



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-15210-TDI/1307

UM PARADIGMA PARA RE-SEGMENTAÇÃO DE IMAGENS DE ALTA RESOLUÇÃO

Thales Sehn Korting

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelos Drs. Leila Maria Garcia Fonseca e Luciano Vieira Dutra, aprovada em
20 de dezembro de 2007.

O original deste documento está disponível em:
<<http://urlib.net/sid.inpe.br/mtc-m17@80/2008/02.12.12.07>>

INPE
São José dos Campos
2007

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6911/6923

Fax: (012) 3945-6919

E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO:

Presidente:

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Membros:

Dr. Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Haroldo Fraga de Campos Velho - Centro de Tecnologias Especiais (CTE)

Dra. Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Jefferson Andrade Anselmo - Serviço de Informação e Documentação (SID)

Simone A. Del-Ducca Barbedo - Serviço de Informação e Documentação (SID)

Vinicius da Silva Vitor - Serviço de Informação e Documentação (SID) - bolsista

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Marilúcia Santos Melo Cid - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva e Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Viveca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-15210-TDI/1307

UM PARADIGMA PARA RE-SEGMENTAÇÃO DE IMAGENS DE ALTA RESOLUÇÃO

Thales Sehn Korting

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelos Drs. Leila Maria Garcia Fonseca e Luciano Vieira Dutra, aprovada em
20 de dezembro de 2007.

O original deste documento está disponível em:

<http://urlib.net/sid.inpe.br/mtc-m17@80/2008/02.12.12.07>

INPE
São José dos Campos
2007

Dados Internacionais de Catalogação na Publicação (CIP)

K967u Korting, Thales Sehn.

Um paradigma para re-segmentação de imagens de alta resolução/ Thales Sehn Korting. – São José dos Campos: INPE, 2007.

64p. ; (INPE-15210-TDI/1307)

1. Segmentação de imagens. 2. Segmentação baseada em grafos. 3. Re-segmentação. 4. Imagens urbanas. 5. Sensoriamento remoto. I. Título.

CDU 621.376.5

Copyright © 2007 do MCT/INPE. Nenhuma parte desta publicação pode ser reproduzida, armazenada em um sistema de recuperação, ou transmitida sob qualquer forma ou por qualquer meio, eletrônico, mecânico, fotográfico, microfilmico, reprográfico ou outros, sem a permissão escrita da Editora, com exceção de qualquer material fornecido especificamente no propósito de ser entrado e executado num sistema computacional, para o uso exclusivo do leitor da obra.

Copyright © 2007 by MCT/INPE. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use of the reader of the work.

Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de Mestre em
Computação Aplicada

Dr. Camilo Daleles Rennó



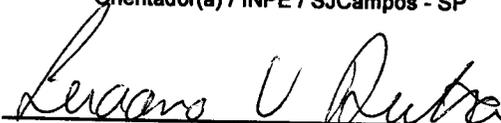
Presidente / INPE / SJC Campos - SP

Dra. Leila Maria Garcia Fonseca



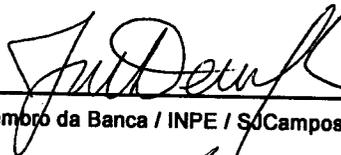
Orientador(a) / INPE / SJC Campos - SP

Dr. Luciano Vieira Dutra



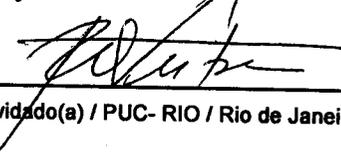
Orientador(a) / INPE / SJC Campos - SP

Dr. José Demisio Simões da Silva



Membro da Banca / INPE / SJC Campos - SP

Dr. Raul Queiroz Feitosa



Convidado(a) / PUC- RIO / Rio de Janeiro - RJ

Aluno (a): Thales Sehn Körting

São José dos Campos, 20 de dezembro de 2007

*“Glória, louvor e gratidão
Honra, força e potência
Pelos séculos a Deus!”*

*Diante de Ti
Martín Valverde*

RESUMO

A segmentação de imagens é uma das mais importantes tarefas na área de processamento de imagens. Utilizada em diversas áreas da ciência, como reconhecimento de caracteres e faces humanas, detecção e classificação de imagens, a segmentação tem recebido grande atenção em sensoriamento remoto, devido aos sensores apresentarem resoluções espectrais e espaciais cada vez melhores. Este trabalho propõe uma metodologia para re-segmentação de imagens urbanas de alta resolução baseada em formas retangulares. O método tem como entrada uma ou mais bandas espectrais, e um conjunto de polígonos resultantes de uma segmentação tradicional. Polígonos adjacentes são conectados em uma estrutura de grafo. Sobre essa estrutura realizam-se buscas por agregamentos cujos formatos sejam retangulares, para objetos comuns ao ambiente urbano, como por exemplo telhados. Para os demais objetos urbanos, *e.g.* praças, árvores, corpos d'água, etc., são utilizadas outras heurísticas de agregamento. Visando facilitar a conexão dos segmentos uma etapa de classificação não supervisionada é realizada. O algoritmo empregado foi o do Mapa Auto-Organizável de Kohonen, ou *Self Organizing Map* (SOM). A estrutura de conexão dos polígonos classificados é apresentada na literatura como Grafo de Regiões Adjacentes, ou *Region Adjacency Graph* (RAG). O RAG possui informação sobre a topologia dos segmentos, que de acordo com sua classe serão agregados para formar novas regiões mais adequadas ao contexto urbano. O resultado da re-segmentação é um novo conjunto de polígonos, que deve melhor representar os elementos de cena presentes na imagem. Um atributo que calcula o grau de retangularidade de um segmento foi proposto para possibilitar o encontro de segmentos com essa característica. Resultados são apresentados e discutidos, de forma a comprovar a acurácia da técnica apresentada, quando comparada com outros métodos de segmentação.

A RE-SEGMENTATION PARADIGM FOR HIGH RESOLUTION IMAGERY

ABSTRACT

Image segmentation is one of the most important tasks in the image processing area. The segmentation problem has received a special attention by the remote sensing community since spectral and spatial resolutions of the optical sensors have been improving continuously. Therefore, this work proposes a re-segmentation method based on rectangular shape applied to high-resolution imagery. The method gets as input one or more spectral channels, and a set of polygons from a previous segmentation. Adjacent polygons are connected in a graph to produce objects of rectangular shape through merging operations. Such operations are applied to objects that typically have rectangular shapes, such as roofs. For other urban objects, *e.g.* parks, trees and lakes, different merging heuristics are applied. To make the polygons connection easy, an unsupervised classification based on Kohonen's Self Organizing Map (SOM) is previously performed. Classified regions compose the graph, called Region Adjacency Graph (RAG), which stores topological relations between polygons. According the polygon classes, a merging strategy is performed. The result is a new set of polygons that better represents the urban environment. In addition, a rectangularity measure to identify rectangular objects in polygons was proposed. The method has been tested on high-resolution images, and it has provided good results when compared with other approaches.

SUMÁRIO

Pág.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

1	INTRODUÇÃO	21
1.1	Objetivo do Trabalho	22
1.2	Organização da Dissertação	25
2	PROBLEMA DE SEGMENTAÇÃO	27
2.1	Métodos de Segmentação	27
2.1.1	Baseados em <i>Pixels</i>	28
2.1.2	Baseados em Área	29
2.1.3	Baseados em Bordas	30
2.1.4	Baseados em Objetos	31
2.1.5	Resumo das Metodologias	32
2.2	Interpretação da Segmentação	33
2.3	Avaliação dos Resultados	33
3	METODOLOGIA	35
3.1	Extração de Atributos	35
3.1.1	Média da Vizinhança	36
3.1.2	Área	36
3.1.3	Perímetro	36
3.1.4	Dimensão Fractal	37
3.1.5	Compacidade	37
3.1.6	Ângulo Principal	37
3.1.7	Grau de Retangularidade	37
3.1.8	Escolha do Atributo de Retangularidade	39
3.2	Classificação Não Supervisionada	40
3.2.1	<i>Self Organizing Map</i> – SOM	41
3.2.2	Treinamento dos Neurônios	42
3.2.3	Processo de Classificação	43

3.3	Rotulação Semântica	44
3.4	Geração do RAG (<i>Region Adjacency Graph</i>)	44
3.5	União dos Polígonos	45
3.5.1	Pré-processamento dos Polígonos	45
3.5.2	Procura por Formas Retangulares	46
4	RESULTADOS	49
4.1	Experimentos Realizados	50
4.1.1	Primeiro Experimento	50
4.1.2	Segundo Experimento	52
4.1.3	Terceiro Experimento	53
4.1.4	Quarto Experimento	54
4.2	Discussão Sobre os Resultados	55
4.2.1	Avaliação dos Experimentos	55
4.3	Interface do Sistema	56
4.3.1	<i>Plugin</i> de Re-Segmentação	57
4.3.2	Detalhes da Implementação	58
5	CONCLUSÕES	61
5.1	Recomendações	62
	REFERÊNCIAS BIBLIOGRÁFICAS	63

LISTA DE FIGURAS

	<u>Pág.</u>
1.1	Passos fundamentais em PDI. 21
1.2	Cena <i>QUICKBIRD</i> de São José dos Campos, Dezembro de 2005. 24
1.3	Segmentação não baseada em formas: a) recorte da Figura 1.2, b) algumas áreas retangulares em destaque e c) resultado da segmentação. 24
3.1	Fluxograma da metodologia proposta. 35
3.2	Cálculo do ângulo principal. 38
3.3	Objeto retangular com $RET \approx 1$: a) região de entrada, b) região em detalhe com seu retângulo envolvente e c) região rotacionada com seu novo retângulo envolvente. 39
3.4	Formas de referência para o comparativo de atributos. 39
3.5	Comparativo de atributos. 40
3.6	Imagem com 5 regiões e o RAG associado. 41
3.7	Exemplo de classificação utilizando SOM: a) imagem de entrada, b) imagem super-segmentada e c) resultado da classificação. 43
3.8	Etapas envolvidas na união dos polígonos. 46
3.9	Exemplo de erro na classificação: a) regiões com erro em destaque e b) regiões após o pré-processamento. 46
3.10	Pré-processamento de polígonos vizinhos de mesma classe: a) regiões de entrada, b) polígonos em destaque e c) resultado do pré-processamento. 47
3.11	Separação de telhados em classes diferentes: a) imagem original com telhados em destaque e b) telhados separados em classes diferentes. 47
4.1	Recortes da área de estudo: a) primeiro experimento, b) segundo experimento, c) terceiro experimento e d) quarto experimento. 49
4.2	Detalhe da super-segmentação para as áreas destacadas da Figura 4.1. 50
4.3	Primeiro experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação. 51
4.4	Segundo experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação. 52
4.5	Terceiro experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação. 53

4.6	Quarto experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação.	54
4.7	Segmentação tradicional para os recortes da área de estudo.	55
4.8	Sistema TerraPixel – Interface padrão.	57
4.9	<i>Plugin</i> para o TerraPixel.	58
4.10	Classes implementadas.	59

LISTA DE TABELAS

	<u>Pág.</u>
2.1 Comparação entre os métodos de segmentação.	32
3.1 Pesos do grafo gerado a partir da Figura 3.6.	44
4.1 Resultados da avaliação qualitativa dos experimentos, comparados com uma segmentação tradicional.	56

LISTA DE ABREVIATURAS E SIGLAS

- IFT – *Image Foresting Transform*
- IHS – *Intensity, Hue e Saturation*
- PDI – *Processamento Digital de Imagens*
- RAG – *Region Adjacency Graph*
- SOM – *Self Organizing Map*

1 INTRODUÇÃO

De acordo com Gonzalez e Woods (2000), há um conjunto de passos fundamentais em qualquer sistema de compreensão de cenas. Apresentados na Figura 1.1, esses passos envolvem diversas etapas de Processamento Digital de Imagens (PDI), desde a aquisição dos dados até a interpretação dos resultados. A etapa de interesse deste trabalho é a segmentação, considerada uma das mais complexas principalmente em relação ao desempenho computacional (CARSON et al., 2002). Extraídos os segmentos da imagem, segue-se à extração de atributos, que caracterizam cada objeto. Estes atributos podem ser de forma, espectrais, estatísticos, etc. Finalmente, as imagens são classificadas, gerando uma saída no formato de um gráfico, tabela, mapas temáticos, entre outros.

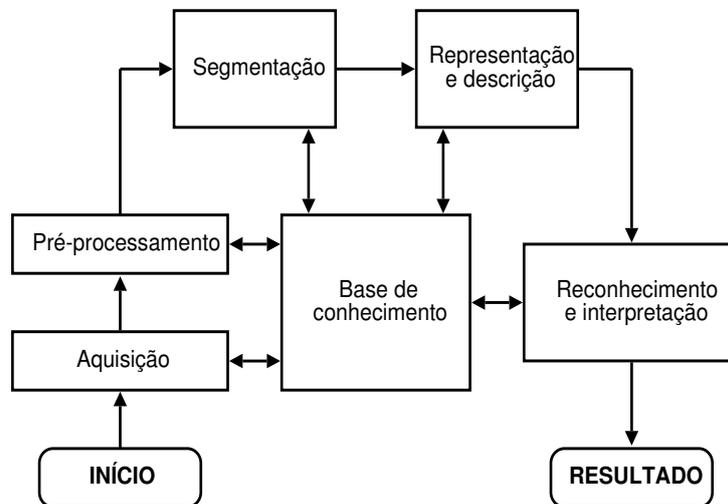


Figura 1.1 - Passos fundamentais em PDI.

Fonte: Adaptação de Gonzalez e Woods (2000).

Pode-se definir a segmentação como a operação que particiona uma imagem em regiões, de modo que os elementos pertencentes a cada região sejam similares em relação a uma ou mais propriedades (JUNG, 2007). Geralmente, a expectativa do resultado de uma segmentação é que os objetos de interesse sejam extraídos automaticamente. Para atingir esse objetivo, normalmente, muitos algoritmos são direcionados apenas a uma classe reduzida de problemas ou dados.

A segmentação é considerada uma das fases fundamentais para a análise de objetos urbanos. Tipicamente o ambiente urbano é constituído de estruturas construídas (prédios,

ruas, telhados, casas, piscinas), vegetação (parques, jardins), áreas de solo exposto e corpos d'água. Imagens de alta resolução, como as obtidas pelos satélites *IKONOS*, *SPOT-5* e *QUICKBIRD* têm sido usadas para obter informação de tais áreas. Estas imagens apresentam alta heterogeneidade espacial e espectral, o que torna o processo de segmentação mais difícil. Com a melhoria da qualidade das imagens, novas ferramentas e conceitos de análise têm surgido, visando aprimorar o mapeamento do espaço urbano e do uso do solo (HEROLD et al., 2002).

A classificação baseada em objetos, usada no sistema *eCognition* (BAATZ; SCHÄPE, 1999), é fundamentada na segmentação multi-resolução, cujo objetivo é identificar áreas com alta homogeneidade. Tal estratégia realiza diferentes segmentações na mesma imagem, gerando resultados que podem ser caracterizados como **super-segmentação** e **sobre-segmentação**¹. Junto a esses resultados são criadas redes semânticas e hierárquicas, contemplando os diversos níveis de segmentação. A hierarquia auxilia na posterior classificação dos objetos, dadas as relações semânticas de pertinência entre os conjuntos de segmentos.

Objetos do ambiente urbano, tais como ruas e habitações, podem ser mais bem avaliados por meio de atributos de forma, textura e contexto. Consequentemente, as análises podem ser melhoradas quando se observam os objetos como grupos de *pixels* com algum grau de homogeneidade. Segundo Antunes e Cortese (2007), os processos de segmentação e classificação de objetos urbanos são mais eficientes quando utilizam atributos espaciais (forma, tamanho e orientação), além dos atributos espectrais. Em sistemas de classificação baseada em objetos, a segmentação é o passo inicial e um dos mais importantes, considerando que se o resultado não for adequado, todos os processamentos subsequentes podem ser prejudicados. Além disso, heurísticas baseadas na própria visão humana podem ser utilizadas para buscar elementos de interesse em imagens.

