

Luiz Felipe Marco Eiterer

**Uma Abordagem Fractal
para Localização do Bloco de Endereço
em Segmentação de Envelopes Postais**

Dissertação apresentada ao Programa de Pós-Graduação em
Informática Aplicada da Pontifícia Universidade Católica do
Paraná como requisito parcial para obtenção do título de
Mestre em Informática Aplicada

Curitiba

2005

Luiz Felipe Marco Eiterer

**Uma Abordagem Fractal
para Localização do Bloco de Endereço
em Segmentação de Envelopes Postais**

Dissertação apresentada ao Programa de Pós-Graduação em
Informática Aplicada da Pontifícia Universidade Católica do
Paraná como requisito parcial para obtenção do título de
Mestre em Informática Aplicada

Área de Concentração : Ciência da Imagem

Orientador: Prof. Dr. Jacques Facon

Curitiba

2005

Eiterer, Luiz Felipe Marco

Uma Abordagem Fractal para Localização do Bloco de Endereço em Segmentação de Envelopes Postais. Curitiba, 2005. 93p.

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

1. Segmentação. 2. Envelopes Postais. 3. Dimensão Fractal. 4. Bloco Endereço. I. Pontifícia Universidade Católica do Paraná. Centro de Ciências Exatas e de Tecnologia. Programa de Pós-Graduação em Informática Aplicada.

Ao João, meu filho.

Agradecimentos

A Deus, *ad aeternum*.

A Maristela, pelo impulso.

Pela amizade e apoio, minha mãe Aparecida, aos meus irmãos Maria Elisa e Jose Alberto e cunhados Rafael e Maria Luiza.

Ao meu pai, Professor Eiterer, pelos retos valores.

Um agradecimento especial ao Professor Orientador Jacques Facon, pela amizade, paciência, orientação e apóio.

Ao doutorando David Menoti, solidário a todos, pela amizade.

Ao Professor Carlos Maziero, pela possibilidade deste projeto no Programa de Pos Graduação em Informática Aplicada.

Ao colegas, Professores e Funcionários do PPGIA.

Sumário

Capítulo 1

Considerações Gerais

1.1. Introdução	1
1.2. Objetivos	3
1.3. Justificativas	3

Capítulo 2

Teoria Fractal

2.1. Introdução ao Caos	4
2.2. Teoria Fractal	7
2.2.1 Teoria Fractal de Durer	7
2.2.2 Teoria Fractal de Cantor	8
2.2.3 Teoria Fractal de Peano, Hilbert e von Kock	9
2.2.4 Teoria Fractal de Sierpinski	9
2.2.5 Teoria Fractal de Julia e Fatou	10
2.2.6 Teoria Fractal de Mandelbrot	12

Capítulo 3

Segmentação por Teoria Fractal

3.1. Segmentação	13
3.2. Dimensão Fractal	15
3.3. Definição da Dimensão Fractal	18
3.4. Cálculo Prático da Dimensão Fractal	18
3.4.1 Técnicas de <i>Box-Counting</i>	19
3.4.2 Coeficiente de Hurst	23
3.4.3 Técnicas Morfológicas	25

Capítulo 4

Aglomerção: Reconhecimento de Padrões não supervisionado

4.1. Introdução	29
-----------------	----

4.2. Algoritmo <i>K-means</i>	34
4.3. Algoritmo <i>Isodata</i>	36
 Capítulo 5	
Método Proposto	
5.1. Método Proposto: uma abordagem fractal	41
5.2 Cálculo da Dimensão Fractal	42
5.3 Algoritmo de Aglomeração <i>K-means</i>	42
5.4. Matriz de Co-ocorrência	43
5.5 Identificação das Classes	45
 Capítulo 6	
Testes, Resultados e Análises	
6.1. Base de Imagens	47
6.2. Imagem de Referência <i>ground truth</i>	48
6.3 Avaliação Metodologia de Segmentação	49
6.4. Resultados e Análises	52
6.4.1 Análise dos Resultados da imagem Carta040 da Base Real	54
6.4.2 Análise dos Resultados para a Base de Imagens Reais	64
6.4.3 Análise dos Resultados da imagem Carta026fundo12 da Base Sintetizada	67
6.4.4 Análise dos Resultados para a Base de Imagens Sintetizadas	74
6.4.5 Comentários Finais	76
 Capítulo 7	
Conclusões e Trabalhos futuros	78
 Referências Bibliográficas	80
 Anexos	
A. Imagem Real carta040 com variações de parâmetro	84
B. Variedades de fundos de envelopes complexos	96

Lista de Figuras

Figura 2.1	Botafumeiro da Catedral de Santiago de Compostela	5
Figura 2.2	Pentágono de Durer	8
Figura 2.3	Triângulo ABC do pentágono de Durer	8
Figura 2.4	Pentágono de Durer depois de 4 sub divisões	8
Figura 2.5	Pente de Cantor depois de 4 sub divisões	9
Figura 2.6	Curva de Kock após 4 gerações	9
Figura 2.7	Triângulo de Sierpinski	10
Figura 2.8	Conjunto de Julia conectado	11
Figura 3.1	Várias Texturas	15
Figura 3.2	2D Variation procedure para vários r	19
Figura 3.3	Sinal fBm	28
Figura 4.1	Ilustração de medida de similaridade	32
Figura 4.2	Evolução do processo de <i>aglomeração</i> em função do número de aglomerados	36
Figura 5.1	Esquema da Abordagem de Segmentação Fractal	41
Figura 5.2	Imagem de dimensões 4 x 4 e níveis de cinza correspondentes	43
Figura 5.3	Resultados da matriz de co-ocorrência para diversos ângulos e distancia 1 para o exemplo da imagem da figura 5.2	44
Figura 6.1	Imagem de envelope postal da base real fundo complexo e branco	47
Figura 6.2	Imagem de envelope postal da base sintetizada	48
Figura 6.3	Imagem de referência <i>ground truth</i>	49
Figura 6.4	Esquema de avaliação das imagens de envelopes postais	51
Figura 6.8	Imagem de envelope postal da base real	53
Figura 6.9	Imagem de envelope postal da base real [3,5]	56
Figura 6.10	Imagem de envelope postal da base real [3,7]	57
Figura 6.11	Imagem de envelope postal da base real [3,9]	58
Figura 6.12	Imagem de envelope postal da base real [3,11]	60
Figura 6.13	Imagem de envelope postal da base real [3,13]	61

Figura 6.14	Imagem de envelope postal da base real [3,15]	62
Figura 6.15	Gráfico das Taxas de Segmentação da Imagem Real carta040	64
Figura 6.16	Gráfico dos Valores Médios das Taxas de Segmentação da Base Real	66
Figura 6.17	Imagem da Base Sintetizada e sua correspondente <i>ground truth</i>	67
Figura 6.18	Imagem da Base Sintetizada com parâmetros [3,5]	68
Figura 6.19	Imagem da Base Sintetizada com parâmetros [3,7]	69
Figura 6.20	Imagem da Base Sintetizada com parâmetros [3,9]	70
Figura 6.21	Imagem da Base Sintetizada com parâmetros [3,11]	71
Figura 6.22	Imagem da Base Sintetizada com parâmetros [3,13]	72
Figura 6.23	Imagem da Base Sintetizada com parâmetros [3,15]	73
Figura 6.24	Gráfico dos Valores Médios das taxas de segmentação das imagens da Base Sintetizada	75
Figura A.1	Imagem Envelopes Postal carta040 com parâmetros [3,5]c2	84
Figura A.2	Imagem Envelopes Postal carta040 com parâmetros [3,5]i001.....	85
Figura A.3	Imagem Envelopes Postal carta040 com parâmetros [3,7]c2	86
Figura A.4	Imagem Envelopes Postal carta040 com parâmetros [3,7]i001.....	87
Figura A.5	Imagem Envelopes Postal carta040 com parâmetros [3,9]c2	88
Figura A.6	Imagem Envelopes Postal carta040 com parâmetros [3,9]i001	89
Figura A.7	Imagem Envelopes Postal carta040 com parâmetros [3,11]c2	90
Figura A.8	Imagem Envelopes Postal carta040 com parâmetros [3,11]i001	91
Figura A.9	Imagem Envelopes Postal carta040 com parâmetros [3,13]c2	92
Figura A.10	Imagem Envelopes Postal carta040 com parâmetros [3,13]i001.....	93
Figura A.11	Imagem Envelopes Postal carta040 com parâmetros [3,15]c2	94
Figura A.12	Imagem Envelopes Postal carta040 com parâmetros [3,15]i001	95
Figura B.1	Variedades de fundos de objetos complexos	96

Lista de Tabelas

Tabela 3.1	Dados do Coeficiente de Hurst	24
Tabela 3.2	Dados do Cálculo do Coeficiente de Hurst	25
Tabela 6.1	Exemplo Hipotético para avaliação da Metodologia de Segmentação	52
Tabela 6.2	Taxas de Segmentação dos testes do envelope postal carta040 da base Real com intervalo de tamanho de janelas [3,5] e variação nos parâmetros do <i>K-means</i>	55
Tabela 6.3	Taxas de Segmentação dos testes do envelope postal carta040 da base Real com intervalo de tamanho de janelas [3,7] e variação nos parâmetros do <i>K-means</i>	57
Tabela 6.4	Taxas de Segmentação dos testes do envelope postal carta040 da base Real com intervalo de tamanho de janelas [3,9] e variação nos parâmetros do <i>K-means</i>	59
Tabela 6.5	Taxas de Segmentação dos testes do envelope postal carta040 da base Real com intervalo de tamanho de janelas [3,11] e variação nos parâmetros do <i>K-means</i>	59
Tabela 6.6	Taxas de Segmentação dos testes do envelope postal carta040 da base Real com intervalo de tamanho de janelas [3,13] e variação nos parâmetros do <i>K-means</i>	61
Tabela 6.7	Taxas de Segmentação dos testes do envelope postal carta040 da base Real com intervalo de tamanho de janelas [3,15] e variação nos parâmetros do <i>K-means</i>	63
Tabela 6.8	Valores médios das Taxas de Segmentação dos testes da base de imagens Reais	65
Tabela 6.9	Valores médios das Taxas de Segmentação dos testes da base de imagens Sintetizadas	75
Tabela 6.10	Quadro comparativo de resultados obtidos por varias abordagens	77

Lista de Simbolos

fBm	<i>movimento Browniano fracionário</i>
DF	Dimensão Fractal
DB	Dimensão <i>box counting</i>
CB	Correios Brasileiros
PUCPR	Pontifícia Universidade Católica do Paraná
DH	Dimensão Fractal de Hausdorff
DT	Dimensão Topológica
DM	Dimensão de Minkowski-Bouligand

Resumo

Com o aumento do desempenho e da eficiência dos equipamentos atuais de computação, os desafios de informatizar e modernizar a Automação Postal estão se tornando cada vez mais uma realidade. As ferramentas computacionais herdadas da Área de Processamento e Análise de Imagens permitem adquirir, armazenar e executar aplicativos com alto grau de complexidade. A seguir, é descrita uma abordagem baseada na dimensão fractal para segmentação de envelopes postais no intuito de localizar automaticamente o bloco-endereço. Busca-se separar e diferenciar automaticamente do envelope postal o seu fundo, o bloco-endereço, o(s) carimbo(s) e os selo(s). Inicialmente é atribuída a cada *pixel* da imagem, a dimensão fractal calculada na sua vizinhança. Em seguida, o algoritmo de aglomeração *K-means* divide a imagem em três grupos, a saber: o fundo do envelope, o ruído e a parte da imagem que contem as informações semânticas relevantes (bloco-endereço, carimbo(s) e selo(s)). A identificação de cada grupo faz-se estatisticamente empregando a matriz de co-ocorrência e certas heurísticas. A avaliação quantitativa da segmentação é realizada comparando o resultado obtido para cada envelope com o resultado ideal (estratégia *ground truth*). A abordagem proposta foi testada sobre duas bases de imagens de envelopes postais - uma real e outra sintetizada – que reúnem vários tipos de fundos e *lay-outs*. Os resultados obtidos mostraram-se promissores com uma taxa de acerto de 96% no bloco de endereço com menos de 7% de ruído. Conclusões e sugestão de futuras pesquisas são apresentadas ao final deste trabalho.

Palavras-chaves: 1. Segmentação. 2. Envelopes Postais. 3. Dimensão Fractal. 4. Bloco Endereço

Abstract

With the increase of the computation equipment performance and efficiency, the challenges of automatizing and modernizing the Postal Automation are becoming each time more realistic. The computational tools inherited of the Area of Processing and Analysis of Images allow to acquire, to store and to execute applications with high degree of complexity. An approach based on the fractal dimension to automatically segment the block-address in postal envelopes is described. This approach aims to automatically separate and differentiate the block-address, and stamps and postmarks in postal envelopes. Initially, the fractal dimension computed in pixel-centered windows is attributed to each pixel of the image. After that, the K-means algorithm divides the fractal dimension image in three clusters: the background one, the noise one and the block-address, stamp, postmark one. The identification of each cluster is based on co-occurrence matrix. The quantitative evaluation of the segmentation is carried through comparing the result obtained for each envelope with the ideal result (strategy ground truth). This approach was tested on two postal envelope image databases - one Real and other synthesized - that aggregate some types of deep and layouts. The obtained results have shown that this approach is promising with an accuracy rate of 96% for the address-block with less than 7% of noise. Conclusions and future research are presented at the end of this work.

Key-words: 1. Segmentation. 2. Postal Envelope. 3. Fractal Dimension. 4. Address Block.

Capítulo 1

Considerações Iniciais

1.1 Introdução

Desde os desenvolvimentos na primeira metade do século 20, a mecanização postal cresceu muito nos anos sessenta e setenta, primeiramente solicitada pelas iniciativas das administrações postais de diferentes nações. A necessidade de padrões comuns na distribuição do correio forçou a cooperação internacional e mais recentemente, a remoção de barreiras nacionais e abriu uma competição em todo o mundo envolvendo as companhias mais qualificadas em Tecnologia de Informação. Desde a metade dos anos setenta, o uso crescente dos faxes e da transferência de dados, e mais recentemente do e-mail, conduziu às predições que dentro de 20 anos relativamente poucos povos se comunicarão por carta. Apesar de tais predições, o volume do correio cresceu 200 por cento nos Estados Unidos, para alcançar um nível de aproximadamente 177 bilhão de pecas por ano.

Com um volume diário de mais de 34 milhões de correspondências, a Automação Postal dos Correios Brasileiros (CB) representa hoje um enorme desafio. Existem máquinas para processar e despachar essas correspondências, que podem atingir a taxa de 36.000 cartas por hora. Porém, esse tipo de automação exige que as cartas tenham um esquema padronizado, estabelecido pelos CB, isto é, que tenham o mesmo tamanho, diagramação padronizada dos dados (bloco de endereço, selo), que o fundo do envelope tenha um contraste adequado para adquirir as imagens e que sejam usadas certos tipos de fontes. Todas as outras correspondências que não obedecem a estes padrões são processadas manualmente. O que representa um enorme volume a ser despachado diariamente. Para o ser humano, esta atividade é extremamente monótona e aflitiva pela repetição, o que justifica a necessidade de desenvolver um sistema automático que seja capaz de processar e despachar envelopes postais de “layouts” e padronizações diversas.

Esta dissertação tem como foco a segmentação dos objetos semânticos relevantes de envelopes postais que não sigam necessariamente os formatos padronizados dos CB. Apesar da existência por parte dos CB de padrões de diagramação de um envelope postal do bloco de endereço, no que diz respeito à posição do nome do destinatário, da rua ou da avenida com número e complemento, bairro, do CEP (Código de Endereçamento Postal), da cidade e do estado, a população Brasileira, de uma maneira geral, não as conhece e/ou não as respeita. O que gera, um volume diário de mais de milhões de correspondências sem padronização. Acrescentando uma grande variabilidade de tamanho e de fundo, reafirma-se a necessidade de desenvolver um sistema robusto de localização automática do bloco-endereço independente da diagramação, do tipo de envelope postal, do selo e carimbo.

Nesta dissertação é apresentada uma nova abordagem de segmentação de envelopes postais brasileiros manuscritos, baseada num mecanismo de classificação por aglomeração da Dimensão Fractal dos diversos conjuntos de informações semânticas relevantes de um envelope (bloco-endereço, carimbo(s) e selo(s)). Objetiva-se localizar e segmentar o bloco de endereço manuscrito, do carimbo(s) e selo(s) sem utilizar de antemão do conhecimento sobre os envelopes postais. A avaliação quantitativa da precisão da abordagem proposta é realizada comparando o resultado obtido para cada envelope com o resultado ideal (estratégia *ground truth*).

Para segmentação de imagens naturais a textura é considerada um descritor regional importante. Tendo o envelope postal uma composição caótica, utilizou-se a Geometria Fractal que determina a rugosidade de uma textura através da Dimensão Fractal. Neste caso foi atribuído a um pixel a Dimensão Fractal da região(janelas) que o envolve , agindo como um filtro. Outro fator importante é a sua simplicidade e sua eficiência.

As bases de imagens usadas são oriundas de um convênio entre a PUCPR e os CB. Os envelopes postais possuem 256 tons de cinza e para sua apresentação, segundo cláusula de sigilosidade do convênio, é necessário esconder os nomes dos destinatários sob tarjas pretas.

1.2 Objetivos

Definir uma abordagem de segmentação de imagens baseada na teoria do Caos empregando a teoria fractal, mais especificamente a dimensão fractal para segmentar imagens complexas.

1.3 Justificativas

A segmentação automática de imagens geralmente representa um campo árduo, difícil e de grande importância no que diz respeito à interpretação das imagens. A segmentação determina o eventual sucesso ou fracasso da análise dessas imagens. De fato, uma segmentação eficiente quase sempre garante sucesso no processo de reconhecimento.

Empregando uma teoria herdada da teoria do caos que, segundo [KAPLAN 1999] [VUDUC 1997] [KELLER, CHENG, CROWNOVER 1989] [NEARY 2000] [MARAGOS, SUN 1993] [LEE, CHEN, HSIEH 2002] [CHEN *et al* 2001] [RINGLER 2003], já se mostrou promissora e consistente na segmentação de vários tipos de imagens, espera-se, a partir da dimensão fractal, extrair de forma eficiente tanto dados manuscritos irregulares quanto objetos geométricos de formatação conhecida (carimbo e selo) com uma única ferramenta que, teoricamente, permite quantificar o aspecto regular e irregular de objetos.

Capítulo 2

Teoria Fractal

2.1 Introdução ao Caos

O interesse pela teoria do caos tem sido meteórico. Durante os anos 80, o interesse foi enorme, e alguns cientistas a têm colocado ao lado de duas grandes revoluções da física teórica no século 20 – a relatividade e a mecânica quântica. Enquanto estas teorias desafiavam os sistemas dinâmicos Newtonianos, o caos tem questionado a opinião tradicional da estrutura Newtoniana. Onde começa o caos a ciência clássica pára [GLEICK 1990]. O crescimento do caos, sua fascinante combinação entre a noção de ordem e a de desordem e a sua constante instabilidade são relacionados com o rápido desenvolvimento de computadores mais potentes. Nos últimos 25 anos, as inúmeras inovações da computação gráfica foram fatores que permitiram a cientistas e matemáticos fazerem progressos no campo dos sistemas não lineares, do qual o caos é derivado.

Através das simulações realistas no computador e de técnicas de vídeo, cientistas puderam ver a evolução dos sistemas dinâmicos e os complexos efeitos caóticos das equações diferenciais básicas. O Caos ganhou interesse especial em várias áreas científicas, como a biologia, física, química, matemática, astronomia, economia e geografia e também a arte e o design.

Uma exceção, para a regra geral da linearidade, foi o conjunto de equações diferenciais derivadas do movimento dos fluídos [CRILLY, EARNSHAW, JONRS 1991]. O matemático suíço Leonard Euler publicou um artigo, em 1755, intitulado “Princípios gerais do movimento dos fluídos”, no qual mostrou um conjunto parcial de equações diferenciais para descrever o movimento de fluídos não viscosos. Após seguiu-se várias melhorias, como a do engenheiro francês Claude Navier em 1821 que levou em consideração a viscosidade. Em 1845, o físico matemático britânico George Stokes fez diferentes hipóteses das mesmas equações. Em 1920, o meteorologista Lew Fry Richardson (que influenciou a historia do fractal através da medição da linha costeira) calculou soluções de uma equação especificando as condições iniciais e limítrofes usando métodos numéricos. Após a segunda guerra, John Von Neumann notou que os métodos analíticos existentes dos matemáticos estavam inadequados para problemas não lineares. O problema de não linearidade é de notória dificuldade. Na dinâmica dos fluídos, por exemplo, o fluxo estável dos fluídos é perfeitamente compreendido. Mas os movimentos

turbulentos, como o rápido movimento da correnteza das águas, são um problema de grande complexidade. A relação do caos com a turbulência na dinâmica dos fluídos permite solucionar certos problemas.

A principal característica do caos é que um simples sistema determinístico pode gerar o que parece ser um comportamento randômico. O caos pode ser observado em sistemas mecânicos básicos. O pêndulo, como o *Botafumeiro* da Catedral de Santiago de Compostela [BERGÉ, POMEAU, GANCE 1995], no qual o eixo movimenta para cima e para baixo é um exemplo Figura 2.1. Para muitas frequências do eixo o balanço do pêndulo é normal, mas em determinado ponto o movimento do pêndulo torna-se errático e imprevisível. Neste ponto o movimento periódico regular do pêndulo, se torna um movimento caótico.

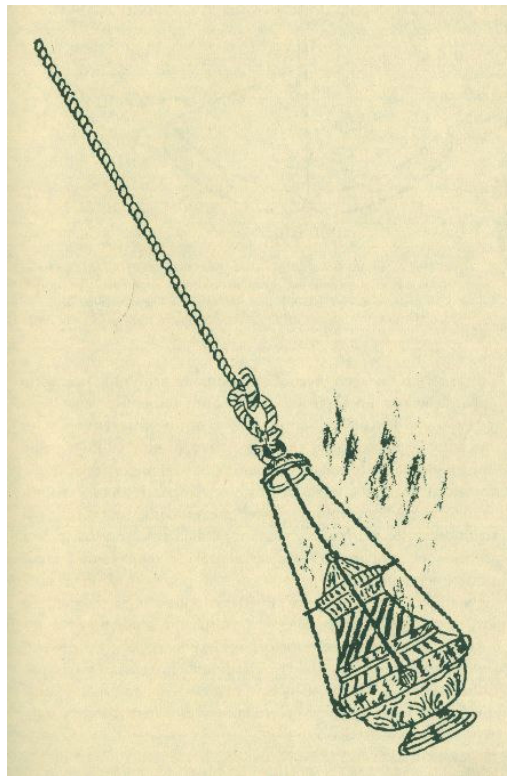


Figura 2.1 Botafumeiro da Catedral de Santiago de Compostela

Uma marca essencial do caos em sistemas não lineares é que eles são extremamente sensíveis às condições iniciais do sistema. Isto quer dizer que dois conjuntos de condições de um sistema que, inicialmente são muito próximos, ao longo do tempo podem fornecer

resultados extremamente diferentes. Um exemplo elementar, o *pinboard* inventado por Sir Francis Galton, pode ilustrar estatisticamente tal fato. De acordo com a lei de Newton, bolas soltas da mesma posição com a mesma velocidade inicial devem seguir o mesmo caminho para baixo. Na prática, é improvável que as bolas serão soltas exatamente na mesma posição e com a mesma velocidade inicial. Quando as bolas batem nos pinos, uma diferença mínima é ampliada e as bolas tomam caminhos diferentes. A sensibilidade de sistemas não lineares a condições iniciais é popularmente conhecida como o “efeito borboleta”, por que um simples bater das asas da borboleta teoricamente alteraria as condições iniciais de um sistema meteorológico e poderia assim aumentar drasticamente padrões meteorológicos ao longo do tempo.

Os sistemas dinâmicos clássicos podem ser pensados como atratores que controlam uma porção do espaço. No caso do pêndulo, todas as trajetórias são “atraídas” para um simples ponto, o atrator, que é a posição quando o pêndulo está parado. Para o sistema descrito pelas equações de Lorenz [LORENZ 1996], existe um atrator, mas ele não é simplesmente um atrator pontual. A maneira que se faz a trajetória para o atrator, neste caso, é mostrado pela projeção bi-dimensional. Se as trajetórias das equações de Lorentz são plotadas no espaço de fase tri-dimensional, vê-se que elas são atraídas em direção a uma delimitada elipsóide, a qual todas as trajetórias entram e saem e na qual elas nunca emergem. Uma propriedade matemática curiosa desta elipsóide é que possui volume zero. É de conhecimento que o atrator de Lorenz é uma estrutura de camadas aninhadas infinitamente, e que o volume zero é consistente com o cruzamento de seções através das camadas sendo fractal nas estruturas. Fractal não tem dimensão inteira como os subconjuntos mais comuns no espaço Euclidiano, mas tem dimensão fracionária. Também, quanto mais de perto você vê os fractais mais detalhes você pode observar. Eles têm a propriedade básica da auto-similaridade. A teoria fractal com suas múltiplas aplicações tem demonstrado utilidade nas aplicações de análise de imagens, é como ver-se-á mais à frente.

2.2 Teoria Fractal

O termo fractal foi cunhado por Benoit Mandelbrot nos anos 70, mas os objetos agora considerados como fractais sempre existiram. Os conjuntos fractais matemáticos têm

alto grau de complexidade geométrica, podem modelar muitos fenômenos e fornecer uma estrutura matemática apropriada para estudar as formas irregulares e complexas encontradas na natureza. Muitos dos objetos que ocorrem naturalmente como árvores, linha da costa ou nuvens, são agora considerados com propriedades fractais, e muito do interesse neste item decorre em simular os fenômenos naturais usando computação gráfica [MANDELBROT 1977]. Outra forma mais abstrata dos objetos fractais foi desenvolvida por artistas e matemáticos, e novamente as técnicas de computação gráfica tem dado novos discernimentos nas estruturas destes objetos.

A definição de Mandelbrot de um conjunto fractal X é que, “a dimensão de Hausdorff $h(X)$ (seção 3.3) não seja um inteiro”. Esta definição formal será aprofundada a seguir, mas é tentador ter agora uma idéia do que compreende um objeto fractal. A propriedade essencial é a da auto-similaridade. Subconjuntos de um objeto fractal têm a mesma forma do objeto todo. Objetos fractais teóricos são infinitesimalmente subdivisíveis deste modo, cada sub-conjunto, pequeno, no entanto contém não menos detalhes do que o conjunto completo. Na realidade objetos fractais são aproximações disto, já que alguns detalhes na subdivisão são perdidos. Estes conceitos de auto-similaridade e subdivisibilidade infinitesimal são vagos, mas deve-se proceder seguindo o espírito de Karl Popper, que disse “Nunca deixe você ser preso tomando seriamente os problemas das palavras e de seus significados... qualquer movimento para aumentar a precisão ou a clareza precisa ser *ad hoc* ou ‘fragmentado’ “.

A seguir faz-se um breve histórico da teoria do caos até Mandelbrot [CRILLY, EARNshaw, JONRS 1991].

2.2.1 Teoria Fractal de Durer

Um dos primeiros artistas que gerou objetos fractais baseados em pentágonos regulares foi Albrecht Durer (1471-1528). Tomando um pentágono regular de lado s e ao redor de cada lado desenhando outros cinco iguais, a forma criada ajusta-se quase exatamente a um pentágono regular de lado S . A Figura 2.2 agora se assemelha a um pentágono maior com cinco incisões na forma de triângulos isósceles com centro em cada lado. Os ângulos iguais destes triângulos são de 72° (o ângulo externo de um pentágono regular é $360^\circ / 5$), o terceiro ângulo sendo metade deste é 36° (como o triângulo). A proporção dos lados deste triângulo é conhecida desde os dias de Pitágoras. Tudo o que se

precisa saber é que o menor lado do triângulo pode ser calculado como sendo $2s \cos 72^\circ$. Isto quer dizer que a relação entre os lados de nossos dois pentágonos é:

$$\frac{s}{S} = \frac{1}{2 + 2 \cos 72} \quad (2.1)$$

Agora imagine começar o procedimento com o pentágono grande. Sabendo do meio de calcular s a partir de S , pode-se desenhar seis pentágonos menores os quais se ajustam dentro do pentágono maior. A subdivisão pode continuar com pentágonos cada vez menores sendo produzidos. Cada sub-pentágono é uma cópia do todo. Se a subdivisão continuar *ad infinitum*, o pentágono de Durer forma realmente um objeto fractal.

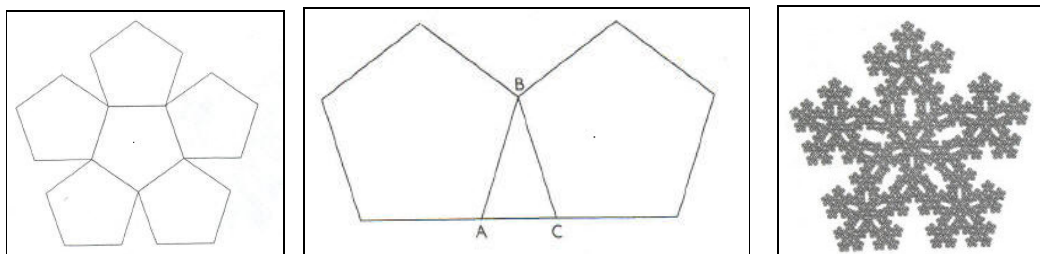


Figura 2.2: Pentágono Figura 2.3: Triângulo ABC Figura 2.4: Pentágono após 4 subdivisões.

2.2.2 Teoria Fractal de Cantor

Durante o século 19 a teoria dos conjuntos foi sendo desenvolvida. Matemáticos se deliciaram produzindo conjuntos com propriedades estranhas, muitas delas agora reconhecidas como de natureza fractal. Um destes é o conjunto desenvolvido por George Cantor (1845-1918). Sua construção é relativamente simples e pode ser ilustrada pelo “pente” de Cantor como mostra a Figura 2.5. Inicie com todos os números reais no intervalo de $[0,1]$. Extraia a terça parte central do intervalo original $(1/3, 2/3)$, deixando as outras duas terça partes $[0, 1/3]$ e $[2/3, 1]$. Este processo continua, sendo extraído a cada estágio a parte central do intervalo, que permanece. Muito do trabalho de Cantor foi usado para solucionar os paradoxos do conceito de infinito. Um dos seus contemporâneos, J.W.R Deekind (1831-1916), proclamou uma propriedade universal dos conjuntos infinitos o qual parece muito com o conceito de objeto fractal de hoje: “Um sistema S é dito ser infinito quando é similar a própria parte dele mesma”.

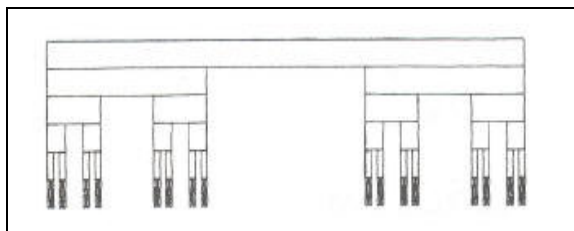


Figura 2.5: Pente de Cantor após 5 sub divisões

2.2.3 Teoria Fractal de Peano, Hilbert e von Kock

Em 1890, Giuseppe Peano (1858-1932) “mostrou como a matemática poderia ultrajar completamente o sentido comum quando construiu o espaço contínuo adicionando curvas. David Hilbert mais tarde desenvolveu uma construção similar, uma curva a qual visita todos os pontos do quadrado e em nenhum lugar é diferenciável. A curva gerada por Kock em 1904 é a curva típica gerada naquele tempo, é um dos objetos fractais mais clássicos, como ilustra a Figura 2.6.

