



SEGMENTAÇÃO E CLASSIFICAÇÃO DE IMAGENS SAR APLICADAS À
DETECÇÃO DE ALVOS ESCUROS EM ÁREAS OCEÂNICAS DE
EXPLORAÇÃO E PRODUÇÃO DE PETRÓLEO

Patrícia Carneiro Genovez

Tese de Doutorado apresentada ao Programa
de Pós-Graduação em Engenharia Civil,
COPPE, da Universidade Federal do Rio de
Janeiro, como parte dos requisitos necessários
à obtenção do título de Doutor em Engenharia
Civil.

Orientadores: Nelson Francisco Favilla Ebecken
Corina da Costa Freitas

Rio de Janeiro
Março de 2010

SEGMENTAÇÃO E CLASSIFICAÇÃO DE IMAGENS SAR APLICADAS À
DETECÇÃO DE ALVOS ESCUROS EM ÁREAS OCEÂNICAS DE
EXPLORAÇÃO E PRODUÇÃO DE PETRÓLEO

PATRÍCIA CARNEIRO GENOVEZ

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA CIVIL.

EXAMINADA POR:

Prof. Nelson Francisco Favilla Ebecken, D.Sc.

Profª. Corina da Costa Freitas, Ph.D.

Profª. Cristina Maria Bentz, D.Sc.

Prof. Fernando Pellon de Miranda, D.Sc.

Prof. Luiz Landau, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

MARÇO DE 2010

Genovez, Patrícia Carneiro

Segmentação e Classificação de Imagens SAR
Aplicadas à Detecção de Alvos Escuros em Áreas
Oceânicas de Exploração e Produção de Petróleo/ Patrícia
Carneiro Genovez - Rio de Janeiro: UFRJ/COPPE, 2010.

XVIII, 235p.: il.; 29,7 cm.

Orientadores: Nelson Francisco Favilla Ebecken

Corina da Costa Freitas

Tese (doutorado) - UFRJ/ COPPE/ Programa de
Engenharia Civil, 2010.

Referências Bibliográficas: p. 179-195.

1. Detecção de Alvos Escuros.
 2. Detecção de Óleo e Ocorrências Ambientais.
 3. Exploração e Produção de Petróleo.
 4. Imagens de Radares Orbitais.
 - .5. Segmentação e Clusterização Automática.
 6. Sistema Híbrido Inteligente.
 7. Áreas Oceânicas.
- I. Ebecken, Nelson Francisco Favilla *et al.*. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Civil. III. Título.

*Dedico este trabalho aos meus pais
Wlademar Genovez e
Neuza Carneiro Genovez*

AGRADECIMENTOS

Ao orientador Dr. Nelson Ebecken pela oportunidade concedida, pela orientação e incentivo dedicados ao longo do desenvolvimento deste trabalho, sempre encorajando e propondo soluções valiosas para a concretização da pesquisa.

À orientadora Dra Corina da Costa Freitas pela especial amizade e compreensão, destacando a orientação dedicada mesclada de valiosas sugestões e questionamentos que foram essenciais para lapidar a pesquisa realizada.

À Cristina Bentz pela amizade que construímos durante anos de trabalho, ao longo dos quais foi possível idealizar o escopo desta pesquisa. Em especial agradeço pela sua orientação informal repleta de diálogos extremamente enriquecedores, os quais despertaram meu profundo interesse pelo tema estudado e, sem os quais, esta pesquisa não teria sido realizada.

Ao amigo Josemá Oliveira de Barros, que desde o início incentivou e acreditou neste trabalho, fornecendo contribuições significativas para a pesquisa desenvolvida, através de sua ampla experiência operacional.

À toda equipe do Controle de Contingência da PETROBRAS, pelo valioso aprendizado e pela amizade que construímos ao longo de 6 anos de “monitoramento orbital” (2003-2009), através dos quais pude iniciar minha trajetória ao longo deste amplo cenário que envolve a detecção automática de óleo e ocorrências ambientais em imagens SAR.

À empresa *Sampling* pelo apoio e coleguismo presentes, sempre que necessário, no desenvolvimento desta pesquisa, agradeço em especial aos colegas Antonio Francisco Corno, Fábio Gomes de Azevedo e José Maria Ferreira de Oliveira Barroco.

À UFRJ/COPPE pela oportunidade concedida de cursar o doutorado vinculado ao programa de Engenharia Civil: Áreas Interdisciplinares, representando um marco em minha formação profissional e acadêmica.

À PETROBRAS por todo apoio concedido, fornecendo dados, softwares, infra-estrutura e, principalmente o incentivo para o desenvolvimento desta pesquisa.

Aos queridos colegas do INPE meus agradecimentos pela sincera amizade e pelo pronto apoio nos momentos de dúvidas e incertezas, em especial aos colegas: Luciano V. Dutra, Sidnei João S. Sant'Anna, Graziela Scoefield, João A. Lorenzetti, Guaracy J. Erthal, Leonardo Sant'Anna Bins, Camilo D. Rennó, Fábio Furlan, José C. Mura, Cláudia Almeida, António

Miguel V. Monteiro, Valdete Duarte, Paulo Roberto Martini, Hellen Borges, Maurício A. Moreira, Joaquim Filho e todos os que de alguma forma auxiliaram.

Aos colegas do Serviço Florestal Brasileiro que auxiliaram e apoiam a finalização do trabalho em especial: Ekena Pinagé, Igor Parreira, Ricardo Vilar, Anna Fanzeres, Ivã Dornellas, Mauricio Vianna (Vulgo Maléfico), Dulce V. Amaral e Maria Alice C. Tocantins.

À querida amiga Neuza Ferreira, meus profundos e sinceros agradecimentos, por todos os momentos especiais que vivemos juntas, sempre me encorajando nos momentos mais difíceis, os quais não teria superado sem a sua amizade. Tudo o que aprendi e vivenciei ao seu lado guardarei com muito zelo.

À minha querida irmã Sarita Carneiro Genovez e ao meu cunhado José Donizeti Costa Junior pelo carinho e amizade, sempre acreditando em meu trabalho e em meus sonhos. Aos meus irmãos Leonardo Carneiro Genovez e Walter Carneiro Genovez e as minhas cunhadas Mariana Genovez e Simone Genovez, agradeço pelo apoio e incentivo, me encorajando a enfrentar os desafios.

Ao Ramon Moraes de Freitas minha especial gratidão pela sua dedicação e apoio, estando ao meu lado nos momentos mais difíceis da tese, dialogando e sugerindo soluções. Em especial agradeço pelo carinho e pelo amor cultivados, sem os quais seria difícil prosseguir.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

**SEGMENTAÇÃO E CLASSIFICAÇÃO DE IMAGENS SAR APLICADAS À
DETECÇÃO DE ALVOS ESCUROS EM ÁREAS OCEÂNICAS DE
EXPLORAÇÃO E PRODUÇÃO DE PETRÓLEO**

Patrícia Carneiro Genovez

Março/2010

Orientadores: Nelson Francisco Favilla Ebecken

Corina da Costa Freitas

Programa: Engenharia Civil

Sistemas para detecção automática de vazamentos de óleo têm sido desenvolvidos para auxiliar a interpretação das imagens SAR, sendo compostos por 4 etapas: a) pré-processamento; b) detecção de alvos escuros; c) cálculo de atributos, e; d) classificação dos alvos em óleo ou assemelhados. A detecção dos alvos escuros é a principal etapa, sem a geometria dos alvos não há como classificá-los em óleo ou assemelhados. Neste contexto, o desenvolvimento de um procedimento automático para a detecção de alvos escuros em imagens SAR, integrando técnicas de segmentação de imagens e de reconhecimento de padrões constitui o objetivo proposto. Os procedimentos envolveram as seguintes etapas: a) definição do espaçamento de *pixel* e do filtro para o pré-processamento; c) partição das imagens em regiões utilizando técnicas de segmentação por crescimento de regiões; d) cálculo, análise exploratória e seleção de atributos; e) detecção de alvos escuros utilizando clusterização de dados, e; f) validação do método proposto. Embora o método proposto tenha detectado alvos escuros com diferentes níveis de complexidade em diferentes recortes de cena, não obteve resultados satisfatórios em todos os exemplos estudados. Como ainda não há um consenso sobre a utilização operacional de métodos plenamente automáticos, a elaboração de um sistema híbrido inteligente capaz de indicar através de regras de decisão as imagens aptas para o processamento automático ou semi-automático foi proposta. Os resultados evidenciaram o potencial destas regras para auxiliar no processo de automatização, sendo necessárias mais amostras para retornar regras mais robustas aplicáveis a toda nova imagem SAR adquirida.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

**SAR IMAGE CLASSIFICATION AND SEGMENTATION FOR DARK SPOT
DETECTION IN OCEANIC PETROLEUM EXPLORATION
AND PRODUCTION AREAS**

Patrícia Carneiro Genovez

March/2010

Advisors: Nelson Francisco Favilla Ebecken

Corina da Costa Freitas

Department: Civil Engineering

Automatic oil detection systems have been developed to improve SAR image interpretation, composed of four principal stages: a) image pre-processing; b) dark spot detection; c) feature extraction, and; d) oil and look-alike classification. The dark spot detection is considered the main step in the processing chain: without the geometry of the spots, the oil and look-alikes classification is unfeasible. In this context, this work aimed to develop an automatic classification procedure able to detect dark spots in SAR images, by the integration of segmentation and pattern recognition techniques. The proposed methodology were organized in six stages as follow: a) images subset selection; b) pixel spacing and filter definition; c) image partition into proper regions using region growing segmentation techniques; d) features extraction, exploratory analyses and feature selection; e) dark spot detection using data clustering, and; f) validation of the proposed method. In spite of the proposed method had been able to detect dark spots with different complexity levels in different image subsets, it didn't achieve good results for all analyzed examples. Considering that in the scientific community there isn't a wide agreement about the operational use of fully automatic methods, the development of an intelligent hybrid system, including decision rules able to conduct the images for one automatic or semi-automatic processing, was an interesting approach. The potential of these rules to improve the automation process was indicated. Nevertheless, more samples to return more robust rules are recommended in order to be widely applied to all SAR images acquired.

SUMÁRIO

	Pág.
RESUMO.....	iv
ABSTRACT.....	v
ÍNDICE DE FIGURAS.....	ix
ÍNDICE DE TABELAS.....	xvi
CAPÍTULO 1 - INTRODUÇÃO.....	1
1.1 – Objetivos.....	5
1.1.1 – Objetivos Específicos.....	5
1.2 - Principais Contribuições	6
1.4 – Organização da Tese.....	7
CAPÍTULO 2 – IMAGEAMENTO SAR EM ÁREAS OCEÂNICAS	8
2.1 - Fatores de Influência na Detecção de Alvos Escuros em Imagens SAR.....	11
2.1.1 - O Ruído Speckle.....	14
CAPÍTULO 3 – DETECÇÃO AUTOMÁTICA DE ÓLEO E OCORRÊNCIAS AMBIENTAIS EM IMAGENS SAR	16
3.1 – Detecção de Alvos Escuros em Imagens SAR.....	17
3.1.1 – Segmentação de Imagens SAR	30
3.1.1.1 – MultiSeg: Informações Técnicas.....	30
3.1.1.2 – SPRING: Informações Técnicas.....	32
3.2 – Clusterização de Dados	32
3.2.1 – Clustan: Informações Técnicas	36
3.3 – Avaliação dos Resultados obtidos para Detecção de Alvos Escuros em Imagens SAR	38
3.3.1 – Métricas calculadas no Nível das Regiões	40

3.3.1 – Métricas Calculadas no Nível dos <i>Pixels</i>	43
CAPÍTULO 4 – METODOLOGIA E RECURSOS UTILIZADOS	46
4.1 – Etapa 1: Conjunto de Dados – Informações Técnicas e Seleção de Exemplos	47
4.2 – Etapa 2: Determinação do Espaçamento de <i>Pixel</i>	51
4.3 – Etapa 3: Determinação do Filtro/Janela	52
4.4 – Etapa 4: Determinação do Procedimento de Segmentação	53
4.5 – Etapa 5: Determinação do Procedimento de Clusterização.....	55
4.6 – Etapa 6: Validação do Método Proposto	56
CAPÍTULO 5 – RESULTADOS E DISCUSSÕES.....	58
5.1 – Etapa 1: Determinação do Espaçamento de <i>Pixel</i>	58
5.1.1 – Determinação dos Parâmetros Adequados para o Software MultiSeg	58
5.1.2 – Determinação do Espaçamento de <i>Pixel</i>	66
5.2 – Etapa 2: Determinação do Filtro/Janela	70
5.3 – Etapa 3: Determinação do Segmentador	73
5.4 – Etapa 4: Determinação do Método de Clusterização.....	89
5.4.1 – Cálculo dos Atributos	90
5.4.2 – Análise Exploratória e Seleção de Atributos	93
5.4.2.1 – Análise de Matriz de Correlação.....	94
5.4.2.2 – Análise de Componentes Principais	101
5.4.2.4 – Seleção de Atributos	104
5.4.3 – Reconhecimento de Padrões por Clusterização Automática	112
5.4.4 – Comparação entre os Métodos Automático e Semi-automático	139
5.5 – Regra de Decisão e Validação do Modelo Proposto	157
CAPÍTULO 6 – CONCLUSÕES E RECOMENDAÇÕES	167

REFERÊNCIAS BIBLIOGRÁFICAS	179
APÊNDICE I	192
APÊNDICE II	193
APÊNDICE III	194
APÊNDICE IV	206
APÊNDICE V	219

ÍNDICE DE FIGURAS

	Pág.
Figura 1.1 – Etapas envolvidas nos algoritmos de detecção automática de óleo e ocorrências ambientais	3
Figura 1.1 – Etapas envolvidas nos sistemas de detecção automática de óleo e ocorrências ambientais	3
Figura 2.2 – Tipos de espalhamento produzidos pela interação do pulso emitido pelo RADAR com os alvos presentes na superfície oceânica, juntamente com os alvos escuros freqüentemente detectáveis nas imagens.	11
Figura 2.3 – Tamanho Relativo dos Comprimentos de Onda na Faixa de Microondas.	12
Figura 2.4 - Geometria de imageamento do SAR e o efeito do ângulo de incidência no <i>near</i> e no <i>far range</i> : Imagem RADARSAT-1.	13
Figura 2.5 – Representação do ruído <i>speckle</i>	15
Figura 3.1 - Diferentes tipos de processamento efetuados no nível do <i>pixel</i> e das regiões encontrados na literatura.....	25
Figura 3.2 – Comparação entre o agrupamento efetuado a partir de regiões e de <i>pixels</i>	28
Figura 3.3 – Ilustração das definições referentes a terminologia adota ao longo da tese.....	39
Figura 3.4 - Esquema de como é efetuado o ajuste de regiões entre as duas imagens: exemplo focado para no ajuste efetuado para os alvos escuros.	42
Figura 4.1 – Etapas de processamento envolvidas no sistema de detecção automática de óleo e ocorrências ambientais.....	46
Figura 4.2 – Etapas integrantes da metodologia proposta na tese.....	47
Figura 4.3 - Histogramas e <i>box plots</i> para uma mesma imagem no formato CEOS e MrSid.	49
Figura 4.4 – Recorte da cena RADARSAT-1 (03/04/2002) utilizada no estudo proposto	51
Figura 4.5 – Metodologia adotada para investigar o espaçamento de <i>pixel</i> adequado para o processamento dos dados RADARSAT no sistema de detecção automático .	52

Figura 4.6 – Metodologia adotada para investigar o filtro/janela adequado para o processamento dos dados RADARSAT no sistema de detecção automático.	53
Figura 4.7 - Fluxo de procedimentos metodológicos planejado para investigar a etapa de segmentação em imagens de SAR.....	54
Figura 4.8 - Fluxo de procedimentos planejado para investigar o método de clusterização mais adequado a ser utilizado após a segmentação das imagens SAR.....	56
Figura 5.1 - Resultados obtidos para segmentações realizadas nos modos/formato: <i>Cartoon/Amplitude</i> , <i>Cartoon/Intensidade</i> e <i>Texture/ Intensidade</i>	59
Figura 5.2 – Exemplos de segmentações nos modos <i>Cartoon/Amplitude</i> , e <i>Cartoon/Intensidade</i>	60
Figura 5.3 - Resultados obtidos para segmentações realizadas com Nel 2, 3 e 4, considerando todos os espaçamentos de pixel avaliados.....	62
Figura 5.4 – Exemplo de <i>box plot</i> calculado para as regiões claras e escuras e método proposto para o cálculo do Limiar de Similaridade em dB.....	64
Figura 5.5 – Metodologia padrão proposta para calcular o Limiar de Similaridade....	65
5.6 – Resultados obtidos para os diferentes espaçamentos de <i>pixel</i> avaliados.	67
Figura 5.7 – Resultados obtidos para a imagem sem filtro e com filtro Lee 3x3, Mediana 3x3, Sigma 3x3 e Kuan 3x3.....	71
Figura 5.8 - Recortes de cena utilizados juntamente com os <i>Phantons</i> elaborados manualmente indicando os alvos escuros de referência.	73
Figura 5.9 - Gráficos comparando as métricas calculadas para os segmentadores MultiSeg e SPRING com e sem a aplicação de filtro.....	78
Figura 5.10 – Efeito da aplicação do Filtro Lee (3X3) no Potencial de Detecção de Alvos Escuros para os softwares utilizados.	82
Figura 5.11 – Síntese dos resultados obtidos para o agrupamento das segmentações realizadas nos softwares SPRING e MultiSeg.	84
Figura 5.12 – Resultados do agrupamento das segmentações realizadas para os diferentes recortes de cena com e sem filtragem para os softwares SPRING e MultiSeg.....	85
Figura 5.13 – Etapas integrantes da metodologia proposta nesta última etapa da tese.....	90

Figura 5.14 – Método proposto para a definição do Atributo Classe do <i>Phantom</i>	92
Figura 5.15 – Atributos e procedimentos utilizados na análise exploratória e seleção de atributos.....	94
Figura 5.16 - Gráficos resultantes da ACP efetuadas para as 11 cenas analisadas.	103
Figura 5.17 - Espacialização dos 10 atributos mais indicados pelo <i>ranking</i> para as 11 cenas estudadas.....	108
Figura 5.18 - Metodologia utilizada para a validação do conjunto adequado de atributos.....	113
Figura 5.19 – Influência da utilização simultânea e individual dos atributos de Máximo e Mínimo na clusterização.	114
Figura 5.20 – Propriedades dos atributos de Máximo e Mínimo que influenciam na clusterização.....	117
Figura 5.21 – Gráfico e espacialização das clusterizações realizadas para a Cena 1	119
Figura 5.22 – Gráfico e espacialização das clusterizações realizadas para a Cena 2	119
Figura 5.23 – Gráfico e espacialização das clusterizações realizadas para a Cena 3	120
Figura 5.24 – Gráfico e espacialização das clusterizações realizadas para a Cena 4	120
Figura 5.25 – Gráfico e espacialização das clusterizações realizadas para a Cena 5	121
Figura 5.26 – Gráfico e espacialização das clusterizações realizadas para a Cena 6	121
Figura 5.27 – Gráfico e espacialização das clusterizações realizadas para a Cena 7	122
Figura 5.28 – Gráfico e espacialização das clusterizações realizadas para a Cena 8	122
Figura 5.29 – Gráfico e espacialização das clusterizações realizadas para a Cena 9	123
Figura 5.30 – Gráfico e espacialização das clusterizações realizadas para a Cena 10	123

Figura 5.31 – Gráfico e espacialização das clusterizações realizadas para a Cena 11.....	124
Figura 5.32 - Espacialização dos melhores resultados da clusterização considerando o universo de 9 e 35 atributos.....	129
Figura 5.33 - Sobreposição dos <i>Phantons</i> com os resultados das clusterizações....	131
Figura 5.34 - Sobreposição dos <i>Phantons</i> com os resultados das clusterizações....	132
Figura 5.35 – Exatidão, Omissão e Inclusão nos resultados obtidos pelas clusterizações para as 5 cenas consideradas satisfatórias.....	134
Figura 5.36 - Diferenças encontradas entre os <i>clusters</i> utilizando 4 e 5 atributos para a cena 11.....	137
Figura 5.37 – Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na detecção de alvos escuros em imagens SAR	141
Figura 5.38 – Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na detecção de alvos escuros em imagens SAR	145
Figura 5.39 – Ampliação da malha de filamentos provenientes de óleos biogênicos existentes nas cenas 1, 10 e 11.....	148
Figura 5.40 – Gráfico com os Erros de Omissão, de Inclusão e as <i>ExAE</i> obtidas para os métodos automático e semi-automáticos em cada recorte de cena.	148
Figura 5.41 – Detalhes da cena 4, ampliando os resultados das clusterizações para os alvos escuros.....	150
Figura 5.42 – Detalhes da cena 5, ampliando os resultados das clusterizações para os alvos escuros relacionados a baixas intensidades de vento.	150
Figura 5.43 – Detalhes da cena 5, ampliando os resultados das clusterizações para os alvos escuros relacionados a óleo.	151
Figura 5.44 – Detalhes da cena 5, ampliando os resultados das clusterizações para os alvos escuros relacionados a óleo.	151
Figura 5.45 - Comparação entre o agrupamento realizado: a) pelo <i>Clustan</i> , considerando as regiões da segmentação, e; b) pelo <i>K-means</i> tradicional aplicado diretamente sobre os pixels.	152
Figura 5.46 – Comparação entre resultados obtidos por diferentes métodos automáticos de detecção de alvos escuros disponíveis na literatura e os obtidos na pesquisa desenvolvida.....	153

Figura 5.47 – Comparação entre resultados obtidos por diferentes métodos automáticos de detecção de alvos escuros disponíveis na literatura e os obtidos na pesquisa desenvolvida.....	155
Figura 5.48 – Método para criação e validação da regra de decisão para definir o processamento como automático ou semi-automático.....	159
Figura 5.49 – Regra de decisão obtida pela ACR para definir o processamento como automático ou semi-automático.	160
Figura 5.50 – Recortes de Cena escolhidos para validação do método proposto para a detecção automática de alvos escuros em imagens SAR.....	161
Figura 5.51 – Resultados obtidos para validação do método de detecção automática de alvos escuros proposto na tese.....	162
Figura 5.52 – Dimensões do recorte de cena B, evidenciando o conjunto de pixels nas regiões escuras delimitadas pela segmentação.....	163
Figura 5.53 – Validação da regra de decisão para definir o processamento como automático ou semi-automático.	165

ÍNDICE DE TABELAS

	Pág.
Tabela 3.1 - Matriz de confusão para o agrupamento das regiões em alvos escuros e oceano.....	43
Tabela 4.1 - Configuração dos modos de imageamento do satélite RADARSAT-1 adequados ao monitoramento costeiro e oceânico	47
Tabela 4.2 - Data das cenas e parâmetros de aquisição das imagens pelo sensor para os recortes de cena selecionados.....	49
Tabela 4.3 – Características associadas à seleção dos recortes de cena utilizados..	50
Tabela 5.1 – Métricas utilizadas para avaliação dos diferentes espaçamentos de pixel	66
Tabela 5.2 – Métricas utilizadas para avaliação dos diferentes filtros/janela testados.	70
Tabela 5.3 – Variáveis calculadas com a amostragem do <i>buffer</i> e LS calculados para o SPRING e para o MultiSeg, considerando a diferença entre os quartis do <i>box plot</i> .	76

Tabela 5.4 – Métricas utilizadas para avaliação dos diferentes segmentadores aplicados nos recortes de cena com e sem filtragem.....	77
Tabela 5.5 - Síntese dos resultados obtidos para as métricas de avaliação por recorte de cena estudado após a aplicação do filtro Lee/3X3.	81
Tabela 5.6 - Teste de Wilcox efetuado para os resultados obtidos com o segmentador MultiSeg com e sem filtragem.....	83
Tabela 5.7 - Teste de Wilcox efetuado para os resultados obtidos com o segmentador SPRING com e sem filtragem.	83
Tabela 5.8 - Teste de Wilcox efetuado para os resultados obtidos com o segmentador MultiSeg e SPRING com filtragem.	88
Tabela 5.9 – Coeficientes de Correlação dos atributos estatísticos em relação ao <i>Phantom</i>	95
Tabela 5.10 - Análise Síntese da Matriz de Correlação para as 11 Cenas: Atributos Estatísticos.	96
Tabela 5.11 – Coeficientes de Correlação dos atributos Geométricos Genéricos e Poligonais em relação ao <i>Phantom</i>	97
Tabela 5.12 - Análise síntese da Matriz de Correlação para as 11 Cenas: Atributos Geométricos Genéricos e Poligonais	98
Tabela 5.13 – Coeficientes de Correlação dos atributos Geométricos <i>Skeletons</i> em relação ao <i>Phantom</i>	98
Tabela 5.14 - Análise síntese da Matriz de Correlação para as 11 Cenas: Atributos Geométricos <i>Skeletons</i>	99
Tabela 5.15 – Coeficientes de Correlação dos atributos texturais em relação ao <i>Phantom</i>	100
Tabela 5.16 - Análise síntese da Matriz de Correlação para as 11 Cenas: Atributos Texturais.....	100
Tabela 5.17 - Síntese do número de componentes necessários em cada cena para manter pelo menos 90% da representatividade dos dados.....	104
Tabela 5.18 – <i>Ranking</i> dos 10 atributos mais importantes para cada cena.....	105
Tabela 5.19 – <i>Ranking</i> dos 10 atributos mais importantes indicados pelo chi-quadrado do 1º ao 10º lugar.	107

Tabela 5.20 – Exatidão dos Alvos Escuros (<i>ExAE</i>) obtida para as clusterizações efetuadas, considerando os atributos de máximo e mínimo simultaneamente e individualmente.....	118
Tabela 5.21 - Síntese dos resultados obtidos considerando as clusterizações efetuadas com os 9 e 35 atributos	128
Tabela 5.22 – Índices de <i>ExAE</i> obtidos para as 11 cenas com 4 e 5 atributos.	137
Tabela 5.23 - Métricas de avaliação calculadas para os resultados das clusterizações efetuadas pelos métodos automático e semi-automático.....	144
Tabela 5.24 – Variável dependente e atributos preditores utilizados na ACR para as 11 cenas.....	158
Tabela 5.25 - Métricas de avaliação calculadas para os resultados das clusterizações efetuadas nos 3 novos recortes de cena para validação do método proposto na tese	161
Tabela 5.26 - Métricas de avaliação calculadas para os resultados das clusterizações automáticas efetuadas nos 3 novos recortes de cena para validação do método proposto.....	164

CAPÍTULO 1

INTRODUÇÃO

Sistemas de vigilância e monitoramento são considerados medidas essenciais para o controle e a prevenção de áreas suscetíveis à poluição por óleo. Este tópico compõe grande parte das deliberações constantes na IMO (*International Maritime Organization*) regulamentada internacionalmente pela MARPOL (*International Convention for the Prevention of Pollution from Ships*) e pela UNCLOS (*United Nations Convention on the Law of the Sea*). Tendo como base estas regulamentações, diversos países têm implantado sistemas de monitoramento costeiros e oceânicos utilizando como instrumento imagens provenientes de sensores orbitais e/ou aerotransportados (ENGELHARDT, 1999).

O sucesso de uma operação de contingenciamento está diretamente ligado à rapidez com que uma mancha de óleo é detectada e monitorada (ASSILZADEH & MANSOR, 2001). Nestas situações, o sensoriamento remoto é considerado estratégico, sendo amplamente utilizado para o monitoramento de ocorrências ambientais envolvendo derramamento de óleo no mar (ENGELHARDT, 1999). Sensores remotos produzem informações consistentes sobre locais de difícil acesso, fornecendo uma visão sinótica sob áreas dotadas de grande extensão – como é o caso das regiões costeiras e oceânicas. O potencial de recobrimento sistemático das áreas de interesse, bem como a capacidade para disponibilizar, em tempo próximo do real, informações sobre eventos imprevisíveis - como é o caso de acidentes envolvendo derramamento de óleo (GADE *et al.*, 2000; BENTZ *et al.*, 2005; BENTZ *et al.*, 2007a; GENOVEZ *et al.*, 2007) - constituem subsídios essenciais à tomada de decisão durante emergências ambientais.

Sistemas sensores operando em diversas faixas do espectro eletromagnético têm sido utilizados para detecção de óleo e ocorrências ambientais diversas no mar (BENTZ *et al.*, 2007a; BROWN & FINGAS, 2001; FINGAS *et al.*, 2000a e 2000b). Em especial, imagens provenientes de RADAR¹ de Abertura Sintética (SAR²) têm sido amplamente utilizadas devido à possibilidade de fornecimento dos dados em tempo próximo do real, bem como, à capacidade de cobrir extensas áreas adquirindo imagens durante o dia ou à noite sem interferência de nuvens (STAPLES &

¹ RADAR: é a sigla resultante de *Radio Detection And Ranging*, referente a sensores que utilizam o intervalo do espectro eletromagnético entre as freqüências de 0,3 GHz a 300 GHz (microondas), operando em plataformas orbitais ou aeronaves.

² SAR: *Synthetic Aperture RADAR*

HODGINS, 1998; GADE *et al.*, 2000; BENTZ & MIRANDA, 2001; MIRANDA *et al.*, 2004; FINGAS & BROWN, 2005; CHANG *et al.*, 2005; BENTZ *et al.*, 2005; GENOVEZ *et al.*, 2007). Nas imagens de RADAR adquiridas em áreas oceânicas, eventos dotados de baixo retroespalhamento são visualmente identificados como alvos escuros, os quais podem estar associados à presença de óleo.

Desde a década de noventa, sistemas operacionais para detecção de vazamento de óleo no mar utilizam de maneira integrada imagens de RADAR – orbital e/ou aerotransportado – e monitoramento aéreo das áreas alvo (BREKKE & SOLBERG, 2005; BENTZ *et al.*, 2007a). A freqüência de aquisição, o processamento e o tempo de entrega das imagens, bem como a experiência dos intérpretes na detecção de óleo e ocorrências ambientais que se assemelham ao óleo nas imagens de RADAR, são variáveis determinantes para a eficiência dos sistemas de monitoramento. Geralmente, as imagens são interpretadas visualmente e as coordenadas geográficas das manchas suspeitas de ser óleo direcionam as verificações de campo e as ações de contingenciamento, quando necessário (STAPLES & HODGINS, 1998; BENTZ & MIRANDA, 2001).

Em alguns casos, especialmente quando um grande número de cenas SAR têm de ser interpretadas, o processamento das imagens e a distinção visual entre manchas de óleo e ocorrências ambientais pode tornar-se uma tarefa difícil e com significativa demanda de tempo (GADE *et al.*, 1998; BREKKE & SOLBERG, 2005;

Mozilla Firefox.lnk SOLBERG *et al.*, 2005; KERAMITSOGLOU *et al.*, 2006). Por outro lado, mesmo que intérpretes devidamente treinados sejam capazes de detectar visualmente vazamentos de óleo e diferenciá-los de ocorrências ambientais, a experiência de cada intérprete confere subjetividade ao processo, influenciando diretamente os resultados da interpretação (INDREGARD, 2004; SOLBERG *et al.*, 2005; SOLBERG, 2005; BENTZ *et al.*, 2007a).

Sistemas para detecção automática de vazamentos de óleo têm sido desenvolvidos para otimizar e auxiliar a interpretação das imagens SAR, buscando reduzir a intervenção humana e, consequentemente, o grau de subjetividade associado ao processo (KUBAT *et al.*, 1998; ESPEDAL & WAHL, 1999; SOLBERG *et al.*, 1999, 2003 e 2005; SOLBERG, 2005; KERAMITSOGLOU *et al.*, 2006). A cadeia de processamentos considera prioritariamente as áreas oceânicas e tem como

objetivo a emissão de um alerta quando um alvo escuro for identificado como uma possível mancha de óleo (KERAMITSOGLOU *et al.*, 2006).

De maneira geral, os sistemas para detecção automática de óleo são compostos por 4 etapas essenciais ao processo (Figura 1.1): a) pré-processamento das imagens; b) detecção de alvos escuros; c) cálculo de atributos, e; d) classificação dos alvos escuros em óleo e assemelhados (BJERDE *et al.*, 1993; DEL FRATE *et al.*, 2000; SOLBERG *et al.*, 1996, 1997 e 2003; ZHENG *et al.*, 2005; BREKKE & SOLBERG, 2005; ANGIULI *et al.*, 2006; CHANG *et al.*, 2005 e 2008; GENOVEZ *et al.*, 2009). Na literatura, há uma diversidade de métodos propostos, o que diferencia cada um é a combinação de técnicas, atributos e algoritmos utilizada em cada uma das etapas citadas. Pesquisas nesta área têm sido desenvolvidas por mais de uma década (BREKKE & SOLBERG, 2005; TOPOUZELIS, 2008) e constituem uma extensa e promissora área de desenvolvimento (SOLBERG *et al.*, 1999; DEL FRATE *et al.*, 2000; GALLAND *et al.* 2004).

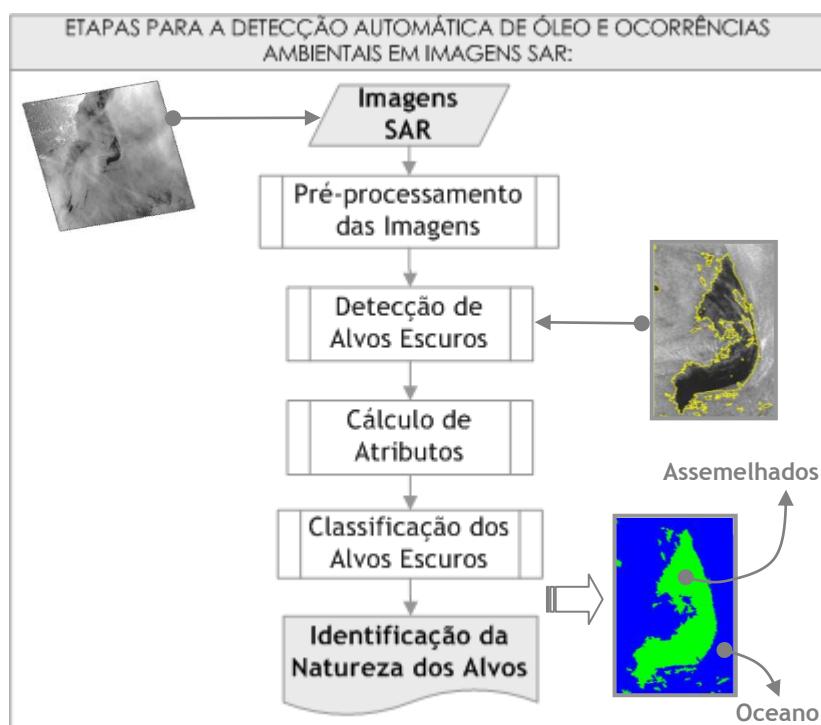


Figura 1.1 – Etapas envolvidas nos sistemas de detecção automática de óleo e ocorrências ambientais.

Dentre as etapas envolvidas, a detecção dos alvos escuros é a principal etapa da cadeia de processamentos. Sem a detecção e individualização da geometria dos alvos não há como prosseguir na etapa de classificação dos eventos em óleo ou assemelhados (KANAA *et al.*, 2003; PELIZZARI & DIAS, 2003; SOLBERG & BREKKE, 2008; TOPOUZELIS *et al.*, 2007a; TOPOUZELIS, 2008). Diversas abordagens para a detecção de alvos escuros têm sido propostas combinando técnicas de segmentação

de imagens e agrupamento de regiões (BREKKE & SOLBERG, 2005; KARATHANASSI *et al.*, 2006; TOPOUZELIS, 2008; SOLBERG & BREKKE, 2008).

Analisando a literatura produzida nas últimas décadas sobre o tema, é perceptível que ainda não há um consenso sobre a utilização operacional de sistemas automáticos para a detecção de óleo e ocorrências ambientais em imagens SAR (SEASAR, 2010). Parte dos autores classifica sua metodologia como automática, mas não operacional (LOPEZ *et al.*, 2005; KARATHANASSI *et al.*, 2006; TOPOUZELIS *et al.* 2007). A utilização de métodos híbridos, mesclando procedimentos automáticos e semi-automáticos, foi sugerida em algumas abordagens (CALABRESI *et al.*, 1999; FISCELLA *et al.*, 2000; PELIZARRI & DIAS, 2003; LOPEZ *et al.*, 2005). Trabalhos que indicam suas metodologias como automáticas e operacionais, enfatizam a necessidade de aprofundar os testes, considerando mais imagens, com diferentes tipos de alvos, em cenários com intensidade de vento variada (SOLBERG & BREKKE, 2008; TOPOUZELIS *et al.*, 2009).

Neste contexto, a formulação de uma hipótese que busca investigar se a definição de um método automático para detecção de alvos escuros em imagens SAR é factível, constitui uma preocupação de relevante interesse no meio acadêmico, que ainda não foi plenamente respondida.

Os trabalhos consultados utilizaram inúmeras técnicas de segmentação de imagens para particionar a imagem em regiões. Entretanto, para agrupar estas regiões em alvos escuros e oceano, a diversidade de técnicas utilizadas é restrita, sendo a limitação a mais comumente utilizada (TOPOUZELIS *et al.*, 2007b; GALLAND *et al.*, 2003 e 2004).

Embora a utilização de regiões, ao invés de *pixels*, possibilite a associação de múltiplos atributos aos segmentos, ampliando as possibilidades de investigação e de resultados passíveis de serem obtidos (NAVULUR, 2006; BAATZ *et al.*, 2003), na bibliografia consultada o cálculo de atributos foi utilizado somente na etapa de classificação dos alvos em óleo ou assemelhados e não na detecção de alvos escuros (TOPOUZELIS, 2008; KERAMITSOGLOU *et al.*, 2006; BREKKE & SOLBERG, 2005).

Considerando o pressuposto de que não existe um algoritmo de segmentação que garanta desempenho plenamente satisfatório em todas as imagens (BEAUCHEMIN & THOMSON, 1997; ESPINDOLA, 2006; BEAULIEU & TOUZI, 2002; SOUSA, 2005), a proposta de uma nova abordagem, que utiliza o cálculo de atributos

e clusterização de dados na etapa de detecção de alvos escuros, foi considerada promissora.

Com base nas evidências descritas, a tese considerou a hipótese de que um método para a detecção automática de alvos escuros em imagens SAR, integrando técnicas de segmentação de imagens e reconhecimento de padrões, pode ser implementado com resultados satisfatórios.

No âmbito dos sistemas automáticos, foi considerada como premissa a possibilidade de padronizar: a) os métodos de pré-processamento a serem aplicados nas imagens; b) os parâmetros associados ao segmentador, tendo como foco a detecção de alvos escuros em diferentes escalas, priorizando eventos associados a óleo; c) um conjunto adequado de atributos, capaz de retornar *clusters* representativos para detectar alvos escuros e oceano.

1.1 – Objetivos

O principal objetivo foi propor um procedimento automático para a detecção de alvos escuros em imagens SAR. Para tal, a influência dos métodos de pré-processamento das imagens, o desempenho de diferentes segmentadores e a utilização de atributos no processo de clusterização em alvos escuros e oceano foram investigados.

Como resultado final, pretende-se disponibilizar um método com desempenho satisfatório para detectar automaticamente a geometria dos alvos escuros nas imagens SAR. Posteriormente, este procedimento será acoplado ao sistema de classificação automática desenvolvido por BENTZ (2006) a fim de integrar toda a cadeia de processamentos envolvida em um sistema de detecção automático de óleo.

1.1.1 – Objetivos Específicos

Para alcançar os objetivos propostos pela pesquisa, têm-se como objetivos específicos:

- a) Com base em um banco de dados estudado e validado previamente (BENTZ, 2006), selecionar imagens com níveis de complexidade de cena diferenciados;
- b) Determinar a configuração da resolução espacial, do filtro e da janela a ser utilizada como padrão na etapa de pré-processamento das imagens SAR;

- c) Verificar o desempenho integrado de técnicas de segmentação e clusterização no processo de detecção de alvos escuros;
- d) Avaliar a influência da utilização de diferentes atributos no processo de agrupamento de regiões em alvos escuros e oceano, indicando um conjunto padrão de atributos;
- e) Avaliar qualitativamente e quantitativamente os resultados provenientes dos diferentes procedimentos testados;
- f) Definir o procedimento automático mais adequado para a detecção de alvos escuros em imagens SAR, considerando os resultados obtidos nas etapas de pré-processamento, segmentação, cálculo de atributos e clusterização, e;
- g) Validar o método proposto.

1.2 - Principais Contribuições

Como principal contribuição, espera-se que um esforço futuro para implementar, de forma automática, o método proposto minimize: a) o tempo gasto na interpretação visual das imagens, especialmente quando houver um grande volume de dados; b) o grau de subjetividade associado à interpretação visual e à delimitação manual da geometria dos alvos escuros, e; c) a quantidade de falsos alarmes, otimizando as rotas efetuadas pelas inspeções aéreas.

Outras contribuições de relevante importância, decorrentes da rápida e automática extração da geometria dos alvos identificados nas imagens, é a possibilidade de criação e atualização constante de um banco de dados contendo: a) a geometria dos alvos escuros; b) o registro da interpretação destes alvos realizada pelos especialistas, e; c) os resultados provenientes das verificações de campo, confirmando ou não estas interpretações. Este banco de dados pode permitir a realização de análises estatísticas e espaciais que, ao longo do tempo, forneceriam diagnósticos históricos sobre a variação da freqüência de ocorrência dos diversos fenômenos identificados, permitindo avaliar e controlar a qualidade do sistema de monitoramento.

Este banco de dados também pode viabilizar o teste e a validação de diferentes algoritmos para detecção automática, bem como constituir um importante repositório de dados, permitindo o armazenamento de todo conhecimento gerado pelo sistema com base nos exemplos obtidos. Desta forma, o próprio sistema pode servir

como um guia, um tutorial para treinar intérpretes inexperientes, fornecendo uma interpretação prévia das feições detectadas.

Finalmente, outra expectativa é de que o método proposto consiga retornar geometrias com acurácia suficiente para alimentar o classificador desenvolvido por BENTZ (2006) sem prejudicar a performance do mesmo.

1.4 – Organização da Tese

A tese abordará os principais problemas envolvidos no desenvolvimento de sistemas para detecção automática de óleo em imagens SAR, tendo como foco a definição de uma metodologia para a detecção da geometria de alvos escuros capaz de ser utilizada operacionalmente com o mínimo de intervenção humana possível.

Neste contexto, os capítulos subseqüentes abordam: a) uma revisão sobre o imageamento SAR em áreas oceânicas, destacando os fatores que influenciam na detecção de óleo (Capítulo 2); b) o estado da arte relacionado aos sistemas desenvolvidos para detecção automática de óleo em imagens SAR (Capítulo 3); c) a indicação dos procedimentos metodológicos utilizados ao longo da tese (Capítulo 4); d) a apresentação dos resultados obtidos (Capítulo 5); e) conclusões (Capítulo 6).

CAPÍTULO 2

IMAGEAMENTO SAR EM ÁREAS OCEÂNICAS

A maioria dos sistemas de monitoramento voltados para detecção de poluição por óleo no oceano utilizam RADARes de Abertura Sintética (SAR) instalados a bordo de satélites e/ou aeronaves (FERRARO *et al.*, 2007). Nas imagens de RADAR, o oceano é representado pela intensidade do sinal de retorno que é retroespelhado pela superfície do mar e detectado pelo sensor. A intensidade média do sinal de retorno atribuída a cada *pixel*, ou célula de resolução, é proporcional à rugosidade da superfície na escala de comprimento de onda emitida pelo RADAR (ELACHI, 1987; BREKKE & SOLBERG, 2005; MONTALI *et al.*, 2006).

Este mecanismo de espalhamento das microondas na superfície do mar pode ser descrito pelo modelo de espalhamento Bragg (BREKKE & SOLBERG, 2005; TOPOUZELIS, 2008) e ocorre devido à rugosidade causada pela presença de pequenas ondas induzidas pelo vento (na escala de milímetros a centímetros) denominadas ondas capilares de gravidade ou ondas Bragg (MERCIER & ARDHUIN, 2005a; ANGIULI *et al.*, 2006). Sistemas SAR são particularmente sensíveis a estas ondas, à medida que os pulsos de microondas emitidos pela antena entram em ressonância com os comprimentos de onda na superfície do oceano dotados de semelhante magnitude (BENTZ, 2006; MERCIER & ARDHUIN, 2005a; MACCANDLESS & JACKSON, 2004; VALENZUELA, 1978).

De acordo com o modelo de espalhamento Bragg, o potencial de retroespelhamento do sinal emitido pelo RADAR é proporcional à densidade de energia espectral das ondas curtas oceânicas, as quais se propagam através ou ao longo da direção de visada da antena. Estas ondas capilares de gravidade são produzidas e propagadas pelo vento na superfície do mar com comprimentos de onda (λ_B) similares aos do RADAR (λ_r) e podem ser representados através da equação $\lambda_B = \lambda_r / 2 \operatorname{sen} \theta$ (HOLT, 2004; ALPERS & ESPEDAL, 2004), onde λ_B corresponde ao comprimento de onda proveniente das ondas oceânicas, λ_r corresponde ao comprimento de onda do pulso de RADAR e θ corresponde ao ângulo de incidência. A Figura 2.1 ilustra o modelo de espalhamento Bragg na superfície do oceano, relacionando λ_B , λ_r e θ .

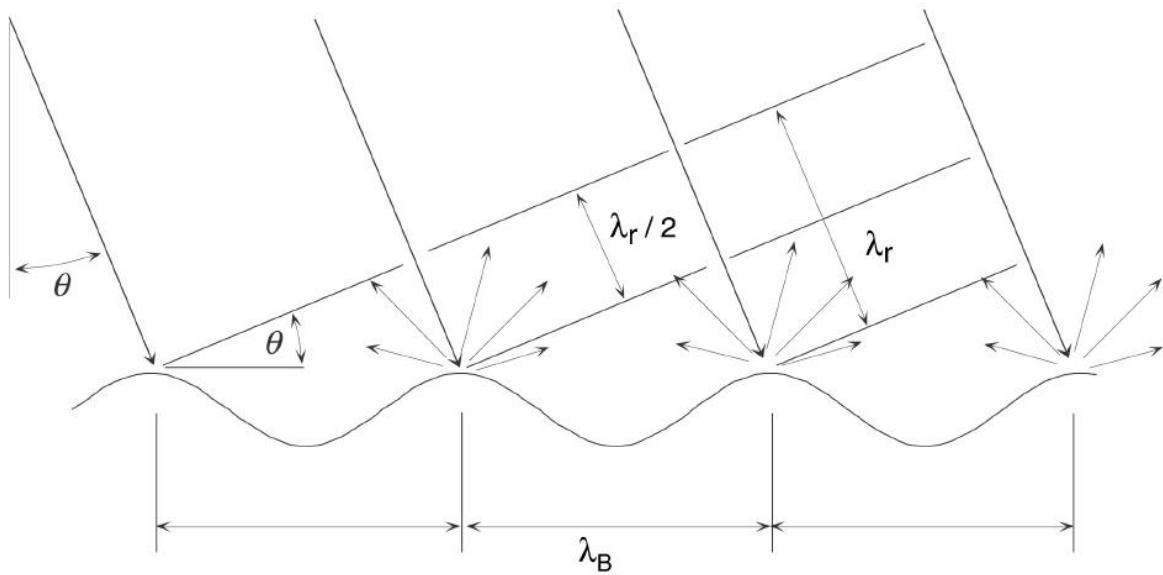


Figura 2.1 - Modelo de espalhamento Bragg na superfície do oceano, relacionando λ_B , λ_r e θ .
FONTE: HOLT, 2004.

A presença de óleo na superfície do mar atenua as ondas capilares de gravidade em função do grau de viscosidade e elasticidade presente nos hidrocarbonetos (MARANGONI, 1872; ALPERS & HUHNERFUSS, 1988; ALPER *et al.*, 1988; ALPERS & HUHNERFUSS, 1989; ESPEDAL & WAHL, 1999; CALABRESI *et al.*, 1999; DEL FRATE *et al.*, 2000; GASULL *et al.*, 2002; MERCIER & ARDHUIN, 2005a e b.; ANGIULI *et al.*, 2006; ALPERS *et al.*, 1991; TOPOUZELIS *et al.*, 2004; MIGLIACCIO & TRANFLAGLIA, 2004; BREKKE & SOLBERG, 2005; ZHENG *et al.*, 2005; MONTALI *et al.*, 2006; BENTZ *et al.*, 2006; FERRARO *et al.*, 2007; SOLBERG & BREKKE, 2008; TOPOUZELIS, 2008). WISMANN *et al.* (1998) evidenciaram que a taxa de atenuação das ondas capilares aumenta com o aumento da viscosidade do óleo e da espessura da camada de óleo.

Este mecanismo de atenuação ocorre de maneira eficiente porque o óleo sobre a água reduz a interface de interação existente entre o ar e a água (MERCIER *et al.*, 2003), provocando uma resistência local à ação do vento, capaz de reduzir a formação e a propagação das ondas capilares (MERCIER & ARDHUIN, 2005a e b). Portanto, este aumento de estabilidade local na superfície do mar provocado pela presença de óleo reduz a amplitude das cristas das ondas (ALMEIDA *et al.*, 2005), diminuindo a rugosidade na superfície do mar.

Superfícies menos rugosas reduzem o coeficiente de retroespalhamento, refletindo a maior parte da energia incidente na direção oposta à antena do RADAR. Este tipo de espalhamento, denominado especular, produz regiões escuras com baixo retroespalhamento nas imagens, as quais podem ser associadas à presença de óleo

(GADE *et al.*, 1998; MERCIER *et al.*, 2003; TOPOUZELIS *et al.*, 2004; BREKKE & SOLBERG, 2005; ZHENG *et al.*, 2005; MERCIER & ARDHUIN, 2005a; ANGIULI *et al.*, 2006; BENTZ, 2006).

O contraste existente entre uma menor rugosidade presente nas manchas de óleo e uma maior rugosidade relacionada ao oceano possibilita a detecção de manchas de óleo no RADAR (TOPOUZELIS *et al.*, 2004; ZHENG *et al.*, 2005; BREKKE & SOLBERG, 2005; RAMALHO & MEDEIROS, 2007; AKKARTAL & SUNAR, 2008). Superfícies rugosas espalham a energia em várias direções, comportando-se como refletores difusos, produzindo como resultado tons mais claros nas imagens devido a um maior retroespalhamento do sinal emitido em direção à antena (LEWIS & HENDERSON, 1998).

A detecção de óleo na superfície do mar é uma tarefa complexa devido à presença de inúmeros fenômenos meteo-oceanográficos que, da mesma forma que o óleo, atenuam as ondas capilares superficiais originando regiões escuras nas imagens de RADAR (BJERDE *et al.*, 1993; SLOGGETT & JORY, 1994; HOVLAND *et al.*, 1994; ESPEDAL & WAHL, 1999; DEL FRATE *et al.*, 2000; MERCIER *et al.*, 2003; SOLBERG, 2005b; BREKKE & SOLBERG, 2005; ZHENG *et al.*, 2005; BENTZ *et al.*, 2005a; BENTZ *et al.*, 2006; MONTALI *et al.*, 2006).