1.1 Objetivo do Trabalho

Os segmentadores comumente detectam regiões com uniformidade espectral, o que nem sempre resulta em objetos com formas geométricas básicas, como retângulos ou círculos. No entanto, em algumas aplicações é desejável que o algoritmo identifique formas pré-definidas. Desta maneira, um método para detecção de objetos em cenas urbanas através da operação de re-segmentação se faz desejável. A re-segmentação tem como objetivo reorganizar um conjunto de polígonos, agregando elementos deste conjunto, para gerar objetos com características morfológicas e geométricas adequadas ao contexto da aplicação.

¹Ver mais detalhes no Capítulo 2.

Dadas as características das imagens de áreas urbanas e dos objetos de interesse, os atributos de forma são parâmetros importantes para caracterizá-los. Geralmente, os resultados de uma segmentação apresentam formas que não correspondem às que o sistema visual humano percebe, mas que são importantes para a aplicação. Portanto, tomando como hipótese que a busca por formas pré-definidas na re-segmentação é capaz de melhorar os resultados de uma segmentação, este trabalho apresenta um paradigma de re-segmentação para aprimorar o resultado de uma segmentação inicial, no contexto da aplicação.

Considerando que detecção e modelagem de feições urbanas são importantes tópicos de pesquisa e apresentam diversas aplicações, tais como planejamento urbano e navegação virtual por cidades (YONGGUAN et al., 2004), este estudo será focado em imagens urbanas. Isto não impede que o método proposto seja aplicado a outros tipos de imagens (agricultura, floresta, etc.). A Figura 1.2 mostra um exemplo de uma área urbana, de 512×512 *pixels*, composição colorida R3G2B1 *QUICKBIRD* da cidade de São José dos Campos (SP – Brasil), obtida em Dezembro de 2005. A idéia é usar imagens que contemplem a realidade das cidades do Brasil, onde a estrutura de arruamento é bastante complexa. Essa Figura ressalta ainda uma área de 128×128 *pixels*, exibida na Figura 1.3a, com algumas formas retangulares destacadas na Figura 1.3b. Nos capítulos seguintes mostra-se que um sistema de visão computacional baseado em objetos pode obter resultados superiores aos demais sistemas que buscam similaridade dentro de uma vizinhança delimitada apenas pelos *pixels*.

A Figura 1.3c apresenta o resultado da segmentação da Figura 1.3a, obtida pelo método de crescimento de regiões, implementado no sistema SPRING (CAMARA et al., 1996). Pode-se observar que os objetos destacados como retangulares são fragmentados em diversos segmentos, ainda que semanticamente representem uma só região. Isso acontece devido às diferenças espectrais inerentes ao processo de geração das imagens. Considerando que a união de duas ou mais regiões pode resultar em uma nova região com um significado mais adequado para o contexto urbano, a hipótese da re-segmentação deve ser investigada. A chamada re-segmentação baseada em objetos retangulares consiste em realizar operações entre os polígonos resultantes de uma segmentação, de modo a maximizar os atributos de forma dos novos segmentos. Elementos tipicamente retangulares, como telhados e prédios, deveriam ser representados por segmentos com formas retangulares e assim por diante.

Os resultados experimentais apresentados neste trabalho foram gerados utilizando o aplicativo TerraPixel², no qual podem ser acoplados módulos de PDI. Este aplicativo e os algoritmos desenvolvidos usam as classes e funções da biblioteca TerraLib (CAMARA et al., 2000) como plataforma de desenvolvimento. A seguir, um resumo das principais contri-

²Sistema disponível no endereço <http://www.dpi.inpe.br/terrapixel/>.



Figura 1.2 - Cena *QUICKBIRD* de São José dos Campos, Dezembro de 2005.

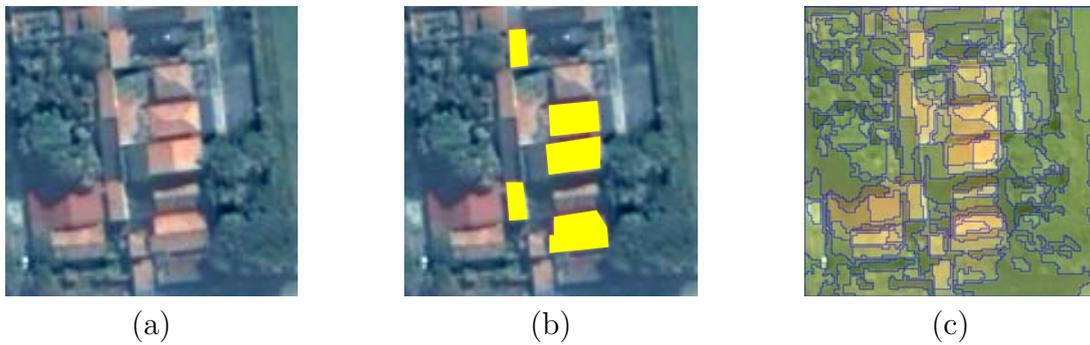


Figura 1.3 - Segmentação não baseada em formas: a) recorte da Figura 1.2, b) algumas áreas retangulares em destaque e c) resultado da segmentação.

buições deste trabalho é apresentado:

- Algoritmo de Re-Segmentação: foi desenvolvido um algoritmo que recebe como entrada uma imagem e um conjunto de polígonos (também referenciados como regiões), e que realiza operações de agregamento destes polígonos. A estrutura de conexão dos polígonos é armazenada em um grafo, onde as regiões adjacentes possuem conexões, com pesos definidos conforme a aplicação. O conjunto resultante representa melhor as estruturas identificadas pelo sistema visual humano, as quais compreendem formas retangulares para telhados, e formatos contínuos

para ruas, árvores e demais elementos tipicamente urbanos.

- Módulo de Classificação: As regiões de entrada pertencem a diferentes classes, e deverão ser re-segmentadas considerando a vizinhança. Essas classes, numa imagem urbana, compreendem telhados, ruas, árvores, entre outros. Um algoritmo de classificação não supervisionada baseado no *Self Organizing Map* proposto por Kohonen (2001) foi implementado para realizar esta etapa. Sua estrutura de dados foi otimizada para ser genérica a qualquer conjunto de entrada e para ser integrada à biblioteca TerraLib, como mais uma alternativa de classificação de padrões. A estrutura desenvolvida está detalhada na Subseção 4.3.2.
- Atributo de “Retangularidade”: para encontrar formas retangulares, a bibliografia recomenda alguns atributos, como a compacidade, a dimensão fractal (MCGARRIGAL; MARKS, 1995), ou ainda a transformada de Hough híbrida (JUNG; SCHRAMM, 2004). Os atributos de compacidade e dimensão fractal não representam de forma acurada todos os retângulos encontrados e submetidos à avaliação. A transformada de Hough, além de ser um método que utiliza todos os *pixels* da imagem, apresenta alto custo computacional e não se aplica a um conjunto de polígonos, conforme a metodologia deste trabalho. Um novo atributo de retangularidade invariante à rotação foi desenvolvido. Seu cálculo envolve a descoberta do ângulo principal do objeto, o qual é rotacionado para esse ângulo. A seguir, gera-se um novo retângulo envolvente. Considerando sua área original e a área deste novo retângulo, pode-se medir quão retangular é o objeto. Sua formulação está na Subseção 3.1.7.
- Módulo de Extração de Atributos: uma classe de extração de atributos de regiões de imagens foi implementada. Essa classe recebe como entrada um polígono e os valores dos *pixels* internos a ele. Com estes dados, é possível extrair diversos atributos espaciais e espectrais. Tais medidas são utilizadas para caracterizar cada região e classificá-las para posterior união via re-segmentação.

1.2 Organização da Dissertação

Este documento está estruturado da seguinte forma: o Capítulo 2 apresenta uma revisão sobre os principais métodos de segmentação, além dos aspectos práticos de implementação e avaliação dos resultados. No Capítulo 3, é apresentada a metodologia de re-segmentação, descrevendo todas as etapas implementadas. Para avaliar a técnica proposta, alguns resultados experimentais e a avaliação destes resultados são apresentados no Capítulo 4. Finalmente, o Capítulo 5 apresenta um resumo do trabalho, conclusões e recomendações para trabalhos futuros.

2 PROBLEMA DE SEGMENTAÇÃO

Antes de propor o método de re-segmentação desenvolvido neste trabalho, uma revisão do problema de segmentação é apresentada. Os métodos mais utilizados e com maior número de citações bibliográficas encontradas na literatura são discutidos aqui. Em seguida, aspectos gerais de implementação, validação dos resultados e percepção humana são apresentados. Tais aspectos devem ser contemplados para validar qualquer método de segmentação de imagens.

2.1 Métodos de Segmentação

Pal e Pal (1993) apresentam uma definição formal para segmentação de imagens. Seja $P()$ um predicado de homogeneidade definido em um grupo de *pixels* conectados. A segmentação é uma partição de um conjunto I em subconjuntos conectados, ou regiões (R_1, R_2, \dots, R_M) de modo que

$$\bigcup_{i=1}^M R_i = I, \quad \text{com } R_i \cap R_j = \phi \quad (2.1)$$

O predicado $P(R_i)$ é válido para todas as regiões, e $P(R_i \cup R_j)$ é falso quando $i \neq j$ e R_i e R_j são vizinhos.

De acordo com Cheng et al. (2001), os métodos de segmentação podem ser divididos em quatro grandes grupos:

- a) baseados em *pixels*;
- b) baseados em área;
- c) baseados em bordas;
- d) baseados em objetos.

Os métodos pertencentes aos três primeiros grupos têm sido amplamente utilizados, além das propostas híbridas, que aplicam diferentes métodos em diversas escalas da mesma imagem. A seguir apresenta-se uma descrição destas metodologias, dando ênfase aos métodos baseados em objetos.

2.1.1 Baseados em *Pixels*

Um método baseado em *pixels* é o de “fatiamento do histograma”, utilizando limiares. Neste caso, supõe-se que as imagens são formadas por regiões contendo diferentes níveis de intensidade de modo que, observando o histograma, os objetos sejam facilmente detectados por um algoritmo de busca por picos e vales (CHENG et al., 2001). Cada pico encontrado corresponde a uma região, que se expande até seus respectivos vales. Segundo Sezgin e Sankur (2004), em diversas aplicações os *pixels* dos objetos contém valores bem diferentes dos *pixels* do fundo, tornando as técnicas de limiarização simples e eficazes na identificação das regiões.

Um dos problemas dos segmentadores baseados em *pixels* é que a definição de uma região é baseada principalmente na similaridade das cores. Isso dificulta a separação de objetos que contenham sombras, diferenças de iluminação ou texturas, tornando as superfícies não homogêneas (CHENG et al., 2001). Em alguns casos, tais problemas podem ser resolvidos utilizando o espaço de cores IHS – *Intensity, Hue e Saturation*, conforme descrito em Gonzalez e Woods (2000). Considerando o espaço urbano, Guo et al. (2007) argumentam que objetos pavimentados apresentam baixo valor na componente Saturação (S), podendo assim ser segmentados no espaço IHS.

Outra abordagem é o algoritmo de *Mean shifts*, um procedimento não paramétrico para estimar gradientes de densidade, proposto por Comaniciu e Meer (1997). As imagens coloridas são convertidas para o espaço $L^*u^*v^*$. Tal espaço possui a propriedade de que a diferença entre o brilho e a cromaticidade entre duas cores é proporcional à distância euclidiana entre elas (CHENG et al., 2001). Desta forma, este sistema de cores facilita a busca de objetos distintos nas imagens. Assim, características significativas dos objetos correspondem às regiões de alta densidade neste espaço.

O método proposto por Shi e Malik (2000), chamado *Normalized cuts*, computa um grafo cujos pesos são as distâncias entre os *pixels* em uma determinada vizinhança. Essas distâncias estão relacionadas com os atributos de similaridade. Dado o grafo, os melhores nodos para se realizarem os cortes na imagem são identificados através do cálculo das principais componentes. Cada corte gera um segmento. Dentro de cada segmento, novas operações no grafo são realizadas, gerando novos subsegmentos até atingir um limiar pré-estabelecido. Os cortes são normalizados por uma função que considera todo o grafo (daí o termo *normalized*) objetivando eliminar a influência de mínimos locais.

Também baseada em grafos, a teoria chamada IFT – *Image Foresting Transform*, apresentada por Falcão et al. (2004), define um caminho de custo mínimo em um grafo cujos

nodos são os *pixels* da imagem e os arcos são as relações de adjacência entre eles. Essas relações são definidas com base na aplicação. Quando alguns nodos são muito distantes, pode-se definir uma separação em dois subgrafos. A partir dessas premissas são realizados cortes no grafo gerando a chamada estrutura de floresta (várias árvores de custo mínimo). No caso da segmentação, cada subgrafo representa um segmento da imagem.

2.1.2 Baseados em Área

Os métodos baseados em área agrupam os *pixels* em regiões homogêneas segundo uma medida de similaridade. São expoentes deste grupo os algoritmos de crescimento de regiões, *watersheds*, *split-and-merge*, entre outros.

Algoritmos de crescimento de regiões geram conjuntos de *pixels*, iniciando com um número limitado de pontos, chamados sementes (BINS et al., 1996). Tais sementes são comparadas aos vizinhos e, segundo um critério de similaridade mínima especificado pelo usuário, grupos de *pixels* semelhantes são agregados. Se um agregamento possuir uma área superior a um limiar, também definido pelo usuário, um segmento na imagem será criado. Essa técnica apresenta resultados satisfatórios quando algum critério de homogeneidade é intrínseco aos objetos presentes na imagem. Segundo Baatz e Schäpe (2000), uma desvantagem é a dependência entre os resultados e o conjunto de sementes que comumente é gerado de maneira aleatória. Outro problema é a ordem na qual as regiões e os *pixels* são examinados, pois conforme Cheng et al. (2001), essa ordem não prioriza a semântica das imagens.

Baseado na morfologia matemática, o algoritmo de *watersheds* é útil para obter partições da imagem e apresenta soluções interessantes para o problema de segmentação quando aplicado a mapas de bordas. De acordo com Jung (2007), os mapas são processados como barreiras topográficas à inundação, iniciando em mínimos locais. Quando a inundação é realizada em duas direções e estas se encontram, uma barreira é criada. No final do processo de inundação, o método gera um conjunto de regiões conectadas e separadas por barragens. O conjunto de regiões dá origem aos objetos segmentados, e as barragens são seus contornos. O Sistema SPRING (CAMARA et al., 1996), disponível para *download* gratuito em <http://www.dpi.inpe.br/spring/>, tem implementados os algoritmos de crescimento de regiões e *watershed*.

Uma outra abordagem, proposta por Sousa (2005), é o segmentador *SegSAR*. Ele pode ser aplicado tanto para imagens de radar quanto ópticas. Este algoritmo usa uma estratégia híbrida, combinando diversas técnicas para segmentar as imagens. O modelo de manchas, também conhecido por modelo *cartoon* (MOREL; SOLIMINI, 1995), é aplicado nas regiões

homogêneas, juntamente com o modelo de textura, aplicado nas regiões não homogêneas, comuns em imagens de radar.

Quando não há divisão entre regiões adjacentes, ou quando os objetos são sobrepostos em uma imagem, o método de separação de regiões (*region splitting*) pode ser empregado. No entanto, em alguns casos as regiões separadas devem ser unidas em uma só, segundo um critério de similaridade pré-definido. A essa operação dá-se o nome de agregamento (*merging*). A união das duas operações mencionadas origina o método de *split-and-merge*. Essa idéia pode ser útil quando os limites entre objetos próximos não são facilmente identificados. Liu e Sclaroff (2004) apresentam duas maneiras de separar e agregar as regiões: o método de detecção de segmentos côncavos, e o da distância entre o ponto e a região limite do objeto. Com relação à estrutura de representação, a técnica de *quadtrees* é comumente usada neste tipo de método.

Em alguns casos, técnicas híbridas para uma segmentação mais precisa são mais adequadas. Uma característica comum desta classe de segmentadores é que os algoritmos, geralmente, iniciam o processamento a partir de uma região não homogênea da imagem. Iterativamente, a imagem é dividida até que todas as regiões homogêneas sejam identificadas (LUCCHESI; MITRA, 2001).