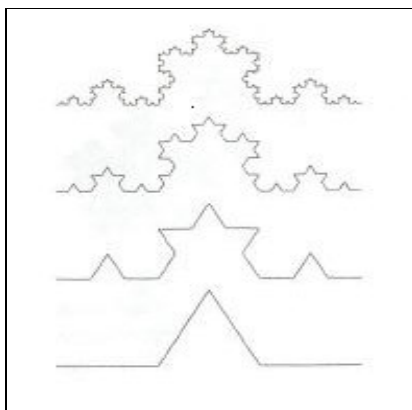


Figura 2.6: Curva de Koch após 4 gerações

2.2.4 Teoria Fractal de Sierpinski

Waclaw Sierpinski (b 1882) deu seu nome a um grande número de objetos fractais. O triângulo de Sierpinski e o tapete de Sierpinski, tem como espaço subjacente o R^2 . A esponja e o tetraedro de Sierpinski tem como espaço subjacente tri-dimensional. Para construir um triângulo de Sierpinski (Figura 2.7) é extraída do triângulo original a cópia invertida central em meia escala dele mesmo e junta-se ao triângulo extraído aos pontos médios dos três lados do triângulo original. Três triângulos com lado igual a metade do

anterior são criados e um quarto da área do triângulo original foi removida. Este processo é repetido para cada triângulo restante enquanto houver áreas não nulas.

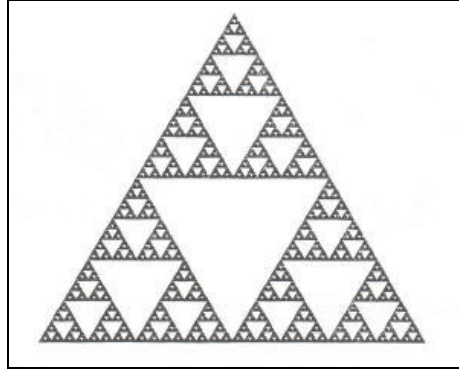


Figura 2.7: Triângulo de Sierpinski

2.2.5 Teoria Fractal de Julia e Fatou

É importante lembrar da relevância dos trabalhos, no plano complexo, de Gaston Julia (1893-1978) e Pierre Fatou (1878-1929) para a teoria do caos. Números complexos tomam a forma de $z = x + i y$ onde o número imaginário i é definido como

$$i^2 = -1 \quad (2.2)$$

Usando $i^2 = -1$, pode definir

$$z^2 = (x + i y) (x + i y) = (x^2 - y^2) + i(2xy) \quad (2.3)$$

Suponha que $c = a + i b$ é uma constante complexa. Tem-se então

$$z^2 + c = (x^2 - y^2 + a) + i(2xy + b) \quad (2.4)$$

Para um par de valores (x_0, y_0) determina-se um ponto no plano que é a representação de $z_0 = x_0 + i y_0$, pode-se usar a função

$$z_1 = f(z_0) = z_0^2 + c \quad (2.5)$$

para criar um “ponto de imagem” de (x_0, y_0) em

$$(x_1, y_1) = (\{x_0^2 - y_0^2 + a\}, \{2 x_0 y_0 + b\}) \quad (2.6)$$

Já vimos que a criação de um objeto fractal frequentemente envolve repetições de uma regra particular, emendando o ponto de início em cada estágio. Este caso não é exceção. Repete-se a regra de criação de (x_1, y_1) para encontrar o segundo ponto da imagem (x_2, y_2) . Este processo é repetido. Há frequentemente dois resultados deste processo. A seqüência de pontos diverge do ponto original, isto é, as distâncias dos pontos das imagens do ponto (X_0, Y_0) tornam-se incrivelmente grandes, eventualmente tendendo a infinito; ou a seqüência de pontos converge para um ponto fixo, ficando perto e cada vez mais perto do ponto. Há outras possibilidades, como um ciclo finito de pontos sendo repetidos, mas a seqüência de divergência e convergência é mais frequentemente observada. Se c é fixo, pontos (x_0, y_0) que agem como ponto de início para tal operação, podem ser divididos em dois conjuntos, o conjunto de pontos cuja seqüência diverge, e o conjunto de ponto cuja seqüência não diverge. Estes pontos que conduzem para as seqüências não divergentes, mas cuja posição esta próxima dos pontos que criam as seqüências divergentes constituem o conjunto de Julia para um valor particular de c . O conjunto de Julia (Figura 2.8) é a fronteira entre valores não divergentes de (x_0, y_0) e valores divergentes de (x_0, y_0) . Este conjunto é de natureza fractal.

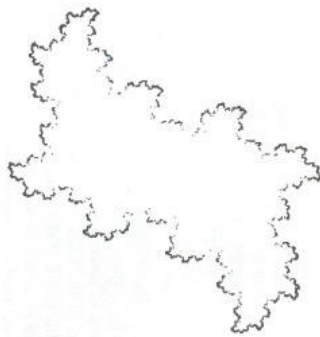


Figura 2.8 Conjunto conectado de Julia para $c = 0,1 - 0,6i$, usando 20000 pontos selecionados randomicamente.

2.2.6 O Fractal de Mandelbrot

O conjunto de Mandelbrot pode ser definido da sua relação com o conjunto de Julia. Na teoria de Julia e Fatou acima, são descritas as duas formas do conjunto de Julia dependendo da escolha do valor da constante c : conectado e desconectado.

Há uma definição equivalente e alternativa a qual é usada para criar imagens do conjunto de Mandelbrot. Adrian Douady explica que esta equivalência é o resultado do teorema provado independentemente por Julia e Fatou em 1919. O método resultante é similar ao usado para ilustração do conjunto de Julia num equipamento de *raster*, quando a função

$$f(z) = z^2 + c \quad (2.7)$$

é usada para criar seqüências de valores complexos $z_0, z_1, z_2, \dots$ correspondente para os pontos no plano (x, y) para um dado valor de c , sendo o *pixel* correspondente ao ponto inicial z_0 sendo colorido conforme as propriedades de seqüência. Para criar uma imagem do conjunto de Mandelbrot, o ponto inicial z_0 tem valor zero, correspondente a origem do plano complexo $(0, 0)$. As seqüências $z_0, z_1, z_2, \dots$ são avaliadas da mesma maneira como para o conjunto de Julia, exceto que a posição do *pixel* é feita para corresponder a constante c . Então, se a seqüência definida acima converge, o ponto correspondente a c pertence ao conjunto de Mandelbrot. Se ela diverge, o ponto correspondente a c esta fora do conjunto de Mandelbrot.

Capítulo 3

Segmentação pela Teoria Fractal

3.1 Segmentação

Segundo Facon [FACON 2002] *“alguns psicólogos alemães no início do século (Kohler, Wertheimer, Koffka) mostraram que o ser humano, ao nível de visão, efetua agrupamentos sobre o que ele percebe, baseados na proximidade, na similaridade, e na continuidade. O objetivo da segmentação é obter, a partir de uma imagem, um conjunto de “primitivas” ou elementos significativos que contém a informação semântica relativa a imagem”*

É importante ter um método de segmentação de uma imagem em regiões sem a intervenção humana. De fato, é mister que o computador esteja apto a diferenciar por algum meio, as diferentes regiões de uma imagem.

Algoritmos de segmentação já foram propostos. Geralmente é necessário ter certas características e encontrar um bom algoritmo de classificação, mapeando os resultados no espaço de características, com a intenção de saber se um ponto está em determinada região, sob exatos limites de probabilidades. A maioria dos métodos deste tipo precisa da intervenção humana para saber o número de regiões.

A análise da textura, de maneira geral, tem o objetivo de descobrir características sobre elas, as quais ajudarão a descrever completamente a textura ou a família de texturas. Isto é muito útil em muitos tipos de aplicações, particularmente no reconhecimento de objetos por computador. Muitos dos trabalhos feitos em classificação de textura são relevantes em segmentação de imagens.

Segundo Facon [FACON 2004] *“A textura é uma propriedade importante das superfícies que o sistema de percepção visual humano utiliza para segmentar e classificar objetos”*. E segundo Gonzalez e Woods [GONZALES, WOODS 2000] *“Embora não*

exista nenhuma definição formal de textura, este descritor intuitivamente fornece medidas de propriedades como suavidade, rugosidade e regularidade”.

Pesquisadores têm observado quatro características da superfície:

1. Direcionabilidade: relaciona-se com o grau de diferença ao longo de diferentes eixos da imagem.
2. Contraste: a extensão da escala do nível de cinza da imagem.
3. Aspereza: se uma textura é fina ou granulada.
4. Regularidade: isto concerne à uniformidade da textura.

Na Figura 3.1 mostram-se vários tipos de texturas com características mencionadas aqui, incluindo alto contraste, aspereza e direcionabilidade.

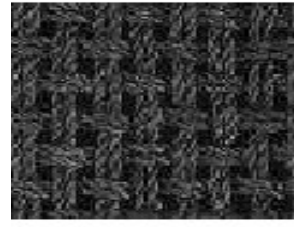
As principais abordagens usadas em análise de imagens para a descrição de texturas são a estatística, a estrutural e a espectral. A abordagem estatística caracteriza a textura como suave, áspera, granular, regular ou randômica, etc. A abordagem estrutural caracteriza arranjos das primitivas de imagens, como a descrição de textura baseada em linhas paralelas regularmente espaçadas. A abordagem espectral baseia-se no espectro de Fourier, identificando picos de alta energia no espectro caracterizando a periodicidade global da imagem.

Para representar texturas com aparência de nevoa como fumaça, mamografia, pinturas de fogo, etc, Mandelbrot popularizou o auto movimento Browniano fracionário (fBm), assim, pesquisadores desenvolveram métodos de segmentação de texturas usando métodos fractais.

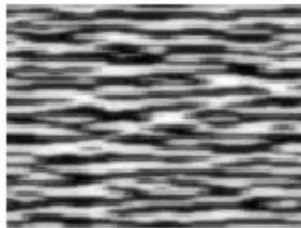
Já que a maioria das texturas do mundo real exhibe detalhes em muitas escalas e que existe um grau estatístico de auto-similaridade, para classes de texturas que tenham estas características, a dimensão fractal pode servir como um bom classificador de texturas.



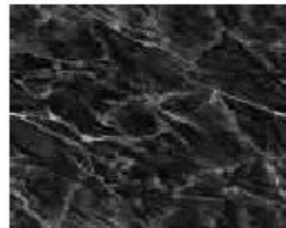
(a) casca de arvore



(b) aniagem



(c) dunas de areia



(d) granito



(e) barro seco pedra áspera



(f) pedra áspera



(g) casca de nos

Figura 3.1: Vários tipos de texturas

3.2 Dimensão Fractal

No universo Euclidiano e Newtoniano, a dimensionalidade é claramente definida. Um ponto tem uma dimensão zero. Uma linha, reta ou curva têm dimensão um, pois qualquer ponto na linha pode ser representado por um simples parâmetro, como sua distância ao longo da linha a um ponto fixo. Similarmente, pontos do plano ou qualquer superfície curva podem ser representados por dois valores, por exemplo, as coordenadas

(x, y) no plano Cartesiano, ou latitude ou longitude na superfície do globo, fazendo estas superfícies bidimensionais. A descrição Newtoniana do espaço em que vivemos é tri-dimensional. Qualquer ponto em nosso universo pode ser localizado por três valores, por exemplo, suas coordenadas Cartesianas (x, y, z) . Matemáticos têm construído outros espaços estendendo analogamente as leis dos espaços de baixa dimensionalidade para os de alta dimensionalidade. Espaços de muitas dimensões não podem ser visualizados, mas eles têm sido úteis em vários contextos, sendo a dimensão destes espaços inteiros não negativos $0, 1, 2, 3, \dots$

Alguns dos conjuntos mais complicados que foram mencionados anteriormente não se encaixam naturalmente nestes espaços restritos. Considere o conjunto de Cantor apresentado na figura 2.5. Ele consiste num conjunto de valores individuais separados que se pode pensar como tendo dimensão zero. Apesar disso, ele tem uma correspondência um para um no conjunto de todos os valores da linha real no intervalo $[0, 1]$, o qual é evidentemente um conjunto de dimensão um. A curva de Kock, apresentada na figura 2.6, aparece ser de uma dimensão pelos argumentos acima. Apesar dos pontos ao longo da curva não poderem ser rotulados pela maneira usual das distâncias de um ponto fixo, pois a distância entre quaisquer dois pontos da curva de Kock é infinita. A curva de Kock não se preenche no espaço bi-dimensional no qual ela existe, então ele não pode ser bi-dimensional. Mas, ela parece ser de mais do que uma dimensão. Problemas similares ocorrem quando considerados objetos fractais no espaço tri-dimensional, como a esponja e o tetraedro de Sierpinski.

Para melhor exemplificar a dimensão fractal, usar-se-ão figuras usuais como o triângulo e quadrado que existem no espaço bi-dimensional, assim como propriedades de auto-similaridade e escala. Desta maneira triângulo e quadrados podem ser subdivididos para conter em escala menor, cópias deles mesmos. Ambos podem ser subdivididos em 4 cópias em $1/2$ escala, nove cópias em $1/3$ de escala, 16 cópias em $1/4$ de escala. Em cada caso, se o número de cópias é N e o fator de escala é f , tem-se a seguinte relação:

$$N = (1/f)^2 \quad (3.1)$$

Um cubo pode ser subdividido em 8 cópias, cada uma delas com lado igual ao anterior em $1/2$ escala, 27 cópias em $1/3$ de escala. A relação entre N e f tornou-se:

$$N = (1/f)^3 \quad (3.2)$$

Este fator $1/f$ indica a relação entre dimensões características do objeto. Pode ser generalizado. Se o objeto pode ser subdividido em N cópias dele mesmo na escala f , então sua dimensão é o valor de D que satisfaz:

$$N = (1/f)^D \quad (3.3)$$

D pode ser obtido diretamente usando-se a propriedade da função do logaritmo

$$D = \frac{\log(N)}{\log(1/f)} \quad (3.4)$$

Considere isto agora aplicado a algum objeto fractal mostrado anteriormente. O conjunto de Cantor contém duas cópias dele mesmo cada uma com $1/3$ de escala, assim sua dimensão é:

$$D = \frac{\log(2)}{\log(3)} = 0,6309 \quad (3.5)$$

ou seja, uma dimensão não inteira entre 0 e 1.

Para a curva de Koch têm-se 4 cópias cada uma com $1/3$ de escala. Assim,

$$D = \frac{\log(4)}{\log(3)} = 1,2619 \quad (3.6)$$

ou seja, uma dimensão não inteira entre 1 e 2.

O triângulo de Sierpinski contém 3 cópias com $1/2$ escala. Assim,

$$D = \frac{\log(3)}{\log(2)} = 1,5850 \quad (3.7)$$

Este método é preciso para objetos que contêm cópias exatas deles mesmos, todos com mesma escala, mas, muitos objetos não têm este comportamento. A dimensão de Hausdorff (seção 3.3) não apresenta essa restrição.

3.3 Definição da Dimensão Fractal

Define-se a dimensão fractal D (ou ainda dimensão de Hausdorff) de um conjunto F da seguinte maneira:

$$\forall \delta > 0, H_s^\delta(F) = \text{Min} \{ \sum |U_i|^s \text{ onde } |U_i| \text{ é uma cobertura } \delta\text{-cobertura de } F \} \quad (3.8)$$

onde $|U|$ é o diâmetro de U e onde δ -cobertura de F é uma cobertura por conjuntos de diâmetro inferior ou igual a δ . Portanto tem-se:

$$D = \text{Min} \{ s / \lim_{\delta \rightarrow 0} H_s^\delta = 0 \} = \text{Max} \{ s / \lim_{\delta \rightarrow 0} H_s^\delta = +\infty \} \quad (3.9)$$

Esta definição permite medir de forma precisa a dimensão de um conjunto.

A abordagem de Wen-Li [LEE, CHEN, HSIEH 2002] mostra que a propriedade de auto-similaridade, como mostrado pela curva de Von Kock, na Figura 2.6, é um dos conceitos centrais da geometria fractal. Pelo conceito da auto-similaridade, a dimensão fractal pode ser estimada. Considerando um conjunto limitado A no espaço- n Euclidiano, o conjunto A é auto similar se A é a união de N_r cópias não sobrepostas dele mesmo em escala menor pela taxa r . A dimensão fractal D_f de A é derivada da relação:

$$D_f = \log(N_r) / \log(1/r) \quad (3.10)$$

3.4 Cálculo Prático da Dimensão Fractal

Através do estudo bibliográfico realizado, percebeu-se que as técnicas de cálculo da Dimensão Fractal podem ser agrupadas em dois grupos: as técnicas não morfológicas do tipo *box-counting* usando coberturas com caixas cúbicas e as técnicas morfológicas usando coberturas morfológicas com elementos estruturantes convexas. Aqui serão apresentadas algumas das técnicas de cálculo da Dimensão Fractal para imagens digitais em níveis de cinza.

3.4.1 Técnicas do tipo *box-counting*

Muitos métodos têm sido propostos para o cálculo da dimensão fractal de imagens. O mais usado é o *Box-counting*, o qual é baseado na cobertura da superfície cuja dimensão se deseja calcular por *boxes* de tamanho fixo.

Pentland em [PENTLAND 1984] foi um dos primeiros a apresentar um método de cálculo da dimensão fractal para imagens digitais, e demonstra que, para texturas homogêneas, a superfície de um objeto 3D é fractal se e somente se sua imagem é fractal e se a sua dimensão fractal não se altera com a escala. A dimensão fractal de uma imagem tem um limite inferior e superior (limitada entre 2 e 3). Quanto mais lisa for a textura, mais a sua dimensão fractal será próxima de 2; quanto mais rugosa for sua textura mais a sua dimensão fractal será próxima de 3.

3.4.1.1 Abordagem *2D Variation Procedure*

A abordagem *2D Variation Procedure* desenvolvida por Roland Kraft e Josef Kauer [KRAFT, KAUER 1995] consiste em uma análise da vizinhança de um *pixel* para diferentes tamanhos de janela centrada nele. As janelas usadas são quadradas de tamanhos diversos. A figura 3.2 ilustra o exemplo para $r = 5$, $r = 7$ e $r = 9$.

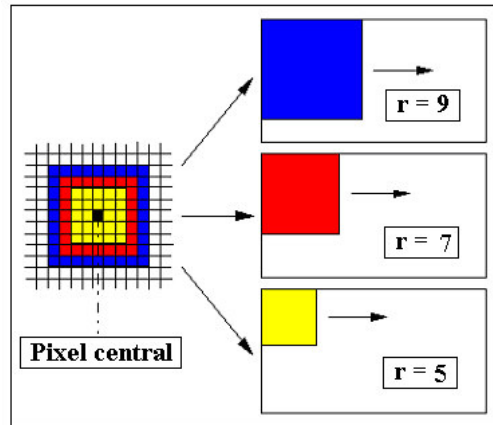


Figure 3.2: Exemplo de análise da *2D variation procedure* para $r = 5$, $r = 7$ e $r = 9$

Essa abordagem pode ser aplicada tanto para imagens binárias quanto para imagens em níveis de cinza. Para imagens em níveis de cinza, a imagem é percorrida *pixel a pixel* e

o algoritmo procura os valores mínimo e máximo de nível de cinza na janela quadrada de tamanho r e os atribui ao *pixel* central respectivamente. Deste modo, define-se uma função bidimensional relacionando os valores máximo e mínimo para cada tamanho de janela e assim, a diferença volumétrica $V(r)$, para a imagem inteira, entre esses valores máximo e mínimo. No modelo de representação de Richardson-Mandelbrot, o volume $V(r)$ mostra-se linearmente dependente do tamanho r da janela, assim:

$$V(r) = \text{constante } r^s \quad (3.11)$$

Obtêm-se a dimensão fractal $D2d$ avaliando a inclinação s da reta no modelo de representação de Richardson-Mandelbrot de $V(r)$ versus r .

$$D2d = 3 - s / 2 \quad (3.12)$$

3.4.1.2 Técnica de Voss

Voss [VOSS 1986] apresentou um método de cálculo da dimensão fractal baseado em caixas cúbicas. Este método busca estimar $N(r)$, o número médio de caixas cúbicas de lado de tamanho r necessárias para cobrir a imagem considerada como sendo uma superfície 3D. Para isto, é preciso avaliar a probabilidade $P(m, r)$ que uma caixa cúbica de tamanho r , centrada num ponto arbitrário da superfície, contenha m pontos do conjunto. Tem-se:

$$\forall r, \sum_{m=1}^{Np} P(m, r) = 1 \quad (3.13)$$

onde Np é o número de pontos neste cubo.

Se $N(m, r)$ representa o número médio de caixas que contem m pontos e K o número de posições (tamanho da região), então:

$$P(m, r) = mN(m, r) / K \quad (3.14)$$

A avaliação do número médio $N(r)$ de caixas disjuntas necessárias para cobrir a superfície é:

$$N(r) = \sum_{m=1}^{Np} N(m, r) = K \sum_{m=1}^{Np} P(m, r) \quad (3.15)$$

A avaliação pelo método dos Mínimos Quadrados da inclinação da melhor reta passando pela nuvem de pontos no sistema de coordenadas $(\ln(r), \ln(N(r)))$, obtida com caixas cúbicas de tamanho r crescente, permite avaliar a dimensão fractal D .

Entretanto, este método usualmente sub-estima a dimensão fractal real, principalmente devido a discretização do domínio da imagem e da quantização dos níveis de cinza.

3.4.1.3 Abordagem de Vuduc

Foi proposto, na abordagem de Vuduc [VUDUC 1997], o seguinte método: basicamente, para cada *pixel* (i, j) na imagem, calcula-se a dimensão fractal de uma pequena janela W ao redor do *pixel*. Atribui-se a dimensão fractal local para este *pixel*.

Seja I = uma matriz $n \times n$ de níveis de cinza

Seja DF = uma matriz de dimensões fractais, um elemento por *pixel* da matriz I

Para cada *pixel* $I[i, j]$

Passo 1: Seja W uma pequena janela da imagem I de tamanho $m \times m$ ao redor de $I[i, j]$

Passo 2: Calcule a dimensão fractal D para W

Passo 3: Seja $DF[i, j] = D$

Faça um histograma dos valores de DF e em seguida atribui uma cor a imagem baseado no número da região. Diferentes cores correspondem a diferentes regiões.

O Passo 2 acima citado pode ser implementado de varias maneiras. A seguir serão apresentados os métodos de Sarkar, de Sarkar e Chaudhuri e de Beaver.

3.4.1.4 Abordagem de Sarkar

Na abordagem de Sarkar [VUDUC 1997] calcula-se $N(r)$ da seguinte maneira. A partir de uma janela J de tamanho $n \times n$, gera-se um conjunto de janelas menores j sem superposição de tamanho $d \times d$ tal que $1 < d \leq n/2$. Para cada uma das janelas, calcula-se o número $n_r(j)$ de caixas de tamanho $d \times d \times d$, necessárias para conter a porção de superfície correspondente a j . O número $n_r(j)$ não é nada mais que a extensão +1 da imagem I na janela j :

$$n_r(j) = \text{Max}_j \{I[s]\} - \text{Min}_j \{I[s]\} + 1 \quad (3.16)$$

Avalia-se $N(r)$ somando para todas as janelas j :

$$N(r) = \sum_j n_r(j) \quad (3.17)$$

Por aumento do tamanho d , ou ainda do fator de escala r , avalia-se $N(r)$ para várias escalas. A avaliação pelo método dos Mínimos Quadrados da inclinação da melhor reta passando pela nuvem de pontos no sistema de coordenadas de $\ln[N(r)]$ versus $\ln(r)$, obtida com caixas cúbicas de tamanho r crescente, permite avaliar a dimensão fractal DF .

3.4.1.5 Abordagem de Sarkar e Chaudhuri

Nessa abordagem [VUDUC 1997], a dimensão fractal DF de um pixel (i, j) da janela W de tamanho $m \times m$ é calculada da seguinte maneira:

Para vários tamanhos de $r \in [0,1]$

1. Divide W em $(1/r)^2$ quadrados
2. Divide a extensão de intensidades $[0..255]$ em $1/r$ níveis, ou bins, numerados de $1 \dots 1/r$
3. Para cada célula do grid (quadrado) $(p, q) \in W$
 - a) Sejam Min e Max os valores mínimo e máximo, respectivamente, naquele grid de célula;
 - b) Calcula $n_{p,q}(r) = \text{Max}_r(p, q) - \text{Min}_r(p, q) + 1$

A avaliação pelo método dos Mínimos Quadrados da inclinação da melhor reta passando pela nuvem de pontos no sistema de coordenadas de $\ln[N(r)]$ versus $\ln(r)$, obtida com caixas cúbicas de tamanho r crescente, permite avaliar a dimensão fractal DF .

3.4.1.6 Abordagem de Beaver

A abordagem de Beaver [VUDUC 1997] utiliza uma variação do *box-counting*. É usado de novo em um *pixel* (i, j) num sub-janela de $m \times m$ *pixels* da janela I . Essencialmente, é usado o mesmo algoritmo de Sarkar com a diferença que a abordagem de Beaver re-escala a intensidade localmente antes de proceder com o *box-counting*. Isto é, para um dado *grid* (p, q) , o método re-escala as intensidades num intervalo $[-1/2r, 1/2r]$ e, subdividem o espaço de intensidade em $1/r$ regiões:

Para vários tamanhos de $r \in (0, 1)$

1. Divide W em $(1/r)^2$ quadrados
2. Para cada célula (p, q) do grid (quadrado) $\in W$
 - a) Re-escala a intensidade na extensão $(-1/2r, 1/2r)$ com $I(i, j)$ na origem
 - b) Divida a extensão das intensidades em $1/r$ níveis, ou bins (regiões, caixas), numeradas $1..1/r$
 - c) Sejam Min e Max os valores mínimo e máxima, respectivamente, daquele grid ;
 - d) calcula $n_{p,q}(r) = Max_r(p, q) - Min_r(p, q) + 1$

A avaliação pelo método dos Mínimos Quadrados da inclinação da melhor reta passando pela nuvem de pontos no sistema de coordenadas de $\ln[N(r)]$ versus $\ln(r)$, obtida com caixas cúbicas de tamanho r crescente, permite avaliar a dimensão fractal DF .

3.4.2 Coeficiente de Hurst

A dimensão fractal pode ser descrita pela relação,

$$D = \log(N) / \log(1/r) \quad \text{como na equação 3.4 .}$$

Isto mostra que há um relacionamento entre o log de N e r. Se for traçado um gráfico de $\log(N)$ x $\log(r)$ o resultado deve ser uma linha reta cuja inclinação seja aproximadamente D.

O coeficiente de Hurst é uma aproximação que utiliza este relacionamento [Parker 1997]. Considere uma figura em que uma região de 7 x 7 pixels é marcada de acordo com a distância de cada pixel ao pixel central. Há oito grupos de pixels, correspondendo às oito distâncias diferentes que são possíveis. Dentro de cada grupo a diferença a maior no nível de cinza é encontrada; é o mesmo que subtrair o nível de cinza do menor no grupo do maior.

O pixel central é ignorado, e uma reta é composta pelo log da diferença máxima (coordenada de y) e o log da distância do pixel central (coordenada de x). A inclinação desta reta é o coeficiente de Hurst, e substitui o pixel no centro da região.

Como exemplo, considere a seguinte região de 7 x 7 pixels de uma imagem:

```

85  70  86  92  60 102 202
91  81  98 113  86 119 189
96  86 102 107  74 107 194
101 91 113 107  76 107 194
99  68 107 107  76 107 194
107 94  93 115  83 115 198
94  98  98 107  81 115 194

```

A primeira etapa em calcular o coeficiente de Hurst é determinar a diferença máxima do nível de cinza para cada classe de distância dos pixels. Começando nos pixels na distância um ou menos do centro, o nível máximo é 113 e o mínimo é 83, para uma distância de 30. A classe seguinte tem a escala $113 - 74 = 39$, e a distância = a classe 2 tem uma escala de $118 - 74 = 44$.

Tabela 3.1 Dados do Coeficiente de Hurst

Classe	$d = 1$	$d = \sqrt{2}$	$d = 2$	$d = \sqrt{5}$	$d = \sqrt{8}$	$d = 3$	$d = \sqrt{10}$
Numero	30	39	44	50	51	130	138

Com estes dados, usando a relação de log x log construa uma reta, assim a etapa seguinte fazer o log da distância e da diferença do nível de cinza, como a mostra na tabela 3.2.

Tabela 3.2 Dados do Cálculo do Coeficiente de Hurst

Ln(Distancia)	0,0	0,347	0,693	0,805	1,040	1,099	1,151
Ln(Delta G)	3,401	3,664	3,784	3,912	3,932	4,868	4,927

A estes pontos, usando a equação dos mínimos quadrados, é construída uma reta, que neste caso tem a equação:

$$y = 1.145 x + 3.229$$

A inclinação desta reta, $m = 1,145$, é o coeficiente de Hurst.

3.4.3 Técnicas morfológicas usando coberturas morfológicas

As técnicas morfológicas de cálculo da Dimensão Fractal baseiam-se no uso de coberturas morfológicas com elementos estruturantes convexos.

Os autores Maragos e Sun [MARAGOS, SUN 1993] propõem uma abordagem morfológica para sinais. Inicialmente eles diferenciam a dimensão fractal de Hausdorff DH da dimensão topológica DT, da dimensão de Minkowski-Bouligand DM, da dimensão do *box counting* DB. No caso contínuo para , tem-se que:

$$0 \leq DT \leq DH \leq DM = DB \leq 2 \quad (3.18)$$

A dimensão de Minkowski-Bouligand DM baseia-se no conceito de Minkowski de encontrar o comprimento de curvas irregulares: dilate as com discos de raio ε para formar a união desses discos e assim crie a cobertura de Minkowski. Calculando a área $A(\varepsilon)$ do conjunto dilatado para todos os raios ε , Bouligand define a dimensão fractal DM como sendo:

$$D_M = \lim_{\varepsilon \rightarrow 0} \left(2 - \frac{\log[A(\varepsilon)]}{\log(\varepsilon)} \right) \quad (3.19)$$

Os autores Maragos e Sun comprovam que essa dimensão fractal DM de Bouligand pode ser determinada por morfologia matemática. No caso de empregar elementos estruturantes compactos que contem a origem e que tem uma distancia mínima não nula da

origem até seu limite, a formulação de Bouligand é ainda respeitada. Os autores Maragos e Sun demonstram que a dimensão fractal DM pode ser obtida como segue:

$$\log\left(\frac{A(\varepsilon)}{\varepsilon^2}\right) \approx D_M \log\left(\frac{1}{\varepsilon}\right) + \text{constate}, \text{ quando } \varepsilon \rightarrow 0 \quad (3.20)$$

Esta formulação permite o cálculo prático da dimensão fractal DM no caso discreto. No caso de sinais discretos de tamanho finito no tempo $f[n] = 0, 1, \dots, N$, a dimensão fractal DM pode ser avaliada usando coberturas morfológicas com elementos estruturantes convexos de tamanho $\varepsilon = 1, 2, 3, \dots, \varepsilon_{\max}$. Isto consiste nos seguintes passos:

- Passo 1: Selecionar um conjunto de elementos estruturantes discretos B que satisfaçam as hipóteses de conterem a origem e de terem uma distância mínima não nula da origem até seu limite e além disto de serem convexos. Preferencialmente, este conjunto deve ter um raio unitário em ambas as direções x e y . Há somente três escolhas para tal unidade do raio de B : o quadrado $3 \times 3 \text{ pixel}$, o *rhombus* de 5 pixels , e o segmento horizontal de 3 pixels .

i) Se B é o quadrado de $3 \times 3 \text{ pixel}$, o g correspondente é de forma retangular.

$$g_r[-1] = g_r[0] = g_r[1] = h \geq 0$$

$$\text{e } g_r[n] = -\infty, \text{ para } n \neq -1, 0, 1.$$

ii) Se B é o *rhombus* de 5 pixels , então g é de forma triangular, definida por

$$g_r[-1] = g_r[0] = g_r[1] = h \geq 0$$

$$\text{e } g_r[n] = -\infty, \text{ para } n \neq -1, 0, 1.$$

iii) Se B é o segmento horizontal de 3 pixels , então o g correspondente não pode ser visto como resultado nem de g ou de g_r por $h = 0$. Neste caso g é uma função retangular igual a zero no seu domínio.

