avaliação de qualidade de segmentação de imagens. Estas abordagens utilizam imagens *ground truth*, como meio de comparação dos resultados.

## 3.2 Imagens *ground truth*

Uma imagem *ground truth* é uma imagem que contém o resultado ideal do processo de segmentação de imagens mas que foi modificada para parecer com uma imagem real. A imagem *ground truth* é a referência de qualidade no resultado da segmentação. As imagens *ground truth* são usadas para a avaliação da segmentação por ser conhecido, à priori, o resultado esperado da segmentação da mesma. Uma imagem *ground truth* geralmente é criada manualmente por um especialista humano devido a grande dificuldade de automação deste processo.

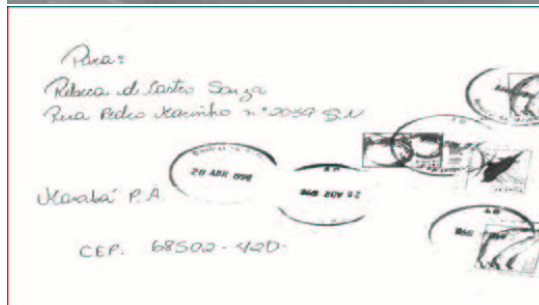
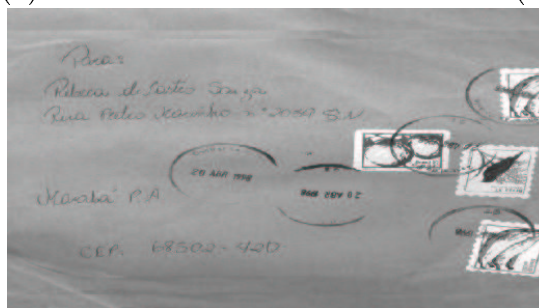
A figura 3.1 demonstra alguns exemplos de imagens *ground truth*.





(a)

(b)



(c)

(d)

Figura 3.1: (a)Imagem real de objetos poliédricos, (b)Imagem *ground truth* gerada artificialmente para a imagem de objetos poliédricos, (c)Imagem real de um envelope postal, (d)Imagem *ground truth* gerada artificialmente a partir do envelope postal

### 3.3 Discrepância baseada no número de pixels mal-segmentados de Zhan [Zha96]

Considerando a segmentação de imagens como um processo de classificação de pixels, a porcentagem de pixels mal classificados é a medida de discrepância que prontamente vem a mente Yasnoff [YMB77]. Supondo que uma imagem consiste de  $N$  classes de pixels, uma matriz de confusão  $C$  com dimensão  $N$  pode se construída, onde cada entrada representa o numero de pixels da classes  $j$  classificados como pixels da classe  $i$  pelo algoritmo de segmentação. Dois tipos de erro podem assim ser computados para cada pixel da classe  $k$ , o qual ambos podem ser usados para descrever a performance de classe-por-classe deste algoritmo Yasnoff [YMB77]. O erro multi-classe I pode ser definido como:

$$M_I^{(k)} = 100 \times \left[ \left( \sum_{i=1}^N C_{ik} \right) - C_{kk} \right] / \left[ \sum_{i=1}^N C_{ik} \right] \quad (3.1)$$

onde o numerador representa o número de pixels da classe  $k$  que não foram classificados como  $k$  e o denominador é o número total de pixels da classe  $k$ . O erro multi-classe II pode ser definido como:

$$M_{II}^{(k)} = 100 \times \left[ \left( \sum_{i=1}^N C_{ki} \right) - C_{kk} \right] / \left[ \left( \sum_{i=1}^N \sum_{j=1}^N C_{ij} \right) \right] \quad (3.2)$$

onde o numerador representa o número de pixels de outras classes que foram denominados como sendo da classe  $k$ . O denominador é o número de pixels total das outras classes.

Nas equações erro multi-classe I e II, cada classe de pixel é igualmente importante. Mesmo sendo uma abordagem simples, a avaliação de discrepância baseada na posição dos pixels mal segmentados, serve como "ponto de partida" para a avaliação de qualidade da segmentação.

### 3.4 Discrepância baseada na posição dos pixels mal-segmentados de Zhan [Zha96]

Mensurar a discrepância baseada somente no número de pixels mal-segmentados não leva em conta a informação espacial destes pixels. Assim é possível que imagens diferentemente segmentadas tenham os mesmos valores de medida de discrepância, se contarmos somente o número de pixels mal-segmentados. Para evitar este problema, uma medida de discrepância baseada na posição dos pixels mal-segmentados foi proposta. Um caminho é usar a distância entre os pixels mal-segmentados e o pixel mais próximo que pertencem de fato a mesma classe de pixels mal-segmentados. Seja  $N$  o número de pixels mal-segmentados de toda a imagem e  $d(i)$  seja a distância métrica do  $i$ -ésimo pixels mal-segmentado e o pixel mais próximo de fato sejam da mesma classe mal-classificada; uma medida de discrepância  $D$  baseada nesta distância é definida por Yasnoff em [YMB77].

$$D = \sum_{i=1}^N d^2(i) \quad (3.3)$$

Na equação (11), a cada distância é ao quadrado. Esta métrica  $ND$  é normalizada mais adiante, isenta a influência do tamanho da imagem e dá uma gama satisfatória de valores.

$$ND = 100 \times \sqrt{D}/A \quad (3.4)$$

Onde  $A$  é o numero total de pixels na imagem (i.e. uma métrica de área).

### 3.5 Nova metodologia para avaliar a qualidade de segmentação de envelopes postais

A revisão de abordagens de avaliação de algoritmos de segmentação mostrou que, basicamente, avalia-se um algoritmo de segmentação por

avaliação empírica de qualidade (*AEQ*) e por avaliação empírica de discrepância (*AED*), sendo o segundo grupo o mais interessante [Zha96].

Por serem imagens complexas, segmentar imagens de envelopes postais Brasileiros é uma tarefa árdua. Conseqüentemente, avaliar a eficiência de métodos de segmentação para tais imagens exige rigor e objetividade.

Para atingir tal objetivo, propõe-se basear a metodologia numa avaliação empírica de discrepância (*AED*) empregando imagens *ground truth* de envelopes postais.

### 3.5.1 Métrica de classificação

Seja  $A$  um algoritmo de segmentação a ser avaliado. Seja  $G_i, i = 1, 2, \dots, G$ , onde  $G$  é o número total de regiões *ground truth* de uma imagem e  $S_j, j = 1, 2, \dots, S$ , onde  $S$  é o número total de regiões obtidas pelo algoritmo de segmentação  $A$ . Seja  $n_{G_i}$  o número de pixels de uma região  $G_i$  e  $n_{S_j}$  o número de pixels de uma região segmentada  $S_j$ . Seja ainda  $m_{ij} = n_{G_i} \cap n_{S_j}$  o número de pixels que ocupam as mesmas posições nas regiões  $G_i$  e  $S_j$ . A medida de discrepância  $D_i$  é definida para cada região  $G_i$ , tem-se:

- segmentação correta: uma única região  $S_j$  segmentou uma única região  $G_i$ . A medida  $D_i$  pode ser definida como sendo  $D_i = m_{ij}$ . No caso de sobreposição completa da região,  $D_i = m_{ij} = n_{G_i} = n_{S_j}$ ;
- segmentação excessiva: há uma fragmentação de uma única região  $G_i$  em um conjunto de regiões  $S_j$ . A medida  $D_i = m_{ij}/s$ ,  $s$  sendo o número total de regiões  $S_j$ ;
- segmentação insuficiente: há uma fusão de  $g$  regiões *ground truth*  $G_i$  em uma única região  $S_j$ . A medida  $D_i = m_{ij}/g$ , sendo  $g$  o número total de regiões  $G_i$ ;
- segmentação ausente: onde uma região  $G_i$  não foi segmentada. Assim uma penalidade foi definida, neste caso  $D_i = -n_{G_i}$ .

Finalmente define-se uma métrica geral de avaliação  $\Upsilon(A)$  do algoritmo  $A$  de segmentação como sendo:

$$\Upsilon(A) = \frac{\sum_{i=1}^n D_i}{\sum_{i=1}^n G_i} \quad (3.5)$$

Percebe-se que a métrica de qualidade de segmentação  $\Upsilon(A)$  do algoritmo  $A$  em envelopes postais tem as seguintes características:

- $\Upsilon(A) = -1$  quando a segmentação é totalmente ausente (o algoritmo  $A$  "ignora" todas as regiões *ground truth*);
- $\Upsilon(A) = 0$  quando o número de pixels corretamente ou excessivamente ou ainda insuficientemente segmentados iguala o número de pixels "esquecidos";
- $\Upsilon(A) = 1$  quando o algoritmo  $A$  segmenta todas e somente as regiões *ground truth*  $G_i$  como o número exato de pixels para cada uma delas;
- Em caso de segmentação excessiva ou insuficiente,  $\Upsilon(A) \leq 1$ ;
- Logo tem-se que  $-1 \leq \Upsilon(A) \leq 1$ ;

Desta forma pode-se avaliar a qualidade de segmentação de imagens utilizando quatro tipos de classificação para cada região  $G_i$  e uma métrica final  $\Upsilon(A)$  para toda a imagem onde:

- Uma segmentação correta ocorre quando o algoritmo  $A$  de segmentação gera uma única região  $S_j$  para uma única região  $G_i$ ;
- Uma segmentação excessiva ocorre quando o algoritmo  $A$  de segmentação gera um número de regiões maior que o esperado, ou seja, ocorre uma divisão de uma região  $G_i$  em várias regiões  $S_j$ ;
- Uma segmentação insuficiente ocorre quando o algoritmo  $A$  de segmentação gera um número de regiões menor que o esperado, ou seja, ocorre uma fusão entre regiões  $G_i$  em uma única região  $S_j$ ;
- Uma segmentação ausente ocorre quando uma região  $G_i$  não foi segmentada, ou seja, quando o algoritmo  $A$  não segmentou uma determinada região.



# Capítulo 4

## EXPERIMENTOS E RESULTADOS

### 4.1 Introdução

A avaliação do desempenho de métodos de avaliação de qualidade da segmentação será realizada de duas maneiras diferentes: A primeira consiste em:

- gerar uma imagem de referência *ground truth* como solução ideal da segmentação;
- gerar através de diversas degradações, outras imagens;
- representar e simular as degradações encontradas em segmentações reais;
- avaliar se os mecanismos de avaliação da qualidade detectam e penalizam corretamente essas degradações de qualidade nas suas métricas finais.

Este tipo de avaliação tem grande importância, pois pode-se saber como os métodos de avaliação funcionam e quais são os parâmetros que eles realmente avaliam.

A segunda maneira consiste em:

- aplicar os métodos de avaliação de qualidade sobre resultados obtidos de processos reais de segmentação;

- avaliar a qualidade destes algoritmos a partir de uma base de imagens *ground truth*.

A importância desta abordagem é avaliar e mensurar o resultado de algoritmos reais de segmentação de envelopes postais.

## 4.2 Avaliação do desempenho de métodos de avaliação de qualidade da segmentação.

### 4.2.1 Introdução

Para facilitar a compreensão e a visualização das tabelas e gráficos apresentados neste capítulo, será apresentado abaixo a denominação empregada para cada método:

- Discrepância baseada no número de pixels mal segmentados – erro multi-classe I de [Zha96] que avalia a porcentagem dos pixels da classe "preto" que não foram classificados como preto, denominado **DBNPMS Tipo  $I_f$** ;
- Discrepância baseada no número de pixels mal segmentados – erro multi-classe de [Zha96] que avalia a porcentagem dos pixels da classe "branco" que não foram classificados como branco, denominado **DBNPMS Tipo  $I_b$** ;
- Discrepância baseada na posição dos pixels mal segmentados –  $ND$  de [Zha96], denominado **DBPPMS  $I_f$** ;
- Discrepância baseada na posição dos pixels mal segmentados –  $ND$  de [Zha96] para avaliar os pixels do *background*, denominado **DBPPMS  $I_b$** ;
- A nova metodologia de avaliação de envelopes postais, denominada de  $\Upsilon_f(A)$ ;
- A nova metodologia de avaliação de envelopes postais, para avaliar os pixels do *background* denominada de  $\Upsilon_b(A)$ ;



- A avaliação de discrepância baseada na posição dos centróides mal segmentados, denominado de **DBPPMS  $II_f$** ;
- A avaliação de discrepância baseada na posição dos centróides mal segmentados para avaliar os pixels do *background*, denominado de **DBPPMS  $II_b$** ;

Conforme visto na seção anterior, a avaliação da qualidade foi realizada de duas formas. Os testes iniciais dos métodos de avaliação de qualidade da segmentação implementados puderam revelar que somente o método erro multi-classe de [Zha96] avalia tanto o excesso como a falta de pixels do *foreground* (**pixels pretos**) da imagem como se pode observar nas tabelas 4.1 e 4.2, onde se pode observar a pequena variação do novo método quanto ao acréscimo de pixels pretos atribuídos ao ruído pimenta e invariação dos métodos DBPPMS I e II quanto ao acréscimo de pixels brancos atribuídos ao ruído sal respectivamente. Desta forma foi implementado a avaliação complementar dos demais métodos.

Pimenta	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $II_f$	$\Upsilon_f(A)$
1%	1,00	1,00	0,99	0,90	<b>1,00</b>
5%	0,98	1,00	0,97	0,90	<b>1,00</b>
10%	0,95	1,00	0,96	0,89	<b>1,00</b>
15%	0,93	1,00	0,96	0,89	<b>1,00</b>
20%	0,91	1,00	0,95	0,89	<b>1,00</b>
25%	0,89	1,00	0,95	0,88	<b>0,99</b>
30%	0,87	1,00	0,94	0,88	<b>0,99</b>
35%	0,85	1,00	0,94	0,88	<b>0,99</b>
40%	0,84	1,00	0,93	0,88	<b>0,99</b>
45%	0,82	1,00	0,93	0,88	<b>0,97</b>
50%	0,80	1,00	0,93	0,88	<b>0,95</b>

Tabela 4.1: Resultados iniciais para ruído pimenta.

Outra constatação dos testes iniciais foi a igualdade dos valores apresentados pelo método DBNPMS Multi-classe I e II, onde o erro Multi-

Sal	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $II_f$	$\Upsilon_f(A)$
1%	1,00	1,00	<b>1,00</b>	<b>1,00</b>	0,98
5%	1,00	0,98	<b>1,00</b>	<b>1,00</b>	0,91
10%	1,00	0,95	<b>1,00</b>	<b>1,00</b>	0,83
15%	1,00	0,93	<b>1,00</b>	<b>1,00</b>	0,75
20%	1,00	0,91	<b>1,00</b>	<b>1,00</b>	0,67
25%	1,00	0,89	<b>1,00</b>	<b>1,00</b>	0,57
30%	1,00	0,87	<b>1,00</b>	<b>1,00</b>	0,50
35%	1,00	0,85	<b>1,00</b>	<b>1,00</b>	0,43
40%	1,00	0,84	<b>1,00</b>	<b>1,00</b>	0,28
45%	1,00	0,82	<b>1,00</b>	<b>1,00</b>	0,11
50%	1,00	0,80	<b>1,00</b>	<b>1,00</b>	0,00

Tabela 4.2: Resultados iniciais para ruído sal.

classe I e II diferem apenas na interpretação dos resultados, como pode-se observar na tabela 4.3, onde **Type  $I_b$**  representa a porcentagem dos pixels da classe branco (*background*) que não foram classificados como branco; **Type  $I_f$**  representa a porcentagem dos pixels da classe preto (*foreground*) que não foram classificados como preto; **Type  $II_f$**  representa a porcentagem dos pixels da classe preto (*foreground*) que foram classificados como branco (*background*); e **Type  $II_b$**  representa a porcentagem dos pixels da classe branco (*background*) que foram classificados como preto (*foreground*). Desta forma serão aqui apresentados apenas os valores do erro Multi-classe I. Assim temos oito métricas de avaliação, onde quatro métricas avaliam os pixels pretos mal segmentados e as outras quatro avaliam os pixels brancos mal segmentados, uma vez que implementamos também uma modificação do método de [Zha96] onde calcula-se a distância entre os centróides dos pixels mal segmentados.

```

abcdefghijklmnop
nopqrstuvwxyz
ABCDEFGHIJK
LMNOPQRSTU
VWXYZ
0123456789 ?!+/*\|:
!@#%'^~`&()[]{}<>.,;

```

(a)

<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>	<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>	<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>
---	---	---

(b)

(c)

(d)

<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>	<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>	<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>
---	---	---

(e)

(f)

(g)

<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>	<pre> <b>abcdefghijklmnop</b> <b>nopqrstuvwxyz</b> <b>ABCDEFGHIJK</b> <b>LMNOPQRSTU</b> <b>VWXYZ</b> <b>0123456789 ?!+/*\ :</b> <b>!@#%'^~`&amp;()[]{}&lt;&gt;.,;</b> </pre>	<pre> abcdefghijklmnop nopqrstuvwxyz ABCDEFGHIJK LMNOPQRSTU VWXYZ 0123456789 ?!+/*\ : !@#%'^~`&amp;()[]{}&lt;&gt;.,; </pre>
---	--	---

(h)

(i)

(j)

Figura 4.1: (a) Imagem *ground truth*, (b) Pimenta(15%), (c) Sal(20%), (d) Sal-Pimenta(50%), (e) Dilatação( $EE_{cruz}$  1 iteração), (f) Dilatação( $EE_{quadrado}$  3 iterações), (g) Dilatação( $EE_{cruz}$  2 iterações) + Sal-Pimenta(20%), (h) Erosão( $EE_{cruz}$  2 iterações), (i) Erosão( $EE_{quadrado}$  5 iterações), (j) Erosão( $EE_{cruz}$  2 iterações) + Sal-Pimenta(35%).