2.1.3 Baseados em Bordas

Os métodos baseados em bordas detectam a descontinuidade dos valores de níveis de cinza, encontrando pontos com mudanças abruptas, que indicam a presença de bordas na imagem. A detecção de bordas não é o resultado da segmentação em si, porém é a entrada para a identificação dos objetos contidos nas suas intersecções (CHENG et al., 2001). A desvantagem deste grupo de segmentadores é a dificuldade em identificar contornos fechados dos objetos. Portanto, além dos detectores de bordas é necessário aplicar outros algoritmos de pós-processamento para fechar as bordas, o que não é uma tarefa trivial (GONZALEZ; WOODS, 2000). Diversos operadores para detecção de bordas em imagens têm sido propostos na literatura, *e.g.* trabalhos de Perona e Malik (1990) e Marr e Hildreth (1980). Muitos deles são baseados em descontinuidades locais presentes nas imagens. Assim, métodos de filtragem que calculam gradientes em várias direções são os mais comuns.

Um algoritmo eficiente de detecção de bordas, proposto por Canny (1986), tem sido bastante utilizado para segmentação. Sua idéia é baseada em alguns critérios básicos:

- a) baixa taxa de erros na identificação das bordas;
- b) localização correta dos pontos de borda;

c) obtenção de bordas de apenas 1 *pixel* de largura.

Esses métodos são eficientes quando as imagens de entrada possuem alto contraste, delimitando as regiões. No entanto, o desempenho dos métodos de segmentação pode ser melhorado através do uso de critérios que reflitam um conhecimento maior sobre o tipo de objeto a ser segmentado.

2.1.4 Baseados em Objetos

Objetos em uma imagem são grupos de *pixels* conectados que possuem características semelhantes. Tais características os distinguem de outras entidades presentes na imagem. Informações sobre o formato dos objetos de interesse não são consideradas na maioria dos métodos propostos na literatura. Os *pixels* são unidades que isoladamente não representam os objetos tais como o ser humano identifica ou observa. Uma maneira de contornar essa situação é iniciar o processo de agregamento de regiões utilizando os chamados *blobs* ao invés de *pixels*. Segundo Hay et al. (2005), *blobs* são pequenas regiões homogêneas significativamente diferentes dos seus elementos vizinhos.

Técnicas de segmentação baseadas em objetos consideram não somente as características espectrais dos *pixels* mas também a sua vizinhança ou informação contextual. O resultado é a criação de objetos que possuam homogeneidade espectral e de forma (JENSEN, 1995). De acordo com Blaschke (2003), os objetos em uma imagem contém outros atributos, além da informação espectral, que os caracterizam. Assim, resultados mais promissores derivam da utilização destes algoritmos baseados em objetos. Segundo Tao et al. (2002), a extração automática de objetos “construídos pelo homem” (*e.g.* casas, prédios) é ainda uma tarefa desafiadora em aplicações na área de sensoriamento remoto. Telhados, por exemplo, podem apresentar marcas causadas por diversos motivos, incluindo as intempéries do tempo. Por isso as marcas, sombras e outros artefatos podem dificultar o processo de extração automática dos objetos.

Um eficiente algoritmo de segmentação baseado em objetos é o implementado no sistema *eCognition*¹. Neste algoritmo, cada objeto possui informações sobre sua relação intrínseca com outros nos quais ele se divide ou do qual ele pertence. Também conhece suas relações com os objetos vizinhos na mesma escala (BLASCHKE, 2003). Um dos atributos usados é a homogeneidade da forma dos objetos, que leva em conta (BENZ et al., 2001):

- a relação entre o comprimento de uma borda e o número total de *pixels* de um objeto (*compactness*);

¹Ver <http://www.definiens.com/> e o trabalho de Baatz e Schäpe (1999).

- a relação entre o comprimento da borda e o comprimento do retângulo envolvente, ou *bounding box*, do referido objeto (*smoothness*).

O sistema *eCognition* usa a estratégia multi-escala. Isto significa que existem várias dimensões espaciais nas quais as entidades, os padrões ou processos podem ser observados e medidos. Os objetos de interesse geralmente aparecem simultaneamente nas diferentes escalas em uma imagem (BAATZ; SCHÄPE, 2000). Por isso, este processo de segmentação caracteriza-se como uma técnica *bottom-up* de agrupamento de regiões, iniciando com objetos de 1 *pixel* de tamanho. Durante as diversas iterações, esses objetos pequenos são unidos aos outros para formar objetos cada vez maiores (GAMANYA et al., 2007). A decisão sobre a união ou não dos *pixels* é baseada no parâmetro de homogeneidade, o qual descreve a semelhança entre objetos adjacentes. O significado e a dimensão dos segmentos mudam de acordo com a escala em que o processamento está sendo realizado, pois segundo Hay et al. (2005) não há uma escala ótima para caracterizar a diversidade de objetos presentes em uma cena.

2.1.5 Resumo das Metodologias

Finalizando a descrição dos métodos, a Tabela 2.1 apresenta uma análise comparativa entre as classes de métodos de segmentação, considerando suas vantagens e desvantagens.

Tabela 2.1 - Comparação entre os métodos de segmentação.

Baseados em	Características	Exemplos	Vantagens	Desvantagens
<i>Pixels</i>	Busca informações entre a vizinhança, ou no espaço de atributos.	Fatiamento de histograma, <i>Mean shifts</i> e <i>Normalized cuts</i> .	Bons resultados em imagens contendo regiões com homogeneidade espectral.	Resultados pobres em imagens com textura.
Área	Agrupamento dos <i>pixels</i> em regiões homogêneas.	Crescimento de regiões, <i>Watersheds</i> e <i>Split-and-merge</i> .	Melhores resultados em imagens com homogeneidade intrínseca.	Dependência das sementes e geração de regiões pouco suavizadas.
Bordas	Detecta discontinuidades e usa as bordas como informação para o encontro das regiões.	Canny, Sobel, Laplaciano.	Bons resultados em imagens de alto contraste.	Bordas geralmente não são fechadas. Necessidade de algoritmos para fechamento de bordas.
Objetos	Analisa regiões como objetos, e sua relação com a vizinhança.	Multi-escala, <i>blobs</i> .	Encontra regiões em diversas escalas. Explora relações semânticas de vizinhança e pertinência.	Implementação de alto custo computacional. Dependência do algoritmo aos dados empregados. Geração de métodos <i>ad hoc</i> .

2.2 Interpretação da Segmentação

Em muitos casos, informações importantes necessárias para interpretar uma imagem não são representadas apenas por *pixels*, mas por objetos na imagem e suas relações. Uma regra geral é que a escala dos objetos de uma imagem que se deseja analisar deve ser maior do que a escala do ruído, das distorções ou borramento inseridos pelo sensor (BAATZ; SCHÄPE, 2000).

Segundo Comaniciu e Meer (1997), uma segmentação tradicional pode ser caracterizada pelo nível de detalhes que apresentar, se dividindo em dois tipos; a sobre-segmentação e a super-segmentação:

- Sobre-segmentação: O nível tolerável de homogeneidade dos elementos é baixo, de modo que apenas feições mais significativas ou de interesse à aplicação são detectadas. É apropriado para gerar mapas de bordas e resulta em poucos segmentos;
- Super-segmentação: As diferenças entre os níveis de cinza são grandes o suficiente para que a imagem seja quebrada em várias regiões, de modo que as informações possam ser fundidas utilizando conhecimentos especializados. Neste nível, muitos segmentos são gerados.

No caso do presente trabalho, é desejável que as regiões de entrada sejam resultantes de uma super-segmentação, já que o objetivo principal é unir segmentos. O ser humano não vê o mundo composto de *pixels* individuais, mas como um conjunto de objetos cujos atributos de tamanho, forma, arranjo espacial e contexto dependem da escala na qual são observados (HAY et al., 2005). Conforme Blaschke (2003), o sistema visual humano usa implicitamente conhecimento estrutural em seus processos de classificação, levando em conta as informações sobre a forma e as relações espaciais entre as regiões presentes nas imagens.

2.3 Avaliação dos Resultados

A avaliação dos resultados da segmentação pode ser qualitativa ou quantitativa. A avaliação qualitativa consiste em uma inspeção visual através da análise do número de regiões geradas e esperadas, a coerência das bordas e a preservação das feições lineares. A avaliação quantitativa considera características do conjunto de segmentos, como a sua cardinalidade, os valores médios dos seus elementos, tamanho das regiões, entre outros (SOUSA, 2005). Para Gamanya et al. (2007) um bom avaliador das técnicas de segmentação é sem dúvida o ser humano. Nenhum resultado de segmentação, mesmo se houver

uma prova quantitativa, será convincente se não satisfizer a percepção humana (BAATZ; SCHÄPE, 2000; ANTUNES; CORTESE, 2007).

Dois métodos qualitativos, descritos a seguir, foram empregados na avaliação dos resultados obtidos neste trabalho.

Seja G o conjunto de regiões resultantes da re-segmentação, e E o conjunto de regiões segmentadas manualmente. Define-se $\#(C)$ como um operador de cardinalidade, o qual retorna o número de elementos de um conjunto C . Uma medida de avaliação ($QUANT$) é dada pela relação entre o número de elementos de cada um destes conjuntos:

$$QUANT = \frac{\#(E)}{\#(G)} \quad (2.2)$$

É desejável que o valor de $QUANT$ seja próximo de 1. Neste caso, o número de regiões geradas pelo algoritmo e pelo especialista são semelhantes.

Outra métrica de avaliação que pode ser calculado entre uma referência e o resultado do algoritmo é o erro médio quadrático ($RMSE$) da área dos segmentos. Dado um conjunto de referência R com N elementos, onde cada elemento R_i representa a área da região i , e outro conjunto, O , contendo as áreas (em $pixels^2$) das regiões re-segmentadas, o erro médio quadrático é calculado como:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (R_i - O_i)^2} \quad (2.3)$$

Neste caso, é desejável que o valor do $RMSE$ seja próximo de 0, visto que essa situação evidencia uma reduzida taxa de erros.

3 METODOLOGIA

Este capítulo apresenta o método de re-segmentação proposto neste trabalho. Para facilitar a compreensão da metodologia, esta é apresentada conforme o fluxograma da Figura 3.1, que indica as diversas fases envolvidas no processo.

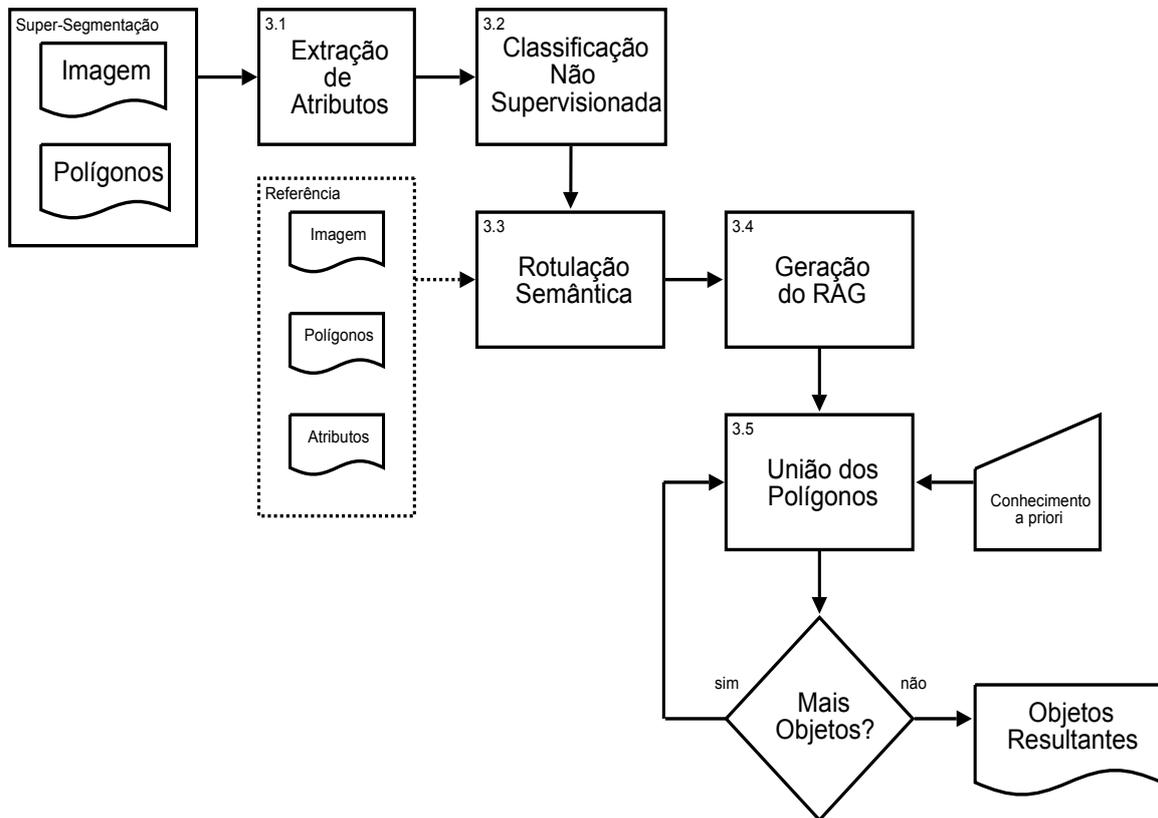


Figura 3.1 - Fluxograma da metodologia proposta.

A entrada é composta por uma ou mais imagens devidamente registradas e por um conjunto de polígonos resultantes da super-segmentação. A saída é dada por um conjunto de objetos cujo significado é adequado ao contexto da aplicação. A seguir os componentes do fluxograma são explicados.

3.1 Extração de Atributos

Seja P o conjunto formado pelos N segmentos de entrada, onde cada elemento P_i delimita um certo número de *pixels*. Dados os valores dos níveis de cinza pertencentes a este elemento de cena e a informação de suas bordas, seus atributos são calculados. Esse conjunto de parâmetros, ou atributos, é composto pelas seguintes medidas:

- a) média dos valores de níveis de cinza do polígono para cada banda;
- b) média dos polígonos da vizinhança;
- c) matriz de covariância entre as bandas;
- d) área (*AREA*);
- e) perímetro (*PERIM*);
- f) dimensão fractal (*FRAC*);
- g) compacidade (*COMP*);
- h) ângulo principal (*ANG*);
- i) grau de retangularidade (*RET*).

Tal conjunto de dados, que é normalizado para evitar diferenças de escala entre os valores, é a entrada do módulo de *Classificação Não Supervisionada*. A escolha do conjunto de atributos foi realizada de maneira empírica, com base na inspeção visual dos resultados obtidos. Esse módulo agrupa os polígonos semelhantes, gerando um número reduzido de classes de objetos ou temas, tais como telhados, árvores, ruas, etc. Essas classes são atribuídas pelo módulo *Rotulação Semântica*, com base em amostras de referência. Os atributos citados anteriormente são descritos a seguir, juntamente com uma avaliação do atributo de retangularidade.

3.1.1 Média da Vizinhança

Com o objetivo de incluir alguma informação de contexto no processo de classificação, as médias dos *pixels* dos polígonos vizinhos são utilizadas. De posse da vizinhança de cada polígono é possível obter as médias de cada vizinho. Assim é gerada uma “média das médias”, atributo que representa a influência da vizinhança sobre o segmento a ser processado.

3.1.2 Área

A medida da área de um polígono, $AREA(P_i)$, é calculada como a quantidade de *pixels* pertencentes a ele, não considerando a resolução da imagem.

3.1.3 Perímetro

O perímetro de uma região, definido por $PERIM(P_i)$, corresponde ao número de *pixels* pertencentes a sua borda.

3.1.4 Dimensão Fractal

A dimensão fractal de um polígono é calculada com base nas medidas de perímetro e área, conforme a equação a seguir:

$$FRAC(P_i) = \frac{2 \ln(0.25 PERIM(P_i))}{\ln(AREA(P_i))} \quad (3.1)$$

O resultado está no intervalo $[1, 2]$, onde valores próximos de 1 correspondem a formas simples, e se aproximam de 2 quando a complexidade da forma aumenta.