- Passo 2: Realizar recursivamente a dilatação e erosão de f por $g^{\oplus \varepsilon}$ para elementos estruturantes de tamanho $\varepsilon = 1, 2, 3, \dots, \varepsilon_{\max}$. Seja $G = \{-1, 0, 1\}$ e $S = \{0, 1, \dots, N\}$, então:

$$f \oplus_s g[n] = \text{Max}_{-1 \leq I \leq 1} \{f[n+I] + g[I]\}, \quad \text{para } \epsilon = 1 \quad (3.21)$$

$$\text{e } f \oplus_s g^{\oplus(\epsilon+1)} = (f \oplus_s g^{\oplus\epsilon}) \oplus_s g, \quad \text{para } \epsilon \geq 2 \quad (3.22)$$

Do mesmo modo para a erosão $f \ominus_s g^{\oplus\epsilon}$. Para $n = 0, N$, as operações locais de máximo e de mínimo tomam lugar somente para as amostras disponíveis (linhas tracejadas na figura 3.3

- Passo 3: Calcular a área de cobertura

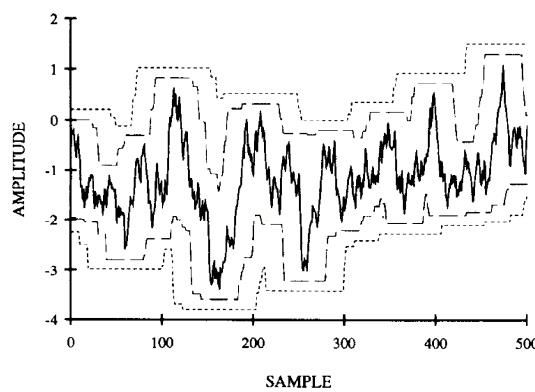
$$A_g[\epsilon] = \sum_{n=0}^N (f \oplus_s g^{\oplus\epsilon} - (f \ominus_s g^{\oplus\epsilon}))[n], \quad \text{para } \epsilon = 1, \dots, \epsilon_{\max} \leq N/2$$

- Passo 4: Ajustar uma linha reta aos mínimos quadrados para o gráfico do $\log(A_g[\epsilon]/\epsilon'^2)$ versus $\log(1/\epsilon')$, para $\epsilon' = 2/N, 4/N, \dots, \epsilon_{\max}$, onde $\epsilon' \cong 2\epsilon/N$ é a escala normalizada, $2/N \leq \epsilon' \leq \epsilon'_{\max} \leq 1$ e $\epsilon'_{\max} = 2\epsilon_{\max}/N$. A inclinação desta linha fornece uma estimativa aproximada da dimensão fractal DM de f , como demonstra a equação (3.20).

Veja que quando $h = 0$, então o passo 2 do algoritmo pode ser mais rápido, especificamente,

$$f \oplus_s g[n] = \max \{f[n-1], f[n], f[n+1]\}, \quad \epsilon = 1$$

$$f \oplus_s g^{\oplus(\epsilon+1)}[n] = \max \{f \oplus_s g^{\oplus\epsilon}[n-1], f \oplus_s g^{\oplus\epsilon}[n+1], f \oplus_s g^{\oplus\epsilon}[n]\}, \quad \epsilon \geq 2$$



(a)

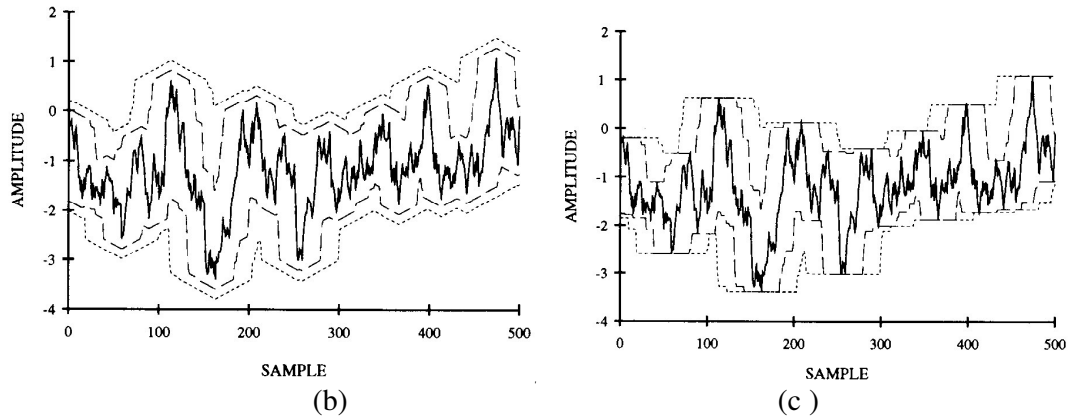


Figura 3.3 Sinal fBm (linha s3lida) com $D = 1.5$, $N = 500$ e sua eros3o /dilata33o por $g \oplus^\epsilon$ (linha tracejada) na escala $\epsilon = 20, 40$: (a) retangular $g = g_r$, com $h = 0,01$, (b) Triangular $g = g_t$, com $h = 0,01$, (c) Retangular g com $h = 0$

Infelizmente os autores Maragos e Sun n3o prop3em nenhuma formula33o no caso de estima33o da dimens3o fractal morfol3gica para imagens em n3veis de cinza.

Capítulo 4

Aglomerção: Reconhecimento de Padrões não supervisionado

4.1 Introdução

O mecanismo de *aglomeração* pode ser visto como um problema de reconhecimento de padrões não supervisionado. Suponha que esteja disponível um conjunto de dados sem qualquer informação quanto ao número de classes presentes nesse conjunto. A aprendizagem não supervisionada pode ser indicada para identificar as classes. Se for aceita a idéia do centro do aglomerado ser um método de representação, um meio obvio de caracterizar esse conjunto de dados é identificar os aglomerados por seus centros. [THEODORIDIS, KOUTROUMBAS 1999]

A aplicação do mecanismo de aglomeração para aprendizagem não supervisionada é, em principio, recomendada. Suponha que um conjunto de dados $\{x_1, x_2, \dots, x_N\}$ de classificação desconhecida. Estes padrões podem ser submetidos em um ou mais algoritmos para identificar os centros de aglomerados representativos. O resultado dos aglomerados pode então ser interpretado como classes de diferentes padrões. Uma vez criadas essas classes podem ser usadas para determinar funções de decisão por meio de um ou mais algoritmos de treinamento. Alternativamente, o centro de um aglomerado identificado na fase da aprendizagem não supervisionada pode ser usado como base para um classificador de distância mínima.

O projeto de um sistema de reconhecimento de padrões não supervisionado, em geral, requerer uma grande dose de intuição e experimentação.

O foco de interesse neste capítulo é a classificação não supervisionada onde a rotulação das classes dos padrões de treinamento não é disponível. Assim, o interesse principal agora é agrupar os padrões em aglomerados ‘sensíveis’, os quais permitirão descobrir similaridades e diferenças entre padrões e tirar conclusões úteis sobre eles. A aglomeração é uma das atividades mentais mais primitiva do homem usada para manusear

a quantidade enorme de informações recebidas todos os dias. Processar toda a informação como uma entidade simples seria impossível. Assim, os humanos tendem a categorizar as entidades (por ex: objeto, pessoa, eventos) em aglomerados. O fato de agrupar os padrões em aglomerados é comum em muitas áreas, tais como: biologia, zoologia, sociologia, arqueologia, geografia, psiquiatria, patologia e engenharia. O mecanismo de aglomeração pode ser encontrado sob diferentes nomes em diferentes contextos, tais como aprendizagem não supervisionada e auto-aprendizagem (em reconhecimento de padrões), taxonomia numérica (em biologia, ecologia), tipologia (em ciências sociais), e partição (na teoria dos grafos). O exemplo a seguir inspirado da biologia vai permitir esclarecer o desafio do processo de Aglomeração.

Considere os seguintes animais: cabrito, cão, boi (mamíferos), pardal, sabia (pássaros), cobra, lagarto (répteis), dourado, salmão, tubarão azul (peixes) e rã (anfíbio). Para organizar estes animais em aglomerado, precisa-se definir um critério de aglomeração. Assim, se empregar o meio como estes animais carregam sua prole como um critério de aglomeração, o cabrito, o boi, o cão e o tubarão azul estarão no mesmo aglomerado, enquanto todo o resto irá formar o segundo aglomerado. Se o critério de aglomerado for a existência de pulmões, o dourado, o salmão, e o tubarão azul estarão no mesmo aglomerado dos “Sem Pulmão”, enquanto todo o resto estará no segundo aglomerado dos “Com Pulmão”. De outra maneira, se o critério de aglomeração for o ambiente onde os animais vivem, o cabrito, o cão, o boi, o pardal, o sabia, a cobra e o lagarto irão formar um aglomerado (dos animais que vivem fora d’água), o dourado, o salmão e o tubarão azul irão formar um segundo aglomerado (dos animais que vivem somente na água), e a rã irá formar o terceiro aglomerado, ela pode viver dentro e fora d’água. Vale a pena assinalar se a coluna vertebral for um critério de aglomeração, todos os animais ficarão no mesmo aglomerado. Finalmente, pode usar critérios de aglomeração compostos, por exemplo: animais que possuem bico e o ambiente em que vivem. Os quatros aglomerados gerados são nesse caso: pardal e sabia (com bico e o ambiente é o ar) ; cabrito, cão, boi, cobra e lagarto (sem bico e o ambiente é a terra); dourado, salmão e tubarão azul (sem bico e o ambiente é a água); e a rã (sem bico e o ambiente pode ser a água e a terra).

Este exemplo mostra que o processo de atribuição de objetos aos aglomerados pode conduzir a resultados muito diferentes, dependendo do critério específico usado para a aglomeração.

Assume-se a seguir que todos os padrões são representados em termos de características, as quais formam o vetor de características unidimensional.

Os passos fundamentais que um especialista precisa seguir para desenvolver uma atividade de aglomeração são:

- *Seleção de Características*: Características precisam ser selecionadas apropriadamente para poder conter toda a informação necessária para o interesse da atividade. O maior objetivo é minimizar as redundâncias entre as características.

- *Medida de Similaridade*: Representa a medida que quantifica quanto similares são ou não são dois vetores de características. É natural assegurar que todas as características selecionadas contribuam igualmente para o cálculo da medida de similaridade e que não há características que dominem outras. Precisa-se tomar cuidado com este aspecto durante o pré-processamento. A medida de similaridade mais comumente usada é a distância Euclidiana. A distância Euclidiana entre dois padrões x e z unidimensionais é definida como segue:

$$D = \|x - z\| \quad (4.1)$$

Quanto menor for essa distância, maior a similaridade. Existe ainda outra medida de distância significativa, a distância de Mahalanobis. A distância de Mahalanobis de x para m é definida como segue:

$$D = (x - m)^T C^{-1} (x - m) \quad (4.2)$$

e representa uma medida de similaridade útil quando propriedades estatísticas estão sendo explicitamente consideradas. C transposto é a matriz de covariância do conjunto de dados, m é o vetor médio, e x representa uma variável desse conjunto de dados.

Medidas de similaridade não precisam restringir-se a métricas. Por exemplo, a função de similaridade não métrica:

$$s(x, z) = x'z / \|x\| \|z\| \quad (4.3)$$

representa o co-seno do ângulo entre os vetores x e z , e é máximo quando x e z são orientados na mesma direção em relação a origem. Esta medida de similaridade é útil quando a região do aglomerado tende a alongar-se na direção dos eixos principais, como na Figura 4.1. Ver-se-á nesta figura, por exemplo, que o padrão z_1 é mais similar a x do que o padrão z_2 visto que $s(x, z_1)$ é menor que $s(x, z_2)$. Notar-se-á, entretanto, que o uso desta medida de similaridade é governada por certas restrições, como por exemplo, uma separação adequada entre regiões de *aglomerados*.

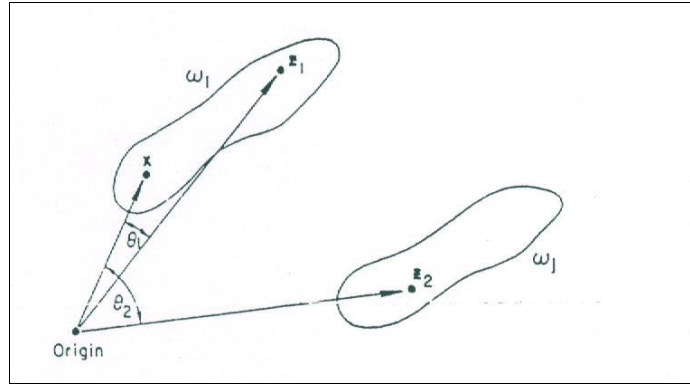


Figura 4.1 Ilustração de medida de similaridade

Quando o conjunto de dados é constituído de valores binários, a função de similaridade da equação 4.3 pode fornecer uma interpretação geométrica interessante. Diz-se que o padrão binário x possui o i ésimo atributo se $x_i = 1$. Então o termo $x'z$ na equação 4.3 é simplesmente o número de atributos compartilhados por x e z , enquanto $\|x\| \|z\| = \sqrt{(x'x)(z'z)}$ é a média geométrica do número de atributos que x e z possuem. Neste caso a função de similaridade $s(x, z)$ é, conseqüentemente, vista como uma função de atributos comuns entre x e z .

A variação binária da equação 4.3 tem recebido atenção das áreas de recuperação de informação, nozologia e taxonomia, é chamada de medida de Tanimoto

$$s(x, z) = x'z / x'x + z'z - x'z \quad (4.4)$$

- *Critério de Aglomeração*: Depois de ter adotado uma medida de similaridade, o desafio consiste em dividir o conjunto de dados em aglomerados sensíveis. Isto depende da interpretação que o especialista dá ao termo ‘sensível’, baseado no tipo de aglomerados esperados para cobrir o conjunto de dados. Por exemplo, um aglomerado compacto de vetores de características no espaço de dimensão 1, pode ser sensível a um critério, visto que um aglomerado alongado pode ser sensível em função de outro aglomerado. O critério de aglomeração pode ser expresso via uma função de custo ou pode representar uma heurística, ou ainda pode ser baseado na minimização (ou maximização) de certos índices de performance.

A abordagem heurística é guiada pela intuição e experiência. Ela consiste na especificação de um conjunto de regras cujo uso explora a escolha da medida de similaridade para ordenar a atribuição dos padrões para o domínio do aglomerado. A distância Euclidiana, equação 4.1, presta-se a esta abordagem por causa da sua interpretação familiar como uma medida de similaridade. Entretanto, já que a proximidade de dois padrões é uma medida relativa de similaridade, é usualmente necessário estabelecer um limiar para definir o grau de aceitação de similaridade no processo de atribuição do aglomerado. O algoritmo mostrado a seguir ilustra esta abordagem.

A abordagem por índice de performance é guiada pelo desenvolvimento do procedimento que minimiza ou maximiza a escolha do índice de performance. Um dos índices mais usados é o somatório dos erros quadráticos, dados por:

$$J = \sum_{j=1}^{N_c} \sum_{x \in S_j} \|x - m_j\|^2 \quad (4.5)$$

onde N_c é o número de *aglomerado*, S_j é o conjunto de amostras pertencendo ao j *ésimo* aglomerado e

$$m_j = \frac{1}{N_j} \sum_{x \in S_j} x \quad (4.6)$$

é o vetor médio das amostras do conjunto S_j . Na equação 4.6 N_j represente o número de amostras em S_j . O índice da equação 4.5 representa a soma total dos erros quadráticos entre as amostras do *aglomerado* e suas médias correspondentes. Há, naturalmente,

inúmeros índices de performances como a média da distância quadrada entre amostras do aglomerado, a média da distância quadrada entre amostras de diferentes aglomerados, o índice da máxima variância e mínima variância.

Não é incomum encontrar algoritmo de *clustering-seeking* que representa uma combinação da abordagem heurística e índices de performance. O algoritmo *Isodata* é um destes.

- *Algoritmo de Aglomeração*: Tendo adotado uma medida de proximidade e um critério de agrupamento, este passo refere-se à escolha de um algoritmo específico que cria a estrutura de *aglomeração* de um conjunto de dados;

- *Validação dos Resultados*: Uma vez obtidos os resultados do algoritmo de *aglomeração*, tem-se que verificar sua exatidão. Isto é geralmente realizado usando testes apropriados;

- *Interpretação dos Resultados*: Em muitos casos, o especialista do campo da aplicação precisa integrar o resultado do *aglomeração* com outras evidências experimentais e analisar o conjunto todo para obter uma solução conclusiva.

4.2 Algoritmo K-Means

O algoritmo descrito abaixo [TOU, GONZALEZ 1974] é baseado na minimização do índice de performance que é definido como a soma das distâncias quadráticas de todos os pontos do aglomerado em relação a seu centro. Este procedimento consiste dos seguintes passos:

- *Passo 1*: Escolha K centros iniciais dos aglomerados $z_1(1), z_2(1), \dots, z_K(1)$. Esta seleção é arbitrária e usualmente seleciona aleatoriamente os centros iniciais;
- *Passo 2*: A k ésima iteração distribui as amostras $\{x\}$ entre os K aglomerados, usando a relação:

$$X \in S_j(k) \text{ se } \|x - z_j(k)\| < \|x - z_i(k)\| \quad (4.7)$$

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

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

$$J_j = \sum_{x \in S_j(k)} \|x - z_j(k+1)\|^2, \quad j = 1, 2, \dots, K \quad (4.8)$$

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

$$z_j(k+1) = \frac{1}{N_j} \sum_{x \in S_j(k)} x \quad j = 1, 2, \dots, K \quad (4.9)$$

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 aglomerados 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 *aglomerados* especificados, pela escolha inicial dos centros, pela ordem de escolha das amostras e, naturalmente, pelas propriedades geométricas dos dados.

A Figura 4.2 ilustra a evolução do processo de *aglomeração* por *K-means* em função do número K de aglomerados.

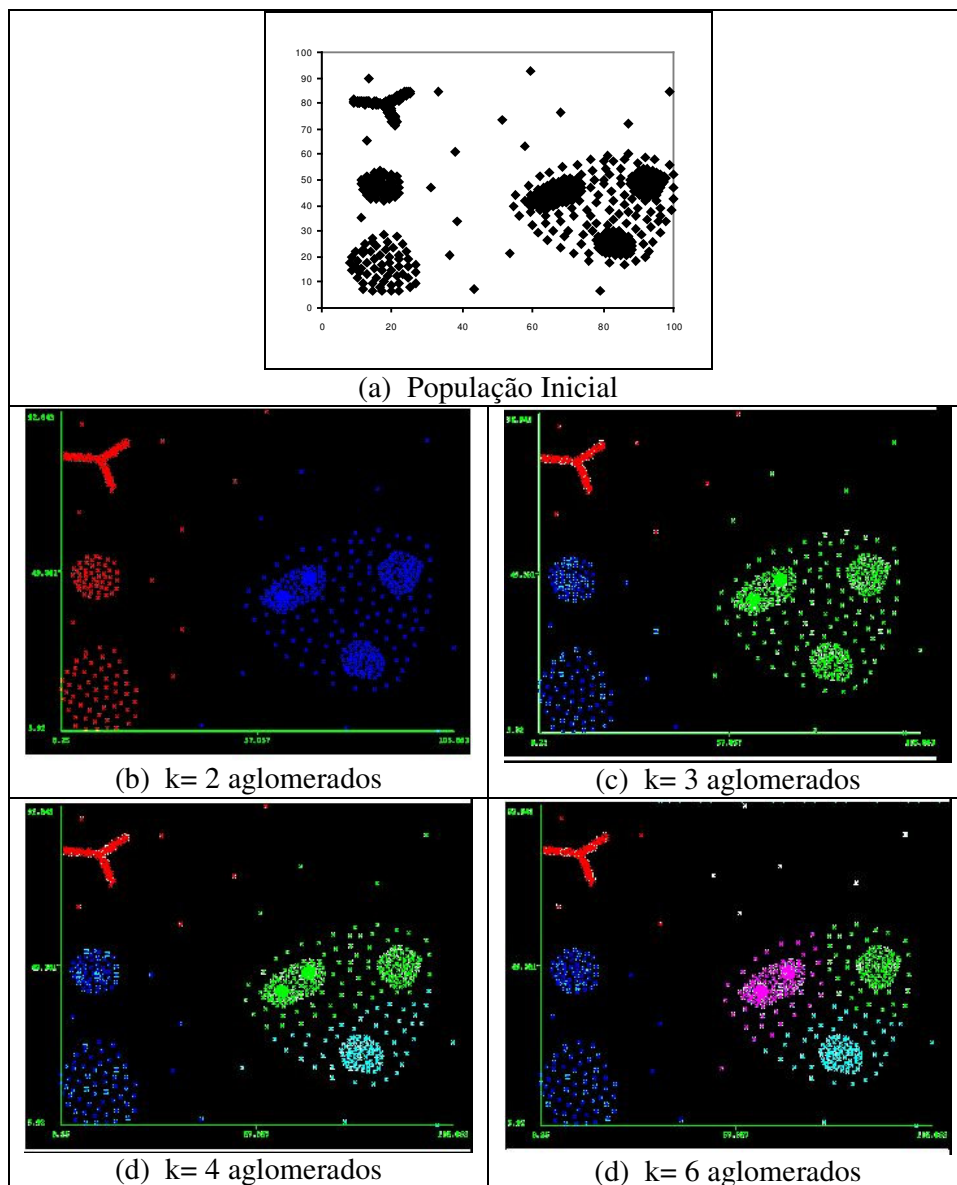


Figura 4.2: Exemplo de evolução do processo de *aglomeração* por *K-means* em função do número K de aglomerados

4.3 Algoritmo Isodata

O algoritmo *Isodata* apresentado aqui é similar, em princípio, ao procedimento do *K-means* no sentido que os centros dos *aglomerados* são iterativamente determinados pela média das amostras[TOU, GONZALEZ 1974]. Ao contrário do *K-means*, *Isodata* apresenta um conjunto adicional de procedimentos heurísticos os quais foram incorporados em um esquema iterativo. A Heurística, neste sentido, deve se entendido

com sendo o resultado de experiências obtidas através de vários testes e, portanto significando um progresso, um avanço.

Antes de executar o algoritmo, é necessário especificar N_c o conjunto dos centros de aglomerado $z_1, z_2, \dots, z_{N_c}$. Este conjunto, que não precisa ser necessariamente igual ao número aos centros dos aglomerados desejados, pode ser formado pela seleção de amostras de um conjunto de dados.

Para um conjunto de N amostras, $\{x_1, x_2, \dots, x_N\}$, o algoritmo *Isodata* consiste nos seguintes passos.

- **Passo 1:** Especificar os seguintes parâmetros:

K = número de centros de aglomerado desejado

θ_N = um parâmetro contra o qual o número de amostras do domínio do *aglomerado* é comparado

θ_s = desvio padrão

θ_c = *lumping*

L = número máximo de pares de centros de aglomerado o qual podem ser agregados

I = número de iterações permitidas

- **Passo 2:** Distribuir as N amostras entre os centros dos aglomerados, usando a relação

$$x \in S_j \text{ se } \|x - z_j\| < \|x - z_i\|, \quad i = 1, 2, \dots, N_c; \quad i \neq j \quad (4.10)$$

para todo x do conjunto de amostras. Nesta notação, S_j representa o sub conjunto de amostras correspondentes ao centro de *aglomerado* z_j

- **Passo 3:** Descartar as amostras do sub conjunto com menos do que θ_N membros; isto é, se para qualquer j , $N_j < \theta_N$, descarte S_j e reduza N_c de 1.

- **Passo 4:** Atualize cada centro de aglomerado z_j , $j = 1, 2, \dots, N_c$ pela média das amostras do seu conjunto correspondente S_j , que é,

$$z_j = \frac{1}{N_j} \sum_{x \in S_j} x, \quad j = 1, 2, \dots, N_c \quad (4.11)$$

onde N_j é o número de amostras em S_j .

- **Passo 5:** Calcule a distância média D_j das amostras do aglomerado S_j ao seu correspondente centro de aglomerado, usando a relação

$$D_j = \frac{1}{N_j} \sum_{x \in S_j} \|x - z_j\|, \quad j = 1, 2, \dots, N_c \quad (4.12)$$

- **Passo 6:** Calcule a distância média total das amostras de seus respectivos centros dos aglomerados, usando a relação

$$D = \frac{1}{N} \sum_{j=1}^{N_c} N_j D_j \quad (4.13)$$

- **Passo 7:** (a) Se esta é a ultima iteração, faça $\theta_c = 0$ e pule para o passo 11. (b) Se $N_c \leq K/2$, siga ao passo 8. (c) Se esta é uma iteração par, ou se $N_c \geq K/2$, pule para o passo 11; senão, continue.

- **Passo 8:** Ache o vetor desvio padrão $\sigma_j = (\sigma_{1j}, \sigma_{2j}, \dots, \sigma_{nj})$ para o sub conjunto de amostras usando a relação

$$\sigma_{ij} = \sqrt{\frac{1}{N_j} \sum_{x \in S_j} (x_{ik} - z_{ij})^2}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, N_c \quad (4.14)$$

onde n é a dimensionalidade da amostra, x_{ik} é o i ésimo componente da k ésima amostra em S_j , z_{ij} é o i ésimo componente de z_j , e N_j é o número de

amostras em S_j . Cada componente de σ_j representa o desvio padrão das amostras em S_j ao longo do eixo de coordenada principal.

- **Passo 9:** Ache o componente máximo de cada de σ_j , $j = 1, 2, \dots, N_c$, e faça-o $\sigma_{j \max}$

- **Passo 10:** Se para qualquer $\sigma_{j \max}$, $j = 1, 2, \dots, N_c$, tem-se $\sigma_{j \max} > \theta_s$ e

(a) $D_j > D$ e $N_j > 2(\theta_N + 1)$ ou

(b) $N_c \leq K/2$

divida z_j em dois novos centros de aglomerado z_{j+} e z_{j-} , apague z_j , e adicione 1 a N_c . O centro de aglomerado z_{j+} é formado pela adição de uma dada quantidade γ_j para o componente de z_j o qual corresponde ao componente máximo de σ_j ; z_{j-} é formado subtraindo γ_j do mesmo componente de z_j . Uma maneira de especificação de γ_j é fazê-lo igual a alguma fração de $\sigma_{j \max}$, isto é, $\gamma_j = k \sigma_{j \max}$, onde $0 < k \leq 1$. A condição básica para escolha de γ_j , é que ele seja suficiente grande para detectar uma diferença na distância de uma determinada amostra e os dois novos centros de aglomerados, mas não tão grande assim para mudar os aglomerados.

Se a divisão ocorrer neste passo, siga para o passo 2; senão continue.

- **Passo 11:** Calcule a distância D_{ij} entre pares (z_i, z_j) de centros para todos os centros dos *aglomerados*:

$$D_{ij} = \|z_i - z_j\|, \quad i = 1, 2, \dots, N_c - 1; \quad j = i + 1, \dots, N_c \quad (4.15)$$

- **Passo 12:** Compare as distâncias D_{ij} com o parâmetro θ_c . Liste as menores distâncias L menores que θ_c , em ordem ascendente, onde $D_{i_1 j_1} < D_{i_2 j_2} < \dots < D_{i_L j_L}$ e L é o número máximo de pares de centros dos aglomerados que podem ser agregados.
- **Passo 13:** Em cada distância D_{iljl} , há associado um par de centros de aglomerados z_{i_l} e z_{j_l} . Começando pela menor destas distâncias, execute uma operação de agregação do par, de acordo com a seguinte regra :

$$z_l^* = 1 / N_{il} + N_{jl} [N_{il}(z_{il}) + N_{jl}(z_{jl})] \quad (4.16)$$

Apague z_{il} e z_{jl} e diminua 1 de N_c .

Percebe-se que somente um par agrupado é permitido e que um centro de um aglomerado consolidado é obtido por ponderação do centro de cada *aglomerado* velho pelo número de amostras. Evidências experimentais indicam que um agrupamento mais complexo pode produzir resultados insatisfatórios. O mecanismo acima citado faz com que centros agrupados sejam mais representativos do ponto médio verdadeiro do conjunto. É importante ressaltar também que, já que um centro pode ser consolidado uma única vez, esta etapa pode nem sempre resultar em L centros consolidados.

- **Passo 14:** Se esta é a última iteração, o algoritmo termina. Senão, volte no passo 1 se qualquer dos parâmetros do procedimento requer mudanças pelo critério do usuário, ou volte no passo 2 se os parâmetros permanecem os mesmos para a próxima iteração. Uma iteração é contada a cada vez que o procedimento retorna ao passo 1 ou 2.

Capítulo 5

Método Proposto

5.1 Uma Abordagem Fractal

A abordagem proposta consiste na segmentação do bloco de endereço usando o cálculo da dimensão fractal, rotulando as classes através da técnica de *aglomeração* do *K-means* e identificando-as através do conceito da matriz de co-ocorrência (Figura 5.1).

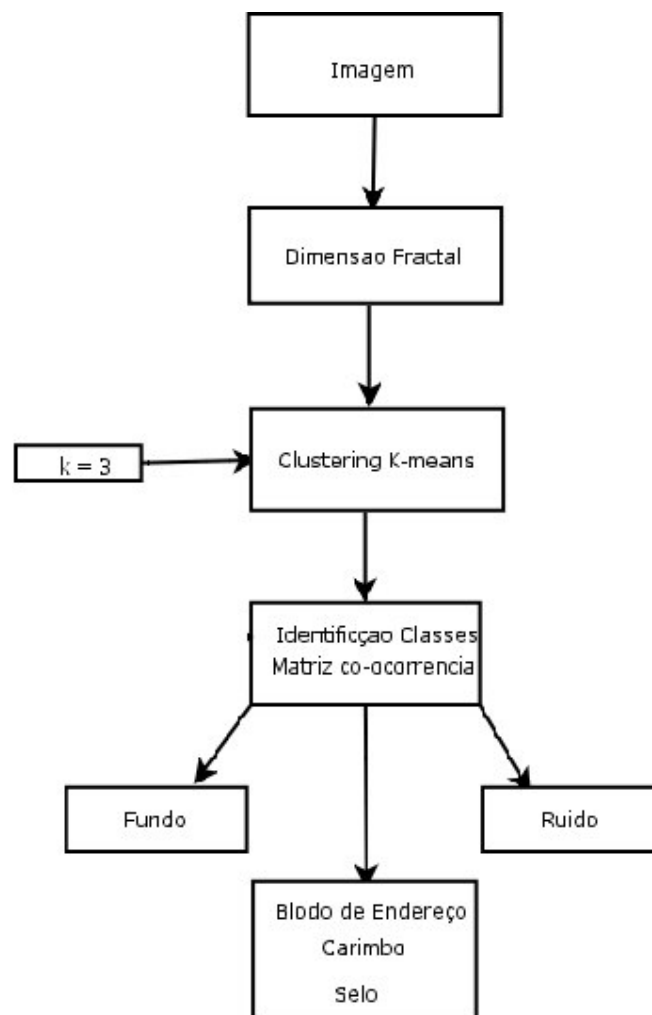


Figura 5.1 Esquema da Abordagem de Segmentação Fractal

5.2 Cálculo da Dimensão Fractal

Observou-se que a técnica de cálculo da dimensão fractal *2D variation procedure* é menos complexa do que outras disponíveis (por exemplo, a técnica baseada em coberturas morfológicas) e tem mostrado ser promissora em imagens complexas [CHEN *et al* 2001]. Por estes motivos foi a escolhida. Aplicou-se o método *2D variation procedure* (seção 3.4.1) em imagens de envelopes postais da base de imagens reais para seis intervalos de tamanho de janela : $r = \{3, 5\}$, $r = \{3, 5, 7\}$, $r = \{3, 5, 7, 9\}$, $r = \{3, 5, 7, 9, 11\}$, $r = \{3, 5, 7, 9, 11, 13\}$, $r = \{3, 5, 7, 9, 11, 13, 15\}$; e para a base de imagens sintetizadas as faixas foram sete : $r = \{3, 5\}$, $r = \{3, 5, 7\}$, $r = \{3, 5, 7, 9\}$, $r = \{3, 5, 7, 9, 11\}$, $r = \{3, 5, 7, 9, 11, 13\}$, $r = \{3, 5, 7, 9, 11, 13, 15\}$, $r = \{3, 5, 7, 9, 11, 13, 15, 17\}$. O resultados da DF da região das janelas atribuído a cada *pixel* num intervalo de [2.0, 3.0], é fornecido pela equação 3.12.