Sal	Type $I_b$	Type $I_f$	Type $II_f$	Type $II_b$
Pimenta				
1%	0,94	1,00	1,00	0,94
5%	0,92	0,98	0,98	0,92
10%	0,90	0,95	0,95	0,90
15%	0,88	0,93	0,93	0,88
20%	0,86	0,91	0,91	0,86
25%	0,85	0,89	0,89	0,85
30%	0,83	0,87	0,87	0,83
35%	0,81	0,85	0,85	0,81
40%	0,80	0,84	0,84	0,80
45%	0,78	0,82	0,82	0,78
50%	0,77	0,80	0,80	0,77

Tabela 4.3: DBNPMS Multi-classe I e II - erosão ( $EE_{cruz}$  uma iterações) com % de ruído sal-pimenta

#### 4.2.2 Avaliação utilizando imagens artificialmente modificadas

Na primeira abordagem foi utilizada uma imagem *ground truth* de ( $640 \times 480$  pixels) que foi artificialmente deteriorada para simular as situações ocorridas em processos reais de segmentação. As modificações feitas na imagem original foram geradas a partir da combinação dos ruídos sal, pimenta e sal-pimenta ( $\{1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50\%$ ),  $i$  ( $\in [1, 5]$ ) erosões e dilatações (ambos com  $EE_{cruz}$  e  $EE_{quadrado}$ ), como pode-se observar na figura 4.1. Desta forma pode-se avaliar o comportamento dos métodos de avaliação da segmentação de acordo com o tipo de erro apresentado, tanto isoladamente como combinados. Como o método  $\Upsilon(A)$  varia de 1 à -1 este foi normalizado e foi utilizado o valor complementar dos demais métodos na geração das tabelas e gráficos.

Os métodos de avaliação serão utilizados para avaliar a qualidade das imagens deterioradas, comparadas com suas respectivas imagens *ground truth* e desta forma poder acompanhar a sensibilidade de cada método de avaliação, com relação a cada tipo de erro encontrado na segmentação. Os gráficos e tabelas a seguir demonstram os resultados obtidos, bem como auxiliam na análise dos métodos implementados.

### 4.2.2.1 Erosão

Os métodos DBNMSP  $I_b$  e  $\Upsilon_f(A)$  têm se mostrado sensíveis ao acréscimo de informações geradas através da erosão binária, utilizando os elementos estruturantes  $EE_{cruz}$  e  $EE_{quadrado}$ , já os demais métodos tem se mostrado pouco sensível, como pode-se observar nas tabelas 4.4 e 4.5 e também nos gráficos 4.2, 4.3, 4.4, 4.5, 4.6 e 4.7.

Erosão	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
1	<b>0,94</b>	1,00	0,96	1,00	0,89	1,00	<b>1,00</b>	0,02
2	<b>0,88</b>	1,00	0,94	1,00	0,89	1,00	<b>0,95</b>	0,01
3	<b>0,83</b>	1,00	0,93	1,00	0,88	1,00	<b>0,86</b>	0,01
4	<b>0,78</b>	1,00	0,92	1,00	0,88	1,00	<b>0,84</b>	0,00
5	<b>0,72</b>	1,00	0,91	1,00	0,88	1,00	<b>0,84</b>	0,00

Tabela 4.4: Resultados médios para erosão utilizando EE cruz

Erosão	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
1	<b>0,93</b>	1,00	0,96	1,00	0,89	1,00	<b>1,00</b>	0,02
2	<b>0,86</b>	1,00	0,94	1,00	0,89	1,00	<b>0,95</b>	0,01
3	<b>0,79</b>	1,00	0,92	1,00	0,88	1,00	<b>0,86</b>	0,01
4	<b>0,72</b>	1,00	0,91	1,00	0,88	1,00	<b>0,84</b>	0,01
5	<b>0,66</b>	1,00	0,90	1,00	0,89	1,00	<b>0,81</b>	0,00

Tabela 4.5: Resultados para erosão utilizando EE quadrado

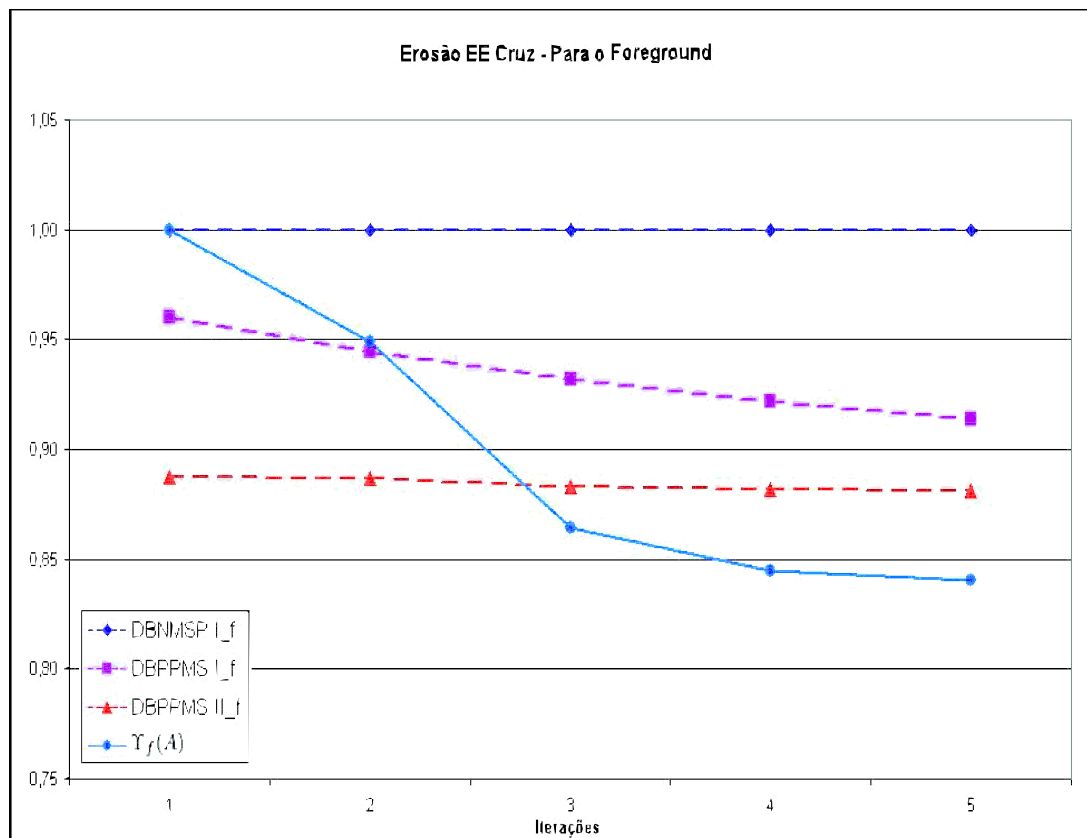


Figura 4.2: Gráfico referente a normalização dos valores do *foreground* para Erosão utilizando EE cruz

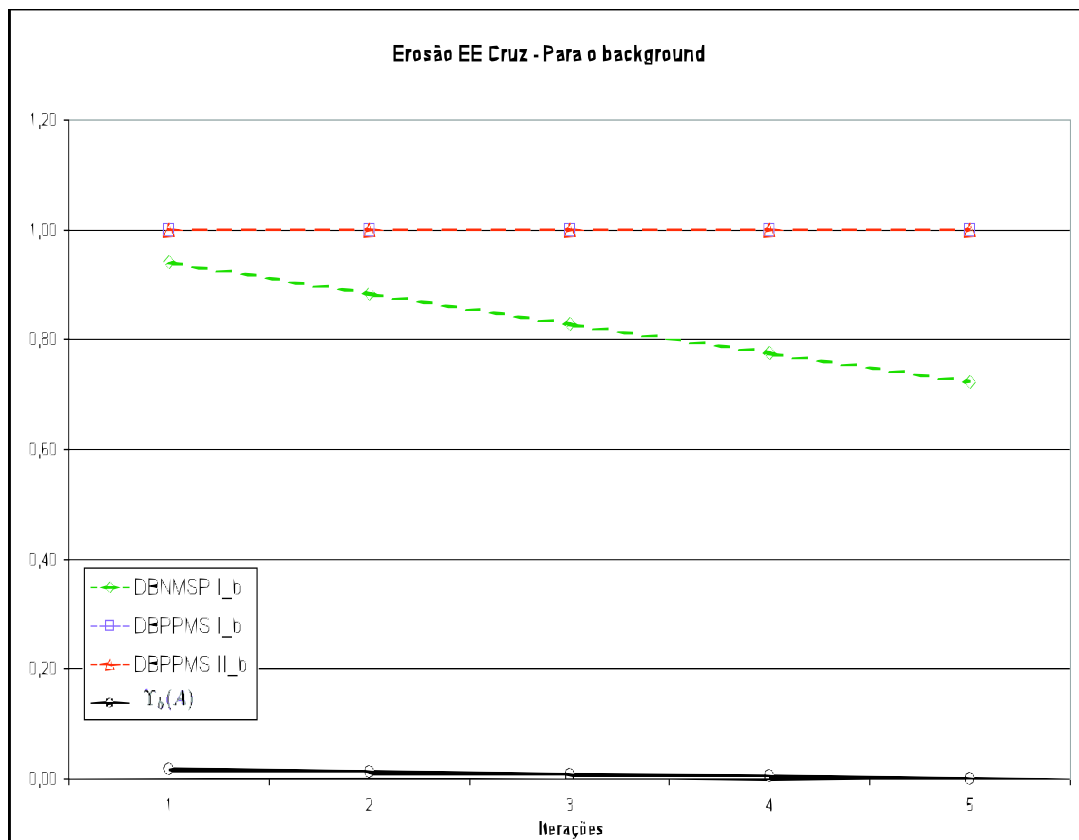


Figura 4.3: Gráfico referente a normalização dos valores do *background* para erosão utilizando EE cruz

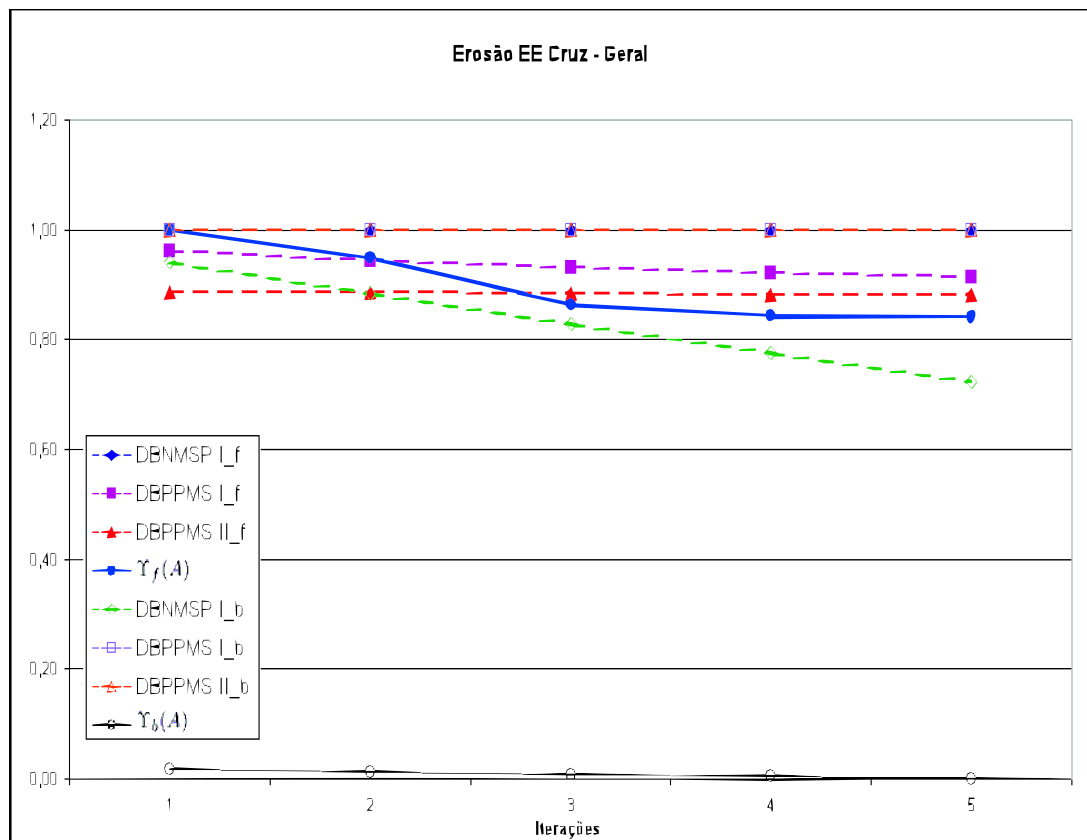


Figura 4.4: Gráfico referente a normalização dos valores gerais para erosão utilizando EE cruz



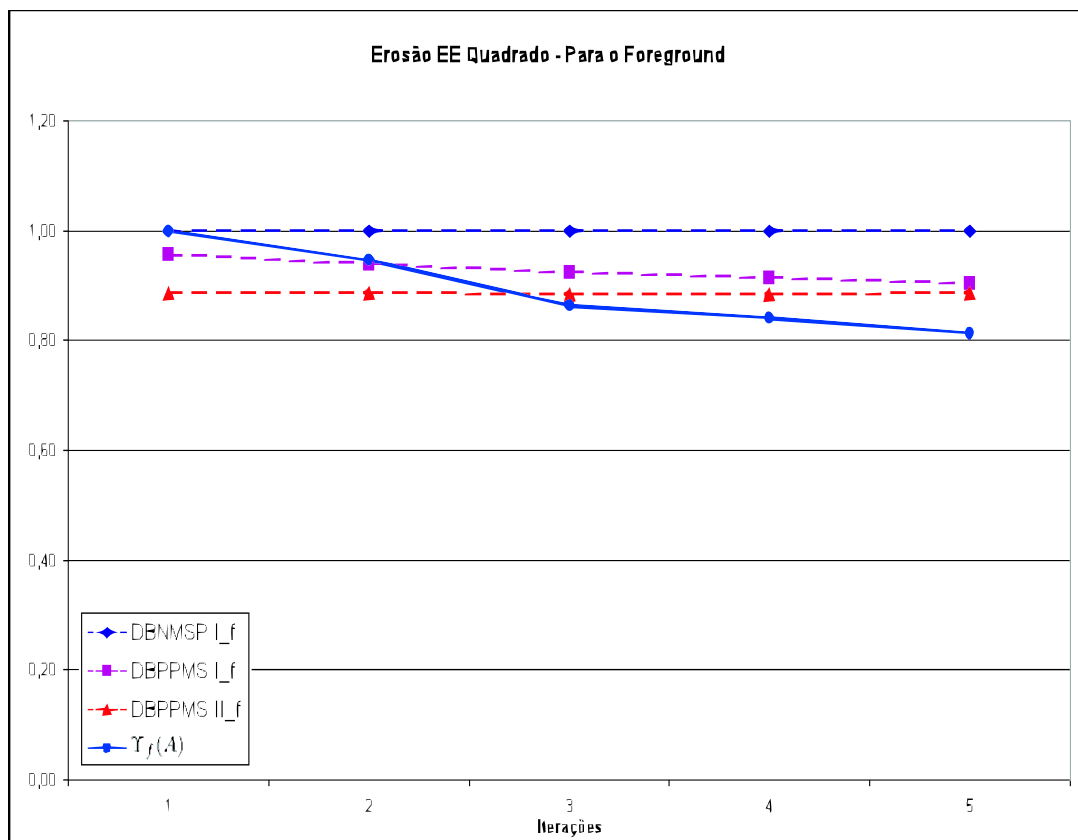


Figura 4.5: Gráfico referente a normalização dos valores do *foreground* para erosão utilizando EE quadrado