3.1.5 Compacidade

A medida de compacidade é calculada pela razão entre o perímetro de uma região $PERIM(P_i)$ e a raiz quadrada de sua área $AREA(P_i)$, conforme a seguinte equação:

$$COMP(P_i) = \frac{PERIM(P_i)}{\sqrt{AREA(P_i)}} \quad (3.2)$$

Essa medida visa representar a quantidade de agrupamento dos elementos de uma região (BAATZ; SCHÄPE, 2000).

3.1.6 Ângulo Principal

Para calcular o ângulo principal de uma região, $ANG(P_i)$, as coordenadas espaciais de todos os pontos internos a ela $C = \{\{x, y\} | \{x, y\} \in P_i\}$ são obtidas. A seguir é calculada a matriz de covariância do conjunto C . Desta matriz são extraídos os auto-vetores. O ângulo do primeiro auto-vetor \mathbf{AV} representa a angulação de P_i , calculada da seguinte maneira:

$$ANG(P_i) = \arctan\left(\frac{AV_y}{AV_x}\right). \quad (3.3)$$

A Figura 3.2 ilustra o cálculo do atributo $ANG(P_i)$.

3.1.7 Grau de Retangularidade

Imagens de áreas urbanas adquiridas por satélites de alta resolução geralmente apresentam objetos com formas padronizadas. Muitos destes objetos, tais como ruas e telhados, são retangulares. Entretanto, devido às propriedades espectrais das imagens, aliadas às

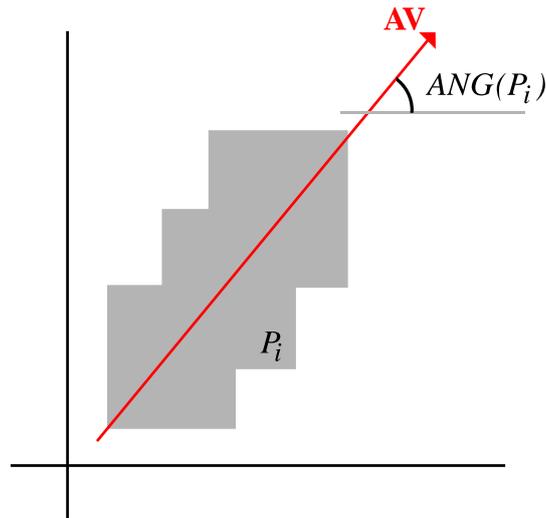


Figura 3.2 - Cálculo do ângulo principal.

deficiências intrínsecas dos sensores, e às condições atmosféricas no momento em que a imagem é gerada, os objetos não apresentam a mesma resposta espectral em toda sua extensão. Assim, a segmentação de imagens urbanas geralmente produz regiões não retangulares, cuja união pode gerar objetos maiores com formato retangular, representando de maneira mais adequada o objeto de interesse.

Para medir o grau de retangularidade de um segmento P_i , pode-se calcular a relação entre sua área $AREA(P_i)$ e a área de seu retângulo envolvente $BOX(P_i)$, definido aqui por $AREA(BOX(P_i))$. Uma estratégia semelhante, que relaciona os perímetros do objeto e de seu retângulo envolvente, é usada em Benz et al. (2001). Entretanto, devido à rotação do objeto, esta medida pode não representar corretamente o grau de retangularidade. Para criar um atributo invariante à rotação, utilizou-se o ângulo principal do polígono, conforme descrito na Equação 3.3. De posse deste ângulo, o objeto P_i é rotacionado, obtendo o novo objeto R_i . Assim, o grau de retangularidade RET , invariante à rotação, é dado por:

$$RET(P_i) = \frac{AREA(P_i)}{AREA(BOX(R_i))}, \quad (3.4)$$

com valores definidos no intervalo $[0, 1]$. Quanto mais próximo de 1 for $RET(P_i)$, mais retangular é o objeto original. A Figura 3.3 mostra um objeto retangular cujo grau de retangularidade é próximo de 1 ($RET \approx 1$).

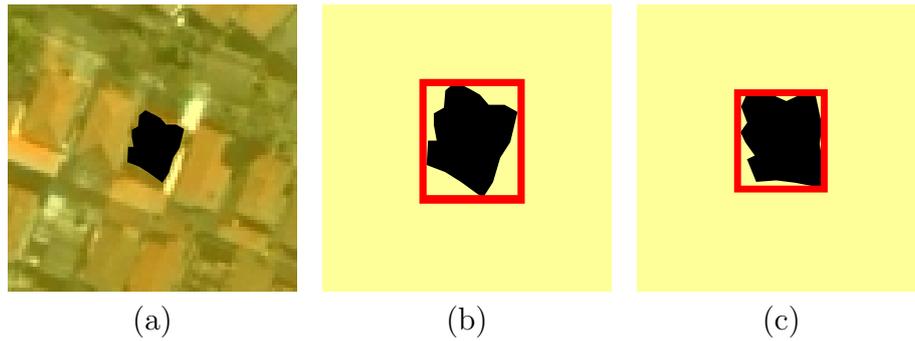


Figura 3.3 - Objeto retangular com $RET \approx 1$: a) região de entrada, b) região em detalhe com seu retângulo envolvente e c) região rotacionada com seu novo retângulo envolvente.

3.1.8 Escolha do Atributo de Retangularidade

Uma avaliação do atributo de retangularidade RET proposto neste trabalho foi realizada. RET é comparado aos atributos de dimensão fractal $FRAC$ e compacidade $COMP$, extraídos de objetos com formas diferentes. Tais formas incluem quadrados, retângulos e elipses em diversas rotações, formatos irregulares e retângulos com ruído nas bordas. Os objetos de referência podem ser visualizados na Figura 3.4. Todos os valores dos atributos estão normalizados no intervalo $[0, 1]$, onde resultados próximos de 1 correspondem a objetos mais retangulares.

1	2	3	4	5	6	7	8	9	10
irr1	irr2	irr3	irr4	irr5	rot1	rot2	rot3	rot4	rot5
11	12	13	14	15	16	17	18	19	20
eli1	eli2	eli3	eli4	eli5	ret1	ret2	ret3	ret4	ret5
21	22	23	24	25					
qua1	qua2	qua3	qua4	qua5					

Figura 3.4 - Formas de referência para o comparativo de atributos.

A Figura 3.5 apresenta os valores dos atributos para as formas de referência. Pode-se

observar que o atributo *RET* representa corretamente o comportamento destes formatos. *FRAC* também apresenta valores próximos de 1 para formatos retangulares. Observa-se ainda que o atributo *RET* apresenta uma pequena instabilidade com objetos quadrados em diferentes rotações, e que o atributo *FRAC* apresentou maior estabilidade. Entretanto, para objetos de forma elíptica, a dimensão fractal também é alta, impossibilitando tal atributo de realizar a distinção entre as duas formas em questão. O atributo *COMP* apresentou os piores resultados nessa comparação. Diante dos resultados apresentados pode-se concluir que o atributo *RET* proposto para caracterizar objetos com forma retangular é mais adequado do que os outros avaliados nesta comparação.

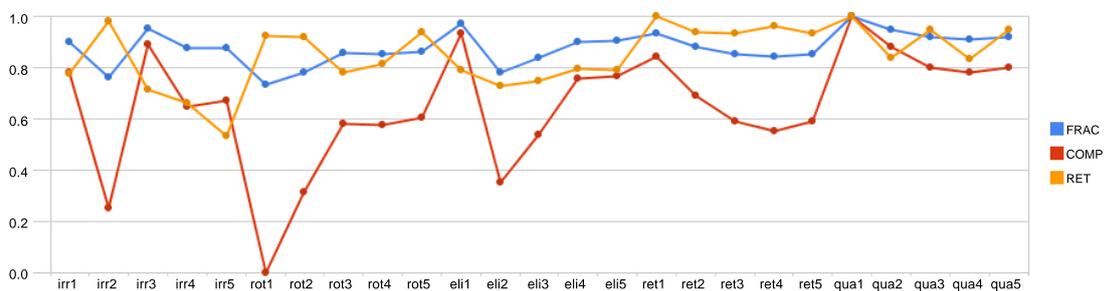


Figura 3.5 - Comparativo de atributos.

3.2 Classificação Não Supervisionada

A entrada deste módulo é um conjunto de segmentos e seus atributos espectrais e de forma. Quando há um número grande de segmentos de entrada e de atributos, a segmentação pode se tornar muito lenta. O custo computacional ainda envolve os processos de extração dos atributos, armazenamento e cálculo das distâncias entre as regiões.

A idéia é reduzir a dimensionalidade do espaço de atributos a único atributo, a classe do elemento que representa os padrões de interesse presentes na imagem. Assim, partindo de amostras que caracterizam os objetos de interesse na busca por formas retangulares, a classificação é realizada. O resultado desta classificação é utilizado na construção do Grafo de Regiões Adjacentes, mais conhecidos como *Region Adjacency Graph* (RAG), descritos em Tremeau e Colantoni (2000). Dado o grafo, o algoritmo leva em conta a forma das regiões vizinhas para re-segmentar a entrada. A Figura 3.6 mostra uma imagem simples, com 5 regiões bem definidas, e o RAG associado a ela.

O algoritmo de classificação usado nesta etapa é o Mapa Auto-Organizável de Kohonen, ou *Self Organizing Map* (SOM). A seguir é apresentada a descrição de como esse algoritmo funciona, seguida dos motivos pelos quais este classificador foi escolhido.

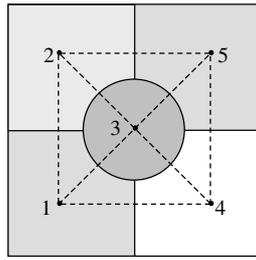


Figura 3.6 - Imagem com 5 regiões e o RAG associado.

3.2.1 *Self Organizing Map* – SOM

Proposto por Kohonen (2001) como uma ferramenta de visualização e análise de dados com várias dimensões, o algoritmo SOM tem sido utilizado para classificação supervisionada e não supervisionada, redução de dimensionalidade, amostragem, mineração de dados, entre outras aplicações (KASKI, 1997). O emprego deste algoritmo no presente trabalho se deu pela comparação de desempenho com outras propostas não supervisionadas, sendo consideradas a velocidade de processamento e a inspeção visual dos resultados.

O funcionamento do SOM é baseado na estrutura das redes neurais humanas. Um conjunto de neurônios é capaz de se especializar na identificação de certos tipos de padrões mediante uma etapa de treinamento. Um importante critério incorporado à estrutura deste algoritmo é o de que informações semelhantes sejam armazenadas por neurônios próximos entre si. Esse conceito confere ao SOM a relação espacial entre os neurônios.

Segundo Bação e Lobo (2007), uma característica relevante do SOM é o desenvolvimento de uma rede ordenada na qual neurônios próximos devem compartilhar semelhanças. Desta maneira, padrões semelhantes ativam áreas próximas no SOM. Pode-se também dizer que partes diferentes são ativadas por tipos específicos de padrões de entrada. Isso leva a uma representação baseada na organização global e na especialização local. O SOM é treinado iterativamente através de um certo número de épocas. Uma época é definida como o processamento de todos os padrões de entrada.

A estrutura de neurônios é composta de diversos pesos. O número de pesos de cada neurônio é definido pela quantidade de atributos dos padrões de entrada. Deve-se mencionar que, além do número de atributos envolvidos, existe outra dimensão relacionada à estrutura da rede de neurônios. Esta diz respeito à localização de cada neurônio e as relações com sua vizinhança. Assim, ela poderá ser $1D$ para neurônios estruturados em uma linha, ou maior quando formar planos ($2D$), superfícies ($3D$) ou hiper-planos (nD).

3.2.2 Treinamento dos Neurônios

O treinamento ocorre através das épocas, para todos os neurônios existentes. Para cada padrão de entrada, o neurônio vencedor é aquele mais próximo ao padrão (VESANTO; ALHONIEMI, 2000). A distância euclidiana foi a medida empregada neste trabalho. Os pesos do neurônio vencedor, que mais se aproximam dos pesos do padrão de entrada, são atualizados. Além disso, levando em conta o parâmetro de vizinhança, neurônios próximos ao vencedor também são atualizados, porém em uma escala reduzida. Esse processo é repetido durante um número de épocas pré-estabelecido, de modo que os neurônios vão se movendo no espaço de atributos, sempre mantendo a relação topológica de vizinhança.

Seja \mathbf{x} um conjunto de N padrões de treinamento, e \mathbf{w} o conjunto de neurônios aqui definido com duas dimensões, de modo que w_{ij} é o neurônio na posição (i, j) . Seja $0 \leq \alpha \leq 1$ a taxa de aprendizado, e $h(w_{ij}, w_{mn})$ a função de vizinhança, que assume valores no intervalo $[0, 1]$, sendo próxima de 0 para neurônios mais distantes. A cada padrão de entrada imposto em uma época, tem-se o neurônio vencedor definido por w_{mn} . Na prática, segundo Kaski (1997), a taxa de aprendizado deve ser ampla nas épocas iniciais, para garantir a ordem global do mapa. Ao longo das épocas ela deve ser diminuída, de modo a proporcionar uma convergência suavizada.

De posse dessas informações, o algoritmo realiza o treinamento do mapa nas etapas de competição entre neurônios (etapas a e b) e atualização (etapa c), da seguinte maneira:

- a) Calcula a distância entre o padrão x_k e todos os neurônios:

$$d_{ij} = |x_k - w_{ij}|$$

- b) Descobre o neurônio vencedor:

$$w_{mn} \rightarrow w_{ij} : d_{ij} = \min(d_{mn})$$

- c) Atualiza todos os neurônios:

$$w_{ij} = w_{ij} + \alpha h(w_{mn}, w_{ij}) |x_k - w_{ij}|$$

- d) Repete o processo até que um número de épocas tenha sido alcançado.

A função para a atualização da vizinhança mais comum é a gaussiana. Ela mede a distância entre as coordenadas dos pesos, ponderadas por um raio de influência r :

$$h(w_{ij}, w_{mn}) = e^{-\frac{1}{2} \frac{(i-m)^2 + (j-n)^2}{r^2}}. \quad (3.5)$$

Os valores de α e r são reduzidos ao longo das épocas, visando auxiliar a convergência dos neurônios.

3.2.3 Processo de Classificação

Dados os neurônios treinados, os novos padrões de entrada são classificados como pertencentes à classe do neurônio mais próximo. Na metodologia de re-segmentação aplicada a imagens urbanas, onde se deseja encontrar objetos retangulares para os telhados, utiliza-se como referência um conjunto de amostras, de modo a isolar a classe de interesse, e sobre ela realizar a busca por formas retangulares. Optou-se por este tipo de classificação para tornar o processo automático, visto que a realização de uma classificação supervisionada acarretaria em uma nova etapa na metodologia, para incluir a interação com o usuário. Além disso, pelo uso de uma referência, é preciso apenas a construção de um conjunto de amostras, aproveitando-o para as demais utilizações do algoritmo. Para as demais classes, onde os segmentos não apresentam retangularidade, a estratégia adotada para a re-segmentação é a simples união dos segmentos adjacentes. A etapa de *Rotulação Semântica* é a responsável por associar um significado a cada classe resultante da aplicação do SOM, de modo a buscar formas retangulares apenas na classe de interesse.

A Figura 3.7 mostra um exemplo de classificação e rotulação semântica de uma cena urbana usando o algoritmo SOM. Em todos os exemplos e experimentos apresentados neste trabalho, foi utilizado um mapa unidimensional, com o número de pesos igual ao número de classes de interesse. O conjunto de atributos de entrada é composto pelas médias espectrais de cada polígono e dos atributos de compacidade e dimensão fractal.