A imagem chamada dimensão fractal gerada foi normalizada entre 0-255.

5.3 Algoritmo de Aglomeração *K-means*

A escolha deste algoritmo de separação das classes fez –se pelo critério da independência, do automatismo, ou seja, por não necessitar de treinamento.

Os parâmetros iniciais (padrões) usados no *K-means* são:

- número de características igual a 1, a própria DF
- número de aglomerados 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

É bem conhecido que o algoritmo *k-means* varia de acordo com as condições iniciais de processamento. Encontra-se na literatura algumas pesquisas sobre melhoras, otimizações na inicialização iterativa do algoritmo [FAYAD, REINA, BRADLEY 1998]. Neste trabalho usou-se min, $(\max + \min)/2$ e max, para inicializar o valor dos centróides dos aglomerados, onde max e min são respectivamente os valores da máxima e da mínima

DF do conjunto. A convergência do algoritmo *K-means* em uma imagem I_{DF} resulta em uma imagem rotulada $I_{k-means}$ onde cada *pixel* é rotulado em um dos três aglomerados.

5.4 Matriz de Co-Ocorrência

A matriz de co-ocorrência de uma imagem é uma matriz $T = [t_{ij}(v, h)]_{L \times L}$ de dimensão $(L \times L)$ que contém informação a respeito das transições de intensidade entre *pixels* adjacentes [PAL 1989]. Em outros termos, $t_{ij}(v, h)$ representa o número de vezes que o par de níveis de cinza (i, j) ocorre para dois *pixels* (k, l) e (m, n) distantes de h e v , respectivamente na horizontal e na vertical.

Pode-se formalizar $t_{ij}(v, h)$ da seguinte maneira:

$$t_{ij}(v, h) = \sum_{l=0}^{M-v} \sum_{k=0}^{N-h} \delta \quad (5.5)$$

em que:

$$\delta = 1 \text{ se } f(k, l) = i \text{ e } f(m, n) = j \text{ para } |k - m| = h \text{ e } |l - n| = v \quad (5.6)$$

$\delta = 0$, caso contrário.

h e v representam a distância entre os dois pontos respectivamente na horizontal e na vertical.

5.4.1. Definição de [HARALICK 1992]

Segundo Haralick [HARALICK 1992], que define a matriz de co-ocorrência como sendo simétrica e como função do relacionamento angular e da distância entre *pixels*. Abaixo, segue um exemplo de uma imagem de dimensões 4x4 com 4 níveis de cinza (0 a 3):

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figura 5.2 Imagem de dimensões 4x4

e níveis de cinza correspondentes

Pode-se formalizar a matriz de co-ocorrência de acordo com diversos relacionamentos angulares bem como a distância entre os *pixels*:

$$\begin{aligned} C(i, j, d, 0^\circ) = \\ \# \{ [(k, l), (m, n)] \mid \\ k - m = 0, |l - n| = d \\ I(k, l) = i, I(m, n) = j \} \end{aligned}$$

$$\begin{aligned} C(i, j, d, 90^\circ) = \\ \# \{ [(k, l), (m, n)] \mid \\ |k - m| = d, l - n = 0 \\ I(k, l) = i, I(m, n) = j \} \end{aligned}$$

$$\begin{aligned} C(i, j, d, 45^\circ) = \\ \# \{ [(k, l), (m, n)] \mid \\ (k - m = d, l - n = -d) \\ \text{ou} \\ (k - m = -d, l - n = d) \\ I(k, l) = i, I(m, n) = j \} \\ C(i, j, d, 135^\circ) = \\ \# \{ [(k, l), (m, n)] \mid \\ (k - m = d, l - n = d) \\ \text{ou} \\ (k - m = -d, l - n = -d) \\ I(k, l) = i, I(m, n) = j \} \end{aligned}$$

em que:

o símbolo # denota o número de elementos do conjunto

0° = Horizontal, 90° = Vertical, 45° = Diagonal superior e 135° = Diagonal inferior

d é a distância entre os *pixels*

C denota as frequências de ocorrência do par (i, j) da matriz de co-ocorrência

$$0^\circ \quad \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$45^\circ \quad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$90^\circ \quad \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

$$135^\circ \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

Figura 5.3 Resultados da matriz de co-ocorrência para diversos ângulos e distância 1 para o exemplo de imagem da Figura 5.2

5.5. Identificação das Classes

Para segmentar o aglomerado do bloco de endereço, o carimbo e os selo é necessário sua identificação entre os três *aglomerados* 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 3×3 que contem 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 :

$$c_{ij}(dx; dy) = \# \{ (x, y) / I_{K-means}(x, y) = i, I_{K-means}(x + dx; y + dy) = j \} \quad (5.7)$$

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

$$C = C_{hor} + C_{ver} + C_{diag} = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{bmatrix} \quad (5.8)$$

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

$$Acc_k = \sum_{i=1}^3 c_{i,k} + \sum_{j=1}^3 c_{k,j}, \text{ onde } 1 \leq k \leq 3 \quad (5.9)$$

O aglomerado do fundo do envelope é identificado pela seleção do *aglomerado* k o qual Acc_k é máximo:

$$K_{fundo} = \text{Arg } k \max (Acc_k) \quad (5.10)$$

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

$$K_{blocoendereço} = \text{Arg } k \min [c_k, k_{fundo}, c_{k_{fundo}}, \kappa] \quad (5.11)$$

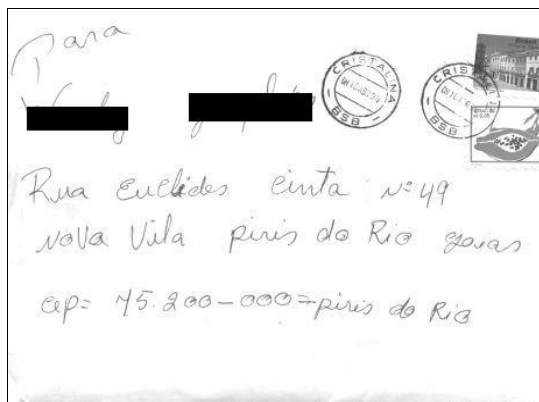
Onde $1 \leq k \leq 3$ e $k \neq k_{fundo}$ e $N_l > 1,5\% \times N_I$. N_l é o numero de *pixels* que pertence ao aglomerado l , e N_I é o numero total de *pixels* na imagem original I . O fator 1,5% foi obtido depois de estudos empíricos das imagens de envelopes. Se o aglomerado é muito pequeno, provavelmente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

Verificou-se após os testes como mostrado detalhadamente a seguir no item 5.6, que a variação dos parâmetros iniciais do *K-means* não aumenta significativamente a qualidade da segmentação. No parâmetro relacionado ao número de iterações praticamente não há alteração do resultado. Nos parâmetros valor de melhoria mínima e número de classes os resultados apresentam uma sensível diminuição na qualidade da segmentação.

Testes, Resultados e Análises

6.1 Base de Imagens

A base de imagens reais foi adquirida a 200 *dpi* e é composta de 101 imagens com fundo complexo e 99 com fundo branco (figura 6.1). Desta base foram segmentados manualmente o carimbo, o selo e o bloco de endereço criando o resultado esperado do processo de segmentação, gerando assim as imagens *ground truth*.



(b)

Figura 6.1 Imagens Reais (a) fundo complexo (b) fundo branco

Segundo Menoti [MENOTI 2003] a base de imagem sintetizada (Figura 6.2), foi criada, no intuito de possibilitar maior volume de dados para os testes, a partir das imagens de referências (*ground truth*) segmentadas manualmente, bloco de endereço, do selo e do carimbo da base de imagens reais e dos tipos de envelopes postais brasileiros com fundo complexo (Figura 6.3).

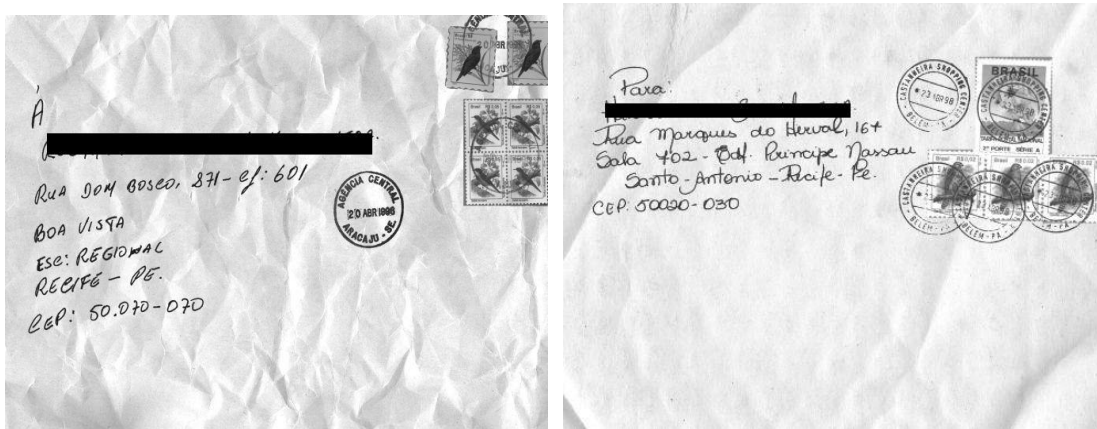


Figura 6.2 Imagens Sintetizadas

As imagens reais e sintetizadas (com quantidade 10 vezes maior) têm as seguintes dimensões médias mais o desvio padrão: largura de $1.857,8 \pm 20,5$ pixels, altura de $1.340,0 \pm 41,9$ pixels e um número total de pixels de $2.549.180,2 \pm 56.822,6$.

O algoritmo de segmentação propõe tratar e diferenciar os diversos tipos de informações contidas nos envelopes postais, de tal maneira a facilitar a localização e extração do bloco de endereço, dos selos e carimbos. As características encontradas nos envelopes são os fundos variados, envelopes amassados, variação do tamanho e espessura das letras, variação de iluminação durante a aquisição, textos manuscritos adicionais, selos de tamanho e posição variável.

6.2 Imagem de Referência (*ground truth*)

Menoti [MENOTI 2003] ressalta a necessidade de um método de avaliação precisa do resultado para a área de segmentação de imagens, como um todo. Entre os poucos e particulares exemplos de avaliação de resultados, a abordagem *ground truth* vem para

ocupar esta lacuna de forma abrangente e precisa. A tradução do termo *ground truth* é base de referência ou abordagem verdadeira. A imagem *ground truth* é a imagem ideal que se espera como resultado de um algoritmo de segmentação (Figura 6.3(a)).

As imagens *ground truth* foram geradas manualmente, e é vital para a avaliação quantitativa da metodologia proposta, pois são comparadas *pixel a pixel* com o resultado da segmentação por classe.

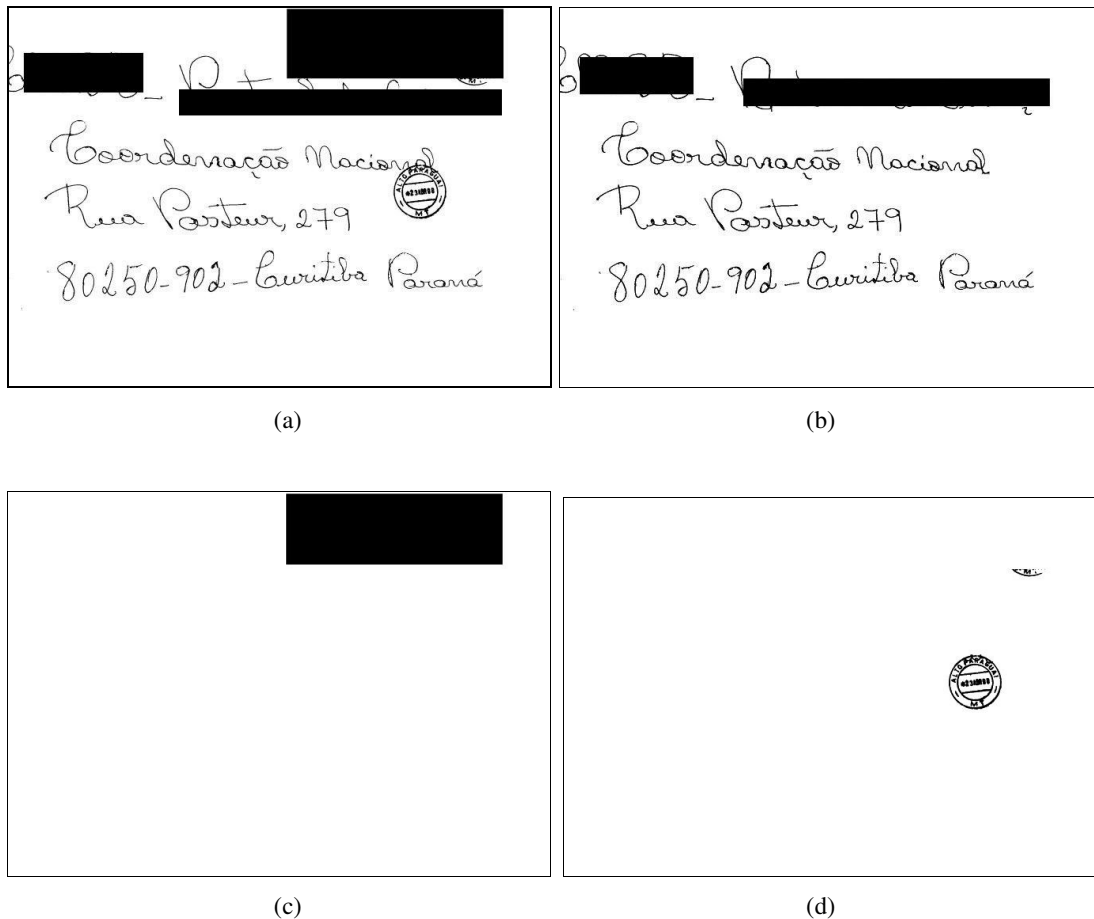


Figura 6.3 Imagens *ground truth* (referência) (a) Resultado da Segmentação Desejada (b) Imagem do bloco de endereço (c) Imagem do selo (d) Imagem do carimbo

6.3 Avaliação da Metodologia de Segmentação

Pelo uso da abordagem *ground truth* é possível uma avaliação precisa dos resultados dos experimentos. Uma avaliação *pixel a pixel* comparando a segmentação obtida com a

solução ideal *ground truth* é a metodologia de avaliação usada. Foi gerada manualmente para cada imagem a ser segmentada uma imagem para cada classe com os *pixels* que devem ser segmentados, ilustrados pela Figura 6.3 (b) (c) e (d).

Por motivo de avaliação classe a classe dos resultados da segmentação, estas imagens foram geradas em separado; entretanto o resultado final mostrado pelo algoritmo de segmentação e o desejado mostram somente uma classe, como mostra a Figura 6.3(a).

Menoti [MENOTI 2003] apresenta na Figura 6.4 o esquema de avaliação *pixel a pixel* das imagens dos envelopes postais. A imagem que é fornecida pelo algoritmo de segmentação I_{Final} é comparada com as imagens de referência (imagens esperadas) Assim cada *pixel* de I_{Final} que foi segmentado é comparado com o *pixel* respectivo em I_{GT1} , I_{GT2} , I_{GT3} . Para pertencer a uma determinada classe, o *pixel* I_{Final} , deve ser encontrado em uma dessas três imagens e em só uma delas. De outro modo, o *pixel* é entendido como ruído e pertencente ao fundo do envelope. As imagens I_{Ob1} , I_{Ob2} , I_{Ob3} e $I_{Ruído}$ são geradas a partir da comparação dos *pixels* de I_{Final} com I_{GT1} , I_{GT2} , I_{GT3} . A taxa de retorno das classes é a razão entre o número de *pixels* existentes nas imagens observadas (geradas pelo algoritmo de segmentação) I_{Ob1} , I_{Ob2} , I_{Ob3} e as imagens de referência (*ground truth*) I_{GT1} , I_{GT2} , I_{GT3} . Para se calcular a medida do ruído na Segmentação, dividi-se o número de *pixels* que foram segmentadas em I_{Final} e não pertencem a I_{GT1} , I_{GT2} e I_{GT3} e o número de *pixels* que pertencem ao fundo do envelope. Estas taxas de retorno e ruído podem ser definidos como:

$$TR_{Classe\ 1} = \frac{n^{\circ} \text{pixel} (I_{FINAL} \cap I_{GT1})}{n^{\circ} \text{pixel} (I_{GT1})} = \frac{n^{\circ} \text{pixel} (I_{Ob1})}{n^{\circ} \text{pixel} (I_{GT1})} \quad (6.1)$$

$$TR_{Classe\ 2} = \frac{n^{\circ} \text{pixel} (I_{FINAL} \cap I_{GT2})}{n^{\circ} \text{pixel} (I_{GT2})} = \frac{n^{\circ} \text{pixel} (I_{Ob2})}{n^{\circ} \text{pixel} (I_{GT2})} \quad (6.2)$$

$$TR_{Classe\ 3} = \frac{n^{\circ} \text{pixel} (I_{FINAL} \cap I_{GT3})}{n^{\circ} \text{pixel} (I_{GT3})} = \frac{n^{\circ} \text{pixel} (I_{Ob3})}{n^{\circ} \text{pixel} (I_{GT3})} \quad (6.3)$$

$$Ruído = \frac{n^{\circ} \text{pixel} (I_{FINAL} - (I_{GT1} \cup I_{GT2} \cup I_{GT3}))}{n^{\circ} \text{pixel} (I_{GT1} \cup I_{GT2} \cup I_{GT3})} \quad (6.4)$$

Onde $TR_{Classe\ 1}$, $TR_{Classe\ 2}$, $TR_{Classe\ 3}$ são as taxas de retorno das classes 1 (bloco de endereço), 2 (selo) e 3 (carimbo) respectivamente; $I_{Final} \cap I_{GT1}$, $I_{Final} \cap I_{GT2}$, $I_{Final} \cap I_{GT3}$

são os conjuntos de intersecção entre a imagem segmentada I_{Final} e as imagens *ground truth* (de referência) I_{GT1} , I_{GT2} e I_{GT3} , respectivamente que representam os conjuntos I_{Ob1} , I_{Ob2} , I_{Ob3} ; $(I_{GT1} \cup I_{GT2} \cup I_{GT3})$ é o conjunto de *pixels* que representam todos os objetos de segmentação; $\overline{(I_{GT1} \cup I_{GT2} \cup I_{GT3})}$ é o complemento do conjunto $(I_{GT1} \cup I_{GT2} \cup I_{GT3})$; e o número de *pixel* (X) é o número de *pixels* existentes no conjunto X .

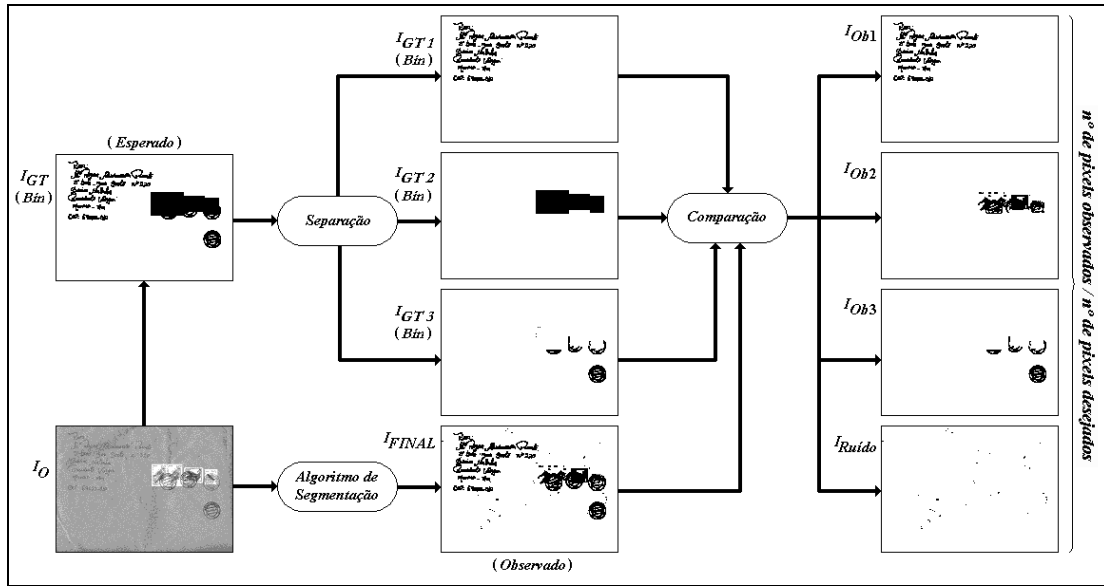


Figura 6.4: Esquema de avaliação das imagens de envelopes postais. I_O é a imagem original; I_{GT} é a segmentação esperada (ou desejada) para I_O ; I_{GT1} , I_{GT2} e I_{GT3} são as imagens *ground-truth* (de referência) das classes bloco endereço, selo e carimbo, respectivamente, previamente separadas; I_{FINAL} é o resultado final do algoritmo de segmentação; I_{Ob1} , I_{Ob2} e I_{Ob3} são as imagens que contém os *pixels* segmentados de cada uma das classes: bloco endereço, selo e carimbo, respectivamente; I_{Ruido} é a imagem que contém os *pixels* segmentados que não pertencem a nenhuma das três classes.

Segue exemplo usado em Menoti [MENOTI 2003] para ilustrar o método de Avaliação. Uma imagem contém 10.000 *pixels* distribuídos conforme a tabela 6.1. Na coluna “*pixels* observados” estão os *pixels* presentes na imagem I_{Final} obtidos pelo algoritmo de segmentação, enquanto que na coluna “*pixels* desejados” estão os *pixels* das imagens de referência. A taxa de retorno de cada uma das classes é a razão entre as duas colunas. São considerados ruídos e de propriedade do fundo do envelope os *pixels* da imagem segmentada I_{Final} que não pertencem a nenhuma das três classes.

Classes	nº de <i>pixels</i> desejados	nº de <i>pixels</i> observados	nº observados / nº desejados
Objetos	650	375	57,69
Bloco Endereço	150	145	96,57
Selo	400	150	37,50
Carimbo	100	80	80,00
Fundo do Envelope/ Ruído	9.300	93	01,00
Total	10.000	763	07,63

Tabela 6.1 Exemplo Hipotético para avaliação

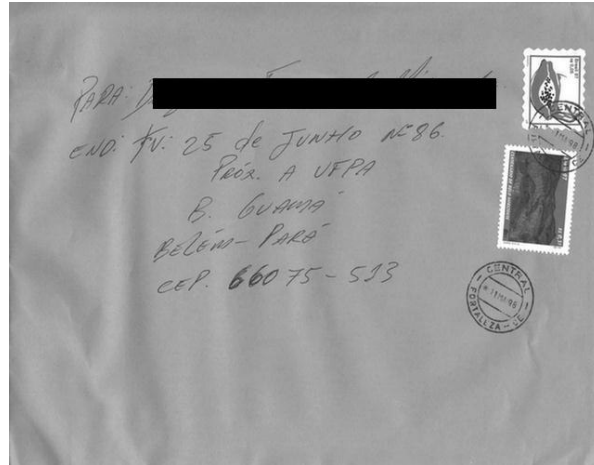
Foi necessário mostrar a abordagem *ground truth* e o descritivo das bases de imagens usadas, para apresentar a metodologia de avaliação quantitativa que permitirá uma análise dos testes por um critério objetivo.

6.4 Resultados e Análises

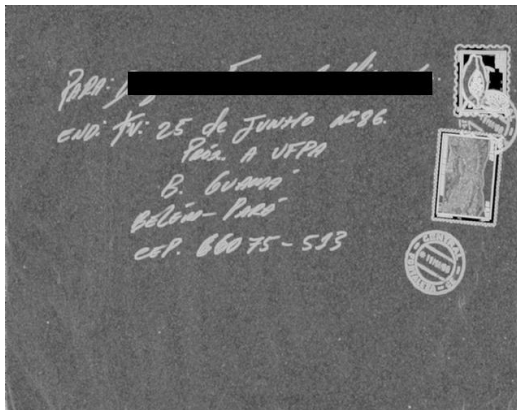
São apresentados neste item gráficos, tabelas e imagens dos resultados dos testes. Também são feitas análises realizadas a partir do método de avaliação dos resultados obtidos *pixel a pixel*, que fornece com precisão uma avaliação da segmentação obtida através do algoritmo de segmentação proposto.

O algoritmo de segmentação deve tratar as variações de diagramação do envelope e não estar preso a esta. Deve ainda tratar todas as possibilidades de variações de informação existente nas imagens do envelope, tais como:

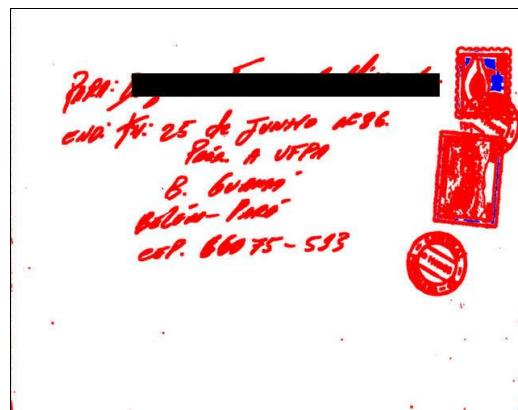
- a complexidade do fundo do envelope
- o tamanho e a posição do bloco de endereço
- a posição, a quantidade de carimbos, assim como a possibilidade de sobreposição destes
- a quantidade e a posição dos selos



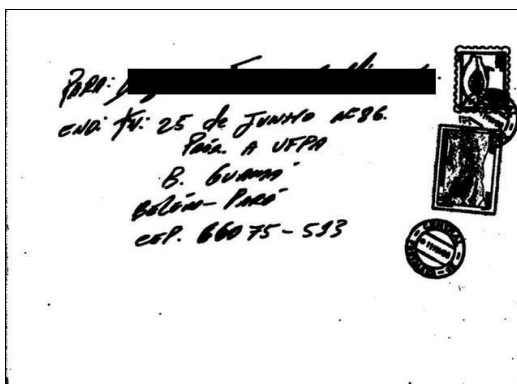
(a)



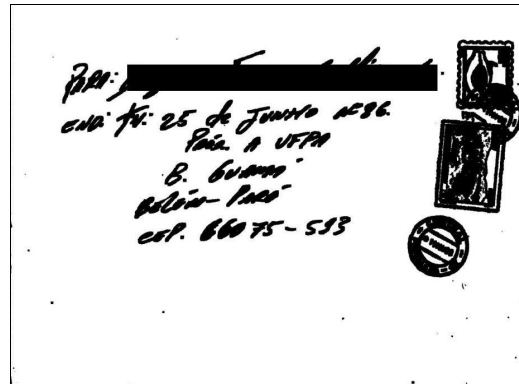
(b)



(c)



(d)



(e)

Figura 6.8 Imagem do Envelope Postal da Base Real (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, (e) Imagem *ground truth* (ideal)

Neste trabalho, busca-se a segmentação do bloco de endereço, que é a informação essencial na imagem do envelope de correio. Igualmente importante é a taxa de segmentação do selo e do carimbo, por mostrarem a abrangência e robustez do algoritmo proposto, pois ele lida com tipos diferentes de imagens, como as manuscritas (bloco de endereço), as pré-impressas (carimbo) e as figuras, paisagens, texturas, padrões (selo) contidos na imagem de envelope postal dos Correios. A taxa do ruído é fator importante para qualidade dos resultados da segmentação como um todo.

As conclusões obtidas a partir dos testes foram tomadas considerando a média dos valores das taxas de segmentação e ruído do processamento de todas as imagens reais e sintetizadas. A Figura 6.8 mostra uma imagem típica que tem valores relativamente próximos a esta média.

6.4.1 Análise dos Resultados da imagem Carta040 da Base Real

Apresenta-se o processamento da imagem carta040 da base de imagens reais como exemplo dos passos realizados pelo algoritmo de segmentação (Figura 6.8), e as imagens resultantes da variação dos parâmetros usados até o ponto de convergência. A imagem carta040 foi escolhida por ser uma imagem complexa, apresentando o fundo bem amassado, e uma diagramação fora do padrão dos correios. Os selos apresentam-se retangulares de tamanho variável, com figuras, padrões e texturas diferentes, sobrepostos por carimbo. A parte manuscrita mostra-se bem inclinada. Os dois carimbos redondos apresentam-se, um inteiramente sobre o fundo do envelope e o outro sobre dois selos diferentes e o fundo do envelope. Nas tabelas 6.2 a 6.7 estão todos os dados e resultados dos testes para a imagem real carta040. A coluna Parâmetros apresenta:

- entre colchetes a faixa do tamanho da janela em *pixels* para extração da DF
- o parâmetro do *K-means* c , que indica o número de *aglomerados*
- o parâmetro do *K-means* k , que indica o número de iterações
- o parâmetro do *K-means* i , que indica o valor de melhoria mínima

Quando omitidos os valores padrões (iniciais) usados no *K-means* são: $c = 3$, $k = 10$ e $i = 0,0001$.

A Tabela 6.7 mostra o melhor valor médio da taxa de segmentação para o intervalo [3,15]. São utilizadas para mostrar os passos do processamento do algoritmo de segmentação e a evolução dos seus resultados, apenas as imagens das variações da faixa do tamanho da janela que utilizam os parâmetros padrões (iniciais) do *K-means* que são: $c = 3$, $k = 10$ e $i = 0,0001$. No caso específico da imagem carta040, os melhores valores convergem quando se usa somente duas classes no *K-means* ($c = 2$) e o intervalo do tamanho de janelas é [3,15]. Exemplos das imagens com parâmetros $c = 2$ e $i = 0,001$ para o intervalo de tamanho de janelas [3,5] a [3,15] estão disponíveis no Anexo A.1

Análise dos Resultados do intervalo de tamanho de janelas[3,5] da imagem Carta040

Na tabela 6.2 verifica-se que os valores das taxas de segmentação do bloco de endereço, selo, carimbo e ruído, aumentam e diminuem nos mesmos grupos de parâmetros. Apresentam-se bem constante com valor igual a 99,38%, 48,07%, 96,02% e 2,84% respectivamente. Há uma pequena melhora quando diminui o número de classes c , para dois no parâmetro [3,5]c2. Há um pequeno decréscimo nas taxas de segmentação ocorrido nos parâmetros [3,5]k55 e [3,5]k40i001. Os melhores valores das taxas são obtidos para os parâmetros [3,5]c2, o que nos mostra que para este intervalo de janelas a diminuição do número de classe para dois no *K-means* foi proveitosa, apesar de um pequeno crescimento da taxa de ruído, pois tem uma classe com os objetos de segmentação e outra com o fundo do envelope e ruído. Os altos valores obtidos para as taxas do bloco de endereço e carimbo no primeiro teste indicam o acerto no método escolhido. A Figura 6.9 exemplifica o processamento para o intervalo de janelas [3,5].