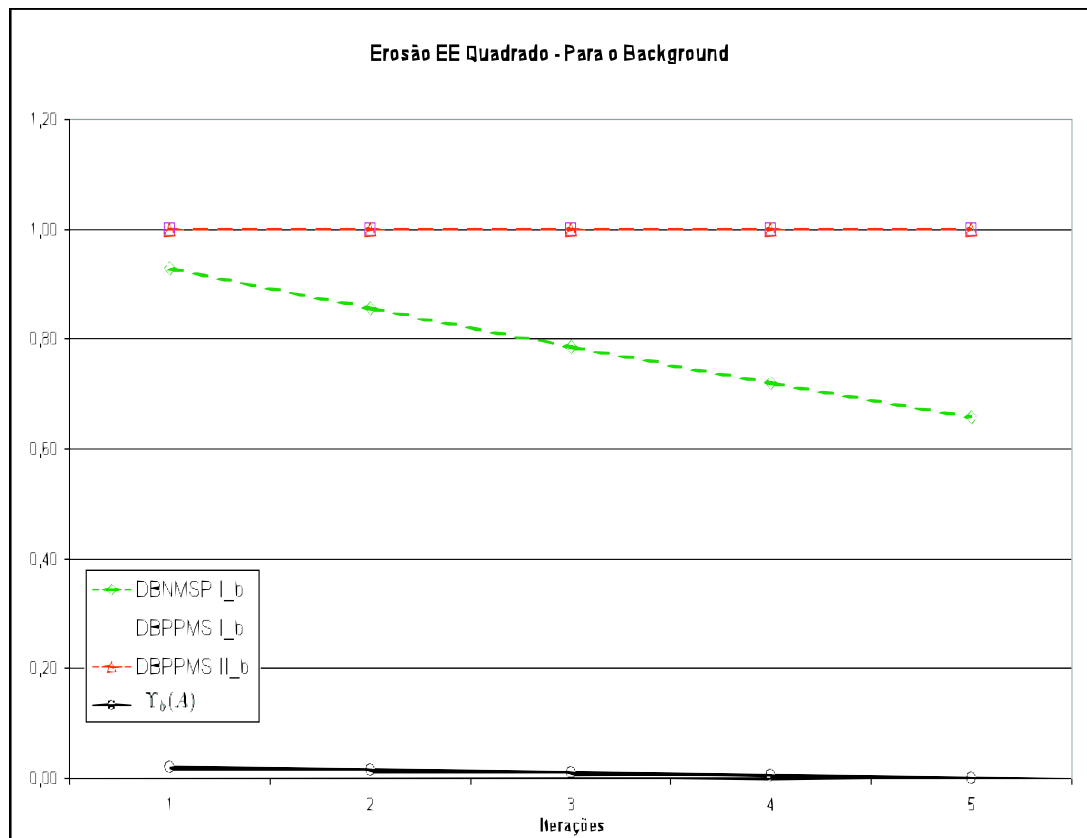


Figura 4.6: Gráfico referente a normalização dos valores do *background* para erosão utilizando EE quadrado

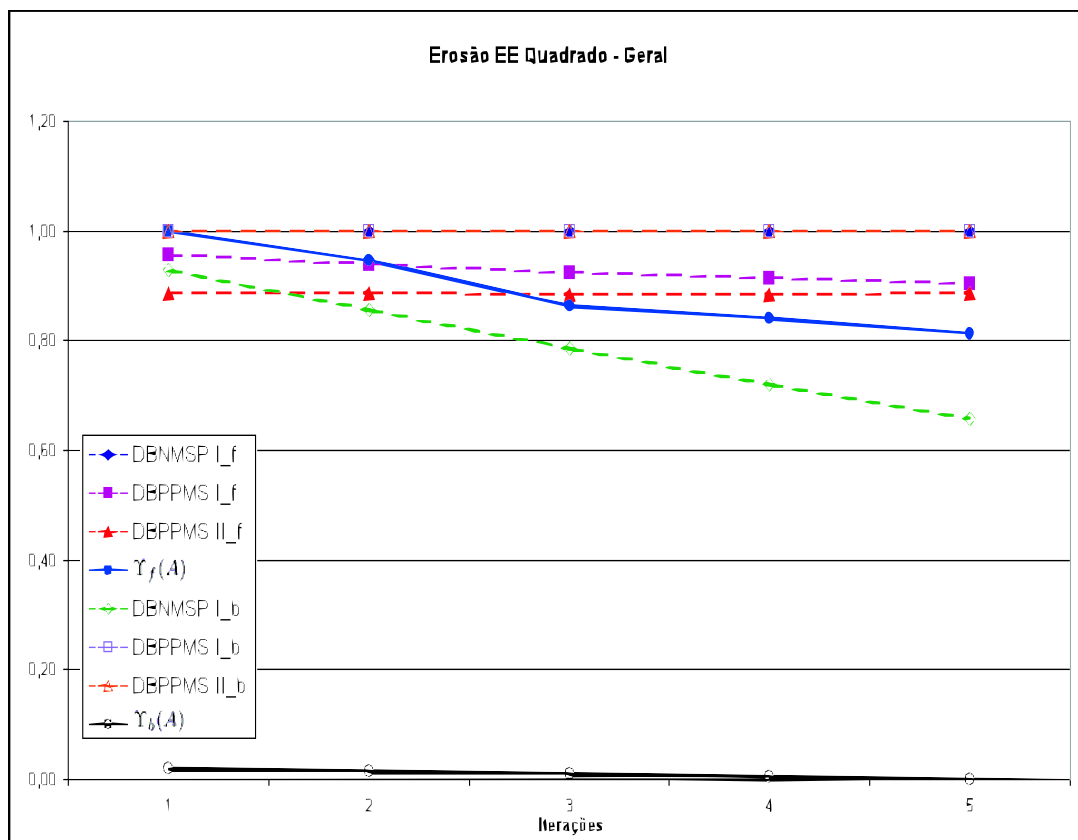


Figura 4.7: Gráfico referente a normalização dos valores gerais para erosão utilizando EE quadrado

### 4.2.2.2 Dilatação

Os métodos  $\Upsilon_f(A)$  e DBNMSP  $I_f$  têm se mostrado muito sensíveis ao decréscimo de informações geradas através da dilatação binária, utilizando elementos estruturantes  $EE_{cruz}$  e  $EE_{quadrado}$ . Cabe ressaltar que as diferenças de resultado obtidos utilizando elementos estruturantes foi considerável, demonstrando a grande sensibilidade quanto a perda de informação do *foreground*. Já os demais métodos têm se mostrado pouco sensível, como pode-se observar nas tabelas 4.6 e 4.7 e nos gráficos 4.8, 4.9, 4.10, 4.11, 4.12 e 4.13.

Dilatação	DBNPMS	DBNPMS	DBPPMS	DBPPMS	DBPPMS	DBPPMS	$\Upsilon_f(A)$	$\Upsilon_b(A)$
	Type $I_b$	Type $I_f$	$I_f$	$I_b$	$II_f$	$II_b$		
1	1,00	<b>0,73</b>	1,00	0,96	1,00	0,89	<b>0,80</b>	1,00
2	1,00	<b>0,47</b>	1,00	0,95	1,00	0,88	<b>0,60</b>	1,00
3	1,00	<b>0,23</b>	1,00	0,93	1,00	0,88	<b>0,38</b>	0,99
4	1,00	<b>0,05</b>	1,00	0,93	1,00	0,87	<b>0,19</b>	0,99
5	1,00	<b>0,01</b>	1,00	0,93	1,00	0,87	<b>0,00</b>	0,99

Tabela 4.6: Resultados médios para dilatação utilizando EE cruz

Dilatação	DBNPMS	DBNPMS	DBPPMS	DBPPMS	DBPPMS	DBPPMS	$\Upsilon_f(A)$	$\Upsilon_b(A)$
	Type $I_b$	Type $I_f$	$I_f$	$I_b$	$II_f$	$II_b$		
1	1,00	<b>0,66</b>	1,00	0,96	1,00	0,88	<b>0,82</b>	1,00
2	1,00	<b>0,34</b>	1,00	0,94	1,00	0,88	<b>0,63</b>	1,00
3	1,00	<b>0,10</b>	1,00	0,93	1,00	0,87	<b>0,49</b>	0,99
4	1,00	<b>0,00</b>	1,00	0,93	1,00	0,87	<b>0,11</b>	0,99
5	1,00	<b>0,00</b>	1,00	0,93	1,00	0,87	<b>0,00</b>	0,99

Tabela 4.7: Resultados médios para dilatação utilizando EE quadrado

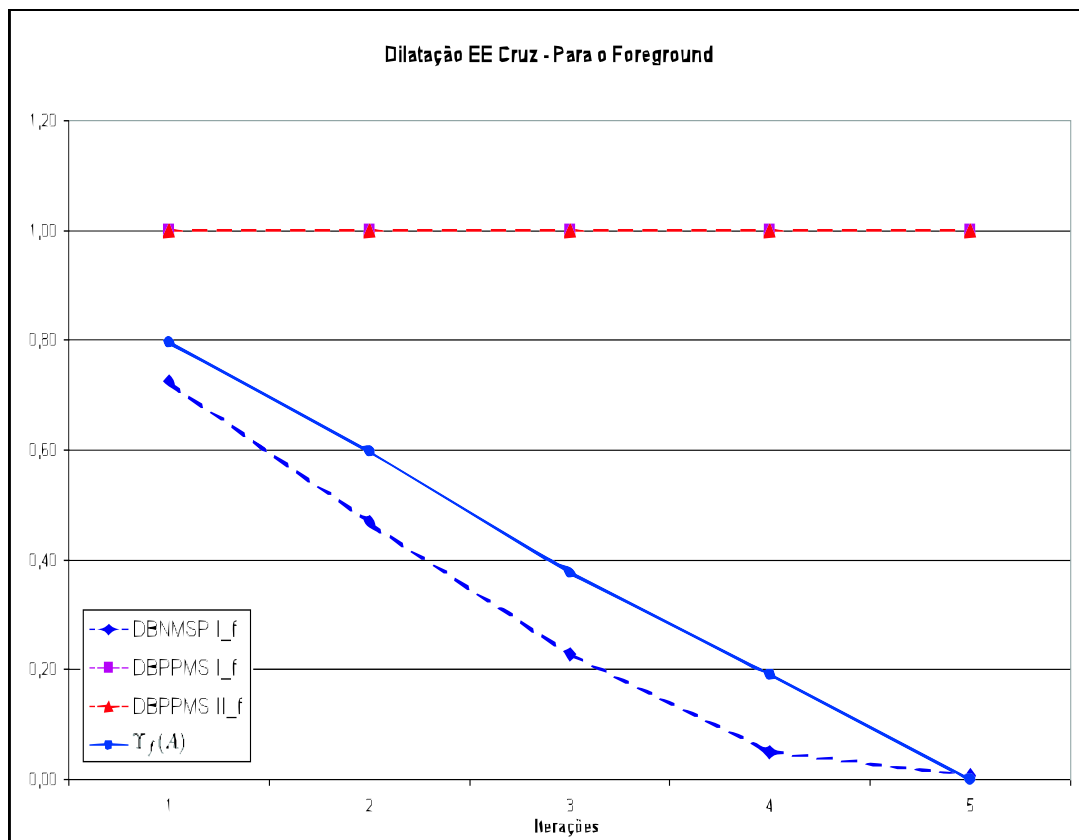


Figura 4.8: Gráfico referente a normalização dos valores do *foreground* para dilatação utilizando EE cruz

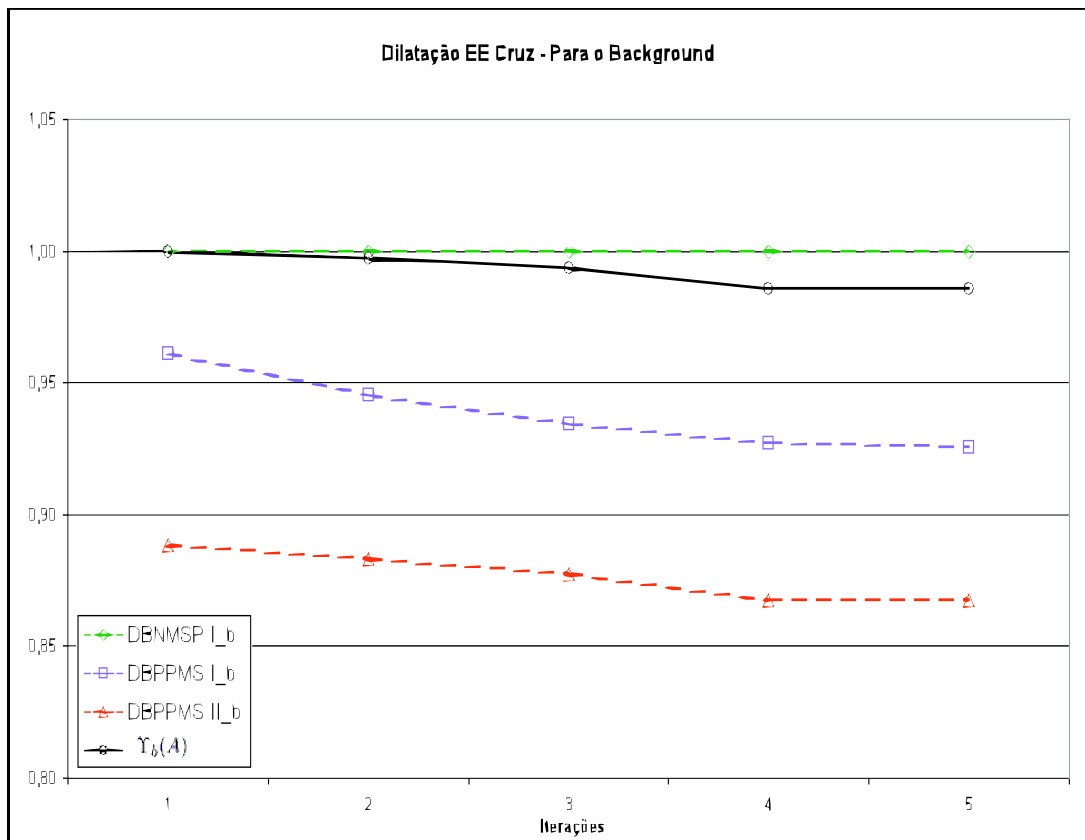


Figura 4.9: Gráfico referente a normalização dos valores do *background* para dilatação utilizando EE cruz

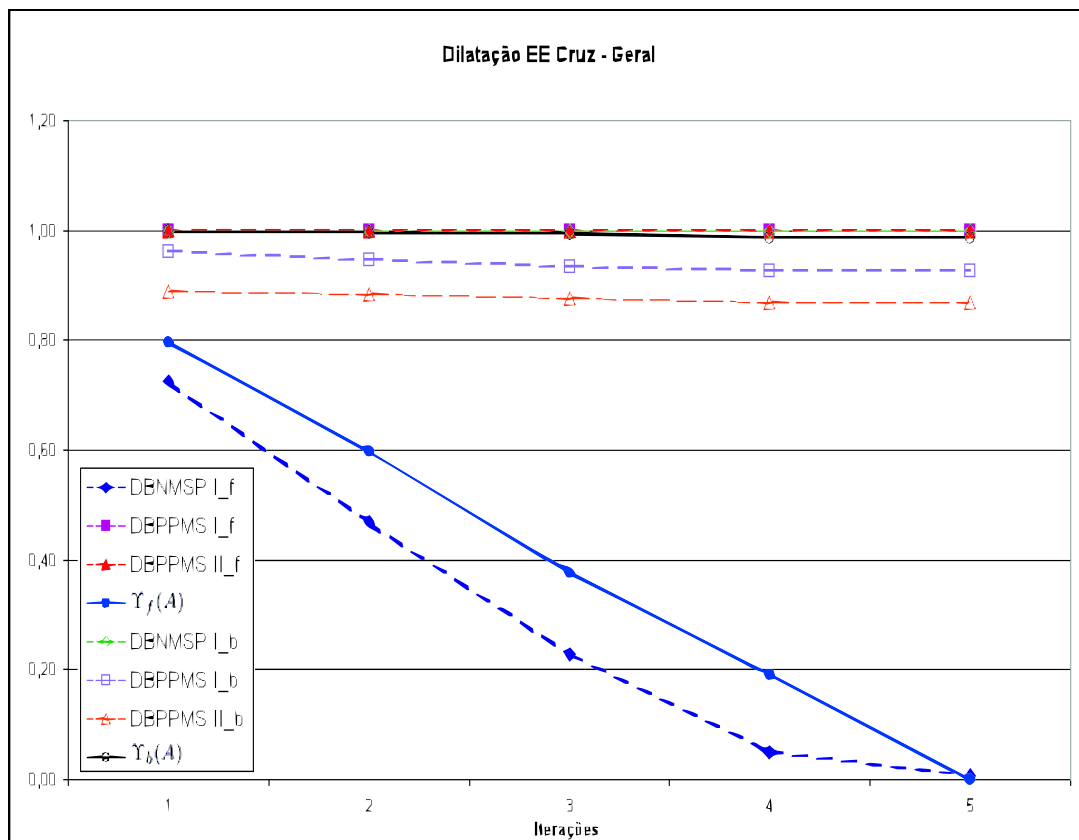


Figura 4.10: Gráfico referente a normalização dos valores gerais para dilatação utilizando EE cruz