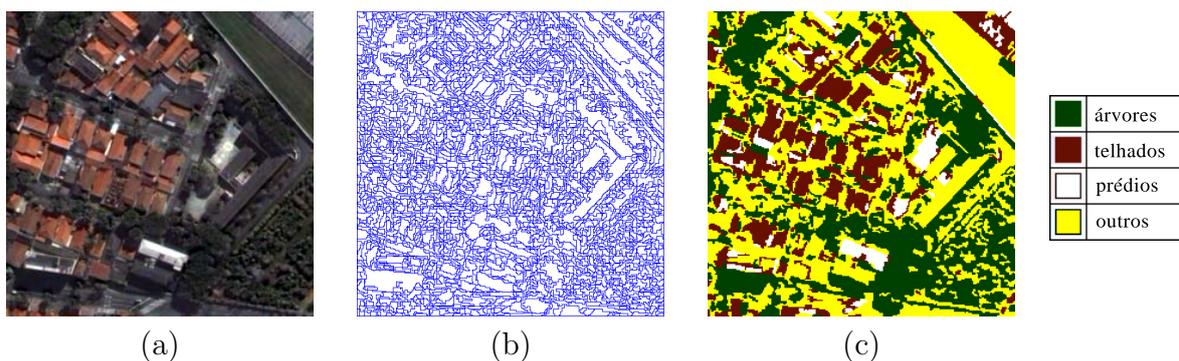


Figura 3.7 - Exemplo de classificação utilizando SOM: a) imagem de entrada, b) imagem super-segmentada e c) resultado da classificação.

3.3 Rotulação Semântica

O presente módulo recebe como entrada um conjunto de referência, contendo amostras das classes de interesse e seus atributos. Além disso, esse módulo recebe o resultado da *Classificação Não Supervisionada*, formado por rótulos atribuídos a cada elemento. A rotulação semântica então associa cada rótulo a um significado, *i.e.* uma classe. Essa classe será utilizada na construção do grafo e na busca por formas retangulares.

3.4 Geração do RAG (*Region Adjacency Graph*)

RAG é uma estrutura de grafo que representa a visão espacial de uma imagem. Ele associa um vértice a cada região e uma aresta a cada par de regiões adjacentes (TREMEAU; COLANTONI, 2000).

Seja I uma imagem e M um grupo de regiões denominadas $R_i, i = 1, \dots, M$ que satisfaça a propriedade básica da segmentação $\bigcup_{i=1}^M R_i = I$. Um grafo é representado pela notação $G = \langle V, A \rangle$, onde $V = \{1, \dots, M\}$ é o conjunto de vértices e $A \subset V \times V$ é o conjunto de arestas, ou conexões entre as regiões adjacentes. Pela notação padrão, cada região R_i é representada por um vértice, ou seja, $R_i = V_i, i = 1, \dots, M$. O valor das arestas é um tipo de medida de similaridade definida de acordo com a aplicação. Esta medida pode ser, por exemplo, a diferença entre as médias das regiões, ou a relação entre suas áreas.

Para ilustrar a construção de um RAG, a Tabela 3.1 mostra os pesos de um grafo gerado a partir da Figura 3.6. Neste exemplo, os pesos são definidos pelas diferenças entre as médias das regiões conectadas, e o valor -1 significa que não há adjacência entre as regiões.

Tabela 3.1 - Pesos do grafo gerado a partir da Figura 3.6.

	1	2	3	4	5
1	0	20	30	30	-1
2	20	0	40	-1	20
3	30	40	0	60	30
4	30	-1	60	0	30
5	-1	20	30	30	0

Neste trabalho, o RAG construído apresenta 3 possíveis valores para relacionar os segmentos:

- -1 para segmentos não adjacentes
- 0 para segmentos adjacentes de classes diferentes
- 1 para segmentos adjacentes de mesma classe

Alguns métodos têm utilizado a estrutura de grafos no processo de re-segmentação. Lezot et al. (2003) sugerem um algoritmo que, iterativamente, suaviza a informação espectral em cada região, e a seguir une os segmentos semelhantes. O processo continua até que algum critério de parada seja satisfeito, como por exemplo o número de iterações.

3.5 União dos Polígonos

Esta etapa produz o resultado final da re-segmentação. O algoritmo tem como entrada um conjunto de polígonos que pertencem a um número reduzido de classes. O único atributo de cada polígono é agora sua classe. As regiões serão unidas de acordo com os critérios estabelecidos *a priori* e com as heurísticas de busca por formas retangulares.

Para realizar a união, ou não, dos polígonos, um estágio de pré-processamento é necessário. Posteriormente, a busca no RAG por formas retangulares é executada. Visando aprofundar a transição do módulo da *Geração do RAG* para a *União dos Polígonos*, a Figura 3.8 apresenta um fluxograma das etapas envolvidas.

3.5.1 Pré-processamento dos Polígonos

A fase de pré-processamento é necessária pois em alguns casos as regiões não são corretamente classificadas, o que pode prejudicar a re-segmentação. A Figura 3.9a mostra um exemplo de classificação com problemas; os polígonos classificados de forma errada estão destacados por retângulos. Neste exemplo, um polígono isolado está associado a uma classe errada e é cercado por outros polígonos identificados por uma classe diferente. A estratégia utilizada nesse caso é a mudança da classe do polígono em questão para a classe da vizinhança, como mostra a Figura 3.9b.

Outra etapa do pré-processamento envolve a eliminação de polígonos cercados apenas por vizinhos de mesma classe. Quando uma região desse tipo é encontrada, ela é unida a alguma região vizinha que não tenha esta mesma propriedade. Isso reduz o número de polígonos a serem processados posteriormente, e faz com que cada elemento do conjunto

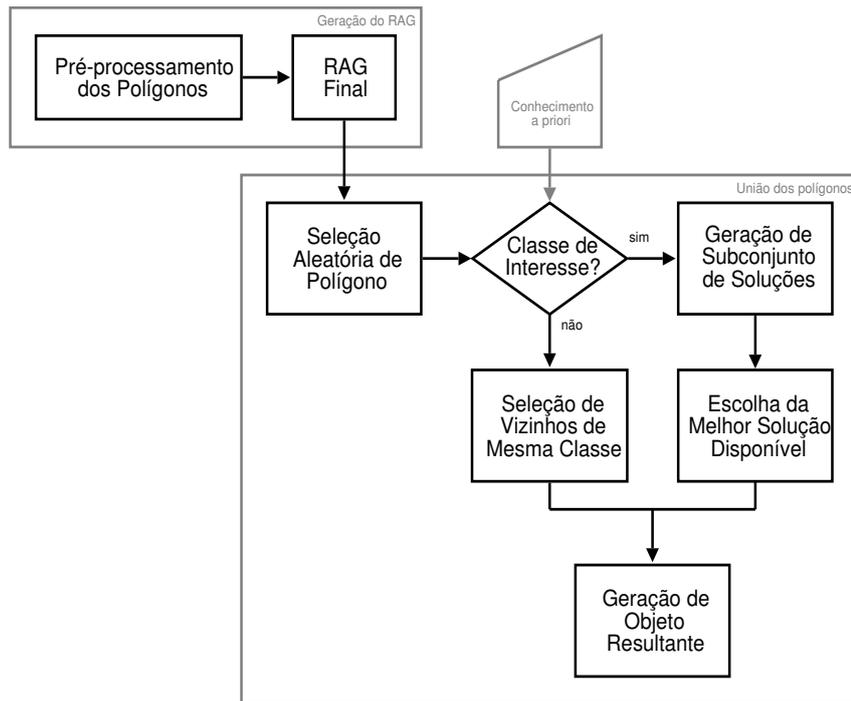


Figura 3.8 - Etapas envolvidas na união dos polígonos.

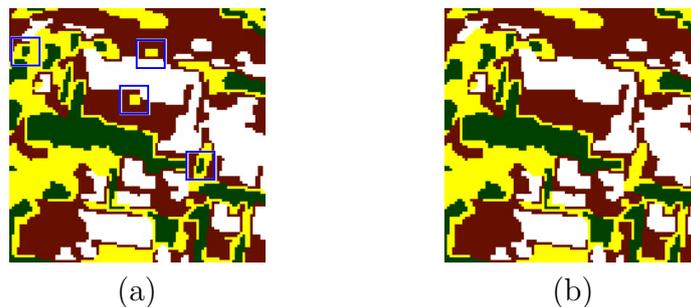


Figura 3.9 - Exemplo de erro na classificação: a) regiões com erro em destaque e b) regiões após o pré-processamento.

de regiões sempre tenha um vizinho de classe diferente. A Figura 3.10 exemplifica esta fase do pré-processamento.

3.5.2 Procura por Formas Retangulares

Dados os polígonos resultantes da etapa anterior, os nós do grafo são percorridos para encontrar os objetos de interesse. Entretanto, alguns grupos de polígonos que deveriam ser unidos para gerar uma região retangular, ainda pertencem a classes distintas. Este problema é ilustrado na Figura 3.11. Desta forma, o algoritmo deve ser capaz de unir regiões cujo agrupamento gere uma nova região com alto grau de retangularidade, levando em conta que alguns polígonos adjacentes pertencem a diferentes classes. A estratégia

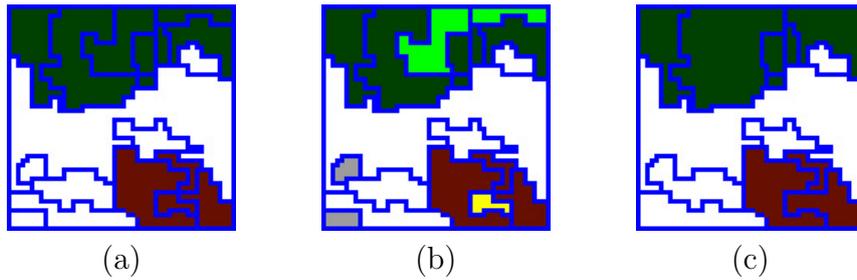


Figura 3.10 - Pré-processamento de polígonos vizinhos de mesma classe: a) regiões de entrada, b) polígonos em destaque e c) resultado do pré-processamento.

consiste em selecionar um grupo de regiões pertencentes a uma mesma classe e além dessas, regiões adjacentes de outras classes. Considerando que gerar e testar todas as combinações possíveis aumenta o custo computacional e inviabiliza o processamento, um subconjunto de soluções é procurado através da inserção e remoção de elementos desse conjunto. Um alto valor do atributo *RET* é um indicativo de uma solução encontrada.



Figura 3.11 - Separação de telhados em classes diferentes: a) imagem original com telhados em destaque e b) telhados separados em classes diferentes.

Durante o processo de união dos segmentos, o grau de retangularidade *RET* é testado, e o maior valor encontrado indica o conjunto de regiões candidato à re-segmentação. Quando o grau de retangularidade for superior a um limiar definido pelo usuário (θ), os segmentos são agregados e formam um novo objeto. As regiões restantes são novamente testadas por outras combinações de polígonos daquela vizinhança. O algoritmo desta estratégia é descrito a seguir.

```

Escolha aleatória de um nó  $v$  (não visitado)
Crie o conjunto  $M$  com  $v$  e vizinhos de mesma classe
Crie o conjunto  $D$  com demais vizinhos
Calcule a retangularidade  $RET$  no conjunto  $M$ 
Repita o procedimento com cada elemento  $D_i$ 
|   Insira o elemento  $D_i$  no conjunto  $M$ 
|   Calcule a retangularidade  $RET_i$  no conjunto  $M$ 
|   Se  $RET_i < RET$ 
|   |   Remova o elemento  $D_i$  do conjunto  $M$ 
|   Senão
|   |   Mantenha o elemento  $D_i$  no conjunto  $M$ 
|   |    $RET \leftarrow RET_i$ 
|   Fim Se
Fim Repita
Se  $RET > \theta$ 
|   Marque como visitados os elementos de  $M$ 
|   Gere um polígono com a união dos elementos de  $M$ 
Fim Se

```

4 RESULTADOS

Este capítulo apresenta os resultados dos experimentos realizados para avaliar o método descrito no Capítulo 3. A área de estudo compreende recortes de imagens *QUICKBIRD* da Cidade de São José dos Campos, adquiridas em Dezembro de 2005, com resolução espacial de $1m$. As imagens mostradas na Figura 4.1 são usadas como referência para avaliação do algoritmo. Os polígonos de entrada do algoritmo de re-segmentação foram obtidos através do algoritmo de Crescimento de Regiões (BINS et al., 1996), implementado na biblioteca TerraLib (CAMARA et al., 2000) e com avaliação visual positiva e superior a diversas outras estratégias, segundo Meinel e Neubert (2002). Os limiares de similaridade e área mínima (d, a) foram ajustados para gerar uma imagem super-segmentada ($d = 10, a = 10$). A Figura 4.2 ilustra a super-segmentação das áreas destacadas por retângulos na Figura 4.1.

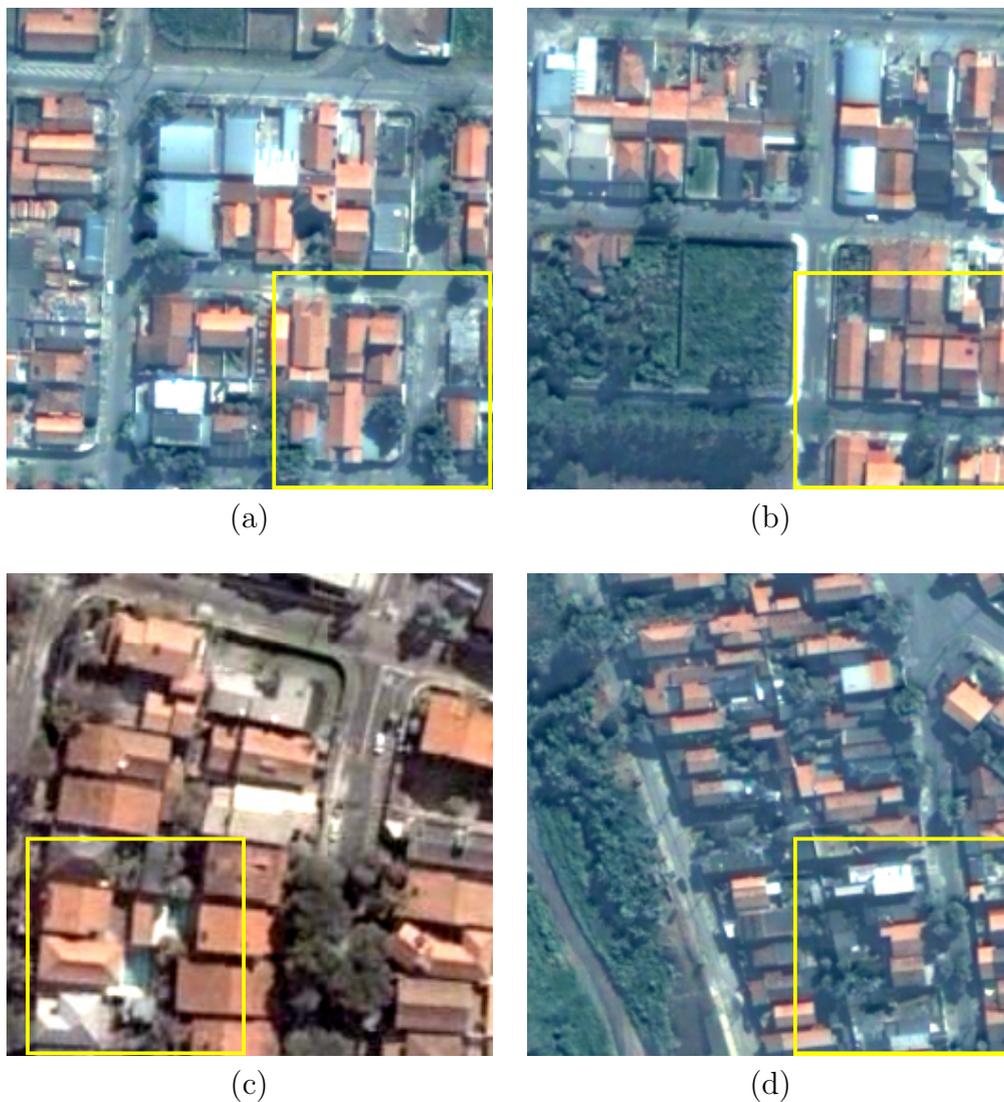


Figura 4.1 - Recortes da área de estudo: a) primeiro experimento, b) segundo experimento, c) terceiro experimento e d) quarto experimento.