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,5]	99,38	48,07	96,02	2,84
[3,5]c2	99,50	50,09	96,54	2,93
[3,5]i001	99,38	48,07	96,02	2,84
[3,5]k25	99,38	48,07	96,02	2,84
[3,5]k25i001	99,38	48,07	96,02	2,84
[3,5]k40	99,38	48,07	96,02	2,84
[3,5]k40i001	98,96	44,41	94,45	2,62
[3,5]k55	99,11	45,43	94,91	2,68
[3,5]k55i001	99,38	48,07	96,02	2,84

Tabela 6.2 taxas de segmentação do intervalo de tamanho de janelas [3,5]
e variação nos parâmetros do *K-means*

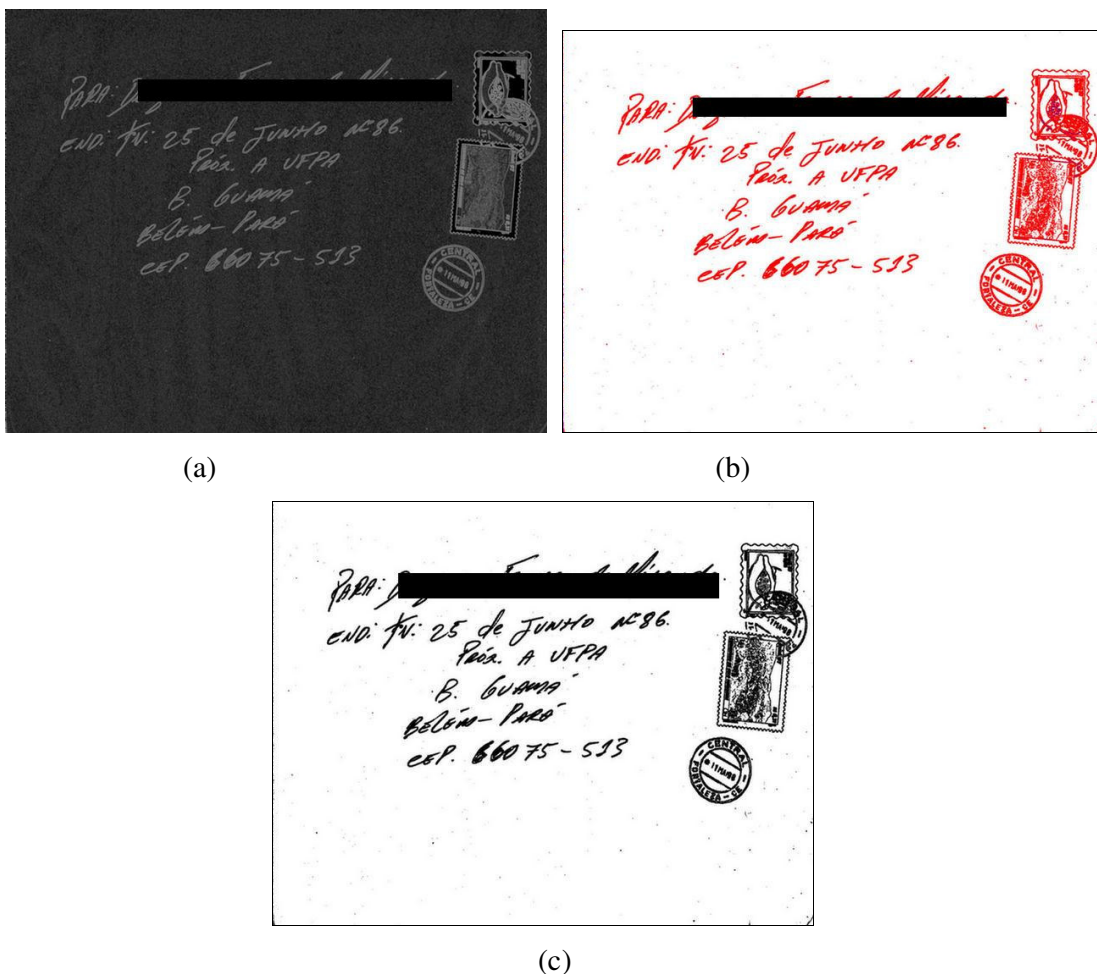


Figura 6.9 imagem real carta040 com parâmetros: intervalo do tamanho de janelas[3,5]; $K\text{-means}$ $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem $K\text{-means}$ (c) imagem segmentada

Análise dos Resultados do intervalo de tamanho de janelas[3,7] da imagem Carta040

Na tabela 6.3 verifica-se que, com o aumento do intervalo do tamanho de janelas para [3,7], há um aumento pequeno dos valores das taxas de segmentação do bloco de endereço, pouco, mas com importância, como exemplificado na Figura 6.10. A taxa do selo cresce expressivamente, próxima aos 20%, de 50,09% para 59,1%. A taxa do carimbo ocorre um significativo acréscimo de 96,54 para 99,63. A taxa do ruído aumenta significativamente, de 2,93% para 3,71%. As taxas do bloco de endereço, do selo, do carimbo e do ruído apresentam-se bem constantes com valores iguais a 99,93%, 58,16%, 99,59% e 3,65% respectivamente. Há uma pequena melhora quando o número de classes c diminui para dois, onde se obtém os melhores valores das taxas, menos para a do ruído.

Com o aumento do intervalo do tamanho da janela, a variação dos valores do K -means não reflete alteração nos valores das taxas de segmentação, a exceção do número de classes c .

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,7]	99,93	58,16	99,59	3,65
[3,7]c2	99,94	59,10	99,63	3,71
[3,7]i001	99,93	58,16	99,59	3,65
[3,7]k25	99,93	58,16	99,59	3,65
[3,7]k25i001	99,93	58,16	99,59	3,65
[3,7]k40	99,93	58,16	99,59	3,65
[3,7]k40i001	99,93	58,16	99,59	3,65
[3,7]k55	99,93	58,16	99,59	3,65
[3,7]k55i001	99,93	58,16	99,59	3,65

Tabela 6.3 taxas de segmentação do intervalo do tamanho de janelas [3,7]
e variação nos parâmetros do K -means

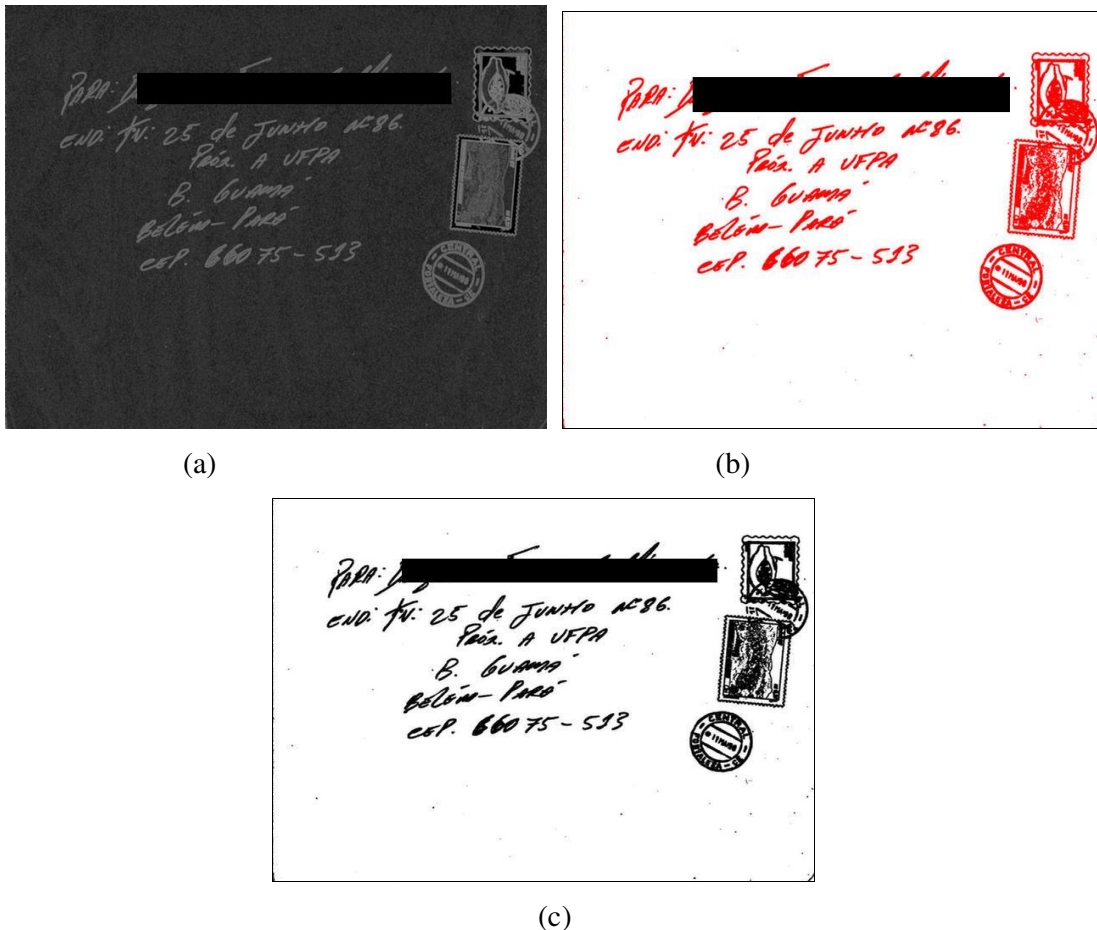


Figura 6.10 imagem real carta040 com parâmetros: intervalo do tamanho de janelas [3,7]; K -means $c = 3$,
 $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem K -means (c) imagem segmentada

Análise dos Resultados do intervalo de tamanho de janelas[3,9] da imagem Carta040

Com o aumento do intervalo do tamanho de janelas para [3,9], a tabela 6.3 apresenta taxas de segmentação excelentes para bloco de endereço (100%) e carimbo (99,96%). Um acréscimo significativo para a taxa do selo, de 59,1% para 67,34% e do ruído de 3,71% para 4,62%. As taxas do bloco de endereço, do selo, do carimbo e do ruído apresentam-se com dois grupos de valores constantes e próximos, sendo o que tem o valor da taxa de melhoria mínima do *K-means* $i = 0,001$ apresenta melhores resultados. Há uma melhora quando diminui o número de classes c para dois, onde se obtém os melhores valores das taxas, menos para a do ruído. Comportamento que repete os anteriores. A Figura 6.11 exemplifica o processamento para o intervalo do tamanho de janelas [3,9].

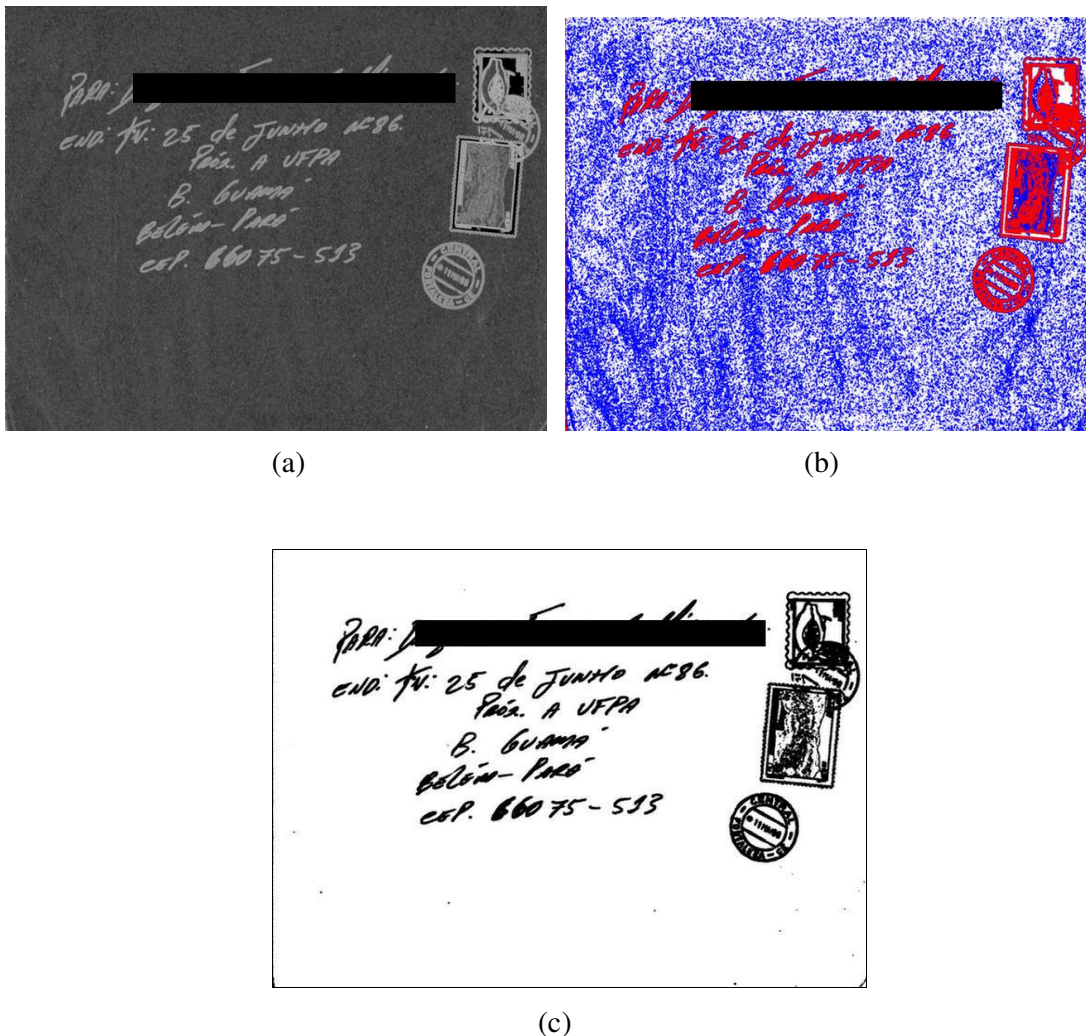


Figura 6.11 imagem real carta040 com parâmetros: intervalo do tamanho de janelas [3,9]; *K-means* $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem *K-means* (c) imagem segmentada

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,9]	99,99	60,65	99,90	4,09
[3,9]c2	100	67,34	99,96	4,62
[3,9]i001	100	66,46	99,94	4,55
[3,9]k25	99,99	60,65	99,90	4,09
[3,9]k25i001	100	66,46	99,94	4,55
[3,9]k40	99,99	60,65	99,90	4,09
[3,9]k40i001	100	66,46	99,94	4,55
[3,9]k55	99,99	60,65	99,90	4,09
[3,9]k55i001	100	66,46	99,94	4,55

Tabela 6.4 taxas de segmentação do intervalo do tamanho de janelas [3,9] e variação nos parâmetros do *K-means*

Análise dos Resultados do intervalo de tamanho de janelas[3,11] da imagem Carta040

Com o aumento do intervalo do tamanho de janelas para [3,11], a Tabela 6.5 nos apresenta as taxas de segmentação perfeitas para o bloco de endereço (100%) e carimbo (100%). Um acréscimo expressivo para a taxa do selo, de 67,34% para 73,88% e do ruído de 4,62% para 5,45%. As taxas do bloco de endereço, do selo, do carimbo e do ruído apresentam-se com dois grupos de valores constantes e próximos, sendo o que considera o valor da taxa de melhoria do *K-means* $i = 0,0001$ (padrão), apresentando melhores resultados. Repetindo o comportamento de testes anteriores. A Figura 6.12 ilustra o processamento para o intervalo do tamanho de janelas [3,11].

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,11]	100	72,95	100	5,35
[3,11]c2	100	73,88	100	5,45
[3,11]i001	100	71,52	100	5,20
[3,11]k25	100	72,95	100	5,35
[3,11]k25i001	100	71,52	100	5,20
[3,11]k40	100	72,95	100	5,35
[3,11]k40i001	100	71,52	100	5,20
[3,11]k55	100	72,95	100	5,35
[3,11]k55i001	100	71,52	100	5,20

Tabela 6.5 taxas de segmentação do intervalo do tamanho de janelas [3,11] e variação nos parâmetros do *K-means*

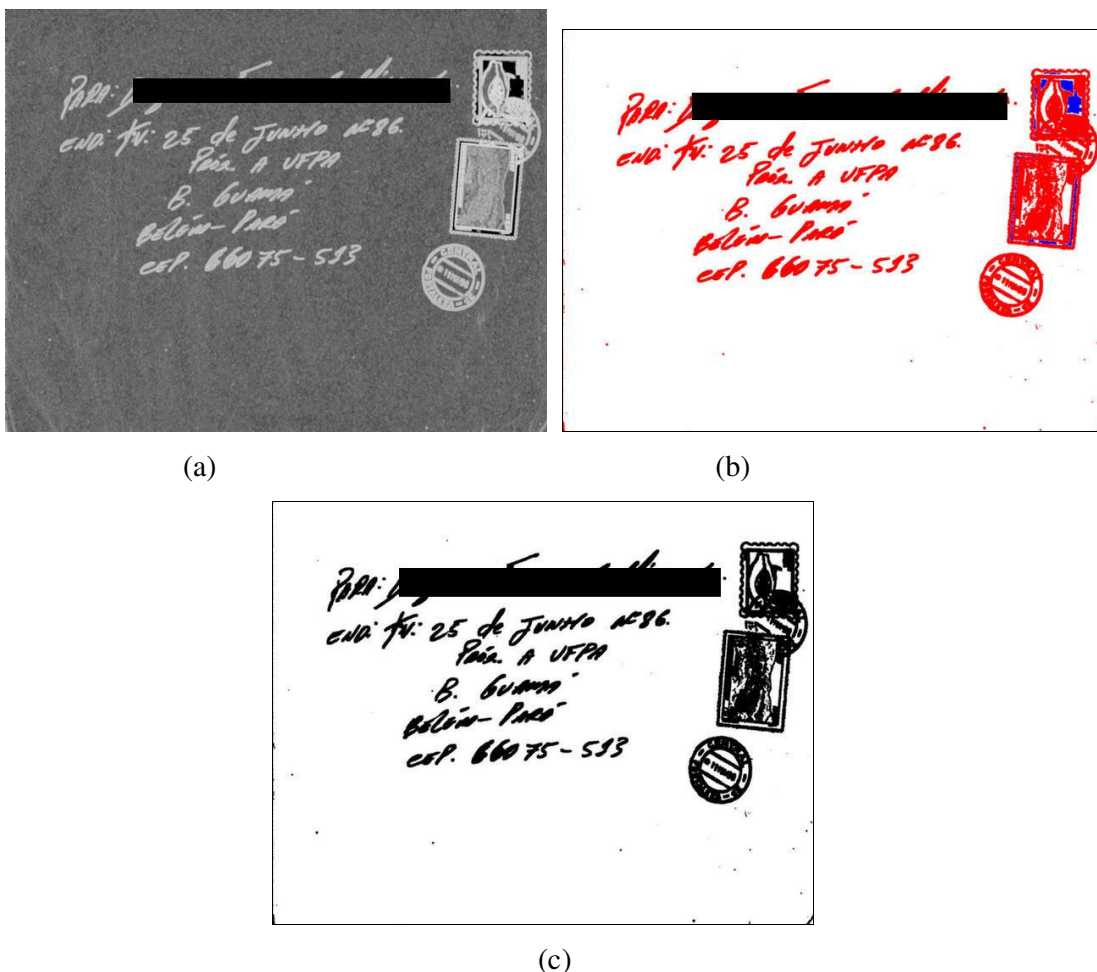


Figura 6.12 imagem real carta040 com parâmetros: intervalo do tamanho de janelas [3,11]; K -means $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem K -means (c) imagem segmentada

Análise dos Resultados do intervalo de tamanho de janelas[3,13] da imagem Carta040

Com o aumento do intervalo do tamanho de janelas para [3,13], a Tabela 6.6 apresenta as taxas de segmentação perfeitas para o bloco de endereço (100%) e carimbo (100%). Um acréscimo expressivo para a taxa do selo, de 73,88% para 78,90% e do ruído de 5,45% para 6,21%. As taxas do bloco de endereço, do selo, do carimbo e do ruído apresentam-se com dois grupos de valores constantes e próximos, apresentando melhores resultados o que tem o valor da taxa de melhoria do K -means $i = 0,0001$ (padrão). Continua a existir uma melhora quando diminui-se o número de classes c para dois, onde se obtém os valores mais expressivos das taxas, menos para a do ruído que também aumenta, repetindo o comportamento de testes anteriores. A Figura 6.13 exemplifica o processamento para o intervalo do tamanho de janelas [3,13].

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,13]	100	78,19	100	6,12
[3,13]c2	100	78,90	100	6,21
[3,13]i001	100	77,43	100	6,02
[3,13]k25	100	78,19	100	6,12
[3,13]K25i001	100	77,43	100	6,02
[3,13]k40	100	78,19	100	6,12
[3,13]K40i001	100	77,43	100	6,02
[3,13]k55	100	78,19	100	6,12
[3,13]K55i001	100	77,43	100	6,02

Tabela 6.6 taxas de segmentação do intervalo de tamanho de janelas [3,13]
e variação nos parâmetros do *K-means*

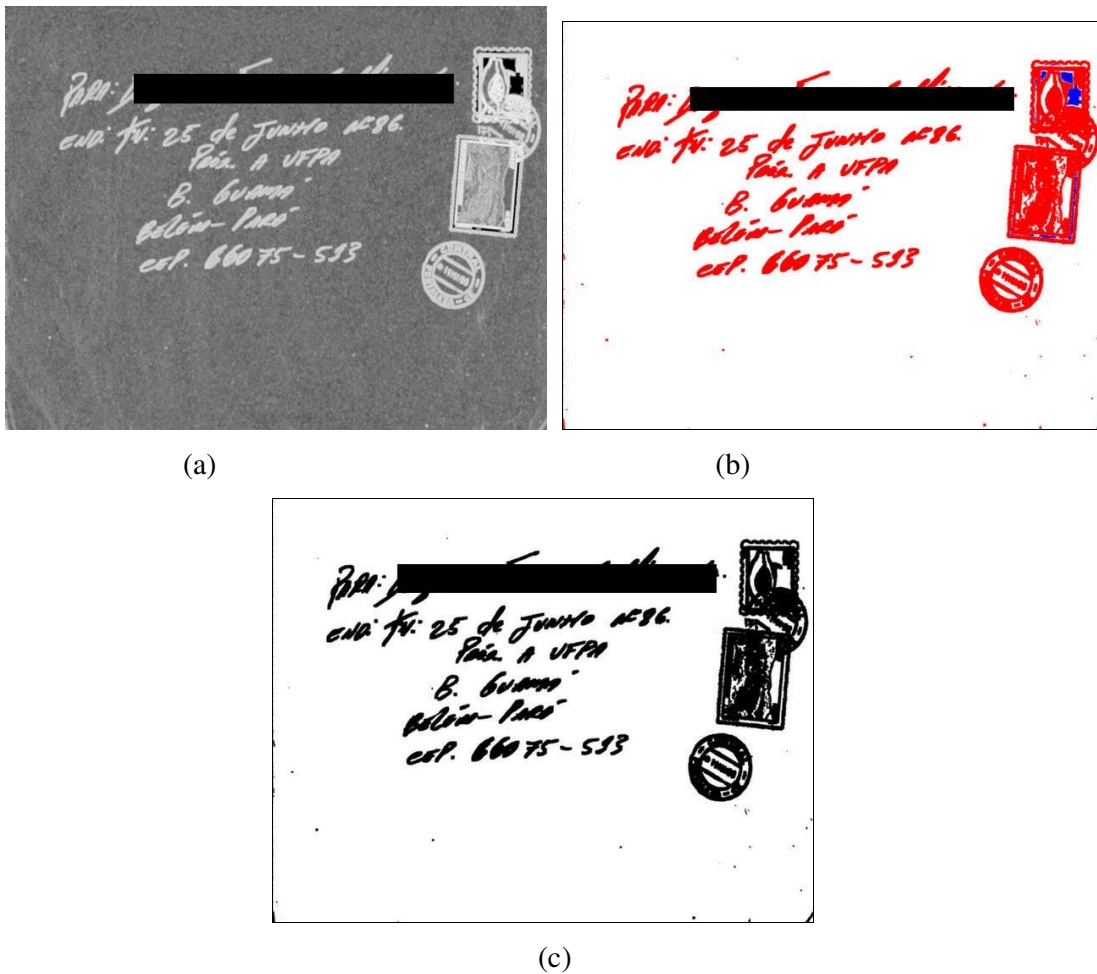


Figura 6.13 imagem real carta040 com parâmetros: intervalo do tamanho de janelas [3,13]; *K-means* $c = 3$,
 $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem *K-means* (c) imagem segmentada

Análise dos Resultados do intervalo de tamanho de janelas[3,15] da imagem Carta040

Com o aumento do intervalo do tamanho de janelas para [3,15], a Tabela 6.7 mostra as taxas de segmentação perfeitas para o bloco de endereço (100%) e carimbo (100%). Um acréscimo expressivo para a taxa do selo, de 78,90% para 82,78% e pouco para a do ruído de 6,21% para 6,91%. As taxas do bloco de endereço, do selo, do carimbo e do ruído apresentam-se praticamente como um só grupo com valores constantes e próximos. O que tem o valor da taxa de melhoria do $i = 0,001$, valor de iteração $k = 40$ e classe $c = 3$ (padrão) do *K-means*, apresenta o menor resultado para a taxa de segmentação do selo, porem a menor taxa de ruído deste intervalo de tamanho de janelas. Continua a existir uma melhora quando diminui o número de classes c para dois, repetindo o comportamento de testes anteriores. Destaca-se que o comportamento médio da base de imagens não é este. A Figura 6.14 ilustra o processamento para o intervalo do tamanho de janelas [3,15].

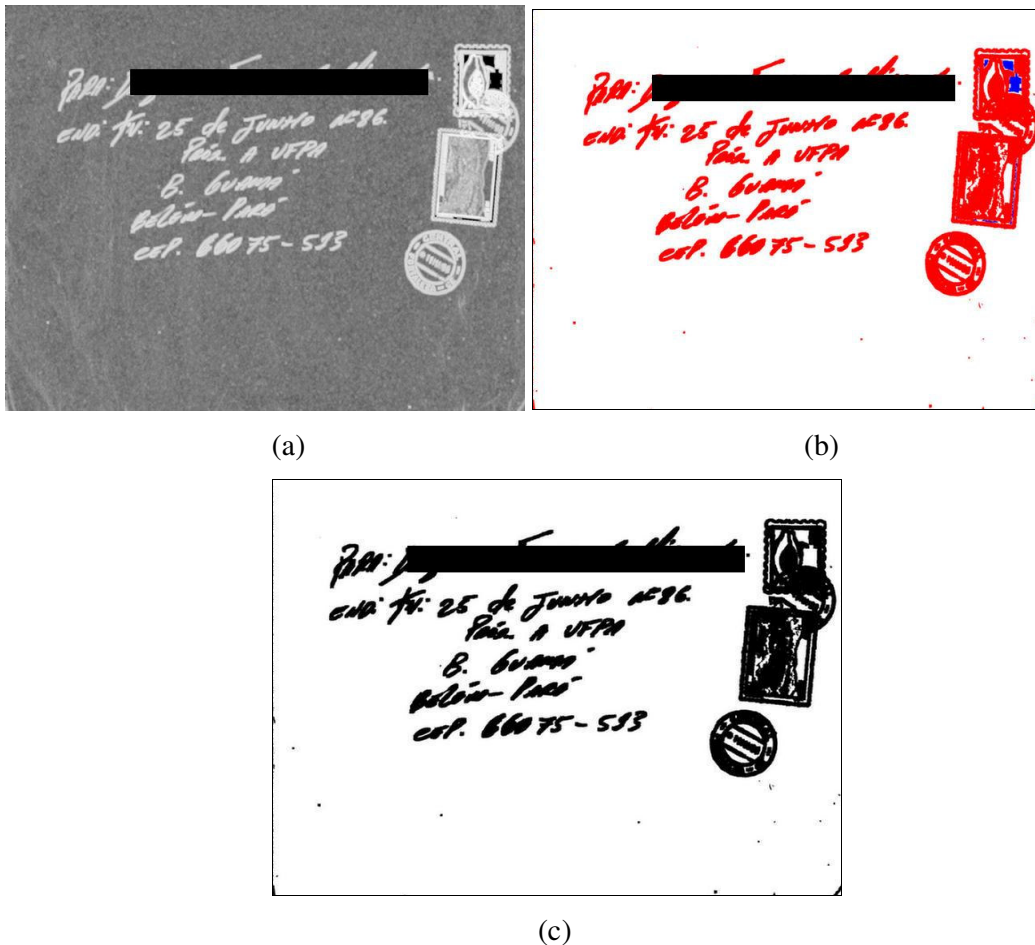


Figura 6.14 imagem real carta040 com parâmetros: intervalo do tamanho de janelas [3,15]; *K-means* $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem *K-means* (c) imagem segmentada

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,15]	100	81,87	100	6,76
[3,15]c2	100	82,78	100	6,91
[3,15]i001	100	81,87	100	6,76
[3,15]k25	100	81,87	100	6,76
[3,15]K25i001	100	81,87	100	6,76
[3,15]k40	100	81,87	100	6,76
[3,15]K40i001	100	77,43	100	6,02
[3,15]k55	100	81,87	100	6,76
[3,15]K55i001	100	81,87	100	6,76

Tabela 6.7 taxas de segmentação do intervalo de tamanho de janelas[3,15]
e variação nos parâmetros do *K-means*

No gráfico da Figura 6.15 nota-se que a partir dos parâmetros [3,7] as taxas de segmentação do bloco de endereço e do carimbo estão máximas e praticamente constantes. A taxa do ruído cresce significativamente e a taxa do selo comporta-se com alternâncias de estabilidade, mas sempre num ímpeto de crescimento. Os melhores valores para esta imagem são obtidos quando o número de classes c do *K-means* é reduzido para 2, uma classe dos objetos de segmentação e a outra do fundo do envelope e do ruído.

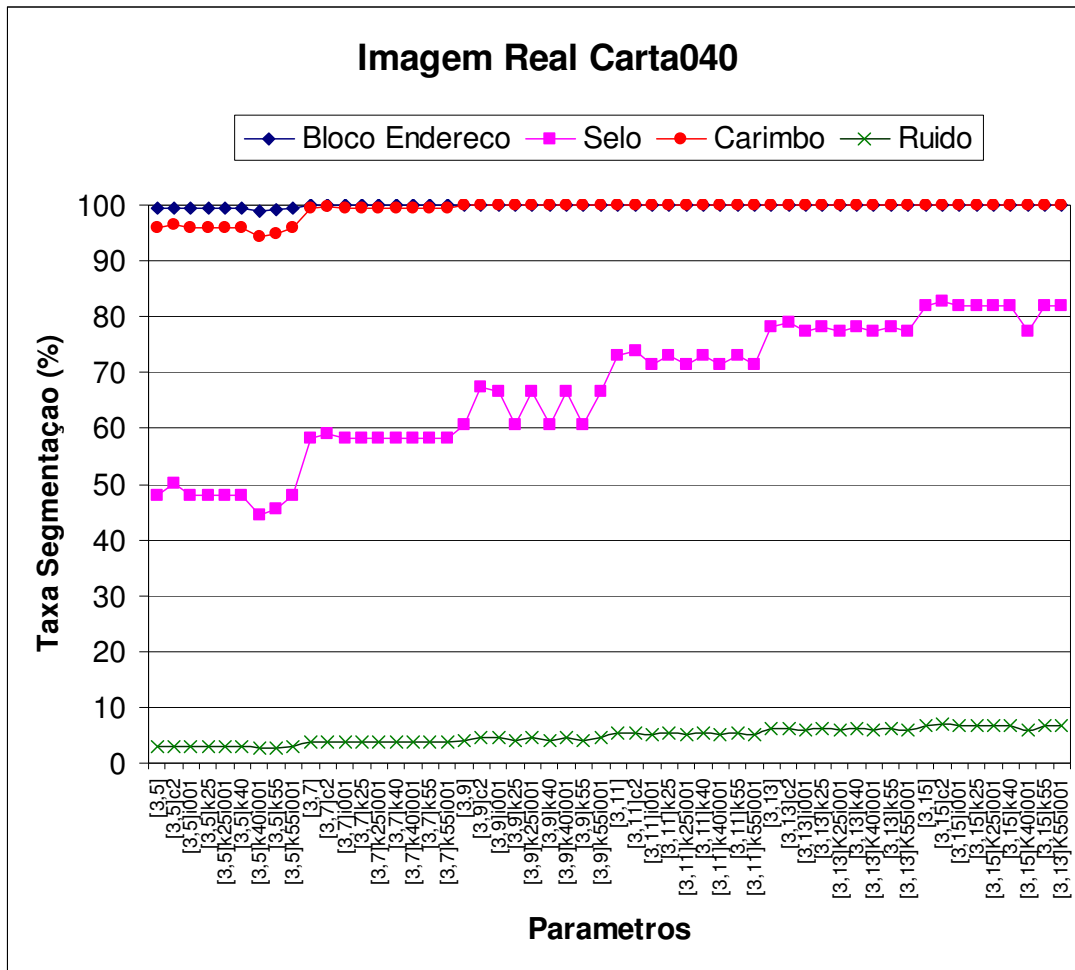


Figura 6.15 Gráfico dos Valores Médios das Taxas de Segmentação da Imagem Real carta040

6.4.2 Análise dos Resultados para a Base de Imagens Reais

Na sequência apresenta-se a Tabela 6.8 e a Figura 6.16 dos valores médios dos resultados dos testes de todas as 200 imagens da base de imagens reais para cada conjunto de parâmetros. É também com base nestes resultados que, avalia-se a eficiência do algoritmo de segmentação proposto.