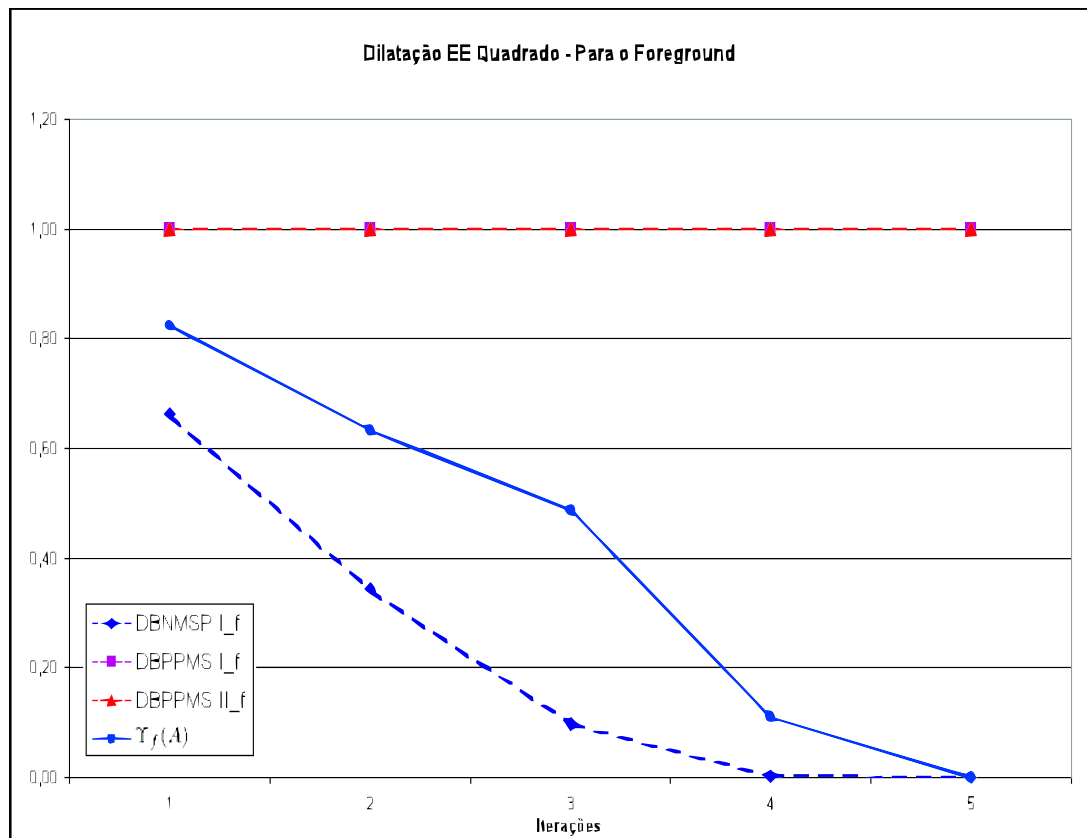


Figura 4.11: Gráfico referente a normalização dos valores do *foreground* para dilatação utilizando EE quadrado



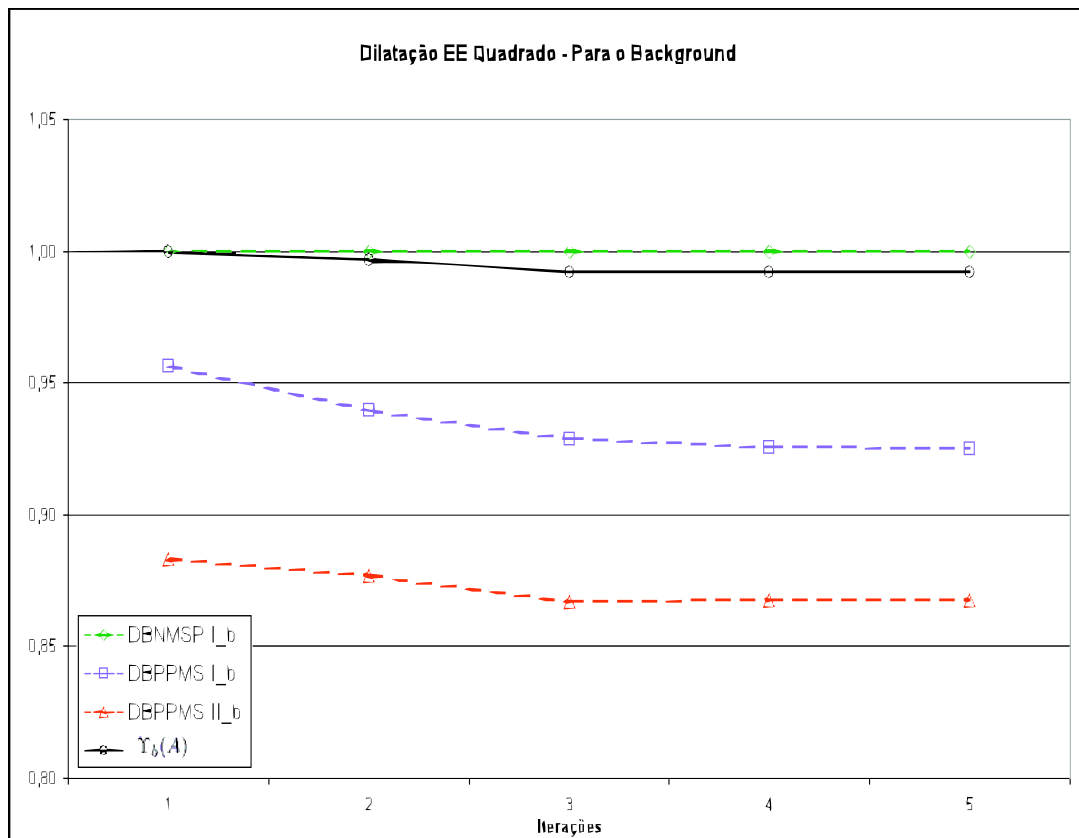


Figura 4.12: Gráfico referente a normalização dos valores do *background* para dilatação utilizando EE quadrado

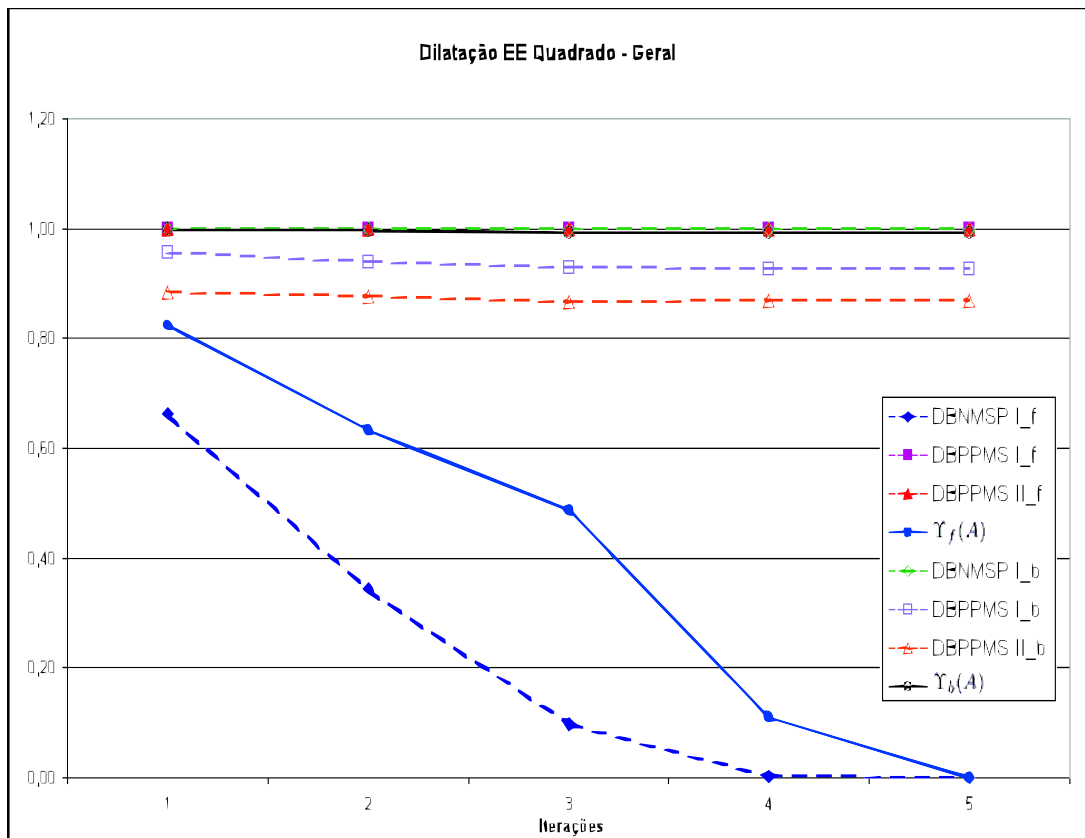


Figura 4.13: Gráfico referente a normalização dos valores gerais para dilatação utilizando EE quadrado

### 4.2.2.3 Ruído pimenta

O método  $\Upsilon_b(A)$  tem se mostrado muito sensível à presença de ruídos no *background* das imagens segmentadas, gerados pelo ruído tipo pimenta. O método DBNMSP  $I_b$  tem mostrado-se sensível ao ruído tipo pimenta. Já os demais métodos têm se mostrado pouco sensíveis, como pode-se observar na tabela 4.8 e nos gráficos 4.14, 4.15 e 4.16

Pimenta	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
1%	<b>1,00</b>	1,00	0,99	1,00	0,90	1,00	1,00	<b>1,00</b>
5%	<b>0,98</b>	1,00	0,97	1,00	0,90	1,00	1,00	<b>0,98</b>
10%	<b>0,95</b>	1,00	0,96	1,00	0,89	1,00	1,00	<b>0,00</b>
15%	<b>0,93</b>	1,00	0,96	1,00	0,89	1,00	1,00	<b>0,00</b>
20%	<b>0,91</b>	1,00	0,95	1,00	0,89	1,00	1,00	<b>0,00</b>
25%	<b>0,89</b>	1,00	0,95	1,00	0,88	1,00	0,99	<b>0,00</b>
30%	<b>0,87</b>	1,00	0,94	1,00	0,88	1,00	0,99	<b>0,00</b>
35%	<b>0,85</b>	1,00	0,94	1,00	0,88	1,00	0,99	<b>0,00</b>
40%	<b>0,84</b>	1,00	0,93	1,00	0,88	1,00	0,99	<b>0,00</b>
45%	<b>0,82</b>	1,00	0,93	1,00	0,88	1,00	0,97	<b>0,00</b>
50%	<b>0,80</b>	1,00	0,93	1,00	0,88	1,00	0,95	<b>0,00</b>

Tabela 4.8: Resultados médios para ruído pimenta

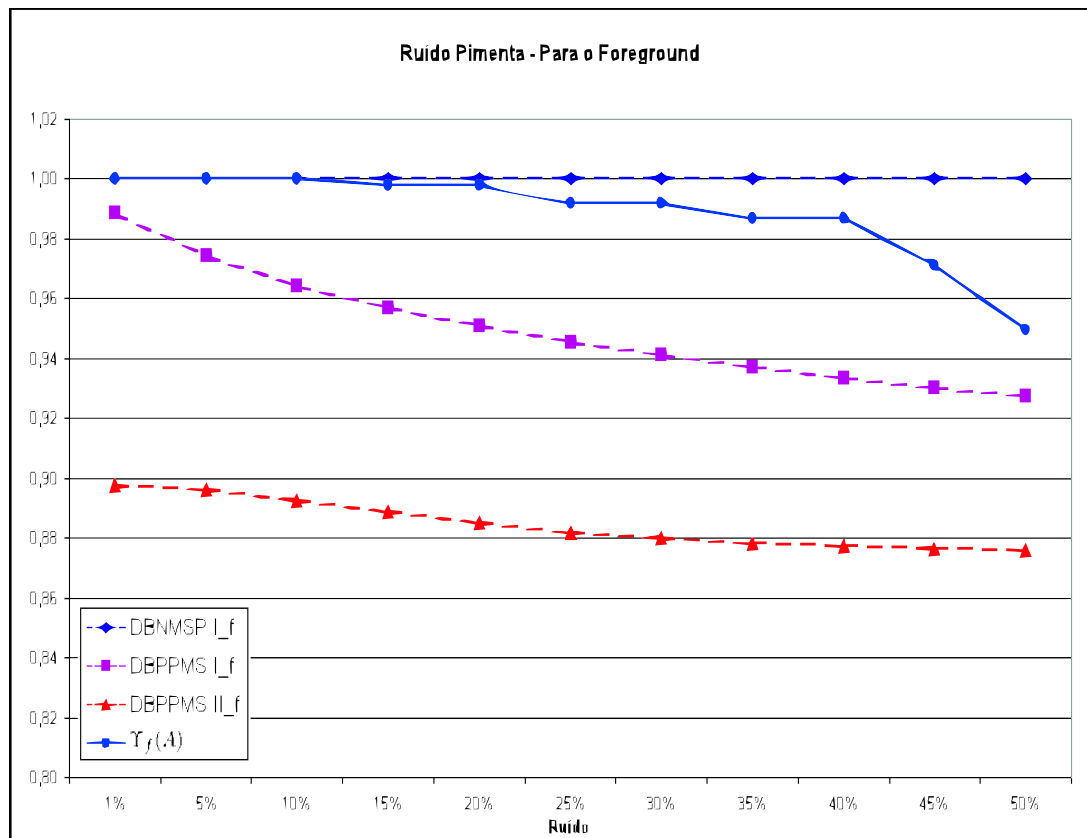


Figura 4.14: Gráfico referente a normalização dos valores do *foreground* para o ruído pimenta.

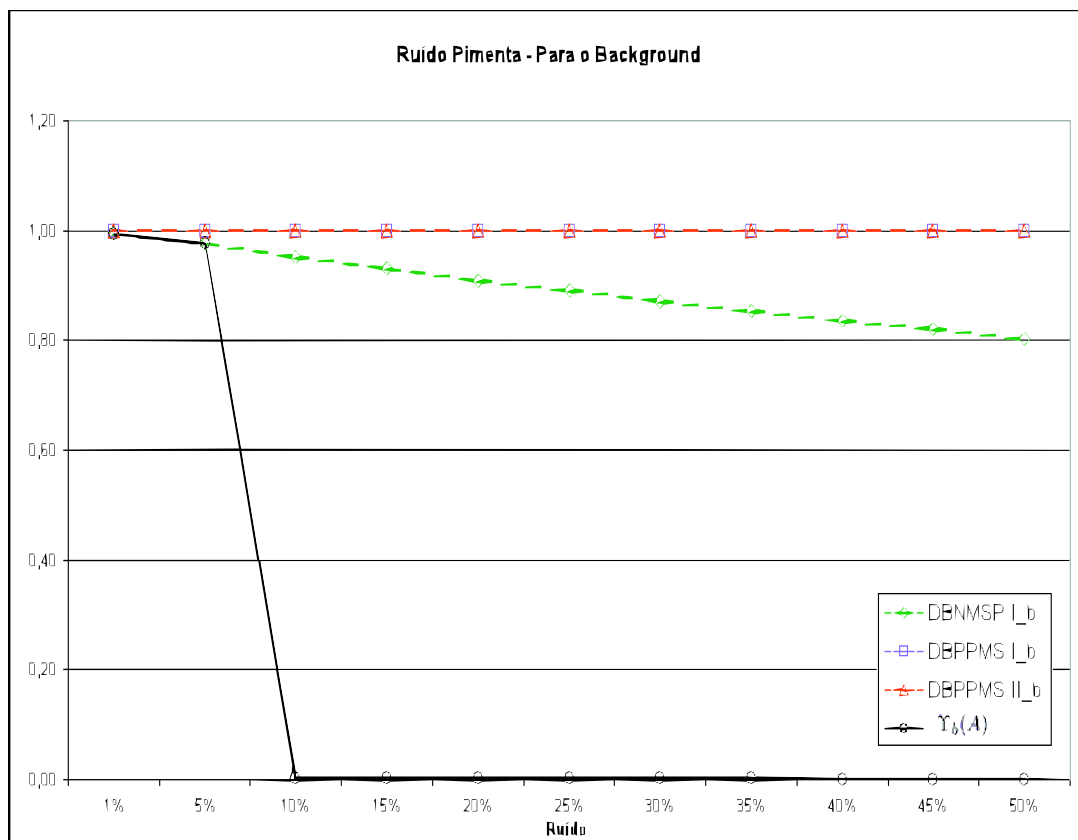


Figura 4.15: Gráfico referente a normalização dos valores do *background* para o ruído pimenta.

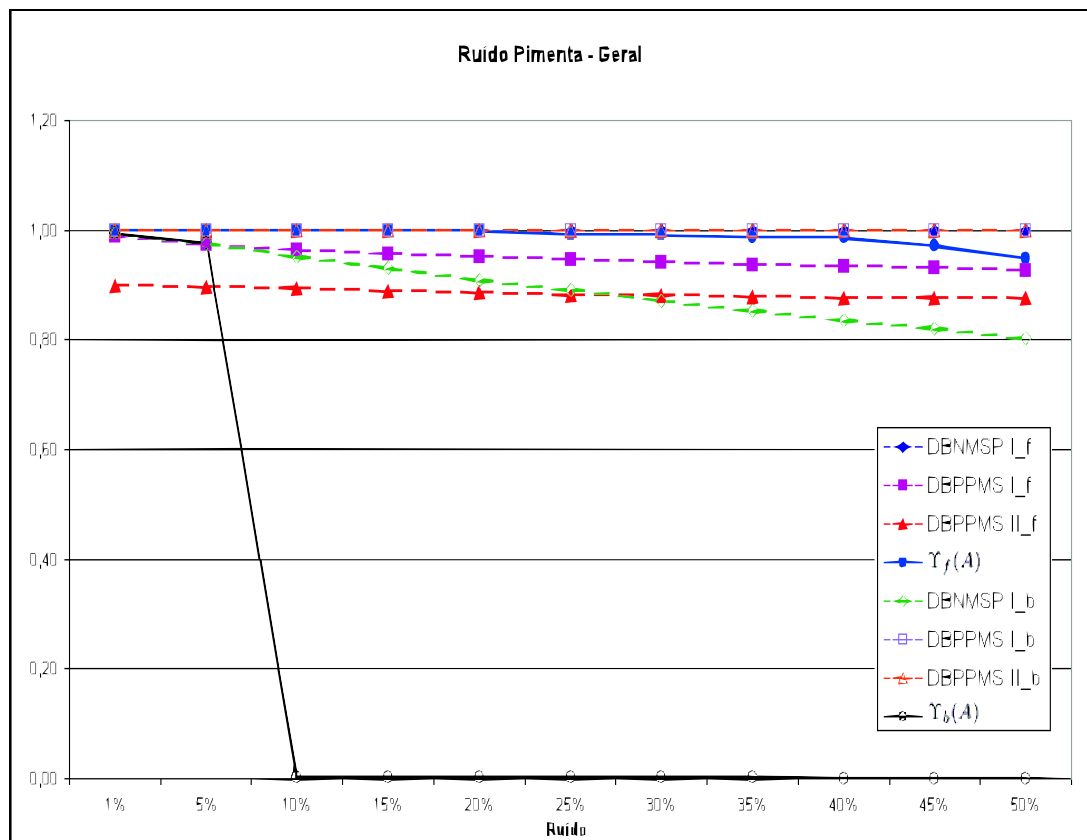


Figura 4.16: Gráfico referente a normalização dos valores gerais para o ruído pimenta.