Fenômenos como florescimento de algas, óleos biogênicos, zonas de ressurgência, áreas com baixa intensidade de vento, sedimentos em suspensão, células de chuva, dentre outros, constituem exemplos de feições passíveis de serem interpretadas erroneamente como óleo (CALABRESI *et al.*, 1999; GASULL *et al.*, 2002; ALPERS & ESPEDAL, 2004; TOPOUZELIS *et al.*, 2004; BENTZ *et al.*, 2005a; BENTZ *et al.*, 2005b; BENTZ *et al.*, 2006; ANGIULI *et al.*, 2006). Estes fenômenos são denominados assemelhados por serem facilmente confundidos com óleo nas imagens de RADAR, resultando em falsos alarmes. Outros tipos de feições meteo-oceanográficas detectáveis em imagens SAR, que podem auxiliar nas operações de resposta em emergências ambientais, são: a) vórtices oceânicos; b) frentes de vento; c) frentes oceânicas; d) ondas superficiais; e) ondas internas, e; f) ondas de gravidade atmosférica.

A Figura 2.2 ilustra exemplos de alvos escuros freqüentemente detectáveis em imagens de RADAR, juntamente com a representação esquemática do mecanismo de imageamento do RADAR.

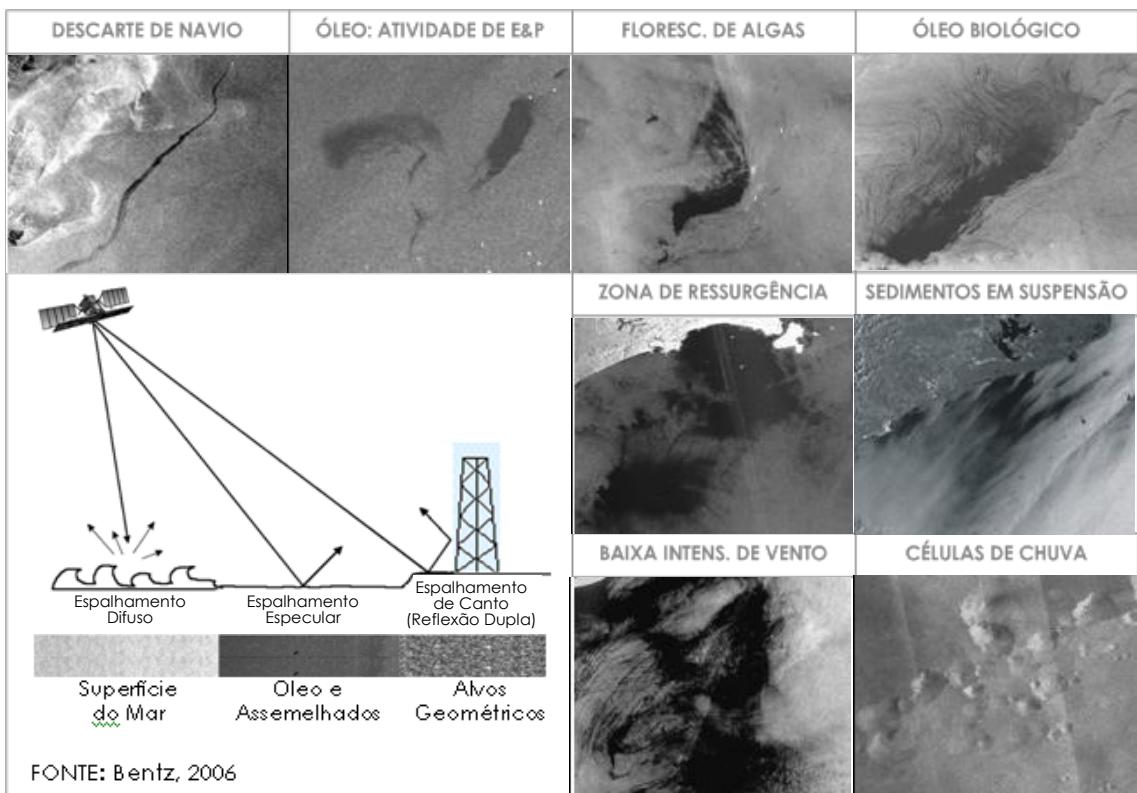


Figura 2.2 – Tipos de esparlhamento produzidos pela interação do pulso emitido pelo RADAR com os alvos presentes na superfície oceânica, juntamente com os alvos escuros freqüentemente detectáveis nas imagens. ADAPTAÇÃO: FONTE - BENTZ, 2006.

Além do óleo, é possível detectar alvos geométricos na superfície do mar em imagens de RADAR, tais como embarcações e plataformas possibilitando, em determinadas situações, identificar ou auxiliar na identificação de potenciais fontes causadoras de vazamentos. Alvos geométricos podem formar ângulos retos voltados para o RADAR produzindo esparlhamento direto ou indireto através de refletores do tipo diédrico ou triédrico, denominados refletores de canto (Figura 2.2). Estes refletores provocam um intenso retorno do sinal retroespalhado para o sensor, produzindo como resposta áreas brilhantes nas imagens de RADAR. A Figura 2.2 ilustra diferentes tipos de esparlhamento produzidos pela interação do pulso emitido pelo RADAR com os alvos presentes na superfície oceânica (BENTZ, 2006; LEWIS & HENDERSON, 1998).

2.2 - Fatores de Influência na Detecção de Alvos Escuros em Imagens SAR

Apesar das imagens de RADAR serem amplamente utilizadas para o monitoramento oceânico (CHANG et al., 2005; BREKKE & SOLBERG, 2005; ANGIULI et al., 2006), a detectabilidade dos alvos escuros depende do seu contraste com o oceano (GASULL et al., 2002). Este contraste é resultado da complexa interação existente entre o pulso de energia eletromagnética emitida por um SAR e a superfície

do oceano, sendo influenciado por diversos fatores como: a) propriedades elétricas da superfície imageada e o comprimento de onda do sinal emitido pelo RADAR; b) variação da intensidade do vento; c) geometria de imageamento; d) tipo de polarização, e; e) o ruído *speckle* (SOLBERG & VOLDEN, 1997; MIGLIACCIO & TRANFAGLIA, 2004; MACCANDLESS & JACKSON, 2004; HOLT, 2004; BREKKE & SOLBERG, 2005; ZHENG *et al.*, 2005; SOLBERG, 2005a; MONTALI *et al.*, 2006; BENTZ, 2006; <http://southport.jpl.nasa.gov/>, 2007).

Propriedades elétricas do alvo e comprimento de onda do RADAR: De acordo com ESPEDAL (1998), LEWIS & HENDERSON (1998) e RANEY (1998) RADARes são instrumentos eficientes para detecção de alvos escuros na superfície do mar devido à alta constante dielétrica da água, que torna a penetração do pulso de RADAR em áreas oceânicas desprezível. A maioria dos RADARes de sensoriamento remoto opera numa faixa de comprimentos de onda que varia de 0,5 cm a 75 cm. Estas freqüências são distribuídas em faixas específicas denominadas Bandas (Figura 2.3), que vão da Banda K, dotada de menor comprimento de onda até a P com maior comprimento de onda, para o caso dos RADARes imageadores (DUTRA *et al.*, 2003; GlobSAR2/CD-ROM).

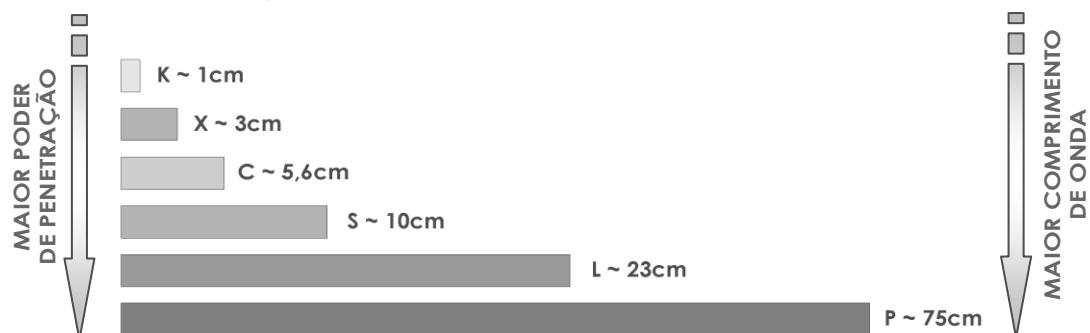


Figura 2.3 – Tamanho Relativo dos Comprimentos de Onda na Faixa de Microondas. Fonte: Material Educativo para Sensoriamento Remoto por Radar GlobeSAR2/CD-ROM.

Apesar da constante dielétrica da água tornar a penetração do pulso eletromagnético desprezível, a Banda X é citada como a mais indicada para a detecção de óleo devido ao seu menor poder de penetração, melhorando a detecção de alvos na superfície do mar (BREKKE & SOLBERG, 2005). Entretanto, satélites como o RADARSAT-1 e o ENVISAT/ASAR, ambos operando na Banda C, têm sido amplamente utilizados com eficiência operacional satisfatória em sistemas de monitoramento de óleo e ocorrências ambientais (BREKKE & SOLBERG, 2005).

Intensidade do Vento: Baixas intensidades de vento permitem a detecção de assemelhados nas imagens de RADAR, aumentando a probabilidade de ocorrência de falsos alarmes (SOLBERG & VOLDEN, 1997). Todavia, existe um limiar máximo e

mínimo a partir do qual o óleo e os assemelhados podem se tornar indetectáveis na superfície do mar (TOPOUZELIS *et al.*, 2004; ALPERS & ESPEDAL, 2004). Isto ocorre porque a intensidade do vento influencia o nível de retroespalhamento e consequentemente a visualização de feições no oceano (BREKKE & SOLBERG, 2005). Velocidades de vento abaixo de 2m/s ou 3m/s não produzem rugosidade suficiente para a detecção de vazamentos de óleo no mar (WISMANN, 1993; SLOGGETT & JORY, 1994; ESPEDAL, 1998; TOPOUZELIS *et al.*, 2004; SOLBERG, 2005a; ZHENG *et al.*, 2005; MONTALI *et al.*, 2006; ANGIULI *et al.*, 2006; FERRARO *et al.*, 2007). Altas velocidades de vento também dificultam significativamente a detecção de óleo. De acordo com a literatura, ventos acima de 12m/s a 14m/s são capazes de dispersar mecanicamente as manchas e aumentar significativamente a rugosidade na superfície do oceano, criando espalhamento suficiente para bloquear a detecção dos alvos escuros (SLOGGETT & JORY, 1994; TOPOUZELIS *et al.*, 2004; ZHENG *et al.*, 2005; MERCIER & ARDHUN, 2005a e b).

Geometria de Imageamento: O coeficiente de retroespalhamento do RADAR também varia em função da geometria de imageamento utilizada, diminuindo com o aumento do ângulo de incidência (MONTALI *et al.*, 2006; BREKKE & SOLBERG, 2005; MIRANDA *et al.*, 2004; SOLBERG, 2005a; BENTZ, 2006). Como consequência, as imagens tendem a apresentar tons mais claros no alcance próximo (*near range*), do que no alcance distante (*far range*). Quanto maior for o ângulo de incidência e a largura da faixa imageada maior a dificuldade para detecção de alvos com baixo sinal de retorno no alcance distante (DUTRA *et al.*, 2003; BENTZ, 2006; <http://southport.jpl.nasa.gov/>, 2007). A Figura 2.4 ilustra a geometria de imageamento do SAR e o efeito do ângulo de incidência no *near* e no *far range*: Imagem RADARSAT-1.

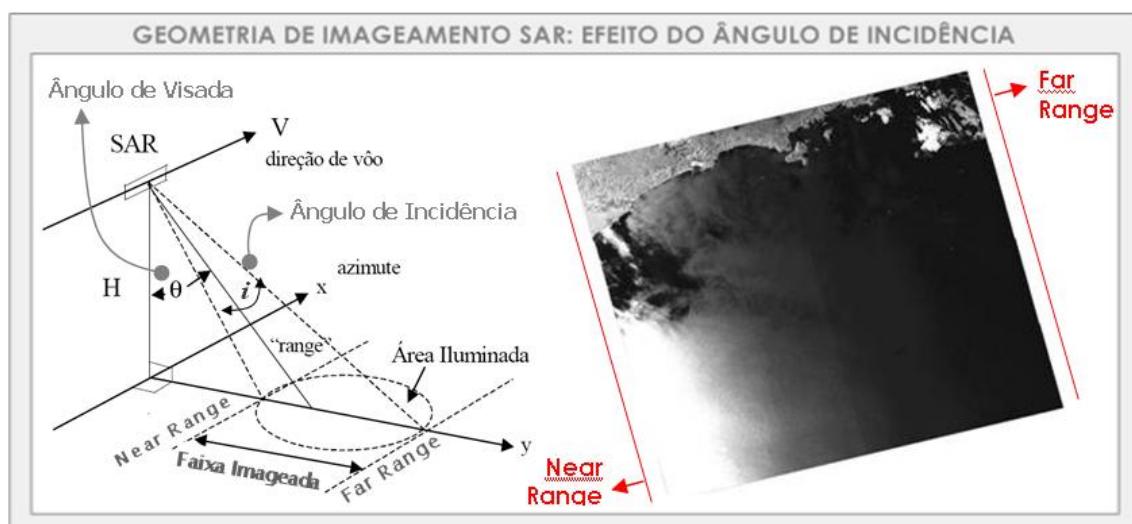


Figura 2.4 - Geometria de imageamento do SAR e o efeito do ângulo de incidência no *near* e no *far range*: Imagem RADARSAT-1.

Polarização da Antena: A definição do tipo de polarização a ser utilizada para a emissão e para a recepção do pulso de microondas pela antena do RADAR também influencia o retroespalhamento (ALPERS & ESPEDAL, 2004; MONTALI *et al.*, 2006; BENTZ, 2006; <http://southport.jpl.nasa.gov/>, 2007). A polarização é definida como a orientação segundo a qual o vetor campo elétrico da onda eletromagnética oscila no tempo. Esta orientação pode determinar polarizações do tipo vertical (V) ou horizontal (H) emitidas de maneira linear, circular ou elíptica. Usualmente, os sistemas de radar utilizam polarizações lineares paralelas, HH³ e VV, ou cruzadas, HV e VH (DUTRA *et al.*, 2003). Embora o tipo de polarização HH, utilizado no sistema RADARSAT-1, seja eficiente para a detecção de alvos escuros (BENTZ & MIRANDA, 2001; BENTZ, 2006), a polarização VV é indicada na literatura como a mais adequada para a detecção de variações na rugosidade da superfície do mar, tais como as provocadas pela presença de óleo (GIRARD-ARDHUIN *et al.*, 2003; BREKKE & SOLBERG, 2005; BENTZ, 2006; SCHMULLIUS & EVASN, 1997). De acordo com ALPERS & ESPEDAL (2004) isto ocorre porque a polarização VV provê um maior contraste entre a mancha e o oceano.

Ruído Speckle: O principal fator de interferência associado à interpretação de imagens de RADAR é a presença do ruído *speckle* (ARVELYNA *et al.*, 2001; GASULL *et al.*, 2002; GALLAND *et al.*, 2004; ZHENG *et al.*, 2005; BREKKE & SOLBERG, 2005; LOPEZ *et al.*, 2005; TELLO *et al.*, 2006; CHANG *et al.*, 2008). Este ruído é inherente ao sistema, sendo causado por interferências locais construtivas e destrutivas (Figura 2.5) no sinal de retorno retroespalhado pela superfície do mar (BENTZ, 2006; GlobSAR2/CD-ROM). O *speckle* gera na imagem uma aparência granular, composta por mudanças bruscas no nível de cinza em *pixels* espacialmente próximos (LEE & HOLT, 1982; DUTRA *et al.*, 2003; BENTZ, 2006). Desta forma, o *speckle* é uma forma de ruído que degrada a qualidade das imagens, interferindo em todas as etapas de detecção dos alvos escuros, sendo recomendado a utilização do processamento em múltiplas visadas (*multi-look*) ou a aplicação de filtros para sua suavização (ARVELYNA *et al.*, 2001; GALLAND *et al.*, 2004; AYED *et al.*, 2006; CHANG *et al.*, 2005; BENTZ, 2006; CCRS, 2007; GlobSAR2/CD-ROM).

³ Na nomenclatura utilizada para definir o sistema de polarização de um sistema RADAR, a primeira letra refere-se à polarização da radiação emitida pela antena, neste exemplo Horizontal, e a segunda letra refere-se à polarização da radiação recebida pela antena, neste exemplo também Horizontal.

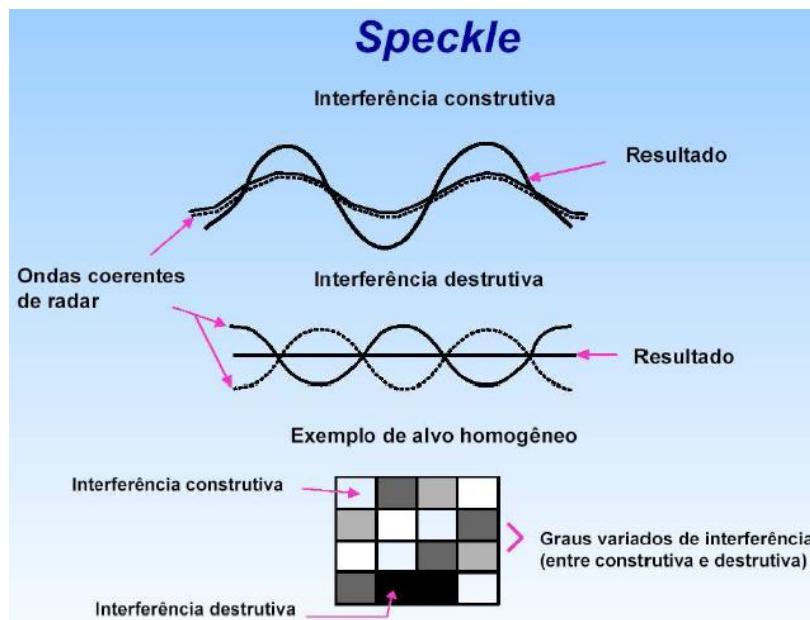


Figura 2.5 – Representação do ruído *speckle*. Fonte: Material Educativo para Sensoriamento Remoto de Radar GlobeSAR2 - CD

Estes inúmeros fatores que influenciam, em maior ou menor escala, no contraste existente entre os alvos escuros e o oceano nas imagens de RADAR fazem com que a detecção dos alvos escuros não seja uma tarefa simples. Especificamente para os casos de vazamento de óleo, o contraste em relação ao oceano também vai variar de acordo com a quantidade e o tipo de óleo vazado, o intervalo de tempo entre o vazamento e o momento da aquisição da imagem, o grau de intemperismo, dentre outros (TOPOUZELIS, 2008; MONTALI *et al.*, 2006; ESPEDAL & WAHL, 1999; SOLBERG & VOLDEN, 1997).

Além destes fatores, imagens utilizadas em diferentes formatos e pré-processadas com diferentes metodologias não retornarão os mesmos resultados para a detecção alvos escuros, bem como para a classificação destes eventos em óleo e assemelhados. De acordo com VACHON *et al.* (1999), até a calibração radiométrica das imagens geradas em diferentes estações de recepção, utilizando diferentes processadores, constitui um fator que interfere no desempenho dos sistemas automáticos para detecção de óleo e assemelhados. Como existem múltiplos fatores envolvidos, a padronização de todos os parâmetros necessários desde o pré-processamento das imagens à etapa final de classificação é essencial para garantir o uso operacional destes sistemas no monitoramento efetivo das regiões costeiras e oceânicas.

CAPÍTULO 3

DETECÇÃO AUTOMÁTICA DE ÓLEO E OCORRÊNCIAS AMBIENTAIS EM IMAGENS SAR

Inúmeros esforços têm sido direcionados para produzir um sistema completamente automático, capaz de reconhecer uma provável mancha de óleo e/ou assemelhados em imagens SAR (BJERDE *et al.*, 1993; SLOGGETT & JORY, 1994; HOVLAND *et al.*, 1994; KUBAT *et al.*, 1998; ESPEDAL & WAHL, 1999; SOLBERG *et al.*, 1999, 2003 e 2005; FISCELLA *et al.*, 2000; GASULL *et al.*, 2002; TOPOUZELIS *et al.*, 2003 e 2004; GALLAND *et al.*, 2004; MIRANDA *et al.*, 2004; BREKKE & SOLBERG, 2005; MERCIER & ARDHUIN, 2005a e b; CHANG *et al.*, 2005; SOLBERG, 2005; KARATHANASSI *et al.*, 2006; ; AYED *et al.*, 2006; KERAMITSOGLOU *et al.*, 2006; MONTALI *et al.*, 2006; BENTZ, 2006; TOPOUZELIS, 2009; HANG *et al.*, 2009, SEASAR, 2010). Estes sistemas são configurados para processar imagens de RADAR, considerando as áreas oceânicas, e emitir um alarme sempre que uma feição for identificada como uma provável mancha de óleo (KERAMITSOGLOU *et al.*, 2006). Esta área do conhecimento reúne problemas complexos, evidenciando a necessidade de considerável investimento em termos de desenvolvimento e pesquisa (TOPOUZELIS *et al.*, 2004; ANGIULI *et al.*, 2006).

Dentre as 4 etapas integrantes dos sistemas para detecção automática de óleo e assemelhados em imagens SAR (Capítulo 1: Figura 1.1), a etapa de detecção dos alvos escuros é a principal etapa da cadeia de processamentos. Sem a detecção e individualização da geometria dos alvos, não há como prosseguir na etapa de classificação dos eventos em óleo ou assemelhados (TOPOUZELIS *et al.*, 2007a; TOPOUZELIS, 2008). Diversas metodologias, combinando técnicas de segmentação e agrupamento, têm sido propostas na literatura, seguindo 2 tipos de abordagem: a) a manual, onde operadores são treinados para interpretar imagens identificando possíveis manchas de óleo, e; b) a semi-automática/automática, onde diferentes técnicas são implementadas através de processamentos automatizados com ou sem a intervenção humana (LUCCA, 1999; TOPOUZELIS, 2008).

As duas abordagens têm como foco detectar regiões dotadas de baixo retroespalhamento nas imagens SAR, as quais podem estar associadas à presença de óleo ou assemelhados (KARATHANASSI *et al.*, 2006; AKKARTAL & SUNAR, 2008; CHANG *et al.*, 2008; TOPOUZELIS, 2008). Na etapa de detecção a preocupação é delimitar, com o máximo de precisão possível, as diferentes geometrias associadas a

qualquer tipo de alvo escuro, sendo a discriminação em óleo ou assemelhado realizada na etapa de classificação.

3.1 – Detecção de Alvos Escuros em Imagens SAR

Alvos escuros são facilmente identificáveis nas imagens SAR. Entretanto, é extremamente difícil, até mesmo para intérpretes experientes, definir os limites, ou seja, as bordas que separam um alvo do oceano (GALLAND *et al.*, 2004; ZHENG *et al.*, 2005). Para auxiliar neste processo, técnicas de segmentação voltadas para a detecção de alvos escuros em imagens SAR têm sido amplamente utilizadas em sistemas automáticos (TOPOUZELIS *et al.*, 2003; PELIZZARI & DIAS, 2003).

A segmentação consiste no particionamento de uma imagem em regiões, considerando critérios de homogeneidade e heterogeneidade (HARALICK & SHAPIRO, 1984; BAATZ *et al.*, 2003; NAVULUR, 2006; NEUBERT *et al.*, 2006; AYED *et al.*, 2006). A partição, ou o agrupamento em regiões de interesse, tem como referência propriedades intrínsecas à imagem como intensidade, contraste ou textura (LE MOIGNE & TILTON, 1995; CAVES *et al.*, 1998; LUCCA, 1999; SOUSA, 2005; NAVULUR, 2006).

Existe uma variedade de métodos automáticos e semi-automáticos para gerar objetos ou regiões em imagens (BAATZ *et al.*, 2003). De maneira geral, os algoritmos de segmentação de imagens podem ser discriminados em 3 categorias (KURZ & BENTEFIFTA, 1997; LUCCA, 1999; PEKKARINEN, 2002; ESPINDOLA, 2006; NAVULUR, 2006): a) ajuste de regiões (*region fitting*); b) detecção de bordas (*edge-based*), e; c) crescimento de regiões (*region-growing*).

Os 3 tipos de algoritmos utilizam os *pixels* da imagem como unidade básica de processamento, a diferença está na forma como os resultados da segmentação são extraídos das imagens e especializados em uma nova imagem.

Os algoritmos de ajuste de regiões, também denominados *pixel-based*, segmentam os objetos de interesse particionando (*splitting*) e agrupando (*merge*) os *pixels* com base em critérios de similaridade ajustados a partir do histograma das imagens. Nos histogramas, altas (picos) e baixas intensidades (vales) dos níveis de cinza (NC) são identificadas e utilizadas para definir agrupamentos nas imagens. A técnica mais comumente utilizada é a limiarização (*threshold*) (NAVULUR, 2006), que consiste no estabelecimento de um limiar (T) para separar os objetos ou regiões de interesse na imagem. De forma simplificada, as operações envolvem testes

estatísticos considerando uma função $T = T[x,y; p(x,y), f(x,y)]$ onde, $f(x,y)$ está relacionado ao NC no ponto (x,y) e $p(x,y)$ representa alguma propriedade local centrada neste ponto, como por exemplo a média dos NC da vizinhança. A imagem limiarizada $g(x,y)$ pode ser definida matematicamente como:

$$g(x,y) = \begin{cases} 1 & \text{se } f(x,y) > T \\ 0 & \text{se } f(x,y) \leq T \end{cases}$$

Desta forma, tendo como base o limiar ou uma função previamente estabelecida com base no histograma das imagens, o processo de segmentação é realizado *pixel a pixel*, retornando uma nova imagem onde cada *pixel* é rotulado como 0 ou 1, separando os alvos de interesse do fundo (*background*) (GONZALES & WINTZ, 1987).

Da mesma forma, os métodos de detecção de bordas operam no nível do *pixel*, retornando uma nova imagem igualmente rotulada em 0 ou 1. Entretanto, a partição é efetuada considerando critérios de dissimilaridade, com o objetivo de evidenciar regiões onde ocorram mudanças abruptas nos NC das imagens, indicando limites ou bordas entre regiões homogêneas com diferentes propriedades (LUCCA, 1999; NAVULUR, 2006). Para efetuar este procedimento é necessário realçar as bordas dos alvos de interesse utilizando filtros como, por exemplo, os de Canny, Robert e Sobel (GONZALES & WINTZ, 1987; LUCCA, 1999; NAVULUR, 2006), aplicados com máscaras de dimensões variadas 2x2, 3x3 e 5x5.

Ambos os métodos de segmentação, ajuste de regiões e detecção de bordas, retornam resultados no nível do *pixel*, sendo um pós-processamento necessário com o objetivo de gerar uma máscara que possibilite o cálculo de atributos a serem utilizados na etapa de classificação em óleo e assemelhados.

De forma diferente dos demais métodos, o crescimento de regiões inicia o processamento no nível dos *pixels* e finaliza retornando regiões resultantes do agrupamento de *pixels* com propriedades similares. O processamento inicia com a escolha de *pixels* de referência denominados *seed points* (BINS *et al.*, 1996; PEKKARINEN, 2002; BAATZ *et al.*, 2003;), a partir dos quais novos *pixels* serão agrupados ou não formando regiões de acordo com um limiar de similaridade e de área mínima previamente estabelecidos. Este agrupamento é efetuado de forma automática até que toda a imagem seja segmentada e não haja mais a possibilidade de fusões em função dos parâmetros estabelecidos pelo usuário no início do processo (BINS *et al.*, 1996; LUCCA, 1999; BAATZ & SCHÄPE, 2000; PINHO, 2005; SOUSA,

2005; NAVULUR, 2006; BENTZ, 2006; ESPINDOLA, 2006).

Nas referências consultadas, voltadas especificamente para a detecção de alvos escuros em imagens de RADAR, pôde-se constatar a utilização dos três tipos de segmentação citados. As pesquisas desenvolvidas, visando à individualização destas feições, abrangem desde a delimitação manual (CALABRESI *et al.*, 1999; FISCELLA *et al.*, 2000) até técnicas automáticas, com ou sem a intervenção humana, com níveis de complexidade diversos (BREKKE & SOLBERG, 2005; BENTZ, 2006; BENTZ *et al.*, 2007b; TOPOUZELIS, 2008).

Como alvos escuros em imagens SAR são caracterizados por baixo retroespalhamento em relação ao oceano, a utilização de técnicas de limiarização (*Threshold*) para a segmentação é comumente observada (BAATZ *et al.*, 2003; BREKKE & SOLBERG, 2005; GASULL *et al.*, 2002). Na literatura são descritos 2 tipos de limiarização: a global (*Threshold*) e a adaptativa (*Adaptative Threshold*). Na abordagem global, o tamanho de uma janela é definido arbitrariamente para calcular ao longo da imagem valores estatísticos que serão comparados a um limiar fixo, previamente estabelecido para a imagem como um todo (KERAMITSOGLOU *et al.*, 2006). Entretanto, como o intervalo que compreende os NC referentes aos alvos escuros não é constante, variando de cena para cena e, inclusive, em uma mesma cena, em função das mudanças na intensidade do vento na superfície do mar, um limiar único, estabelecido para toda a imagem, não é aconselhável (GASULL *et al.*, 2002).

Desta forma, a utilização da limiarização adaptativa (*Adaptative Threshold*) constitui uma alternativa amplamente utilizada em diversos sistemas de detecção automática (TOPOUZELIS, 2008; SOLBERG & BREKKE, 2008; BREKKE & SOLBERG, 2005; SOLBERG *et al.*, 1996). Nesta proposta, um limiar é configurado localmente, tendo como base valores estatísticos calculados considerando uma janela de dimensões pré-definidas, aplicada ao longo da imagem (GASULL *et al.*, 2002). Uma tentativa inicial foi publicada por SKØELV & WAHL (1993) utilizando um algoritmo configurado com base em um histograma simples, bimodal, para uma janela de $N \times N$ pixels ($N = 25$ pixels). Abordagens similares foram descritas por VACHON *et al.* (1998) e MANORE *et al.* (1998). BJORDE *et al.* (1993) primeiramente aplicaram um algoritmo de clusterização para verificar se a área considerada pela janela apresentava características homogêneas, caso o comportamento do histograma fosse bimodal um limiar era calculado e aplicado sobre os *pixels* localizados na janela.

Considerando a significativa influência da intensidade do vento para a detecção de alvos escuros em imagens de RADAR, SOLBERG & VOLDEN (1997), SOLBERG & SOLBERG (1996) e SOLBERG *et al.* (2003) utilizaram um limiar adaptativo configurado como K_{dB} resultante da estimativa de um sinal de retorno médio para uma determinada janela. Neste método, o valor de K é determinado com base nos dados de vento, os quais são configurados manualmente considerando 4 categorias de intensidade: baixo, baixo/médio, médio e alto. Finalmente, uma clusterização foi aplicada sobre os *pixels* rotulados como alvos escuros e oceano para melhorar as bordas e diminuir o ruído da segmentação.

Posteriormente, SOLBERG (2005a e 2005b) e SOLBERG *et al.* (2003) aprimoraram este método de segmentação para ser utilizado quando não existem dados de vento disponíveis. Para tal, um novo método para calcular o limiar utilizando a homogeneidade local, denominado PMR-ratio (*power-to-mean ratio*), foi proposto (SOLBERG *et al.*, 2005; SOLBERG, 2005a; SOLBERG, 2005b). Uma abordagem piramidal, em multi-escala foi utilizada, seguida de uma etapa de clusterização, visando uma melhor separação entre os *pixels* rotulados como alvos escuros e o oceano.

No âmbito da segmentação por ajuste de regiões, existem diversas formas de calcular limiares automaticamente tendo como base o histograma das imagens. FISCELLA *et al.* (2000) utilizou *Normalized RADAR Cross-Section* (NRCS) para estabelecer um limiar a partir do qual os *pixels* considerados escuros na imagem fossem discriminados. Este limiar determina que os *pixels* rotulados como escuros na imagem devem ser menores que a metade do valor Médio obtido pelo NRCS em uma determinada área da imagem. Neste método, pequenas áreas não foram detectadas por serem consideradas irrelevantes e grandes áreas também, por terem a probabilidade de estar associadas a baixas intensidades de vento.

A técnica de morfologia matemática para segmentação de alvos escuros foi utilizada por GASULL *et al.* (2002). Nesta pesquisa foi testada a utilização combinada de filtros morfológicos (*top-hat*) de erosão e dilatação com limiarização adaptativa, alternando janelas de 11x11 para selecionar os alvos maiores e máscara menores de 3x3 para selecionar alvos menores.

A utilização de segmentação por detecção de bordas para extração dos alvos escuros em sistemas de detecção automática de óleo foi primeiramente observada por SLOGGETT & JORY (1994). ESPEDAL (1998) combinou a utilização do filtro Sobel

para detecção de bordas com fatiamento adaptativo, primeiramente utilizado por WEINSTEIN *et al.* (1993) para a individualização das feições.

ARVELYNA *et al.* (2001) utilizaram a técnica de máxima entropia assumindo dois momentos como parâmetro, óleo e mar. De acordo com o autor, o desempenho foi melhor quando comparada com técnicas baseadas na matriz de co-ocorrência dos níveis de cinza da Imagem (*Grey Level Co-occurrence Matrix - GLCM*).

KANAA *et al.* (2003) utilizaram de forma integrada segmentação por detecção de bordas e por ajuste de regiões. Primeiramente suavizaram o ruído das imagens com um filtro Gaussiano, aplicando seqüencialmente o filtro de *Canny* para a detecção de bordas. Posteriormente um limiar DHT (*Directional Hysteresis Threshold*) foi aplicado, considerando um ângulo θ , com o objetivo de realçar as bordas do óleo na mesma direção do vento. Uma operação de clusterização com abordagem probabilística Bayesiana foi efetuada para melhorar os resultados formando agrupamentos mais homogêneos para os *pixels* rotulados como escuros nas imagens.

Ainda no âmbito da detecção de bordas, TELLO *et al.* (2006) utilizaram transformada *Wavelet* para realçar as bordas na imagem, seguida de uma limiarização para separar as bordas dos alvos escuros encontrados do oceano.

BARNI *et al.* (1995) disponibilizaram o primeiro exemplo de utilização de lógica *fuzzy* em um algoritmo para segmentação de alvos escuros em imagens SAR. ZHENG *et al.* (2005) também utilizaram clusterização *fuzzy* para gerar a segmentação. Neste método uma função é definida para mensurar a probabilidade de um determinado *pixel* pertencer ou não a um determinado conjunto de dados, sendo ou não agrupado ao mesmo. A dificuldade encontrada é que o número de clusters nem sempre pode ser definido a priori. Uma vez que a clusterização *fuzzy* é obtida, designando cada *pixel* para um grupo, um mapa dos agrupamentos encontrados é produzido pelo algoritmo. Uma operação final envolve uma pós-filtragem do mapa para obter regiões mais homogêneas.

Ainda no universo do *pixel* alguns autores investem no pré-processamento das imagens, aplicando diferentes filtros ou combinação de imagens de textura, indo direto para a etapa de classificação utilizando algoritmos diversos. MERCIER & ARDHUIN (2005a) aplicaram decomposição *Wavelet* e utilizaram Máquinas de Suporte Vetorial (SVM) supervisionadamente para classificar a imagem em alvos escuros e oceano. TOPOUZELIS *et al.* (2007a) combinou filtro Lee com janela 3x3 e 5x5, normalizou os NC da imagem e posteriormente aplicou uma classificação por Redes Neurais para

separar as regiões associadas aos alvos escuros do oceano. Estas regiões foram vetorizadas, seqüencialmente 10 atributos foram calculados e uma nova classificação, também por Redes Neurais, foi efetuada para classificar os alvos escuros em óleo e assemelhados.

KERAMITSOGLOU *et al.* (2006) e CHANG *et al.* (2005 e 2008) utilizam metodologias que transitam entre o nível do *pixel* e das regiões. Estes autores primeiramente aplicam uma limiarização adaptativa para rotular os pixels em escuros e claros (oceano) e, em uma segunda etapa, utilizam uma operação de interpolação por vizinho mais próximo para gerar regiões relacionadas aos alvos escuros e oceano. A diferença entre estes métodos é que KERAMITSOGLOU utilizou filtro Gaussiano e não detalhou a técnica de limiarização utilizada e CHANG utilizou filtro de Mediana e limiarização pela técnica de *Moment Preserving Method* (MPM). HANG *et al.* (2009), propuseram algo semelhante, entretanto, com a utilização de fatiamento para separar alvos escuros de oceano e, posteriormente, uma vetorização para gerar as regiões. Posteriormente, todos estes autores calcularam atributos e efetuaram classificação em óleo e assemelhados, utilizando técnicas diferenciadas.

MIRANDA *et al.* (2004) aplicaram o filtro de Mediana para a suavização do ruído *speckle*, utilizaram o classificador textural por Semivariograma (STC) para cálculo dos atributos e o algoritmo de clusterização ISODATA para classificar as regiões em óleo e oceano.

Outras técnicas identificadas na literatura para gerar regiões em imagens SAR foram *Minimum Description Length* (MDL) e Campos Markovianos. A utilização da técnica de segmentação por *Minimum Description Length* (MDL), baseada em partições deformáveis, foi proposta por GALLAND *et al.* (2003 e 2004) e consiste na partição da imagem em regiões homogêneas e do agrupamento utilizando limiarização automática MDL (*Minimun Description Length*). A segmentação inicial da imagem é obtida por uma partição deformável, denominada grade ativa (*active grid*). Esta grade contém um conjunto de nós localizados na junção entre os segmentos. O algoritmo procura encontrar o número adequado de nós e regiões minimizando o critério global de MDL. Como, geralmente, a segmentação resultante compõe os alvos com mais de um polígono, uma nova limiarização é utilizada para formar um único polígono para cada alvo escuro da cena.

PELIZARRI & DIAS (2003) desenvolveram um algoritmo supervisionado para segmentação de imagens utilizando como base uma abordagem Bayesiana. A

intensidade dos sinais de retorno observados nas amostras de treinamento para cada classe - alvos escuros e oceano – é modelada através de uma distribuição Gamma (*finite Gamma mixture*), a partir da qual são extraídos a média e a variância. Seqüencialmente, a utilização de *Markov Random Field* (MRF) define uma grade 2d a partir da qual obtém-se a segmentação. A utilização de MRF para segmentação de imagens foi relatada também em LOPEZ *et al.* (2005). Neste trabalho MRF foi utilizado para segmentar uma imagem resultante da fusão de uma imagem original com uma imagem de textura (GLCM: Correlação), gerada a partir da matriz de co-ocorrência dos níveis de cinza.

Um método de segmentação não supervisionado foi descrito por MERCIER *et al.* (2003) utilizando uma extensão vetorial com base no modelo de *Hidden Markov Chain* (HMC), onde a estimativa dos parâmetros é efetuada através do método *General Iterative Conditional Estimation* (ICE). ARDHUIN *et al.* (2003 e 2004) foram os primeiros trabalhos encontrados na literatura que demonstram uma preocupação com o problema das bordas que separam um alvo escuro do oceano. Nesta abordagem ARDHUIN *et al.* compararam o desempenho da utilização de 2 métodos: a) filtro Sobel; b) Cadeias Markov com modelo de distribuição Pearson, e; c) Cadeias Markov com modelo de distribuição Gaussiana. Entretanto, para a detecção por Cadeias de Markov utilizou ao invés de 2 classes, como é comumente considerado na literatura, 3 e 4 classes, com o objetivo de evidenciar que os resultados da segmentação podem representar melhor a transição entre o óleo (alvo escuro) e o oceano com um número maior de classes.

No âmbito dos segmentadores por crescimento de regiões, DEL FRATE *et al.* (2000) e MONTALI *et al.* (2006) mesclaram as técnicas de ajuste de regiões e crescimento de regiões. Inicialmente os *pixels* da imagem foram rotulados como escuros (alvos escuros) ou claros (oceano) através de um limiar calculado com base no histograma da imagem. Numa segunda etapa, os *pixels* mais escuros foram escolhidos como *pixels* sementes a partir dos quais o algoritmo de crescimento de regiões iniciou o processo de segmentação. A classificação foi realizada utilizando redes neurais tendo como base 11 atributos. MONTALI *et al.* (2006) utilizou a mesma lógica de processamento, entretanto cálculo u o limiar utilizando NRC para selecionar os *pixels* sementes e testou diversos classificadores, também com 11 atributos, sendo o que retornou melhor resultado foi o de discriminante linear.

KARATHANASSI *et al.* (2006) desenvolveram uma metodologia para detectar alvos escuros em imagens de RADAR com alta resolução espacial (10m), utilizando

segmentação por crescimento de regiões com uma abordagem em multiresolução. Na primeira etapa, a imagem é segmentada com o objetivo de obter segmentos maiores. Nesta abordagem, como o segmentador cresce as regiões a partir dos *pixels* dividindo a imagem em segmentos teoricamente homogêneos, a probabilidade de incluir um alvo escuro dentro de um segmento associado ao oceano é menor, quando comparada às técnicas de limiarização adaptativa que utilizam janelas com dimensões fixas. Nesta menor escala um limiar é calculado fornecendo informações sobre a distribuição estatística de cada região individualizada. Na segunda etapa, uma nova segmentação é efetuada tendo como base a primeira, com o objetivo de gerar regiões menores. Para estas novas regiões um novo limiar é estimado, comparado com o limiar do super-segmento ao qual pertence, e utilizado para separar os alvos escuros do oceano. Finalmente, um conjunto de atributos foram calculados e utilizados como dados de entrada em um classificador *Fuzzy* para discriminar óleo de assemelhados. Embora os resultados tenham sido satisfatórios KARATHANASSI destaca que o método não é operacional em função do tempo necessário para o processamento (2h). O mesmo método foi utilizado por TOPOUZELIS *et al.* (2007b) e TOPOUZELIS *et al.* (2009), sendo que nesta publicação os autores testaram redes neurais para a classificação.

BENTZ (2006) também utilizou segmentação por crescimento de regiões (BAATZ & SCHÄPE, 1999), entretanto optou por agrupar as regiões resultantes da segmentação manualmente, ao invés de utilizar um limiar adaptativo. Esta escolha foi efetuada com o objetivo de fornecer uma maior acurácia à geometria dos alvos escuros resultantes da segmentação. Como o foco do seu estudo era desenvolver um sistema automático para classificação de óleo e assemelhados, considerando suas subclasses, uma maior acurácia associada à geometria dos alvos escuros foi essencial, pois todos os atributos utilizados na classificação foram calculados com base nestas geometrias. A investigação de um método automático para separar os alvos escuros do oceano, a ser acoplado no sistema de classificação proposto por BENTZ (2006), constitui o principal objetivo da tese.

Conforme evidenciado na bibliografia consultada, as combinações que podem ser efetuadas utilizando diferentes algoritmos, em cada uma das 4 etapas indicadas, são ilimitadas. A Figura 3.1 disponibiliza um fluxograma geral, sintetizando os diferentes métodos manual, semi-automático/automático, efetuados no nível do *pixel* e das regiões, encontrados na literatura.

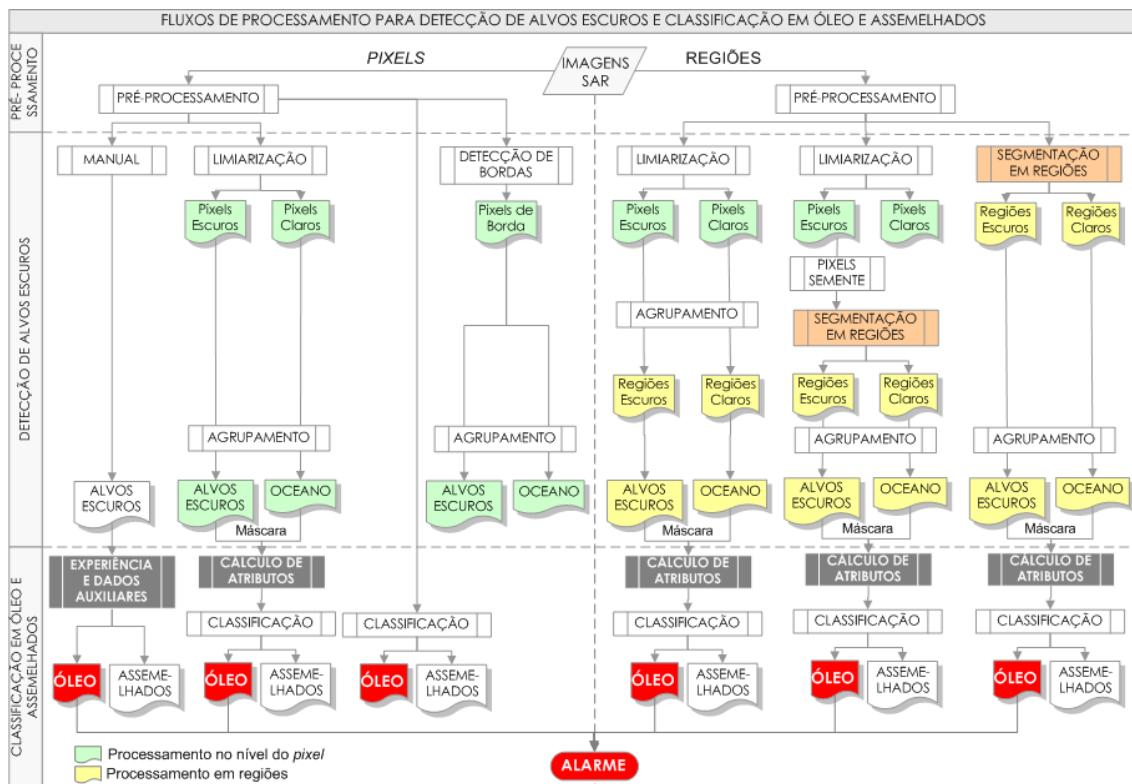


Figura 3.1 - Diferentes tipos de processamento efetuados no nível do *pixel* e das regiões encontrados na literatura.

Apesar da diversidade e relevância das técnicas de segmentação para a melhoria da eficiência dos sistemas de detecção automática, é importante destacar que não existe um algoritmo que apresente desempenho plenamente satisfatório para todos os tipos de imagens, assim como nem todos os algoritmos disponíveis apresentam o mesmo grau de eficiência para uma mesma imagem (BEAUCHEMIN & THOMSON, 1997; BEAULIEU & TOUZI, 2002; SOUSA, 2005; NAVULUR, 2006; ESPINDOLA, 2006).

Desta forma, as diferentes técnicas de segmentação utilizadas, ajuste de regiões, detecção de bordas e crescimento de regiões, apresentam vantagens e desvantagens associadas ao processo de detecção de alvos escuros.

Neste sentido, apesar do delineamento manual dos alvos escuros ser a técnica mais popular e facilmente replicável, a dependência da experiência do intérprete agrava subjetividade ao processo. A “segmentação” manual não é única e um mesmo intérprete pode gerar resultados diferentes para uma mesma imagem (LUCCA, 1999; TOPOUZELIS *et al.*, 2008; TOPOUZELIS, 2009).

No caso da segmentação por ajuste de regiões, há uma boa eficiência computacional. Entretanto, a abordagem por limiarização é sensível à

heterogeneidade dos dados, não sendo plenamente satisfatória nas situações em que o histograma não tem um comportamento tipicamente bimodal (CHANG *et al.*, 2008). Esta variação dos histogramas ocorre porque o comportamento estatístico dos NC nas imagens não é linear e varia e função de fatores como o ruído *speckle*, o ângulo de incidência, dentre outros descritos no Capítulo 2 (SOLBERG & SOLBERG, 1996; ARVELYNA *et al.*, 2001; TOPOUZELIS *et al.*, 2007a). Como os resultados da segmentação são diretamente influenciados pelo limiar estabelecido, um esforço é necessário na tentativa de ajustar limiares satisfatórios (CHANG *et al.*, 2008). Outra desvantagem que os métodos limiarização adaptativa apresentam está relacionada ao tamanho fixo da janela que percorre a imagem (KARATHANASSI *et. al.*, 2006). É possível que um alvo escuro esteja parcialmente contido em diferentes janelas, sendo fragmentado com diferentes tamanhos. Neste caso, as estatísticas calculadas por janela retornarão valores diferentes para caracterizar um mesmo alvo escuro (KARATHANASSI *et. al.*, 2006). Grandes janelas abrangendo regiões oceânicas com homogeneidade nos NC podem deixar de detectar alvos escuros de pequena dimensão (LUCCA. 1999). A detecção por limiar também falha quando o contraste entre os alvos e o oceano é muito fraco (TOPOUZELIZ *et al.*, 2007b). SOLBERG & SOLBERG (1996) evidenciaram 3 situações em que a limiarização não foi efetiva para detectar os alvos escuros: a) manchas de óleo finas e alongadas; b) baixo contraste das manchas em relação ao oceano homogêneo, e; c) manchas de óleo localizadas em oceanos com alta heterogeneidade espectral, dificultando a separação das bordas.

A desvantagem associada ao método de detecção de bordas é que nem sempre é possível obter bordas ideais para imagens com moderada complexidade. Bordas extraídas para imagens complexas, como é o caso das imagens SAR que apresentam o ruído *speckle*, podem ser muito fragmentadas, contendo trechos desconectados entre si, segmentos de borda ausentes, bem como falsas bordas, fatores que complicam a interpretação da imagem. Outra desvantagem está associada à necessidade de uma limiarização após a detecção das bordas para selecionar as bordas referentes aos alvos de interesse. Nesse processo, quanto menor for o limiar, mais bordas serão detectadas, sendo o resultado muito suscetível ao ruído, selecionando feições irrelevantes. De forma contrária, valores de limiar muito altos podem não detectar trechos de borda importantes para compor a geometria de um determinado alvo escuro, não sendo fácil a tarefa de encontrar um equilíbrio entre estas duas situações.

Considerando os algoritmos por crescimento de regiões, a fragilidade está associada à definição do critério de limiar de similaridade e à seleção dos *pixels* semente (*seed points*) que irão iniciar o processo de segmentação (LUCCA, 1999). A agregação dos *pixels* às regiões ou segmentos ocorrerá de acordo com o limiar de similaridade estabelecido. Se este for muito alto poderá ocorrer uma sub-segmentação da imagem e alvos de interesse não serão detectados, de forma contrária, limiares muito baixos super-segmentam a imagem. No caso das imagens de RADAR em que a transição entre um alvo escuro e o oceano não é bem definida, sendo formada por uma graduação de NC, a seleção de *pixels* semente nestas áreas da imagem pode não retornar bons resultados. Desta forma, a tarefa de fixar parâmetros que possam ser utilizados em diferentes imagens não é trivial.

Apesar das desvantagens associadas a cada método, inúmeros autores ressaltam as vantagens de se trabalhar com regiões ao invés de *pixels*. Especialmente no caso das imagens de RADAR, cujo ruído multiplicativo é uma característica inerente deste sensor, o agrupamento dos *pixels* em regiões homogêneas utilizando técnicas de segmentação é capaz de reduzir a interferência deste ruído no processo de detecção de alvos escuros (GALLAND *et al.*, 2003 e 2004; CHANG *et al.*, 2005). Além de melhorar a qualidade dos resultados, a utilização de segmentos, ou regiões, como unidade base de processamento reduz significativamente o custo computacional em relação aos processamentos efetuados *pixel* a *pixel* (BAATZ *et al.*, 2003).