Primeiramente ressalta-se que a taxa de ruído quando o parâmetro c é igual a 2, é praticamente o dobro das outras para a mesma faixa do tamanho da janela. No ponto de convergência ela é de 14,44% o que é extremamente grande, inviabilizando sua aplicação. Para a classe bloco de endereço o uso dos parâmetros iniciais do *K-means* são os mais adequados, gerando as melhores taxas. Para a classe selo o uso dos parâmetros iniciais do *K-means* com c igual a 2 é o que melhor resposta obtém-se. Para a classe carimbo a

melhor taxa é obtida com o uso dos parâmetros iniciais do *K-means*, a exceção dos intervalos dos tamanhos de janelas iniciais onde as melhores taxas são obtidas,

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,5]	92,61±13,61	48,86±13,84	82,02±19,66	3,97±6,10
[3,5]c2	94,75±16,47	61,91±14,07	89,46±17,97	9,91±11,21
[3,5]i001	86,40±19,89	48,18±16,81	77,19±24,27	6,23±9,56
[3,5]k25	92,13±15,10	48,73±14,00	81,82±20,27	3,97±6,10
[3,5]k25i001	86,40±19,89	48,18±16,81	77,19±24,27	6,23±9,56
[3,5]k40	92,13±15,10	48,73±14,00	81,82±20,27	3,97±6,10
[3,5]k40i001	94,10±12,45	49,95±12,46	84,19±17,47	2,76±1,87
[3,5]k55	95,18±11,09	50,81±11,65	85,88±16,14	2,71±0,99
[3,5]k55i001	86,40±19,89	48,18±16,81	77,19±24,27	6,23±9,56
[3,7]	96,05±10,89	59,76±15,68	88,97±17,58	6,20±7,50
[3,7]c2	95,88±17,04	67,78±15,36	91,96±18,01	10,36±11,07
[3,7]i001	91,89±17,91	59,07±18,20	85,42±22,19	7,95±8,97
[3,7]k25	96,05±10,89	59,76±15,68	88,97±17,59	6,20±7,50
[3,7]k25i001	91,89±17,91	59,07±18,20	85,42±22,19	7,95±8,97
[3,7]k40	96,05±10,89	59,76±15,68	88,97±17,59	6,20±7,50
[3,7]k40i001	91,89±17,91	59,07±18,20	85,42±22,19	7,95±8,97
[3,7]k55	96,05±10,89	59,76±15,68	88,97±17,59	6,20±7,50
[3,7]k55i001	91,89±17,91	59,07±18,20	85,42±22,19	7,95±8,97
[3,9]	97,24±13,64	66,34±16,90	91,89±17,22	7,37±8,29
[3,9]c2	94,61±21,34	71,95±17,89	92,12±21,79	11,92±11,38
[3,9]i001	96,65±14,49	64,53±17,24	91,83±16,50	5,20±4,55
[3,9]k25	97,25±13,64	66,45±16,80	91,89±17,22	7,38±8,28
[3,9]k25i001	96,65±14,19	64,53±17,24	91,83±16,50	5,20±4,55
[3,9]k40	97,25±13,64	66,45±16,80	91,89±17,22	7,38±8,28
[3,9]k40i001	96,65±96,65	64,53±17,24	91,83±16,50	5,20±4,55
[3,9]k55	97,25±13,64	66,45±16,80	91,89±17,22	7,38±8,28
[3,9]k55i001	96,65±14,19	64,53±17,24	91,83±16,50	5,20±4,55
[3,11]	99,02±5,97	73,39±13,92	95,64±10,72	7,22±6,04
[3,11]c2	95,31±20,17	76,57±18,08	93,22±21,48	13,46±11,53
[3,11]i001	95,90±15,72	70,30±19,35	92,32±17,65	7,09±6,67
[3,11]k25	99,02±5,97	73,39±13,92	95,64±10,72	7,22±6,04
[3,11]k25i001	95,90±15,72	70,30±19,35	92,32±17,65	7,09±6,67
[3,11]k40	99,02±5,97	73,39±13,92	95,64±10,72	7,22±6,04
[3,11]k40i001	95,90±15,72	70,30±19,35	92,32±17,65	7,09±6,67
[3,11]k55	99,02±5,97	73,39±13,92	95,64±10,72	7,22±6,04
[3,11]k55i001	95,90±15,72	70,30±19,35	92,32±17,65	7,09±6,67
[3,13]	99,25±5,09	77,19±13,79	96,84±8,66	7,85±5,49
[3,13]c2	94,90±21,29	78,82±20,20	93,24±22,37	14,44±11,89
[3,13]i001	95,75±16,35	73,11±20,71	93,11±17,94	7,84±6,81
[3,13]k25	99,25±5,09	77,20±13,79	96,84±8,66	7,85±5,49
[3,13]k25i001	95,75±16,35	73,11±20,71	93,11±17,94	7,84±6,81
[3,13]k40	99,25±5,09	77,20±13,79	96,84±8,66	7,85±5,49

[3,13]k40i001	95,75±16,35	73,11±20,71	93,11±17,94	7,84±6,81
[3,13]k55	99,25±5,09	77,20±13,79	96,84±8,66	7,85±5,49
[3,13]k55i001	95,75±16,35	73,11±20,71	93,11±17,94	7,84±6,81
[3,15]	97,39±14,62	79,05±18,83	95,65±15,58	9,36±7,27
[3,15]c2	95,12±21,22	81,40±20,92	93,58±22,28	14,70±11,64
[3,15]i001	95,04±20,19	77,12±21,66	93,47±19,58	9,36±7,97
[3,15]k25	97,39±14,62	79,08±18,84	95,65±15,58	9,37±7,27
[3,15]k25i001	95,04±20,19	77,12±21,66	93,47±19,58	9,36±7,97
[3,15]k40	97,39±14,62	79,08±18,84	95,65±15,58	9,37±7,27
[3,15]k40i001	95,04±20,19	77,12±21,66	93,47±19,58	9,36±7,97
[3,15]k55	97,39±14,62	79,08±18,84	95,65±15,58	9,37±7,27
[3,15]k55i001	95,04±20,19	77,12±21,66	93,47±19,58	9,36±7,97

Tabela 6.8 dos valores médios da taxa de segmentação dos testes

efetuados sobre a base de imagens reais

O decréscimo do valor de melhoria mínima do *K-means* i para 0,001 resultou em uma piora nas taxas de segmentação e um aumento nas taxas de ruído. O aumento do número de iterações k do *K-means* não apresenta alterações dos valores das taxas de segmentação.

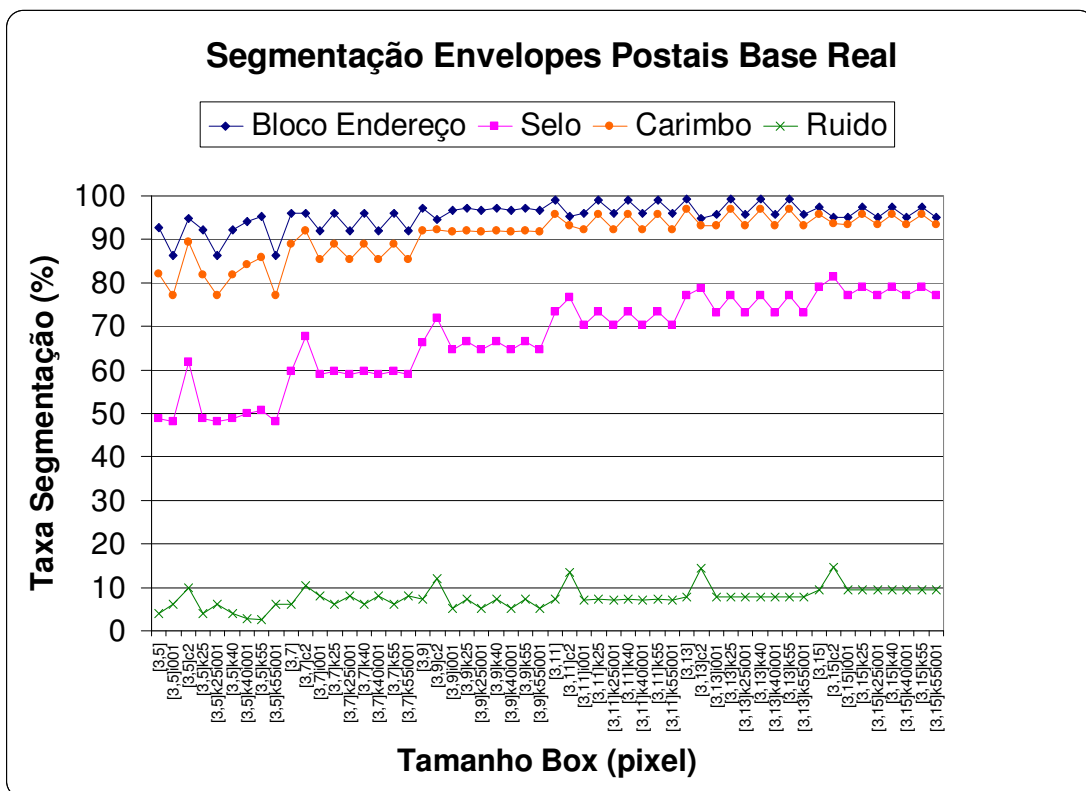


Figura 6.16 Gráfico dos valores médios das taxas de segmentação das imagens da base real

Concluí-se, que o fator preponderante para o aumento das taxas de segmentação é o aumento do valor do intervalo do tamanho de janelas, que para a base de imagens reais o ponto de convergência é a [3,13] com os parâmetros iniciais do *K-means*. Onde a taxa para o bloco de endereço é igual a 99,25%, resultado excelente para o objetivo deste trabalho. Para o selo a taxa igual a 77,19% é um bom resultado e bastante promissor. Para o carimbo a taxa igual a 96,84%, constitui um resultado excelente. Para a taxa do ruído igual a 7,85%, o que representa um valor um pouco alto, mas que não compromete os resultados. Verifica-se que com o aumento da taxa de segmentação aumenta-se também a taxa do ruído, o que é indesejável, pois o ruído pode-se misturar aos objetos de segmentação. Como certifica o método proposto, apresenta para imagens reais complexas, resultados promissores, pois não usou-se nenhum pré nem pós-processamento na base de imagens que possui imagens manuscritas, pré-impressas e figuras (texturas).

6.4.3 Análise dos Resultados para a imagem carta026fundo12 da Base Sintetizada

Agora, os testes efetuados sobre a base de imagens sintetizadas são analisados. Pelo comportamento dos resultados obtidos nos testes da base de imagens reais os parâmetros usados pelo *K-means* na base de imagem sintetizada são os iniciais (padrão). Visto que, as suas variações não apresentam melhoras significativas nos resultados dos testes da base de imagens reais, também, por serem os objetos de segmentação da base de imagens sintetizadas pertencentes todos a base de imagens reais, variando apenas os fundos complexos para cada imagem. Abaixo a Figura 6.17 mostra um exemplo da base de imagens sintetizadas e a imagem *ground truth* correspondente.

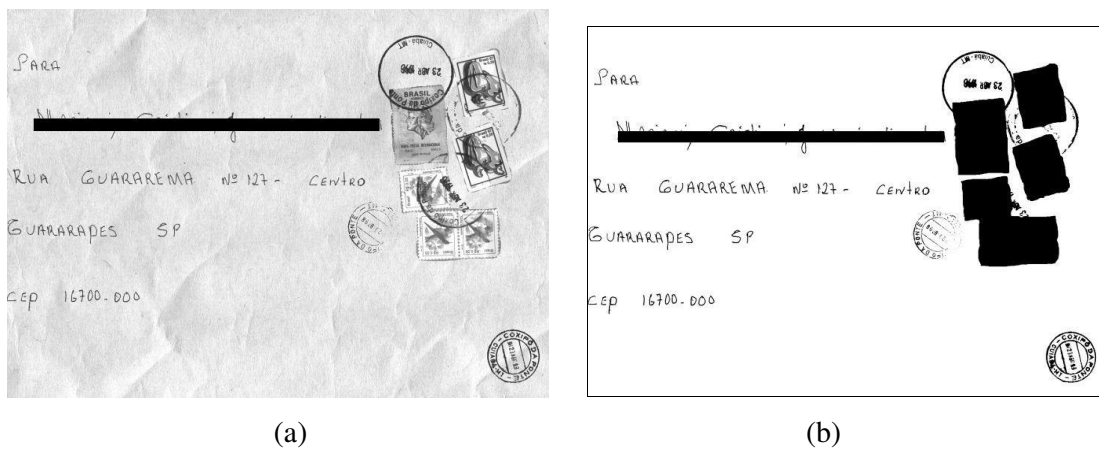
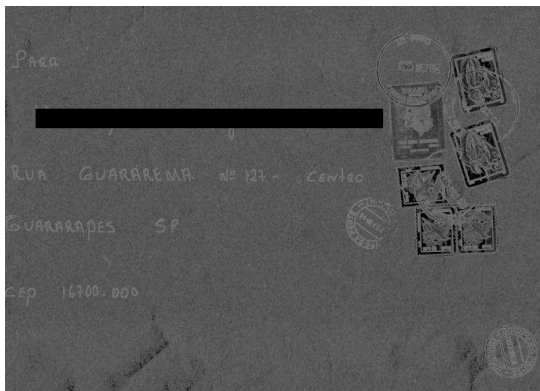


Figura 6.17 (a) Imagem da Base Sintetizada carta026fundo12 (b) Imagem *ground truth*

Apresenta-se a imagem carta026fundo12 como exemplo do processamento da base de Imagem Sintetizada. A escolha desta foi por apresentar um fundo complexo, vários selos e carimbos, uma diagramação bem dispersa; e também por ter resultados acima da media da base o que mostra que se pode alcançar excelentes resultados de segmentação com este método, para este tipo de base.

Como se vê na imagem 6.18 a classe do bloco de endereço não foi segmentada. O *K-means* não consegue identificar pelas características da DF passadas a ele esta classe. Os resultados abaixo comprovam isto,

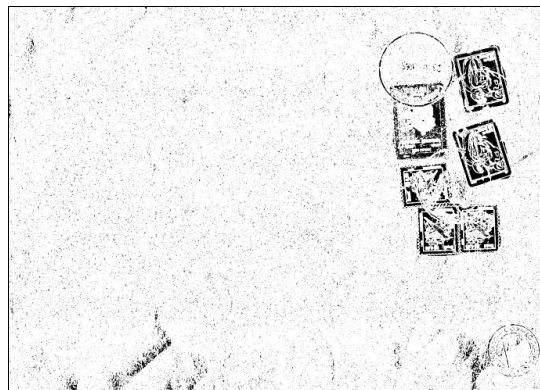
Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,5]	0.43	30.88	12.55	1.55



(a)



(b)



(c)

Figura 6.18 imagem sintetizada carta026fundo12 com parâmetros: intervalo do tamanho de janelas [3,5]; *K-means* $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem *K-means* (c) imagem segmentada

Com o aumento do intervalo do tamanho de janelas para [3,7] (Figura 6.19), a classe do bloco de endereço e selo não foi segmentada. O *K-means* não identifica pelas características da DF passadas a ele estas classes. Os resultados abaixo mostram isto:

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,7]	1.30	0	4.87	0.07

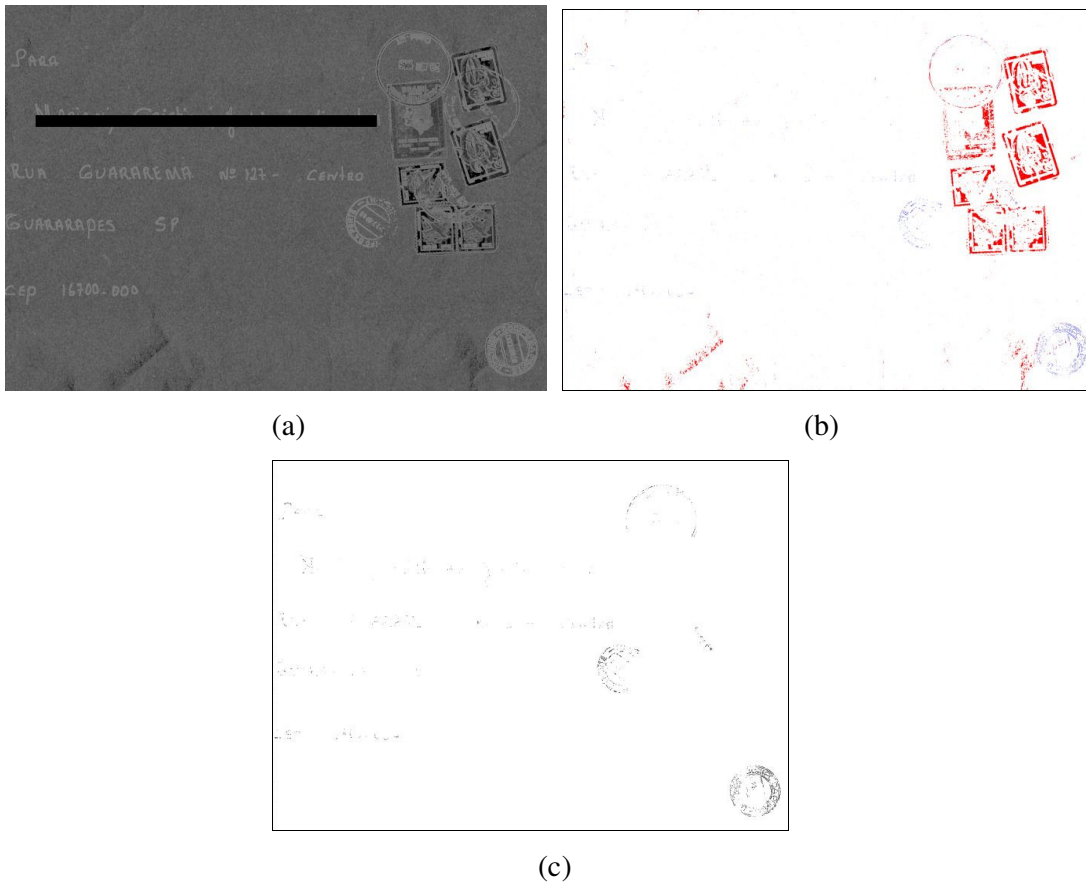
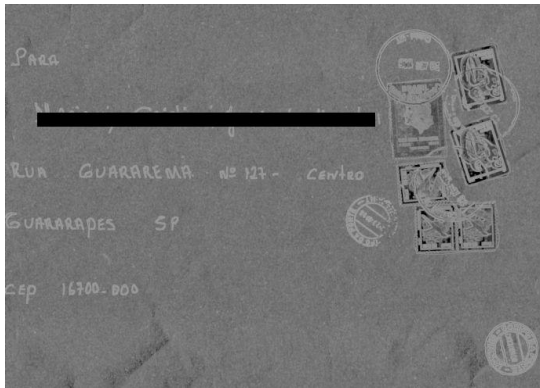


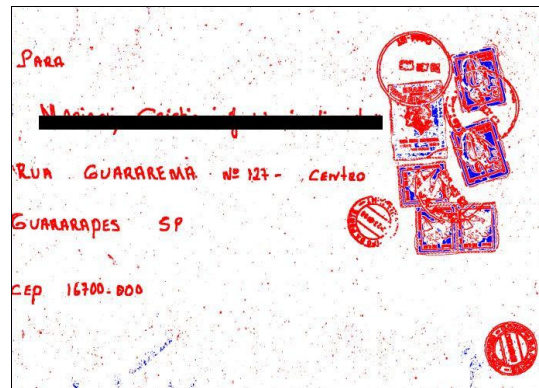
Figura 6.19 imagem sintetizada carta026fundo12 com parâmetros: intervalo do tamanho de janelas [3,7]; *K-means* $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem *K-means* (c) imagem segmentada

Com o aumento do intervalo do tamanho de janelas para [3,9] (Figura 6.20), todos os objetos foram segmentados com expressivas taxas especialmente o bloco de endereço e carimbo. Abaixo são apresentados os resultados,

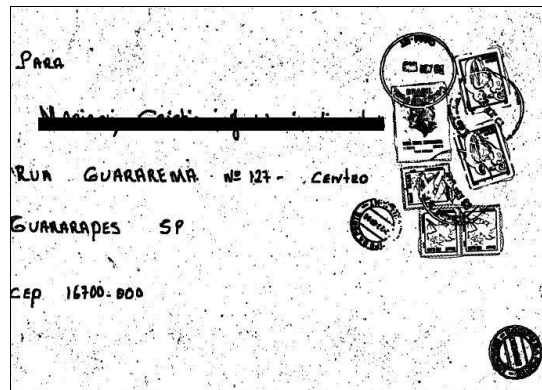
Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,9]	99,20	39,20	79,07	4,02



(a)



(b)



(c)

Figura 6.20 imagem sintetizada carta026fundo12 com parâmetros: intervalo do tamanho de janelas [3,9]; $K\text{-means}$ $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem $K\text{-means}$ (c) imagem segmentada.

Com o aumento do intervalo do tamanho de janelas para [3,11] (Figura 6.21), as taxas das classes do bloco de endereço e carimbo decresceram fortemente, o selo praticamente não foi segmentado. As características da DF enviadas ao $K\text{-means}$, não permitem uma correta identificação das classes.

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,11]	32,47	0,49	22,86	0,53

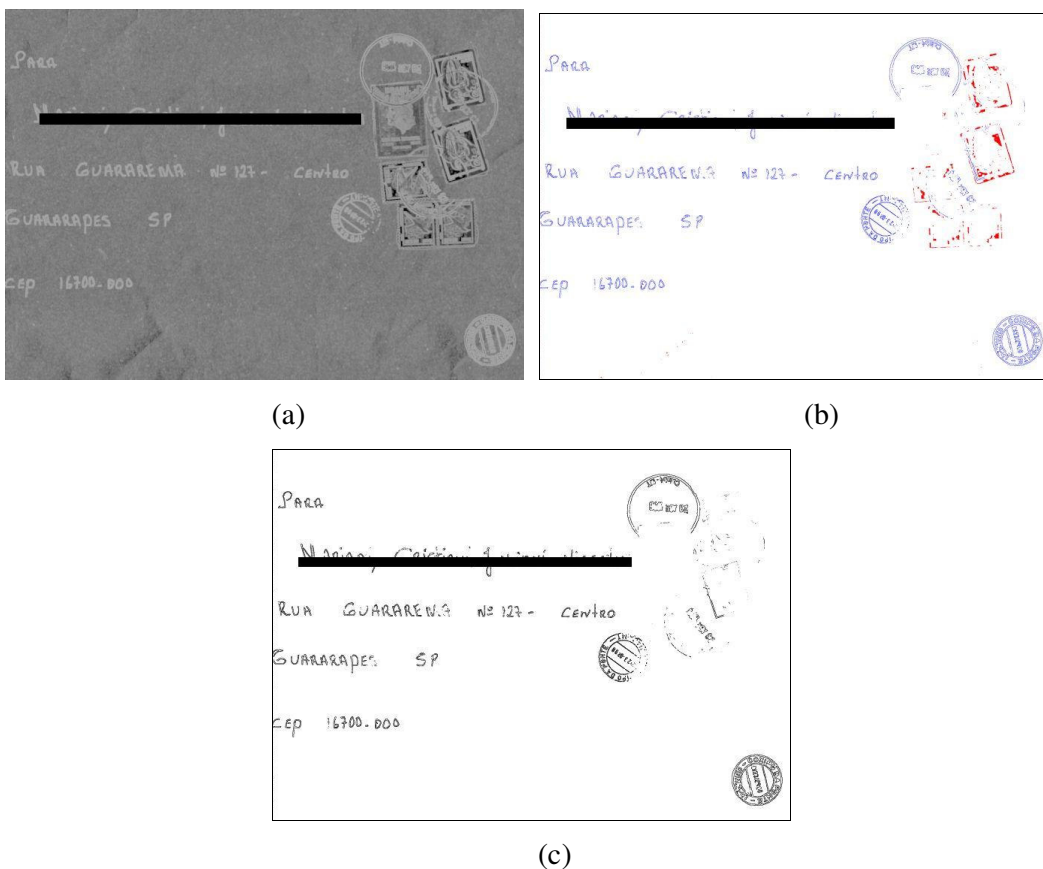
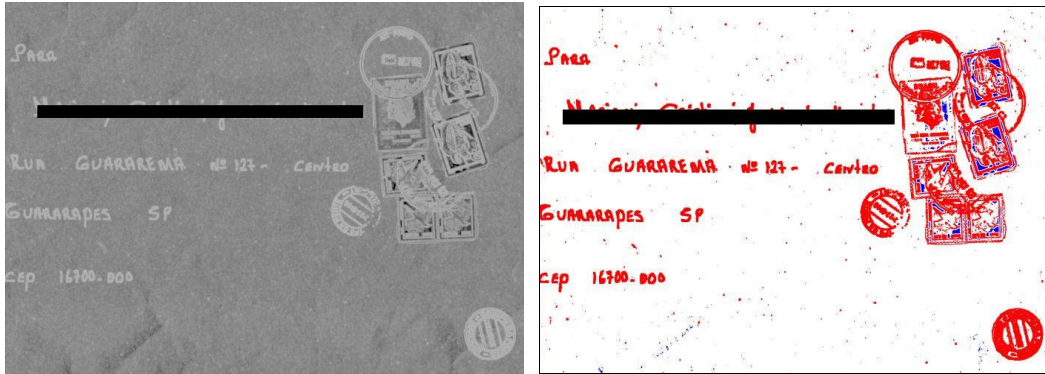


Figura 6.21 imagem sintetizada carta026fundo12 com parâmetros: intervalo do tamanho de janelas [3,11]; $K\text{-means}$ $c = 3$, $k = 10$ e $i = 0,0001$ (a) imagem DF, (b) imagem $K\text{-means}$ (c) imagem segmentada

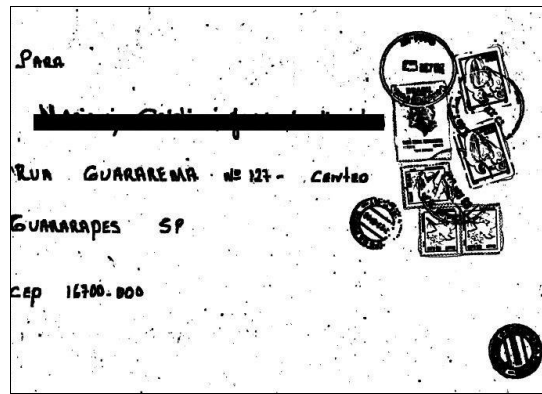
Com o aumento do tamanho do intervalo de janelas para [3,13], todos os objetos foram segmentados com taxas excelentes. Especialmente o bloco de endereço e carimbo que apresentam uma melhora acentuada (Figura 6.22). Abaixo os resultados demonstram isto,

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,13]	99,97	47,19	92,07	4,98



(a)

(b)



(c)

Figura 6.22 imagem sintetizada carta026fundo12 com parâmetros: intervalo do tamanho de janelas [3,13]; $K\text{-means}$ $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem $K\text{-means}$ (c) imagem segmentada.

Com o aumento do intervalo do tamanho de janelas para [3,15] (Figura 6.23), os objetos continuaram a melhorar suas taxas, com mais vigor para o selo. O bloco de endereço é quase 100% e o carimbo encontra-se num patamar excelente com 95,64%. O ruído neste processo continua a crescer também. Com estas taxas atingimos o ponto de convergência, especialmente pela taxa de ruído que atingiu um patamar que para frente poderia comprometer a qualidade da segmentação, ainda mais se o ruído estiver próximo ao bloco de endereço. Ainda há espaço para expressivo crescimento da taxa do selo, mas certamente o aumento da taxa de ruído será um impeditivo.

Parâmetros	Bloco Endereço	Selo	Carimbo	Ruído
[3,15]	99,98	53,08	95,64	5,91

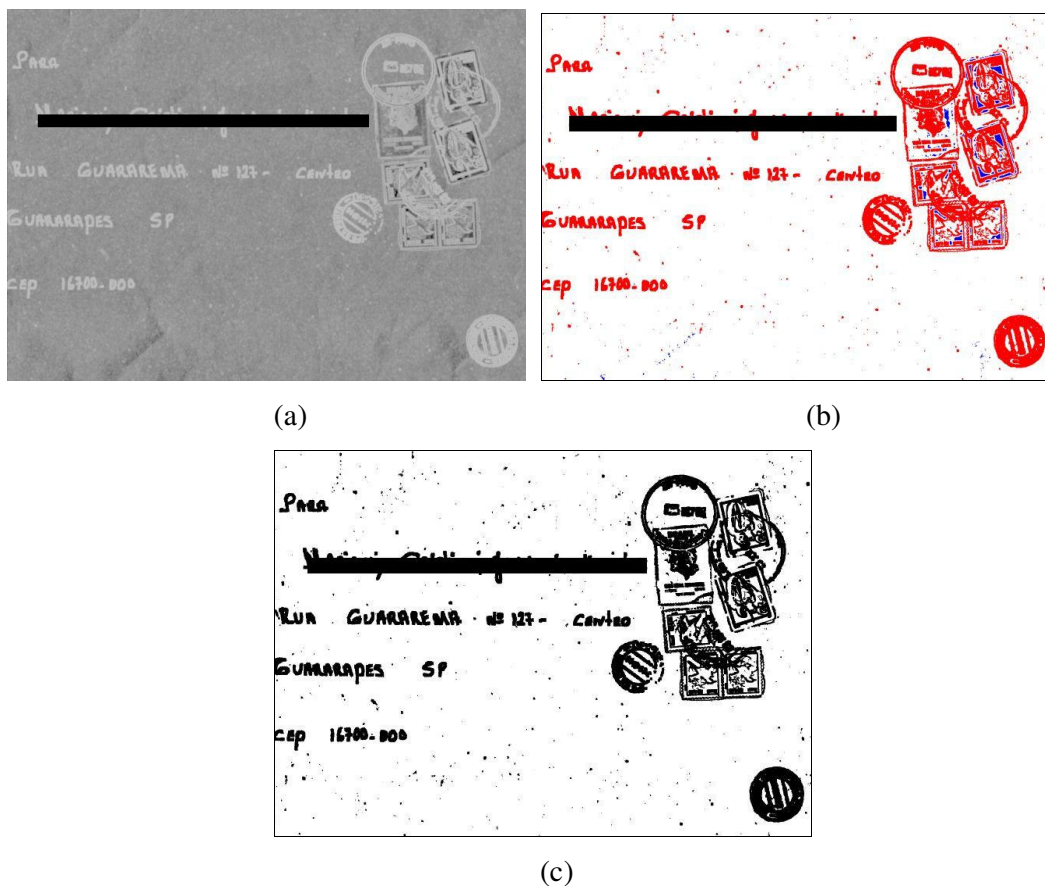


Figura 6.23 imagem sintetizada carta026fundo12 com parâmetros: intervalo do tamanho de janelas [3,15], $K\text{-means}$ $c = 3$, $k = 10$ e $i = 0,0001$. (a) imagem DF, (b) imagem $K\text{-means}$ (c) imagem segmentada.