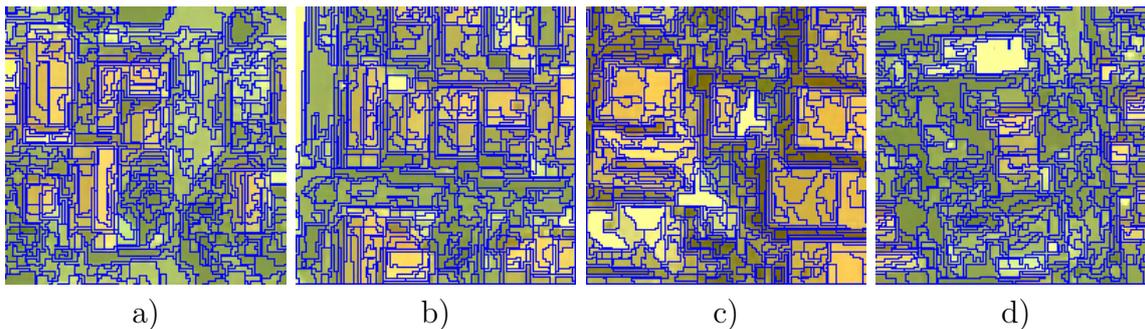


Figura 4.2 - Detalhe da super-segmentação para as áreas destacadas da Figura 4.1.

A Seção 4.1 apresenta os resultados da re-segmentação destes recortes e a Seção 4.2 discute os resultados gerados. Finalmente, os detalhes da estrutura de dados e o sistema computacional são apresentados na Seção 4.3.

4.1 Experimentos Realizados

Quatro experimentos foram realizados neste trabalho. Foi dada uma atenção especial à identificação de objetos retangulares, tais como telhados. Para ilustrar os resultados, cada experimento mostra a imagem com algumas formas em destaque, os polígonos super-segmentados sobrepostos à cena em questão, a classificação da imagem pelo algoritmo SOM e, finalmente, os polígonos resultantes da re-segmentação sobrepostos à imagem de entrada. Os resultados são avaliados em termos da taxa de acertos e erros e tempo de execução. Todos os experimentos foram realizados em uma configuração *AMD Athlontm 3000+*, com *512MB* de memória RAM, no sistema operacional *Linux Mandriva 2006*.

4.1.1 Primeiro Experimento

O primeiro experimento apresenta uma área urbana que pode ser classificada em 5 classes: árvores, telhados, prédios, ruas e outros. Telhados e prédios podem ser representados por formas retangulares, conforme destaca a Figura 4.3a. Alguns dos problemas da classificação puderam ser contornados através da técnica de pré-processamento dos polígonos, proposta na Subseção 3.5.1. Após as etapas de re-segmentação sobre a entrada, formada por 2042 polígonos e 3 bandas espectrais, foram gerados 578 polígonos. O tempo de execução foi de 258 segundos.

Na Figura 4.3d é possível observar que alguns objetos não foram re-segmentados corretamente. Parte deste problema deve-se aos erros obtidos na super-segmentação, que uniu objetos distintos por apresentarem respostas espectrais semelhantes. Esses polígonos de entrada acabaram comprometendo as fases de processamento posteriores. Em outros polígonos com resultados não satisfatórios, pode-se dizer que o subconjunto de soluções

escolhido não contemplou a solução ideal pois, conforme apresentado na Subseção 3.5.2, nem todas as soluções são testadas para não comprometer o custo computacional.

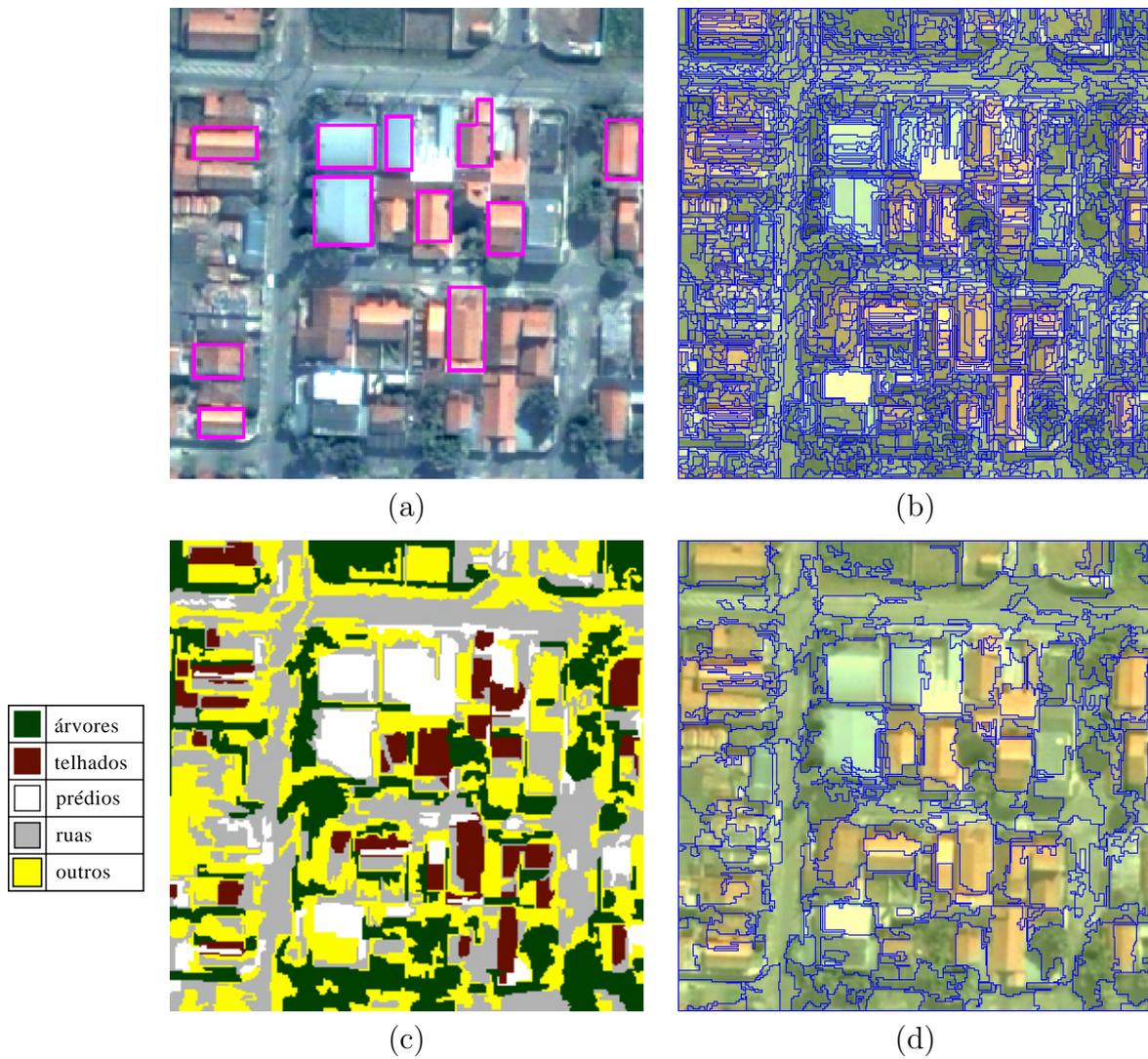


Figura 4.3 - Primeiro experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação.

4.1.2 Segundo Experimento

Este experimento testa uma imagem de área urbana, que contém diversos telhados com formas retangulares. Além disso, há uma área verde que ocupa, aproximadamente, 25% da imagem, e algumas construções que, embora sejam retangulares, não pertencem à mesma classe dos telhados (Figuras 4.4a e 4.4c). O algoritmo recebeu como entrada 1998 polígonos e 3 bandas espectrais, resultando em 634 polígonos. O resultado da re-segmentação é mostrado na Figura 4.4d. As formas retangulares foram identificadas e extraídas corretamente. Além disso, deve-se mencionar que regiões tipicamente irregulares (*e.g.* áreas verdes) foram re-segmentadas de maneira visualmente satisfatória, gerando regiões com maior homogeneidade semântica do que espectral. O tempo de execução foi de 317 segundos.

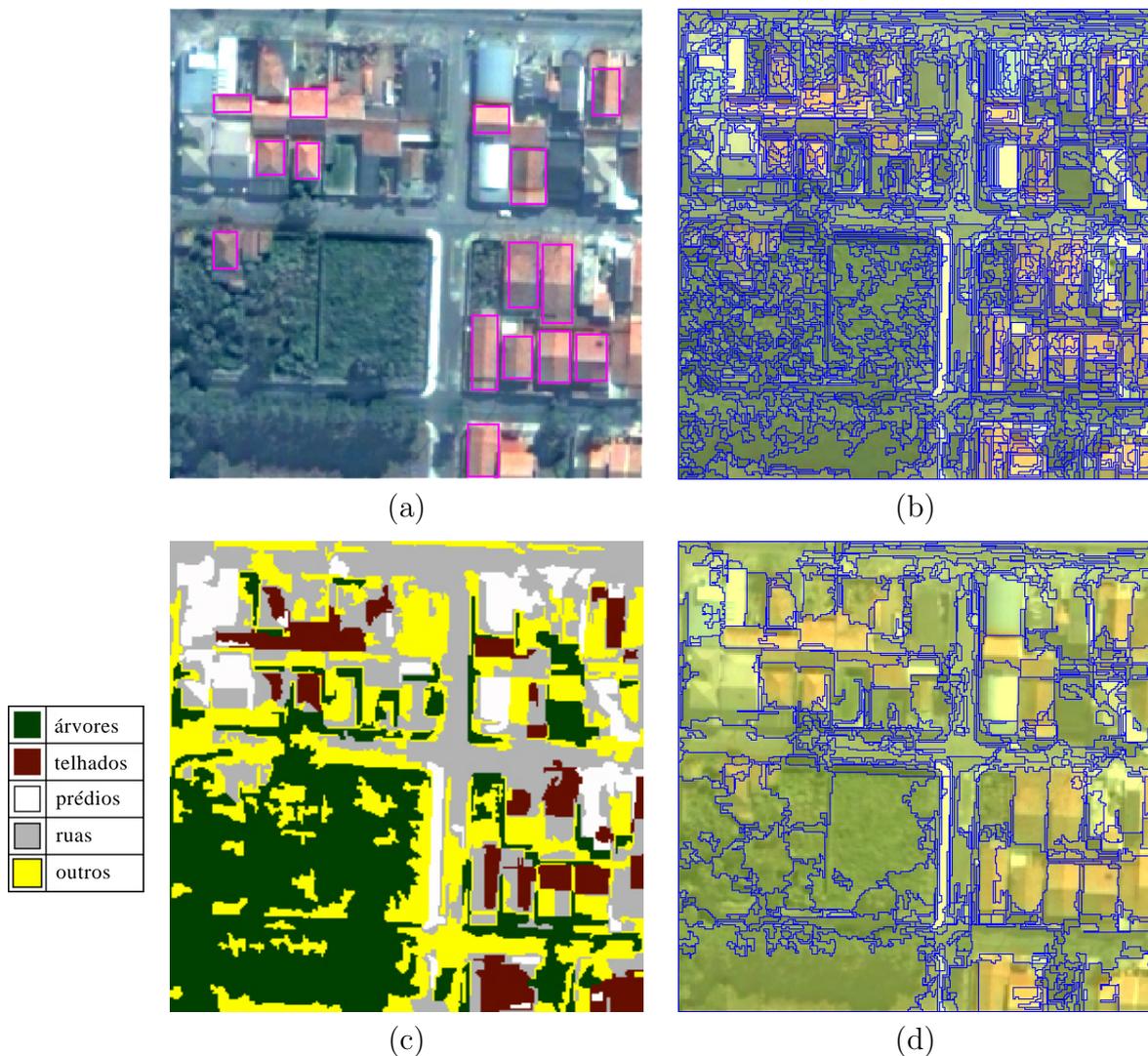


Figura 4.4 - Segundo experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação.

4.1.3 Terceiro Experimento

O terceiro experimento apresenta uma região com telhados, ruas e árvores em destaque. Pode-se observar, na imagem da Figura 4.5a, a presença de telhados não retangulares, rotacionados em relação aos eixos principais da figura. Considerando objetos com formas não retangulares, as fases de pré-classificação e pré-processamento são fundamentais para que se obtenha resultados corretos. Após a re-segmentação, 750 polígonos foram gerados, como mostra a Figura 4.5d. O algoritmo processou os 2264 polígonos de entrada em 196 segundos.

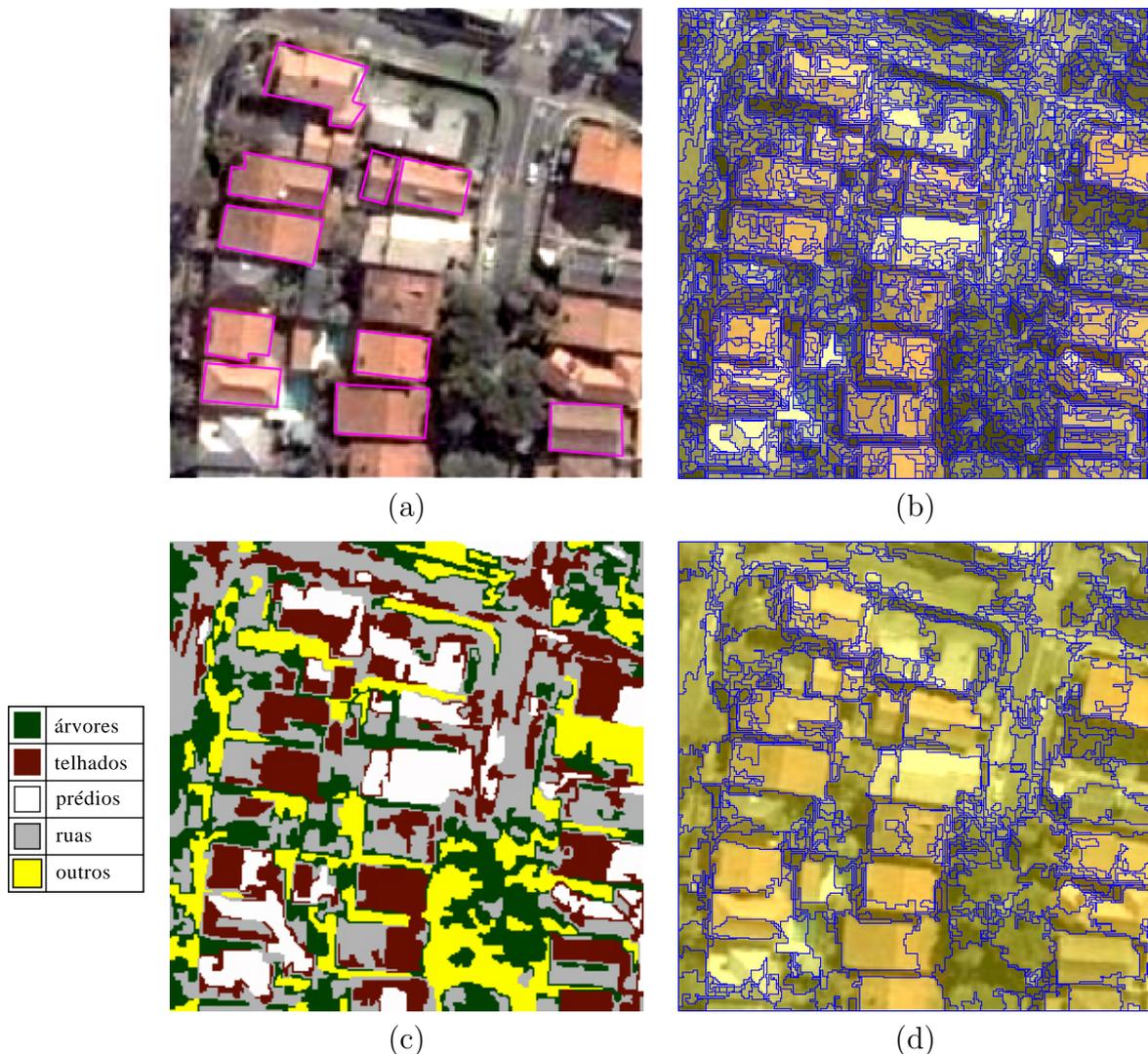


Figura 4.5 - Terceiro experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação.