#### 4.2.2.4 Ruído sal

O método  $\Upsilon_f(A)$  tem se mostrado muito sensível à presença de ruídos no *foreground* das imagens segmentadas, gerado pelo ruído tipo sal. O método DBNMSP  $I_f$  tem mostrado-se sensível ao ruído tipo sal. Já os demais métodos tem se mostrado pouco sensíveis, como pode-se observar na tabela 4.9 e nos gráficos 4.17, 4.18 e 4.19.

Sal	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
1%	1,00	<b>1,00</b>	1,00	1,00	0,90	0,92	<b>0,98</b>	1,00
5%	1,00	<b>0,98</b>	1,00	0,99	1,00	0,92	<b>0,91</b>	1,00
10%	1,00	<b>0,95</b>	1,00	0,98	1,00	0,94	<b>0,83</b>	1,00
15%	1,00	<b>0,93</b>	1,00	0,98	1,00	0,94	<b>0,75</b>	1,00
20%	1,00	<b>0,91</b>	1,00	0,98	1,00	0,94	<b>0,67</b>	1,00
25%	1,00	<b>0,89</b>	1,00	0,98	1,00	0,94	<b>0,57</b>	1,00
30%	1,00	<b>0,87</b>	1,00	0,97	1,00	0,94	<b>0,50</b>	1,00
35%	1,00	<b>0,85</b>	1,00	0,97	1,00	0,94	<b>0,43</b>	1,00
40%	1,00	<b>0,84</b>	1,00	0,97	1,00	0,94	<b>0,28</b>	0,99
45%	1,00	<b>0,82</b>	1,00	0,97	1,00	0,94	<b>0,11</b>	0,99
50%	1,00	<b>0,80</b>	1,00	0,97	1,00	0,94	<b>0,00</b>	0,98

Tabela 4.9: Resultados médios para ruído sal

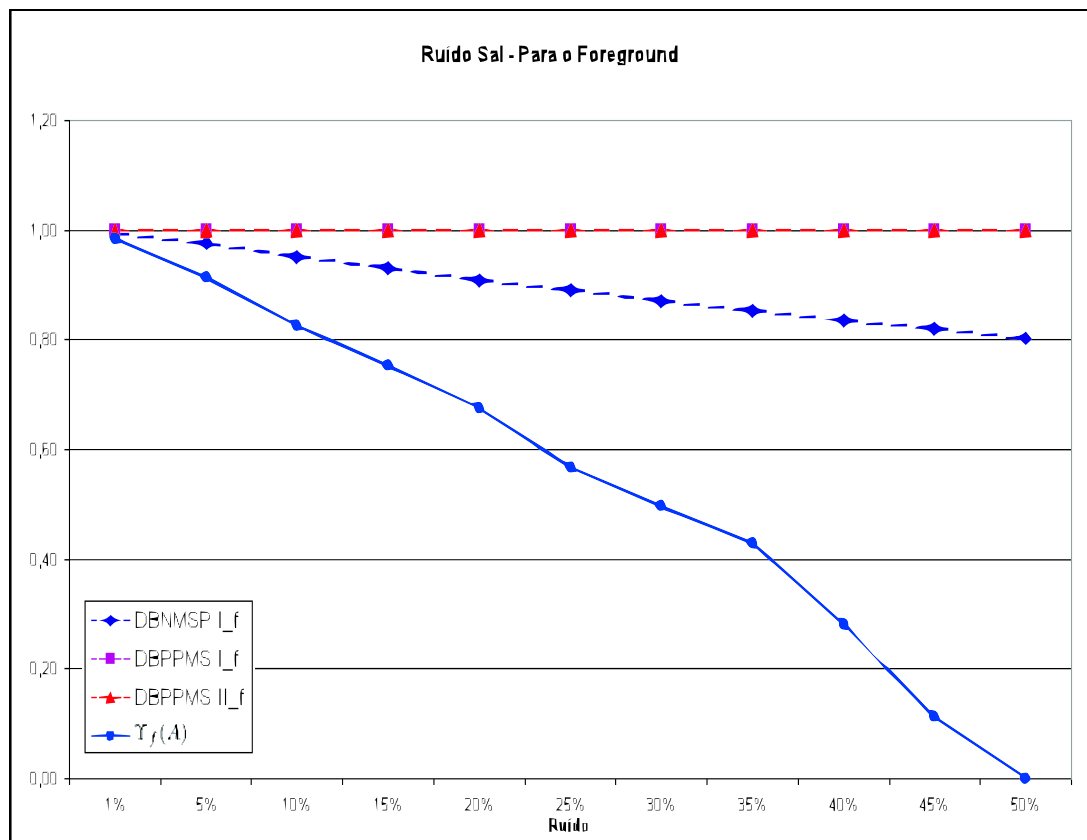


Figura 4.17: Gráfico referente a normalização dos valores do *foreground* para ruído sal



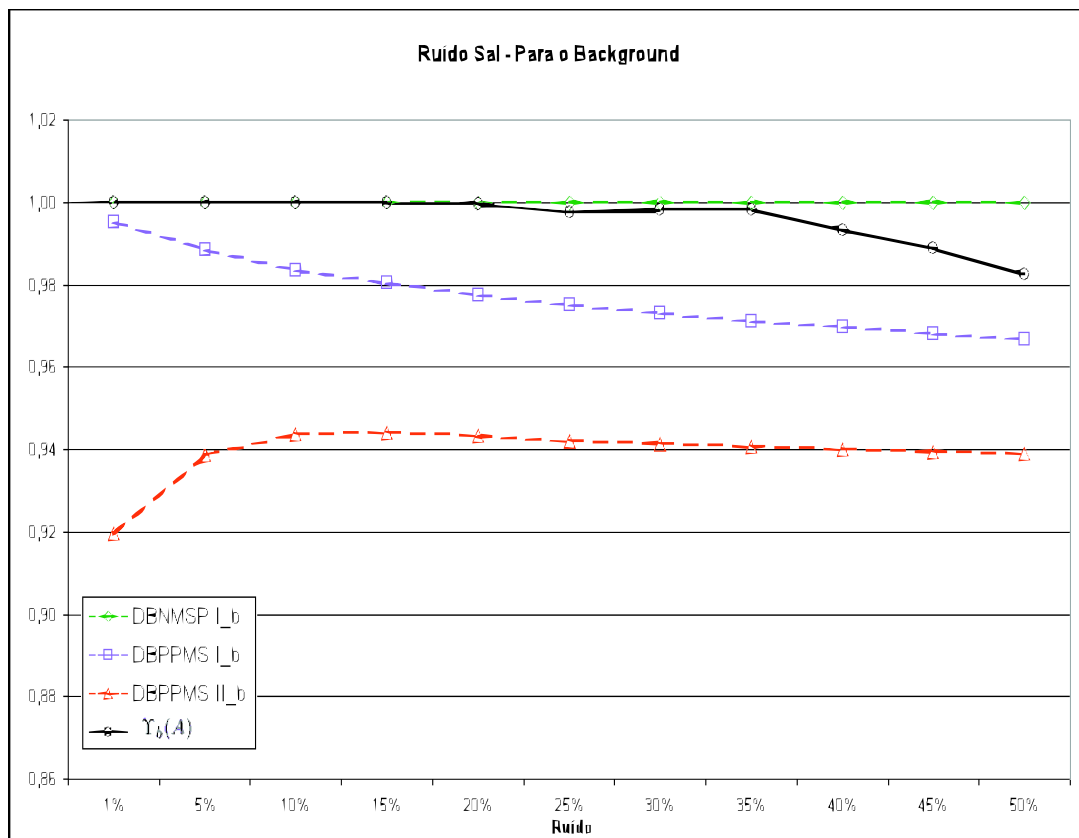


Figura 4.18: Gráfico referente a normalização dos valores do *background* para ruído sal

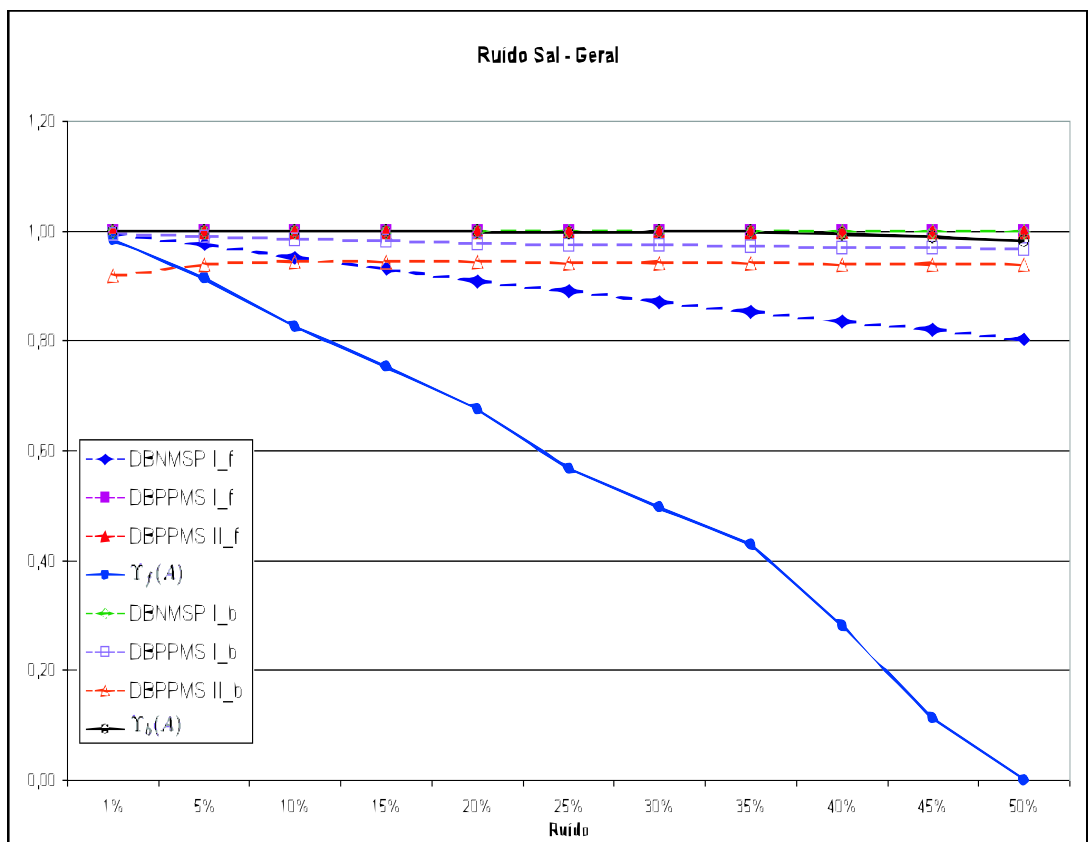


Figura 4.19: Gráfico referente a normalização dos valores gerais para ruído sal

#### 4.2.2.5 Ruído sal-pimenta

O método  $\Upsilon_b(A)$  tem se mostrado muito sensível à presença de ruídos no *foreground* e no *background* das imagens segmentadas, gerado pelo ruído sal-pimenta. Os métodos  $\Upsilon_f(A)$  DBNMSP  $I_f$  e DBNMSP  $I_b$  têm se mostrado sensíveis ao ruído tipo sal-pimenta. Já os demais métodos têm se mostrado pouco sensíveis, como pode-se observar na tabela 4.10 e nos gráficos 4.20, 4.21 e 4.22.

Sal Pimenta	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
1%	1,00	1,00	0,99	1,00	0,90	0,92	<b>1,00</b>	1,00
5%	0,98	0,98	0,97	0,99	0,90	0,94	<b>0,98</b>	0,98
10%	0,95	0,95	0,96	0,98	0,89	0,94	<b>0,96</b>	0,01
15%	0,93	0,93	0,96	0,98	0,89	0,94	<b>0,93</b>	0,01
20%	0,91	0,91	0,95	0,98	0,89	0,94	<b>0,91</b>	0,01
25%	0,89	0,89	0,95	0,98	0,88	0,94	<b>0,88</b>	0,00
30%	0,87	0,87	0,94	0,97	0,88	0,94	<b>0,86</b>	0,00
35%	0,85	0,85	0,94	0,97	0,88	0,94	<b>0,83</b>	0,00
40%	0,84	0,84	0,93	0,97	0,88	0,94	<b>0,80</b>	0,00
45%	0,82	0,82	0,93	0,97	0,88	0,94	<b>0,76</b>	0,00
50%	0,80	0,80	0,93	0,97	0,88	0,94	<b>0,72</b>	0,00

Tabela 4.10: Resultados médios para ruído sal-pimenta

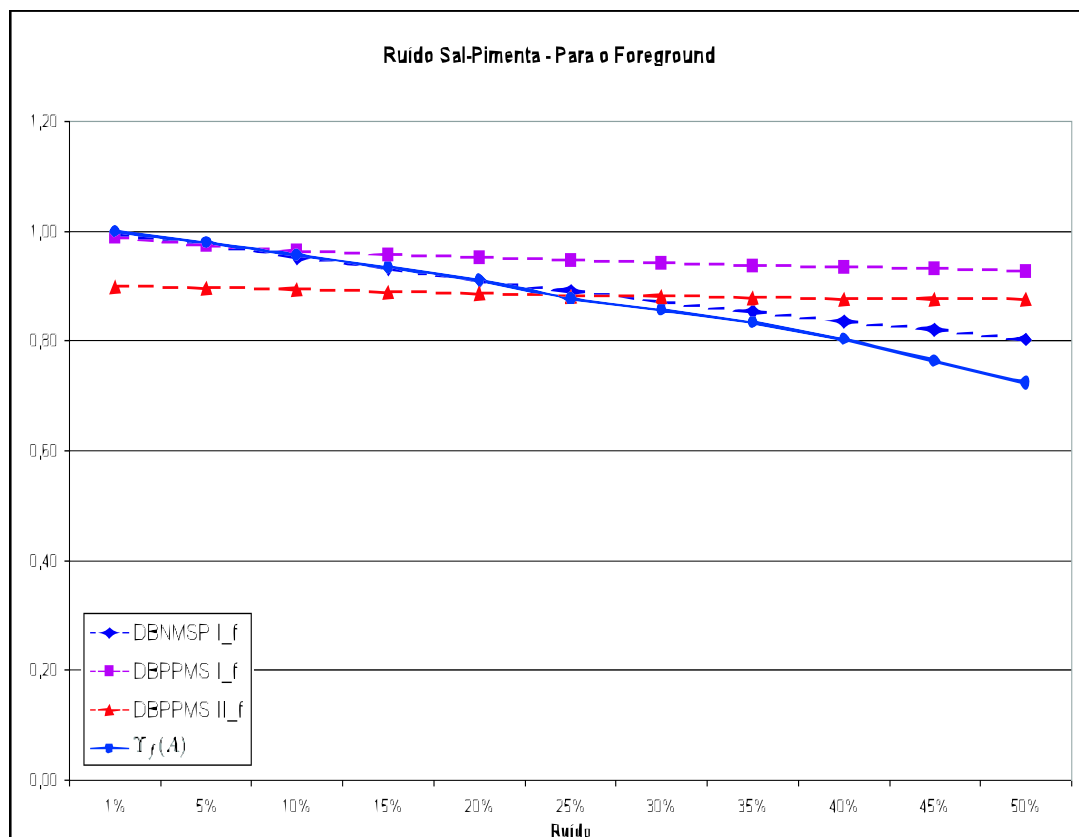


Figura 4.20: Gráfico referente a normalização dos valores do *foreground* para ruído sal-pimenta