O método de crescimento de regiões é considerado mais eficiente devido à sua capacidade de garantir a formação de segmentos fechados e homogêneos, considerando a relação espacial existente entre os dados (ZUCKER, 1976; TILTON & LAWRENCE, 2000; MEINEL & NEUBERT, 2004; SILVA *et al.*, 2005; ESPINDOLA, 2006; SILVA, 2006). GALLAND *et al.* (2004) evidenciaram as diferenças obtidas na detecção de alvos escuros em imagens SAR utilizando o agrupamento de regiões ao invés de *pixels* previamente rotulados por limiarização. A Figura 3.2 disponibiliza os resultados obtidos por GALLAND *et al.* (2004).

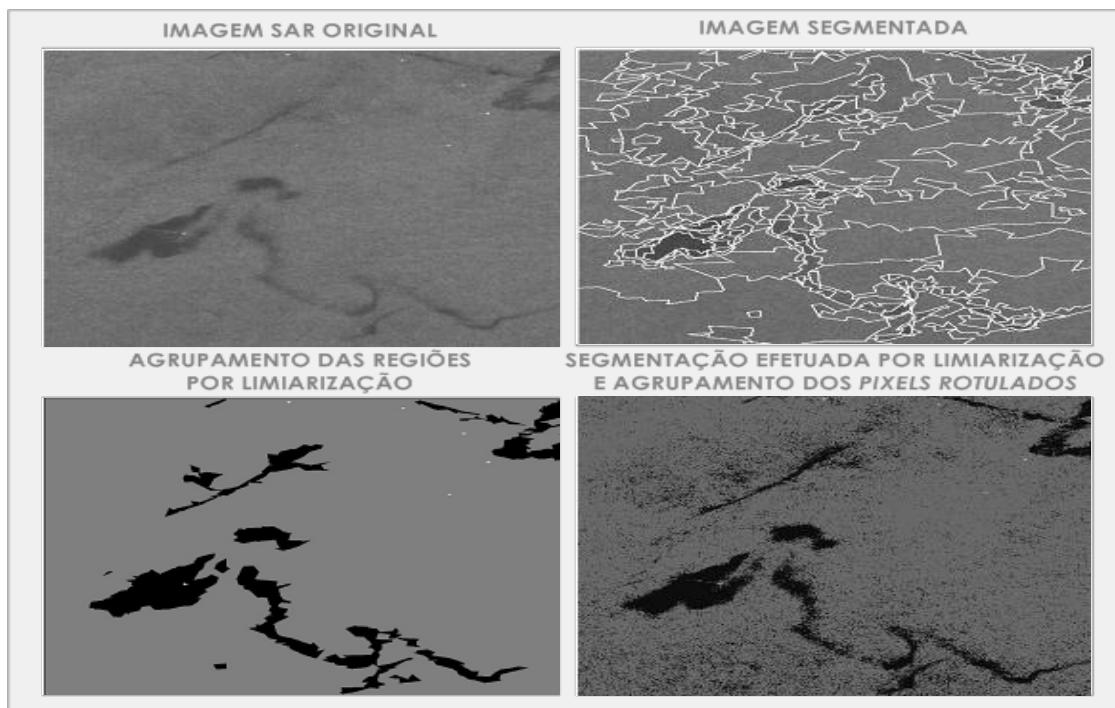


Figura 3.2 – Comparação entre o agrupamento efetuado a partir de regiões e de *pixels*.
FONTE: GALLAND *et al.* (2004)

Com base nestas constatações, no âmbito da tese optou-se pela utilização de algoritmos de segmentação, utilizando a abordagem de crescimento de regiões. Apesar das vantagens de se trabalhar com regiões ao invés de *pixels*, imagens SAR em áreas oceânicas são cenários imprevisíveis, dotados de complexidade variada, onde diferentes tipos de alvos podem ocorrer em diferentes escalas, características que dificultam a delimitação da geometria utilizando apenas a segmentação de imagens. Mesmo concentrando o processamento em uma única mancha de óleo, a segmentação pode partitioná-la em mais de uma região de acordo com os padrões de homogeneidade/heterogeneidade dos NC que a compõe (GALLAND *et al.*, 2003 e 2004; ZHENG *et al.*, 2005), sendo necessário um pós-processamento para agrupar estas regiões em alvos escuros e oceano (GALLAND *et al.*, 2004).

Desta forma, embora a escolha de um método adequado de segmentação seja uma tarefa de relevante importância, o foco da tese foi a investigação de técnicas adequadas para agrupar estas regiões ou segmentos em alvos escuros representativos.

Considerando o estado da arte referente aos métodos propostos para a detecção de alvos escuros, o fluxograma do APÊNDICE I disponibilizou uma síntese da diversidade de métodos e combinações de algoritmos citados na revisão bibliográfica, considerando como referência as 4 etapas integrantes dos sistemas

automáticos para detecção de óleo e assemelhados em imagens SAR. Neste fluxograma é possível comparar o método proposto na tese para a etapa de detecção de alvos escuros, identificando as inovações propostas em relação aos demais métodos.

Desta forma, conforme evidenciado nos fluxogramas da Figura 3.1 e do APÊNDICE I, nos métodos testados até o momento a detecção de alvos escuros é dividida em 2 etapas: a) segmentação das imagens, utilizando um dos 3 tipos de técnicas citadas, e; b) agrupamento dos *pixels* ou das regiões para formar geometrias mais representativas e compactas para os alvos de interesse. A proposta trabalhada ao longo da tese está dividida em 4 etapas: a) segmentação das imagens, utilizando algoritmos de crescimento de regiões; b) cálculo de diversos atributos para cada região originada pela segmentação; c) seleção do conjunto adequado de atributos, e; d) clusterização destes atributos em alvos escuros e oceano.

Comparando as diferentes estruturas propostas no fluxograma do APÊNDICE I, é perceptível que o cálculo de atributos foi utilizado somente na etapa de classificação das imagens em óleo e assemelhados. Neste ponto, a possibilidade de calcular múltiplos atributos a partir das regiões (BAATZ *et al.*, 2003) foi mais um fator que fortaleceu a utilização do método de crescimento de regiões.

Outro fator de destaque foi a preocupação com a utilização de um método de seleção de atributos para definir um conjunto adequado de atributos, capaz de auxiliar o clusterizador no reconhecimento de padrões, retornando geometrias o mais representativas possíveis para os alvos escuros. Em praticamente toda a bibliografia consultada não foi verificada a preocupação com a utilização de um método de seleção de atributos, os primeiros trabalhos que evidenciaram estudos para seleção de atributos foram STATHAKIS *et al.* (2006), BENTZ (2006) e BENTZ *et al.* (2007b). TOPOUZELIS (2009) propôs um método interativo, utilizando algoritmos genéticos para selecionar os atributos a serem seqüencialmente processados em uma rede neural para classificação em óleo e assemelhados.

Outra preocupação recorrente ao longo da tese refere-se à possibilidade de padronizar os parâmetros envolvidos em todas as etapas da detecção, deste a utilização dos segmentadores, ao cálculo dos atributos até a clusterização dos dados. Com relação à padronização dos parâmetros, quesito necessário para automatização de um sistema, praticamente nenhuma referência consultada abordou o tema com profundidade.

3.1.1 – Segmentação de Imagens SAR

Existe uma grande variedade de algoritmos de segmentação implementados em diferentes softwares utilizando conceitos diferenciados. Estes algoritmos, comercializados ou distribuídos gratuitamente, oferecem ampla possibilidade de investigação (NEUBERT *et al.*, 2006). Entretanto, pesquisas recentes indicaram que a abordagem por crescimento de regiões é capaz de gerar segmentações de melhor qualidade (ZUCHER, 1976; SILVA, 2006; MEINEL & NEUBERT, 2004; SILVA *et al.*, 2005).. Com base nestas constatações, foram utilizados softwares gratuitos, testados e considerados eficientes, cujos algoritmos de segmentação utilizassem a abordagem por crescimento de regiões (MEINEL & NEUBERT, 2004; NEUBERT *et al.*, 2006). Os segmentadores testados foram:

- a) o segmentador MultiSeg desenvolvido pelo INPE e implementado no software ENVI 4.0. Este algoritmo tem como principal objetivo a segmentação multi-níveis de imagens de RADAR através da abordagem de crescimento de regiões, sendo inicialmente denominado SegSAR (SOUSA, 2005; SOUSA *et al.*, 2005), e;
- b) o segmentador por crescimento de regiões, também desenvolvido pelo INPE (BINS *et al.*, 1996) e implementado no software SPRING (CÂMARA *et al.*, 1996) disponibilizado gratuitamente na *internet*.

3.1.1.1 – MultiSeg: Informações Técnicas

O MultiSeg é um segmentador híbrido que utiliza de forma integrada técnicas de crescimento de regiões, agrupamento de regiões, ajuste de bordas e separação de regiões não homogêneas (SOUSA, 2005; SOUZA *et al.*, 2005; SCOFIELD *et al.*, 2007). Inicialmente o MultiSeg foi concebido para processar dados SAR em intensidade, sendo um módulo voltado para imagens óticas desenvolvido posteriormente.

As imagens podem ser segmentadas por dois modelos de representação, a serem escolhidos: a) o modelo de manchas (*Cartoon*) o qual supõe homogeneidade das regiões na imagem de RADAR, ou; b) o modelo de textura (*Texture*), que supõe a presença de heterogeneidade nas regiões, caracterizada pela variação do retroespalhamento. Os parâmetros que necessitam de definição para o processamento dos dados SAR no MultiSeg são: a) o processamento das imagens em amplitude ou intensidade; b) o modelo a ser utilizado (*Cartoon* ou *Texture*); c) o Limiar

de Similaridade (LS) e o nível de confiança utilizados nos testes de crescimento e agrupamento de regiões, indicando os *pixels* que devem ou não ser agrupados em uma região; d) a área mínima em *pixels* que as regiões podem na atingir no último nível da segmentação; e) o número equivalente de looks (NEL); f) o número de compressões desejado, e; g) coeficiente de variação utilizado no teste de homogeneidade no modelo de textura.

A segmentação inicia com a compressão piramidal das imagens, onde a base da pirâmide (Nível: C1) contém a imagem original sem compressão. Nos níveis subseqüentes a imagem é comprimida por um fator de 4 em cada nível, formando um único *pixel* a partir da média de 4 *pixels* componentes da imagem anterior sendo $C = \{1, \dots, N\}$. O processo de crescimento de regiões começa no nível de maior compressão da imagem, a partir do qual cada *pixel* selecionado aleatoriamente é considerado como uma região. A partir destes *pixels* uma análise de similaridade é efetuada em relação aos seus vizinhos para indicar quais *pixels* devem ou não ser agrupados crescendo a região. A imagem segmentada é obtida ao final dos testes de agrupamento no nível de maior compactação, considerando a resolução mais degradada. O próximo passo é descompactar as regiões formadas na segmentação sucessivamente, onde em cada nível de descompressão as imagens serão 4 vezes maiores em relação ao nível acima. Efetuar a segmentação no último nível de compressão diminui a interferência do ruído *speckle* nos resultados da segmentação, à medida que a degradação da resolução da imagem nos diferentes níveis de compressão funciona como um filtro. Entretanto, as bordas geradas pela segmentação das imagens com menor resolução no maior nível de compressão (por exemplo, C5), são consideradas grosseiras quando comparadas com a imagem em um nível anterior de compressão (por exemplo, C4). Para suavizar este efeito é necessário efetuar um ajuste de bordas. O processo termina quando as descompressões e ajustes de bordas subseqüentes alcançam o nível C1 da imagem original, sendo efetuado o teste final para verificar se foram produzidas regiões menores do que o limiar de área previamente estabelecido.

Este software foi utilizado nas etapas de pré-processamento e determinação dos métodos de filtragem por ser gratuito e específico para imagens de RADAR. Na etapa de avaliação de segmentadores o MultiSeg foi comparado com o segmentador implementado no SPRING.

3.1.1.2 – SPRING: Informações Técnicas

O segmentador implementado no *software* SPRING considera uma abordagem tradicional de crescimento de regiões, com algumas modificações implementadas para amenizar a sensibilidade associada à ordem dos *pixels* escolhidos para iniciar os processamentos (BINS *et al.*, 1996).

O fluxo de processamento efetuado pelo *software* SPRING comprehende as seguintes etapas: a) seleção aleatória inicial de um conjunto de *pixels* $\{R_i, i = 1, \dots, n\}$, denominados sementes, a partir do quais se inicia o processo de crescimento de regiões; b) inicialmente cada um destes *pixels* é considerado uma região, a partir da qual seus *pixels* vizinhos serão avaliados sendo os mais similares agrupados, crescendo a região; c) a média das regiões é recalculada a cada novo *pixel* agrupado; d) os mesmos procedimentos são efetuados seqüencialmente até que todos os *pixels* da imagem estejam agregados em uma região. Os parâmetros a serem definidos pelo usuário são a área mínima e o limiar de similaridade que corresponde a diferença entre as médias das regiões que se deseja separar.

Como o SPRING não é um segmentador programado especificamente para imagens de RADAR, por este motivo a utilização do *software* MultiSeg foi priorizada nas primeiras etapas da tese, sendo o desempenho do SPRING comparado com o do MultiSeg somente na etapa de determinação do melhor segmentador.

3.2 – Clusterização de Dados

Conforme indicado no Capítulo 3: Item 3.1, diversos métodos são utilizados para a individualização da geometria dos alvos escuros em imagens SAR, sendo as técnicas de segmentação de imagens comumente utilizadas para auxiliar nesta tarefa. Após a divisão das imagens em regiões homogêneas, um pós-processamento para agrupar estas regiões em alvos escuros e oceano é necessário. A limiarização adaptativa é a técnica mais comumente utilizada para o agrupamento das regiões resultantes do processo de segmentação (TOPOUZELIS *et al.*, 2007b; GALLAND *et al.*, 2003 e 2004). CHANG *et al.* (2005) e KERAMITSOGLOU *et al.* (2006) utilizaram interpolação por vizinho mais próximo para agrupar os *pixels*; MIRANDA *et al.* (2004) utilizaram o ISODATA para separar em *clusters* os alvos escuros e o oceano. BJERDE *et al.* (1993), GASULL *et al* (2002) e LOPEZ *et al.* (2005) utilizaram o termo clusterização, mas não informaram qual método e qual(is) o(s) atributo(s) utilizado(s).

Considerando as desvantagens associadas ao método de limiarização indicadas no item 3.1 e o objetivo de propor um método para a detecção de alvos escuros que possa ser implementado operacionalmente de forma automática, o estudo realizado investigou a utilização de clusterização de dados para agrupar as regiões resultantes da segmentação de imagens. Esta proposta é interessante à medida que a performance de diferentes atributos na separação dos *clusters* pode ser investigada.

O processo de agrupamento de objetos físicos ou abstratos em classes é denominado clusterização ou classificação não supervisionada (CHEN *et al.*, 1996). Um *cluster* é um agrupamento de objetos que são similares entre si e dissimilares em relação aos objetos contidos em outros clusters (HAN & KAMBER, 2001). O agrupamento destes objetos é efetuado automaticamente a partir do reconhecimento de padrões intrínsecos ao conjunto de dados utilizado. Neste contexto, a seleção de atributos capazes de discriminar os diferentes *clusters* contidos em um conjunto de dados é essencial para melhorar o desempenho da clusterizações no reconhecimento de padrões (JAIN *et al.*, 1999).

Técnicas voltadas para a análise de *cluster* são capazes de fornecer novos significados, novas informações através da partição de um grande conjunto de dados em componentes menores. A clusterização identifica regiões com alta concentração de objetos dotados de propriedades semelhantes com base em alguma medida de distância dentro de um conjunto de dados multidimensional. Esta técnica fornece como resultado novos padrões de distribuição para o conjunto de dados (CHEN *et al.*, 1996) e interessantes correlações existentes entre os dados (HAN & KAMBER, 2001).

Técnicas de clusterização têm sido estudadas em diversas áreas com ênfases diferentes, dentre elas: estatística, aprendizado de máquina, banco de dados espacial, e mineração de dados. Na literatura existe um grande número de algoritmos de clusterização disponíveis. A escolha do algoritmo a ser utilizado dependerá do objetivo proposto e do tipo de dados a ser analisado. De maneira geral, os métodos de clusterização podem ser classificados em 5 categorias (HAN & KAMBER, 2001): a) métodos de partição; b) métodos hierárquicos; c) métodos baseados em densidade; d) métodos baseados em malhas (*grid-based*), e; e) métodos baseados em modelo (*model-based*).

Métodos de partição: Dado um banco de dados com n objetos, um método de partição produz como resultado k partições, onde cada partição representa um *cluster* sendo $k \leq n$. As k partições devem atender requisitos como: (a) cada grupo deve

conter pelo menos 1 objeto, e; (b) cada objeto deve pertencer a um único grupo.

Dado k , número de partições a ser atingido, um método de partição cria uma partição inicial a partir da qual, através de uma técnica de relocação interativa pode melhorar o particionamento através da movimentação de objetos de um grupo para outro. O critério geral a ser avaliado para uma boa partição do espaço de atributos é o de que objetos constituintes de um mesmo *cluster* sejam semelhantes uns aos outros e diferentes de objetos pertencentes a outros *clusters*. Neste contexto existem diversos critérios para avaliar a qualidade das partições.

Dentre os métodos de partição, os mais conhecidos e comumente utilizados, são “*K-Means*” e “*k-Medoids*” (JAIN *et al.*, 1999; ANDRADE, 2004). O método “*K-Means*” é baseado no conceito estatístico do centróide, onde cada cluster é representado pelo valor médio dos objetos que o compõe, esperando que a semelhança entre os objetos dentro do mesmo grupo ou *cluster* seja alta e a semelhança entre objetos de *clusters* diferentes seja baixa. Em contrapartida, o método “*K-Medoids*” busca diminuir a sensibilidade a objetos isolados, deixando de levar em conta o centro do grupo como ponto de referência para utilizar o medóide, que é o objeto localizado mais próximo ao centro do grupo. Assim, este método pode abrandar a influência de objetos isolados, baseando-se no princípio de minimização da diferença, ou baixa semelhança, entre os objetos de um grupo/*cluster* e o seu ponto de referência (ANDRADE, 2004; HAN & KAMBER, 2001).

Métodos hierárquicos: Um método hierárquico funciona através da criação de uma decomposição hierárquica para um dado conjunto de dados. Métodos hierárquicos podem ser classificados como aglomerativos e divisivos dependendo de como a decomposição é realizada. Métodos aglomerativos (*bottom-up*) inicialmente consideram cada objeto separadamente, a partir do qual após fusões sucessivas com objetos semelhantes formam grupos separados. Estas fusões são realizadas até que um determinado critério de parada seja estabelecido. A abordagem divisiva (*top-down*), ao contrário da aglomerativa, inicia considerando todos os objetos como membros de um mesmo *cluster* e após interações sucessivas os *clusters* vão sendo divididos em *clusters* menores até que os limiares estabelecidos sejam atingidos.

Métodos baseados em densidade: Devido à falta de habilidade que os algoritmos baseados em métodos particionais têm em descobrir grupos com formas arbitrárias, isto é, grupos que apresentam objetos dispersos em uma distribuição espacial bastante irregular, foram desenvolvidos métodos de clusterização baseados

em densidade. Estes métodos consideram tipicamente os grupos como regiões densas de objetos separados no espaço por regiões de baixa densidade – que, normalmente, representam algum tipo de ruído (ANDRADE, 2004).

Métodos baseados em malhas (*grid-based*): Estes métodos quantificam o espaço de objetos em um número finito de células formando uma malha de dados com estrutura em múltipla resolução, a partir da qual são efetuadas operações de clusterização. Esta abordagem tem como ganho seu rápido processamento, pois não depende diretamente do número de objetos a serem analisados e sim do número de células em cada dimensão do espaço quantificado.

Métodos baseados em modelo (*model-based*): Os métodos baseados em modelos criam um modelo hipotético para cada grupo desejado e procuram ajustar os dados da melhor maneira ao modelo criado. Algoritmos baseados em modelos são capazes de descobrir os *clusters* através de funções de densidade que refletem a distribuição espacial dos dados. Este método também permite determinar automaticamente o número de *clusters* baseados em padrões estatísticos, considerando a presença de ruídos ou *outliers* - dados fora do padrão – proporcionando o desenvolvimento de algoritmos robustos.

Apesar da variedade de métodos e algoritmos, o processo de clusterização é considerado subjetivo, uma vez que um mesmo conjunto de dados pode ser particionado de forma diferente a partir da utilização de diferentes métodos (JAIN *et al.*, 1999). Em certos casos, um mesmo método aplicado a um mesmo conjunto de dados pode convergir em diferentes soluções de agrupamento ao final das interações. Esta subjetividade é o que torna o processo de clusterização uma tarefa difícil.

Especificamente para o caso da detecção automática de alvos escuros em imagens SAR, os quais podem estar associados a derramamento de óleo, os resultados do agrupamento das regiões resultantes da segmentação em alvos escuros e oceano precisam ser estáveis. Não seria viável utilizar operacionalmente um método de agrupamento que, a cada novo processamento, retornasse uma solução diferente para os *clusters*.

Desta forma, a capacidade de garantir a integridade e a replicabilidade dos resultados, retornando sempre os mesmos agrupamentos toda vez que um processo de clusterização for efetuado a partir de um mesmo conjunto de dados, é essencial para a automatização do sistema. No âmbito da tese, o método de clusterização denominado *Focal Point*, implementado no software *Clustan* (WISHART, 2004) foi

utilizado por oferecer, de forma diferente dos demais métodos ou programas, estabilidade e replicabilidade dos resultados. O item 3.2.1 disponibiliza os detalhes técnicos sobre o software.

3.2.1 – *Clustan*: Informações Técnicas

O algoritmo de clusterização denominado *Focal Point (FP)*, implementado no software *Clustan*, utiliza o método de partição *k-means* para efetuar o agrupamento dos dados (WISHART, 2004). Além de ser considerado clássico na literatura, o método *K-means* possui fácil implementação e baixo custo computacional, sendo comumente utilizado em grandes conjuntos de dados, produzindo resultados satisfatórios (MACQUEEN, 1967; JAIN *et al.*, 1999). Os padrões de similaridade utilizados para separar as regiões em diferentes *clusters* são obtidos a partir de métricas de distância, sendo a Distância Euclidiana utilizada no software. O *k-means* e outros diversos métodos de clusterização são sensíveis à partição inicial, ou seja, a ordem das regiões escolhidas para iniciar o processo de clusterização (JAIN *et al.*, 1999), o que pode retornar diferentes resultados para um mesmo conjunto de dados a cada nova interação. A garantia de retornar sempre um mesmo resultado através de uma série de testes efetuados ao longo das interações constitui o diferencial do *FP*.

No estudo proposto os procedimentos de clusterização foram efetuados a partir de diferentes atributos ($j = \{1, \dots, \eta\}$) calculados para cada região ($i = \{1, \dots, m\}$) resultante das segmentações, com o objetivo de agrupá-las em p clusters ($p = \{1, \dots, q\}$), retornando como resultado uma imagem binária, separando os alvos escuros do oceano.

Para avaliar a qualidade das clusterizações e garantir o retorno do melhor resultado, a métrica utilizada pelo software equivale à média da soma dos quadrados das distâncias euclidianas (*ME*) obtidas para cada *cluster p* (*Ep*). De acordo com WISHART (2004), *Ep* equivale à soma dos quadrados das distâncias de todas as regiões agrupadas no *cluster p* em relação à média deste *cluster*, sendo:

$$(1) \quad Ep = \sum_{i \neq p} \sum_j \frac{(x_{ij} - \mu_{p,j})^2}{\eta}$$

A notação j se refere às variáveis utilizadas na clusterização, sendo x_{ij} o valor de uma variável j associado a uma região i . A média em um determinado cluster p é representada por $\mu_{p,j}$, e η corresponde ao número total de variáveis. Se os *clusters*

foram bem separados, as distâncias associadas a cada região estarão próximas à média do *cluster* onde a região foi agrupada.

Portanto, a convergência do melhor resultado é garantida pela *ME*, representada pela equação (2) $ME = \sum_p E_p$ onde, uma determinada região (*i*) somente será retirada de um *cluster p* e realocada em um *cluster q* quando (3) $E_{p,i} + E_{q+i} < E_p + E_q$. Desta forma, o *FP* inicia a clusterização pelo método *k-means* agrupando interativamente cada região em um determinado *cluster* buscando minimizar o critério *ME* considerando E_p calculado para todos os *clusters* encontrados. Esse processo termina quando qualquer alteração nos membros dos *clusters* não aumenta nem diminui o valor obtido para *ME*, sendo este o teste de relocação efetuado pelo *software* (WISHART, 2004).

Os processamentos efetuados pelo *FP* estão divididos em duas etapas: a) em um primeiro estágio o *software* utiliza todas as variáveis e seleciona as regiões randomicamente para iniciar as clusterizações, retornando ao final um *ranking* com as melhores soluções encontradas considerando *ME*. Neste processo a reproduzibilidade das diferentes soluções encontradas também é considerada. Neste estágio o *software* também identifica valores inexistentes e as substitui automaticamente pela média da variável no referido *cluster*. Algumas regiões podem não ser agrupadas neste estágio da clusterização por estarem se comportando como *outliers*, evitando a distorção dos *clusters*. Estas regiões são identificadas se estiverem a uma distância (d^2) do *cluster* mais próximo, quando comparadas com a maior distância (d) obtida entre uma região e a média do referido *cluster*. Regiões consideradas duvidosas por estarem localizadas exatamente entre 2 *clusters* também podem não ser classificadas nesta etapa; b) em um segundo estágio, as regiões não agrupadas em alvos escuros ou oceano no primeiro estágio são agrupadas nos *clusters* encontrados, considerando as mínimas distâncias em relação a media dos *clusters* definidos no primeiro estágio.

As características que tornam o método *FP* implementado no *software Clustan* diferente dos demais programas de clusterização são (WISHART, 2004): a) a realização do teste de relocação das regiões com base na *ME*, garantindo que a clusterização convergiu para o melhor resultado possível de ser obtido a partir do conjunto de dados considerado; b) este teste também avalia a sensibilidade do clusterizador com a alteração randômica da ordem com que as regiões são inicialmente consideradas na formação dos *clusters*, compilando ao final as melhores soluções; c) a disponibilização de uma lista com os melhores resultados obtidos, onde o usuário decide qual será o modelo escolhido; d) o cálculo da reproduzibilidade de

cada solução encontrada para a clusterização, considerando a freqüência de ocorrência dos *clusters* em relação ao número de tentativas efetuadas; e) a identificação de variáveis e nulas e *outliers*.

3.3 – Avaliação dos Resultados obtidos para Detecção de Alvos Escuros em Imagens SAR

O uso operacional de Sistemas de Detecção Automática de Óleo com base em imagens de RADAR exige uma rápida performance, bem como resultados dotados de alta confiabilidade produzidos com o mínimo de intervenção humana possível (BREKKE & SOLBERG, 2005). Como o grau de confiabilidade é diretamente influenciado pela qualidade da geometria dos alvos escuros extraídos das imagens, os métodos de segmentação, bem como de agrupamento das regiões, possuem importância significativa por exercerem influência nas demais etapas de processamento. Neste sentido, um método eficiente para avaliar qualitativa e quantitativamente os resultados gerados por diferentes procedimentos de segmentação possui importância estratégica para uma implementação eficiente destes sistemas (MEINEL & NEUBERT, 2004; NEUBERT *et al.*, 2006).

De acordo com LUCCA (1999), SOUSA (2005), NEUBERT *et al.* (2006) e SCOFIELD *et al.*, (2007) existe uma diversidade de métodos de avaliação propostos para comparar o desempenho de algoritmos de segmentação em relação às regiões (DELVES *et al.*, 1992, QUEGAN *et al.*, 1988; CAVES *et al.*, 1998;) e em relação às bordas (PRATT, 1991; ROMÁN-ROLDÁN *et al.*, 2001). Grande parte destes métodos adota abordagens quantitativas, basicamente métodos de discrepância empírica, os quais analisam o número de *pixels* erroneamente classificados em relação a uma segmentação de referência (NEUBERT *et al.*, 2006).

Segundo ZHANG (1996), os algoritmos de segmentação podem ser avaliados por métodos analíticos ou métodos empíricos. Os métodos analíticos examinam diretamente o algoritmo de segmentação analisando seus princípios e suas propriedades (ZHANG, 1996; LUCCA *et al.*, 1998; LUCCA, 1999; SOUSA, 2005). Os métodos empíricos - de qualidade e de discrepância - avaliam indiretamente o(s) algoritmo(s) utilizado(s), sendo que: a) os métodos empíricos de qualidade são fundamentados a partir de uma comparação visual entre os produtos de segmentação, considerando a experiência do intérprete na avaliação e definição dos melhores produtos, e; b) os métodos de discrepancia analisam a qualidade do segmentador comparando os resultados de segmentos gerados utilizando diversas métricas em

relação a uma segmentação de referência e indicam a precisão e eficiência do algoritmo (ZHANG, 1996; LUCCA *et al.*, 1998; LUCCA, 1999; SOUSA, 2005; SCOFIELD *et al.*, 2007).

Neste trabalho, os métodos de avaliação qualitativa (métodos empíricos de qualidade) e quantitativa (métodos de discrepância) (DELVES *et al.*, 1992) foram utilizados para avaliar comparativamente imagens binárias contendo, não as regiões resultantes da segmentação, mas os polígonos resultantes do agrupamento destas regiões em alvos escuros e oceano. No âmbito da tese, o termo regiões foi a nomenclatura adotada para indicar os polígonos resultantes do processo de segmentação. Esta denominação foi escolhida em função do próprio método de segmentação utilizado, o qual trabalha com crescimento de regiões, individualizando áreas com NC homogêneos, caracterizados em regiões claras e escuras na imagem. O termo alvo escuro refere-se ao agrupamento das regiões provenientes da segmentação, individualizando a geometria peculiar dos diferentes tipos de alvos escuros (Capítulo 2) passíveis de ocorrerem em imagens SAR adquiridas em áreas oceânicas. A Figura 3.3 disponibiliza uma ilustração com as diferenças entre estas duas definições utilizadas ao longo da tese.

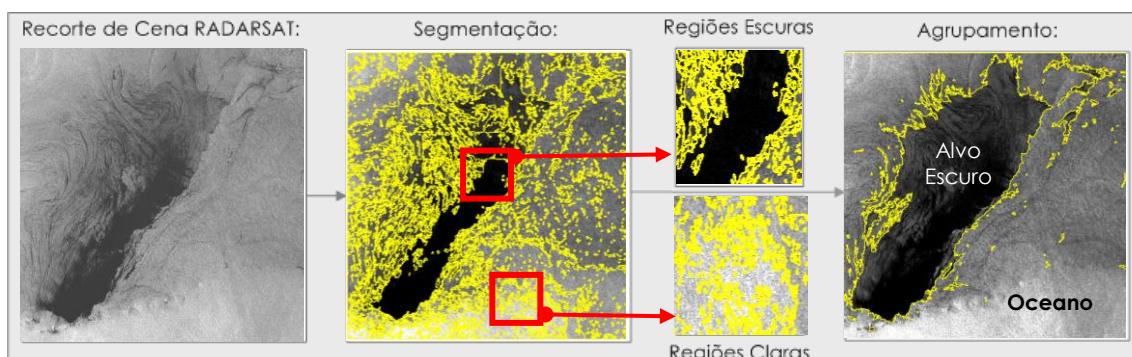


Figura 3.3 – Ilustração das definições referentes a terminologia adota ao longo da tese.

Para calcular as métricas de avaliação, imagens de referência binárias (*Phantons*), considerando as categorias alvos escuros e oceano, foram elaboradas manualmente com o máximo de detalhamento possível por um intérprete experiente. Uma segmentação prévia das imagens para auxiliar na elaboração dos *Phantons* não foi utilizada para evitar que os resultados fossem tendenciosos, uma vez que diferentes segmentadores foram testados (Capítulo 5: item 5.3). Neste contexto, mesmo com as imprecisões associadas à elaboração manual dos *Phantons*, a experiência do intérprete foi valorizada. Importante destacar que esta imprecisão, associada à delimitação das bordas que separam os alvos escuros do oceano,

também ocorre nos processamentos automáticos, em função do ruído *speckle* e outros fatores de interferência descritos no Capítulo 2 (item 2.2).

Foram utilizados 2 tipos de métricas; a) calculadas no nível de regiões, considerando os polígonos resultantes do agrupamento dos segmentos/regiões produzidas pelas segmentações, e; b) métricas calculadas no nível do *pixel* a partir da matriz de confusão.

3.3.1 – Métricas calculadas no Nível das Regiões

Conforme comentado anteriormente, o cálculo das métricas envolve duas imagens: a) a imagem resultante do agrupamento das regiões em alvos escuros e oceano, denominada g_a , e; b) a imagem de referência (*Phantom*), contendo alvos escuros e o oceano, denominada g_p . As regiões, ou polígonos que representam os alvos escuros e o oceano nas imagens g_a e g_p são respectivamente denominadas ra e rp , sendo as funções de segmentação $g_a: S \rightarrow \{ra_1, \dots, ra_n\}$ e $g_p: S \rightarrow \{rp_1, \dots, rp_m\}$, onde n e m são o total de polígonos nas imagens g_a e g_p , respectivamente. O símbolo S refere-se à representação das imagens em uma matriz bidimensional (S) formada por linhas (l) e colunas (c), sendo $S = \{1\dots l\} \times \{1\dots c\}$, onde cada elemento constitui um *pixel*, dotado de coordenadas (x, y) e de um Nível de Cinza (NC) associado.

De acordo com DELVES *et al.* (1992) e LUCCA (1999), considerando que o número de regiões agrupadas em alvos escuros e oceano (n) freqüentemente difere do número de regiões definidas no *Phantom* (m), é necessário: a) sobrepor estas imagens (g_p e g_a); b) comparar os pares de polígonos (rp_j, ra_i) contidos nas duas imagens, onde $j = \{1, \dots, m\}$ e $i = \{1, \dots, n\}$; c) e efetuar uma associação entre estas regiões. O processo de associação entre as regiões contidas em g_p e g_a é efetuado através da construção de duas matrizes (Gf) e (Fit) ambas com dimensões $m \times n$. A notação utilizada foi: $\langle I_{rp_j} \rangle, \langle I_{ra_i} \rangle$ para indicar a média dos NC no k -ésimo polígono, e; $N(k_j), N(k_i)$ para indicar o número de *pixels* contidos no k -ésimo polígono de g_p e g_a respectivamente. As matrizes Gf e Fit foram construídas através da seguintes operações:

$$(1) \quad Gf(j, i) = \frac{N(rp_j \cap ra_i)}{N(rp_j \cup ra_i)}$$

$$(2) \quad Fit(j, i) = \frac{(xd + yd + ((pd + id)/2))}{Gf(rp_j, ra_i)}$$

onde:

$$(3) \quad xd(j,i) = \frac{|\langle x_{rp_j} \rangle - \langle x_{ra_i} \rangle|}{c}$$

$$(4) \quad yd(j,i) = \frac{|\langle y_{rp_j} \rangle - \langle y_{ra_i} \rangle|}{l}$$

$$(5) \quad pd(j,i) = \frac{|N(rp_j) - N(ra_i)|}{N(rp_j) + N(ra_i)}$$

$$(6) \quad id(j,i) = \frac{|\langle Irp_j \rangle - \langle Ira_i \rangle|}{\langle Irp_j \rangle + \langle Ira_i \rangle}$$

Cada um dos componentes das matrizes *Fit* e *Gf* indicados nas equações (1 e 2) efetuam uma comparação entre um par de polígonos ou regiões (rp_j , ra_i), fornecendo: medidas de posicionamento no plano *xy* (xd e yd), intensidade média (id) e tamanho (pd). O componente xd (equação: 3) corresponde à diferença entre a média da posição de *pixels* existentes no eixo x para rp_j e ra_i , dividida pela dimensão da imagem no eixo das abscissas (nº de colunas c). A mesma operação é efetuada para o componente yd (equação: 4), sendo este calculado em relação ao eixo das ordenadas (nº de linhas l). Estas duas métricas fornecem informações sobre o erro de posicionamento entre os centros geométricos (centróides) das duas regiões comparadas. O componente pd (equação: 5) efetua uma relação entre os números de *pixels* que compõe os dois polígonos ou regiões comparadas, fornecendo uma métrica de tamanho relativo entre as mesmas. Já o componente id (equação: 6) analisa as propriedades dos NC entre as regiões comparadas mensurando a diferença entre as intensidades médias dos *pixels* componentes das regiões comparadas (rp_j , ra_i).

Como resultado final deste procedimento, quanto maior for o valor de *Gf* (equação: 1) obtido após a sobreposição de duas regiões rp_j e ra_i , menor será o valor de *Fit* (equação: 2) na posição correspondente (rp_j, ra_i). Desta forma, cada uma das regiões $rp \{rp_1, \dots, rp_m\}$ no *Phantom* será associada a uma única região $ra \{ra_1, \dots, ra_n\}$, considerando o menor valor de *Fit* obtido na matriz (DELVES et al., 1992). A Figura 3.4 disponibiliza um esquema de como é feito o ajuste de regiões entre as duas imagens.

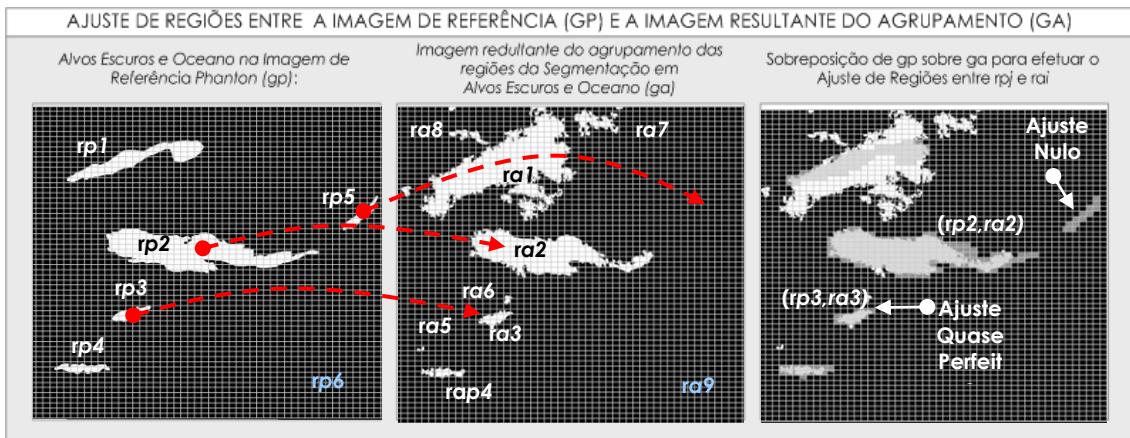


Figura 3.4 - Esquema de como é efetuado o ajuste de regiões entre as duas imagens: exemplo focado para no ajuste efetuado para os alvos escuros.

Tendo como base o ajuste de regiões efetuado (rp_j, ra_i) é possível avaliar cada região ou polígono da imagem agrupada g_a individualmente em relação à imagem de referência g_p , calculando as 4 métricas indicadas a seguir (DELVES *et al.*, 1992; LUCCA, 1999; SCOFIELD *et al.*, 2007):

$$(7) \quad Fitxy = 1 - \frac{xd + yd}{2}$$

$$(8) \quad Fitn = 1 - \frac{|N(rp_j) - N(ra_i)|}{N(rp_j) + N(ra_i)}$$

$$(9) \quad Fit_i = 1 - \frac{|\langle I_{rp_j} \rangle - \langle I_{ra_i} \rangle|}{\langle I_{rp_j} \rangle + \langle I_{ra_i} \rangle}$$

$$(10) \quad Gf = \frac{N(rp_j \cap ra_i)}{N(rp_j \cup ra_i)}$$

Todas as métricas descritas acima foram padronizadas em um intervalo $[0, 1]$, onde o valor 1 representa o melhor ajuste. Para cada polígono da imagem agrupada ajustado em relação ao *Phantom* é obtido um valor referente às 4 métricas calculadas. Para se obter uma avaliação global dos resultados do agrupamento são computadas as médias para cada métrica, considerando o número de regiões/polígonos que compõe o *Phantom* no recorte de cena avaliado. Posteriormente, a distância euclidiana (d_{E2}) entre a média das quatro métricas indicadas nas equações 7, 8, 9 e 10, pode ser obtida para sintetizar os resultados, sendo expressa por:

$$(11) \quad d_{E2} = \sqrt{(Fitxy - 1)^2 + (Fitn - 1)^2 + (Fit_i - 1)^2 + (Gf - 1)^2}$$

A variação de d_{f2} está entre 0 e 2, sendo o melhor resultado a menor distância d_{f2} obtida em relação aos objetos ou polígonos da imagem de referência g_p (NEUBERT *et al.*, 2006).

Outra métrica calculada no nível das regiões foi o número de alvos escuros detectados em relação ao *Phantom* (IPol), calculado a partir da intersecção dos polígonos entre a imagem agrupada (g_a) e a imagem de referência (g_p).

3.3.1 – Métricas Calculadas no Nível dos Pixels

Como o objetivo da tese não foi avaliar diretamente os resultados dos procedimentos de segmentação, mas sim o resultado do agrupamento das regiões resultantes das segmentações, o cálculo de métricas a partir de uma matriz de confusão também foi utilizado. Matrizes de confusão são comumente utilizadas para avaliar a qualidade de uma classificação ou mapeamento. Considerando que o interesse das análises efetuadas foi avaliar a separação dos alvos escuros do oceano, o estudo realizado também remete a um típico problema de classificação no qual se deseja obter duas classes. Neste sentido, a imagem contendo o agrupamento das regiões em alvos escuros e oceano também pode ser interpretada como uma classificação.

Para o cálculo da matriz de confusão foi utilizado o domínio do *pixel*, comparando-se as duas imagens binárias, *Phantom* e imagem segmentada agrupada nas classes Alvo Escuros (E) e Oceano (O). A Tabela 3.1 indica a matriz de confusão aplicada no domínio do *pixel*, depois do agrupamento das regiões.

Tabela 3.1 - Matriz de confusão para o agrupamento das regiões em alvos escuros e oceano.

		Imagem Agrupada		
<i>Phantom</i>	Matriz de Confusão	Alvos Escuros (E)	Oceano (O)	Soma
	Alvos Escuros (E)	NEE	NEO	NEE+NEO = Alvos Escuros no Phantom
	Oceano (O)	NOE	NOO	NOE+NOO = Oceano no Phantom

A nomenclatura adotada na matriz de confusão foi: a) NEE: número de *pixels* escuros associados a alvos escuros corretamente agrupados como alvos escuros; b) NOO: número de *pixels* claros associados ao oceano corretamente agrupados como oceano; c) NEO: número de *pixels* escuros associados aos alvos escuros erroneamente agrupados como oceano, e; d) NOE: número de *pixels* claros associados ao oceano erroneamente agrupados como alvos escuros. Tendo como

base a matriz de confusão as métricas foram computadas considerando somente os acertos e os erros ao redor dos alvos escuros.

Esta escolha foi efetuada, considerando que as amostras de alvos escuros e de oceano em imagens SAR não são balanceadas, proporcionalmente, há muito mais oceano do que alvos escuros. KUBAT *et al.* (1998) abordou o problema das amostras não平衡adas, especificamente no âmbito dos sistemas automáticos para a detecção de óleo no mar, destacando que, nestes casos, métricas de avaliação global não podem ser utilizadas como medidas de desempenho. Por este motivo, ressalta-se que no trabalho não foram utilizados a exatidão global e coeficientes de avaliação como o *Kappa*. Testes foram efetuados para a utilização do *Kappa*, entretanto, os coeficientes gerados eram sempre bons, uma vez que para o oceano os índices de acerto são, proporcionalmente, muito maiores, mascarando as imprecisões observadas na classificação dos alvos escuros.

Desta forma, as métricas utilizadas desconsideraram a porção NOO da matriz de confusão, considerando somente NEE, NEO e NOE para a composição das métricas denominadas Fração de Intersecção entre a Classificada e o *Phantom*, Omissão, Inclusão e Exatidão dos Alvos Escuros, especificadas abaixo:

Fração de Intersecção entre a Classificada e o *Phantom* (*FInt*): indica a proporção de *pixels* de alvos escuros detectados pela classificação em relação aos *pixels* de alvo escuros do *Phantom*. Pode ser representada pelo número de *pixels* da intersecção de alvos escuros do *Phantom* e da imagem classificada, dividido pelo número de *pixels* escuros do *Phantom*, sendo:

$$(12) \quad FInt = \frac{NEE}{NEE + NEO}$$

Essa fração permite visualizar a proporção de alvos escuros do *Phantom* que foram mapeados, porém não evidencia os erros de inclusão.

Omissão (*Om*): indica quanto a classificação errou associando *pixels* de alvos escuros ao oceano, alterando o formato destes alvos em relação à imagem de referência (g_p), sendo:

$$(13) \quad Om = \frac{NEO}{NEE + NEO}$$

Inclusão (In): indica quanto a classificação errou associando *pixels* de oceano a alvos escuros, alterando o formato destes alvos em relação à imagem de referência (g_p), sendo:

$$(14) \quad In = \frac{NOE}{NEE + NEO}$$

Exatidão dos Alvos Escuros ($ExAE$): esta métrica foi definida com o objetivo de sintetizar em um único índice as métricas referentes à matriz de confusão relacionadas unicamente aos alvos escuros. Neste sentido, de forma diferente da $FInt$, a intersecção dos alvos escuros foi dividida pela união de todas áreas escuras (*Phantom* ou imagem classificada). Quando esse índice se aproxima de 1 indica que o alvo escuro na imagem classificada está mais próximo do *Phantom* considerando os dois tipos de erro.

$$(15) \quad ExAE = \frac{NEE}{NEE + NEO + NOE}$$

Além destas métricas, o tempo de processamento, calculado numa mesma plataforma computacional, juntamente com uma avaliação qualitativa, considerando a opinião do intérprete sobre os resultados obtidos ao longo da tese, também foi considerado.

CAPÍTULO 4

METODOLOGIA E RECURSOS UTILIZADOS

Considerando o estado da arte em técnicas de segmentação de imagens e clusterização, a metodologia proposta tem por objetivo estabelecer um procedimento automático para a detecção de polígonos representativos para os diversos tipos de alvos escuros identificáveis em imagens SAR (Capítulo 2). A Figura 4.1 contém um fluxograma das etapas vislumbradas para o funcionamento completo do sistema destacando as etapas a serem desenvolvidas no estudo proposto.

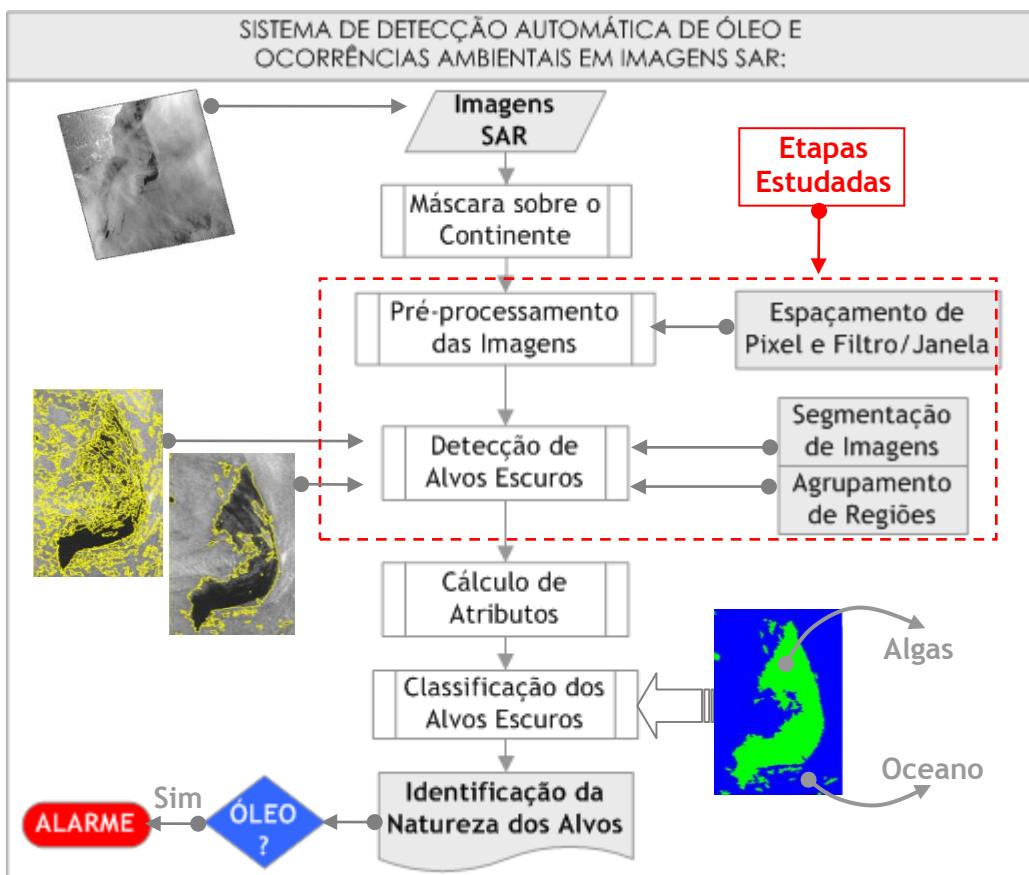


Figura 4.1 – Etapas de processamento envolvidas no sistema de detecção automática de óleo e ocorrências ambientais.

Os procedimentos metodológicos propostos neste trabalho estão organizados em 6 etapas distintas (Figura 4.2): a) definição do conjunto de dados: formato dos dados e seleção dos recortes de cena; b) determinação do espaçamento de *pixel* mais adequado para as imagens; c) determinação do filtro e da janela a serem utilizados no pré-processamento das imagens; d) detecção de regiões escuras e claras através de técnicas de segmentação de imagens de satélite, determinando o segmentador mais adequado; e) determinação de um procedimento de clusterização não supervisionado para a produção de polígonos representativos para alvos escuros detectáveis em

imagens SAR, e; f) validação do método proposto. Estas etapas estão detalhadas nas seções 4.1, 4.2, 4.3, 4.4, 4.5 e 4.6.

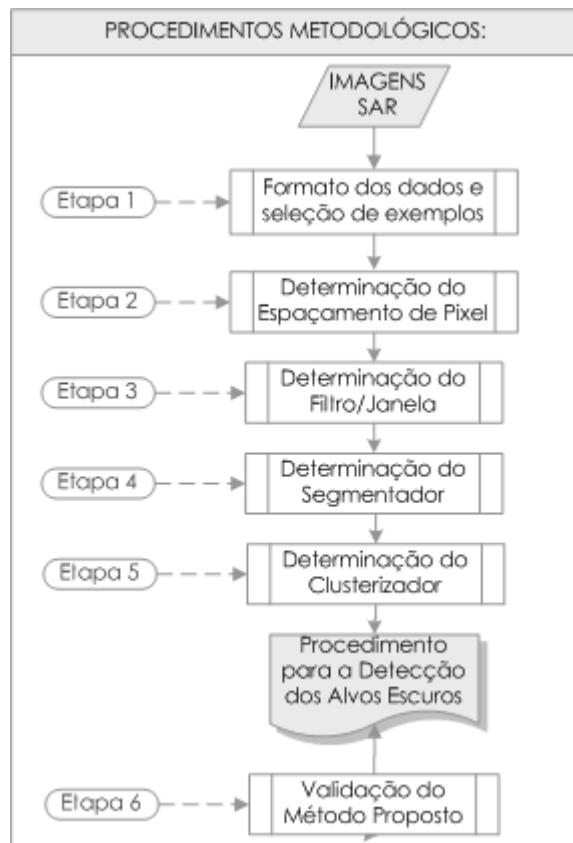


Figura 4.2 – Etapas integrantes da metodologia proposta na tese.