4.1.4 Quarto Experimento

O último experimento testou uma imagem composta por diversos telhados, com áreas inferiores aos encontrados nos experimentos anteriores, além de árvores, ruas, prédios e algumas sombras, como pode ser visto na Figura 4.6a. Um conjunto de 2028 polígonos e 3 bandas espectrais foram processados em 243 segundos, e 695 polígonos de saída foram obtidos. Esse resultado é mostrado na Figura 4.6d.

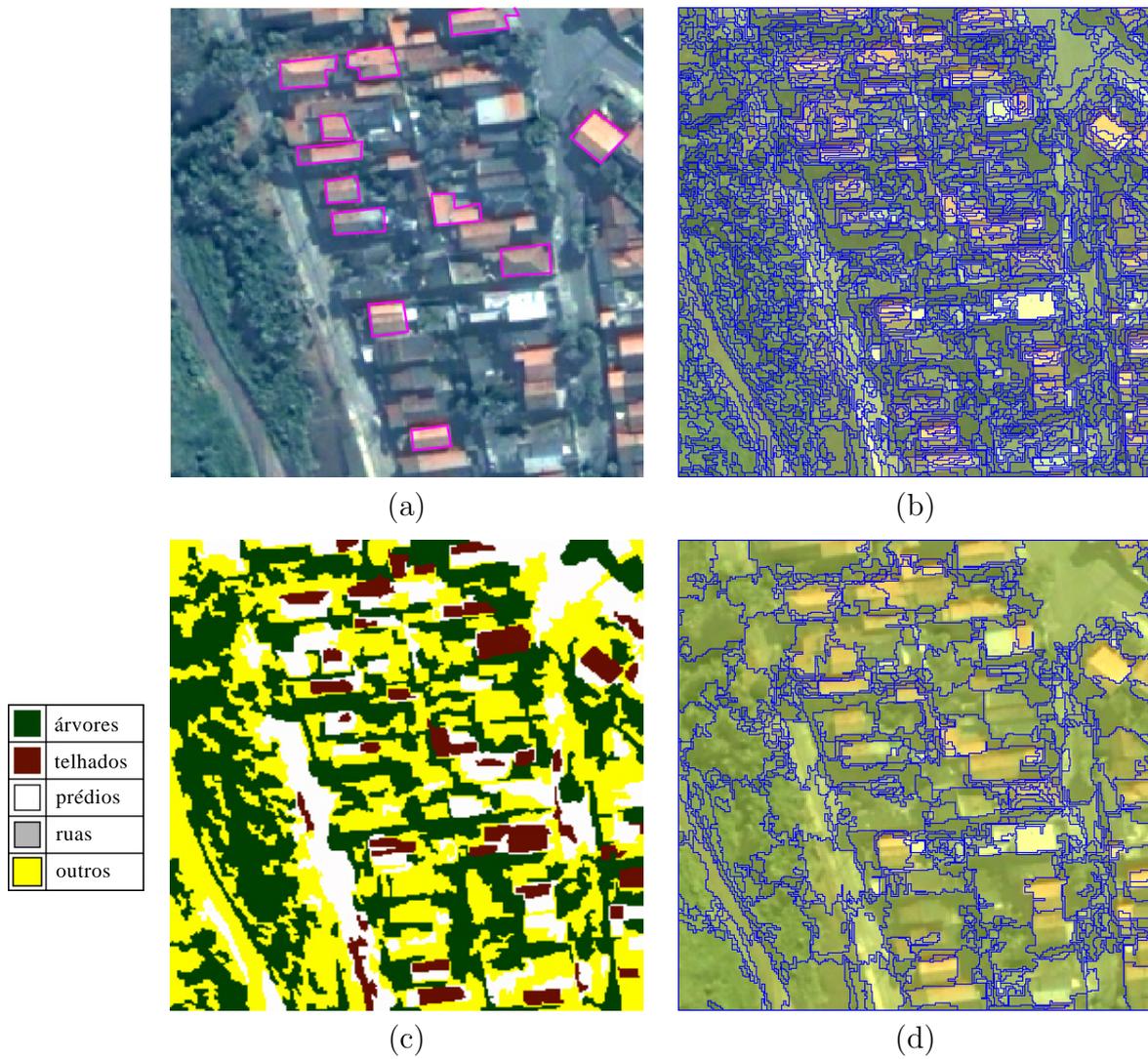


Figura 4.6 - Quarto experimento: a) imagem original com algumas formas em destaque, b) super-segmentação de entrada, c) classificação resultante e d) resultado da re-segmentação.

4.2 Discussão Sobre os Resultados

Esta seção apresenta a avaliação dos resultados da re-segmentação obtidos nos experimentos anteriores. A avaliação do atributo de retangularidade é também realizada e comparada a outras abordagens citadas no Capítulo 2.

4.2.1 Avaliação dos Experimentos

As imagens segmentadas utilizadas como referência na avaliação do método foram geradas pelo segmentador tradicional por crescimento de regiões, com limiares de similaridade ($d = 40$) e área mínima ($a = 40$), ajustados para gerar resultados coerentes com a aplicação. A Figura 4.7 mostra os resultados desta segmentação para os quatro experimentos.

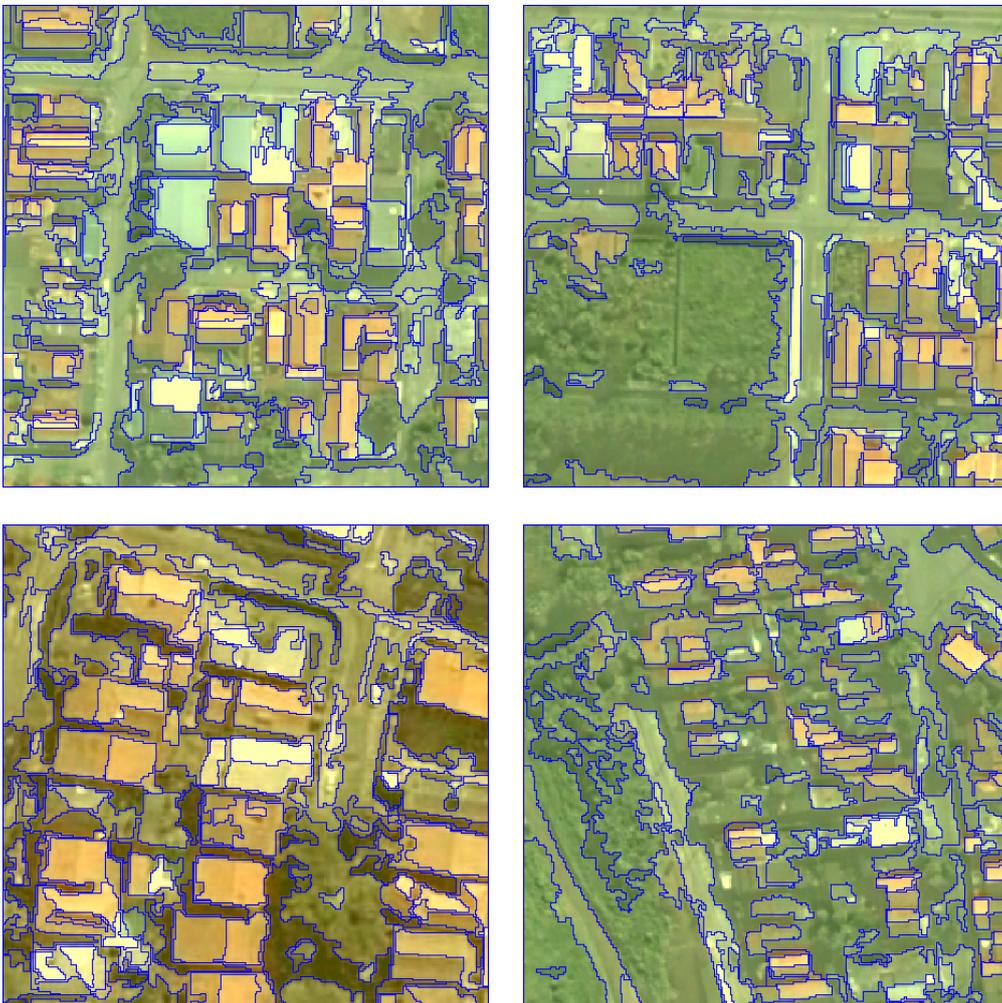


Figura 4.7 - Segmentação tradicional para os recortes da área de estudo.

A Tabela 4.1 apresenta os valores das medidas *QUANT* e *RMSE* (Seção 2.3) calculados para os experimentos, tanto para a re-segmentação proposta quanto para o método de crescimento de regiões tradicional. Para simplificar os cálculos, foi escolhido um subconjunto de polígonos contendo algumas regiões de interesse (formas retangulares). Polígonos com ruído e sombras, além dos localizados nas bordas da imagem foram desconsiderados.

Tabela 4.1 - Resultados da avaliação qualitativa dos experimentos, comparados com uma segmentação tradicional.

Experimentos	Re-Segmentação		Tradicional	
	<i>QUANT</i>	<i>RMSE</i>	<i>QUANT</i>	<i>RMSE</i>
Primeiro	0,834	0,251 <i>pixels</i>	0,416	0,715 <i>pixels</i>
Segundo	0,625	0,354 <i>pixels</i>	0,385	0,631 <i>pixels</i>
Terceiro	0,769	0,574 <i>pixels</i>	0,455	1,357 <i>pixels</i>
Quarto	0,667	0,215 <i>pixels</i>	0,527	0,446 <i>pixels</i>

A média dos coeficientes *QUANT* para os quatro experimentos foi de 0,724. Os resultados obtidos pelo método de crescimento de regiões tradicional obtiveram uma média de *QUANT* = 0,446. O erro médio quadrático das imagens re-segmentadas foi inferior ao da segmentação tradicional em todos os experimentos, como pode ser observado na tabela.

4.3 Interface do Sistema

O algoritmo de re-segmentação proposto foi implementado na biblioteca TerraLib e para a sua execução foi desenvolvido um *plugin* para o sistema TerraPixel, desenvolvido sob a mesma plataforma. Um *plugin* funciona como um aplicativo, rodando internamente ao ambiente principal e compartilhando elementos manipulados por ele (conjuntos de imagens e de polígonos). A Figura 4.8 apresenta a interface do sistema TerraPixel, onde é possível visualizar o conjunto de polígonos e as imagens (*Active Polygons/Rasters*).

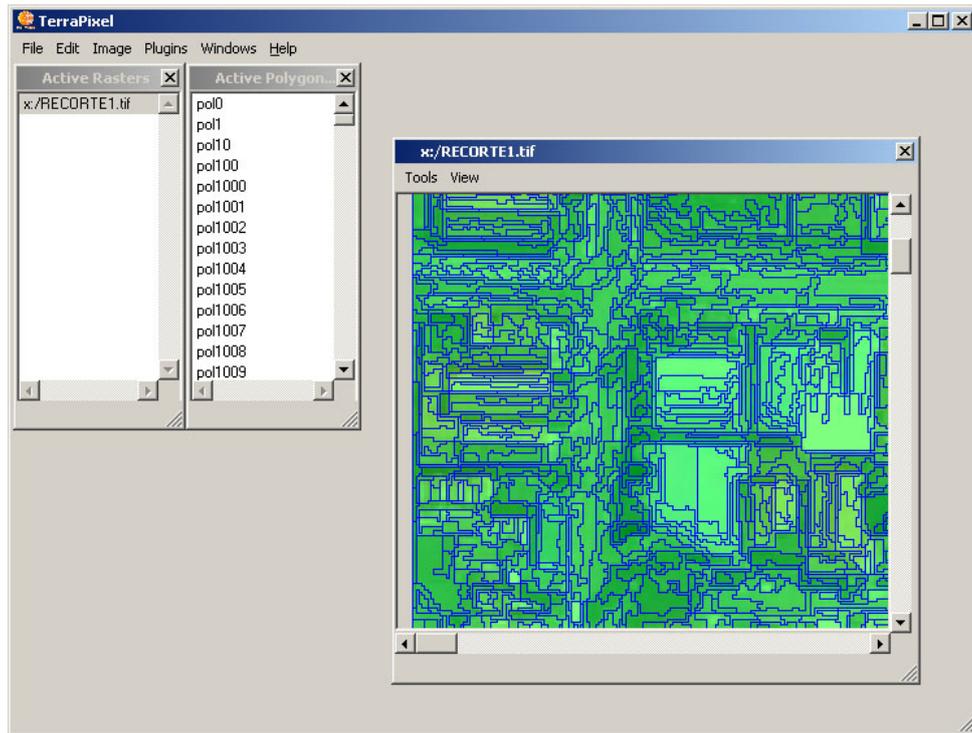


Figura 4.8 - Sistema TerraPixel – Interface padrão.

4.3.1 *Plugin* de Re-Segmentação

O algoritmo proposto neste trabalho tem como dados de entrada os seguintes parâmetros:

- Limiar de retangularidade: A busca no grafo objetiva encontrar polígonos com o maior grau de retangularidade possível. Esse limiar corresponde ao valor acima do qual a união de um conjunto de polígonos define um novo segmento de saída;
- Número de classes: A classificação SOM recebe como entrada um número de neurônios, que nesta implementação corresponde ao número de classes do sistema. Tais classes definem os polígonos de entrada, conforme a etapa de classificação dos segmentos, apresentada na Seção 3.2;
- Imagens: Um ou mais canais podem ser utilizados como parâmetros de entrada. Os valores de níveis de cinza dos *pixels* são utilizados para o cálculo dos atributos espectrais, conforme descrição na Seção 3.1;
- Polígonos: Os polígonos são obtidos pelo processo de super-segmentação. A partir deste conjunto de polígonos, os atributos espectrais e de forma são extraídos, e a união de seus subconjuntos produz os segmentos de saída.

A Figura 4.9 ilustra a interface do *plugin* com o grupo de parâmetros descrito.

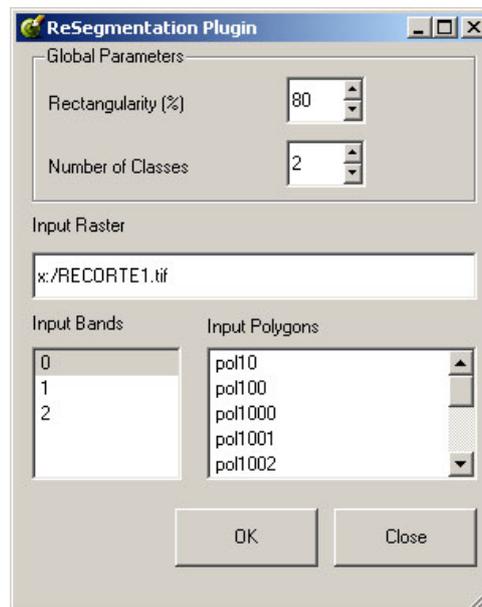


Figura 4.9 - *Plugin* para o TerraPixel.

4.3.2 Detalhes da Implementação

Na implementação do método, três classes foram definidas: `TePDIAttributes`, `TePDISOMClas` e `TePDIReSegmentation`, e implementam os módulos de extração de atributos, classificação utilizando SOM e de re-segmentação. Todas foram integradas ao módulo de PDI da TerraLib. O diagrama destas classes, contendo a interface e o conjunto de atributos, é mostrado na Figura 4.10, para facilitar o entendimento de sua estrutura e programação.