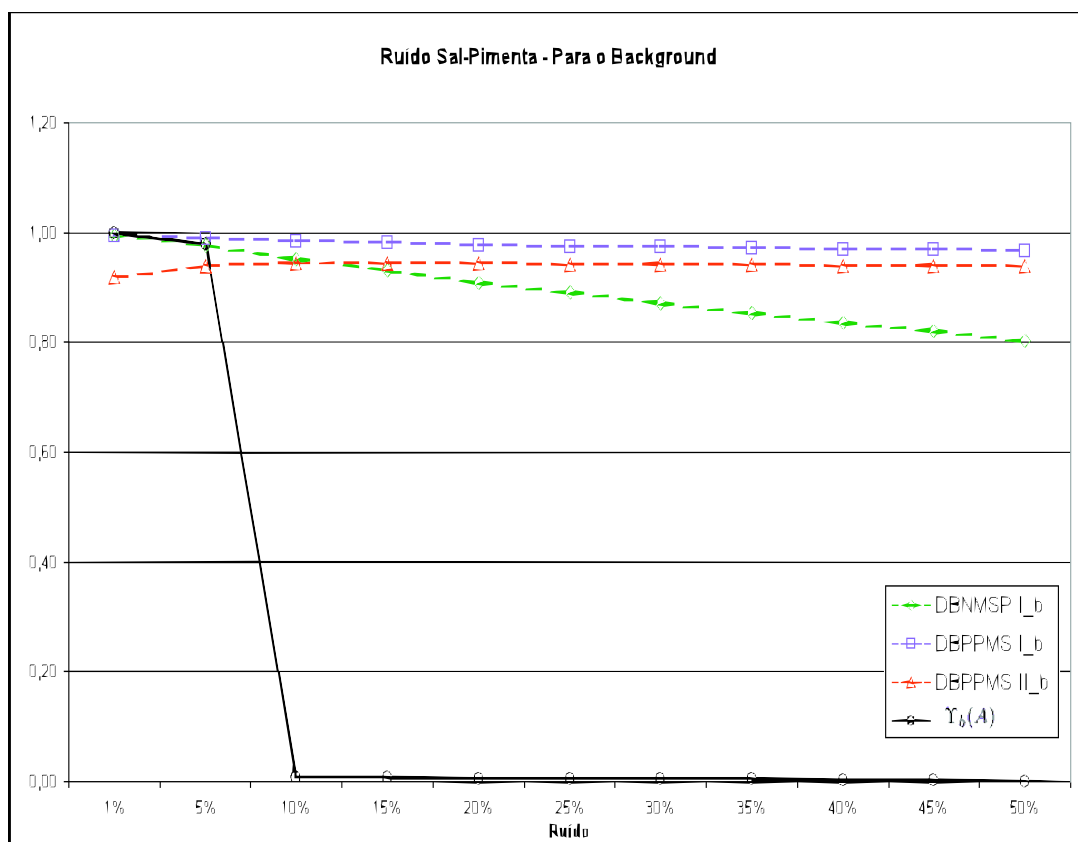


Figura 4.21: Gráfico referente a normalização dos valores do *background* para ruído sal-pimenta

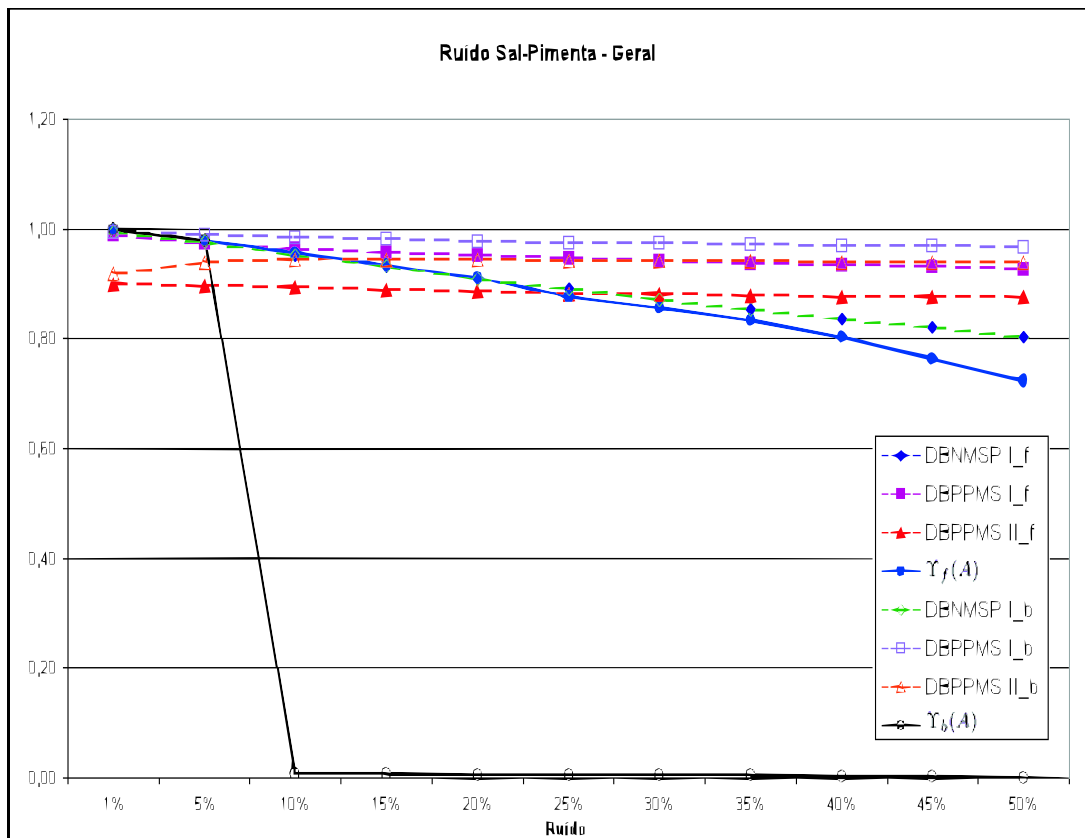


Figura 4.22: Gráfico referente a normalização dos valores gerais para ruído sal-pimenta

#### 4.2.2.6 Erosão + ruído sal-pimenta

Os métodos  $\Upsilon_f(A)$ , DBNMSP  $I_f$  e DBNMSP  $I_b$  têm se mostrado sensíveis ao acréscimo de informações geradas e a presença de ruídos no *foreground* e *background* das imagens segmentadas, gerados através da erosão binária, e pelo ruído sal-pimenta. Já os demais métodos têm se mostrado pouco sensível, como pode-se observar na tabela 4.11 e nos gráficos 4.23, 4.24 e 4.25.

Erosão+ Sal-Pimenta	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
1%	<b>0,94</b>	<b>1,00</b>	0,96	1,00	0,89	0,92	<b>1,00</b>	0,00
5%	<b>0,92</b>	<b>0,98</b>	0,95	0,99	0,90	0,94	<b>0,98</b>	0,00
10%	<b>0,90</b>	<b>0,95</b>	0,95	0,98	0,89	0,94	<b>0,95</b>	0,00
15%	<b>0,88</b>	<b>0,93</b>	0,94	0,98	0,89	0,94	<b>0,93</b>	0,00
20%	<b>0,86</b>	<b>0,91</b>	0,94	0,98	0,88	0,94	<b>0,90</b>	0,00
25%	<b>0,85</b>	<b>0,89</b>	0,94	0,98	0,88	0,94	<b>0,88</b>	0,00
30%	<b>0,83</b>	<b>0,87</b>	0,93	0,97	0,88	0,94	<b>0,85</b>	0,00
35%	<b>0,81</b>	<b>0,85</b>	0,93	0,97	0,88	0,94	<b>0,82</b>	0,00
40%	<b>0,80</b>	<b>0,84</b>	0,93	0,97	0,88	0,94	<b>0,81</b>	0,00
45%	<b>0,78</b>	<b>0,82</b>	0,92	0,97	0,88	0,94	<b>0,78</b>	0,00
50%	<b>0,77</b>	<b>0,80</b>	0,92	0,97	0,88	0,94	<b>0,76</b>	0,00

Tabela 4.11: Resultados médios para erosão 1 iteração + ruído sal-pimenta

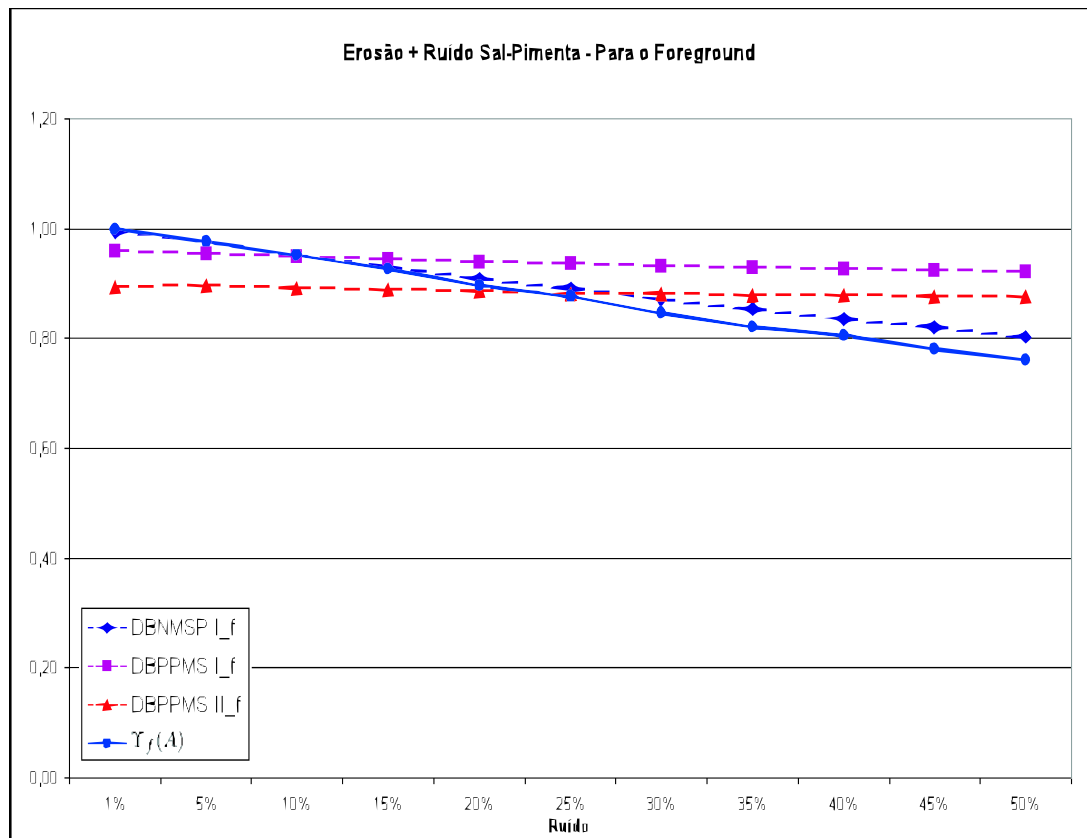


Figura 4.23: Gráfico referente a normalização dos valores do *foreground* para erosão 1 iteração + ruído sal-pimenta



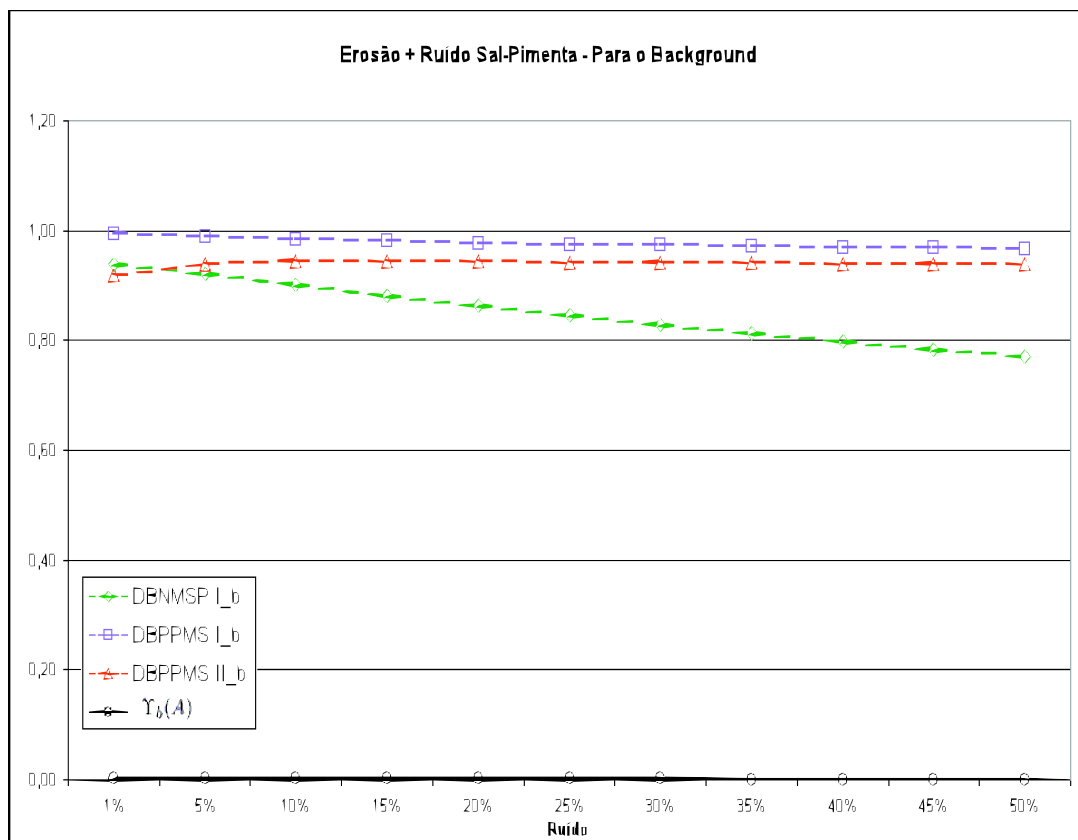


Figura 4.24: Gráfico referente a normalização dos valores do *background* para erosão 1 iteração + ruído sal-pimenta

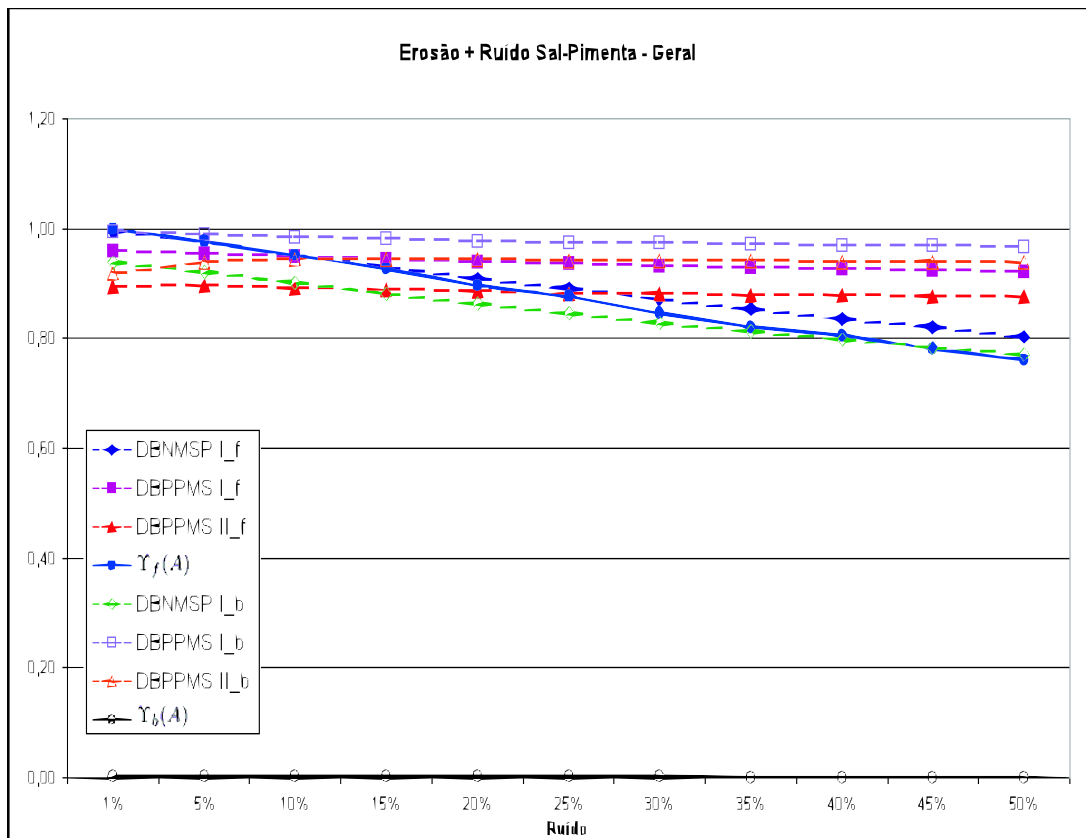


Figura 4.25: Gráfico referente a normalização dos valores gerais para erosão 1 iteração + ruído sal-pimenta

#### 4.2.2.7 Dilatação + ruído sal-pimenta

Os métodos  $\Upsilon_f(A)$  e  $\Upsilon_b(A)$ , têm se mostrado muito sensíveis ao decréscimo de informações e a presença de ruídos no *foreground* e *background* das imagens segmentadas, gerados através da dilatação binária, e pelo ruído sal-pimenta. Os métodos DBNMSP  $I_f$  e DBNMSP  $I_b$  têm se mostrado sensíveis. Já os demais métodos têm se mostrado pouco sensíveis, como pode-se observar na tabela 4.12 e nos gráficos 4.26, 4.27 e 4.28.