4.1 – Etapa 1: Conjunto de Dados - Informações Técnicas e Seleção de Exemplos

Imagens provenientes do satélite RADARSAT-1, Banda C (freqüência de 5,3GHz e comprimento de onda de 5,6cm), polarização *HH*, nos modos de imageamento *ScanSAR Narrow* e *ScanSAR Wide* foram cedidas pela PETROBRAS para os estudos propostos. Como no momento da seleção dos exemplos o satélite RADARSAT-2 ainda não estava operacional, imagens deste satélite não foram utilizadas. A Tabela 4.1.apresenta a configuração dos modos de imageamento do RADARSAT 1 utilizados na tese.

Tabela 4.1 - Configuração dos modos de imageamento do satélite RADARSAT-1 adequados ao monitoramento costeiro e oceânico

Satélite: Sistema Sensor	Ano de Lançamento	Resolução Espacial	Largura da Faixa	Banda e Polarização	Entrega em Tempo Próximo do
RADARSAT-1 - Modo: SCANSAR NARROW A e B		50m	300 km		
RADARSAT-1 - Modo: SCANSAR WIDE	1995	100m	500 km	C (HH)	Operacional

Os produtos utilizados são formados a partir da junção de 2 ou mais modos únicos denominados *single look*. Estas imagens são denominadas *multi-look*, sendo o dado complexo disponível apenas para os modos *single look*. O modo *ScanSAR Narrow* - com número de *looks* igual 4 - é formado a partir da junção de 2 ou 3 modos *single look*, enquanto que os produtos *ScanSAR Wide* - com número de *looks* igual a 8 - são formados pela junção de 4 modos *single look*.

Embora as imagens possam ser fornecidas no formato bruto CEOS, para garantir o caráter operacional da atividade os produtos são entregues em tempo próximo do real no formato compactado *MrSid* (*Multi-resolution Seamless Image Database*) com resolução radiométrica de 8 bits. A possibilidade de utilizar os dados no formato CEOS foi avaliada, entretanto, o tempo médio gasto para o *download* destes dados é alto (35min) quando comparado com o tempo gasto para o *download* dos dados no formato *MrSid* (segundos). Complementarmente, os dados no formato compactado, são entregues georreferenciados, filtrados e com correção do padrão de antena, processamentos necessários para a utilização dos dados no formato CEOS.

As imagens RADARSAT utilizadas pela PETROBRAS são processadas em três estações de recepção, duas localizadas no Canadá (Prince Albert e Gatineau) e uma na Argentina (CONAE). Apesar dos dados serem processados em três locais diferentes, um dos operadores da estação Gatineau garantiu que há um padrão no pré-processamento dos dados utilizado em todas as estações de recepção (GAMBLE, 2008). Esta informação é importante à medida que diferentes pré-processamentos podem influenciar nos resultados obtidos para a detecção de óleo e assemelhados. De acordo com GAMBLE (2008), as imagens são compactadas em 10 vezes (10: 1), reduzindo o tamanho dos arquivos de 100-150 Mb para 10-15 Mb no formato *MrSid* sem reduzir as dimensões da imagem. Embora a compactação seja imprescindível em função do caráter operacional da atividade, é importante ressaltar que o processo de compactação/descompactação altera as propriedades estatísticas das imagens SAR.

O teste de *Kolmogorov-Smirnov*⁴, comparando uma mesma imagem RADARSAT no formato CEOS e descompactada a partir do formato *MrSid*, rejeita a hipótese nula de que as distribuições estatísticas são iguais nos dois formatos com *p-value* igual a zero. A rejeição da hipótese nula comprova que as propriedades estatísticas das imagens são alteradas após a compressão dos dados no formato

⁴ O teste de *Kolmogorov-Smirnov* é utilizado para evidenciar se duas amostras independentes possuem o mesmo tipo de distribuição estatística.

MrSid. A Figura 4.3 disponibiliza os histogramas e os *box plots* para uma mesma imagem no formato CEOS e *MrSid*, juntamente com os atributos estatísticos de média, mediana, mínimo, máximo e desvio padrão, evidenciando as diferenças.

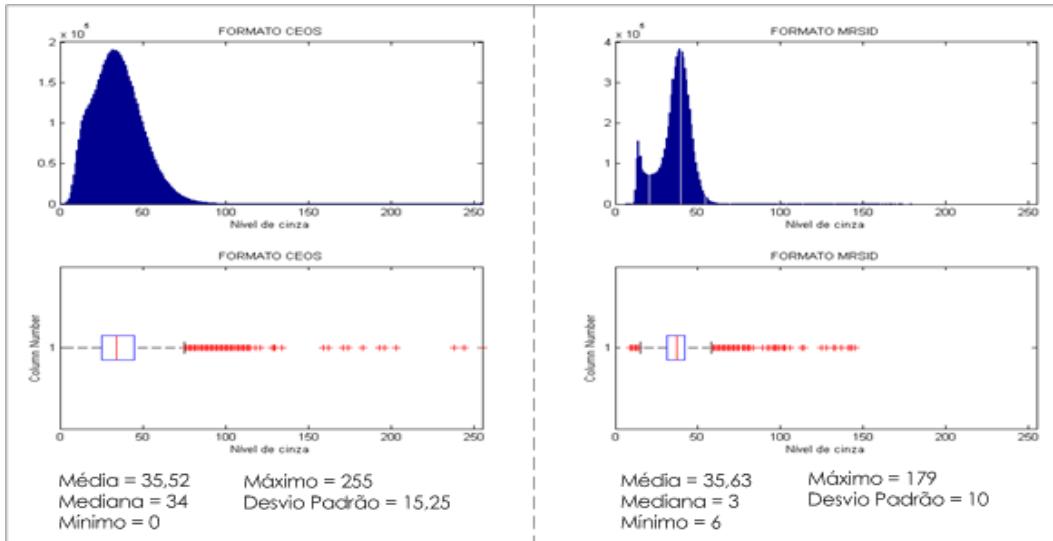


Figura 4.3 - Histogramas e *box plots* para uma mesma imagem no formato CEOS e *MrSid*.

Apesar da compactação alterar as propriedades estatísticas das imagens, esse formato foi utilizado considerando que a hipótese da tese foi desenvolver um método para detecção automática de alvos escuros com potencial para ser utilizado operacionalmente.

Para a tese foram selecionados 14 recortes de cena, 11 para efetuar as análises e 3 para validar o método proposto ao final da tese. A Tabela 4.2 contém a data das cenas e os parâmetros de aquisição das imagens pelo sensor para cada um dos recortes de cena utilizados.

Tabela 4.2 - Data das cenas e parâmetros de aquisição das imagens pelo sensor para os recortes de cena selecionados.

Recortes de Cena:	Datas	Modo de Imageamento	Órbita	Resolução Espacial (m)	Largura de Faixa (km)
Cena: 1	6/7/2001	ScanSAR Narrow	Ascendente	50	300
Cena: 2	6/7/2002	ScanSAR Narrow	Descendente	50	300
Cena: 3	6/3/2006	ScanSAR Wide	Ascendente	100	500
Cena: 4	9/9/2001	ScanSAR Narrow	Ascendente	50	300
Cena: 5	13/6/2008	ScanSAR Wide	Ascendente	100	500
Cena: 6	5/5/2002	ScanSAR Narrow	Descendente	50	300
Cena: 7	19/12/2001	ScanSAR Narrow	Descendente	50	300
Cena: 8	3/4/2002	ScanSAR Narrow	Ascendente	50	300
Cena: 9	4/7/2001	ScanSAR Narrow	Descendente	50	300
Cena: 10	11/3/2005	ScanSAR Narrow	Ascendente	50	300
Cena: 11	14/6/2008	ScanSAR Narrow	Descendente	50	300
Cena Validação A:	2/1/2003	ScanSAR Narrow	Ascendente	50	300
Cena Validação B:	13/7/2007	ScanSAR Wide	Ascendente	100	500
Cena Validação C:	15/10/2001	ScanSAR Narrow	Descendente	50	300

A seleção de exemplos constitui etapa essencial para o desenvolvimento de qualquer metodologia que tenha por objetivo propor um sistema para a detecção automática de alvos escuros em imagens SAR. Considerando os inúmeros fatores peculiares às imagens SAR, os quais dificultam a detecção dos alvos escuros (Capítulo 2), a escolha de recortes de cena individualizando alvos escuros com bom contraste em relação ao oceano pode levar a conclusões precipitadas sobre a possibilidade ou não de automatização do sistema.

Neste contexto, a seleção realizada buscou refletir ao máximo a realidade encontrada nas imagens SAR, incluindo: a) recortes de cena considerados fáceis, nos quais um único alvo escuro, com geometria simples e bom contraste em relação ao oceano foram individualizados; b) recortes de cena considerados difíceis, contendo mais de um tipo de alvo escuro, em diferentes escalas, com geometria variável; c) alvos escuros ocorrendo simultaneamente com baixas intensidades de vento, dificultando a detecção em função da variação do contraste entre estes alvos escuros e o oceano, e; d) recortes de cena com falhas no processamento efetuado para gerar as imagens, como a descontinuidade dos NC entre faixas de imageamento distintas. A Tabela 4.3 apresenta uma descrição das características dos recortes de cena RADARSAT selecionados, indicando: a) o tamanho dos recortes de cena em *pixels*; b) o número de alvos escuros individualizados; c) a complexidade da geometria; d) a existência de alvos escuros de diferentes tamanhos no mesmo recorte; e) se o contraste entre os alvos escuros e o oceano era instável ao longo de todo recorte ou apresentava variações; f) se a imagem apresentava falhas de processamento ou não, e; g) se nos recortes havia apenas um tipo de alvo escuro – óleo(s) ou assemelhado(s) - ou os dois tipos de evento – óleo(s) e assemelhado(s).

Tabela 4.3 – Características associadas à seleção dos recortes de cena utilizados.

Recortes de Cena:	Tamanho da Cena	No de Alvos	Geometria dos Alvos Escuros	Escala dos Alvos Escuros	Contraste na Cena	Falhas no Processamento	Tipos de Alvos Escuros
Cena 1	1.934.781	1	Complexa	Muito Variável	Instável	Ausente	Assemelhados
Cena 2	819.993	1	Simples	Sem Variação	Estável	Ausente	Óleo
Cena 3	151.620	1	Simples	Sem Variação	Instável	Ausente	Óleo
Cena 4	473.770	Multiplos	Simples	Muito Variável	Instável	Presente	Óleo
Cena 5	1.139.223	Multiplos	Complexa e Simples	Muito Variável	Instável	Ausente	Óleo e Assemelhados
Cena 6	103.772	2	Simples	Sem Variação	Estável	Ausente	Óleo
Cena 7	782.222	1	Complexa	Muito Variável	Estável	Ausente	Assemelhados
Cena 8	385.396	Multiplos	Complexa	Muito Variável	Estável	Ausente	Assemelhados
Cena 9	306.783	Multiplos	Simples	Muito Variável	Instável	Ausente	Óleo
Cena 10	1.333.480	Multiplos	Complexa e Simples	Muito Variável	Instável	Ausente	Óleo e Assemelhados
Cena 11	427.025	Multiplos	Complexa	Muito Variável	Instável	Ausente	Óleo e Assemelhados
Cena Validação A:	2/1/2003	1	Simples	Sem Variação	Estável	Ausente	Óleo
Cena Validação B:	13/7/2007	Multiplos	Complexa	Muito Variável	Instável	Presente	Assemelhados
Cena Validação C:	15/10/2001	Multiplos	Complexa	Sem Variação	Estável	Ausente	Assemelhados

4.2 – Etapa 2: Determinação do Espaçamento de Pixel

O estudo realizado utilizou um recorte de uma cena RADARSAT-1 (1671 linhas X 922 colunas) adquirida em 03/04/2002 no modo *ScanSAR Narrow* com espaçamento de *pixel* de 50m, largura de faixa de 300km, polarização *HH*, número de *looks* igual a 4 entregue em 8 bits no formato compactado *MrSid* (Figura 4.4).

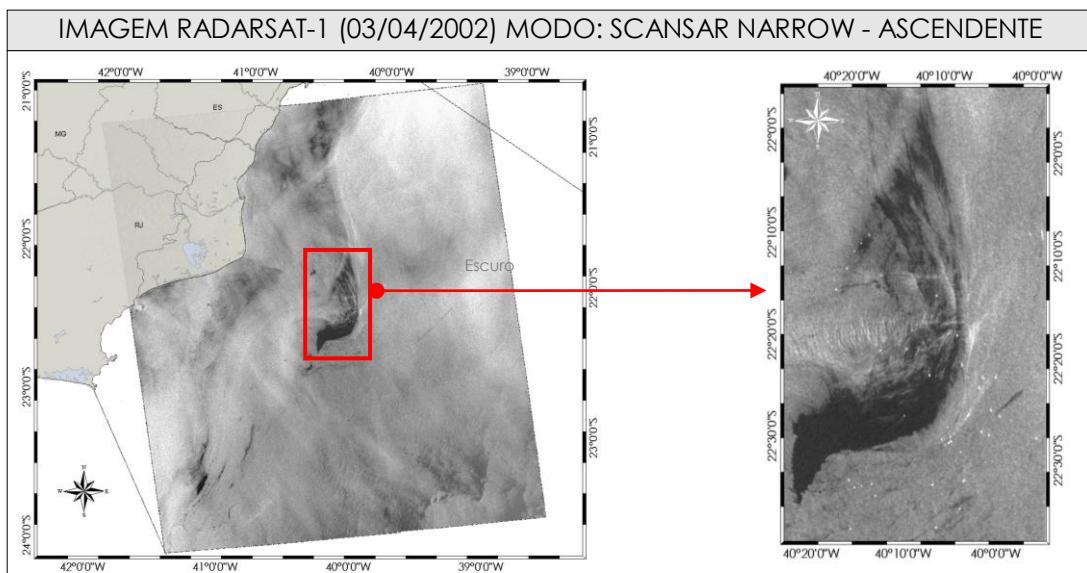


Figura 4.4 – Recorte da cena RADARSAT-1 (03/04/2002) utilizada no estudo proposto.

O recorte da cena RADARSAT-1 com 50m de espaçamento de *pixel* foi reamostrado para 100m, 150m e 200m utilizando o método de reamostragem por vizinho mais próximo no software ENVI, com o objetivo de não alterar as propriedades radiométricas das imagens. Todas as imagens foram segmentadas no software MultiSeg (SOUSA, 2005; SOUSA *et al.*, 2005), escolhido por ser gratuito e programado especificamente para imagens de RADAR. Estudos para identificar qual o melhor modo de processamento e quais os parâmetros adequados para serem utilizados no software MultiSeg foram realizados. Os segmentos, regiões claras e escuras geradas para cada espaçamento de *pixel*, foram agrupados em 2 categorias, alvos escuros e oceano, considerando como atributo a média dos níveis de cinza (NC) para cada segmento ou região. A técnica de agrupamento utilizada foi limiarização por Médias, sendo o limiar definido interativamente pelo intérprete, escolhendo o valor de partição que retornou visualmente o melhor resultado. Finalmente, computando-se as métricas de avaliação descritas na seção 3.3, as quais relacionam a imagem resultante dos agrupamentos com a imagem de referência (*Phantom*), foi determinado o espaçamento de *pixel* mais adequado. A Figura 4.5 ilustra a metodologia adotada para investigar o espaçamento de *pixel* adequado para o processamento dos dados RADARSAT para a metodologia proposta.

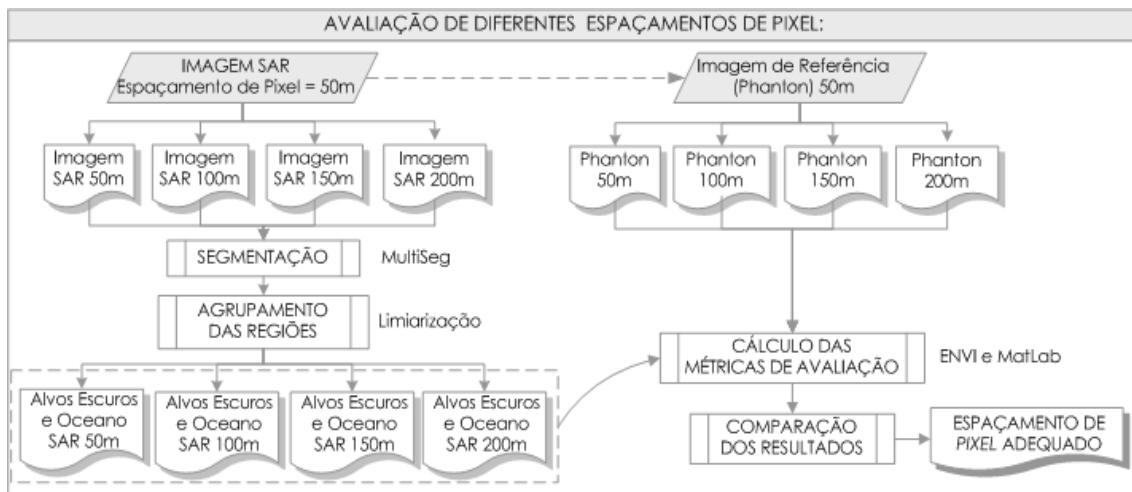


Figura 4.5 – Metodologia adotada para investigar o espaçamento de *pixel* adequado para o processamento dos dados RADARSAT no sistema de detecção automático.

4.3 – Etapa 3: Determinação do Filtro/Janela

Após a determinação do espaçamento de *pixel* mais adequado a ser utilizado foram testados os filtros Frost, Frost Enhanced, Lee, Lee Enhanced, Kuan, Gamma, Sigma e Mediana, todos aplicados considerando as dimensões das janelas de 3x3, 5x5 e 7x7. O mesmo recorte de cena da primeira etapa foi utilizado como dado de entrada para investigação dos diferentes filtros (Figura 4.4). Os recortes de cena sem filtro e filtrados, considerando os diferentes filtros e janelas, foram segmentados no software Multiseg retornando regiões escuras e claras. Estas regiões foram agrupadas por limiarização em alvos escuros e oceano e estes foram avaliados comparativamente com a imagem de referência (*Phantom*), seguindo a mesma metodologia descrita no item 4.2. No total foram produzidos 25 resultados a serem comparados com o *Phantom* elaborado com o espaçamento de *pixel* definido como mais adequado no item 4.2. Definir se a aplicação de um filtro na fase de pré-processamento das imagens oferece melhoria significativa nos resultados da segmentação e qual dos filtros/janela oferece melhor resultado foi o objetivo desta etapa. A Figura 4.6 ilustra a metodologia utilizada para determinar o melhor filtro/janela.

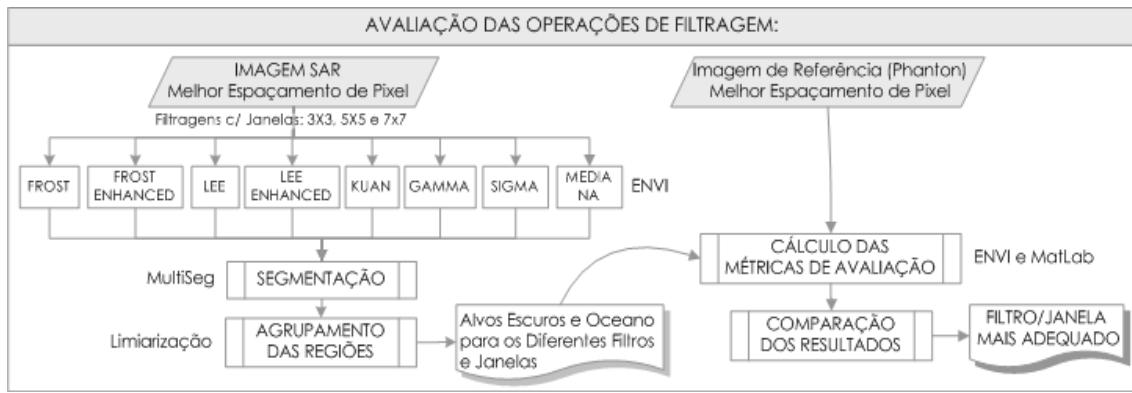


Figura 4.6 – Metodologia adotada para investigar o filtro/janela adequado para o processamento dos dados RADARSAT no sistema de detecção automático.

A aplicação de filtragem nas imagens pode alterar a eficiência da segmentação e, consequentemente, do agrupamento das regiões escuras e claras para os diferentes tipos de alvos escuros dotados de geometria e textura peculiares. Para avaliar se a aplicação do filtro/janela escolhido contribui para a detecção de alvos escuros, na etapa seguinte os diferentes recortes de cena foram processados sem filtro e com o filtro/janela indicado como mais adequado nesta etapa. Este procedimento foi realizado para validar, com maior consistência, a decisão sobre utilizar ou não um filtro/janela padrão na fase de pré-processamento das imagens.

4.4 – Etapa 4: Determinação do Procedimento de Segmentação

Para a determinação do segmentador, além do recorte de cena já utilizado nas etapas anteriores (Figura 4.4), foram selecionadas imagens SAR com diferentes tipos de alvos escuros, passíveis de serem encontrados em áreas oceânicas. A diversidade de alvos escuros foi considerada nesta etapa à medida que a eficiência dos segmentadores é influenciada pela geometria e pela textura peculiar a cada tipo de alvo (Capítulo 4: item 4.1). Nesta etapa da tese, 11 recortes de cena RADARSAT sem filtragem e pré-processados com o espaçamento de *pixel* e o filtro/janela mais adequados – ambos determinados nas etapas anteriores - foram segmentados nos softwares Multiseg e SPRING. As regiões escuras e claras resultantes das segmentações foram agrupadas em alvos escuros e oceano. A técnica de agrupamento de limiarização por Médias também foi utilizada, sendo o limiar definido interativamente pelo intérprete, escolhendo o valor de partição que retornou visualmente o melhor resultado. A escolha dos limiares interativamente foi utilizada nesta etapa da tese para que uma avaliação justa dos segmentadores fosse realizada, indicando sempre o melhor limiar para cada cena, uma vez que a utilização de um limiar fixo não retornaria bons resultados para todas as cenas, com todos os segmentadores.

Os alvos escuros e oceano, resultantes do agrupamento das regiões escuras e claras produzidas nas segmentações, foram avaliados comparativamente com os indicados como referência no *Phantom*, seguindo a mesma metodologia descrita no item 4.2. A Figura 4.7 apresenta os procedimentos metodológicos para a determinação do segmentador.

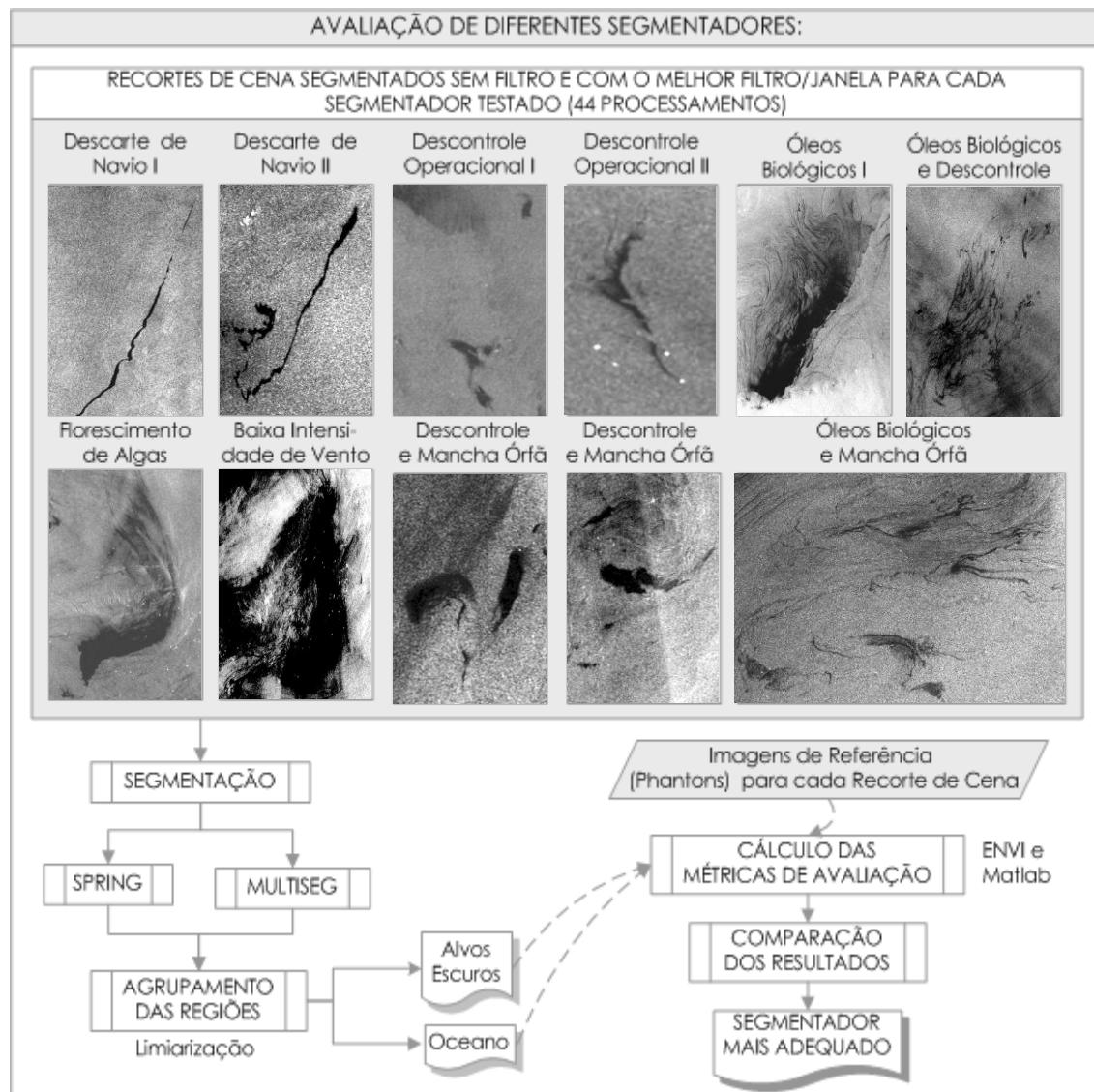


Figura 4.7 - Fluxo de procedimentos metodológicos planejado para investigar a etapa de segmentação em imagens de SAR.

Os segmentadores testados utilizam diferentes parâmetros de entrada, os quais influenciam no número de segmentos e no formato dos segmentos produzidos. Como o foco da tese não esteve centrado na segmentação, mas no agrupamento das regiões resultantes da segmentação, a preocupação foi balizar estes parâmetros de maneira que a imagem não fosse subsegmentada a ponto de perder feições de interesse e nem supersegmentada, exigindo um maior desempenho computacional para o processamento dos resultados. Como o objetivo desta etapa foi identificar o

segmentador adequado para ser acoplado ao sistema de detecção automático foi necessário pesquisar valores que se adequassem como padrão para os parâmetros exigidos pelo segmentador. Para tal, como parte dos resultados gerados, pesquisas foram realizadas a fim de investigar os parâmetros adequados para serem utilizados nos softwares testados.

4.5 – Etapa 5: Determinação do Procedimento de Clusterização

Uma vez determinados, através dos procedimentos descritos nas etapas anteriores, o espaçamento de *pixel*, o filtro/janela e o segmentador mais adequados, o procedimento metodológico para definir o método de clusterização foi composto pelas seguintes etapas: a) segmentação das imagens; b) cálculo de diferentes atributos estatísticos, geométricos e texturais, tendo como base as imagens e as regiões escuras e claras resultantes das segmentações; c) análise exploratória e seleção de atributos; d) agrupamento automático das regiões em alvos escuros e oceano, considerando os diferentes atributos como dados de entrada no software *Clustan*; e) avaliação do(s) atributo(s) mais adequados, comparando os resultados das clusterizações efetuadas cumulativamente inserindo gradualmente cada atributo, com os alvos de referência do *Phantom*; f) considerando o conjunto de atributos selecionado, uma comparação do desempenho entre técnicas automáticas (não supervisionada) e semi-automáticas (supervisionadas) para a detecção de alvos escuros foi efetuada. Nesta última etapa foram comparados: a) os resultados provenientes dos algoritmos supervisionados de Máxima Verossimilhança (Maxver) e Máquinas de Suporte Vetorial (SVM); b) os resultados das operações de limiarização efetuadas nos itens anteriores, e; c) os resultados obtidos pelo clusterizador *Focal Point* implementado no software *Clustan*.

A Figura 4.8 apresenta os procedimentos metodológicos para a determinação do procedimento de clusterização e do(s) atributo(s) a serem utilizados na detecção de alvos escuros em imagens SAR.

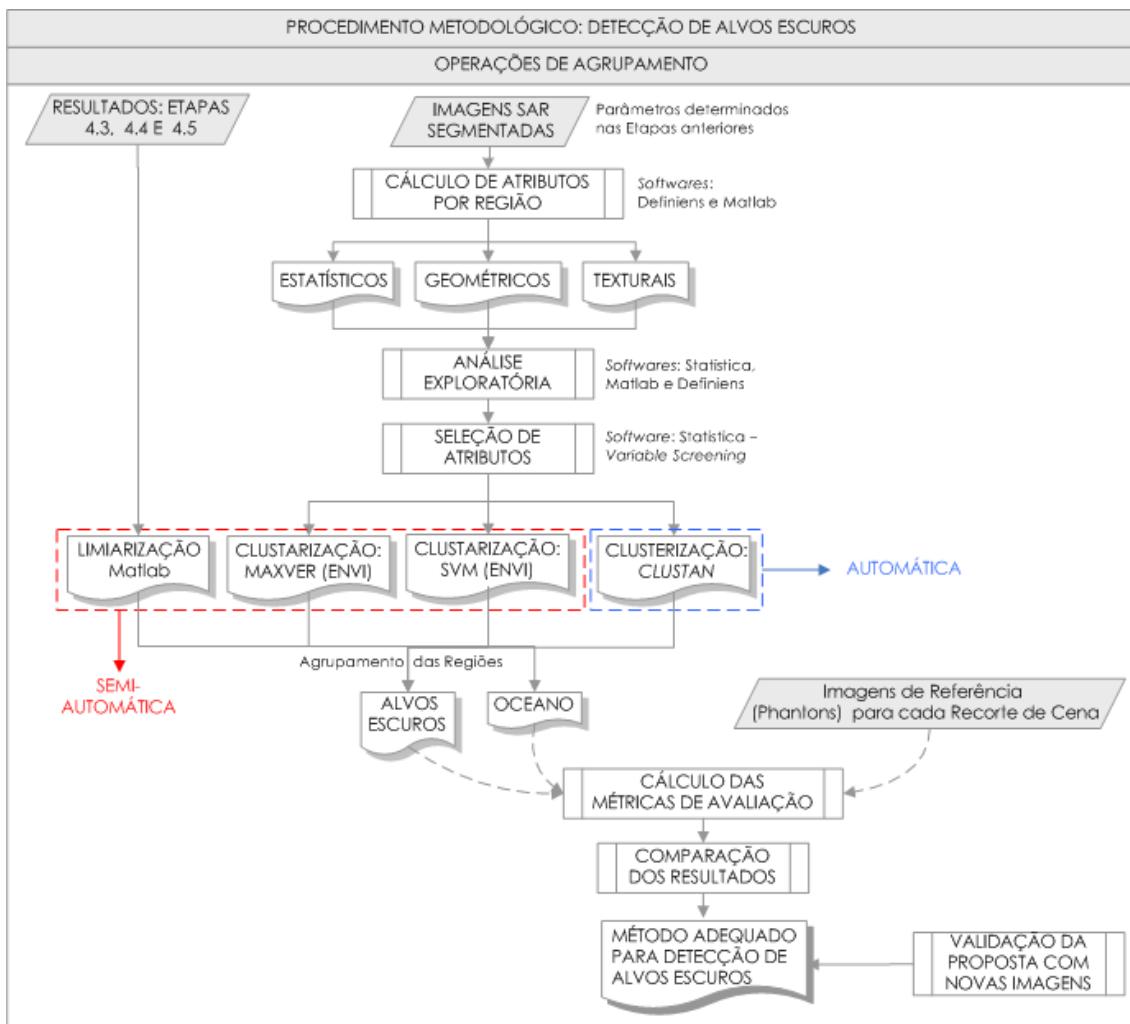


Figura 4.8 - Fluxo de procedimentos planejado para investigar o método de clusterização mais adequado a ser utilizado após a segmentação das imagens SAR.

Conforme comentado, as avaliações foram realizadas através da comparação dos alvos escuros resultantes do agrupamento das regiões escuras e claras, com os polígonos indicados como referência nos *Phantons*, considerando o índice de Exatidão dos Alvos Escuros descrito no item 3.3.

4.6 – Etapa 6: Validação do Método Proposto

A possibilidade de estabelecer uma regra de decisão para o processamento automático das imagens foi considerada. Para tal, o método de Árvores de Classificação e Regressão - ACR (*Classification and Regression Trees* - CART) disponível no módulo *Data Mining* do Software *Statistica*, foi utilizado.

A partir da regra de decisão encontrada, a validação do método proposto foi efetuada a partir do pré-processamento de 3 novas imagens SAR, utilizando o espaçamento de *pixel*, o filtro/janela e o segmentador com os parâmetros definidos nas etapas anteriores como os mais adequados para a detecção de alvos escuros. O

conjunto de atributos adequado, igualmente definido na etapa 4.5 da tese, foi calculado para cada novo recorte de cena e as clusterizações foram efetuadas no software *Clustan*. *Phantons* foram delineados manualmente e as métricas de avaliação calculadas da mesma forma que nas etapas anteriores.

Posteriormente, os atributos chaves utilizados pela ACR para formular as regras de decisão foram igualmente calculados e analisados para os 3 recortes de cena, a fim de verificar se o comportamento encontrado refletiria a regra indicada pelo algoritmo. Finalmente, com base em todos os resultados obtidos ao longo da tese, um método para a detecção de alvos escuros em imagens SAR foi proposto, integrando técnicas de segmentação de imagens e clusterização de dados.

CAPÍTULO 5

RESULTADOS E DISCUSSÕES

5.1 – Etapa 1: Determinação do Espaçamento de Pixel

5.1.1 – Determinação dos Parâmetros Adequados para o Software MultiSeg

Conforme comentado no Capítulo 3: Item 3.1.1.1, o software MultiSeg foi inicialmente utilizado nas primeiras etapas, tendo sido necessário investigar quais os modos e parâmetros adequados para a segmentação. Desta forma, com relação aos modos de segmentação *Cartoon* e *Texture* disponíveis no software MultiSeg, análises preliminares evidenciaram que o tempo de processamento necessário para segmentar o mesmo recorte de imagem foi maior no modo *Texture* (60min) do que no *Cartoon* (em média 3min), utilizando a mesma plataforma computacional. Comparando os resultados com o *Phantom* (Figura 5.1), o modo *Texture* segmentou mais intensamente as regiões escuras (alvos de interesse) em detrimento das regiões claras, resultado que reduziria o custo computacional na etapa de agrupamento dos segmentos em alvos escuros e *oceano*. A análise visual dos resultados indicou que, embora a segmentação no modo *Texture* tenha priorizado a segmentação das regiões escuras, a geometria dos alvos escuros obtida após o agrupamento das regiões não ficou bem delineada. O maior número global de regiões (escuras e claras) produzidas no modo *Cartoon* permitiu formar polígonos mais representativos para os alvos escuros com maior acurácia em relação ao *Phantom*. Além da possibilidade do modo *Texture* detectar com menor acurácia a geometria dos alvos escuros em relação ao modo *Cartoon*, a diferença de custo computacional evidenciada entre estes modos inviabilizaria o uso operacional do modo *Texture*, a medida que o sistema de detecção automática processaria imagens inteiras e não somente recortes de cenas⁵.

Os modos de aquisição das imagens RADARSAT para o monitoramento costeiro e oceânico (*ScanSAR Narrow* e *ScanSAR Wide*) são entregues em amplitude em 8 bits (0~255). Para processá-los no modo *Cartoon* com melhor desempenho foi necessário utilizá-los em intensidade 16 bits (0~65536), pois o tempo de processamento foi menor neste formato. A Figura 5.1 ilustra os resultados obtidos para

⁵ O tamanho do recorte de cena utilizado no experimento é de 922 colunas X 1671 linhas, totalizando 1540662 pixels, e o tamanho da imagem inteira é de 7873 colunas X 8413 linhas, totalizando 66235549 pixels. O recorte de cena equivale a apenas 2% da área total da cena, restando ainda 98% da área da cena a ser processada.

segmentações realizadas nos modos/formato *Cartoon/Amplitude*, *Cartoon/Intensidade* e *Texture/Intensidade*, evidenciando o número de regiões produzidas e o tempo de processamento. O software não conseguiu processar os dados no modo/formato *Texture/Amplitude*.

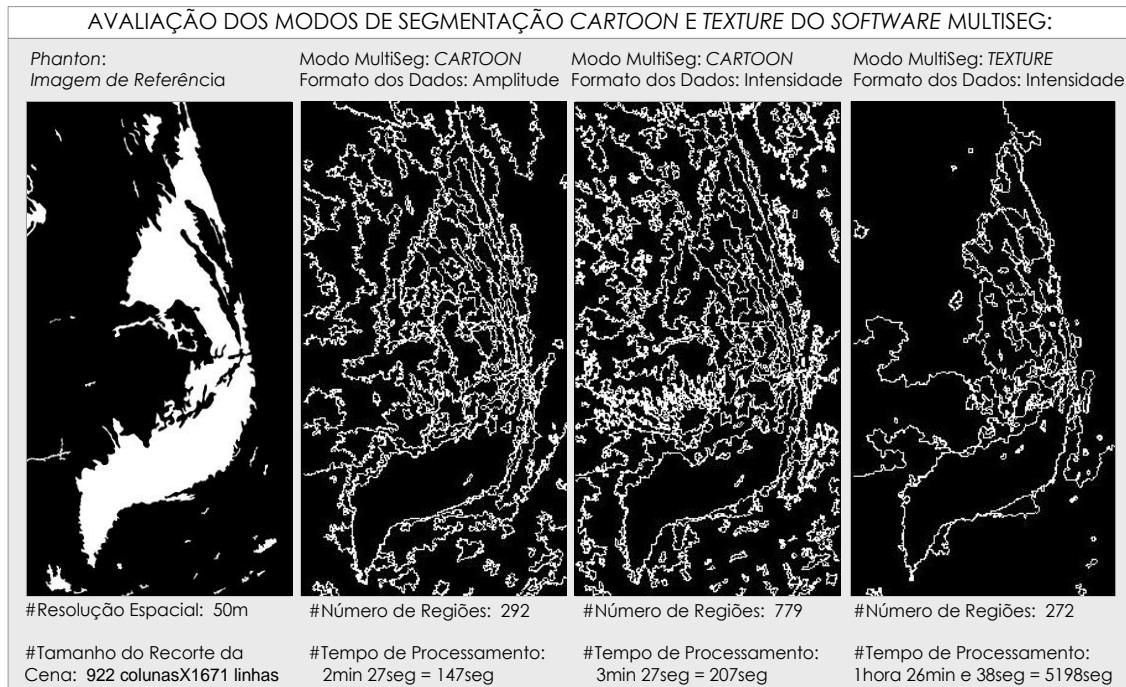


Figura 5.1 - Resultados obtidos para segmentações realizadas nos modos/formato: *Cartoon/Amplitude*, *Cartoon/Intensidade* e *Texture/ Intensidade*.

Conforme observado na Figura 5.1, a diferença no tempo de processamento foi pequena comparando-se os dados em amplitude e em intensidade para o modo *Cartoon*. Desta forma, optou-se pela utilização dos dados de entrada no padrão de configuração do software em intensidade.

Esta escolha considerou também a constatação de que a utilização dos dados em intensidade produziu um número maior de regiões (779) quando comparado com o número de regiões produzidas para os dados em amplitude (292). Nas áreas costeiras e oceânicas há diversos tipos de fenômenos que ocorrem em diferentes escalas sendo detectáveis nas imagens de RADAR como alvos escuros em grande, média e pequena dimensão espacial. A subsegmentação de uma imagem contendo diferentes alvos em múltiplas escalas de representação espacial pode resultar na redução do potencial de detecção dos alvos de pequena dimensão. Neste contexto, quanto maior o número de regiões produzidas maior a chance de detecção de pequenos alvos. O maior número de regiões com dimensões variadas, também pode resultar em um melhor delineamento, ou seja, uma maior acurácia na delimitação das bordas dos polígonos representativos para os alvos escuros que se deseja detectar na cena processada. A Figura 5.2 compara a segmentação produzida para os dados de entrada em amplitude

e em intensidade no modo *Cartoon*, evidenciando exemplos de pequenos alvos escuros que foram melhor detectados em função do maior número de regiões geradas para os dados de entrada em intensidade. Considerando os resultados obtidos para o teste dos modos de processamento existentes no software MultiSeg, optou-se por utilizar - ao longo de toda a tese - o modo *Cartoon* com os dados de entrada em intensidade (0~65535).

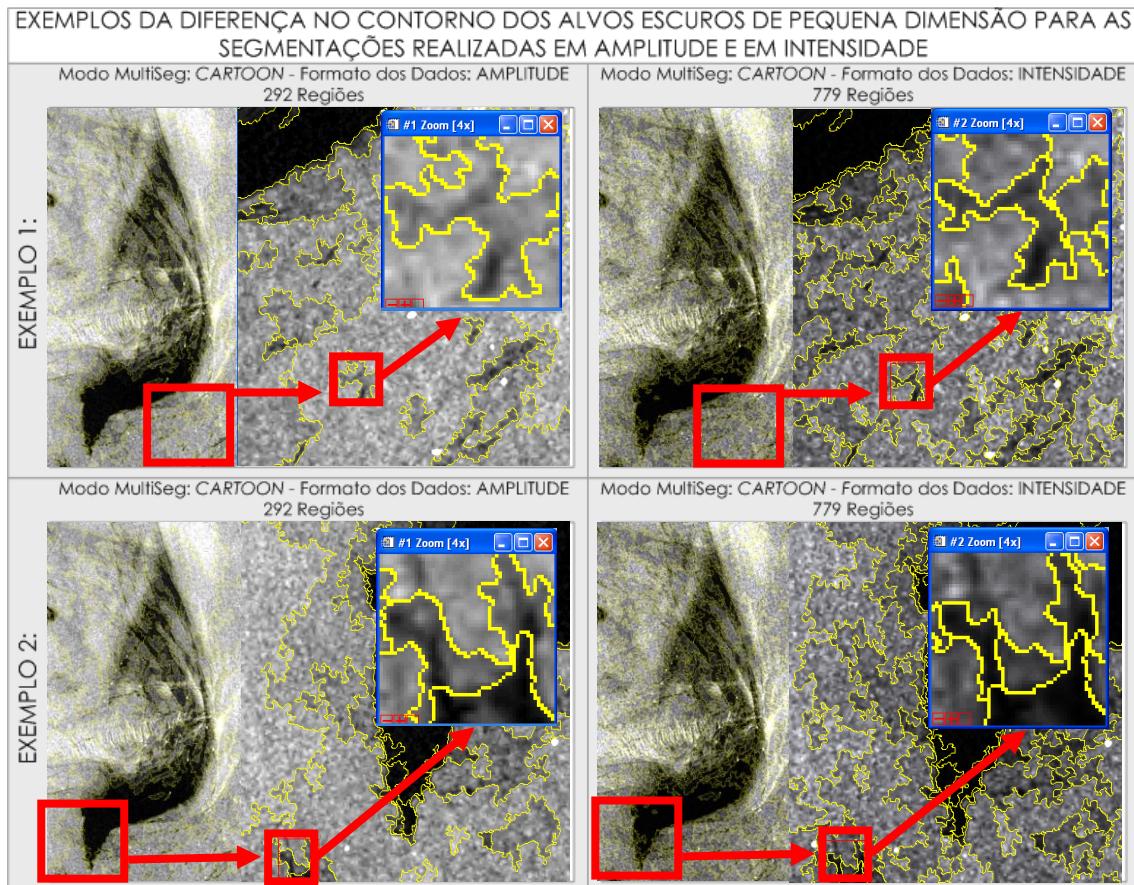


Figura 5.2 – Exemplos de segmentações nos modos Cartoon/Amplitude, e Cartoon/Intensidade.

Para determinar o espaçamento de pixel mais adequado, a ser adotado como padrão para os dados de entrada em um sistema de detecção automática de óleo e ocorrências ambientais, foram investigados os parâmetros de entrada exigidos pelo MultiSeg. Nesta etapa, para analisar o efeito da variação do espaçamento de pixel das imagens nos resultados das segmentações, foram definidos os valores que seriam utilizados para os seguintes parâmetros de entrada: a) área mínima: definição do número mínimo de pixels que o menor segmento gerado pode ter; b) número equivalente de *looks* (*Nel*); e; c) limiar de similaridade (*LS*).

Área Mínima: De acordo com BENTZ (2006), durante dois anos de monitoramento, a mínima área detectada como mancha de óleo em imagens de SAR

foi de $0,45\text{km}^2$. A experiência de campo obtida por inspetores ambientais⁶ apontou que, a mínima área de manchas de óleo identificável visualmente pode chegar a $0,25\text{km}^2$. Com base nesta experiência, buscando um critério mais rigoroso que possibilitasse a máxima detecção de regiões escuras nas imagens, o parâmetro de entrada para a área mínima foi definido como $0,25\text{km}^2$, permitindo a detecção de áreas menores que $0,45\text{km}^2$, as quais podem estar associadas com prováveis vazamentos de óleo no mar. O parâmetro de área mínima em *pixels* variou de acordo com o espaçamento de pixel da imagem a ser segmentada para representar, de maneira equivalente, a mesma área ($0,25\text{km}^2$). Para os espaçamentos de 50m, 100m, 150m e 200m foram necessários, respectivamente, 100 *pixels*, 25 *pixels*, 11 *pixels* e 6 *pixels* como área mínima para detectar uma área de $0,25\text{km}^2$.

Número Equivalente de *Looks* (Nel): O Nel é calculado em função da distribuição estatística dos NC em uma imagem de RADAR (FRERY *et al.*, 2007), sendo sempre menor ou igual ao Número de *Looks* (NI) da imagem. A compactação da imagem no formato *MrSid* altera as propriedades estatísticas das imagens de RADAR (item 4.1), produzindo valores incoerentes, muito maiores do que o NI da imagem. Como não foi possível calcular o Nel, e este parâmetro deve ser sempre menor ou igual ao NI da imagem, segmentações foram testadas com valores de Nel 2, 3 e 4, considerando que o NI da cena estudada era 4. Conforme verificado nos resultados apresentados na Figura 5.3, para um mesmo Nel observou-se que o número de regiões produzidas diminuiu com o aumento do espaçamento de pixel nas imagens. Todavia, para todos os espaçamentos estudados, quanto maior o Nel maior o número de regiões produzidas e quanto maior o número de regiões maior a possibilidade de detecção de alvos escuros de pequena dimensão em imagens SAR. Como o importante para a aplicação em questão é não deixar de detectar uma possível emergência ambiental, o Nel 4 foi considerado como parâmetro mais adequado. Importante observar a diferença no tempo de processamento exigido para um recorte de cena com espaçamento de 50m (cerca de 3min) em relação aos demais espaçamentos de pixel (menos de 1min). A Figura 5.3 ilustra os resultados obtidos para as segmentações realizadas no modo/formato *Cartoon/Intensidade* com Nel 2, 3 e 4, para os espaçamentos de pixel de 50m, 100m, 150m e 200m, evidenciando o número de regiões produzidas e o tempo de processamento exigido em cada caso.

⁶ Informação obtida através de Relatórios Técnicos da PETROBRAS e comunicação pessoal com os inspetores ambientais da PETROBRAS responsáveis pelo monitoramento e pelo combate de manchas de óleo em alto mar ao longo de 6 anos de experiência.

AVALIAÇÃO DO NÚMERO EQUIVALENTE DE LOOKS (NEL = 2, 3 OU 4) PARA O SOFTWARE MULTISEG: MODO/FORMATO CARTOON/INTENSIDADE

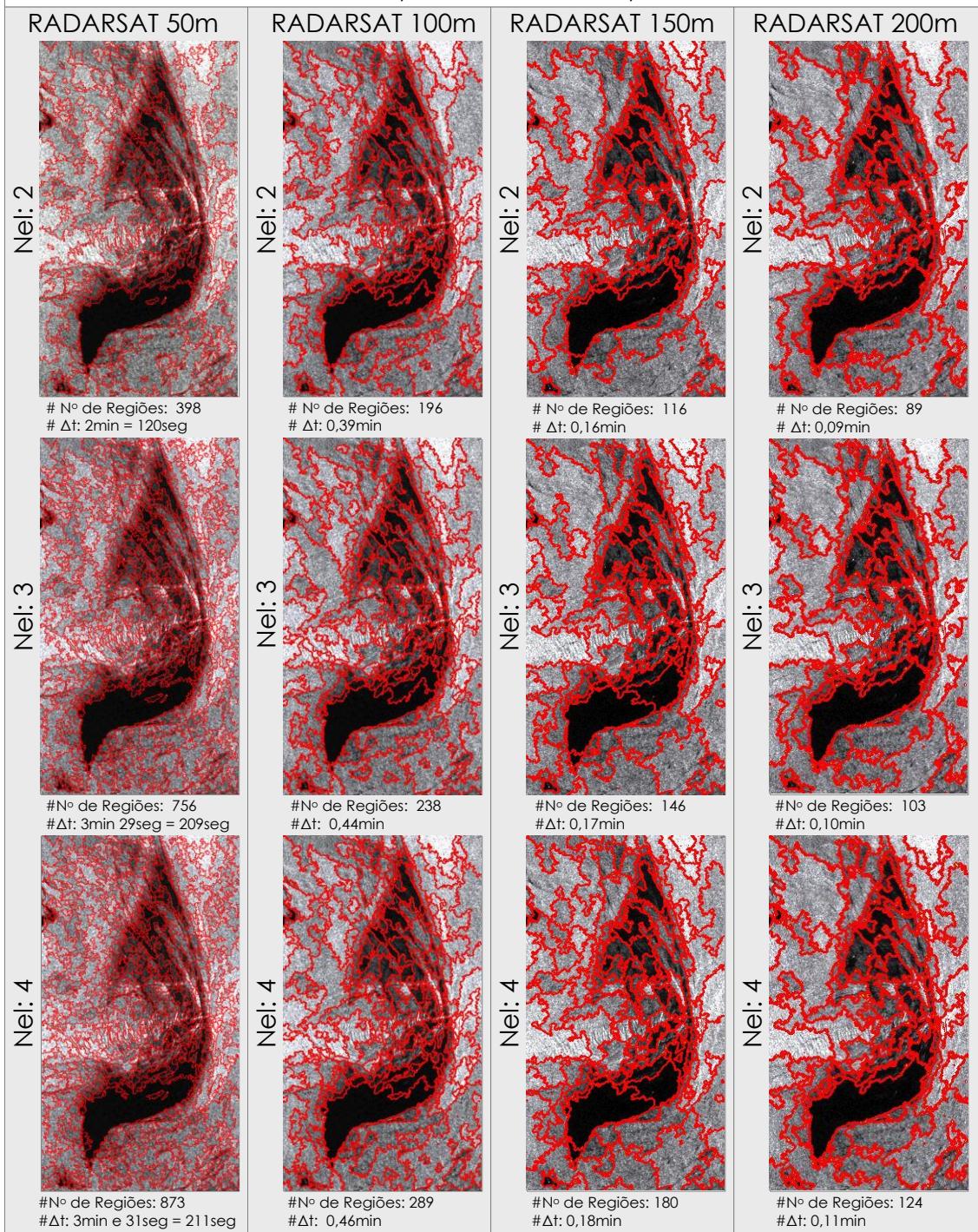


Figura 5.3 - Resultados obtidos para segmentações realizadas com Nel 2, 3 e 4, considerando todos os espaçamentos de pixel avaliados.