TePDIAttributes <i>Spectral and Shape Attributes</i>
<p>-r_ : TeRaster <i>Imagem de entrada</i></p> <p>-bands_ : unsigned <i>Número de bandas</i></p> <p>-W_ : unsigned <i>Largura</i></p> <p>-H_ : unsigned <i>Altura</i></p> <p>-p_ : TePolygon <i>Polígono de entrada</i></p> <p>-centroid_ : TeCoord2D <i>Centróide do Polígono</i></p> <p>-x_ : vector<unsigned> <i>Vetor de posições no eixo x</i></p> <p>-y_ : vector<unsigned> <i>Vetor de posições no eixo y</i></p> <p>-position_covariance_ : TeMatrix <i>Covariância dos vetores de posições</i></p> <p>-angle_ : double <i>Ângulo principal</i></p> <p>-new_p_ : TePolygon <i>Polígono rotacionado</i></p> <p>-N_ : unsigned <i>Número de pixels</i></p>
<p>+TePDIAttributes(p:TePolygon,r:TeRaster)</p> <p>+~TePDIAttributes()</p> <p>+get_angle(deg:bool)</p> <p>+get_area()</p> <p>+get_box_area()</p> <p>+get_new_area()</p> <p>+get_new_box_area()</p> <p>+get_perimeter()</p> <p>+get_frac()</p> <p>+get_compacity()</p> <p>+get_rectangularity()</p> <p>+get_mean(b:unsigned)</p> <p>+get_covariance()</p> <p>+get_distance(a:TePDIAttributes)</p> <p>+rotate_polygon(a:double)</p> <p>+calculate_angle(p:TePolygon)</p>

TePDISOMClas <i>Self Organizing Map</i>
<p>-X_ : double** <i>Vetor de atributos de entrada</i></p> <p>-weights_ : struct vector <i>Vetor de pesos</i></p> <p>-w_size_ : unsigned <i>Número de pesos</i></p> <p>-N_ : unsigned <i>Número de atributos</i></p> <p>-max_epochs_ : unsigned <i>Máximo de épocas</i></p> <p>-radius_ : double <i>Raio de influência</i></p> <p>-decreasing_ : double <i>Redução da taxa de aprendizado</i></p> <p>-alpha_ : double <i>Taxa de aprendizado</i></p>
<p>+TePDISOMClas()</p> <p>+~TePDISOMClas()</p> <p>+Init(params:TePDIParameters)</p> <p>+train()</p> <p>+classify(tmp_input:vector<double>)</p> <p>-h_gauss(a[]:double,b[]:double,r:double)</p> <p>+get_distance(a[]:double,b[]:double)</p>

TePDIReSegmentation <i>Re-Segmentation</i>
<p>-threshold_ : double <i>Limiar de retangularidade</i></p> <p>-n_classes_ : unsigned <i>Número de classes</i></p> <p>-RAG: TeMatrix <i>Region Adjacency Graph</i></p>
<p>-visit(p1:unsigned,level:unsigned)</p> <p>+TePDIReSegmentation()</p> <p>+~TePDIReSegmentation()</p> <p>+CheckParameters(params:TePDIParameters)</p> <p>#ResetState(params:TePDIParameters)</p> <p>#RunImplementation()</p>

Figura 4.10 - Classes implementadas.

5 CONCLUSÕES

Devido à complexidade do problema de segmentação de imagens, não existe um único método que seja adequado para todas as aplicações e tipos de dados. O que se tentou neste trabalho, foi desenvolver um método de re-segmentação com o objetivo de melhorar os resultados de uma segmentação prévia. O método proposto usa atributos espectrais e de forma, e informação contextual através da vizinhança dos segmentos. Para direcionar o método, foi proposto resolver o problema de segmentação de objetos urbanos em imagens de alta resolução. O foco principal consistiu em extrair objetos urbanos de forma retangular.

O algoritmo de re-segmentação foi implementado em C++, dentro do ambiente da biblioteca TerraLib. Para gerar os resultados, um sistema computacional baseado em um *plugin* para o sistema TerraPixel foi desenvolvido e está disponível livremente. A implementação do método foi realizada, basicamente, em três módulos principais: extração de atributos, classificação via SOM e re-segmentação. O método recebe como entrada imagens multi-espectrais e um conjunto de polígonos, dos quais são extraídos os atributos espectrais e de forma. Para a busca dos polígonos retangulares foi usada a estrutura de grafos. Através de um grafo pode-se realizar buscas em um nível de vizinhança pré-definido, influenciando os resultados e o desempenho do algoritmo. O emprego do RAG na metodologia acompanhou a natureza dos dados de entrada, nos quais regiões possuem relações topológicas. Através disso, a etapa de *União dos Polígonos* foi eficaz na geração dos resultados.

Na comparação entre os segmentos de referência e os gerados pela re-segmentação, através de uma avaliação visual e análise qualitativa, pode-se observar bons resultados. Os índices *QUANT* e *RMSE* foram calculados, obtendo valores de até 0,834 e 0,215 respectivamente. Foram também realizadas segmentações utilizando uma metodologia tradicional baseada em crescimento de regiões, e a avaliação apontou a superioridade do algoritmo de re-segmentação proposto.

Alguns resultados apresentaram erros em determinadas regiões, evidenciando uma desvantagem da metodologia, relacionada à escolha de um subconjunto das possíveis soluções. Embora tal subconjunto solucione o problema do desempenho computacional, exclui hipóteses durante a busca por formas retangulares. Como a escolha deste subconjunto é aleatória, em alguns casos o sistema não obteve resultados satisfatórios. Outra fonte de erros nos resultados se deu pela super-segmentação de entrada, que em determinadas ocasiões apresentou segmentos pertencentes a objetos distintos, pelo fato de serem espectralmente próximos.

5.1 Recomendações

A implementação envolve várias etapas de processamento. Assim, modificações podem ser realizadas para tornar o algoritmo mais preciso e eficiente em termos computacionais. A pesquisa no grafo apresenta a ordem de complexidade $O(n^2)$. Uma possível solução seria gerar uma árvore de custo mínimo, e realizar nela a pesquisa sem recursividade, reduzindo a complexidade para $O(n)$. No entanto, a geração desta árvore envolve também um estudo sobre qual atributo será utilizado para conectar o grafo inicial.

O uso de classificação supervisionada na re-segmentação é possível. Porém, como já mencionado anteriormente, isso tornaria o processo mais interativo e portanto, menos automático. Nessa estratégia, um novo módulo de treinamento e rotulação manual dos resultados deve ser implementado. Dessa maneira, pode-se incorporar conhecimento *a priori* das classes presentes na imagem. Em sistemas computacionais, é desejável que o número de parâmetros de entrada seja baixo para facilitar seu uso. No entanto, é importante realizar uma comparação dos resultados obtidos pela re-segmentação através de um classificador supervisionado.

Como a entrada do algoritmo é composta de um conjunto de polígonos resultantes da super-segmentação, um número muito grande de elementos é processado a cada etapa. Uma recomendação para otimizar o processamento é a de trabalhar com subconjuntos de polígonos, que compreendam a vizinhança do elemento em questão. Isso poderia ser obtido com uma janela centralizada em um polígono, que apresente um raio pré-estabelecido, de forma a compreender apenas elementos internos a esta janela. A biblioteca TerraLib contém o suporte a bancos de dados com extensão geográfica, nos quais tal técnica já está implementada. Entretanto, não é recomendável, em um primeiro momento, associar o sistema proposto a estruturas de bancos de dados, visto que o sistema perderia sua característica operacional, agregando novas etapas e tecnologias para sua instalação.

O aplicativo foi desenvolvido para processar imagens urbanas, buscando formas retangulares, tipicamente encontradas nesse ambiente. No entanto, para generalizar seu funcionamento, é recomendável a inclusão de novos atributos de forma. Usando estes atributos, novas buscas por padrões que se encaixem na mesma estrutura de grafos podem ser realizadas. Esses padrões podem ser passados via parâmetro pelo usuário, ou incluírem novas formas básicas, como círculos, triângulos, ou ainda padrões espectrais, como texturas. Com relação ao conhecimento *a priori*, pode-se adaptar o sistema para receber dados de mapas temáticos. Essa informação auxilia na etapa de pré-classificação e também na re-segmentação propriamente dita.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANTUNES, A. F. B.; CORTESE, S. C. Determinação de áreas de ocupação irregulares por meio de análise orientada a objetos. In: INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Anais XIII Simpósio Brasileiro de Sensoriamento Remoto**. São José dos Campos: INPE, 2007. p. 481–487. Disponível em: <<http://dpi.inpe.br/sbsr@80/2006/10.31.14.10>>. 22, 34
- BAATZ, M.; SCHÄPE, A. Object-Oriented and Multi-Scale Image Analysis in Semantic Networks. **2nd International Symposium on Operationalisation of Remote Sensing Conference and Exhibition**, Enschede, ITC, v. 1, p. 16–20, 1999. 22, 31
- _____. Multiresolution Segmentation—an optimization approach for high quality multi-scale image segmentation. In: **Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium**. Salzburg: [s.n.], 2000. v. 200, p. 12–23. 29, 32, 33, 34, 37
- BAÇÃO, F.; LOBO, V. **Introduction to Kohonen’s self-organizing maps**. ISEGI, Portugal, 2007. 41
- BENZ, U.; BAATZ, M.; SCHREIER, G. OSCAR-object oriented segmentation and classification of advanced radar allow automated information extraction. **Geoscience and Remote Sensing Symposium**, v. 4, p. 1913–1915, July 2001. 31, 38
- BINS, L. S.; FONSECA, L. M. G.; ERTHAL, G. J.; II, F. M. Satellite imagery segmentation: a region growing approach. In: INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Anais**. São José dos Campos: INPE, 1996. p. 677–680. ISBN 85-17-00014-5. Disponível em: <<http://urlib.net/sid.inpe.br/deise/1999/02.05.09.30>>. 29, 49
- BLASCHKE, T. Object-based contextual image classification built on image segmentation. **IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data**, p. 113–119, 27-28 Oct. 2003. 31, 33
- CAMARA, G.; SOUZA, R. C. M.; PEDROSA, B. M.; VINHAS, L.; MONTEIRO, A. M. V.; PAIVA, J. A. d. C.; CARVALHO, M. T. d.; GATTASS, M. Terralib: technology in support of gis innovation. In: **Anais**. Sao Paulo: [s.n.], 2000. 23, 49
- CAMARA, G.; SOUZA, R. C. M. d.; FREITAS, U. M. d.; GARRIDO, J. Spring: Integrating remote sensing and gis by object-oriented data modelling. **Computer and Graphics**, v. 20, n. 3, p. 395–403, May - June 1996. ISSN 0097–8493. 23, 29

CANNY, J. A computational approach to edge detection. **IEEE Trans. Pattern Anal. Mach. Intell.**, IEEE Computer Society, Washington, DC, USA, v. 8, n. 6, p. 679–698, 1986. ISSN 0162-8828. 30

CARSON, C.; BELONGIE, S.; GREENSPAN, H.; MALIK, J. Blobworld: image segmentation using expectation-maximization and its application to image querying. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 24, n. 8, p. 1026–1038, 2002. 21

CHENG, H.; JIANG, X.; SUN, Y.; WANG, J. Color image segmentation: Advances and prospects. **Pattern Recognition**, v. 34, n. 12, p. 2259–2281, 2001. 27, 28, 29, 30

COMANICIU, D.; MEER, P. Robust analysis of feature spaces: color image segmentation. **Computer Vision and Pattern Recognition**, IEEE Computer Society, Los Alamitos, CA, USA, v. 00, p. 750, 1997. ISSN 1063-6919. 28, 33

FALCÃO, A. X.; STOLFI, J.; LOTUFO, R. A. The Image Foresting Transform: Theory, Algorithms, and Applications. **IEEE Trans. on Pattern Analysis and Machine Intelligence**, IEEE Press, v. 26, n. 1, p. 19–29, Jan 2004. 28

GAMANYA, R.; MAEYER, P. de; DAPPER, M. de. An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization. **Expert System with Applications**, v. 32, p. 616–624, 2007. 32, 33

GONZALEZ, R.; WOODS, R. **Processamento de imagens digitais**. São Paulo: Edgard Blucher, 2000. 21, 28, 30

GUO, D.; WEEKS, A.; KLEE, H. Robust approach for suburban road segmentation in high-resolution aerial images. **International Journal of Remote Sensing**, Taylor & Francis, v. 28, n. 2, p. 307–318, 2007. Disponível em: <http://www.informaworld.com/10.1080/01431160600721822>. 28

HAY, G.; CASTILLA, G.; WULDER, M.; RUIZ, J. An automated object-based approach for the multiscale image segmentation of forest scenes. **International journal of applied earth observation and geoinformation**, Elsevier, v. 7, n. 4, p. 339–359, 2005. 31, 32, 33

HEROLD, M.; SCEPAN, J.; MULLER, A.; GUNTHER, S. Object-oriented mapping and analysis of urban land use/cover using IKONOS data. **EARSEL Symposium Geoinformation for European-wide integration**, Prague, v. 22, 2002. 22

JENSEN, J. **Introductory digital image processing – a remote sensing perspective**. [S.l.]: Prentice Hall PTR Upper Saddle River, NJ, USA, 1995. 31

- JUNG, C. Unsupervised multiscale segmentation of color images. **Pattern Recognition Letters**, Elsevier Science Inc. New York, NY, USA, v. 28, n. 4, p. 523–533, 2007. 21, 29
- JUNG, C.; SCHRAMM, R. Rectangle detection based on a windowed Hough transform. In: **Brazilian Symposium on Computer Graphics and Image Processing**. Curitiba: [s.n.], 2004. v. 17, p. 113–120. 25
- KASKI, S. Data exploration using self-organizing maps. **Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series**, v. 82, p. 57, 1997. 41, 42
- KOHONEN, T. **Self-organizing maps**. 3. ed. Berlin: Springer, 2001. 25, 41
- LEZORAY, O.; ELMOATAZ, A.; SRC, I.; EA, L.; SAINT-LO, F. Graph based smoothing and segmentation of color images. In: **Seventh International Symposium on Signal Processing and Its Applications**. Paris: [s.n.], 2003. v. 1, n. 7. 45
- LIU, L.; SCLAROFF, S. Deformable model-guided region split and merge of image regions. **Image and Vision Computing**, v. 22, p. 343–354, 2004. 30
- LUCCHESI, L.; MITRA, S. Color image segmentation: A state-of-the-art survey. **Proc. of the Indian National Science Academy (INSA-A)**, v. 67, n. 2, p. 207–221, 2001. 30
- MARR, D.; HILDRETH, E. Theory of Edge Detection. **Proceedings of the Royal Society of London. Series B, Biological Sciences**, JSTOR, v. 207, n. 1167, p. 187–217, 1980. 30
- MCGARIGAL, K.; MARKS, B. **FRAGSTATS – Spatial pattern analysis program for quantifying landscape structure**. Portland: Department of Agriculture, Forest Service, Pacific Northwest Research Station, 1995. 25
- MEINEL, G.; NEUBERT, M. A Comparison of Segmentation Programs for High Resolution Remote Sensing Data. **development**, v. 4, p. 2002, 2002. 49
- MOREL, J.; SOLIMINI, S. Variational methods in image segmentation. **Progress in Nonlinear Differential Equations and their Applications**, v. 14, 1995. 29
- PAL, N.; PAL, S. A review on image segmentation techniques. **Pattern Recognition**, v. 26, n. 9, p. 1277–1294, 1993. 27
- PERONA, P.; MALIK, J. Scale-space and edge detection using anisotropic diffusion. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 12, n. 7, p. 629–639, 1990. 30

SEZGIN, M.; SANKUR, B. Survey over image thresholding techniques and quantitative performance evaluation. **Journal of Electronic Imaging**, SPIE, v. 13, p. 146, 2004. 28

SHI, J.; MALIK, J. Normalized cuts and Image segmentation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE Computer Society, Los Alamitos, CA, USA, v. 22, n. 8, p. 888–905, 2000. ISSN 0162-8828. 28

SOUSA, M. d. A. **Segmentação multi-níveis e multi-modelos para imagens de radar e ópticas**. 131 p. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2005–03–31 2005. Disponível em:

<<http://urlib.net/sid.inpe.br/jeferson/2005/06.01.18.47>>. 29, 33

TAO, W.; TIAN, J.; LIU, J. A new approach to extract rectangle building from aerial urban images. **International Conference on Signal Processing**, v. 1, p. 1–4, 2002. 31

TREMEAU, A.; COLANTONI, P. Regions adjacency graph applied to color image segmentation. **Image Processing, IEEE Transactions on**, v. 9, n. 4, p. 735–744, 2000. 40, 44

VESANTO, J.; ALHONIEMI, E. Clustering of the self-organizing map. **Neural Networks, IEEE Transactions on**, v. 11, n. 3, p. 586–600, 2000. 42

YONGGUAN, X.; SOOKUAN, L.; SENG, T. et al. Feature extraction using very high resolution satellite imagery. In: **Science for Society: Exploring and Managing a Changing Planet**. Anchorage: International Geoscience and Remote Sensing Symposium, 2004. v. 3, p. 2004–2007. 23

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.