Dilatação+ Sal-Pimenta	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
1%	1,00	0,72	0,99	0,96	0,90	0,92	<b>0,73</b>	<b>1,00</b>
5%	0,98	0,72	0,97	0,96	0,90	0,94	<b>0,68</b>	<b>0,98</b>
10%	0,95	0,70	0,96	0,96	0,89	0,94	<b>0,65</b>	<b>0,96</b>
15%	0,93	0,70	0,96	0,96	0,89	0,94	<b>0,63</b>	<b>0,93</b>
20%	0,91	0,69	0,95	0,96	0,89	0,94	<b>0,60</b>	<b>0,91</b>
25%	0,89	0,68	0,95	0,96	0,88	0,94	<b>0,53</b>	<b>0,89</b>
30%	0,87	0,67	0,94	0,96	0,88	0,94	<b>0,50</b>	<b>0,87</b>
35%	0,85	0,66	0,94	0,96	0,88	0,94	<b>0,50</b>	<b>0,01</b>
40%	0,84	0,65	0,93	0,96	0,88	0,94	<b>0,44</b>	<b>0,00</b>
45%	0,82	0,65	0,93	0,96	0,88	0,94	<b>0,43</b>	<b>0,00</b>
50%	0,80	0,64	0,93	0,95	0,88	0,94	<b>0,37</b>	<b>0,00</b>

Tabela 4.12: Resultados médios para dilatação 1 iteração + ruído sal-pimenta

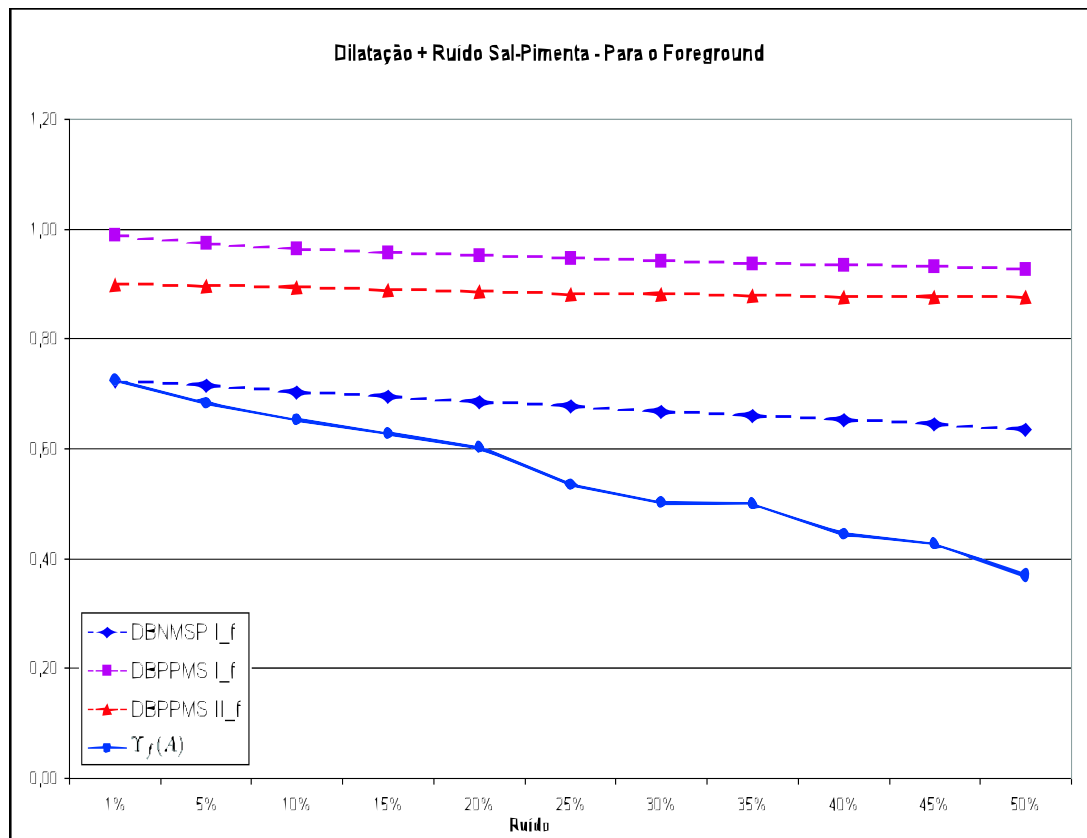


Figura 4.26: Gráfico referente a normalização dos valores do *foreground* para dilatação 1 iteração + ruído sal-pimenta

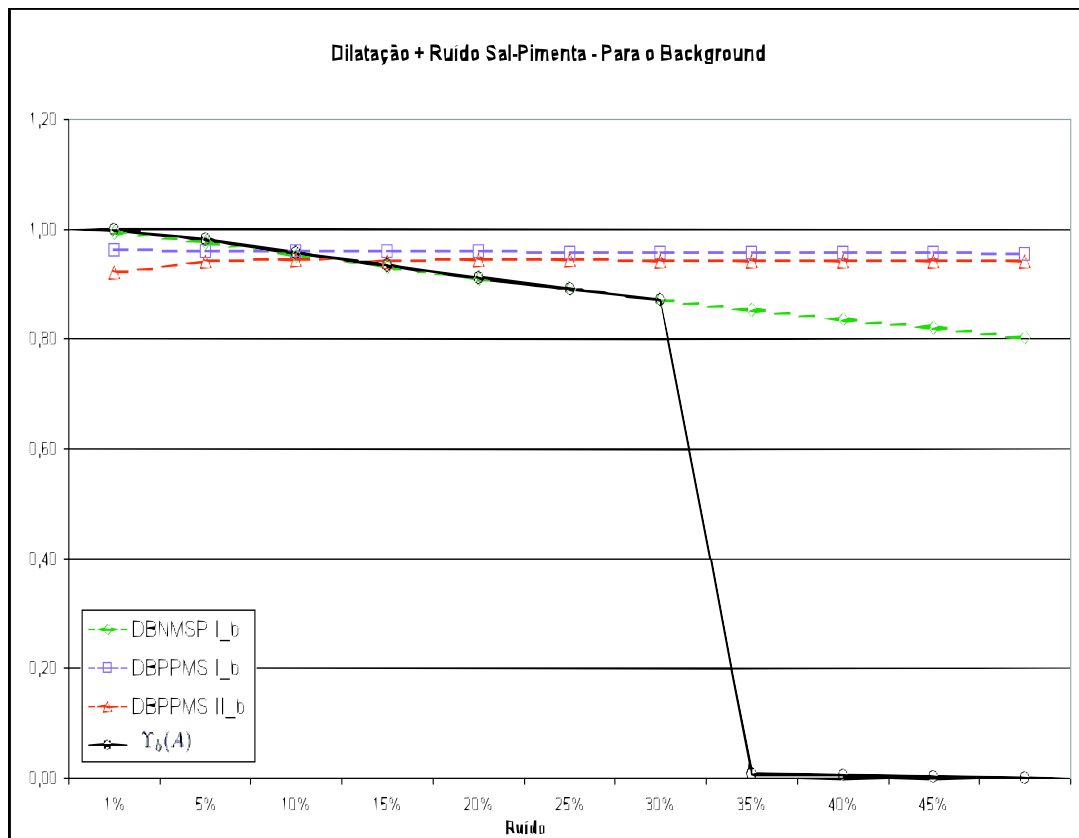


Figura 4.27: Gráfico referente a normalização dos valores do *background* para dilatação 1 iteração + ruído sal-pimenta

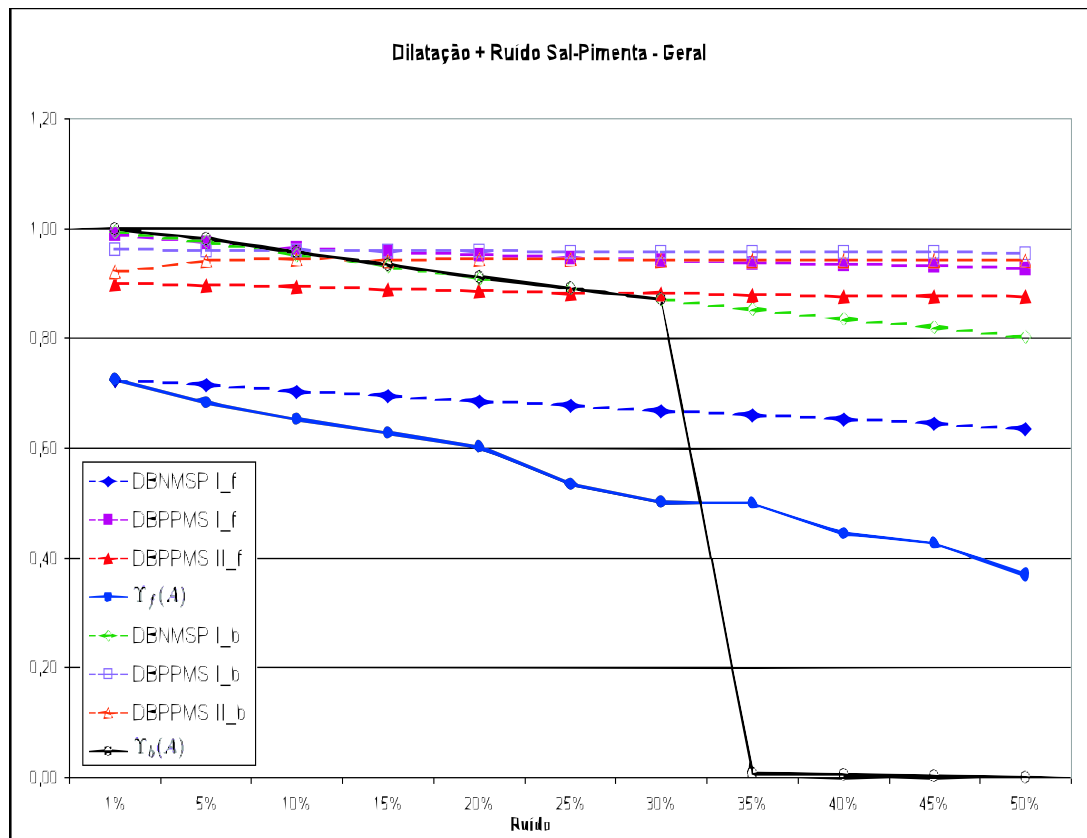


Figura 4.28: Gráfico referente a normalização dos valores gerais para dilatação 1 iteração + ruído sal-pimenta

### 4.2.3 Análise da eficiência e da sensibilidade da métrica proposta

Ao serem analisados os resultados da segmentação pela nova metodologia  $\Upsilon(A)$  pudemos perceber as seguintes características quanto aos tipos de classificação do método  $\Upsilon(A)$ , que são:

- Segmentação correta;
- Segmentação excessiva;
- Segmentação ausente;
- Segmentação insuficiente.

Entretanto é preciso identificar onde ocorrem estes tipos de segmentação nas imagens de testes.

No caso da erosão binária, onde ocorre um acréscimo de pixels do *foreground* poderá ocorrer a presença da segmentação insuficiente.

No caso da dilatação binária, onde ocorre um decréscimo de pixels do *foreground* poderá ocorrer a presença da segmentação excessiva ou da segmentação ausente.

No caso do ruído pimenta, onde ocorre acréscimo de pixels do *foreground* poderá ocorrer a presença da segmentação excessiva.

No caso do ruído sal, onde ocorre um acréscimo de pixels do *background* poderá ocorrer a presença da segmentação ausente.

- a métrica  $\Upsilon(A)$  penaliza a segmentação excessiva imediatamente a partir do momento que uma qualquer segmentação excessiva tenha ocorrido. Porém a penalização não muda significativamente, se a segmentação excessiva for maior, como se pode observar nos gráficos 4.8, 4.11, 4.14.
- a métrica  $\Upsilon(A)$  penaliza a segmentação insuficiente gradativamente mas sem excesso, como se pode observar nos gráficos 4.2, 4.5.
- a métrica  $\Upsilon(A)$  penaliza muito a segmentação ausente. Basta a segmentação ser considerada ausente para um pedaço pequeno da imagem, que a métrica de classificação cai imediatamente. À medida que mais áreas da imagem são

classificadas como ausente a penalização torna-se mais severa, como se pode observar nos gráficos 4.11, 4.17.

A métrica  $\Upsilon(A)$  avalia a segmentação de forma severa, utilizando critérios qualitativos e quantitativos.

No intuito de comparar mais facilmente a sensibilidade dos métodos de avaliação de segmentação perante a simulação de ruídos em imagens artificiais, a tabela de comparação 4.13 foi montada qualificando quais são os métodos muito sensíveis, sensíveis ou pouco sensíveis. Esta classificação será dada de acordo com as faixas de variação:

- De 0% a 10% – Pouco sensível;
- de 11% a 50% – Sensível e
- de 51% a 100% – Muito sensível

	DBNPMS Type $I_b$	DBNPMS Type $I_f$	DBPPMS $I_f$	DBPPMS $I_b$	DBPPMS $II_f$	DBPPMS $II_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
Erosão	Sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Sensível	Pouco sensível
Dilatação	Pouco sensível	Muito sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Muito sensível	Pouco sensível
Pimenta	Sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Muito sensível
Sal	Pouco sensível	Sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Muito sensível	Pouco sensível
Sal-Pimenta	Sensível	Sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Sensível	Muito sensível
Erosão+ Sal-Pimenta	Sensível	Sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Sensível	Pouco sensível
Dilatação+ Sal-Pimenta	Sensível	Sensível	Pouco sensível	Pouco sensível	Pouco sensível	Pouco sensível	Muito sensível	Muito sensível

Tabela 4.13: Tabela de sensibilidade entre metodologias de avaliação e ruídos



#### 4.2.4 Avaliação de qualidade sobre resultados obtidos de processos reais de segmentação

Conforme visto no início deste capítulo, a segunda abordagem de avaliação da qualidade de segmentação consiste em utilizar algoritmos reais de segmentação de imagens de envelopes postais. Para isso serão utilizadas 118 imagens reais de envelopes postais brasileiros com suas respectivas imagens *ground truth*. Os algoritmos que serão utilizados são:

- Segmentação *Wavelet* de envelopes postais de [MFBJ03] e
- Segmentação de envelopes postais de [EFM04] utilizando uma abordagem fractal.

A avaliação dos algoritmos de segmentação de envelopes postais citados, acima, será feita pelas metodologias que foram consideradas como sensíveis ou muito sensíveis a pelo menos um tipo de ruído na tabela 4.13. Nas tabelas 4.14 e 4.15 serão apresentadas as médias entre as 118 imagens para cada metodologia.

Média de [MFBJ03]	DBNPMS Type $I_f$	DBNPMS Type $I_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
	0,64	0,99	0,40	0,14

Tabela 4.14: Média das medidas de avaliação de segmentação para a segmentação *Wavelet* de envelopes postais de [MFBJ03]

Média de [EFM04]	DBNPMS Type $I_f$	DBNPMS Type $I_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
	0,87	0,92	0,38	0,02

Tabela 4.15: Média das medidas de avaliação de segmentação para a segmentação de envelopes postais de [EFM04] utilizando uma abordagem fractal

Pode-se observar que na média das avaliações de segmentações reais a média da nova metodologia de avaliação da segmentação é muito baixa se comparada com a metodologia DBNPMS. Isso deve-se ao fato de as imagens *ground truth* de envelopes postais que contém selos terem os selos preenchidos e de a nova metodologia ter grande sensibilidade ao ruídos Sal, como pode-se observar na imagem 4.29.

Desta forma se fizermos a média separada das imagens que contém selos da imagens que não contém, teremos uma avaliação mas próxima da realidade, como pode-se observar nas tabelas 4.16 e 4.17.