Limiar de Similaridade (LS): O LS é o parâmetro utilizado nos testes de crescimento e agrupamento de regiões. Para o software MultiSeg duas regiões distintas serão consideradas similares e serão agrupadas em uma única região quando a metade da diferença entre a média dos NC for menor do que o LS estabelecido (SOUZA, 2005). Desta forma, quanto menor for o LS maior será a

sensibilidade do segmentador para captar diferenças entre regiões distintas e espacialmente próximas, consequentemente, maior será o número de regiões produzidas pela segmentação, aumentando o potencial de detecção de alvos escuros na imagem, incluindo os de pequena dimensão.

Com base nas considerações acima, foi proposto um método padrão para o cálculo do LS que refletisse a menor distância entre a distribuição das regiões escuras e claras, resultando em um menor valor do LS calculado com base nas propriedades estatísticas e espaciais dos NC componentes da imagem.

Para calcular este parâmetro, considerando o recorte de cena utilizado nesta etapa da tese, uma segmentação supervisionada no software ENVI⁷ foi realizada onde o intérprete definiu, interativamente, o melhor limiar de corte ($NC_{Amplitude}=28$ ou $NC_{Intensidade}=784$) que permitiu ao segmentador detectar os polígonos mais representativos para os alvos escuros identificados. O resultado desta segmentação foi utilizado como uma máscara para gerar uma borda ao redor dos alvos escuros contidos na imagem. Ao redor desta borda um *buffer* com raio de 5 pixels para dentro dos alvos escuros e de 5 pixels para fora dos alvos escuros foi criado. Este *buffer* foi utilizado como máscara para extrair as propriedades estatísticas das regiões limítrofes claras e escuras presentes na imagem.

O cálculo proposto para o LS foi feito como base no *box plot* calculado para as regiões claras e escuras amostradas através do *buffer* gerado ao redor da borda dos alvos escuros. O critério adotado foi a subtração do valor obtido para o quarto quartil do *box plot* ($Q4_{REscura}$) calculado para as regiões escuras do valor obtido para o primeiro quartil do *box plot* ($Q1_{RClara}$) calculado para as regiões claras. A Figura 5.4 e 5.5 disponibilizam: a) o *box plot* calculado para as regiões escuras e claras amostradas pelo *buffer*, e; b) o método proposto como padrão para calcular o LS em dB.

⁷ O segmentador do ENVI utiliza técnicas de *Threshold* e filtragem aplicadas com base em um limiar de corte definido supervisionadamente pelo intérprete.

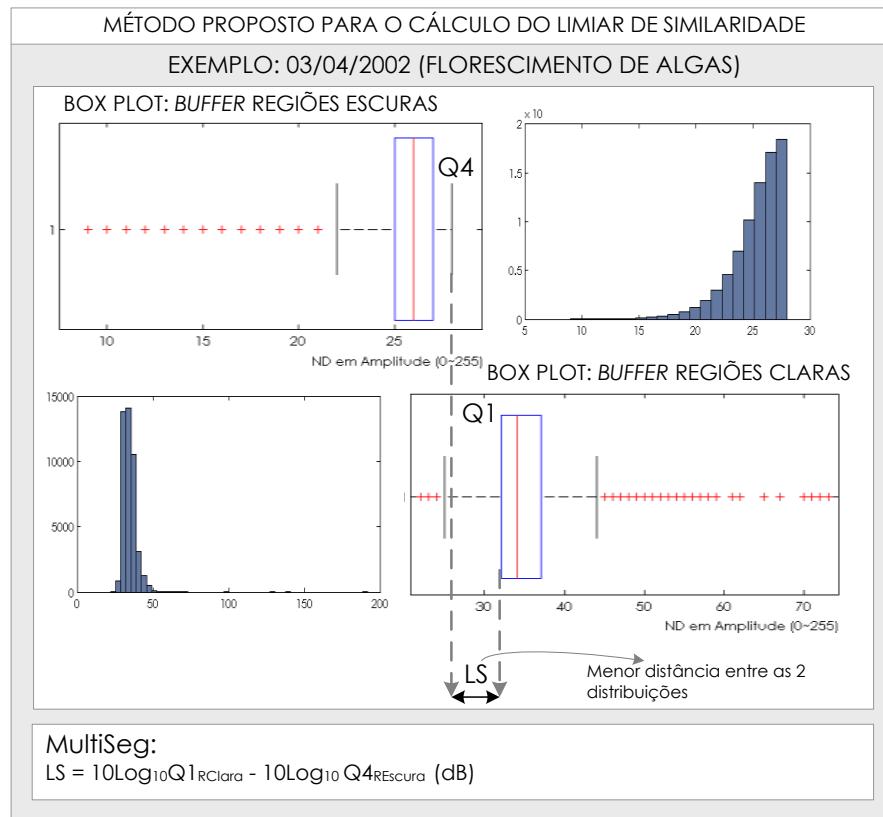


Figura 5.4 – Exemplo de *box plot* calculado para as regiões claras e escuras e método proposto para o cálculo do Limiar de Similaridade em dB.

Este método garante que, utilizando todos os quartis representantes das regiões escuras, o risco de não detectar uma provável mancha de óleo é reduzido sem deixar de considerar a maior parte dos dados referentes às regiões claras (75%). É possível garantir que as regiões escuras estão sendo priorizadas porque o valor do Q4 para o *box plot* das regiões escuras corresponde ao limiar de corte definido interativamente na segmentação do ENVI, o qual forneceu a melhor separação entre os alvos escuros e o oceano.

O critério proposto para cálculo do LS privilegia a detecção de regiões escuras em detrimento das regiões claras. Esta premissa considera que é preferível errar detectando regiões claras como escuras do que o contrário, detectando regiões escuras como claras, as quais podem estar associadas a um possível vazamento de óleo e, consequentemente, a uma situação de emergência ambiental. A Figura 5.5 ilustra o método proposto para o cálculo do LS com base nas amostras dos *buffers*.

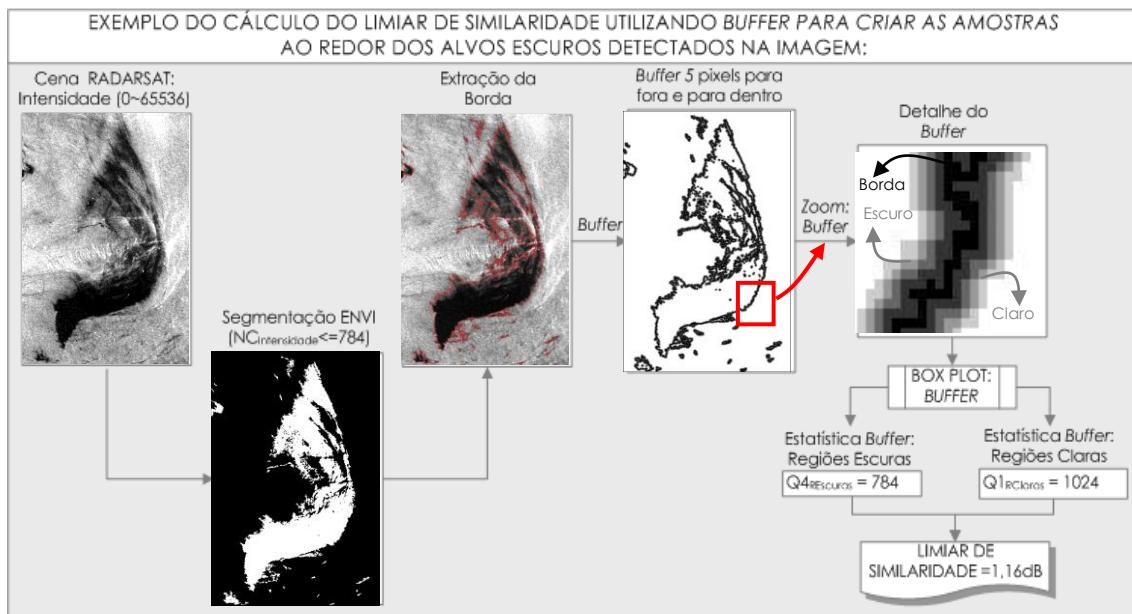


Figura 5.5 – Metodologia padrão proposta para calcular o Limiar de Similaridade.

O LS calculado com base nas amostras obtidas ao redor das bordas dos alvos escuros é mais efetivo do que amostras que consideram toda a área de alvos escuros e oceano existente na imagem. Amostras que consideram a área total das regiões escuras e claras possuem maior variância e não consideram a adjacência espacial das áreas limítrofes entre as distribuições. O cálculo dos LS a partir de amostras ao redor dos *buffers* considera somente as propriedades estatísticas peculiares às áreas de transição que se deseja separar. Na região de fronteira a média para os *pixels* no oceano será sempre menor do que a obtida a partir de amostras considerando toda a área oceânica. De forma contrária, a média para os *pixels* localizados para dentro da borda dos alvos escuros será sempre maior do que as obtidas a partir de amostras selecionadas nos alvos escuros como um todo. Como consequência, a distância entre as médias dos *pixels* claros e escuros diminui à medida que as amostras estão limitadas a região de fronteira. O cálculo do LS, considerando amostras globais e localizadas ao redor dos *buffers* para o recorte de cena utilizado evidenciou as observações acima: a) o LS calculado com base em amostras abrangendo o total de regiões escuras e claras da imagem, sendo a metade da diferença entre as médias, foi 3,16dB em intensidade e 10,34 em amplitude; b) o LS calculado a partir das amostras efetuadas nos *buffers*, sendo a metade da diferença entre as médias foi 1,26dB e 4,35 em amplitude, e; c) o LS calculado a partir das amostras efetuadas nos *buffers*, considerando os quadrantes do *box plot*, foi 1,16dB e 4 em amplitude. A comparação entre estes 3 itens evidenciou que os menores valores obtidos para o LS, tanto em intensidade quanto em amplitude, foram obtidos pelo método dos quartis.

O método proposto foi a solução encontrada para padronizar o cálculo do LS caracterizado como a menor distância entre as distribuições a partir de um critério único, evitando análises tendenciosas. Nas etapas seguintes da tese, onde foram utilizados diferentes recortes de cena e segmentadores, esta padronização foi importante para garantir uma comparação justa entre os resultados, à medida que os parâmetros de entrada dos segmentadores foram calculados com base em um mesmo critério.

5.1.2 – Determinação do Espaçamento de *Pixel*

Os valores indicados no item acima ($LS = 1,16\text{dB}$ e $Nel = 4$) foram utilizados como parâmetros de entrada para as segmentações realizadas no MultiSeg, considerando diferentes valores para o espaçamento de *pixel* 50m, 100m, 150m e 200m para o recorte de cena estudado. Conforme descrito no Capítulo 4, item 4.2, as regiões resultantes da segmentações foram agrupadas considerando como atributo a média dos NC, utilizando o limiar de corte definido interativamente na segmentação do ENVI, abaixo do qual foram detectados os alvos escuros ($NC_{\text{Intensidade}} \leq 784$) e acima do qual foi detectado o oceano ($NC_{\text{Intensidade}} > 784$). Os polígonos resultantes do agrupamento das regiões foram avaliados comparativamente com os polígonos de referência (*Phantons*) através das métricas de avaliação descritas no Capítulo 3, item 3.3.

A Tabela 5.1 contém as métricas de avaliação calculadas para os resultados obtidos para os diferentes espaçamentos de *pixel*. A Figura 5.6 apresenta os resultados obtidos para as segmentações, os agrupamentos efetuados por limiarização, a de intersecção dos alvos escuros entre a imagem classificada e o *Phantom* e os erros de omissão e de inclusão calculados em relação aos alvos escuros.

Tabela 5.1 – Métricas utilizadas para avaliação dos diferentes espaçamentos de *pixel*

Nº Polígonos do <i>Phantom</i> : 34	Resoluções	Métricas no Nível do <i>Pixel</i>				Métricas no Nível das Regiões						
		Tempo de Processamento	Exatidão: AEscuros	Intersecção: AEscuros	Omissão: AEscuros	Inclusão: AEscuros	Intersecção de Polígonos	Taxa de Int. de Polígonos	G forma	Fitxy	Fiti	Fitn
P50:	7,98min	0,81	0,89	0,11	0,10	14	0,41	0,78	0,98	1,00	0,92	0,24
P100:	1,36min	0,80	0,88	0,12	0,09	7	0,21	0,77	0,93	0,90	0,33	0,72
P150:	0,34min	0,78	0,85	0,15	0,10	4	0,12	0,74	0,91	0,90	0,31	0,75
P200:	0,16min	0,76	0,84	0,16	0,10	3	0,09	0,77	0,90	0,83	0,20	0,85

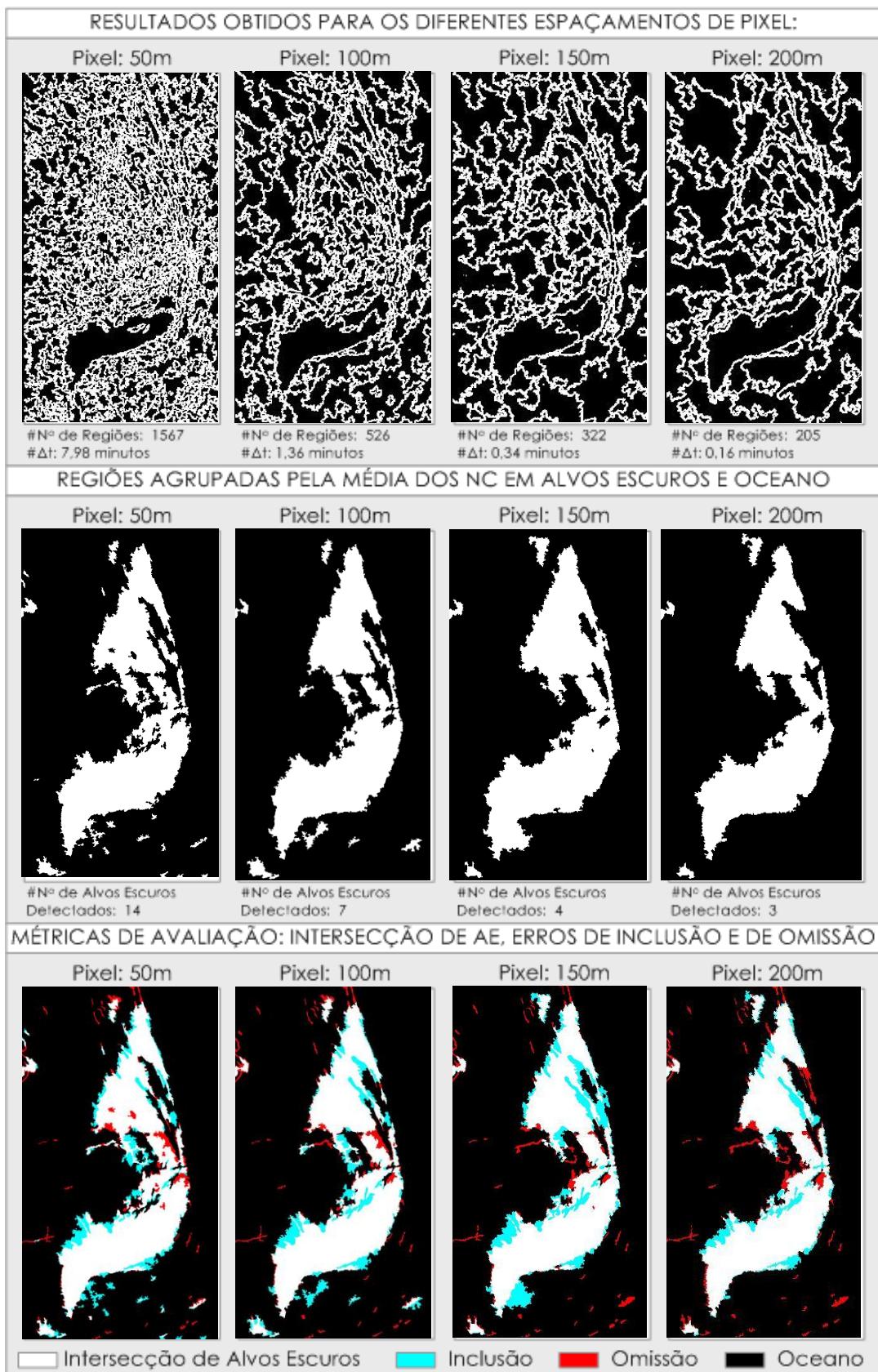


Figura 5.6 – Resultados obtidos para os diferentes espaçamentos de *pixel* avaliados.

As métricas calculadas (Tabela 5.1) evidenciaram que a diminuição gradativa do espaçamento de *pixel*/ resultou em uma redução na capacidade de detecção de

pequenas feições, bem como modificações no contorno das feições. Esta afirmação pôde ser constatada visualmente na Figura 5.6 e através da métrica de intersecção de polígonos, onde foi observada uma redução acentuada do número de polígonos detectados.

A distância euclidiana calculada a partir dos ajustes de forma, tamanho, intensidade e posição aumentou significativamente com o aumento do espaçamento de *pixel*. Este aumento foi causado pela redução gradativa de todos os ajustes, sendo o *Fitn* a métrica que mais apresentou diferença em relação ao espaçamento de 50m. Estes resultados indicaram que houve um leve deslocamento do centróide dos polígonos (*Fitxy*), da intensidade média dos NC que compõe os polígonos (*Fiti*) e da forma dos polígonos associados a alvos escuros e a oceano (*Gforma*). Entretanto, o tamanho destes polígonos, representado pelo *Fitn*, foi significativamente alterado. A análise visual dos resultados apresentados na Figura 5.6 confirmou esta métrica, sendo possível verificar que: a) com o aumento gradativo do espaçamento de *pixel* ocorreu uma perda da capacidade de detectar pequenas feições, o que gerou regiões com ajuste nulo (item 3.3.1), indicando polígonos existentes no *Phantom*, mas que deixaram de ser detectados na imagem classificada; b) o aumento da omissão de regiões que no *Phantom* eram alvos escuros e na classificada viraram oceano, também contribuiu para esta alteração no tamanho dos polígonos.

As métricas calculadas especificamente para os alvos escuros, a partir da matriz de confusão (Tabela 5.1) reforçaram as observações acima indicando, juntamente com o aumento do espaçamento de *pixel*: a) uma diminuição gradativa da Exatidão dos Alvos Escuros e da Fração de Intersecção dos Alvos Escuros em relação ao *Phantom*; b) um aumento gradativo da omissão de regiões escuras, agrupadas erroneamente como parte do oceano. Com relação à taxa de inclusão errônea de *pixels* de alvos escuros como oceano, praticamente não ocorreu variação.

Considerando que a máxima detecção de alvos escuros é essencial para minimizar o risco de que uma possível mancha de óleo não seja detectada, prejudicando as ações de contingenciamento o melhor espaçamento de *pixel* a ser adotado seria 50m, por fornecer o maior potencial de detecção de alvos escuros. Todavia, para que o sistema de detecção seja utilizado operacionalmente o tempo de processamento deve ser minimizado. Os resultados obtidos indicaram que o tempo de processamento para a imagem com 50m foi 6 vezes maior do que o tempo necessário para processar a imagem de 100m (Tabela 5.1). Esta diferença dificultaria o uso operacional das imagens com 50m de espaçamento de *pixel* uma vez que o tamanho

de uma cena RADARSAT completa (8413linhas x 7873colunas) é 42 vezes maior do que o recorte de cena utilizado (1671linhas x 922colunas). Testes preliminares foram efetuados tentando segmentar uma cena completa com 50m e o segmentador não conseguiu concluir o processamento.

Neste contexto, a utilização do espaçamento de *pixel* de 100m como padrão foi indicada por constituir uma alternativa para reduzir o tempo de processamento, considerando a necessidade operacional da aplicação. Esta escolha também foi influenciada pelas métricas disponíveis na Tabela 5.1, as quais indicaram que, dentre os demais espaçamentos (150m e 200m), o de 100m foi o que conseguiu definir melhor as bordas dos alvos escuros e detectar um maior número de pequenos alvos quando comparado com o espaçamento de 50m. O espaçamento de 100m foi o que ofereceu o melhor custo/benefício, mesmo considerando que o numero de polígonos detectados em relação ao *Phantom* diminuiu.

Outra evidência que reforçou a utilização do espaçamento de *pixel* de 100m foi a aquisição de imagens RADARSAT também no modo *ScanSAR Wide* (100m), sendo aconselhável padronizar o espaçamento de *pixel* de entrada das imagens.

A padronização do espaçamento de *pixel* em 100m também foi indicada por diversos autores, como uma solução adequada para a detecção de óleo na superfície do mar, por diminuir o custo computacional e funcionar como um filtro suavizando o ruído inerente às imagens SAR (SLOGGETT & JORY, 1994; CALABRESI *et al.*, 1999; FISCELLA *et al.*, 2000; DEL FRATE *et al.*, 2000; GASULL *et al.*, 2002; BREKKE & SOLBERG, 2005; BENTZ, 2006; KERAMITSOGLOU *et al.*, 2006;). Importante destacar que, apesar das indicações encontradas na literatura reforçarem a padronização do espaçamento de 100m, nenhum dos trabalhos consultados mostrou resultados comparando o ganho e a perda associado à utilização dos diferentes espaçamentos.

5.2 – Etapa 2: Determinação do Filtro/Janela

Para o mesmo recorte de cena utilizado na etapa anterior foram aplicados 8 filtros (Lee, Lee Enhanced, Frost, Frost Enhanced, Kuan, Gamma, Sigma e Mediana) com 3 dimensões de janela (3x3, 5x5 e 7x7), totalizando 27⁸ imagens filtradas com diferentes janelas. Após a aplicação dos filtros no recorte RADARSAT-1 com 100m de espaçamento de pixel - definido como mais adequado no item 5.1.2 – os recortes de cena filtrados foram: a) segmentados no software MultiSeg utilizando os mesmos parâmetros definidos anteriormente; b) agrupados considerando a média dos NC por limiarização, e; c) comparados através das métricas de avaliação definidas no Capítulo 3: item 3.3 em relação ao *Phantom* para 100m.

Após o agrupamento das regiões efetuado por limiarização, o número de alvos escuros detectados em relação aos alvos escuros considerados como referência no *Phantom* variou de 4 a 10 entre os 27 recortes de cena filtrados e 1 sem filtro. Considerando que, para a aplicação em questão, o importante é não deixar de detectar um possível vazamento de óleo, foram analisados os filtros que obtiveram a máxima detecção, 10 alvos escuros (Tabela 5.2). A tabela contendo as métricas obtidas para todas as filtragens realizadas está disponível para consulta no APÊNDICE II.

Tabela 5.2 – Métricas utilizadas para avaliação dos diferentes filtros/janela testados.

Nº Polígonos do <i>Phantom</i> : 34	Métricas no Nível do Pixel				Métricas no Nível das Regiões								
	Filtros/Janelas Avaliados	Tempo de Processamento	Exatidão: AEscuros	Flntersecção: AEscuros	Omissão: AEscuros	Inclusão: AEscuros	Intersecção de Polígonos	Taxa de Int. de Polígonos	G forma	Fitxy	Fiti	Fitn	Dist. Euclidiana
Sem Filtro		1,36min	0,80	0,88	0,12	0,09	7	0,21	0,77	0,93	0,90	0,33	0,72
Lee 3x3		0,83min	0,75	0,94	0,06	0,24	10	0,29	0,92	0,82	0,81	0,96	0,28
Mediana 3x3		0,76min	0,76	0,94	0,06	0,24	10	0,29	0,91	0,82	0,81	0,97	0,28
Sigma 3x3		0,89min	0,76	0,94	0,06	0,24	10	0,29	0,88	0,82	0,82	0,92	0,29
Kuan 3x3		0,75min	0,76	0,93	0,07	0,24	10	0,29	0,90	0,80	0,82	0,91	0,30

A Figura 5.7 apresenta os resultados para as segmentações, os agrupamentos e os erros de omissão e de inclusão para o recorte de cena sem filtro e filtrado com os filtros/janela Lee 3x3, Mediana 3x3, Sigma 3x3 e Kuan 3x3.

⁸ São 27 imagens porque para o filtro Gamma é necessário definir o número de Looks, no experimento aplicamos o filtro Gamma com as 3 janelas e com 1 e 3 looks, aumentando em 3 o número de imagens.

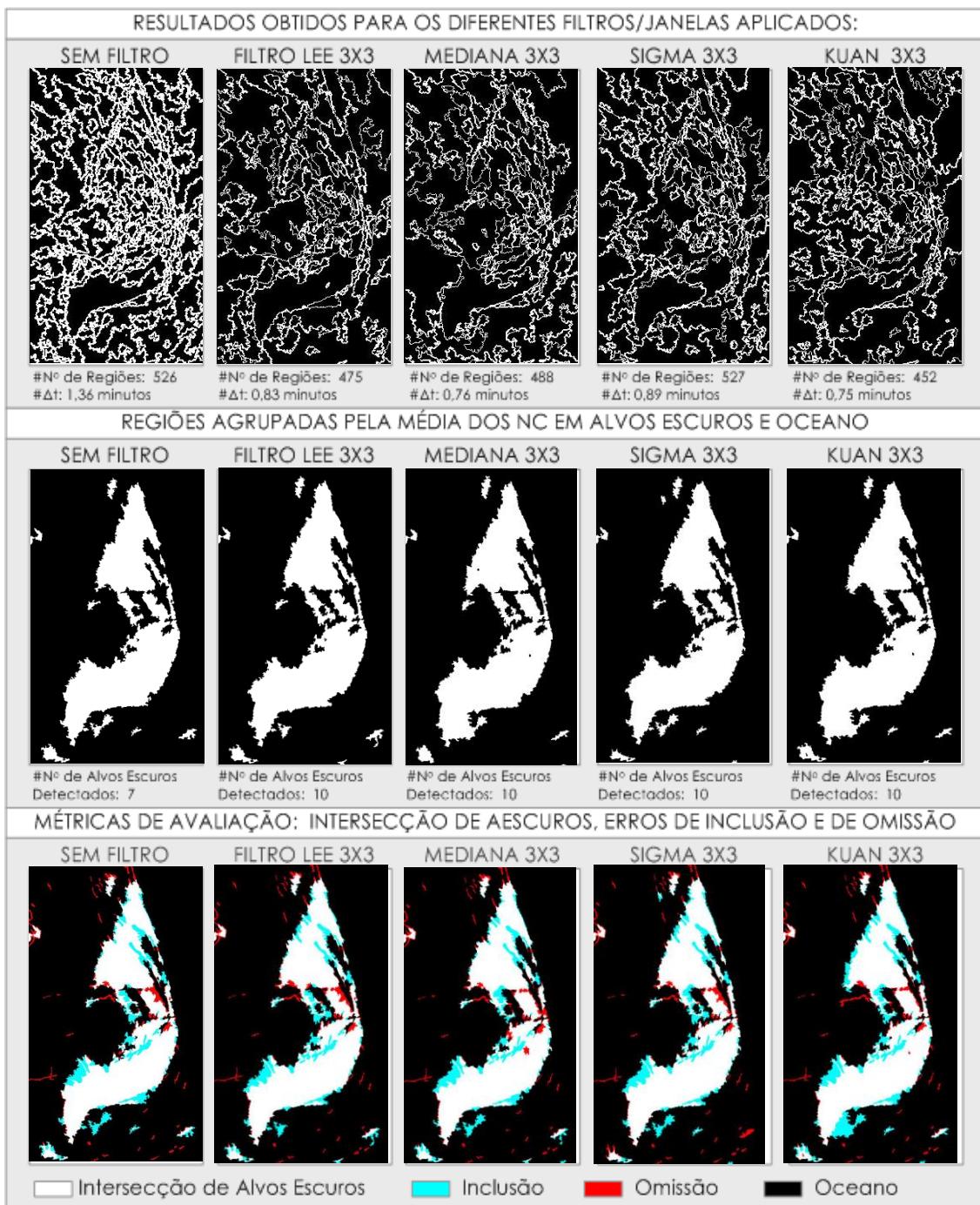


Figura 5.7 – Resultados obtidos para a imagem sem filtro e com filtro Lee 3x3, Mediana 3x3, Sigma 3x3 e Kuan 3x3.

Dentre as métricas consideradas (Tabela 5.2), o tempo de processamento não apresentou variação significativa entre os recortes filtrados, sendo desconsiderado como parâmetro para seleção do filtro/janela mais adequado à aplicação. As métricas de Exatidão dos Alvos Escuros e de Fração de Intersecção de Alvos Escuros aumentaram em relação ao recorte sem filtragem e permaneceram praticamente estáveis para todos os filtros. O erro de omissão apresentou o mesmo comportamento, diminuiu em relação ao recorte sem filtragem e foi praticamente o mesmo para todos

os resultados (0,06%). O erro de inclusão aumentou para os recortes filtrados, permanecendo estável entre os filtros analisados (0,24%). Considerando aos ajustes de forma (*Gforma*), tamanho (*Fitn*), intensidade (*Fitl*) e posição (*Fitxy*) é interessante notar que a aplicação do filtro melhorou os ajustes de forma (*Gforma*) e de tamanho (*Fitn*) em relação ao *Phantom* e piorou os ajustes de posição e de intensidade dos NC. Como a melhoria dos ajustes prevaleceu, a Distância Euclidiana foi bem menor para as imagens filtradas, indicando que a aplicação do filtro na etapa de pré-processamento das imagens pode ser padronizada em sistema automático. Outro resultado que reforça a utilização de filtragem é que, embora a diminuição do espaçamento de *pixel* tenha reduzido o potencial de detecção dos alvos escuros verificado no item 5.1.2, a aplicação de filtragem aumentou de 7 para 10 o número de alvos escuros detectados em relação ao *Phantom*.

Embora as menores distâncias euclidianas (0,28) tenham sido verificadas para os filtros Lee (3x3) e Mediana (3x3), as métricas variaram muito pouco entre os 4 melhores filtros indicados, sendo a determinação do melhor filtro a ser utilizado como padrão no sistema automático injusta considerando somente as métricas indicadas.

Desta forma, considerando a extensa literatura consultada sobre sistemas de detecção automática de óleo, foi constatado que o filtro mais comumente utilizado na etapa de pré-processamento das imagens é o Lee (BJERDE *et al.*, 1993; KARATHANASSI *et al.*, 2006; STATHAKIS *et al.*, 2006; BENTZ, 2006; TOPOUZELIS *et al.*, 2007a e 2007b; HANG *et al.*, 2009, TOPOUZELIS *et al.*, 2009). O filtro de Mediana foi utilizado em três das referências consultadas (MIRANDA *et al.*, 2004; CHANG *et al.*, 2005; HANG *et al.*, 2009), sendo que em uma delas o filtro de Mediana foi combinado com o filtro Lee (HANG *et al.*, 2009). O filtro Kuan foi utilizado em ZHENG *et al.* (2005). No fluxograma disponível no APÊNDICE I é possível verificar a diversidade de filtros utilizados pelos diferentes autores na etapa de pré-processamento.

Esta avaliação da influência dos diferentes filtros no potencial de detecção de alvos escuros nas imagens foi realizada apenas com um recorte de cena e pode ser que outros tipos de alvos apresentem resultado diferenciado. A fim de embasar com maior robustez uma decisão sobre a utilização ou não de filtragem como um procedimento padrão no pré-processamento das imagens, as análises efetuadas para determinação do segmentador mais adequado consideraram diferentes alvos, todos segmentados com e sem filtragem, sendo adotado o filtro Lee (3x3), indicado como o mais adequado.

5.3 – Etapa 3: Determinação do Segmentador

Para determinação do segmentador a ser utilizado como padrão no sistema de detecção automática de alvos escuros foram selecionados 11 recortes de cena RADARSAT contendo alvos escuros com dimensões, geometria e textura variados. Estes recortes de cenas com espaçamento de *pixel* de 100m sem filtragem e pré-processados com o filtro/janela Lee/3x3 – parâmetros determinados nas etapas anteriores - foram segmentados nos softwares MultiSeg e SPRING, resultando em 44 processamentos. Para cada um destes recortes foram elaborados manualmente os *Phantons*, imagens de referência utilizadas para a avaliação comparativa dos resultados produzidos. As regiões escuras e claras resultantes destas 44 segmentações foram agrupadas em alvos escuros e oceano, utilizando a técnica de limiarização, onde para cada imagem um limiar de corte foi definido interativamente pelo intérprete a fim de retornar o melhor resultado. Os alvos escuros resultantes do agrupamento das regiões escuras produzidas na segmentação foram avaliados comparativamente com os alvos escuros indicados como referência no *Phantom*, considerando as mesmas métricas de avaliação utilizadas nos itens anteriores.

A Figura 5.8 apresenta: a) os recortes de cena RADARSAT utilizados; b) os *Phantons* elaborados manualmente para cada recorte de cena; c) o histograma para os alvos escuros; d) o tamanho de cada recorte de cena (linhas x colunas); e) o NC indicado como limiar de corte utilizado no agrupamento definido interativamente, considerando os recortes de cena com e sem filtragem, e; f) a média e o desvio padrão dos alvos escuros.

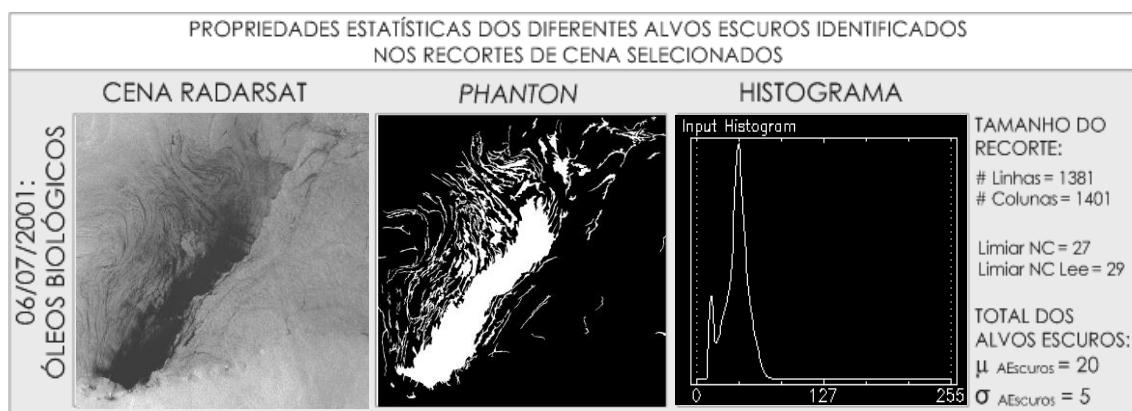


Figura 5.8 - Recortes de cena utilizados juntamente com os *Phantons* elaborados manualmente indicando os alvos escuros de referência.

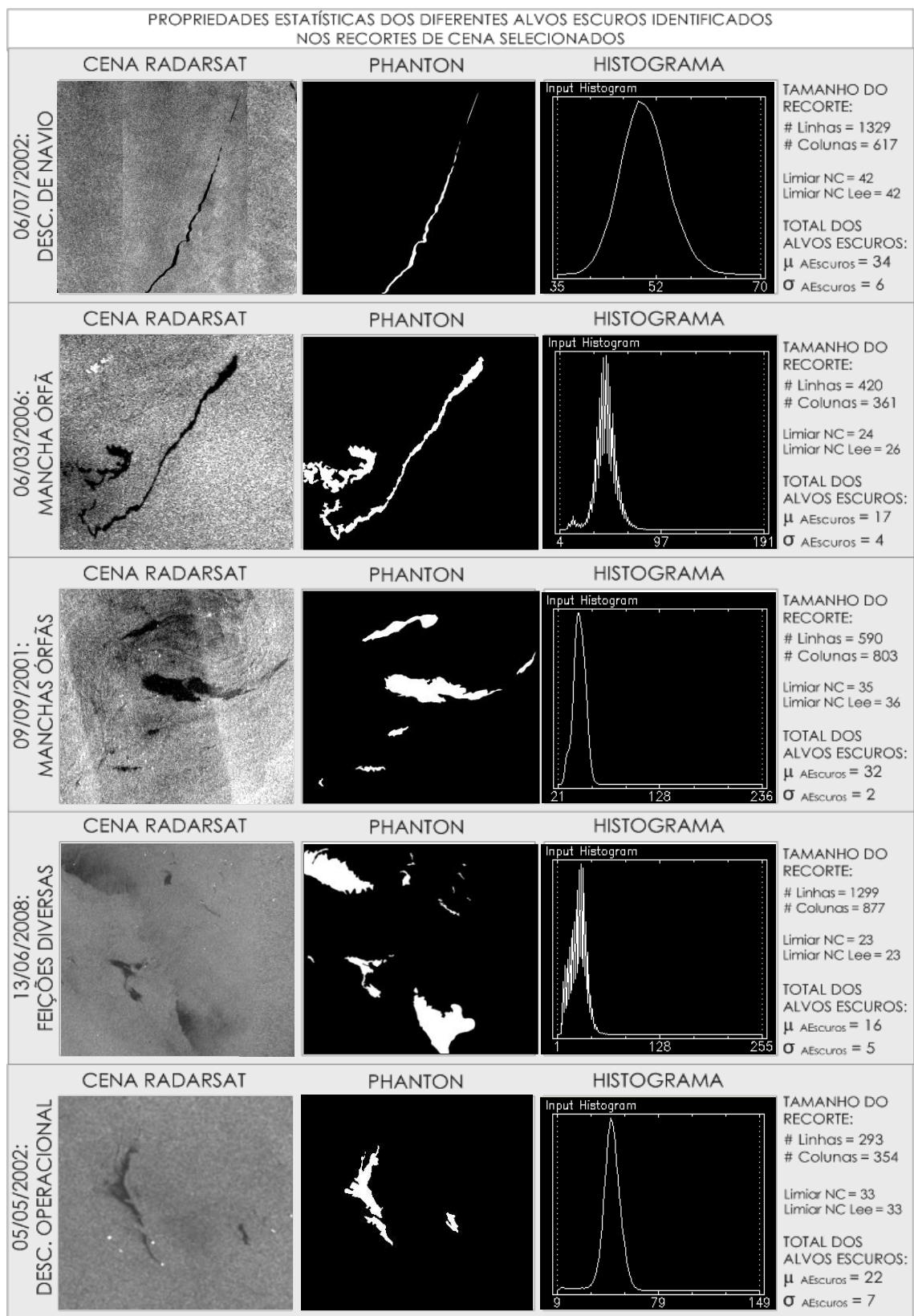


Figura 5.8 (Continuação) - Recortes de cena utilizados juntamente com os *Phantons* elaborados manualmente indicando os alvos escuros de referência.

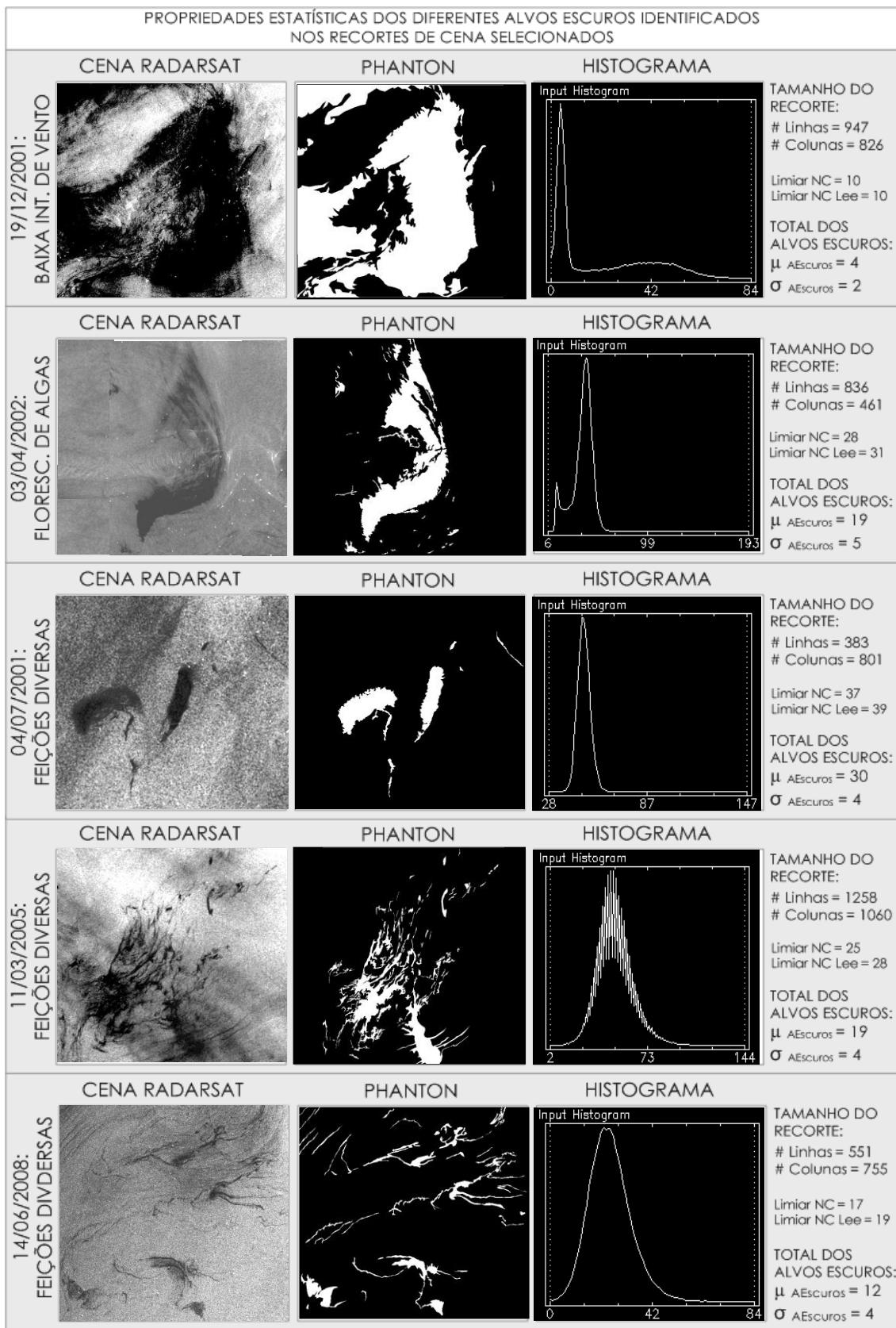


Figura 5.8 (Continuação) - Recortes de cena utilizados juntamente com os *Phantons* elaborados manualmente indicando os alvos escuros de referência.

Para realizar as segmentações o software MultiSeg exige como parâmetros de entrada a Área Mínima (AM) em pixels, o Número Equivalente de Looks (Nel) e o

Limiar de Similaridade (LS). A AM está atrelada ao espaçamento de *pixel* da imagem e foi definida nas etapas anteriores como 25 pixels para o espaçamento de 100m, o Nel foi definido como 4 e o LS foi calculado em dB para cada recorte de cena com e sem filtro de acordo com o método proposto no item 5.1.1. O software SPRING exige como parâmetros de entrada a AM em *pixels*, a mesma utilizada para o MultiSeg e o LS, calculado com o mesmo método em amplitude (0~255).

A Tabela 5.3 apresenta para todas as cenas estudadas com e sem filtro: a) a média dos NC em amplitude para as regiões claras e escuras amostradas pelo *buffer* ao redor da borda dos alvos escuros; b) os valores em amplitude para o 4º quartil dos alvos escuros e para o 1º quartil do oceano extraídos do *box plot* calculado, e; c) o LS calculado com base no *box plot* para o SPRING e para o MultiSeg.

Tabela 5.3 – Variáveis calculadas com a amostragem do *buffer* e LS calculados para o SPRING e para o MultiSeg, considerando a diferença entre os quartis do *box plot*.

Recortes Estudados		Dados Calculados para os <i>Buffers</i> em Amplitude				SPRING: LS em Amplitude	MultiSeg: LS em Decibéis
Data da Cena RADARSAT	Filtro Lee (3X3)	Média das REscuras	Média das RClaras	4º Quartil REscuras	1º Quartil RClaras	LS Box Plot: Q1 _{REscuras} - Q4 _{RClaras}	LS Box Plot: Q1 _{REscuras} - Q4 _{RClaras}
6/7/01	Sem Filtro	25,62	34,62	28	32	4	1,16
	Com Filtro	27,10	34,26	29	32	3	0,86
6/7/02	Sem Filtro	37,07	53,31	42	50	8	1,51
	Com Filtro	37,84	51,58	42	49	7	1,34
6/3/06	Sem Filtro	19,06	39,94	24	34	10	3,03
	Com Filtro	21,30	38,19	26	34	8	2,33
9/9/01	Sem Filtro	33,48	39,86	35	38	3	0,71
	Com Filtro	34,98	39,83	36	38	2	0,47
13/6/08	Sem Filtro	20,46	30,59	23	27	4	1,39
	Com Filtro	21,12	29,74	23	27	4	1,39
5/5/02	Sem Filtro	27,61	42,93	33	40	7	1,67
	Com Filtro	28,34	41,43	33	39	6	1,45
19/12/01	Sem Filtro	4,81	36,13	10	28	18	8,94
	Com Filtro	5,46	31,26	10	24	14	7,60
3/4/02	Sem Filtro	25,80	34,50	28	32	4	1,16
	Com Filtro	28,27	34,98	30	33	3	0,83
4/7/01	Sem Filtro	34,54	45,53	37	42	5	1,10
	Com Filtro	36,99	45,12	39	43	4	0,85
11/3/05	Sem Filtro	19,97	33,32	25	27	2	0,67
	Com Filtro	24,35	35,31	28	32	4	1,16
14/6/08	Sem Filtro	11,70	25,15	16	19	3	1,49
	Com Filtro	16,12	26,54	19	23	4	1,66

A Tabela 5.4 apresenta os resultados obtidos para os 44 processamentos realizados com e sem filtro para a determinação do segmentador mais adequado.

Tabela 5.4 – Métricas utilizadas para avaliação dos diferentes segmentadores aplicados nos recortes de cena com e sem filtragem.

CENAS	NAPP Phantom	Segmentadores SPRING e MULTISEG com e sem Filtro			Métricas no Nível do Pixel				Métricas no Nível das Regiões					
		Filtro	Segmentador	Tempo de Processamento	Exatidão: AEscuros	FInters.: AEscuros	Omissão: AEscuros	Inclusão: AEscuros	Int. de Polig.	G forma	Fitxy	Fiti	Fitn	Dist. Euclidiana
	Cena 1: 06/07/01 NPP = 213	Sem Filtro	SPRING:	1min e 12seg	0,62	0,70	0,30	0,13	0,57	0,92	0,84	0,88	0,15	0,88
		Multiseg:		8min e 30seg	0,62	0,73	0,27	0,19	0,41	0,96	0,82	0,87	0,10	0,93
		Com Filtro	SPRING:	3min e 25seg	0,61	0,79	0,21	0,29	0,62	0,93	0,85	0,89	0,14	0,88
		Multiseg:		1min e 46seg	0,61	0,79	0,21	0,29	0,52	0,95	0,83	0,77	0,11	0,93
	Cena 2: 06/07/02 NPP = 4	Sem Filtro	SPRING:	1min e 30seg	0,78	0,84	0,16	0,08	1,00	0,54	0,92	0,96	0,61	0,61
		Multiseg:		1min e 59seg	0,78	0,83	0,17	0,06	1,00	0,51	0,94	0,97	0,63	0,62
		Com Filtro	SPRING:	9seg	0,78	0,81	0,19	0,07	1,00	0,54	0,92	0,97	0,60	0,61
		Multiseg:		1min e 42seg	0,75	0,82	0,18	0,06	1,00	0,50	0,94	0,94	0,66	0,61
	Cena 3: 06/03/06 NPP = 3	Sem Filtro	SPRING:	4seg	0,46	0,48	0,52	0,03	1,00	0,77	0,81	0,83	0,29	0,78
		Multiseg:		21seg	0,43	0,44	0,56	0,02	1,00	0,78	0,81	0,83	0,29	0,79
		Com Filtro	SPRING:	3seg	0,46	0,52	0,48	0,04	1,00	0,78	0,82	0,85	0,29	0,78
		Multiseg:		19seg	0,50	0,48	0,52	0,03	1,00	0,78	0,81	0,69	0,29	0,83
	Cena 4: 09/09/01 NPP = 6	Sem Filtro	SPRING:	30seg	0,50	0,55	0,45	0,11	0,67	0,72	0,87	0,96	0,36	0,71
		Multiseg:		1min e 26seg	0,53	0,54	0,46	0,03	0,67	0,68	0,87	0,97	0,41	0,69
		Com Filtro	SPRING:	20seg	0,52	0,59	0,41	0,15	0,67	0,71	0,88	0,96	0,37	0,71
		Multiseg:		1min e 49seg	0,51	0,56	0,44	0,08	0,67	0,66	0,88	0,93	0,43	0,68
	Cena 5: 13/06/08 NPP = 18	Sem Filtro	SPRING:	3min e 36seg	0,50	0,63	0,37	0,26	0,78	0,73	0,92	0,91	0,38	0,69
		Multiseg:		8min e 29seg	0,57	0,66	0,34	0,16	0,61	0,73	0,90	0,88	0,36	0,71
		Com Filtro	SPRING:	15min e 17seg	0,58	0,76	0,24	0,24	0,61	0,76	0,90	0,90	0,32	0,73
		Multiseg:		6min e 15seg	0,61	0,68	0,32	0,18	0,50	0,78	0,88	0,87	0,30	0,75
	Cena 6: 05/05/02 NPP = 3	Sem Filtro	SPRING:	2seg	0,31	0,32	0,68	0,03	1,00	0,68	0,95	0,90	0,40	0,69
		Multiseg:		12 seg	0,26	0,26	0,74	0,00	0,67	0,70	0,95	0,89	0,37	0,71
		Com Filtro	SPRING:	2seg	0,30	0,28	0,72	0,02	1,00	0,68	0,95	0,91	0,40	0,69
		Multiseg:		11seg	0,27	0,30	0,70	0,00	1,00	0,69	0,95	0,81	0,38	0,72
	Cena 7: 19/12/01 NPP = 14	Sem Filtro	SPRING:	1min e 43seg	0,87	0,92	0,08	0,07	0,57	0,89	0,85	0,74	0,14	0,92
		Multiseg:		6min e 15seg	0,80	0,82	0,18	0,02	0,93	0,88	0,85	0,74	0,19	0,87
		Com Filtro	SPRING:	1min e 30seg	0,80	0,90	0,10	0,05	0,57	0,92	0,81	0,73	0,10	0,96
		Multiseg:		12min e 19seg	0,86	0,82	0,18	0,02	0,93	0,89	0,85	0,52	0,19	0,96
	Cena 8: 03/04/02 NPP = 36	Sem Filtro	SPRING:	10seg	0,79	0,88	0,12	0,12	0,36	0,86	0,83	0,89	0,24	0,80
		Multiseg:		1min e 36seg	0,80	0,88	0,12	0,09	0,21	0,77	0,93	0,90	0,33	0,72
		Com Filtro	SPRING:	8seg	0,75	0,95	0,05	0,29	0,58	0,82	0,88	0,91	0,31	0,73
		Multiseg:		0,83min	0,75	0,94	0,06	0,24	0,29	0,92	0,82	0,81	0,36	0,28
	Cena 9: 04/07/01 NPP = 10	Sem Filtro	SPRING:	7seg	0,64	0,68	0,32	0,07	0,50	0,76	0,88	0,94	0,31	0,74
		Multiseg:		40seg	0,60	0,63	0,37	0,05	0,50	0,77	0,87	0,94	0,29	0,76
		Com Filtro	SPRING:	7seg	0,68	0,74	0,26	0,10	0,90	0,86	0,92	0,97	0,46	0,65
		Multiseg:		43seg	0,67	0,74	0,26	0,08	0,80	0,71	0,90	0,92	0,39	0,69
	Cena 10: 11/03/05 NPP = 119	Sem Filtro	SPRING:	3min	0,49	0,59	0,41	0,19	0,53	0,93	0,88	0,87	0,17	0,85
		Multiseg:		2min	0,47	0,51	0,49	0,09	0,32	0,95	0,85	0,86	0,09	0,93
		Com Filtro	SPRING:	1min e 10seg	0,57	0,71	0,29	0,23	0,53	0,92	0,88	0,88	0,16	0,86
		Multiseg:		10min	0,57	0,71	0,29	0,25	0,38	0,94	0,86	0,87	0,11	0,91
	Cena 11: 14/06/08 NPP = 77	Sem Filtro	SPRING:	1min e 14seg	0,28	0,35	0,65	0,21	0,36	0,95	0,81	0,84	0,12	0,92
		Multiseg:		5min e 43seg	0,26	0,27	0,73	0,07	0,25	0,97	0,76	0,82	0,07	0,98
		Com Filtro	SPRING:	3seg	0,30	0,40	0,60	0,23	0,40	0,95	0,82	0,88	0,12	0,91
		Multiseg:		1min e 30seg	0,32	0,34	0,66	0,15	0,17	0,97	0,77	0,84	0,06	0,98

A Figura 5.9 apresenta os gráficos comparando as métricas de ajuste entre regiões e de matriz de confusão, ambas calculadas para os segmentadores MultiSeg e SPRING com e sem a aplicação de filtro.