O exemplo do processamento da imagem sintetizada carta026fundo12 indica que, esta imagem esta acima da media dos resultados da base de imagens sintetizadas. Tabela 6.3. mostra que, o método proposto tem excelentes resultados para segmentação do bloco de endereço e carimbo, imagens manuscrita e pré-impressa, respectivamente. A taxa do selo mostrou-se baixa, mas acima dos valores médios da base sintetizada. A taxa do ruído ficou próxima do ponto de convergência da base de imagens sintetizadas e não compromete a segmentação. Outro indicativo deste teste é que o $K\text{-means}$, para este tipo de imagens apresenta um comportamento que às vezes falha para definir os *aglomerados*, prejudicando ou não permitindo uma segmentação adequada, pois dimensão fractal como única característica não é suficiente para o $k\text{-means}$ operar.

6.4.4 Análise dos Resultados da Base de Imagens Sintetizadas

A Tabela 6.3 apresenta os valores médios das taxas de segmentação, resultado dos testes efetuados sobre a base de imagens sintetizadas. Primeiramente, para o intervalo do tamanho de janelas - [3,5] - o método apresenta resultados de baixo significado na segmentação e com alto índice de ruído.

Com o aumento do tamanho do intervalo de janelas para [3,7], há um crescimento expressivo das taxas do bloco de endereço e carimbo, mas um decréscimo na do selo. Indicativo da dificuldade do método para segmentar este tipo de imagem (figura, textura) nesta base.

A partir deste ponto, com o crescimento do tamanho da faixa da janela da DF para [3,9], há um crescimento expressivo nas taxas de segmentação e um pequeno acréscimo na taxa do ruído. Após, há uma melhora com pequenos acréscimos nas taxas do bloco de endereço, um crescimento significativo da taxa do selo e do carimbo.

O ponto de convergência é atingido no parâmetro [3,15] onde a taxa de ruído de 6,14% foi fator preponderante. A taxa do bloco de endereço atinge 90,87%, a do selo 48,49% e a do carimbo 86,48%. Poderia-se ter convergido no parâmetro [3,13], pois a taxa do ruído é menor e a do bloco de endereço é praticamente a mesma, economizando custo computacional. Mas, por uma melhora significativa nas taxas do selo e do carimbo é pertinente a escolha de [3,15], como está indicado na Figura 6.24 que mostra a evolução gráfica dos resultados.

As taxas de segmentação obtidas para esta base são significativas e promissoras, pois são atingidos patamares elevados, mas com espaço para um crescimento significativo. O fator preponderante para não se atingir taxas maiores é a dificuldade do *K-means* em separar os *aglomerado* para algumas imagens, o que faz com que a media caia sensivelmente. Mostra-se que para segmentação dos tipos de imagens manuscritas e pré-impressas o método é bom. A exceção é a do selo que tem uma baixa taxa de segmentação, indicando dificuldade para imagens do tipo figura e textura.

Parâmetros	Bloco de Endereço	Selo	Carimbo	Ruído
[3,5]	15,38±30,71	44,37±12,54	18,93±25,35	19,86±13,86
[3,7]	75,58±35,24	29,04±18,41	64,06±34,73	2,88±1,85
[3,9]	87,52±28,77	35,73±17,34	77,10±30,31	3,76±2,67
[3,11]	87,86±30,32	39,08±19,08	80,00±30,85	4,26±2,40
[3,13]	90,66±27,60	44,62±19,19	84,75±28,58	5,27±3,25
[3,15]	90,87±27,77	48,49±20,42	86,48±58,56	6,14±2,85
[3,17]	93,01±24,65	54,83±20,36	89,78±25,67	9,08±5,72

Tabela 6.9 dos valores médios dos testes da base de imagens sintetizadas

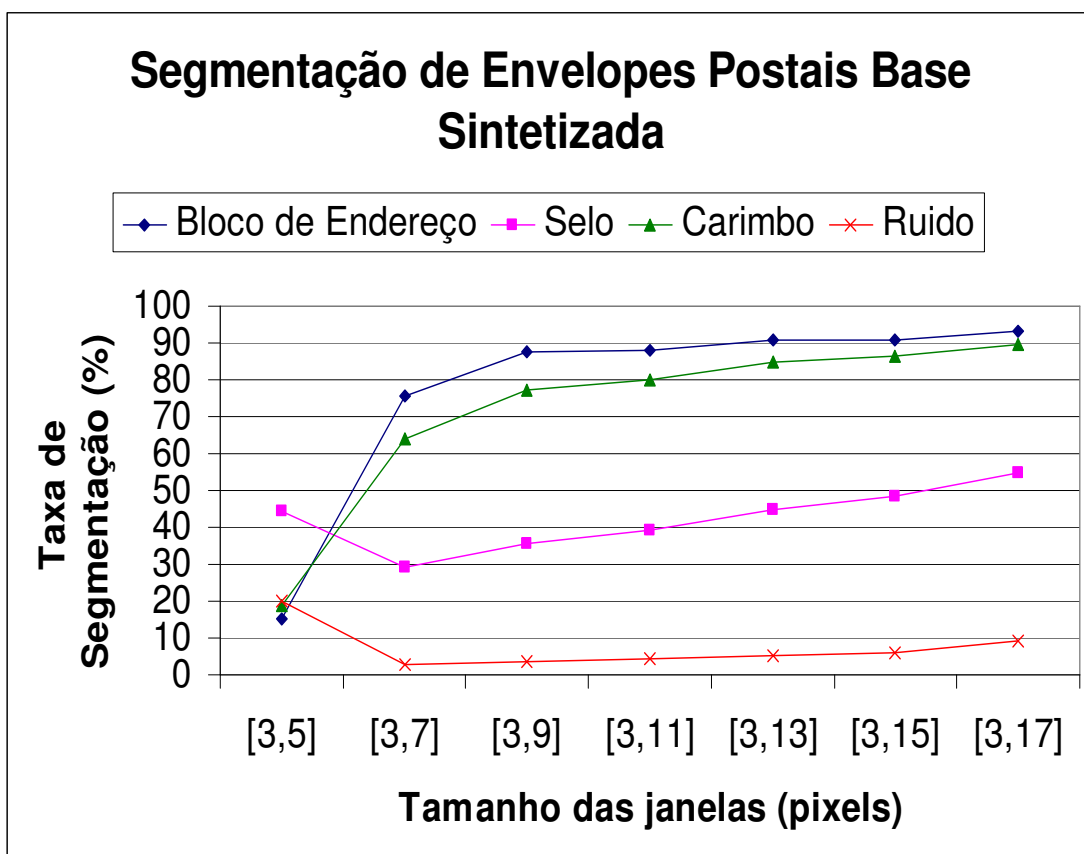


Figura 6.24 Gráfico dos valores médios das taxas de segmentação das imagens da base sintetizada

6.4.5 Comentários Finais

Os resultados apresentados pelos testes mostram que a abordagem fractal para segmentação de imagens demonstra-se robusta e promissora, pois, atingiram excelentes taxas para a base de imagens reais e relevantes taxas para a base de imagens sintéticas. A diferença das taxas entre as bases é relativa às limitações do *K-means* em separar os aglomerados na base de imagens sintetizadas. A taxa do ruído atingido no ponto de convergência é aceitável e não prejudica de uma maneira geral a segmentação, apesar de não se ter a localização dos ruídos.

O método mostrou-se simples, por ter poucas etapas de processamento e poucos parâmetros. Sua automatização é alta, pois a única definição *a priori* é o número de classes para o *K-means*. Sendo não-supervisionado, atinge seu objetivo de segmentar o bloco de endereço sem nenhuma regra, pré ou pós-processamento. Faz o mesmo para o carimbo e o selo. Por apresentar fundos e/ou texturas complexos semelhantes a fundos de envelopes a segmentação da classe selo obteve taxas menores. A metodologia usada não soube resolver ainda este problema.

O tempo de processamento médio para cada imagem com o ponto de convergência na faixa de tamanho de janela [3,13] é de 1 minuto e 40 segundos em uma maquina Pentium III dual com 3gb de memória RAM.

Há relevantes trabalhos desenvolvidos no PPGIA em segmentação de imagens. Akira [AKIRA 2002] desenvolveu um método baseado na Abordagem Morfológica, utilizando matriz de Co-Ocorrência e Histograma 2D. A avaliação foi visual e a base de imagens é diferente obtendo resultados promissores. Menoti [MENOTI 2003] apresentou uma abordagem baseada em uma seleção de características no espaço *Wavelet*, utilizando avaliação *ground truth* dos resultados e as mesmas bases de imagens obtendo 97,13% de taxa de segmentação para o bloco de endereço com 0,63% de ruído. Ayala [AYALA 2004] apresentou uma abordagem baseada em aprendizagem, utilizando a base de imagens reais de envelopes postais brasileiros, obtendo 98,59% de taxa de segmentação para o bloco de endereço com 2,06% de ruído. Na Tabela 6.4 é apresentado um comparativo entre as abordagens.

(% $\mu \pm \sigma$)	Abordagem <i>wavelet</i>	Abordagem por aprendizagem	Abordagem por dimensão fractal [3,13]
Bloco-Endereço	97,13 $\pm$ 05,86	98,62 $\pm$ 06,91	99,25 $\pm$ 05,09
Selo	23,99 $\pm$ 18,65	87,04 $\pm$ 19,67	77,19 $\pm$ 13,79
Carimbo	86,01 $\pm$ 23,31	95,36 $\pm$ 13,64	96,84 $\pm$ 08,66
Ruído	00,63 $\pm$ 00,93	02,06 $\pm$ 01,32	07,85 $\pm$ 05,49

Tabela 6.10 Resultados obtidos por diversas abordagens

Capítulo 7

Conclusões e Trabalhos Futuros

Este trabalho apresenta uma abordagem para segmentação dos objetos semânticos relevantes de envelopes postais (bloco de endereço, carimbo(s) e selo(s)) que necessariamente não seguem os formatos padronizados dos CB. O preenchimento, as vezes incorreto da população gera milhões de correspondências sem padronização. Confirma-se a necessidade de desenvolver um sistema robusto de localização do bloco de endereço independente da diagramação e tipo de envelope postal.

O método proposto é baseado num mecanismo de classificação por aglomeração da Dimensão Fractal dos objetos semânticos de um envelope. Objetiva-se localizar e segmentar o bloco de endereço manuscrito, do carimbo(s) e selo(s) sem conhecimento *a priori* sobre os envelopes postais.

As conclusões após a realização dos testes são que o objetivo principal do método proposto, a localização e segmentação do bloco de endereço por um mecanismo não supervisionado, foi atingido e apresentou resultados promissores, pois :

- Alcançou-se a taxa de segmentação média de 99,25% dos *pixels* do bloco de endereço para a base de imagens reais, 96,84% do carimbo e 77,19% para o selo. Valores eficazes. A taxa do ruído de 7,85% não é um bom valor, mas aceitável, pois o método foi robusto e segmentou envelopes de vários tamanhos, tipos de fundos de envelopes, tipos de diagramação e tipos de imagens.
- Para a base de imagens sintetizadas os valores foram promissores, 90,87% dos *pixels* do bloco de endereço, 86,48% do carimbo, um pouco abaixo dos resultados obtidos para a base de imagens reais. Visto que o *K-means* não teve o desempenho esperado para tratar de toda a complexidade da base sintetizada, os resultados médios obtidos foram menores. A taxa de 48,49% para o selo indica mais uma vez que o método para esta base complexa não teve a robustez desejada, mas vale

lembrar que o objetivo deste trabalho era principalmente localizar e segmentar o bloco de endereço. A taxa do ruído de 6,14%, não é boa, mas não compromete os resultados, pois o método localiza e segmenta o bloco de endereço.

- A base de imagens reais, com a sua quantidade e variedade de imagens, mostrou-se adequada para validar nos testes o método proposto. A base de imagens sintetizadas veio fortalecer os resultados obtidos pela base de imagens reais, apesar dos resultados inferiores, indicando que a construção desta base influencia nos testes.

Nesta área de pesquisa os trabalhos futuros sugeridos são:

- A melhoria do processo de inicialização do *K-means*, que pode resultar em melhor separação dos *aglomerados*, melhorando as taxas de segmentação e reduzindo a taxa de ruído.
- Análise e avaliação de outros métodos de *clustering* [FAYAD, REINA, BRADLEY 1998].
- Inclusão de novas características para que, em conjunto com a DF, possa-se aumentar a qualidade do processo de aglomeração. Por exemplo, a Lacunaridade, “*que é uma medida que se relaciona fortemente com a distribuição e tamanho de espaços vazios em um fractal. A Lacunaridade mede a divergência de um conjunto para uma invariante translacional, ou seja, o quanto uma estrutura se afasta de uma distribuição mais uniforme*” [FERNANDES, CONCI 2004].

Referências Bibliográficas

- [AKIRA 2002] AKIRA, E. Dissertação de Mestrado "Segmentação de Envelopes Postais Brasileiros por Abordagem Morfológica utilizando matriz de Co-Ocorrência e Histograma 2D", Pontifícia Universidade Católica do Paraná, Programa de Pós-Graduação em Informática Aplicada, 2002.
- [AYALA 2004] AYALA, H. Tese de Doutorado "Uma abordagem de Segmentação por Aprendizagem", Pontifícia Universidade Católica do Paraná, Programa de Pós-Graduação em Informática Aplicada, 2004
- [BERGÉ, POMEAU, GANCE 1995] BERGÉ, P; POMEAU, Y; GANCE, M. "Dos Ritmos ao Caos", Editora UNESP, São Paulo, Brasil 1995.
- [CHEN *et al* 2001] CHEN, W.; YUAN, S.; HSIAO, H.; HSIEH, C. "Algorithms to Estimating Fractal Dimension of Textured Images", in *Proceedings International Conference on Acoustics Speech and Signal Processing*, IEEE, 2001, volume 3, páginas 1541-1544.
- [CRILLY, EARNSHAW, JONRS 1991] CRILLY, A.J.; EARNSHAW, R.A.; JONRS, H. "Fractals and Chaos", Springer-Verlag, 1991.
- [ECT 2001] ECT – Empresa Brasileira de Correios e Telégrafos, Correio do Brasil, Banco de Imagens, Departamento de Engenharia, Divisão de Desenvolvimento Tecnológico, Convenio PUCPR-ECT, 2001.
- [FACON 2002] FACON, J. "Estudo sobre fractal", Pontifícia Universidade Católica do Paraná, Programa de Pós-Graduação em Informática Aplicada, Relatório Interno, Setembro 2002.

- [FACON 2004] FACON, J. Processamento e Análise de Imagens, Apostila de aula, 135pp, 2004 //www.ppgia.pucpr.br/~facon/IndexPrincipalBrMaterialGraduacao2004.html
- [FAYAD, REINA, BRADLEY 1998] FAYAD, U.; REINA, C.; BRADLEY, P. "Initialization of iterative refinement agglomerating algorithms", in *Proceedings 4th International Conference on Knowledge Discovery and Data Mining - KDD98*, 1998.
- [FERNANDES, CONCI 2004] FERNANDES, J.L.; CONCI, A. "A Lacunaridade na Caracterização Espacial de Padrões", *Anais do 3º Congresso Temático de Dinâmica e Controle da SBMAC*, 2004, pp 1432-1442
- [GLEICK 1990] GLEICK, J. "Caos, A criação de uma nova ciência", Editora Campus, Rio de Janeiro, Brasil, 1990.
- [GONZALES, WOODS 2000] GONZALES, R.C; WOODS, R.E "Processamento de Imagens Digitais", Editora Edgard Blucher, São Paulo, Brasil, 2000
- [HARALICK 1992] HARALICK, R.M; SHAPIRO L. "*Computer and Robot Vision*", Addison-Wesley Publishing, EUA, 1992
- [KAPLAN 1999] KAPLAN, L. "Extended fractal analysis for texture classification in segmentation" in *Proceedings IEEE Transactions on Image Processing*, Novembro 1999, volume 8, número 11, pp1572-1585.
- [KELLER, CHENG, CROWNOVER 1989] KELLER, J.; CHENG, S.; CROWNOVER, R.M. "Texture description and segmentation through fractal geometry". In *Proceedings Computer Vision Graphics and Image Processing*, 1989, volume 45, pp 150-166.

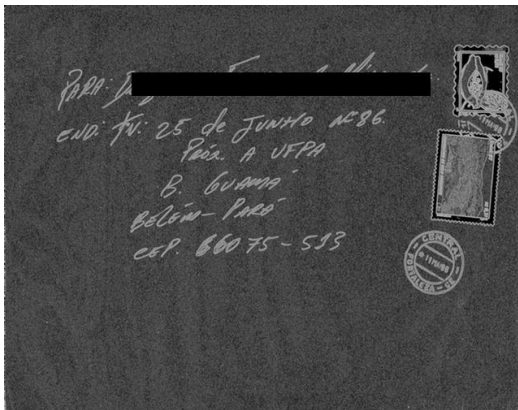
- [KRAFT, KAUER 1995] KRAFT, R.; KAUER, J., “*Estimating the Fractal Dimension from Digitized Images*”, Munich University of Technology, Statistics and Data Processing Institute, <http://www.weihenstephan.de/ane/algorithms/algorithms.html>, acessado em 12/04/2003.
- [LEE, CHEN, HSIEH 2002] LEE, W.; CHEN, Y.; HSIEH, K. “Robust Calculation of Fractal Dimension of Images and Its Applications to Classification of Ultrasonic Liver Images and Texture Images”, in *Proceedings International Symposium on Circuits and Systems*, IEEE, 2002, volume 2, pp 656-659.
- [LORENZ 1996] LORENZ, E. “A Essência do Caos”, Editora UnB, Brasília, Brasil, 1996.
- [MANDELBROT 1977] MANDELBROT, B.B. “*The Fractal Geometry of Nature*”, W.H. Freeman and Company, New York, EUA, 1977
- [MARAGOS, SUN 1993] MARAGOS, P.; SUN, F.K. “Measuring the Fractal Dimension of Signals: Morphological Covers and Iterative Optimization”, in *Proceedings Transactions of Signal Processing*, IEEE, 1993, volume 14, número 1, pp 108-121.
- [MENOTI 2003] MENOTI, D. Dissertação de Mestrado “Segmentação de Envelopes Postais para Localização do Bloco de Endereço: Uma abordagem baseada em uma seleção de características no espaço *Wavelet*”, Pontifícia Universidade Católica do Paraná, Programa de Pós-Graduação em Informática Aplicada, 2003.
- [NEARY 2000] NEARY, D. “Fractal methods in Image analysis and Coding”, thesis MsC, Dublin City University, Dublin, 2000.
- [PAL 1989] PAL, N.R.; PAL, S.K. “*Entropic Thresholding*”, *Signal Processing*, volume 16, pp 97- 108.
- [PARKER 1997] Parker, J.R “*Algorithms for Image Processing and Computer Vision*”, John Willey & Sons, Toronto, 1997

- [PENTLAND 1984] PENTLAND, A.P. "Fractal based description of natural scenes", in *Proceedings Transactions on Pattern Analysis and Machine Intelligence*, IEEE, Novembro 1984, PAMI-6(6), pp 661-672, 1984
- [POPPER 1976] Popper, K.R. "Unended Quest : An Intellectual Autobiography", Fontana, Londres, 1976
- [RINGLER 2003] RINGLER, A. "Texture Segmentation by local fractal dimension as applied to oceanic search and rescue", in *Proceedings International Conference on Information Communications Signal Processing*, IEEE, 2003, volume 2, pp 975-979.
- [SARKAR, CHAUDHURI 1992] SARKAR, N; CHAUDHURI, B. "An efficient approach to estimate fractal dimension of textural images". *Pattern Recognition*, 25(9), pp 1035-1041, 1992
- [THEODORIDIS, KOUTROUMBAS 1999] THEODORIDIS, S.; KOUTROUMBAS, K. "*Pattern Recognition*", Academic Press, EUA, 1999
- [TOU, GONZALEZ 1974] TOU, J.T.; GONZALEZ, R.C. "*Pattern Recognition Principles*", Addison-Wesley Publishing Company, Reading, EUA, 1974.
- [VOSS 1986] VOSS, R. "Random Fractals: Characterization and measurement in scaling Phenomena and disordered Systems". R. Pynn and A. Skjeltorp Editors, Plenum Press, New York, 1986.
- [VUDUC 1997] VUDUC, R. "Image Segmentation using fractal dimension", *Technical Report*, Cornell University, Maio 1997

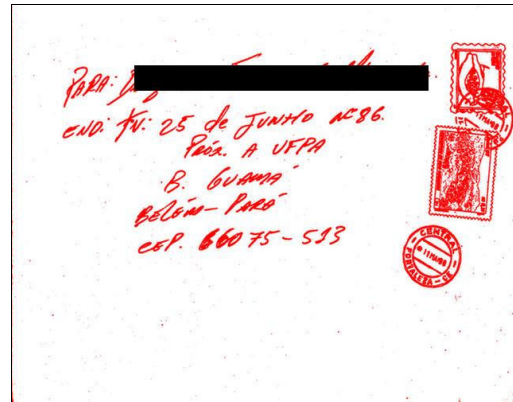
Anexos

A. Exemplos da Imagem Real carta040 com variações dos parâmetros

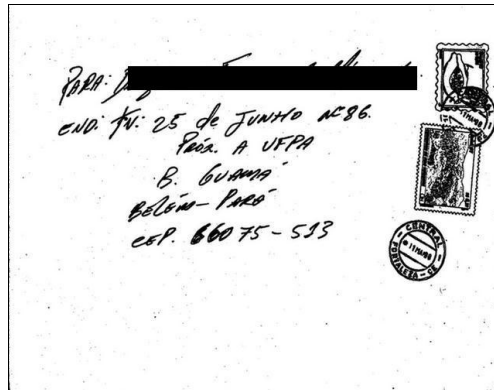
1. Imagem com os seguintes parâmetros: faixa da janela igual a $[3,5]$, número de aglomerados igual a 2 e taxa de melhoria mínima igual a 0,0001



(a)



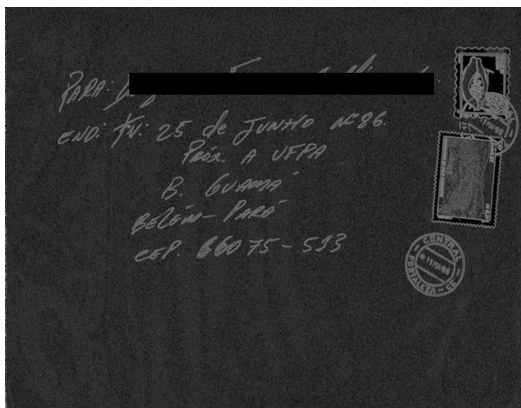
(b)



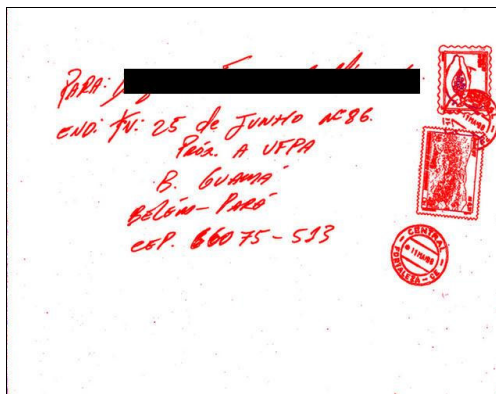
(c)

Figura A.1 carta040 $[3,5]c2$ (a) imagem da característica dimensão fractal, (b) imagem *K-means*, (c) imagem segmentada

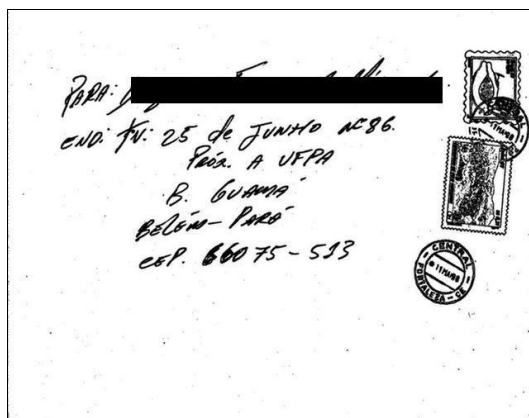
2. Imagem com os seguintes parâmetros: faixa da janela igual a [3,5], número de aglomerados igual a 3 e taxa de melhoria mínima igual a 0,001



(a)



(b)



(c)

Figura A.2 carta040 [3,5]i001 (a) imagem da característica dimensão fractal, (b) imagem *K-means*, (c) imagem segmentada

3. Imagem com os seguintes parâmetros: faixa da janela igual a $[3,7]$, número de aglomerados igual a 2 e taxa de melhoria mínima igual a 0,0001

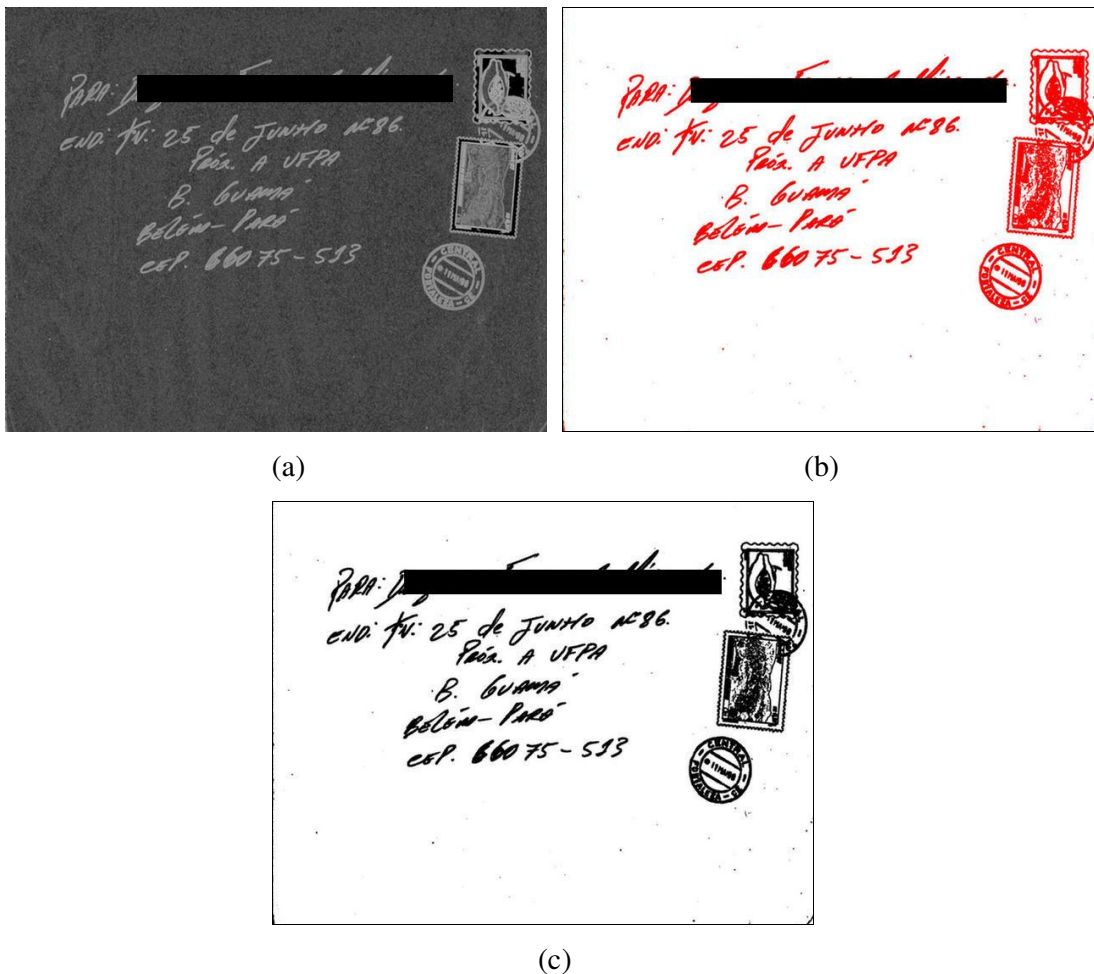
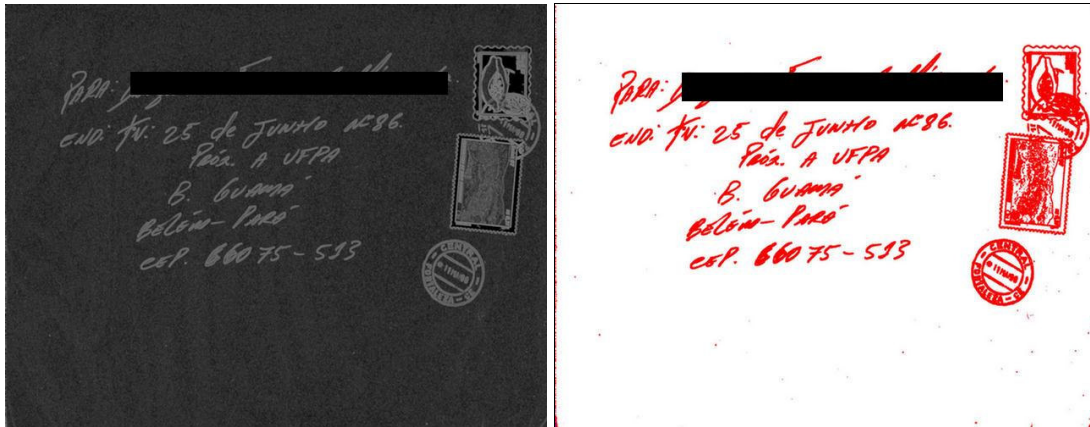


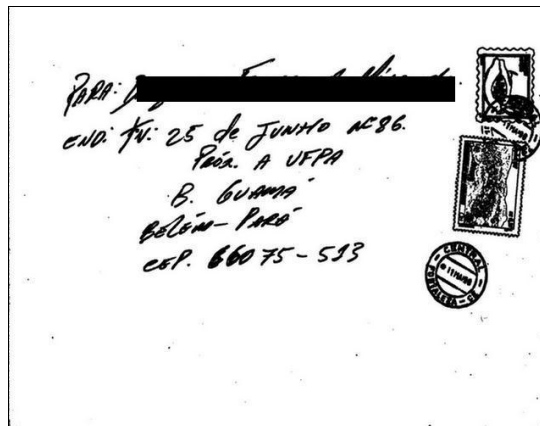
Figura A.3 carta040 $[3,7]c2$ (a) imagem da característica dimensão fractal, (b) imagem *K-means*, (c) imagem segmentada

4. Imagem com os seguintes parâmetros: faixa da janela igual a [3,7], número de aglomerados igual a 3 e taxa de melhoria mínima igual a 0,001



(a)

(b)



(c)

Figura A.4 carta040 [3,7]i001 (a) imagem da característica dimensão fractal, (b) imagem *K-means*,
(c) imagem segmentada