Média de [MFBJ03]	DBNPMS Type $I_f$	DBNPMS Type $I_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
imagens contendo selos	0,55	1,00	0,28	0,16
imagens sem selos	0,98	0,99	0,83	0,02

Tabela 4.16: Média das metodologias de avaliação de segmentação para a segmentação *Wavelet* de envelopes postais de [MFBJ03]

Média de [EFM04]	DBNPMS Type $I_f$	DBNPMS Type $I_b$	$\Upsilon_f(A)$	$\Upsilon_b(A)$
imagem contendo selos	0,85	0,91	0,32	0,02
imagem sem selos	0,97	0,94	0,61	0,02

Tabela 4.17: Média das metodologias de avaliação de segmentação para a segmentação de envelopes postais de [EFM04] utilizando uma abordagem fractal

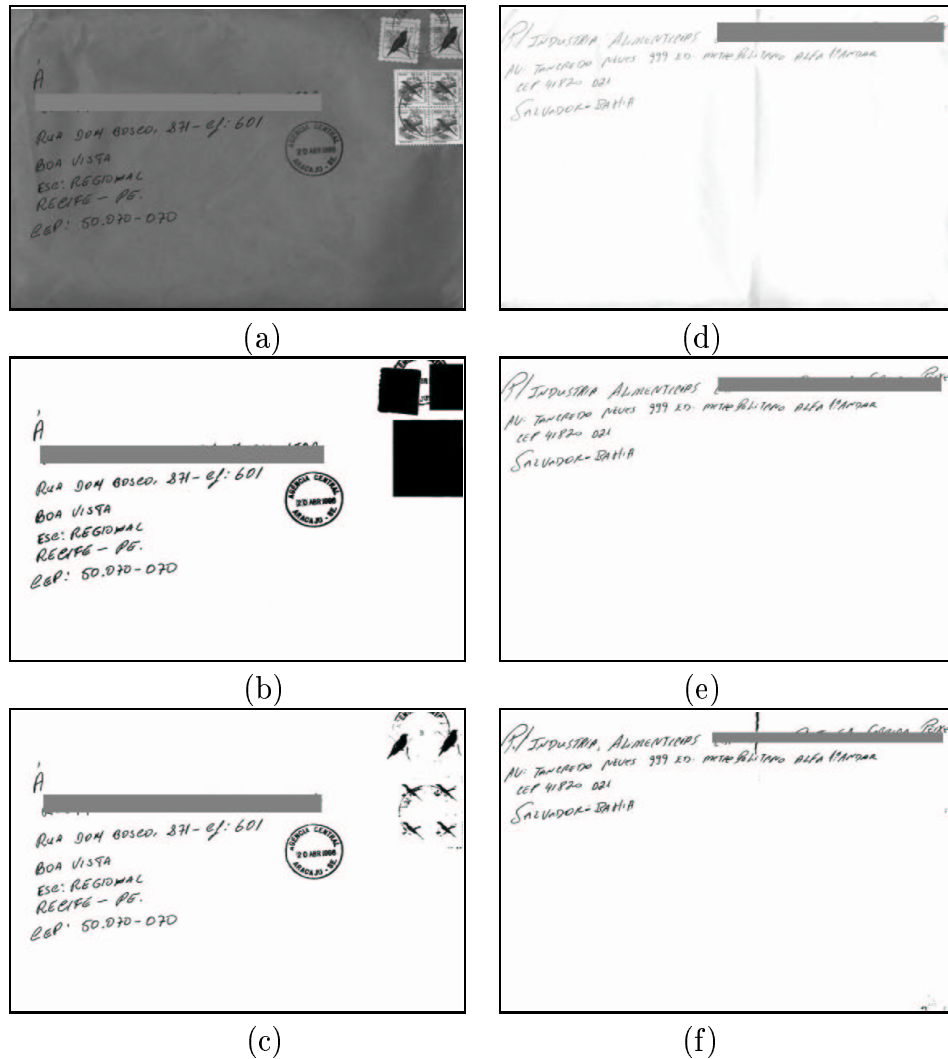


Figura 4.29: Exemplos da segmentação do bloco de endereços: (a) Imagem original No 1 contendo selo, (b) Imagem *ground truth*, (c) Resultado da segmentação, (d) Imagem original No 2 que não contém selo, (e) Imagem *ground truth*, (f) Resultado da segmentação.

Como se pode observar através das tabelas acima e na figura 4.30 a segmentação *Wavelet* de envelopes postais de [MFBJ03] teve resultados melhores que a segmentação de envelopes postais de [EFM04] utilizando uma abordagem fractal.

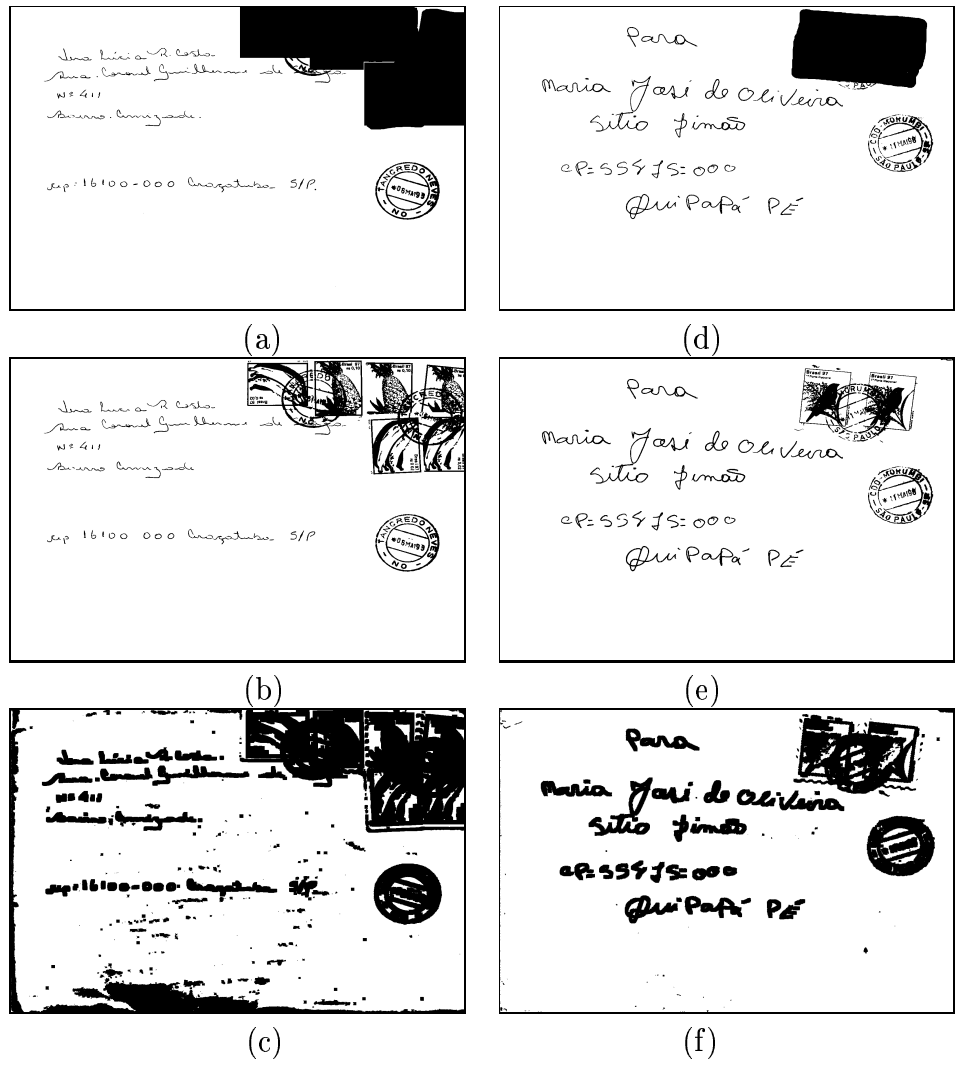


Figura 4.30: Resultado da segmentação de envelopes postais. (a) Imagem *ground truth* do envelope 028, (b) Imagem *ground truth* do envelope 075, (c) resultado da segmentação do envelope 028 pelo método de [MFBJ03], (d) resultado da segmentação do envelope 075 pelo método de [MFBJ03], (e) resultado da segmentação do envelope 028 pelo método de [EFM04], (f) resultado da segmentação do envelope 075 pelo método de [EFM04].



## Capítulo 5

# Conclusões e Trabalhos Futuros

Este trabalho apresentou uma metodologia para avaliar objetivamente e quantitativamente mecanismos de segmentação por região de imagens. Os desafios da presente pesquisa foram:

- propor um novo mecanismo de avaliação de discrepância, capaz de avaliar vários cenários de má segmentação;
- propor um mecanismo de geração de imagens sintetizadas que baseou-se no uso de ruídos sal e pimenta e de deformações geradas por erosão e dilatação. Procurou-se mesclar esses diversos processos para simular diversos cenários que podem ocorrer durante um processo real de segmentação;
- comparar o novo mecanismo de avaliação de discrepância com os existentes na literatura.

O novo mecanismo de avaliação de discrepância proposto buscou quantificar 4 cenários de segmentação: correta, excessiva, insuficiente ou ainda ausente. Também foi proposta uma métrica geral de avaliação  $\Upsilon()$  cujo papel foi de unificar as avaliações obtidas pela quantificação anterior.

A estratégia de comparação foi realizada de duas maneiras: numa primeira etapa, tanto os mecanismos de avaliação de discrepância tradicionais implementados quanto o novo mecanismo foram testados nessa base de imagens sintetizadas. Numa segunda etapa, os mecanismos que foram classificados como sensível ou muito sensível em pelo menos um dos diversos cenários de imagens sintetizadas (ver 4.13),

foram aplicados em dois métodos reais de segmentação de envelopes postais, baseado em segmentação *Wavelet* de [MFBJ03] e em dimensão fractal de [EFM04].

A conclusão geral dessa pesquisa é que os objetivos principais de propor um novo mecanismo de avaliação de discrepância, capaz de avaliar vários cenários de má segmentação; propor um mecanismo de geração de imagens sintetizadas refletindo os diversos cenários que podem ocorrer em uma segmentação; e de comparar o novo mecanismo; foram atingidos e apresentaram resultados interessantes:

- a geração da base de imagens sintetizadas simulando diversos cenários de segmentação, foi de suma importância para que se possa avaliar quantitativamente a influência das perturbações e deformações que podem ocorrer num processo de segmentação;
- o novo critério de avaliação de segmentação de imagens, apesar de sua simplicidade, atendeu aos requisitos de avaliação objetiva e qualitativa. Em diversos cenários de má classificação, mostrou-se muito mais sensível ou tão sensível quanto a maioria dos outros critérios de segmentação (tabela 4.13). Ainda pode-se verificar, no caso de segmentação real de envelopes postais, que ele foi mais severo e mais criterioso na sua avaliação (tabelas 4.14, 4.15, 4.16 e 4.17);
- outra vantagem é que, ao contrário do estudo de [RLA<sup>+</sup>01], que exclui cenários de segmentação ruim, não há nenhuma restrição quanto a aplicação do novo critério de avaliação.
- uma consequência direta das comparações realizadas é que na avaliação de métodos reais de segmentação, pode-se verificar que o método de segmentação *Wavelet* de [MFBJ03] é mais eficiente que o método de segmentação utilizando uma abordagem fractal de [EFM04].

Nesta área de pesquisa os trabalhos futuros propostos são:

- aprimorar o novo critério de avaliação de discrepância englobando mais cenários de segmentação;
- comparar o novo mecanismo de avaliação de discrepância com os mais recentes publicados na literatura;



- prever e estender o novo critério de avaliação de discrepância para outros tipos de segmentação (por contorno ou por textura).

# Referências

- [And00] S. Ando. Consistent gradient operators. *IEEE Trans.on Pattern Analysis and Machine Intelligence*, PAMI-22(3):252–264, March 2000.
- [Bal81] D. H Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.
- [Ber86] J. Bernsen. Dynamic thresholding of gray-level images. *Proc. Eighth Int'l Conf. Conf. on Pattern Recognition, Paris, France*, pages 1251–1255, 1986.
- [BL94] M. Batty and P . Longley. *Fractal Cities*. Academic Press, 1994.
- [CA84] C.H. Chien and J.K. Aggarwal. A normalized quadtree representation. *Computer Vision, Graphics and Image Processing*, 26:331–346, 1984.
- [CH80] R.W. Connors and C.A. Harlow. A theoretical comparison of texture algorithms. *IEEE Trans.on Pattern Analysis and Machine Intelligence*, PAMI-2(3):204–222, 1980.
- [Cla86] K.C. Clarke. Computation of the fractal dimension of topographic surfaces using the triangular prism surface area method. *Computers and Geosciences*, 12(5):713–722, 1986.
- [CMVM86] F. Cheevasuvit, H. Maitre, and D. Vidal-Madjar. A robust method for picture segmentation based on a split and merge procedure. *Computer Vision, Graphics and Image Processing*, 34:268–281, 1986.

- [CSK93] B.B. Chaudhuri, N. Sarkar, and P. Kundu. Improved fractal geometry based texture segmentation technique. *IEEE Proceedings*, 140:233–241, 1993.
- [DQRC<sup>+</sup>89] B. Dubuc, J.F. Quiniuo, C. Roques-Carnes, C. Tricot, and S.W. Zucker. Evaluating the fractal dimensions of profile. *Phys. Rev.*, 39:1500–1512, 1989.
- [EFM04] L. F. Eiterer, J. Facon, and D. Menoti. Fractal-based approach for segmentation of address block in postal envelopes. *IX CIARP 2004 (9TH IBEROAMERICAN CONGRESS ON PATTERN RECOGNITION), 2004, Puebla - Mexico. LNCS Lecture Notes in Computer Science. Puebla -Mexico: Springer-Verlag - LNCS Lecture Notes in Computer Science*, 1, 2004.
- [ETM91] L. Eikvil, T. Taxt, and K. Moen. A fast adaptive method for binarization of document images. *Proc First Int 'l Conf Document Analysis and Recognition, Saint Malo, France*, pages 435–443, 1991.
- [Fac01] Jacques Facon. Metodologia de avaliação de abordagens de segmentação de imagens. Technical report, Universidade Pontifícia Universidade Católica do Paraná, PUCPR, Curitiba-Brazil, October 2001.
- [GBGP94] H. Greenspan, S. Belongie, R. Goodman, and P. Perona. Rotation invariant texture recognition using a steerable pyramid. *IEEE International Conference on Pattern Recognition*, pages 162–167, October 1994.
- [GW92] Rafael C. Gonzales and Richard E. Woods. *Processamento Digital de Imagens*. Editora Edgard Blücher Ltda., 1992.
- [HJBJ<sup>+</sup>96] A. Hoover, G. Jean-Baptiste, X. Jiang, P.J. Flynn, H. Bunke, D.B. Goldof, K. Bowyer, D.W. Eggert, A. Fitzgibbon, and R.B. Fisher. An experimental comparison of range image segmentation algorithms. *IEEE Trans.on Pattern Analysis and Machine Intelligence*, PAMI-18(7):673–689, July 1996.

- [KPKPR99] B. Kangar-Parsi, B. Kangar-Parsi, and A. Rosenfeld. Optimally isotropic laplacian operator. *IEEE Transactions on Image Processing*, 10(8):1467–1472, October 1999.
- [LL93] C.H. Li and C.K. Lee. Minimum cross entropy thresholding. *Pattern Recognition*, 26(4):616–626, 1993.
- [Mal89] S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 11, número7, julho de 1989:674–693, 1989.
- [Man83] B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman And Company, New York, 1983.
- [MFBJ03] D. Menoti, J. Facon, D. L. Borges, and A. S. Brito Júnior. Segmentation of postal envelopes for address block location: an approach based on feature selection in wavelet. *ICDAR 2003-7th International Conference on Document Analysis and Recognition. Los Alamitos, California, USA: IEEE Computer Society*, 2:699–703, 2003.
- [MM96] W. Y. Ma and B. S. Manjunath. Texture feature and learning similarity. *Computer Vision and Pattern Recognition*, pages 425–430, June 1996.
- [Mor87] H.P. Moravec. Towards automatic visual obstacle avoidance. *Proc. 5th Int. Joint. Conf. Artificial Intelligence, Cambridge, MA*, page 584, August 1987.
- [Ots79] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, SMC 9(1):62–66, April 1979.
- [PaRHA84] S. Peleg, J. Naor and R. Hartley, and D. Avnir. Multiple resolution texture analysis and classification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-6(4):518–523, 1984.

- [Pen84] A.P. Pentland. Fractal based description of natural scenes. *IEEE Trans.on Pattern Analysis and Machine Intelligence*, PAMI-6(6):661–672, November 1984.
- [PTVF96] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B. Flannery. *Numerical Recipes in C: the art of scientific computing - Second Edition*. Cambridge University Press - Reino Unido, 1996.
- [Rin03] A. Ringler. Texture segmentation by local fractal dimension as applied to oceanic search and rescue. *IEEE ICICS, International Conference on Information Communications Signal Processing*, 2:975–979, 2003.
- [RLA<sup>+</sup>01] R.R. Roldán, J.F.G. Lopera, C.A. Allah, J.M. Aroza, and P.L.L. Escamilla. A measure of quality for evaluating methods of segmentation and edge detection. *Pattern Recognition*, 34:969–980, 2001.
- [Rut01] G.A. Rutledge. Wavelets in remote sensing - texture scale mapping. Technical report, Department of Mechanical Engineering - University of Victoria - British Columbia, <http://www.geog.uvic.ca/geog322/htmls/index.html>, 2001.
- [SN96] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, EUA, 1996.
- [SS01] Mehmet Sezgin and Bülent Sankur. Selection of thresholding methods for non-desctructive testing applications. *IEEE Computer Society Press*, March 2001.
- [Uns95] M. Unser. Texture classification and segmentation using wavelets frames. *IEEE Trans.on Image Processing*, 4(11):1549–1560, November 1995.
- [Voo95] Hubert F.J.M. Voogd. *Error Measurement for Segmentation Techniques*. Master Thesis of Katholieke Universiteit Nijmegen, 1995.
- [Wel93] P.D Wellner. Adaptative thresholding for the digitaldesk. *Technical Report EPC-1993-110, Rank Xerox Ltd*, July 1993.

- [YALG95] L. Yang, F. Albrechtsen, T. Lonnestad, and P. Grottum. A supervised approach to the evaluation of image segmentation methods. *Proceedings of CAIP-Prague- Computer Analysis of Images and Patterns*, pages 759–765, 1995.
- [YCC95] J.C. Yen, F.J. Chang, and S. Chang. A new criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*, 4(3):370–378, March 1995.
- [YMB77] W.A. Yasnoff, J. K. Mui, and J. W. Bacus. Error measures for scene segmentation. *Pattern Recognition*, 9:217–231, 1977.
- [Zha96] Y.J. Zhang. A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8):1335–1346, 1996.