Figura 5.9 - Gráficos comparando as métricas calculadas para os segmentadores MultiSeg e SPRING com e sem a aplicação de filtro.



Figura 5.9 – (Continuação) Gráficos comparando as métricas calculadas para os segmentadores MultiSeg e SPRING com e sem a aplicação de filtro.

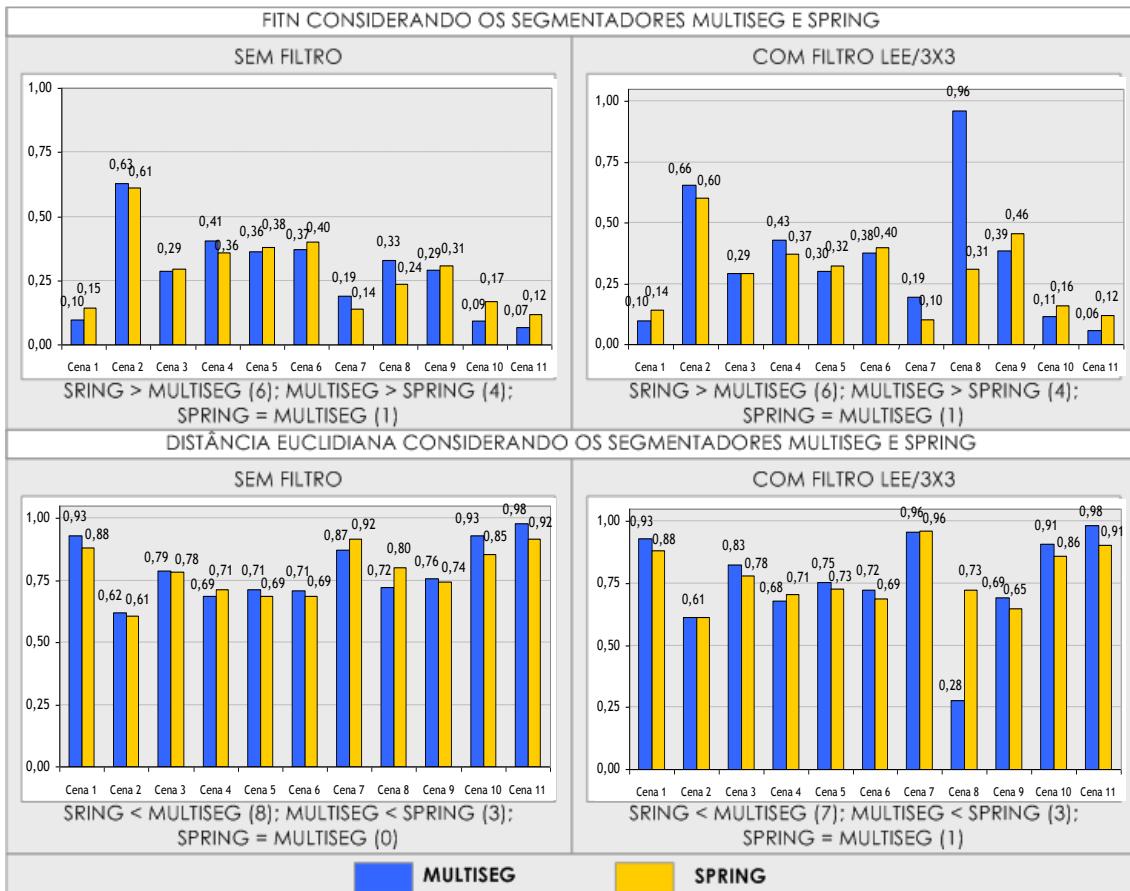


Figura 5.9 – (Continuação) Gráficos comparando as métricas calculadas para os segmentadores MultiSeg e SPRING com e sem a aplicação de filtro.

Com base nos resultados apresentados, primeiramente foi avaliado o efeito da aplicação do filtro/janela Lee/3x3 no potencial de detecção de alvos escuros nas imagens SAR e, posteriormente, foi avaliada a performance dos segmentadores MultiSeg e SPRING. Considerando as métricas referentes à Exatidão dos Alvos Escuros, à Fração de Intersecção de Alvos Escuros, Intersecção de Polígonos, Gforma, Fitxy, Fiti e Fitn, quanto maior o valor obtido, sendo mais próximo de 1, melhor o resultado. Para as métricas de Omissão, Inclusão e Distância Euclidiana quanto menor o valor obtido melhor o resultado.

O critério para a avaliação do efeito da aplicação do filtro/janela Lee/3X3 nos recortes de cena estudados considerou que a detecção de alvos escuros melhorou com a aplicação do filtro nos casos em que: a) houve um aumento no número de alvos escuros detectados após a aplicação do filtro, e; b) o número de alvos escuros permaneceu o mesmo e pelo menos uma das demais métricas de avaliação melhorou. A aplicação do filtro teve um efeito considerado nulo, ou neutro, nos casos em que o número de alvos escuros detectados não foi alterado mesmo que as demais métricas de avaliação tenham piorado. O critério considerou que a aplicação do filtro piorou os

resultados nos casos em que houve uma diminuição no número de alvos escuros detectados, mesmo que as demais métricas de avaliação tenham melhorado. A Tabela 5.5 apresenta a síntese dos resultados obtidos por recorte de cena estudado após a aplicação do filtro.

Tabela 5.5 - Síntese dos resultados obtidos para as métricas de avaliação por recorte de cena estudado após a aplicação do filtro Lee/3X3.

MÉTRICAS DE AVALIAÇÃO PARA A APLICAÇÃO DO FILTRO LEE (JANELA 3X3)								
Data da Imagem	Segmentador	Exatidão: AEscuros	Intersecção: AEscuros	Omissão: AEscuros	Inclusão: AEscuros	Intersecção de Polígonos	Distância Euclidiana	Efeito da Aplicação do Filtro
6/7/2001	Multiseg:	aumentou	aumentou	diminuiu	aumentou	aumentou	igual	MELHOROU
	SPRING:	diminuiu	aumentou	diminuiu	aumentou	aumentou	igual	MELHOROU
6/7/2002	Multiseg:	diminuiu	diminuiu	aumentou	igual	igual	diminuiu	NEUTRO
	SPRING:	igual	diminuiu	aumentou	diminuiu	igual	igual	NEUTRO
6/3/2003	Multiseg:	aumentou	aumentou	diminuiu	aumentou	igual	aumentou	MELHOROU
	SPRING:	igual	aumentou	diminuiu	aumentou	igual	igual	MELHOROU
9/9/2001	Multiseg:	diminuiu	aumentou	diminuiu	aumentou	igual	diminuiu	MELHOROU
	SPRING:	aumentou	aumentou	diminuiu	aumentou	igual	igual	MELHOROU
13/6/2008	Multiseg:	aumentou	aumentou	diminuiu	aumentou	diminuiu	aumentou	PIOROU
	SPRING:	aumentou	aumentou	diminuiu	diminuiu	diminuiu	aumentou	PIOROU
5/5/2002	Multiseg:	aumentou	aumentou	diminuiu	igual	aumentou	aumentou	MELHOROU
	SPRING:	diminuiu	diminuiu	aumentou	diminuiu	igual	igual	NEUTRO
19/12/2001	Multiseg:	aumentou	igual	igual	igual	igual	aumentou	NEUTRO
	SPRING:	diminuiu	diminuiu	aumentou	diminuiu	igual	aumentou	NEUTRO
3/4/2002	Multiseg:	diminuiu	aumentou	diminuiu	aumentou	aumentou	diminuiu	MELHOROU
	SPRING:	diminuiu	aumentou	diminuiu	aumentou	aumentou	diminuiu	MELHOROU
4/7/2001	Multiseg:	aumentou	aumentou	diminuiu	aumentou	aumentou	diminuiu	MELHOROU
	SPRING:	aumentou	aumentou	diminuiu	aumentou	aumentou	diminuiu	MELHOROU
11/3/2005	Multiseg:	aumentou	aumentou	diminuiu	aumentou	igual	diminuiu	MELHOROU
	SPRING:	aumentou	aumentou	diminuiu	aumentou	aumentou	aumentou	MELHOROU
14/6/2008	Multiseg:	aumentou	aumentou	diminuiu	aumentou	diminuiu	igual	PIOROU
	SPRING:	aumentou	aumentou	diminuiu	aumentou	aumentou	diminuiu	MELHOROU

█ Comportamento positivo
█ Comportamento negativo

De acordo com os resultados apresentados a aplicação do filtro antes da segmentação das imagens melhorou a detecção de alvos escuros nas imagens de RADAR na maioria dos casos. Para o software MultiSeg, 7 dentre os 11 recortes de cena processados tiveram a qualidade da detecção melhorada após a aplicação do filtro, 2 tiveram efeito neutro e em 2 casos o efeito foi negativo. Para o software SPRING, 7 dentre os 11 recortes de cena processados tiveram a qualidade da detecção melhorada após a aplicação do filtro, 3 casos tiveram efeito negativo e em 1 caso a aplicação do filtro não influenciou nos resultados. A Figura 5.10 apresenta um gráfico com a síntese dos resultados por software.

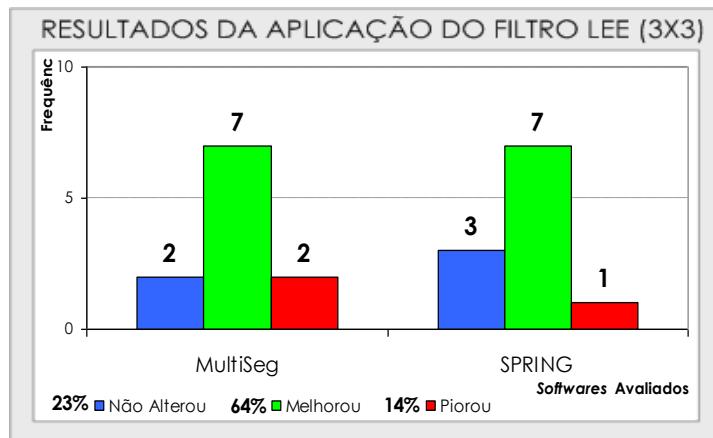


Figura 5.10 – Efeito da aplicação do Filtro Lee (3X3) no Potencial de Detecção de Alvos Escuros para os softwares utilizados.

Para validar os resultados obtidos o teste estatístico não paramétrico de Wilcox foi efetuado para verificar se a diferença entre as métricas de avaliação para os processamentos realizados para cada segmentador, com e sem filtragem, foi estatisticamente significativa. Valores de $p<0,05$ indicam que as variáveis comparadas são estatisticamente diferentes e, consequentemente, a aplicação ou não de filtragem no pré-processamento das imagens SAR influenciou na detecção de alvos escuros. A Tabela 5.6 disponibiliza o teste de Wilcox efetuado para os resultados obtidos com o segmentador MultiSeg com e sem filtragem e a Tabela 5.7 disponibiliza os resultados obtidos com o segmentador SPRING com e sem filtragem.

De acordo com os testes, dentre as 6 métricas de avaliação obtidas para o segmentador MultiSeg e para o segmentador SPRING, 3 foram consideradas estatisticamente diferentes comparando a aplicação ou não de filtragem na etapa de pré-processamento. Desta forma, para as métricas de erros de inclusão, erros de omissão e fração de intersecção dos alvos escuros, a aplicação do filtro faz diferença. Como o resultado da análise das métricas indicou que tanto para o MultiSeg, quanto para o SPRING a aplicação do filtro/janela Lee/3X3, na maioria dos recortes de cena utilizados, melhorou ou não alterou a detecção dos alvos escuros, optou-se por padronizar a aplicação do filtro Lee 3x3 na etapa de pré-processamento.

Tabela 5.6 - Teste de Wilcox efetuado para os resultados obtidos com o segmentador MultiSeg com e sem filtragem.

Teste Não Paramétrico de WilcoX: Processamentos com e sem Filtragem são considerados estatisticamente significativos os valores de p <0,05000				
Segmentador: MULTISEG	Valid N	T	Z	p-level
Exatidão dos AEscuros: Sem Filtro X Com Filtro	11	14,00	1,69	0,0912
FIntersecção de AEscuros: Sem Filtro X Com Filtro	11	2,00	2,76	0,0058
Omissão: Sem Filtro X Com Filtro	11	2,00	2,76	0,0058
Inclusão: Sem Filtro X Com Filtro	11	2,00	2,76	0,0058
Nº de Polígonos: Sem Filtro X Com Filtro	11	5,00	1,52	0,1282
Distância Euclidiana: Sem Filtro X Com Filtro	11	30,00	0,27	0,7897

Tabela 5.7 - Teste de Wilcox efetuado para os resultados obtidos com o segmentador SPRING com e sem filtragem.

Teste Não Paramétrico de WilcoX: Processamentos com e sem Filtragem são considerados estatisticamente significativos os valores de p <0,05000				
Segmentador: SPRING	Valid N	T	Z	p-level
Exatidão dos AEscuros: Sem Filtro X Com Filtro	11	24,00	0,80	0,42
FIntersecção de AEscuros: Sem Filtro X Com Filtro	11	1,00	2,85	0,0044
Omissão: Sem Filtro X Com Filtro	11	1,00	2,85	0,0044
Inclusão: Sem Filtro X Com Filtro	11	0,00	2,93	0,0033
Nº de Polígonos: Sem Filtro X Com Filtro	11	5,00	1,52	0,1282
Distância Euclidiana: Sem Filtro X Com Filtro	11	20,00	1,16	0,2477

Numa segunda fase foi avaliado o segmentador mais adequado considerando apenas os processamentos efetuados nos softwares MultiSeg e SPRING com aplicação de filtragem, uma vez que este procedimento foi indicado como padrão para o pré-processamento das imagens SAR.

O número de polígonos detectados foi o mesmo para os 2 softwares em 4 situações e em apenas em 1 caso o MultiSeg teve um desempenho melhor que o SPRING. Para a Fração de Intersecção dos Alvos Escuros, em 6 casos o SPRING forneceu as maiores taxas, em 3 situações os resultados foram iguais para os 2 softwares e em 2 o MultiSeg superou o SPRING (Figuras 5.10 e 5.11).

As menores taxas de Omissão foram obtidas pelo software SPRING em 6 das 11 situações analisadas. A Distância Euclidiana foi menor em 8 das 11 situações para o software SPRING. Importante destacar que embora o SPRING tenha omitido menos regiões escuras ele incluiu uma maior taxa de regiões claras erroneamente como alvos escuros em relação ao MultiSeg. O MultiSeg obteve os menores índices de inclusão errônea de regiões de oceano como alvos escuros, sendo melhor que o SPRING em 9 das 11 situações analisadas, todavia omitiu mais regiões escuras do que o SPRING.

Considerando as métricas por ajuste de regiões, os resultados do MultiSeg apresentaram menor ajuste de posição (*Fitxy*), de tamanho (*Fitn*) e de intensidade (*Fiti*). Entretanto o ajuste de forma (*Gforma*) foi melhor em relação ao SPRING, provavelmente porque o MultiSeg incluiu erroneamente menos regiões de oceano, o que melhorou seu ajuste de forma em relação ao *Phantom*.

Para a aplicação em questão o importante é não deixar de detectar um alvo escuro, sendo as métricas de número de polígonos detectados e taxa de omissão quesitos mais importantes do que a métrica de inclusão. Como o crítico é deixar de identificar uma possível emergência ambiental os erros de inclusão foram considerados toleráveis e os resultados obtidos indicaram o segmentador do software SPRING como mais adequado por apresentar as maiores taxas de detecção de alvos escuros e as menores taxas de omissão em relação ao MultiSeg. A Figura 5.11 sintetiza os resultados obtidos para o agrupamento das regiões resultantes das segmentações realizadas nos softwares SPRING e MultiSeg.

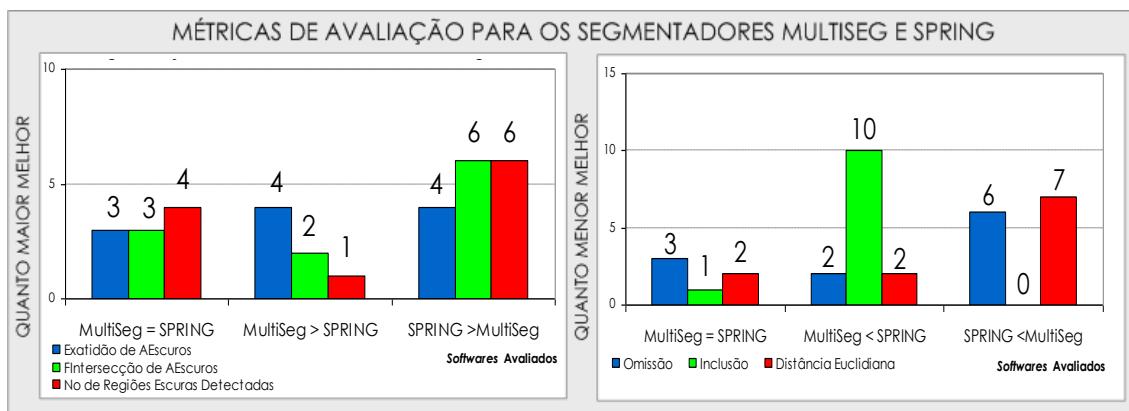


Figura 5.11 – Síntese dos resultados obtidos para o agrupamento das segmentações realizadas nos softwares SPRING e MultiSeg.

A Figura 5.12 apresenta a visualização dos resultados do agrupamento das regiões escuras e claras em alvos escuros e oceano após as segmentações realizadas nos softwares SPRING e MultiSeg com e sem filtragem.

RESULTADOS DAS SEGMENTAÇÕES REALIZADAS PARA OS DIFERENTES TIPOS DE ALVOS ESCUROS

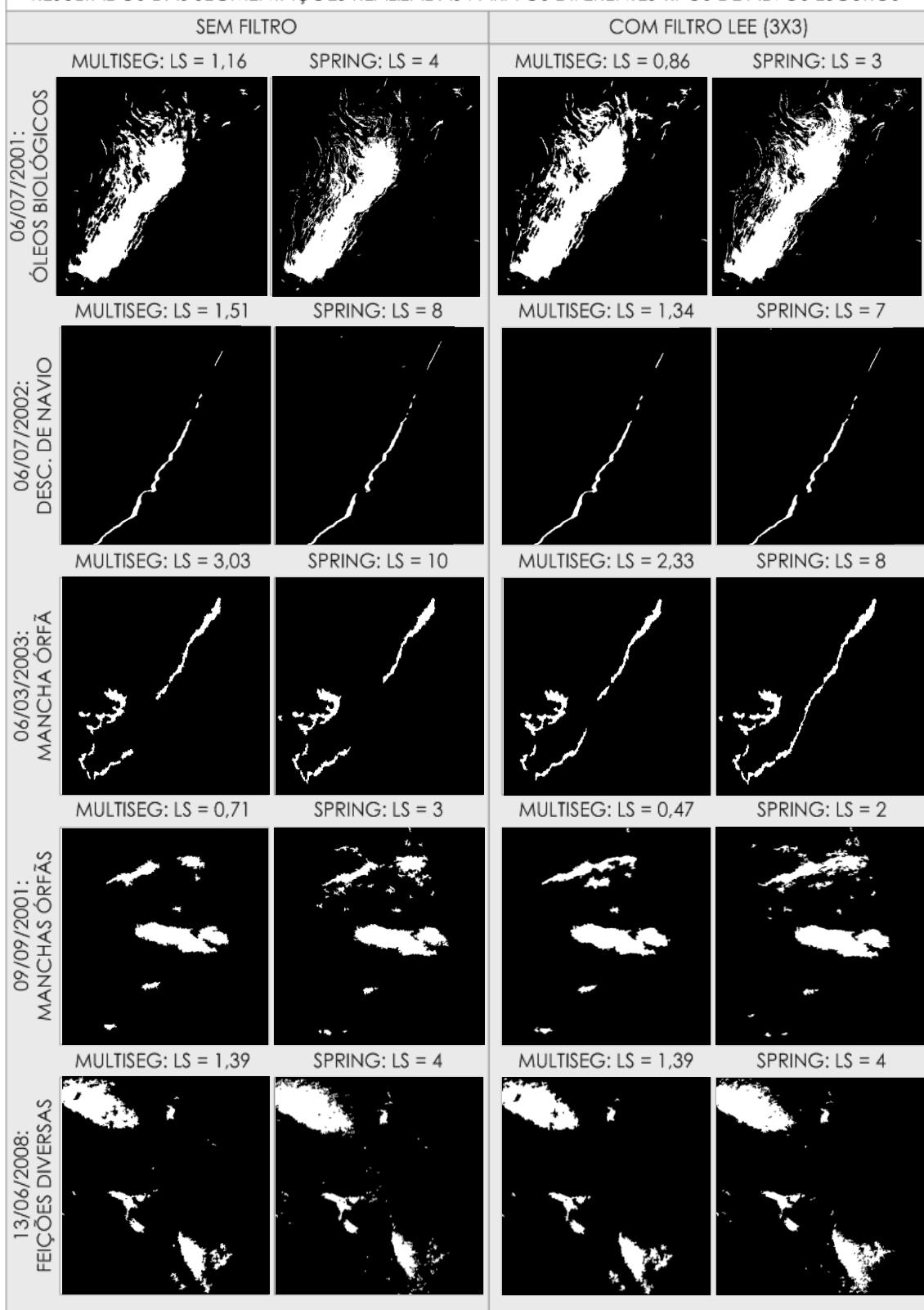


Figura 5.12 – Resultados do agrupamento das segmentações realizadas para os diferentes recortes de cena com e sem filtragem para os softwares SPRING e MultiSeg.

RESULTADOS DAS SEGMENTAÇÕES REALIZADAS PARA OS DIFERENTES TIPOS DE ALVOS ESCUROS

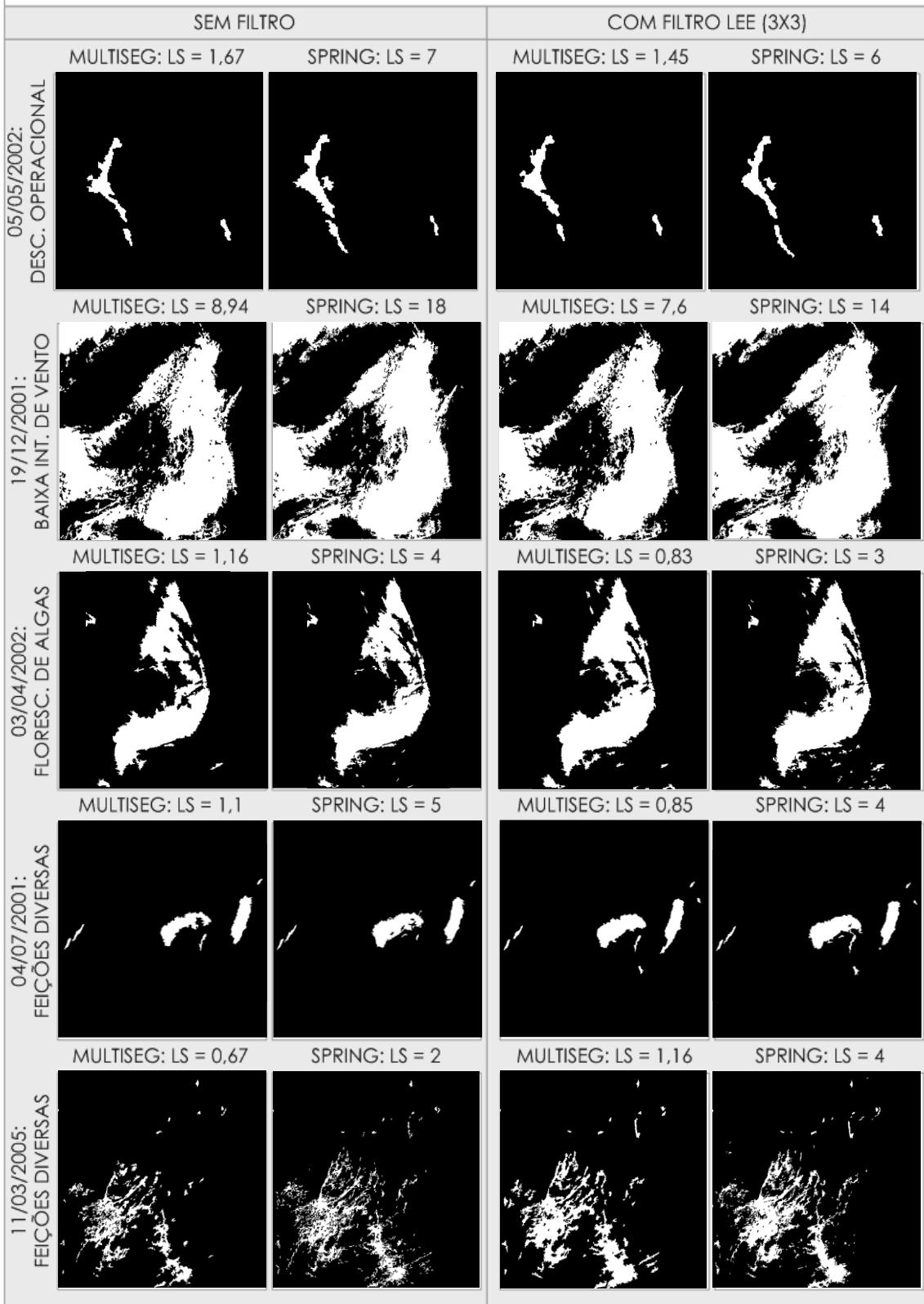


Figura 5.12 – (Continuação) Resultados do agrupamento das segmentações realizadas para os diferentes recortes de cena com e sem filtragem para os softwares SPRING e MultiSeg.

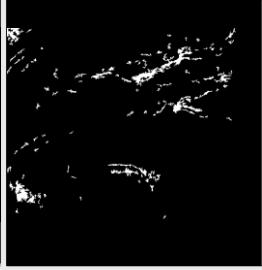
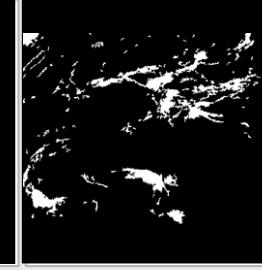
RESULTADOS DAS SEGMENTAÇÕES REALIZADAS PARA OS DIFERENTES TIPOS DE ALVOS ESCUROS			
SEM FILTRO		COM FILTRO LEE (3X3)	
MULTISEG: LS = 1,49	SPRING: LS = 3	MULTISEG: LS = 1,66	SPRING: LS = 4
			

Figura 5.12 – (Continuação) Resultados do agrupamento das segmentações realizadas para os diferentes recortes de cena com e sem filtragem para os softwares SPRING e MultiSeg.

A análise visual dos resultados obtidos para as cenas processadas com e sem filtro, utilizando os 2 segmentadores, reforçou que a aplicação de filtragem nas imagens melhora ou, praticamente não influencia na detecção de alvos escuros, para a maioria dos casos. A diminuição do ruído após a passagem do filtro foi perceptível visualmente no resultado final do agrupamento das segmentações para alguns exemplos.

Comparando-se visualmente os resultados obtidos pelos processamentos com filtragem realizados para os 2 segmentadores para cada recorte de cena estudado, foi identificado que em 6 casos o SPRING forneceu a geometria dos alvos escuros com maior semelhança em relação aos *Phantons*. O MultiSeg foi visualmente melhor em 5 casos. Desta forma, embora a análise das métricas de avaliação tenha indicado o SPRING como o segmentador mais adequado, a análise visual indica que tanto o SPRING como o MultiSeg foram eficazes na detecção dos alvos escuros em imagens SAR. Esta percepção foi reforçada pelo teste de Wilcox (Tabela 5.8), o qual evidenciou que não há diferença estatisticamente significativa entre as métricas de avaliação obtidas para os dois segmentadores.

Considerando o escopo da aplicação em questão para todos os casos, as diferenças verificadas não afetariam as ações de resposta, pois, com maior ou menor precisão, alvos escuros foram identificados e uma verificação de campo poderia ser acionada mesmo que os pequenos alvos não tivessem sido detectados. Estas evidências levariam a conclusão de que os segmentadores testados são equivalentes e tanto faz escolher um ou outro como mais adequado. Todavia, a acurácia da geometria obtida na etapa de detecção de alvos escuros influencia diretamente na eficácia do processo de classificação em um sistema de detecção automática de óleo e ocorrências ambientais. Neste contexto o SPRING forneceu as maiores taxas de

detecção de números de polígonos, ou seja, de alvos escuros, bem como, as menores Distâncias Euclidianas para as métricas de ajuste em relação aos *Phantons*.

Tabela 5.8 - Teste de Wilcox efetuado para os resultados obtidos com o segmentador MultiSeg e SPRING com filtragem.

Teste Não Paramétrico de WilcoX: Processamentos com Filtragem são considerados estatisticamente significativos os valores de p <0,05000				
MULTISEG X SPRING	Valid N	T	Z	p-level
Exatidão dos AEscuros: SPR X MSEG	11	23,00	0,89	0,3739
FIntersecção de AEscuros: SPR X MSEG	11	24,00	0,80	0,4236
Omissão: SPR X MSEG	11	24,00	0,80	0,4236
Inclusão: SPR X MSEG	11	19,00	1,24	0,2132
Nº de Poligonos: SPR X MSEG	11	8,00	1,01	0,3105
Distância Euclidiana: SPR X MSEG	11	18,00	1,33	0,1823

Portanto, apesar da proximidade dos resultados obtidos para ambos os segmentadores, o segmentador que detectou as regiões escuras nos recortes de cenas processados com maior proximidade das referências elaboradas nos *Phantons* foi o SPRING. Este resultado não era esperado, a medida que o SPRING é um segmentador desenvolvido com foco em imagens ópticas e o MultiSeg possui um módulo específico para imagens de RADAR. Provavelmente este fato ocorreu porque os recortes de cena utilizados perderam as características típicas de uma imagem RADAR devido a compressão realizada pelo formato *MrSid*, fato evidenciado no item 4.1. Embora os resultados apresentados sejam um primeiro indicativo, não são conclusivos quanto à indicação do segmentador mais adequado para a detecção automática de regiões escuras em imagens SAR. Para as demais etapas da tese o segmentador SPRING será utilizado para testar os diferentes métodos de clusterização, uma vez que apresentou resultados melhores em relação ao MultiSeg. O tempo de processamento foi outra evidência que reforçou esta escolha, sendo menor para o SPRING em 8 dos 11 recortes de cena processados.

Como seria inviável reprocessar todos os dados novamente nesta etapa da tese, a perspectiva futura é testar o segmentador MultiSeg em seu modo para imagens ópticas, comparando com os resultados do SPRING.

5.4 – Etapa 4: Determinação do Método de Clusterização

Nesta etapa da tese o objetivo foi avaliar a performance de técnicas automáticas de clusterização aplicadas à detecção de alvos escuros em imagens SAR. Neste contexto, a utilização de regiões, ao invés de *pixels*, possibilitou o cálculo e a associação de atributos diversos aos segmentos, ampliando significativamente as possibilidades de investigação e de resultados passíveis de serem obtidos através da clusterização (BAATZ *et al.*, 2003). Esta abordagem difere das metodologias disponíveis na literatura (APÊNDICE I), as quais utilizam o cálculo de diferentes atributos somente na etapa de classificação dos alvos em óleo ou assemelhados e não no agrupamento dos alvos escuros (TOPOUZELIS, 2008; KERAMITSOGLOU *et al.*, 2006; BREKKE & SOLBERG, 2005).

Conforme detalhado no Capítulo 3 (item 3.2), o processo de agrupamento ou clusterização pode ser considerado como uma “classificação” não supervisionada de padrões reconhecidos no conjunto de dados analisado. O reconhecimento destes padrões, intrínsecos ao espaço *n*-dimensional de atributos, é efetuado automaticamente pelo algoritmo sem o auxílio de amostras previamente classificadas ou de informações sobre o conjunto de dados (CHEN *et al.*, 1996; JAIN *et al.*, 1999). Desta forma, o processo de clusterização consiste numa metodologia capaz de explorar automaticamente a inter-relação existente entre o conjunto de atributos, reconhecendo padrões e agrupando-os, respectivamente, em diferentes grupos ou *clusters* de acordo com um critério de similaridade previamente definido.

As etapas integrantes de uma análise de clusterização são (JAIN & DUBES, 1988; JAIN *et al.*, 1999): a) o cálculo e a seleção de atributos; b) o reconhecimento de padrões; c) a clusterização considerando os padrões reconhecidos; d) a abstração dos resultados obtidos; e) a avaliação dos resultados e a validação do método indicado.

Tendo como base os itens componentes de uma típica análise de clusterização, a metodologia proposta foi dividida em 4 etapas: (Etapa I) cálculo de diferentes atributos, seguido da análise exploratória e da seleção dos atributos mais indicados para a detecção de alvos escuros; (Etapa II) experimentos de clusterização para extração de um conjunto de atributos padrão, o qual propicie a melhor predição automática possível para delinear a geometria dos avos escuros; (Etapa III) determinação do método padrão, definindo, o conjunto adequado de atributos, o método de clusterização adequado e uma regra de decisão que possibilite a automatização do sistema, e; (Etapa IV) validação dos resultados e do método

proposto com a utilização de um novo conjunto de dados. A Figura 5.13 ilustra de maneira sintética as etapas da metodologia proposta nesta última etapa da tese.



Figura 5.13 – Etapas integrantes da metodologia proposta nesta última etapa da tese.

Como não há um consenso entre os pesquisadores sobre qual abordagem, manual, semi-automática ou automática, representa o melhor custo/benefício para operacionalizar a detecção de óleo e assemelhados em imagens SAR (SOLBERG & BREKKE, 2008; TOPOUZELIS, 2008; SEASAR, 2010), uma comparação entre estas diferentes abordagens foi realizada para definir com maior rigor o procedimento mais eficaz.

5.4.1 – Cálculo dos Atributos

Dentre as etapas integrantes do processo de clusterização (item 5.4), o cálculo juntamente com a seleção de atributos, é considerada a etapa mais importante (JAIL *et al.*, 1999). A qualidade dos resultados da clusterização é completamente dependente da qualidade dos dados de entrada, o que confere importância estratégica à definição do conjunto de atributos (TOPOUZELIS, 2008; SOLBERG & BREKKE, 2008; STATHAKIS *et al.*, 2006; BENTZ, 2006).

O cálculo de atributos, denominado “*feature extraction*” na literatura especializada, consiste na conversão ou transformação dos dados de entrada em novos atributos, inserindo novas dimensões no espaço de atributos. Esta maior variabilidade dos dados de entrada confere ao clusterizador mais possibilidades de investigação e de combinações para convergir e encontrar uma solução ótima para o problema estudado.

Os atributos freqüentemente utilizados na etapa de classificação de óleo e assemelhados são agrupados em 4 categorias (BENTZ *et al.*, 2009; TOPOUZELIS, 2008; STATHAKIS *et al.*, 2006; KARATHANASSI *et al.*, 2006; MONTALI *et al.*, 2006; BENTZ, 2006; BREKKE & SOLBERG, 2005; MARGHANY, 2001; DEL FRATE *et al.*, 2000; FISCELLA *et al.*, 2000; TOPOUZELIS *et al.*, 2000; ESPEDAL, 1999; SOLBERG, 1999; SOLBERG & VOLDEN, 1997): a) Espectrais ou Estatísticos: extraídos a partir dos NC das imagens de satélite; b) Geométricos: extraídos da própria geometria dos

polígonos delineados pela segmentação sendo divididos em genéricos/poligonais e *skeletons*. Os geométricos genéricos/poligonais constituem métricas extraídas dos polígonos como um todo como área, perímetro, dentre outras. Os *skeletons* são métricas que indicam a estrutura interna dos polígonos, sendo calculadas a partir de uma triangulação efetuada internamente a partir das bordas dos polígonos, utilizando o método de Delaunay (BAATZ et al., 2003); c) Texturais: calculados a partir da matriz de co-ocorrência dos níveis de cinza (HARALICK et al., 1979), indicando a correlação espacial existente entre *pixels* vizinhos, e; d) Contextuais: relacionados com o cenário existente no momento da aquisição da imagem (ex: proximidade dos alvos de interesse de fontes poluidoras como plataformas, dutos e/ou embarcações, dentre outros). Atributos contextuais e relacionados à heterogeneidade e ao contraste dos alvos escuros em relação ao oceano não foram utilizados considerando que, nesta etapa da tese, a localização e a dimensão dos alvos escuros não são conhecidas.

Desta forma, para cada região clara ou escura individualizada no processo de segmentação, 36 atributos foram calculados tendo como base os 11 recortes de cena pré-processados com o filtro Lee 3X3⁹. Dos 36 atributos, 16 foram calculados a partir da sobreposição das regiões sobre os *pixels* da imagem, sendo 8 estatísticos e 8 texturais, e 20 foram calculados a partir das propriedades geométricas dos polígonos resultantes das segmentações. Detalhamento técnico sobre a definição dos atributos geométricos pode ser encontrado no manual do DEFINIENS (2006) e sobre os atributos texturais em HARALICK et al. (1973, 1979).

Os atributos estatísticos - calculados no Matlab - e os texturais e geométricos - calculados no *Definiens* - foram exportados no formato tabular e compatibilizados em um mesmo arquivo através de identificadores únicos associados a cada região. Os atributos foram normalizados, sendo cada valor obtido para um determinado atributo em uma região, subtraído pelo mínimo valor obtido e dividido pela diferença entre o máximo e o mínimo.

Para que os processamentos fossem controlados, a análise exploratória, a seleção dos atributos e a avaliação dos resultados obtidos pelas clusterizações foram efetuadas considerando a geometria dos *Phantons* como referência. A proposta considerou que, avaliando os atributos em relação a uma verdade previamente conhecida, seria possível identificar com maior eficácia o conjunto de dados mais adequado para retornar *clusters* representativos e, consequentemente, mais similares às imagens de referência (*Phantons*).

⁹ Detalhes sobre o formato dos dados utilizados estão disponíveis no Capítulo 4: item 4.1.

Para tal foi necessário criar um atributo categórico dependente denominado Classe do *Phantom*. Este atributo foi calculado, com o objetivo de associar a cada região resultante da segmentação, um código indicando a sua localização em relação aos alvos escuros ou ao oceano definidos no *Phantom*. Entretanto, em imagens SAR, o limiar que separa um alvo escuro do oceano não é exato (ZHENG *et al.*, 2005), as bordas não apresentam contornos bem definidos, consequentemente, a sobreposição dos limites das regiões resultantes das segmentações não necessariamente coincide com o limite das regiões definidas no *Phantom*. Desta forma, tanto o *Phantom* elaborado manualmente com base na experiência do intérprete quanto as regiões resultantes das segmentações apresentam imprecisão com relação à definição das bordas que separam um alvo escuro do oceano.

Na pesquisa realizada foi valorizada a experiência do intérprete e o atributo Classe do *Phantom* foi calculado sobrepondo o *Phantom* às regiões da segmentação. O resultado numérico foi uma graduação de 0 a 1 que representa quanto uma determinada região se encontra dentro ou fora de um alvo escuro definido no *Phantom*. Para gerar o atributo categórico foi necessário mapear estes valores [0 a 1] em 4 classes, considerando os respectivos intervalos de sobreposição: a) região clara (Código 1) – sobreposição maior que 0 e menor ou igual a 0,25; b) região de transição entre o claro e o escuro (Código 2) – sobreposição maior que 0,25 e menor que 0,50; c) região de transição entre o escuro e o claro (Código 3) – sobreposição maior ou igual a 0,50 e menor que 0,75, e; d) região escura (Código 4) - sobreposição maior ou igual a 0,75 e menor ou igual a 1. A Figura 5.14 ilustra o método proposto para a definição do Atributo Classe do *Phantom*.

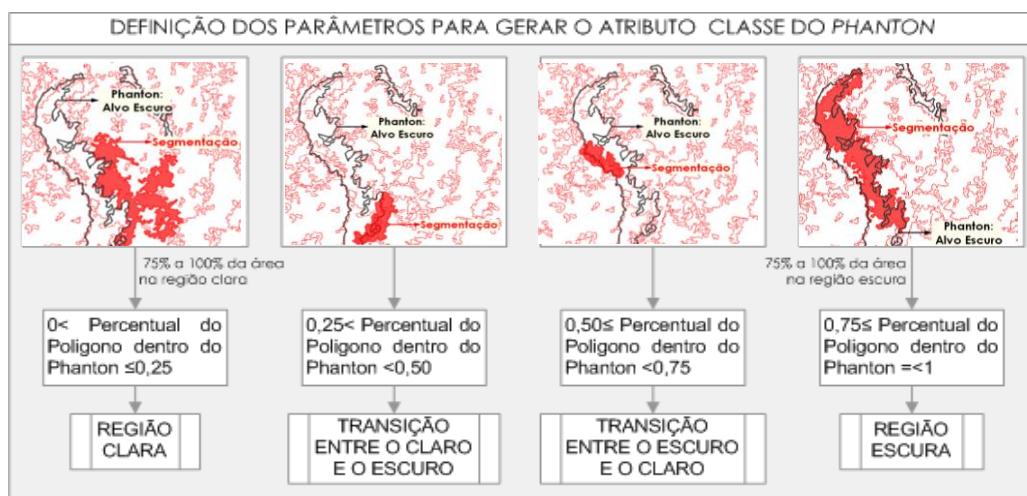


Figura 5.14 – Método proposto para a definição do Atributo Classe do *Phantom*.

Embora os atributos estatísticos e texturais tenham maior chance de contribuir para um bom resultado na clusterização, os atributos geométricos também foram

considerados. Os diferentes tipos de alvos escuros - associados a derramamento de óleo ou a eventos meteo-oceanográficos - possuem particularidades relacionadas com sua forma e tamanho capazes de discriminar estes diferentes tipos de eventos. Estas propriedades descritas por BENTZ (2006) suscitaram a importância de pesquisar se algum dos inúmeros atributos de forma, calculados pelo software *Definiens*, poderia contribuir para melhorar os resultados da clusterização. Como o foco da tese foi justamente a detecção da geometria de alvos escuros, uma investigação sobre a influência dos atributos associados à forma no processo de clusterização não poderia deixar de ser realizada.

5.4.2 – Análise Exploratória e Seleção de Atributos

Não existe uma solução pronta capaz de indicar quais os melhores atributos para serem utilizados em um processo de clusterização automática (JAIN *et al.*, 1999). De acordo com HUGHES (1968), a partir de certo número de atributos, a inserção de novas variáveis ao modelo insere pouco ou nenhuma informação podendo, em casos específicos, confundir e até mesmo piorar a performance do clusterizador. Por outro lado, independente do número de atributos e da sua ordem de inserção no modelo, a manutenção de atributo(s) estatisticamente não significativo(s) no conjunto de dados também prejudica a performance do clusterizador. O número e os tipos de atributos não são fixos e variam de acordo com a natureza dos dados e do problema estudado, sendo a análise exploratória e a seleção de atributos procedimentos fundamentais para auxiliar na definição destes parâmetros.

Apesar da diversidade de atributos utilizada na literatura consultada, pesquisas sistemáticas comparando a eficiência de cada atributo para a detecção de óleo e assemelhados são raras (TOPOUZELIS *et al.*, 2009; BENTZ *et al.*, 2007; RAMALHO & MEDEIROS, 2007; STATHAKIS *et al.*, 2006; KARATHANASSI *et al.*, 2006; BENTZ, 2006). A seleção dos atributos é feita de modo arbitrário, com base em conhecimento teórico sem utilizar um método para a seleção de atributos (TOPOUZELIS, 2008; STATHAKIS *et al.*, 2006; KUBAT *et al.*, 1998). Desta forma, muitos dos atributos utilizados pelos especialistas não contribuem efetivamente para otimizar os resultados da classificação (KUBAT *et al.*, 1998).

O conhecimento prévio do usuário sobre as propriedades do conjunto de dados e a natureza do problema estudado também é fundamental para a seleção de um conjunto adequado de atributos (JAIN & DUBES, 1988). Neste contexto, técnicas de análise exploratória de dados e seleção de atributos foram utilizadas para reduzir o

espaço de atributos, com o objetivo de identificar os parâmetros mais adequados para a detecção de alvos escuros em imagens SAR.

Para avaliar a significância estatística da relação existente entre os 36 atributos e a variável dependente Classe do *Phantom*, as técnicas estatísticas mono e multivariadas utilizadas foram: a) histograma, *box plot* e espacialização de atributos; b) matriz de correlação, e; d) análise de principais componentes. O módulo *Variable Screening*, implementado no software *Statistica* especificamente para a seleção de atributos, foi utilizado para indicar a ordem de importância das variáveis a serem utilizadas nas clusterizações. A Figura 5.15 detalha os atributos e procedimentos estatísticos utilizados na análise exploratória e seleção de atributos.

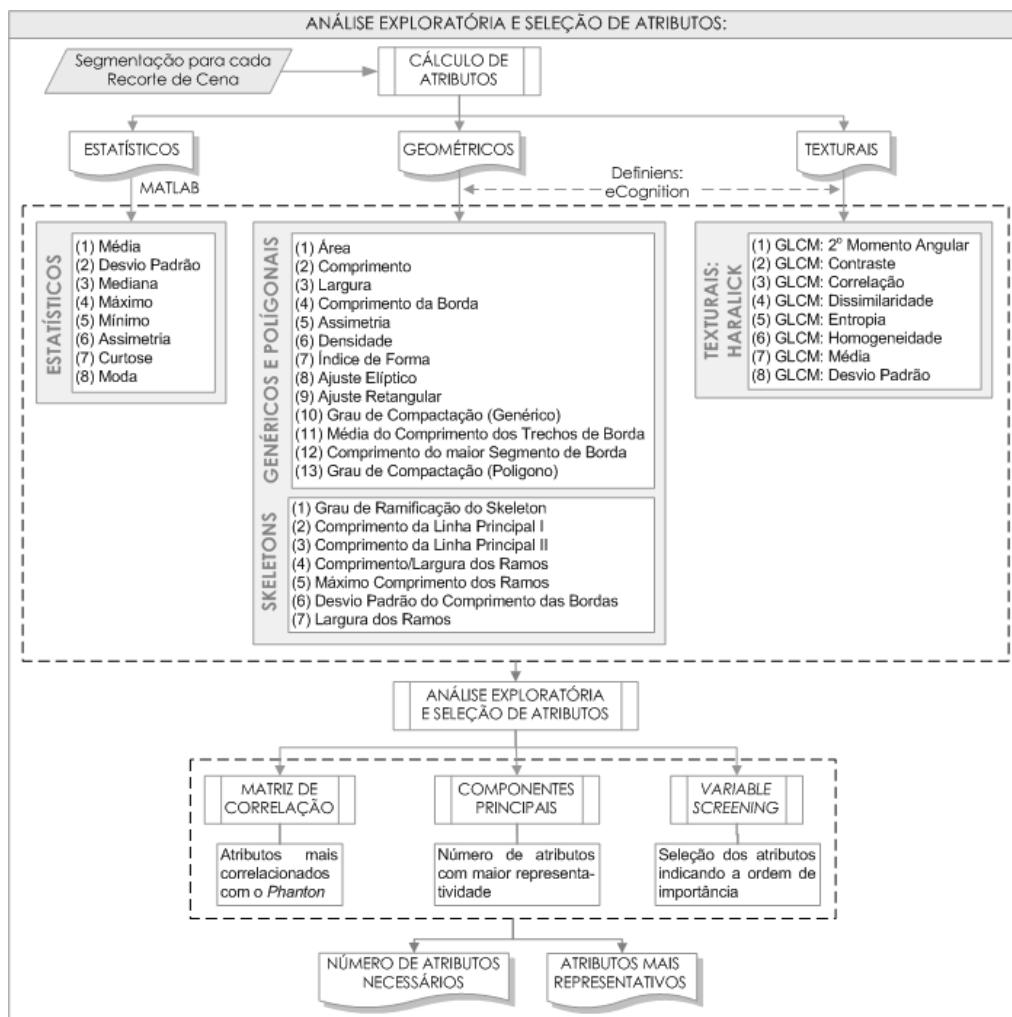


Figura 5.15 – Atributos e procedimentos utilizados na análise exploratória e seleção de atributos.

5.4.2.1 – Análise de Matriz de Correlação

O coeficiente de correlação de *Spearman* foi considerado mais adequado por ser não-paramétrico e avaliar a relação entre duas variáveis sem assumir

comportamento linear e distribuição normal como no caso da correlação de Pearson (SIEGEL, 1977). De forma diferente da correlação de Pearson, aplicável somente a dados numéricos, a correlação de Spearman pode ser aplicada a conjuntos de dados que apresentem também variáveis categóricas ordinais, como é o caso do atributo Classe do *Phantom*.