5. Imagem com os seguintes parâmetros: faixa da janela igual a [3,9], número de aglomerados igual a 2 e taxa de melhoria mínima igual a 0,0001

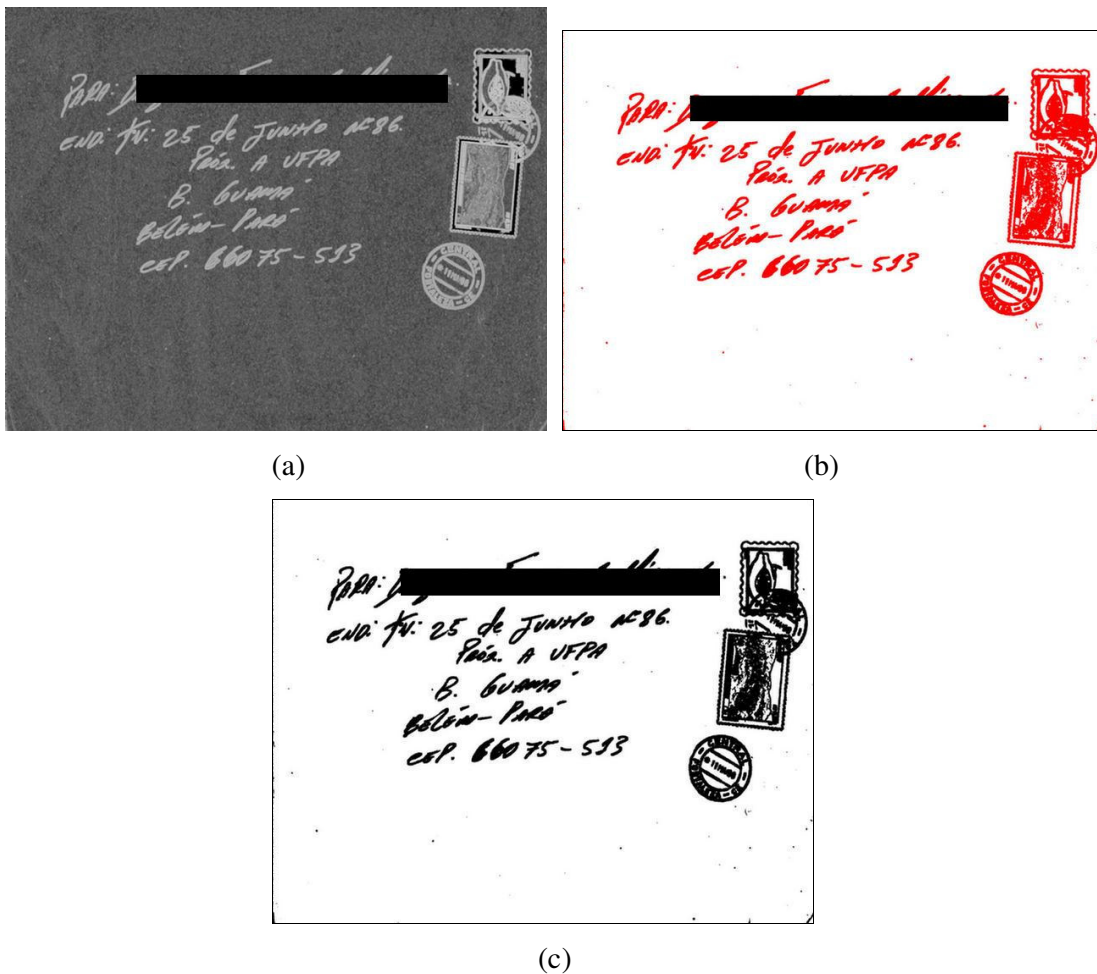
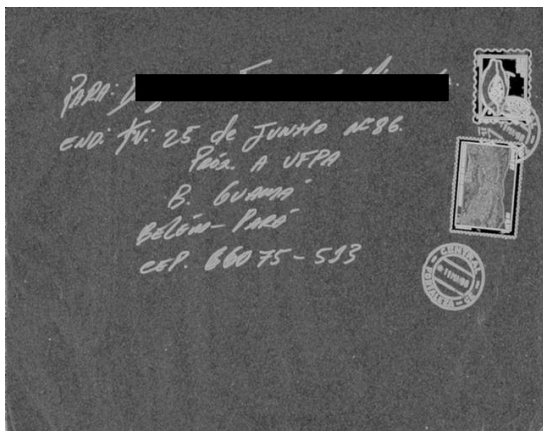
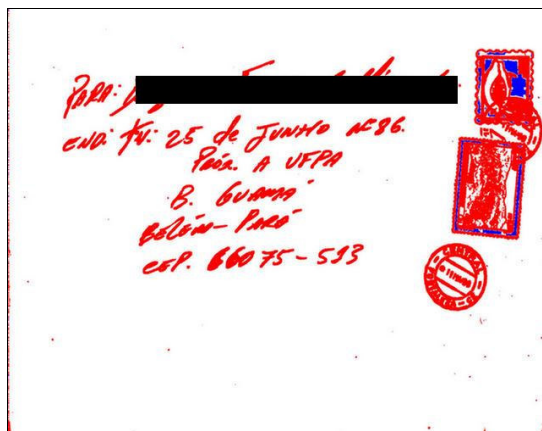


Figura A.5 carta040 [3,9]c2 (a) imagem da característica dimensão fractal, (b) imagem *K-means*, (c) imagem segmentada

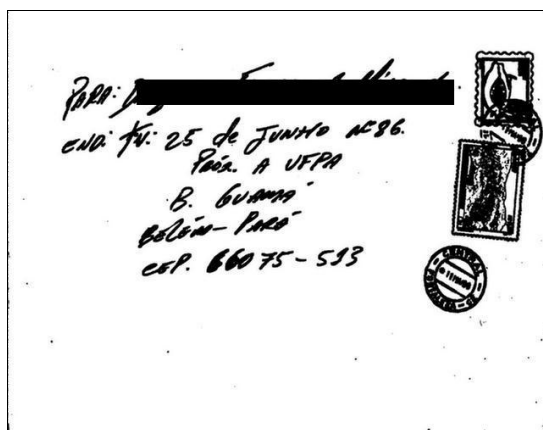
6. Imagem com os seguintes parâmetros: faixa da janela igual a [3,9], número de aglomerados igual a 3 e taxa de melhoria mínima igual a 0,001



(a)



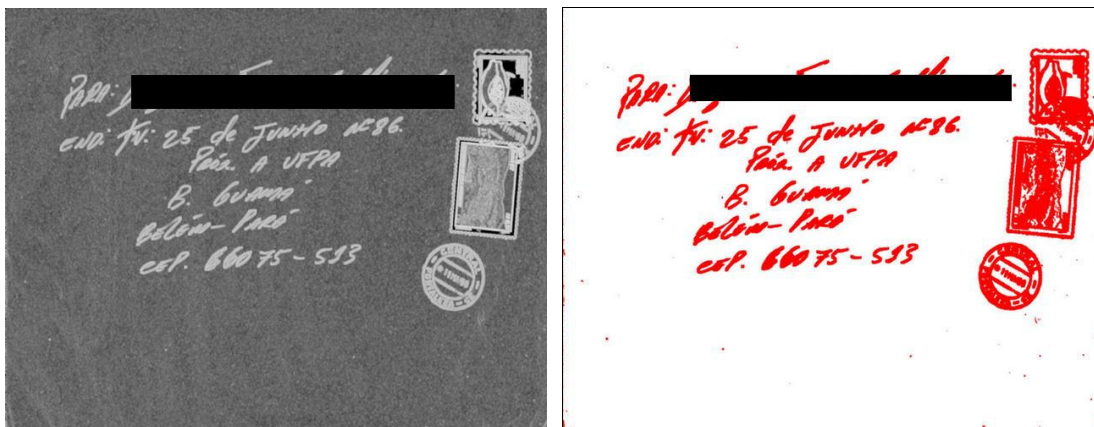
(b)



(c)

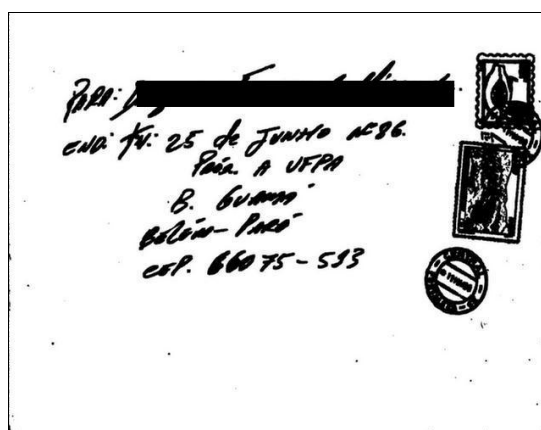
Figura A.6 carta040 [3,9]i001 (a) imagem da característica dimensão fractal, (b) imagem *K-means*,
(c) imagem segmentada

7. Imagem com os seguintes parâmetros: faixa da janela igual a [3,11], número de aglomerados igual a 2 e taxa de melhoria mínima igual a 0,0001



(a)

(b)



(c)

Figura A.7 carta040 [3,11]c2 (a) imagem da característica dimensão fractal, (b) imagem *K-means*, (c) imagem segmentada

8. Imagem com os seguintes parâmetros: faixa da janela igual a [3,11], número de aglomerados igual a 3 e taxa de melhoria mínima igual a 0,001

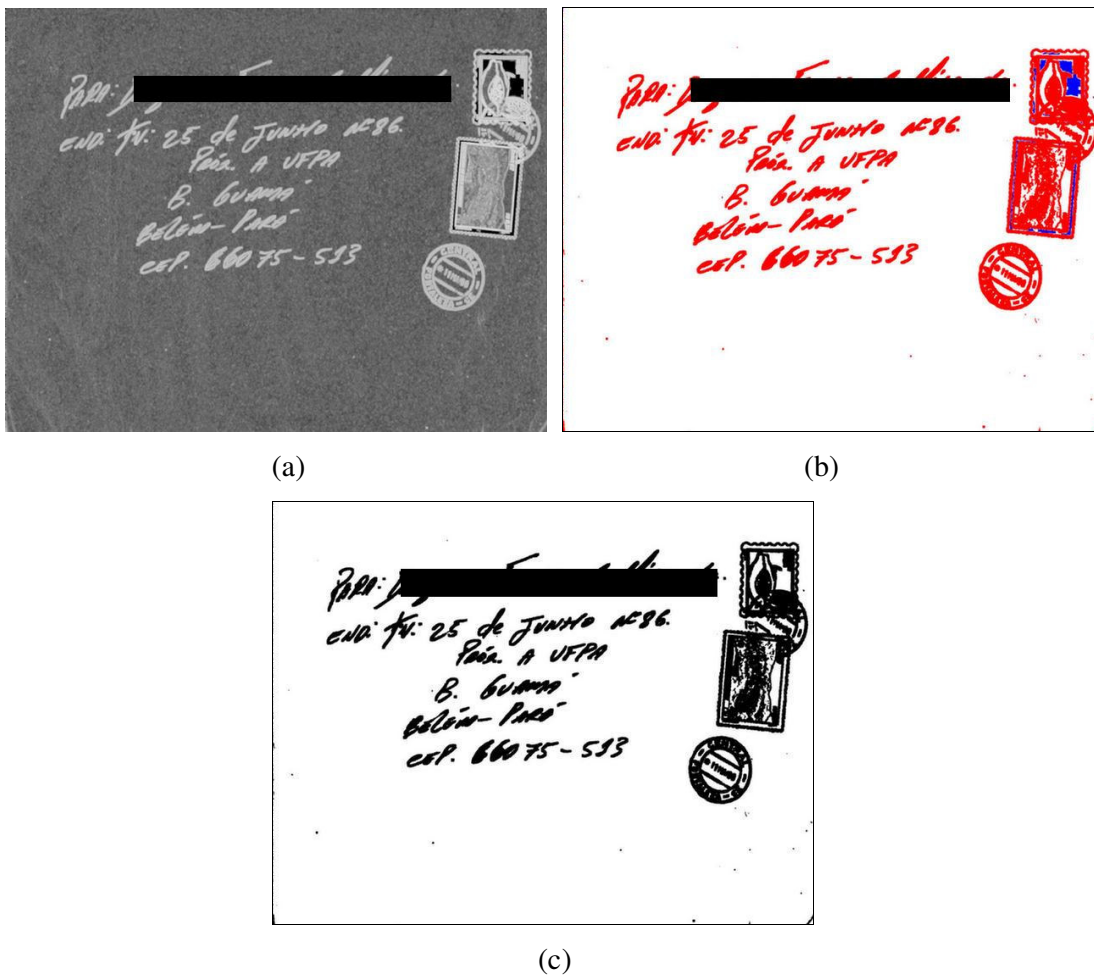


Figura A.8 carta040 [3,11]i001 (a) imagem da característica dimensão fractal, (b) imagem *K-means*, (c) imagem segmentada

9. Imagem com os seguintes parâmetros: faixa da janela igual a [3,13], número de aglomerados igual a 2 e taxa de melhoria mínima igual a 0,0001

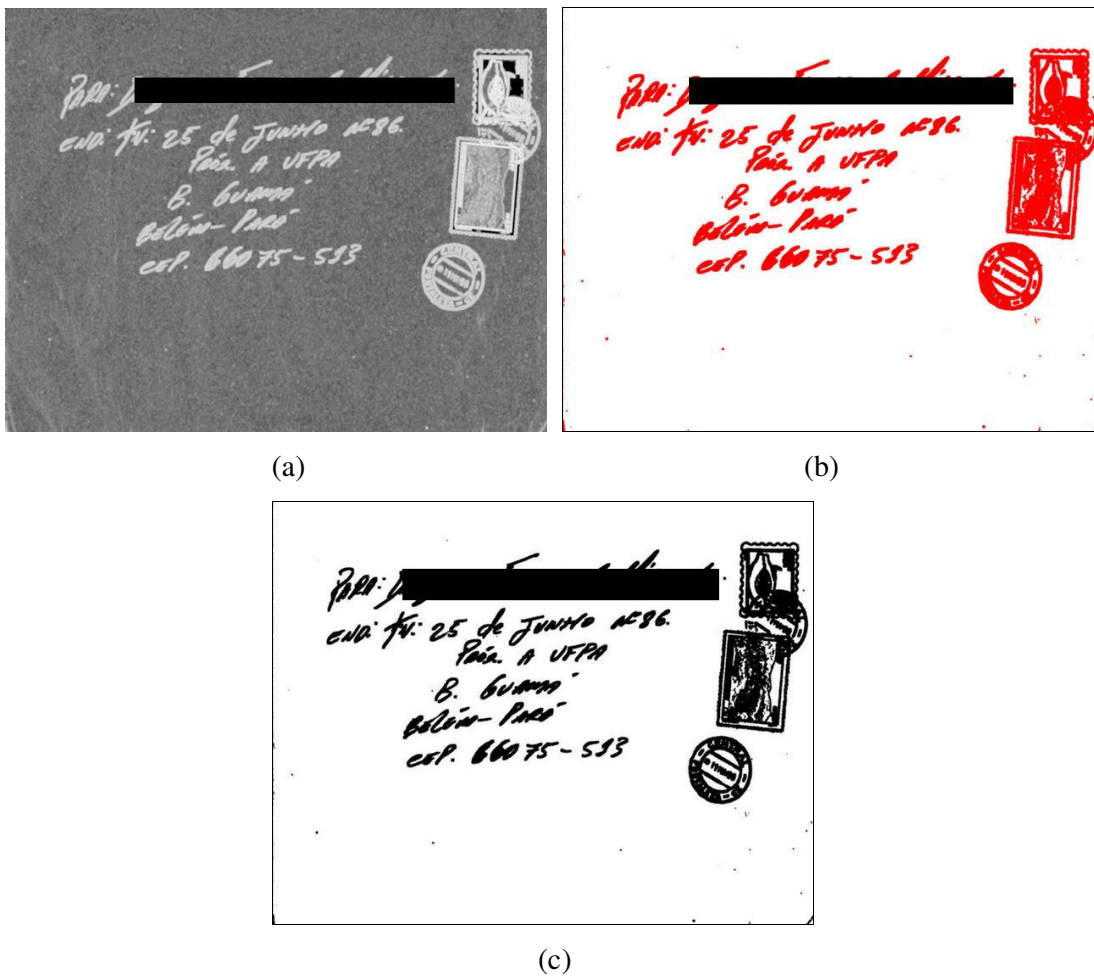
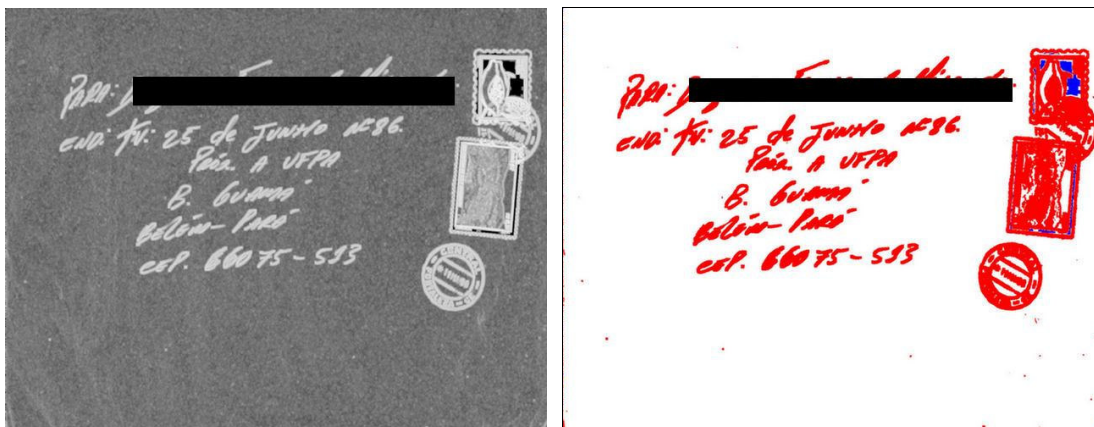


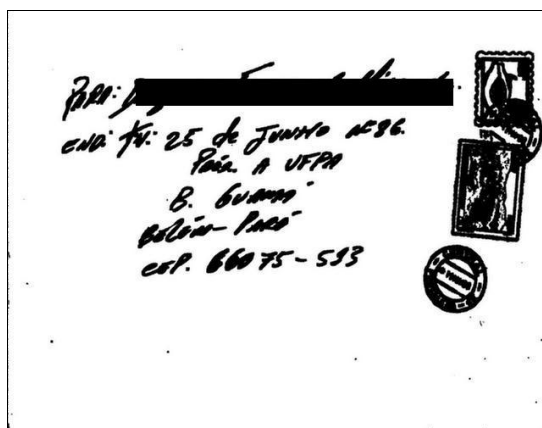
Figura A.9 carta040 [3,13]c2 (a) imagem da característica dimensão fractal, (b) imagem *K-means*, (c) imagem segmentada

10. Imagem com os seguintes parâmetros: faixa da janela igual a [3,13], número de aglomerados igual a 3 e taxa de melhoria mínima igual a 0,001



(a)

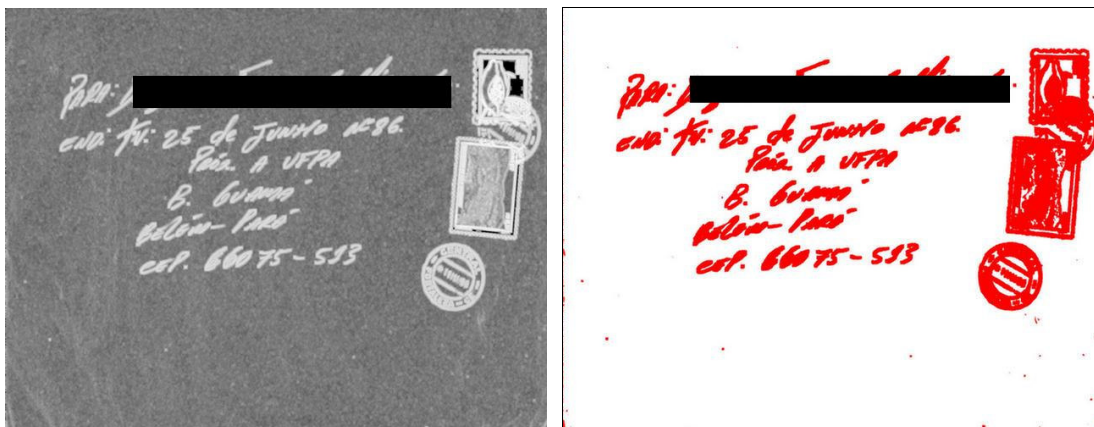
(b)



(c)

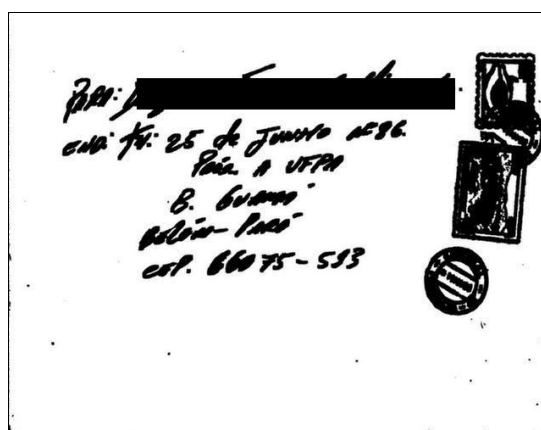
Figura A.10 carta040 [3,13]i001 (a) imagem da característica dimensão fractal, (b) imagem *K-means*,
(c) imagem segmentada

11. Imagem com os seguintes parâmetros: faixa da janela igual a [3,15], número de aglomerados igual a 2 e taxa de melhoria mínima igual a 0,0001



(a)

(b)



(c)

Figura A.11 carta040 [3,15]c2 (a) imagem da característica dimensão fractal, (b) imagem *K-means*,
(c) imagem segmentada

12. Imagem com os seguintes parâmetros: faixa da janela igual a $[3,15]$, número de aglomerados igual a 3 e taxa de melhoria mínima igual a 0,001

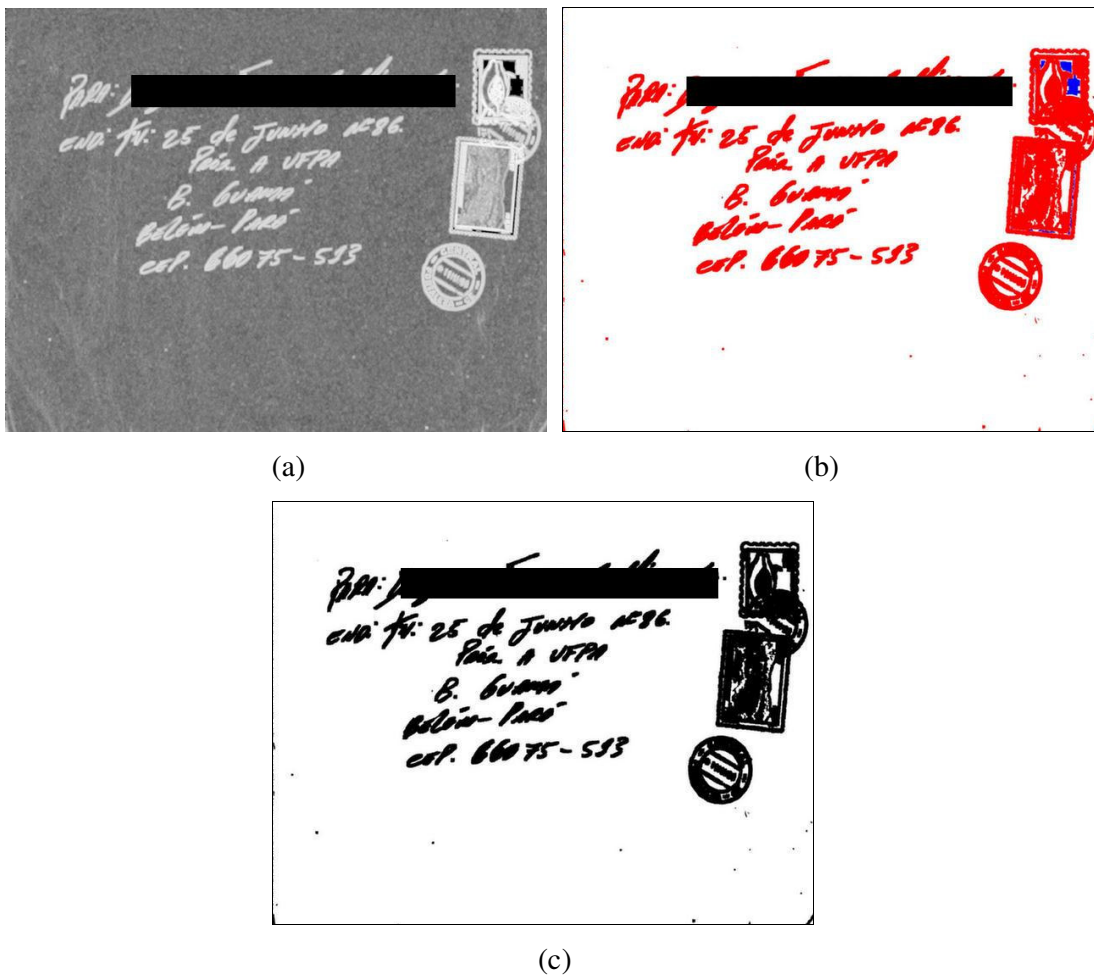
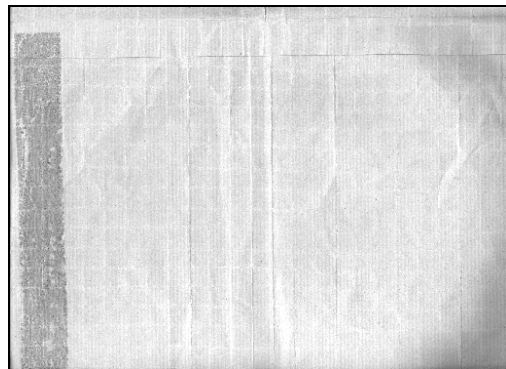
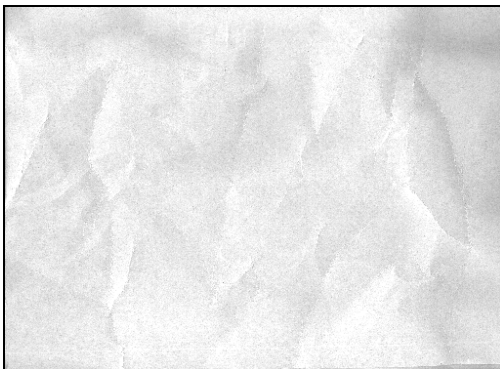
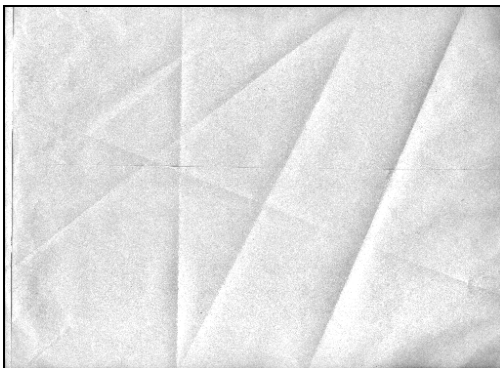


Figura A.12 carta040 $[3,15]i001$ (a) imagem da característica dimensão fractal, (b) imagem K-means, (c) imagem segmentada

B. Variedade de fundos de envelopes complexos

A variedade de fundos de envelopes apresentada nesta seção corresponde aos tipos de fundos de envelopes complexos sem objetos de segmentação que foram utilizados para gerar a base de imagens sintetizadas. São considerados complexos porque, apresentam uma variação expressiva dos tons cinza e, também apresentam na sua superfície bastantes irregularidades. Estes fundos correspondem à face reversa de alguns envelopes de correspondências comerciais. Uma parte da face reversa foi escolhida, pois a mesma não contém nenhum objeto que se assemelhe aos objetos do envelope que poderiam vir a prejudicar (ou interferir) na avaliação do método proposto.



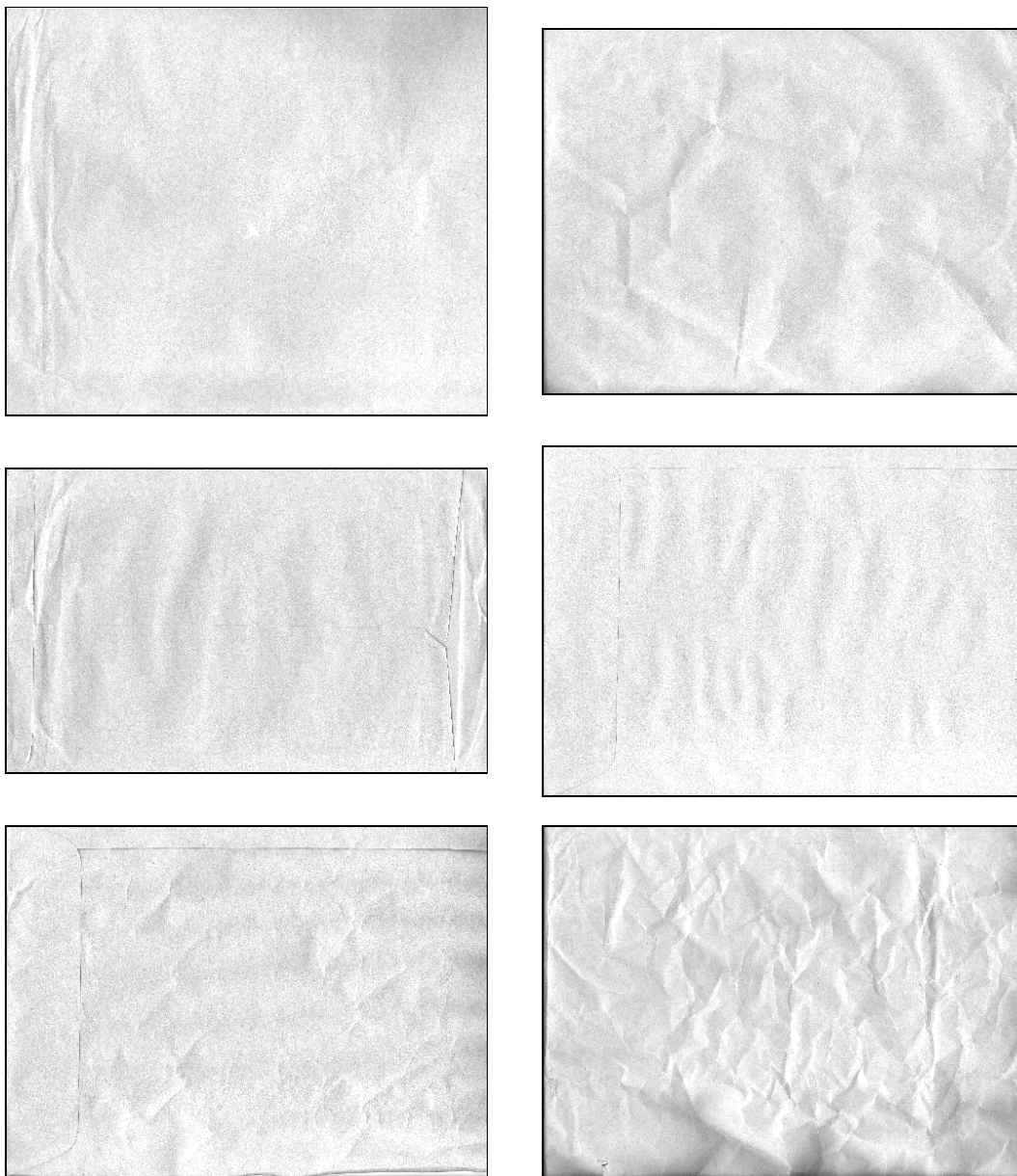


Figura B.1: Variedade de Fundos de envelopes complexos da base de imagens sintetizadas.