A matriz de correlação foi processada para cada uma das 11 cenas estudadas, considerando separadamente os diferentes tipos de atributos estatísticos, geométricos (subdivididos em gerais/poligonais e *skeletons*) e texturais em relação ao atributo categórico Classe do *Phantom*. No total foram calculadas 44 matrizes de correlação, com o objetivo de identificar quais os atributos mais correlacionados com o *Phantom*. Este capítulo apresenta somente a síntese dos resultados obtidos, as matrizes de correlação calculadas estão disponíveis para consulta no APÊNDICE III.

Atributos Estatísticos

A Tabela 5.9 contém: a) os coeficientes de correlação dos atributos estatísticos em relação ao *Phantom* para as 11 cenas processadas; b) a máxima correlação obtida por atributo dentre as 11 cenas; c) a média das correlações obtidas por atributo, e; d) a frequência com que a correlações não foram significativas em relação ao *Phantom*.

Tabela 5.9 – Coeficientes de Correlação dos atributos estatísticos em relação ao *Phantom*.

Coeficientes de Correlação dos Atributos Estatísticos em relação ao <i>Phantom</i> para as 11 Cenas Estudadas								
Classe <i>Phantom</i>	Média	Desvio Padrão	Mediana	Máximo	Mínimo	Assimetria	Curtos e	Moda
<i>Phantom: Cena 1</i>	-0,6456*	0,2632*	-0,6443*	-0,6098*	-0,6508*	-0,08167*	0,07862*	-0,6419*
<i>Phantom: Cena 2</i>	-0,5175*	0,3629*	-0,5247*	-0,5139*	-0,5259*	-0,2605*	-0,2638*	-0,5222*
<i>Phantom: Cena 3</i>	-0,6928*	0,2068*	-0,6938*	-0,6375*	-0,6857*	-0,198*	-0,2726*	-0,6838*
<i>Phantom: Cena 4</i>	-0,3308*	-0,003535	-0,3323*	-0,2883*	-0,3324*	0,009045	0,05319*	-0,3294*
<i>Phantom: Cena 5</i>	-0,3326*	-0,1637*	-0,3342*	-0,3054*	-0,3246*	-0,02625	0,04689*	-0,3369*
<i>Phantom: Cena 6</i>	-0,6103*	0,5661*	-0,6066*	-0,548*	-0,6171*	-0,08852	-0,2073	-0,6095*
<i>Phantom: Cena 7</i>	-0,05068	0,1218*	-0,1006*	-0,01522	0,07788	0,2887*	0,09038*	-0,2171*
<i>Phantom: Cena 8</i>	0,6792*	-0,08605*	0,6847*	0,5971*	0,6751*	-0,004693	0,09213*	0,6834*
<i>Phantom: Cena 9</i>	-0,5079*	0,3258*	-0,5118*	-0,4329*	-0,5085*	-0,01858	-0,1593*	-0,5126*
<i>Phantom: Cena 10</i>	-0,5154*	-0,2751*	-0,5185*	-0,4573*	-0,4941*	-0,0741*	-0,1455*	-0,5166*
<i>Phantom: Cena 11</i>	-0,5585*	-0,3267*	-0,557*	-0,5037*	-0,55*	-0,2556*	-0,2536*	-0,5528*
Correlação Máxima	0,6928	0,5661	0,6938	0,6375	0,6857	0,2887	0,2726	0,6838
Correlação Média	0,4947	0,2456	0,5008	0,4463	0,4947	0,1187	0,1512	0,5097
Correlações Não Significativas	1	1	0	1	1	5	1	0
* As Correlações Máxima e Média foram calculadas com os valores de (<i>r</i>) em módulo								
* As Correlações significativas com o <i>Phantom</i> (<i>p</i> < 0,05) estão assinaladas em vermelho								

Dentre as 88 medidas de correlação obtidas em relação ao *Phantom*, nenhuma atingiu (*r*) acima de 0,7, 39 (44%) apresentaram valores entre 0,5 e 0,7 e a maioria dos casos (49: 56%) tiveram valores de (*r*) menores que 0,5. De maneira geral, estes valores representam uma baixa relação dos atributos estatísticos com o *Phantom*.

Apesar destes baixos valores de (r), 78 casos (89%) foram considerados estatisticamente significativos em relação ao *Phantom*, tendo sido possível identificar quais os atributos que melhor se relacionaram com o *Phantom*.

A maior máxima correlação (em módulo) foi obtida pela Mediana em relação ao *Phantom*, seguida dos atributos de Mínimo, Moda e Média. As maiores médias de correlação foram obtidas pela Moda, Mediana, Média e Mínimo. Consequentemente, estes atributos apresentaram os maiores valores de (r) observados em relação ao *Phantom*, sendo estes estatisticamente não significativos em apenas duas cenas para os atributos de Média e de Mínimo.

Os atributos de Assimetria e Curtose foram os menos correlacionados com o *Phantom* com coeficientes médios de correlação de 0,12 e 0,15, respectivamente, sendo a Assimetria o atributo que apresentou maior número de correlações não significativas (5). A Tabela 5.10 sintetiza os resultados obtidos, indicando: a) a freqüência acumulada com que cada um dos atributos foi correlacionado com o *Phantom* do 1º (maiores coeficientes) ao 8º lugar (menores coeficientes), e; b) o ranking final dos atributos mais indicados do 1º ao 8º lugar.

Tabela 5.10 - Análise Síntese da Matriz de Correlação para as 11 Cenas: Atributos Estatísticos.

Freqüência Acumulada com que Cada um dos Atributos Estatísticos foi Correlacionado com o <i>Phantom</i> do 1º ao 8º lugar								
Atributos	1º	2º	3º	4º	5º	6º	7º	8º
Média	1	4	8	10	0	0	11	0
Desvio Padrão	0	0	1	0	2	8	10	11
Mediana	3	8	9	11	0	0	0	0
Máximo	0	0	0	0	9	10	0	11
Mínimo	4	0	6	10	0	11	0	0
Assimetria	1	0	0	0	0	0	4	11
Curtosi	0	0	0	0	1	4	9	11
Moda	2	5	8	11	0	0	0	0
Atributos mais Indicados	Mínimo	Mediana	Moda	Média	Máximo	Desvio Padrão	Curtosi	Assimetria

* Os números em vermelho indicam os atributos que já foram selecionados anteriormente

Dentre os primeiros atributos indicados pelo ranking a Mediana, a Moda, a Média e o Mínimo foram os que tiveram maior correlação com o *Phantom* com coeficientes médios de correlação de *Spearman* (r) de 0,5, sendo, provavelmente, os mais adequados para serem utilizados na etapa de clusterização.

Atributos Geométricos: Genéricos e Poligonais

A Tabela 5.11 contém: a) os coeficientes de correlação dos atributos geométricos (genéricos/poligonais) em relação ao *Phantom* para as 11 cenas processadas; b) a máxima correlação obtida por atributo dentre as 11 cenas; c) a

média das correlações obtida por atributo, e; d) a frequência com que a correlações não foram significativas em relação ao *Phantom*.

Tabela 5.11 – Coeficientes de Correlação dos atributos Geométricos Genéricos e Poligonais em relação ao *Phantom*.

Coeficientes de Correlação dos Atributos Geométricos (Genéricos/Poligonais) em relação ao <i>Phantom</i> para as 11 Cenas Estudadas													
Classe <i>Phantom</i>	Área	Assimetria	Compr./t de Borda	Compacitação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	µdos Comp.de Borda	Compactação: Polígono	Comp.da maior Borda
<i>Phantom: Cena 1</i>	0,1009*	0,2282*	0,116*	0,08913*	-0,1979*	-0,03811*	0,1914*	-0,04711*	0,1088*	-0,003691	0,1795*	-0,1219*	0,2087*
<i>Phantom: Cena 2</i>	0,1424*	0,483*	0,2921*	0,2858*	-0,4578*	-0,1598*	0,4121*	-0,2125*	0,3667*	-0,1528*	0,2756*	-0,3555*	0,3452*
<i>Phantom: Cena 3</i>	0,384*	0,2658*	0,4166*	0,3376*	-0,2718*	-0,2678*	0,4611*	-0,2789*	0,3272*	0,2656*	0,3669*	-0,3247*	0,3658*
<i>Phantom: Cena 4</i>	0,1128*	0,06577*	0,127*	0,08723*	-0,0498	-0,06752*	0,1328*	-0,06516*	0,1272*	0,09374*	0,1075*	-0,1171*	0,1122*
<i>Phantom: Cena 5</i>	0,08485*	-0,003589	0,08095*	0,05683*	0,01575	-0,03095*	0,07964*	0,03125*	0,06309*	0,08946*	0,07541*	-0,05256*	0,03827*
<i>Phantom: Cena 6</i>	0,317*	0,4962*	0,3354*	0,3401*	-0,4566*	-0,2656*	0,4102*	-0,2562*	0,3064*	0,1958	0,2399	-0,3401*	0,316*
<i>Phantom: Cena 7</i>	-0,04268	0,1202*	-0,01117	0,01804	-0,1403*	-0,01955	0,01563	-0,0227	0,0356*	-0,08723*	-0,02499	-0,03775	0,0111
<i>Phantom: Cena 8</i>	-0,07063	-0,1704*	-0,09499*	-0,09526*	0,1356*	0,05924	-0,1379*	0,08007*	-0,09583*	0,003394	-0,1086*	0,08639*	-0,071
<i>Phantom: Cena 9</i>	0,1641*	0,2952*	0,1791*	0,0573	-0,1298*	0,03583	0,2608*	0,01542	0,1056	0,02072	0,2069*	-0,1221*	0,2068*
<i>Phantom: Cena 10</i>	0,1324*	0,1626*	0,08267*	-0,01478	-0,0274	0,03669*	0,1597*	0,04059*	-0,001485	0,03959*	0,2004*	-0,01016	0,1717*
<i>Phantom: Cena 11</i>	0,1068*	0,2921*	0,07704*	0,02249	-0,2053*	0,01801	0,2047*	0,009502	0,03435	-0,0513	0,1987*	-0,0456	0,2114*
Correlação Máxima	0,3840	0,4962	0,4166	0,3401	0,4578	0,2678	0,4611	0,2789	0,3667	0,2656	0,3569	0,3555	0,3658
Correlação Média	0,1508	0,2348	0,1648	0,1276	0,1898	0,0910	0,2242	0,0963	0,1429	0,0912	0,1795	0,1467	0,1871
Correlações Não Significativas	2	1	1	4	3	3	1	3	4	5	2	3	2

* As Correlações Máxima e Média foram calculadas com os valores de (*r*) em módulo
* As Correlações significativas com o *Phantom* (*p* < 0,05) estão assinaladas em vermelho

Os coeficientes obtidos indicaram que os atributos geométricos não apresentaram bons coeficientes de correlação em relação ao *Phantom*, tendo sido a média das máximas correlações 0,37 e a média global das correlações 0,16. Dentre as 143 correlações obtidas em relação ao *Phantom*, nenhuma apresentou valores de (*r*) acima de 0,7, apenas 1 (1%) apresentou valor 0,5 e a maioria dos casos (142 – 99%) tiveram valores de (*r*) menores que 0,5. Estes valores reforçam a baixa relação dos atributos geométricos (genéricos e poligonais) com o *Phantom*. Apesar destes baixos valores de (*r*), a maioria dos casos 109 (76%) apresentou correlação considerada estatisticamente significativa em relação ao *Phantom*.

Os atributos de Ajuste Elíptico, Ajuste Retangular e Largura tiveram os piores índices de correlação, sendo praticamente nulos. Coerentemente, estes atributos estão entre os que tiveram os maiores números de correlações não significativas em relação ao *Phantom*.

A Tabela 5.12 sintetiza dos resultados obtidos, indicando: a) a freqüência acumulada com que cada um dos atributos foi correlacionado com o *Phantom* do 1º (maiores coeficientes) ao 13º lugar (menores coeficientes), e; b) o ranking final dos atributos mais indicados do 1º ao 13º lugar.

Tabela 5.12 - Análise síntese da Matriz de Correlação para as 11 Cenas: Atributos Geométricos Genéricos e Poligonais

Frequência Acumulada com que Cada um dos Atributos Geométricos (Genéricos/Poligonais) foi Correlacionado com o Phantom do 1º ao 13º lugar													
Atributos	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º	11º	12º	13º
Área	0	1	2	3	5	7	8	0	9	0	10	0	11
Assimetria	6	7	8	0	0	0	0	0	0	0	9	10	11
Compr./to de Borda	0	1	3	0	4	6	10	0	0	0	0	11	0
Compactação	0	0	0	1	0	3	4	5	6	9	11	0	0
Densidade	1	3	6	0	0	0	7	0	0	9	0	10	11
Ajuste Elíptico	0	0	0	0	0	0	0	0	2	4	8	11	0
Comprimento	2	4	6	10	0	0	0	0	0	0	11	0	0
Ajuste Retangular	0	0	0	0	0	0	1	2	4	6	8	9	11
Índice de Forma	0	1	0	2	3	5	6	7	9	10	0	0	11
Largura	1	0	2	0	0	0	0	5	0	0	0	7	11
μ dos Comp. de Borda	1	0	2	3	7	0	9	0	10	0	0	11	0
Compactação (Polig.)	0	0	0	1	3	2	0	6	7	0	0	8	0
Comp. da maior Borda	0	3	0	5	0	7	0	8	9	10	0	0	11
Atributos mais Indicados	Assimetria	Comprimento	Densidade	Comp. da maior Borda	μ dos Comp. de Borda	Área	Compr./to de Borda	Índice de Forma	Compactação (Polig.)	Compactação	Ajuste Retangular	Ajuste Elíptico	Largura

* Os números em vermelho indicam os atributos que já foram selecionados anteriormente

A freqüência acumulada indicou que os atributos geométricos que mais se relacionaram com o *Phantom* foram a Assimetria dos Polígonos, o Comprimento e a Densidade, sendo estes os atributos com maiores valores médios de (r) observados em relação ao *Phantom*.

Atributos Geométricos: *Skeletons*

A Tabela 5.13 contém: a) os coeficientes de correlação dos atributos geométricos (*skeletons*) em relação ao *Phantom* para as 11 cenas processadas; b) a máxima correlação obtida por atributo dentre as 11 cenas; c) a média das correlações obtida por atributo, e; d) a frequência com que a correlações não foram significativas em relação ao *Phantom*.

Tabela 5.13 – Coeficientes de Correlação dos atributos Geométricos *Skeletons* em relação ao *Phantom*.

Coeficientes de Correlação dos Atributos Geométricos (<i>Skeletons</i>) em relação ao <i>Phantom</i> para as 11 Cenas Estudadas							
Classe <i>Phantom</i>	Grau de Ramificação do <i>Skeleton</i>	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Compriment/o/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Comp. das Bordas	Largura dos Ramos
<i>Phantom: Cena 1</i>	0.02956*	0.1312*	0.1327*	0.09948*	0.05811*	0.2279*	0.07899*
<i>Phantom: Cena 2</i>	0.04258	0.3348*	0.3348*	0.3188*	0.08628	0.3762*	-0.03204
<i>Phantom: Cena 3</i>	0.106	0.4015*	0.4015*	0.337*	0.3013*	0.3338*	0.1234
<i>Phantom: Cena 4</i>	0.07272*	0.1195*	0.1196*	0.1019*	0.1034*	0.09485*	0.06598*
<i>Phantom: Cena 5</i>	0.04102*	0.08049*	0.08104*	0.0599*	0.05487*	0.01429	0.06685*
<i>Phantom: Cena 6</i>	0.2247	0.3518*	0.3518*	0.3704*	0.2774*	0.2632*	0.05125
<i>Phantom: Cena 7</i>	-0.06519	0.003264	0.003505	0.02469	-0.07464	0.02576	-0.07542
<i>Phantom: Cena 8</i>	-0.003393	-0.11*	-0.11*	-0.09488*	-0.0056	-0.09107*	-0.02162
<i>Phantom: Cena 9</i>	-0.004729	0.2123*	0.2127*	0.1244*	0.01427	0.2354*	0.1548*
<i>Phantom: Cena 10</i>	-0.01082	0.1078*	0.1082*	0.01872	0.05999*	0.1752*	0.1772*
<i>Phantom: Cena 11</i>	-0.01684	0.1383*	0.1402*	0.06308	0.02108	0.2158*	0.1578*
Correlação Máxima	0.2247	0.4015	0.4015	0.3704	0.3013	0.3762	0.1772
Correlação Média	0.0552	0.1810	0.1814	0.1467	0.0943	0.1867	0.0914
Correlações Não Significativas	8	1	1	3	5	2	5

* As Correlações Máxima e Média foram calculadas com os valores de (r) em módulo

* As Correlações significativas com o *Phantom* ($p < 0,05$) estão assinaladas em vermelho

Os coeficientes máximos e médios obtidos indicaram que os atributos de geométricos (*Skeletons*), não foram bem relacionados com o *Phantom*, sendo todos os (r) menores do que 0,5. Estes valores reforçam a baixa relação dos atributos geométricos (*Skeletons*) com o *Phantom*. Em concordância com estes resultados, além dos baixos valores de (r), 52 casos (72%) apresentaram correlação considerada estatisticamente não significativa em relação ao *Phantom*.

A Tabela 5.14 sintetiza os resultados obtidos, indicando: a) a freqüência acumulada com que cada um dos atributos foi correlacionado com o *Phantom* do 1º (maiores coeficientes) ao 7º lugar (menores coeficientes), e; b) o *ranking* final dos atributos mais indicados do 1º ao 7º lugar.

Tabela 5.14 - Análise síntese da Matriz de Correlação para as 11 Cenas: Atributos Geométricos *Skeletons*.

Freqüência Acumulada com que Cada um dos Atributos Geométricos (<i>Skeletons</i>) foi Correlacionado com o <i>Phantom</i> do 1º ao 7º lugar							
Atributos	1º	2º	3º	4º	5º	6º	7º
Grau de Ramificação do Skeleton	0	0	1	0	0	5	11
Comp. da Linha Principal I	2	6	8	10	0	0	11
Comp. da Linha Principal II	2	6	10	0	0	11	0
Comp./Largura dos Ramos	1	0	3	7	10	11	0
Máximo Comp. dos Ramos	0	1	2	3	7	11	0
Desv. Padrão do Comp. das Bordas	4	6	0	8	10	0	11
Largura dos Ramos	2	3	4	5	7	8	11
Atributos mais Indicados	Desv. Padrão do Comprimento das Bordas	Comp. da Linha Principal II	Comp. da Linha Principal I	Comp./Largura dos Ramos	Máximo Compr./to dos Ramos	Largura dos Ramos	Grau de Ramificação do Skeleton

* Os números em vermelho indicam os atributos que já foram selecionados anteriormente

O Desvio Padrão do Comprimento das Bordas foi o atributo com maiores valores de (r), sendo o mais indicado em 1º lugar, seguido do Comprimento da Linha Principal I e II em 2º e 3º lugar, sendo estes os atributos que tiveram o maior número de casos estatisticamente significativos em relação ao *Phantom*.

Atributos Texturais

A Tabela 5.15 contém: a) os coeficientes de correlação dos atributos texturais em relação ao *Phantom* para as 11 cenas processadas; b) a máxima correlação obtida por atributo dentre as 11 cenas; c) a média das correlações obtida por atributo, e; d) a frequência com que a correlações não foram significativas em relação ao *Phantom*.

Dentre as 88 medidas de correlação obtidas em relação ao *Phantom*, nenhuma apresentou valores de (r) acima de 0,7, apenas 8 (9%) apresentaram valores entre 0,5 e 0,7 e a maioria dos casos (80 – 91%) tiveram valores de (r) menores que 0,5. Estes resultados reforçam a baixa relação dos atributos texturais com o *Phantom*. Apesar

destes baixos valores de (r), a maioria dos casos 72 (81%) apresentou correlação considerada estatisticamente significativa em relação ao *Phantom*.

Tabela 5.15 – Coeficientes de Correlação dos atributos texturais em relação ao *Phantom*.

Classe <i>Phantom</i>	GLCM 2º Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão
<i>Phantom: Cena 1</i>	-0,2721*	-0,1219*	0,1849*	-0,1234*	0,2695*	-0,1877*	0,2628*	-0,02967*
<i>Phantom: Cena 2</i>	-0,3781*	-0,005071	0,2725*	0,02077	0,3939*	-0,2474*	0,5177*	0,2521*
<i>Phantom: Cena 3</i>	-0,2358*	-0,3809*	0,5736*	-0,3374*	0,287*	0,1889*	0,51*	0,01254
<i>Phantom: Cena 4</i>	0,01343	-0,09717*	0,1688*	-0,09441*	-0,006817	0,1272*	0,0667*	0,009623
<i>Phantom: Cena 5</i>	0,1559*	-0,1267*	0,1926*	-0,1046*	-0,1471*	0,2*	0,06359*	-0,068
<i>Phantom: Cena 6</i>	-0,601*	-0,5241*	0,5521*	-0,5078*	0,601*	-0,431*	0,3052*	-0,3168*
<i>Phantom: Cena 7</i>	-0,06447	0,08551	-0,06998	0,08744*	0,05463	-0,1805*	-0,4932*	0,1039*
<i>Phantom: Cena 8</i>	0,0917*	0,07448*	-0,147*	0,07101	-0,09353*	0,01345	-0,199*	-0,03931
<i>Phantom: Cena 9</i>	-0,3193*	-0,1801*	0,2757*	-0,1708*	0,3209*	-0,1882*	0,2119*	-0,02395
<i>Phantom: Cena 10</i>	0,1292*	-0,1348*	0,297*	-0,1137*	-0,1125*	0,2615*	0,1913*	0,1486*
<i>Phantom: Cena 11</i>	0,1222*	-0,01359	0,3378*	0,0308	-0,1214*	0,2243*	0,4238*	0,3214*
Correlação Máxima	0,6010	0,5241	0,5736	0,5078	0,6010	0,4310	0,5177	0,3214
Correlação Média	0,2166	0,1586	0,2793	0,1511	0,2189	0,2046	0,2950	0,1144
Correlações Não Significativas	2	2	1	3	2	1	0	5
* As Correlações Máxima e Média foram calculadas com os valores de (r) em módulo								
* As Correlações significativas com o <i>Phantom</i> ($p < 0,05$) estão assinaladas em vermelho								

A Tabela 5.16 sintetiza os resultados obtidos, indicando: a) a freqüência acumulada com que cada um dos atributos foi correlacionado com o *Phantom* do 1º (maiores coeficientes) ao 8º lugar (menores coeficientes), e; b) o ranking final dos atributos mais indicados do 1º ao 8º lugar.

Tabela 5.16 - Análise síntese da Matriz de Correlação para as 11 Cenas: Atributos Texturais.

Frequência Acumulada com que Cada um dos Atributos Texturais foi Correlacionado com o <i>Phantom</i> do 1º ao 8º lugar								
Atributos	1º	2º	3º	4º	5º	6º	7º	8º
GLCM: 2º Momento Angular	2	3	5	6	7	10	11	0
GLCM: Contraste	0	0	2	3	7	8	9	11
GLCM: Correlação	3	6	8	9	10	11	0	0
GLCM: Dissimilaridade	0	0	0	3	4	7	11	0
GLCM: Entropia	1	4	5	6	7	8	0	11
GLCM: Homogeneidade	1	4	0	6	7	9	10	11
GLCM: Média	4	5	7	8	9	0	10	11
GLCM: Desv. Padrão	0	0	2	3	4	0	7	11
Atributos mais Indicados	GLCM: Média	GLCM: Correlação	GLCM: 2º Momento Angular	GLCM: Entropia	GLCM: Homogeneidade	GLCM: Contraste	GLCM: Dissimilaridade	GLCM: Desv. Padrão

* Os números em vermelho indicam os atributos que já foram selecionados anteriormente

De acordo com o ranking do 1º ao 8º lugar os 3 primeiros atributos mais indicados foram o GLCM da Média, Correlação e 2º Momento Angular. Estes resultados mantêm coerência com os coeficientes de correlação que indicaram como atributos mais relacionados ao *Phantom* o GLCM da Média, da Correlação e da Entropia, sendo estes os atributos com maiores valores de (r) e maior número de casos estatisticamente significativos em relação ao *Phantom* seguidos do GLCM do 2º Momento Angular.

5.4.2.2 – Análise de Componentes Principais

Conforme diagnosticado pelas análises de correlação, no universo dos 36 atributos estudados alguns apresentaram baixa correlação com o *Phantom*. Estes resultados forneceram indícios de que seria necessário reduzir a dimensionalidade dos dados excluindo atributos fracamente correlacionados com o *Phantom*, os quais não explicam o problema e podem prejudicar os resultados das clusterizações. Neste contexto, a análise de componentes principais (ACP) foi utilizada para indicar o conjunto mínimo de atributos, capazes de preservar a maior representatividade contida no conjunto de dados original.

A ACP pode ser efetuada utilizando como base uma matriz de correlação ou de variâncias-covariâncias. Como resultado, o espaço n -dimensional de atributos é combinado linearmente resultando em n -componentes principais, denominados autovetores, cada qual com um autovalor associado. Os autovalores correspondem à variância total dos dados originais distribuída de forma decrescente entre cada um dos autovetores (JOHNSON & WICHERN, 1992). A análise efetuada verificou quantos autovetores eram necessários para atingir um percentual de representatividade de pelo menos 90% em relação aos dados originais. No total foram processadas 11 ACP, as tabelas resultantes com o valor dos autovalores, da variância total e cumulativa para cada autovetor estão disponíveis para consulta no APÊNDICE IV. A Figura 5.16 disponibiliza os gráficos resultantes da ACP para cada recorte de cena, indicando o número de autovetores responsáveis por pelo menos 90% da representatividade dos dados.

O número de componentes, ou autovetores, são disponibilizados em ordem decrescente de acordo com os autovalores obtidos, sendo o número de componentes igual ao número de variáveis. Os gráficos da Figura 5.16 evidenciaram que a partir de certo número de autovetores a variância acumulada estabiliza, aumentando muito pouco em cada componente. O maior percentual da representatividade dos dados ficou concentrado nas primeiras componentes.

No caso em estudo, o número de componentes necessários para atingir uma representatividade de pelo menos 90% variou entre 4 e 10. Em média foram necessários apenas 8 atributos em relação ao conjunto de dados inicial, 22% do total de atributos (36). A Tabela 5.17 contém a síntese das ACP indicando o número de componentes necessários para manter pelo menos 90% da representatividade dos dados em cada cena.

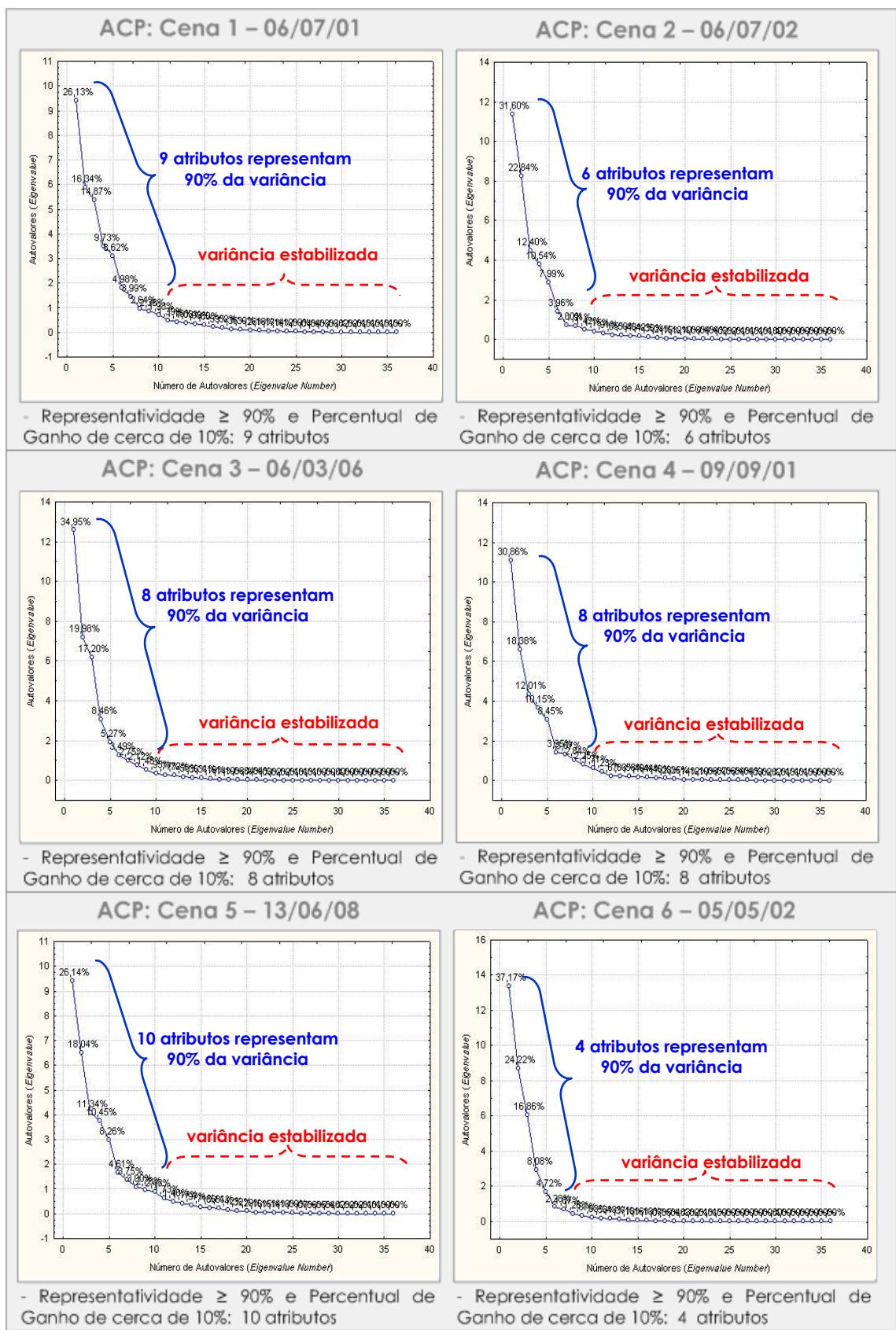


Figura 5.16 - Gráficos resultantes da ACP efetuadas para as 11 cenas analisadas.

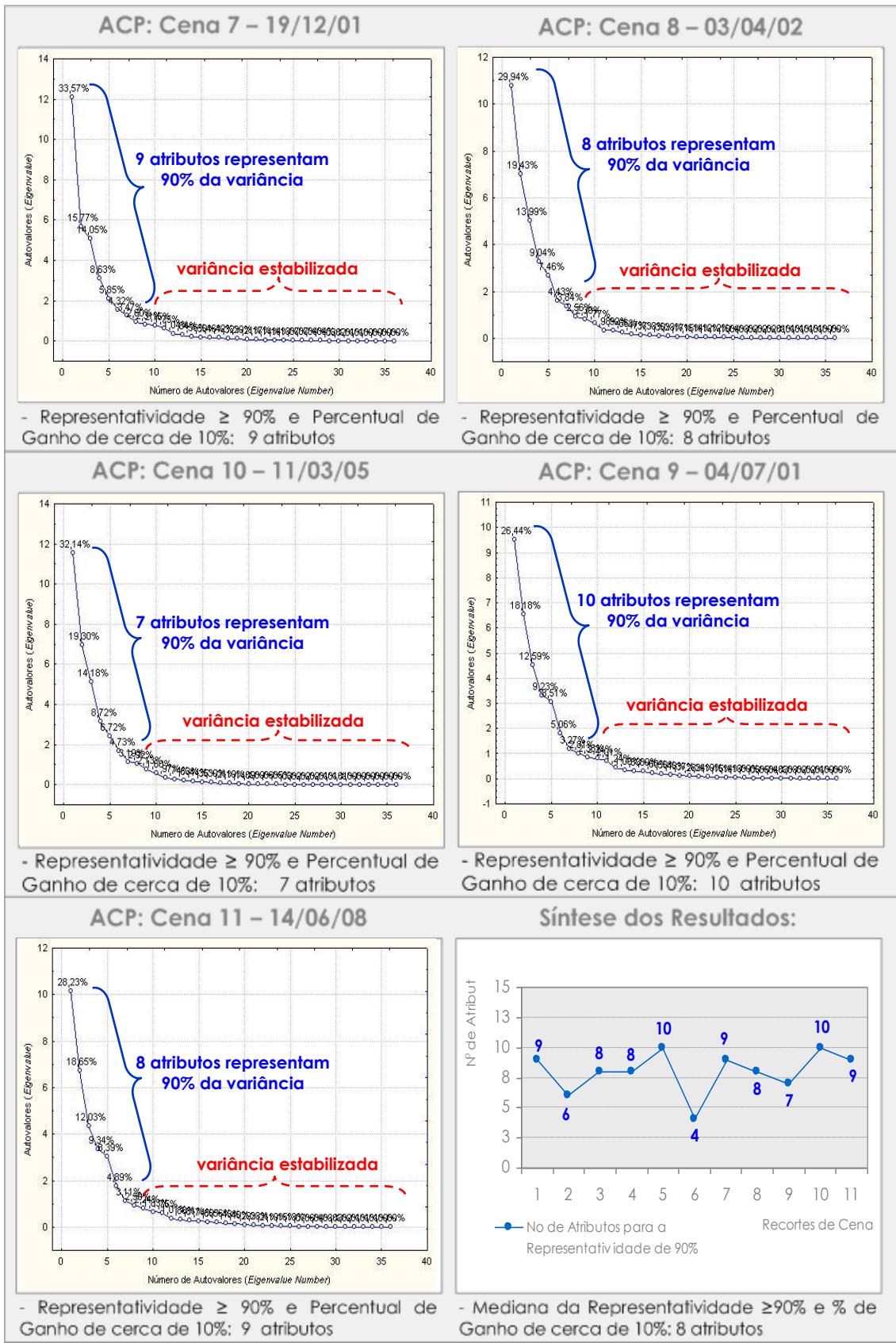


Figura 5.16 – Continuação: Gráficos resultantes da ACP efetuadas para as 11 cenas analisadas.

Tabela 5.17 - Síntese do número de componentes necessários em cada cena para manter pelo menos 90% da representatividade dos dados.

Cenas	Número de Variáveis Necessário para a Representatividade ser $\geq 90\%$	Percentual em Relação ao Total de Atributos: 36
1	9	25%
2	6	17%
3	8	22%
4	8	22%
5	10	28%
6	4	11%
7	9	25%
8	8	22%
9	7	19%
10	10	28%
11	9	25%
Média	8	22%
Mediana	8	22%

Como o número de atributos não é fixo e varia a cada nova cena estudada, as análises subseqüentes são importantes para verificar se uma padronização do número e dos tipos de atributos é viável para automatizar o sistema.

5.4.2.4 – Seleção de Atributos

O processo de seleção de atributos tem como objetivo identificar um subconjunto de atributos que melhor caracterize o conjunto original de dados (LIU & SETIONO, 1995; LEHMANN, 2005; STATHAKIS *et al.*, 2006) removendo variáveis irrelevantes, redundantes ou ruidosas. A redução do número de atributos diminui o custo computacional envolvido nos processamentos, melhora a qualidade do conjunto de dados e, consequentemente, o potencial de predição das variáveis, tornando os resultados mais precisos e fáceis de serem interpretados (DUDA & HARD, 1973; JAIN *et al.*, 1999; NISBET *et al.*, 2009).

Conforme comentado no item 5.4.2 a seleção de atributos é um procedimento essencial para melhorar o desempenho da clusterização. A seleção de um subconjunto de atributos mais representativo para ser utilizado em análises subseqüentes pode melhorar significativamente o reconhecimento de padrões e, consequentemente, a representatividade dos *clusters*. Especialmente nos casos em que se tem um grande volume de dados, avaliar a importância relativa de cada atributo pode torna-se difícil e até mesmo inviável sem a utilização de um método automático de seleção de atributos (NISBET *et al.*, 2009).

Existe uma diversidade de algoritmos para seleção de atributos (LEHMANN, 2005), sendo os métodos mais comumente utilizados os heurísticos e os de busca

exaustiva (BLUM & LANGLEY, 1997; WU & FLACH, 2002). Considerando que o número de combinações (C) possíveis em um conjunto de dados com M atributos equivale a $C=2^M$, uma busca exaustiva pelo melhor subconjunto de atributos pode ser computacionalmente inviável dependendo da dimensão do espaço de atributos. Nestes casos, os métodos de busca heurística são amplamente utilizados, sendo as abordagens mais comuns: a) ganho de informação; b) índice de gini; c) medidas de discrepância, e; c) teste chi-quadrado (WU & FLACH, 2002; LEHMANN, 2005).

No âmbito da tese, o método de seleção de atributos baseado no teste de chi-quadrado foi considerado adequado por tratar amplos conjuntos de dados contendo variáveis independentes numéricas e dependentes categóricas sem assumir que a relação entre os preditores e a variável dependente é linear ou monovariada. O chi-quadrado tem sido indicado na literatura como um método efetivo e confiável para selecionar variáveis numéricas e ordinais, como é o caso do estudo proposto (LIU & SETIONO, 1995). O algoritmo calcula um valor de chi-quadrado para cada atributo independente (36 preditores) em relação a variável dependente (Classe do *Phantom*). Como resultado final o algoritmo fornece uma tabela com o *ranking* dos atributos ordenados de maneira decrescente, considerando a significância estatística (p valor) calculada para cada valor de chi-quadrado (*Statistica*). As tabelas contendo os resultados fornecidos pelo chi-quadrado para as 11 cenas estudadas estão disponíveis para consulta no APÊNDICE V.

Como a ACP indicou que no máximo 10 atributos foram necessários para manter 90% da representatividade do conjunto de dados original, a análise dos resultados obtidos pela seleção de atributos considerou somente os 10 primeiros atributos. A Tabela 5.18 disponibiliza o *ranking* dos 10 atributos mais importantes para cada cena, de acordo com o chi-quadrado.

Tabela 5.18 – *Ranking* dos 10 atributos mais importantes para cada cena

<i>Variable Screening</i>	<i>Cena: 1</i>	<i>Cena: 2</i>	<i>Cena: 3</i>	<i>Cena: 4</i>	<i>Cena: 5</i>
<i>Ranking</i> considerando a ordem de importância fornecida pelo Chi-Quadrado para os 10 atributos com maior importância por Cena	Média Mediana Moda Mínimo Máximo GLCM Média GLCM 2º Momento Angular Desvio Padrão Assimetria dos Polígonos GLCM Homogeneidade	Densidade Máximo Mínimo GLCM Média Assimetria dos Polígonos Comprimento/Largura GLCM 2º Momento Angular GLCM Entropia Ajuste Retangular Índice de Forma	Mediana Média Mínimo Moda Máximo GLCM Média Ajuste Retangular GLCM Correlação Densidade Índice de Forma	Moda Mediana Média Mínimo Máximo GLCM Média Comprimento Ajuste Retangular Densidade Média do Comprimento das Bordas	Mediana Moda Média Mínimo Máximo GLCM Média Ajuste Retangular GLCM Homogeneidade Compactação Densidade
<i>Cena: 6</i>	<i>Cena: 7</i>	<i>Cena: 8</i>	<i>Cena: 9</i>	<i>Cena: 10</i>	<i>Cena: 11</i>
Mínimo Média Mediana Desvio Padrão Moda GLCM Dissimilaridade GLCM Correlação GLCM Média GLCM Contraste Densidade	GLCM Média Mínimo Média Mediana Moda Máximo Assimetria Compactação GLCM Homogeneidade Índice de Forma	Mediana Moda Média Mínimo Máximo GLCM Média Ajuste Retangular GLCM Homogeneidade Ajuste Elíptico Densidade	Mediana Mínimo Moda Média Máximo Densidade GLCM Média Ajuste Retangular Assimetria dos Polígonos Desvio Padrão	Moda Mediana Média Mínimo Máximo GLCM Correlação GLCM Média GLCM Homogeneidade Desvio Padrão Média do Comprimento das Bordas	Mediana Média Moda Mínimo Máximo GLCM Média GLCM Correlação Assimetria dos Polígonos Desvio Padrão Assimetria

Embora haja uma recorrência de certos atributos nos 11 recortes de cena estudados, o conjunto dos 10 melhores atributos indicados pelo método de seleção variou de cena para cena. Este resultado é esperado uma vez que o conteúdo das imagens, ou seja, os tipos de alvos, sua dimensão e geometria são diferentes em cada cena. Os parâmetros relacionados ao modo de imageamento do sensor como ângulo de incidência, largura da faixa e resolução espacial utilizados no momento da aquisição também originam estas diferenças entre as cenas (KARATHANSSI *et al.*, 2006; BAATZ *et al.*, 2003; GASULL *et al.*, 2002). Mesmo imagens adquiridas por um mesmo sensor em um mesmo modo de imageamento apresentarão diferenças, sendo estas influenciadas predominantemente pelas condições meteo-oceanográficas estabelecidas na superfície do mar no momento da aquisição dos dados (Capítulo 2, item 2.2).

O desafio proposto foi encontrar um conjunto de atributos padrão, capaz de retornar bons resultados na clusterização, sendo utilizado a cada nova cena processada, fornecendo como resultado *clusters* representativos para os alvos escuros contidos na cena. Para que esta análise fosse factível, o conjunto de cenas teste selecionadas para a tese buscaram abranger esta diversidade de tipos de alvos e complexidade de cenários (Capítulo 4, item 4.1).

Esta heterogeneidade encontrada entre os 10 atributos mais importantes (chi-quadrado) para cada recorte de cena é consequência desta seleção. Se as cenas tivessem sido recortadas sempre individualizando um único alvo com um bom contraste em relação ao oceano (priorizando eventos relacionados a óleo) - como frequentemente observado na literatura disponível - seria possível ter obtido uma maior homogeneidade entre os atributos selecionados. Esta maior recorrência dos atributos selecionados facilitaria a padronização deste parâmetro (tipos e número de atributos), entretanto, não refletiria a realidade das imagens SAR adquiridas nas áreas oceânicas, provavelmente resultando em uma grande oscilação no desempenho da clusterização.

Para indicar esse conjunto padrão de atributos, ao invés de se trabalhar individualmente com a performance obtida utilizando os 10 melhores atributos indicados para cada cena, a freqüência acumulada de indicação dos atributos foi avaliada, originando um *ranking* com os 10 mais indicados (Tabela 5.19).

Tabela 5.19 – *Ranking* dos 10 atributos mais importantes indicados pelo chi-quadrado do 1º ao 10º lugar.

Atributos	Ordem de Indicação pelo Chi-Quadrado (Considerando os 10 Primeiros)									
	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º
Média	1	4	9	10						11
Mínimo	1	3	5	11						
GLCM Média	1		2			8	10	11		
Mediana	5	8	10	11						
Moda	2	5	8	9	11					
Máximo					9	11				
Densidade	1	1				2			4	7
Assimetria dos Polígonos					1		2	3	5	6
Ajuste Retangular							3	2	1	
Desvio Padrão				1				2	4	5
GLCM Homogeneidade						1	4		4	5
GLCM Correlação						2	3			
GLCM 2º Momento Angular						2		1	2	
Compactação										2
Índice de Forma										2
Média do Comprimento dos Trechos de Borda										
Comprimento/Largura (Linha Principal)						1				
GLCM Dissimilaridade						1				
Comprimento							1			
GLCM Entropia								1		
Ajuste Elíptico									1	
GLCM Contraste									1	
Os 10 Primeiros atributos mais indicados pelo Chi-Quadrado	Mediana	Média	Moda	Mínimo	Máximo	GLCM Média	Ajuste Retangular	GLCM Homogeneidade	Assimetria (Forma)	Densidade

* Os números em vermelho indicam atributos que já foram selecionados anteriormente

O método de seleção utilizado indicou os atributos estatísticos como os mais relacionados com o *Phantom*, ocupando os 5 primeiros lugares no *ranking*, respectivamente: Mediana, Média, Moda, Mínimo e Máximo. Embora a ordem de importância dos atributos seja diferente da indicada pelos coeficientes de correlação de *Spearman*, os atributos são praticamente os mesmos. Em concordância com os resultados obtidos anteriormente, estes atributos foram os que tiveram os maiores coeficientes ($(r) \approx 0,5$) e a menor freqüência de coeficientes estatisticamente não significativos ($p < 0,05$) em relação ao *Phantom*.

Conforme esperado, os atributos indicados nas análises anteriores como pouco relacionados com o *Phantom*, os geométricos e os texturais foram os menos indicados entre os 10 primeiros do *ranking*. Apesar dos baixos coeficientes, a Assimetria dos Polígonos, a Densidade e o GLCM da Média estão entre os atributos geométricos e textuais mais correlacionados significativamente com o *Phantom*.

A Figura 5.17 contém a espacialização dos 10 atributos normalizados indicados pelo *ranking*, sendo visualmente perceptível que os atributos estatísticos apresentam maior potencial para auxiliar na detecção dos alvos escuros. Os atributos estatísticos ressaltaram com maior contraste, maior estabilidade e menor ruído as geometrias referentes aos alvos escuros em todas as cenas, sendo os atributos de Mínimo e de Máximo os menos estáveis.

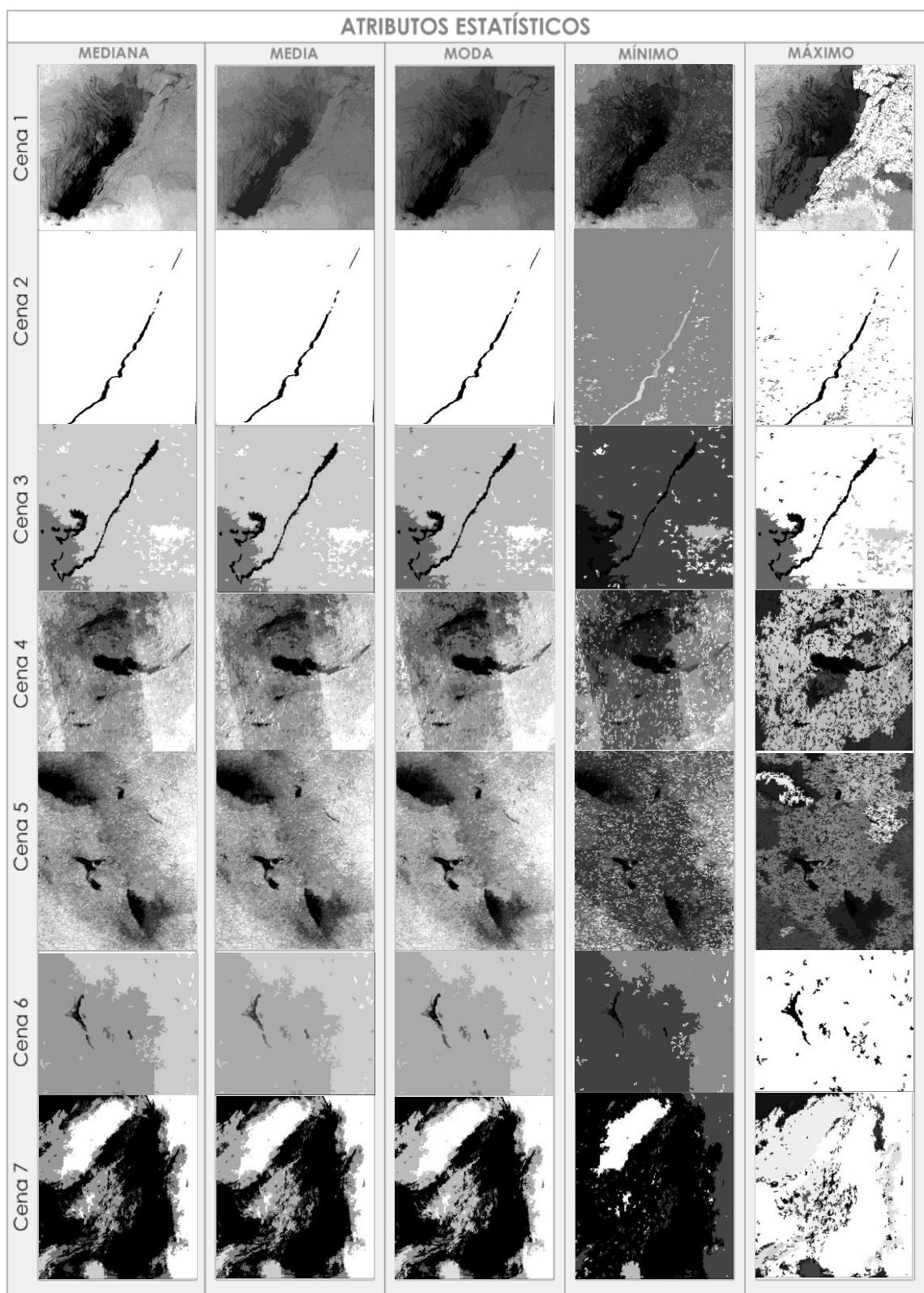


Figura 5.17 - Espacialização dos 10 atributos mais indicados pelo *ranking* para as 11 cenas estudadas.

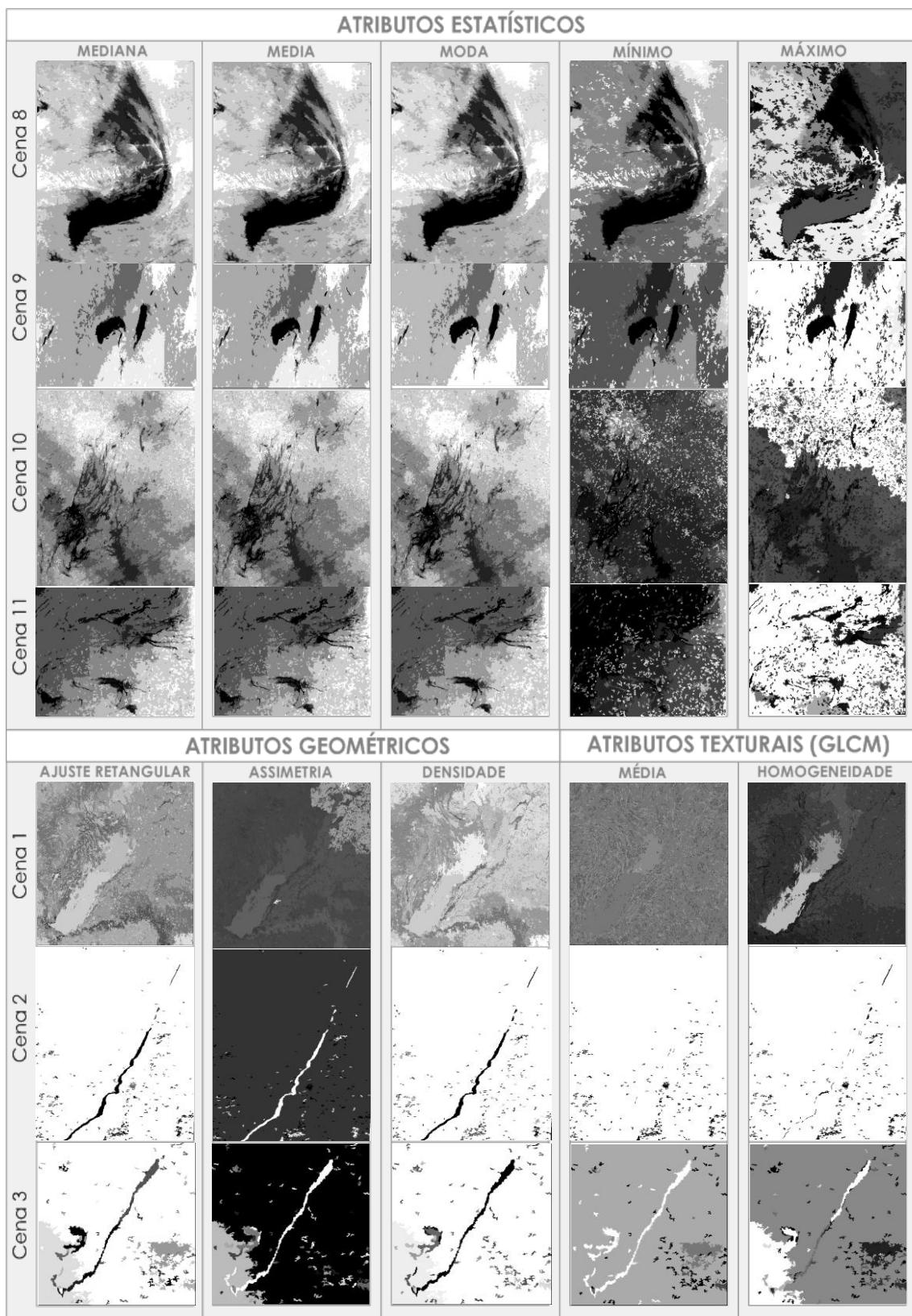


Figura 5.17 – Continuação: Espacialização dos 10 atributos mais indicados pelo *ranking* para as 11 cenas estudadas.

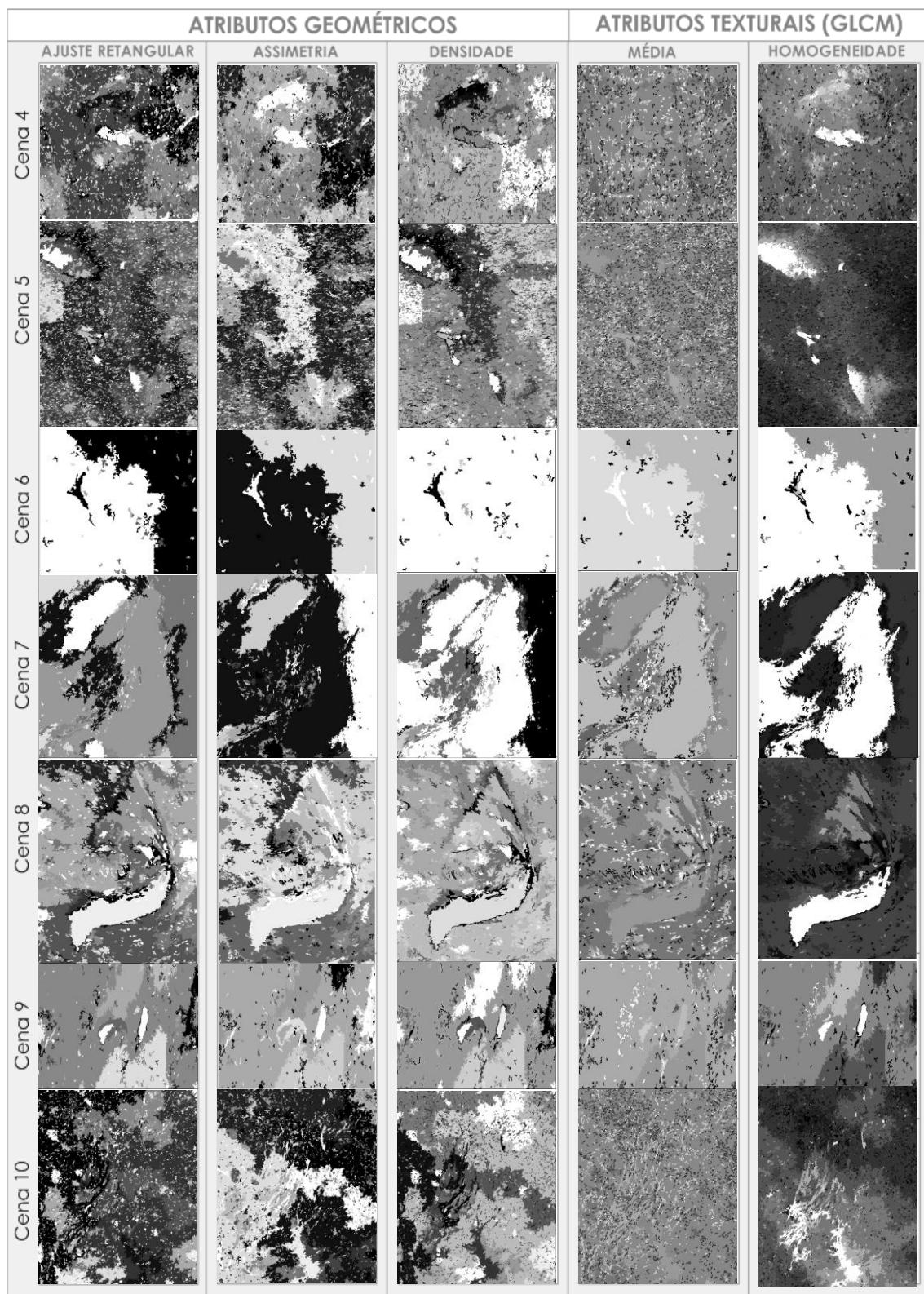


Figura 5.17 – Continuação: Espacialização dos 10 atributos mais indicados pelo *ranking* para as 11 cenas estudadas.

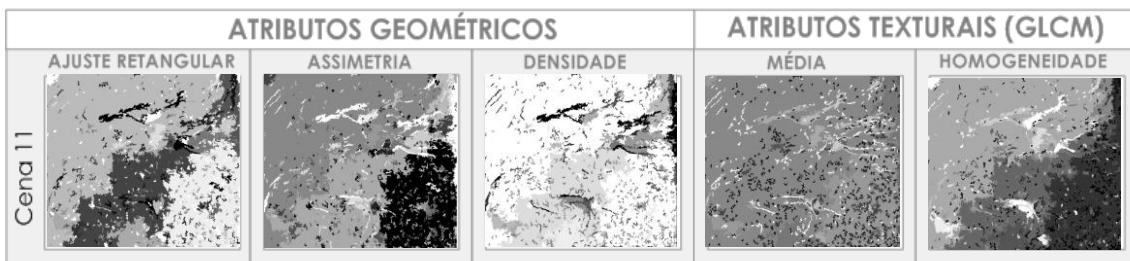


Figura 5.17 – Continuação: Espacialização dos 10 atributos mais indicados pelo *ranking* para as 11 cenas estudadas.

Embora os atributos geométricos tenham apresentado um comportamento instável e ruidoso para a maioria das cenas, de forma contrária, nas cenas 2, 3 e 6, em que manchas de óleo com geometrias bem definidas e alongadas estavam presentes, estes atributos contribuíram significativamente para aumentar o contraste com o oceano. Visualmente, o atributo de Densidade parece ter contribuído também para ressaltar assemelhados produzidos pela presença de óleos biológicos nas cenas 1 e 11, e de baixa intensidade de vento na cena 7. Com exceção da cena 2, o atributo textural que aparentou ter o maior potencial para auxiliar no reconhecimento de padrões foi o GLCM da Homogeneidade. Mesmo abrangendo um número menor de cenas, o GLCM da Média visualmente contribuiu para destacar feições em relação ao oceano nos recortes de cenas 3, 6, 7, 10 e 11.

Os trabalhos que utilizaram seleção de atributos na etapa de classificação confirmaram a importância de alguns dos atributos indicados entre os 10 melhores pelo método de seleção utilizado. Os atributos geométricos de Assimetria dos Polígonos, Ajuste Retangular e Densidade e o atributo textural GLCM da Média, também foram selecionados por TOPOUZELIS *et al.* (2009) e STATHAKIS *et al.* (2006). KARATHANASSI *et al.* (2006) também indicou o GLCM da Média e BENTZ (2006) a Assimetria dos Polígonos. O GLCM da Homogeneidade de HARALICK *et al.* (1979) também foi um atributo indicado na literatura (AKKARTAL & SUNAR; 2008, MARGHANY, 2001; ASSILZADEH & MANSOR, 2001) com potencial para auxiliar a detecção automática de alvos escuros em imagens SAR, indicando as áreas com máxima homogeneidade na cena. A indicação dos atributos estatísticos de primeira ordem não foi explicitamente verificada na literatura porque os autores utilizam métricas calculadas a partir da combinação destes atributos.

Portanto, para a detecção de alvos escuros, os atributos selecionados como mais importantes pelo chi-quadrado foram, respectivamente: Mediana, Média, Moda, Mínimo, Máximo, GLCM da Média, Ajuste Retangular, GLCM da Homogeneidade, Assimetria dos Polígonos e Densidade. Importante destacar que os atributos

selecionados pelo chi-quadrado como mais importantes são multicorrelacionados entre si, as matrizes de correlação estão disponíveis no APÊNDICE III. Entretanto, é notável que estes atributos foram os que apresentaram as melhores correlações com o *Phantom*, sendo os demais atributos muito pouco relacionados. Desta forma, a seleção efetuada não é considerada definitiva, pois a contribuição efetiva de cada um destes atributos para a detecção de alvos escuros também será avaliada através dos resultados das clusterizações realizadas nas próximas etapas da tese.

5.4.3 – Reconhecimento de Padrões por Clusterização Automática

Nesta etapa da tese o objetivo foi avaliar a performance de técnicas automáticas de clusterização aplicadas à detecção de alvos escuros em imagens SAR. Conforme definido (Capítulo 5.4), o processo de clusterização é capaz de explorar automaticamente a inter-relação existente entre o conjunto de atributos, reconhecendo padrões e agrupando-os em *clusters*, de acordo com um critério de similaridade previamente estabelecido. O reconhecimento de padrões é considerado essencial no processo de clusterização e está intrinsecamente ligado à qualidade dos atributos selecionados (JAIL *et al.*, 1999).

O padrão (X) é representado por um vetor multidimensional ($X = X_1, \dots, X_d$) onde cada dimensão corresponde a um atributo, sendo a dimensionalidade do padrão igual ao número total de atributos considerados (DUDA & HART, 1973). Espera-se que as propriedades estatísticas do conjunto de atributos extraídos das imagens SAR e da geometria dos polígonos apresentem comportamento diferenciado nas regiões escuras e claras das imagens. Este comportamento, quando identificado através de padrões discriminantes entre si, permitiria ao algoritmo detectar *clusters* representativos separando, sem ou com reduzida confusão, os alvos escuros do oceano.

Padronizar os parâmetros envolvidos nas etapas de processamento é necessário para implementar um sistema automático, sendo a definição dos tipos e da quantidade de atributos fundamental para a etapa de clusterização. Este quesito é raramente abordado na literatura especializada em detecção automática de óleo. Cada autor utiliza um número diferente de atributos (TOPOUZELIS *et al.*, 2002; 2008, e; 2009; STATHAKIS *et al.*, 2006; KARATHANASSI *et al.*, 2006), sem apresentar estudos evidenciando que a escolha representa a melhor performance para os classificadores (APÊNDICE I).

As clusterizações foram efetuadas no software *Clustan* módulo *Focal Point (FP)* inserindo os 10 melhores atributos um a um, na ordem de importância indicada (Tabela 5.19), com o objetivo de verificar a parcela de contribuição de cada atributo nos resultados da clusterização. A validação do melhor conjunto de atributos foi efetuada comparando os alvos escuros resultantes das inúmeras clusterizações efetuadas, com os alvos escuros definidos como referência nos *Phantons* elaborados para cada cena. A metodologia utilizada para a validação do conjunto adequado de atributos está detalhada no fluxograma da Figura 5.18.

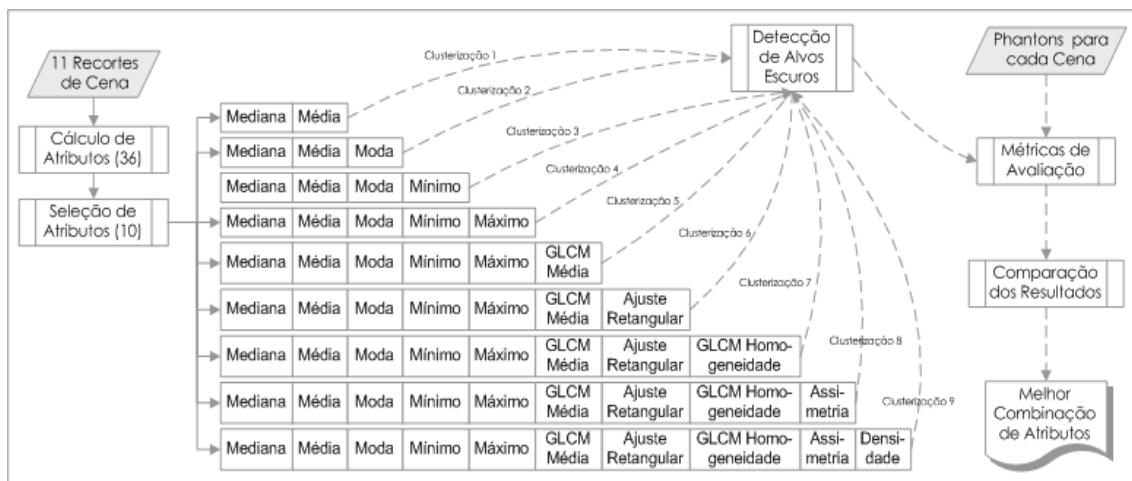


Figura 5.18 - Metodologia utilizada para a validação do conjunto adequado de atributos.

A avaliação da qualidade dos alvos escuros produzidos pelas clusterizações em relação aos *Phantons* não considerou todas as métricas de avaliação utilizadas nas etapas anteriores da tese. Índices como a Fração de Intersecção de Alvos Escuros (*FInt*), a Intersecção de Polígonos (*IPol*) e a taxa de Omissão (*Om*) tornaram-se insensíveis para caracterizar as diferenças entre os resultados obtidos. Isto ocorreu porque, dentre as inúmeras clusterizações efetuadas considerando diferentes combinações de atributos, existem resultados bons e ruins para serem avaliados. Para os resultados em que ocorreu a inclusão exagerada de regiões que seriam oceano como alvos escuros, uma sobreposição com o *Phantom* indicaria que todos os alvos escuros foram detectados, sem considerar a imprecisão da forma dos *clusters*. Neste contexto, as métricas de avaliação *FInt* e *IPol* não poderiam ser consideradas para avaliar todos os resultados obtidos, sendo indicado não utilizá-las a fim de evitar interpretações errôneas. A mesma recomendação aplica-se para o índice de omissão (*Om*), sendo praticamente nulo nos casos em que os erros de inclusão (*In*) foram exagerados, tornando-se inexpressivo para a avaliação.

Como o cálculo dos atributos de ajuste Gf , $Fiti$, $Fitn$ e $Fitxy$ exige um alto custo computacional e o volume de resultados produzidos foi muito grande (cerca de 550 imagens agrupadas), tornou-se inviável utilizá-los nesta etapa da tese.

O atributo indicado como mais coerente para ser utilizado na avaliação dos resultados foi a Exatidão dos Alvos Escuros ($ExAE$), métrica que considera a intersecção dos alvos escuros dividida pela união da intersecção, com os erros de omissão e de inclusão (Capítulo 3: Item 3.3). Como esta métrica considera os erros de Om e de In adquire sensibilidade para indicar situações em que o clusterizador falhou incluindo ou omitindo regiões erroneamente nos alvos escuros, sendo o agrupamento perfeito obtido quando $ExAE$ for igual a 1.

Os resultados iniciais evidenciaram que, mesmo utilizando um método para a seleção de atributos, o processamento simultâneo dos atributos selecionados pode causar oscilações na performance do clusterizador, o que nem sempre retorna resultados aceitáveis. Na maioria das cenas estudadas a utilização do atributo de Mínimo isolado ou em conjunto com o Máximo piorou os resultados da clusterização e, consequentemente, a detecção de alvos escuros. A capacidade de discriminação associada a um determinado atributo muda se o mesmo estiver sendo analisado isoladamente ou em conjunto (TOPOUZELIS *et al.*, 2009). A Figura 5.19 disponibiliza os *clusters* resultantes da utilização simultânea e individual dos atributos de máximo e mínimo nos resultados da clusterização, juntamente com a $ExAE$ obtida em cada caso.

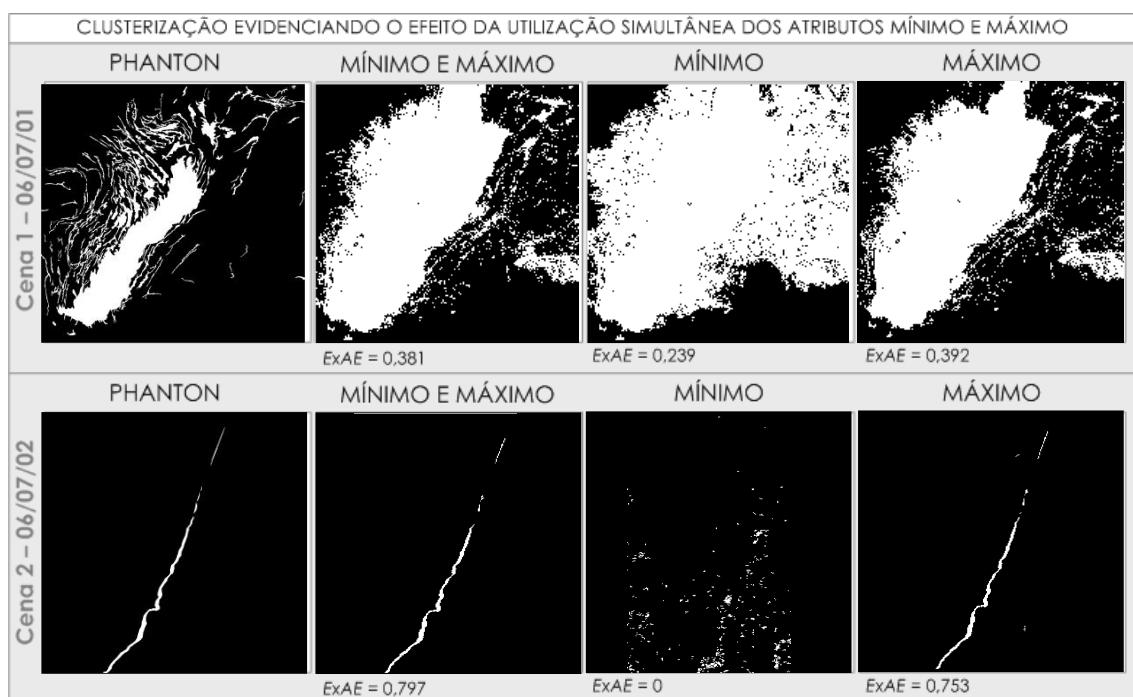


Figura 5.19 – Influência da utilização simultânea e individual dos atributos de Máximo e Mínimo na clusterização.

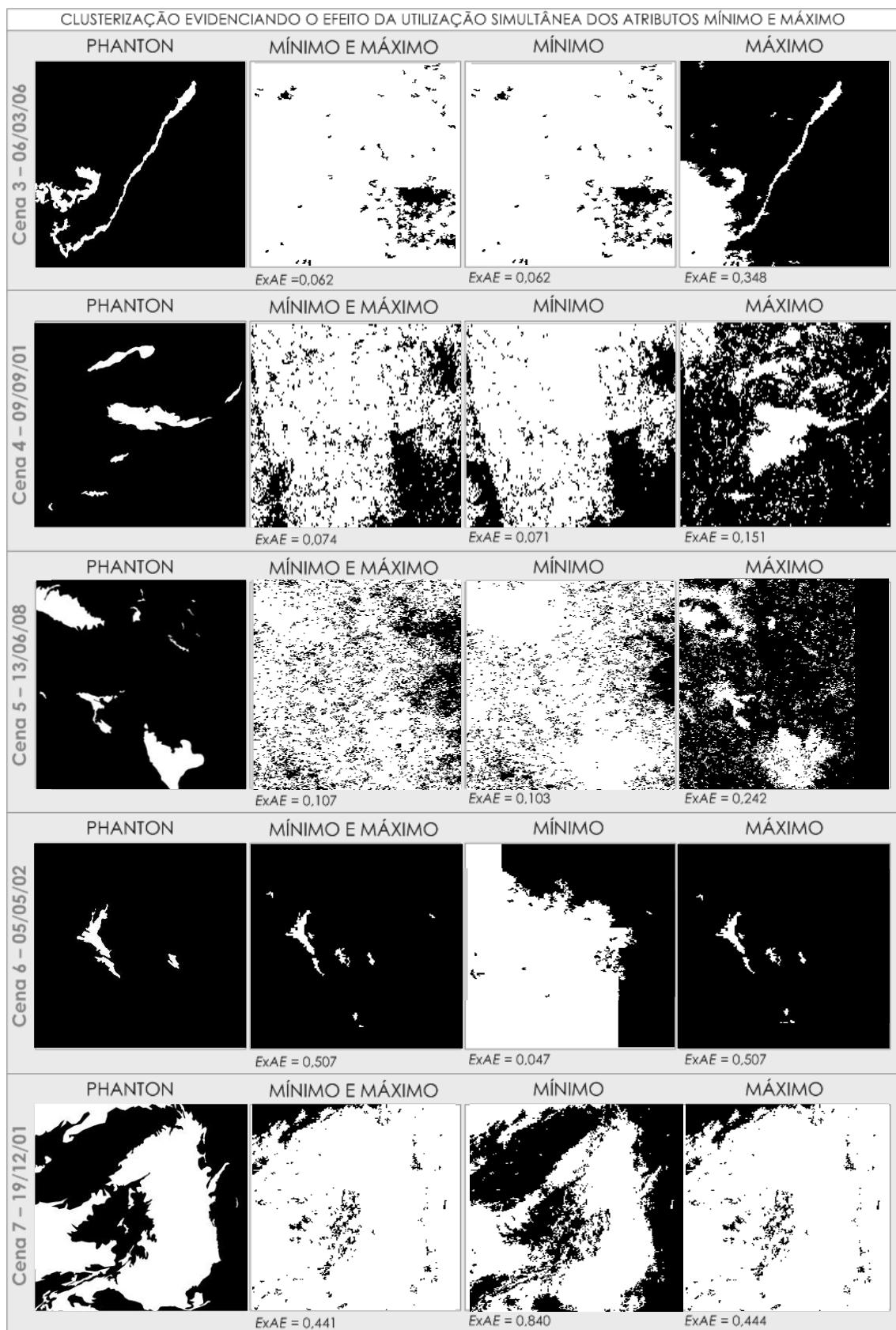


Figura 5.19 – Influência da utilização simultânea e individual dos atributos de máximo e mínimo na clusterização.

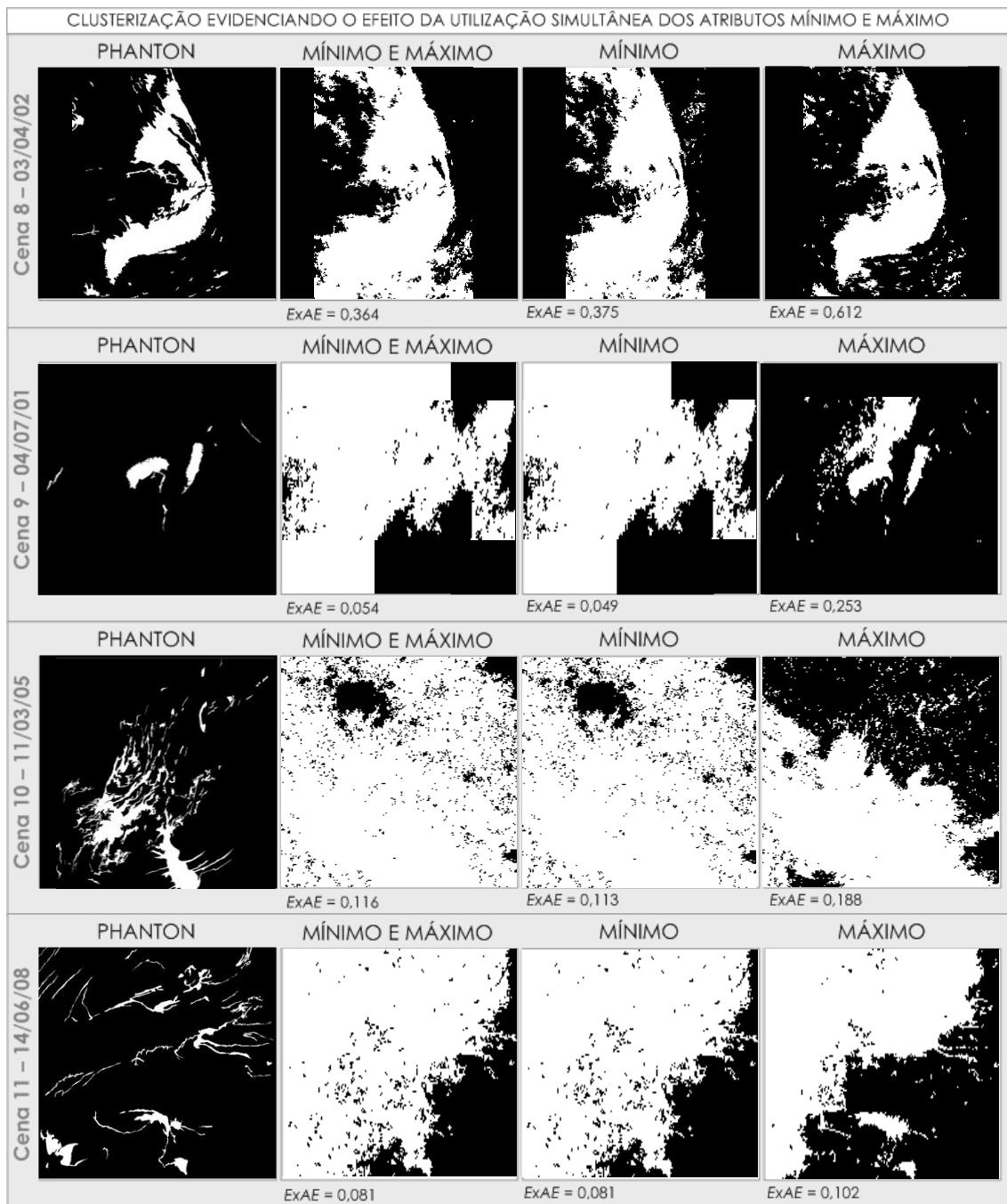


Figura 5.19 – Continuação: Influência da utilização simultânea e individual dos atributos de máximo e mínimo na clusterização.

A espacialização dos *clusters* detectados, juntamente com a *ExAE* obtida para cada recorte de cena em relação ao *Phantom*, confirmou que o atributo de Mínimo produz um efeito negativo na grande maioria dos casos, prejudicando a performance da clusterização. Para entender os motivos pelos quais este efeito ocorre a Figura 5.20 disponibiliza, para 3 exemplos: a) as imagens de RADAR filtradas com o Lee 3x3 sem nenhum contraste aplicado; b) a espacialização dos atributos de Mínimo e Máximo, e; c) os *box plots* gerados com a distribuição dos atributos de Máximo e

Mínimo em relação às 4 classes do atributo categórico Classe do *Phantom* definidas na Figura 5.14.

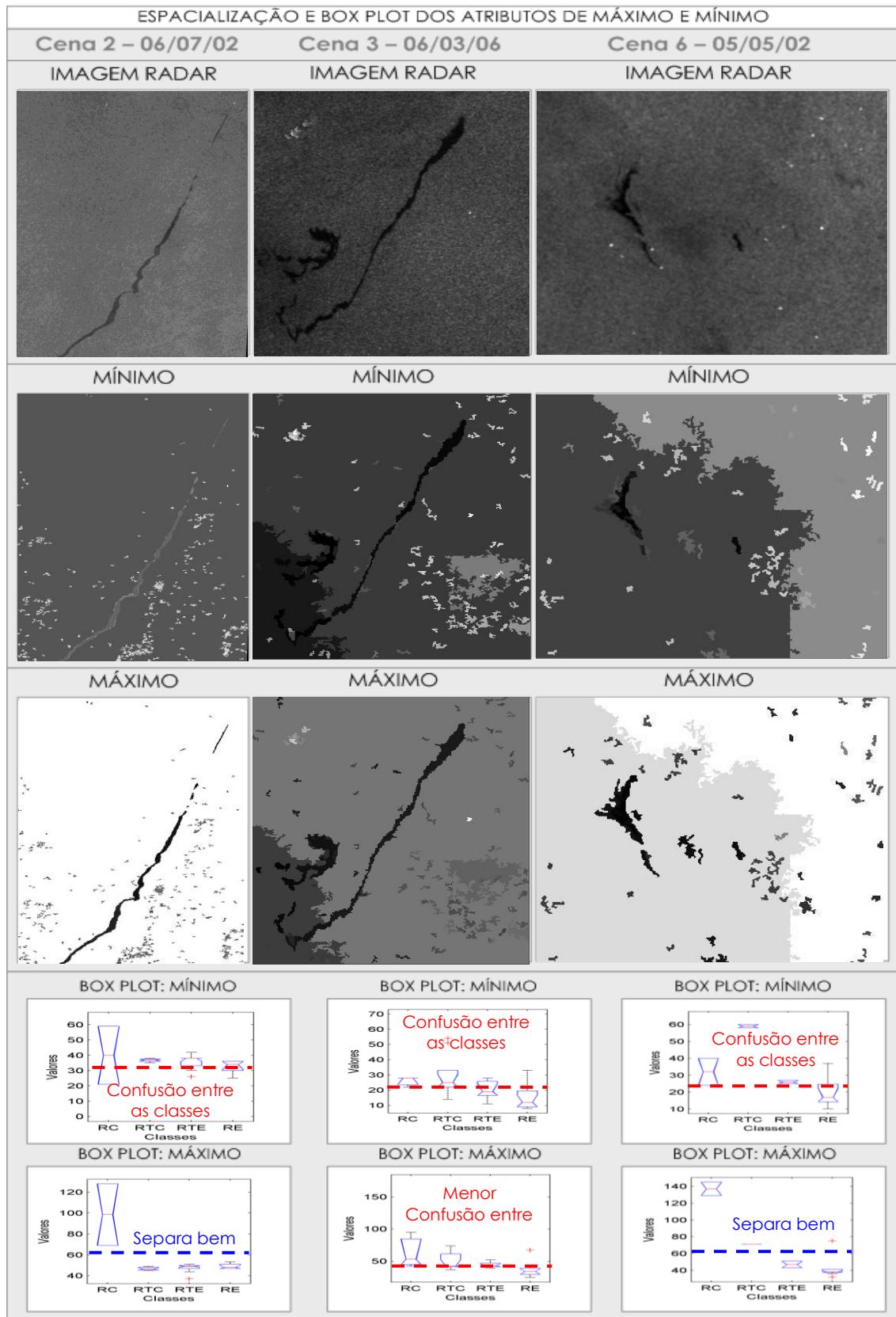
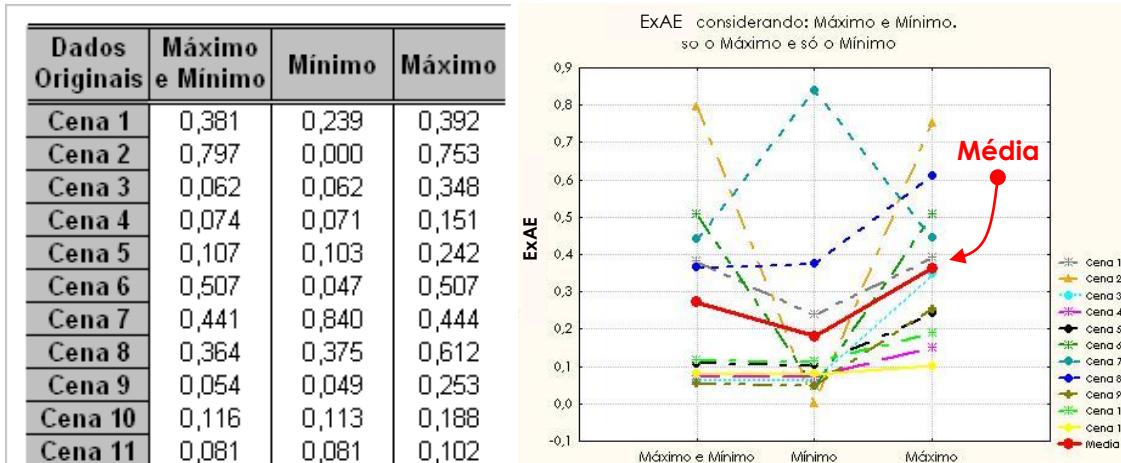


Figura 5.20 – Propriedades dos atributos de Máximo e Mínimo que influenciam na clusterização.

Nos *box plots* foi possível verificar que a distância entre as Regiões Claras (RC) e Escuras (RE), e de Transição entre o Claro (RTC) e o Escuro (RTE) foi maior para o atributo de Máximo. Quanto maior for a diferença existente entre as Regiões (RC, RTC, RTE e RE) maior será o contraste dos alvos escuros em relação ao oceano e, consequentemente, o potencial do clusterizador para identificar padrões que possibilitem um melhor particionamento do conjunto de dados. Este melhor contraste entre os alvos escuros e o oceano fornecido pelo Máximo é visível na espacialização dos atributos na Figura 5.22. Desta forma, o atributo de Mínimo não constitui uma boa métrica, sendo indicada a sua exclusão do conjunto de atributos selecionados pelo chi-quadrado. Esta recomendação é reforçada com os dados disponíveis na Tabela 5.20, que disponibiliza a *ExAE* obtida pelas clusterizações, considerando os atributos de Máximo e Mínimo simultaneamente e individualmente.

Tabela 5.20 – Exatidão dos Alvos Escuros (*ExAE*) obtida para as clusterizações efetuadas, considerando os atributos de máximo e mínimo simultaneamente e individualmente.



O gráfico confirmou as análises indicando que na média a clusterização considerando somente o atributo de Máximo teve um melhor desempenho caracterizado pelas maiores *ExAE*. Para a grande maioria das cenas clusterizadas, quando a *ExAE* não foi maior considerando somente o atributo de Máximo, foi igual à *ExAE* obtida com a combinação do Máximo e do Mínimo.

Com base nos resultados apresentados, a série de Figuras 5.21 a 5.31 disponibiliza os resultados das clusterizações realizadas para as 11 cenas, inserindo um a um os 9 melhores atributos indicados pela análise exploratória e pelo método de seleção de atributos, desconsiderando o atributo de Mínimo. Os gráficos elaborados disponibilizam conjuntamente a variação da *ExAE* e da Média da Soma dos Quadrados da Distância Euclidiana (*ME*), obtidos para as clusterizações realizadas a cada novo atributo considerado. A *ME* foi analisada por ser o critério utilizado pelo *FP*

para convergir ao final das interações em uma solução considerada ótima para a clusterização (WISHART, 2004).

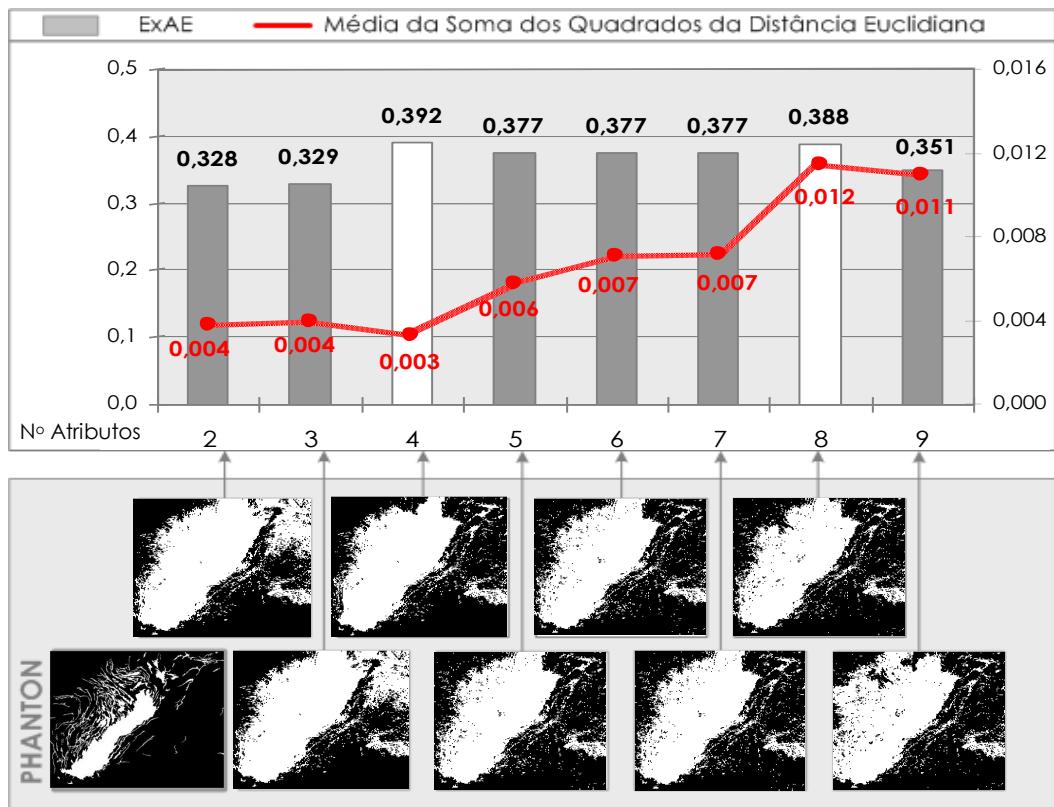


Figura 5.21 – Gráfico e espacialização das clusterizações realizadas para a Cena 1.

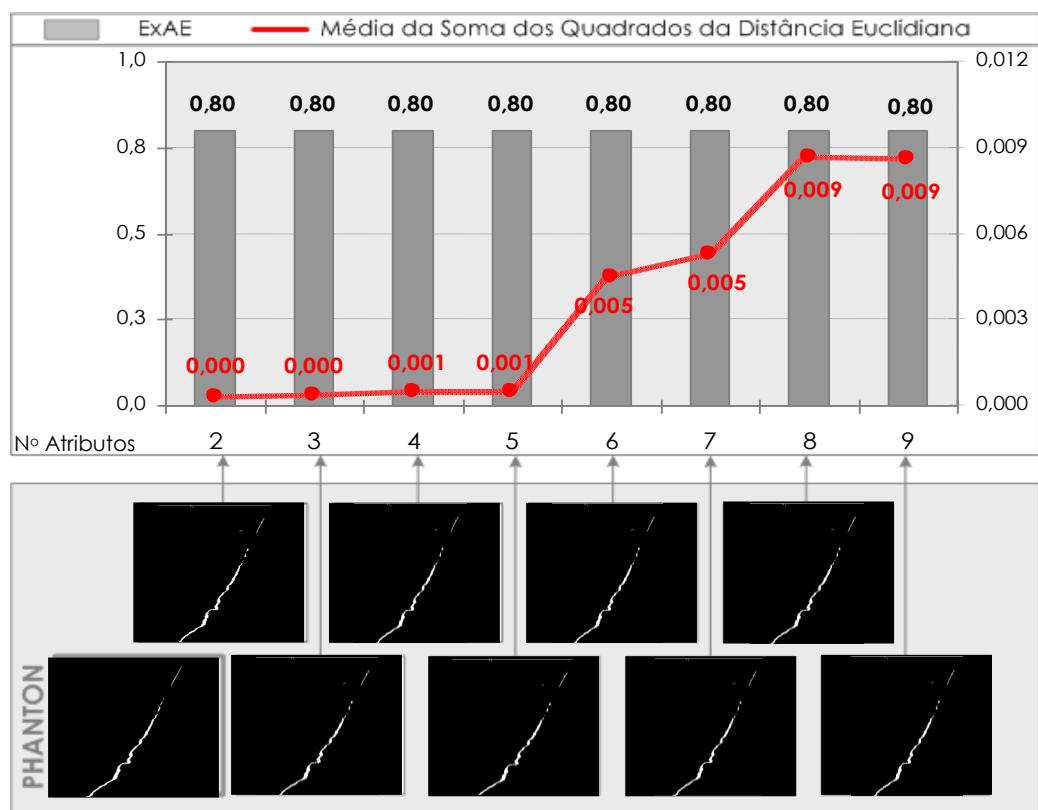


Figura 5.22 – Gráfico e espacialização das clusterizações realizadas para a Cena 2.

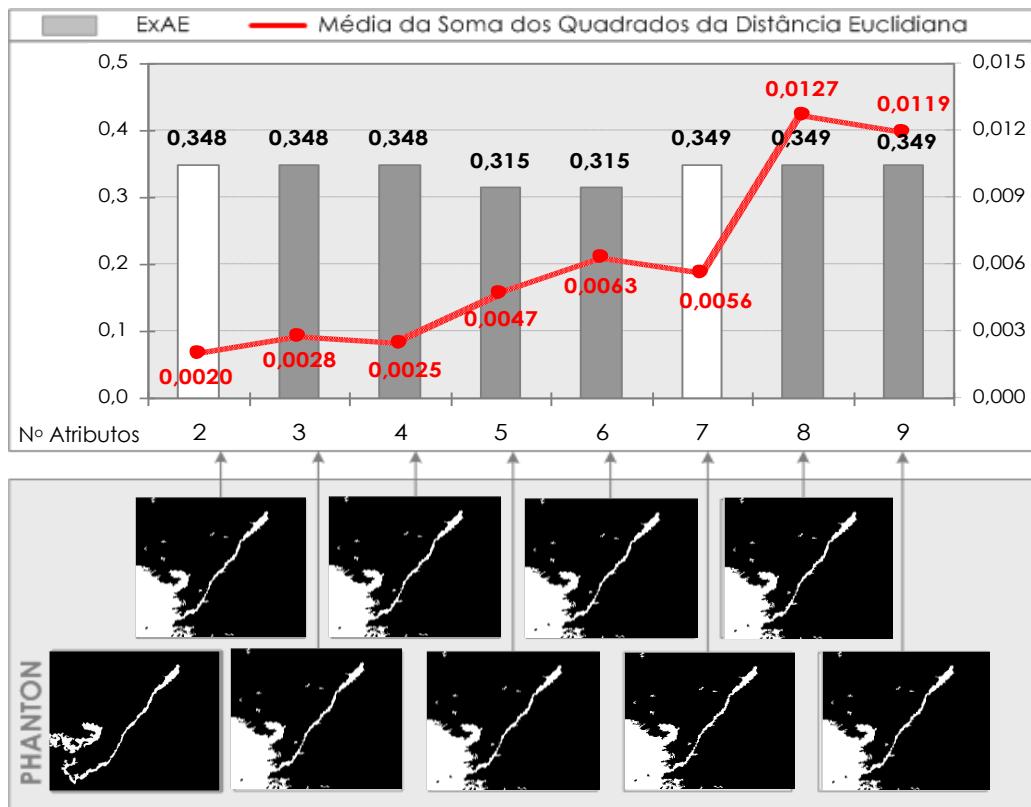


Figura 5.23 – Gráfico e espacialização das clusterizações realizadas para a Cena 3.

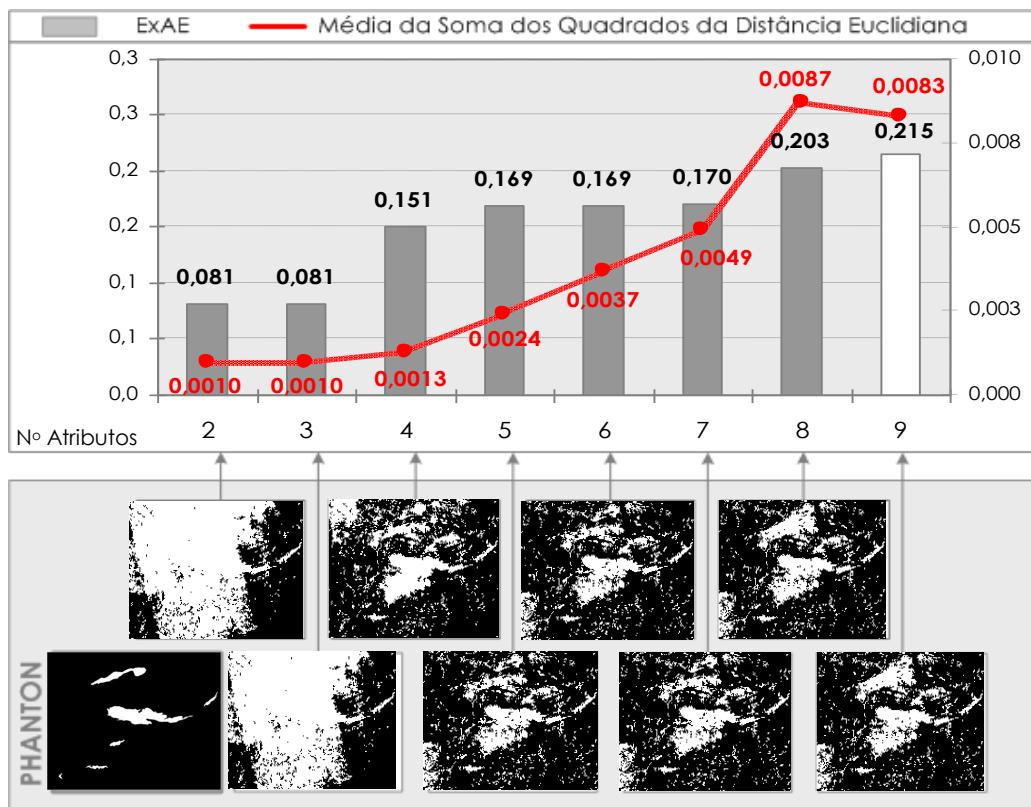


Figura 5.24 – Gráfico e espacialização das clusterizações realizadas para a Cena 4.

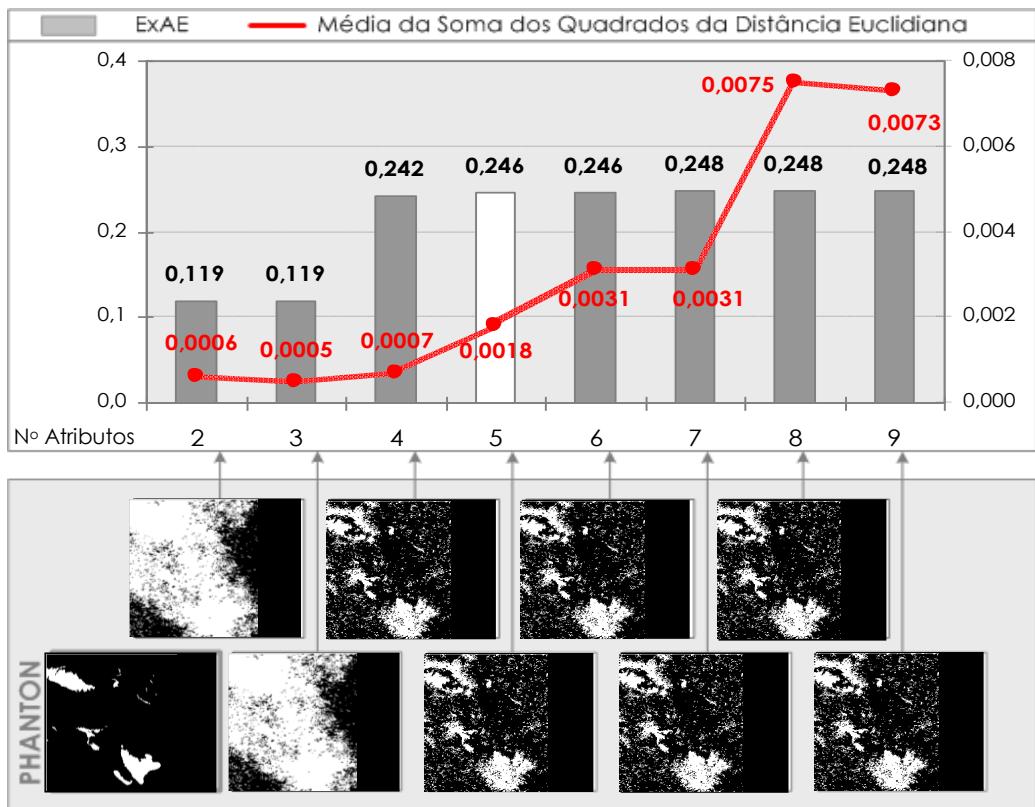


Figura 5.25 – Gráfico e espacialização das clusterizações realizadas para a Cena 5.

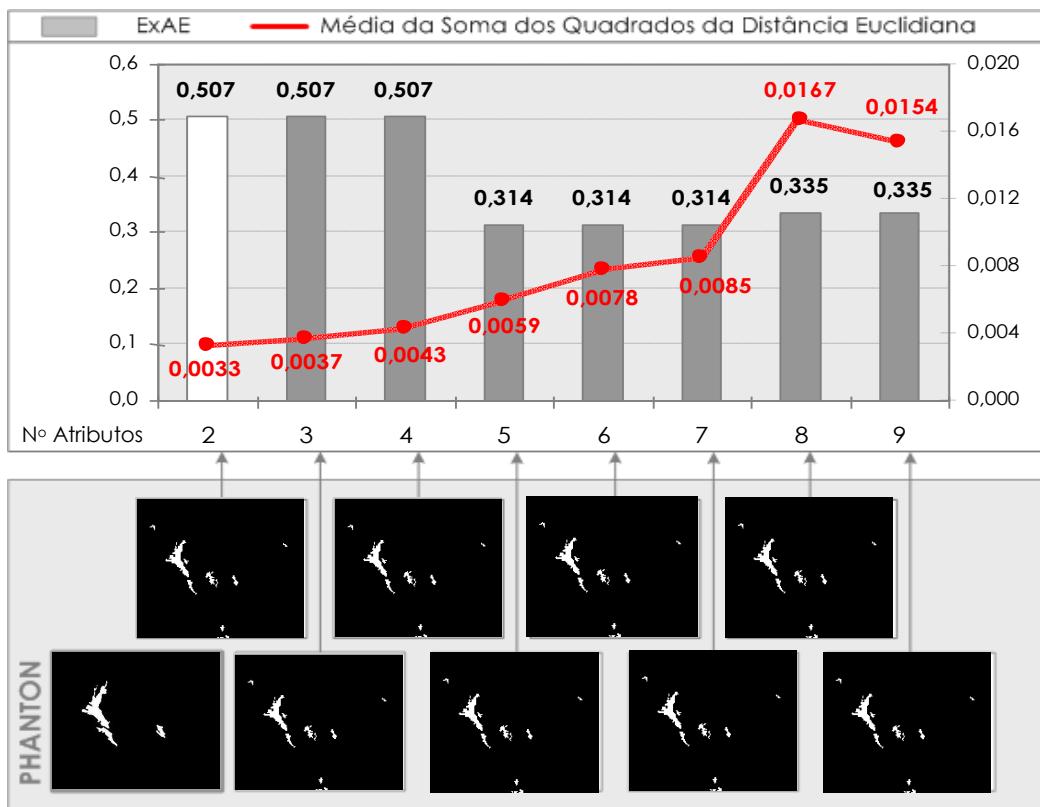


Figura 5.26 – Gráfico e espacialização das clusterizações realizadas para a Cena 6.

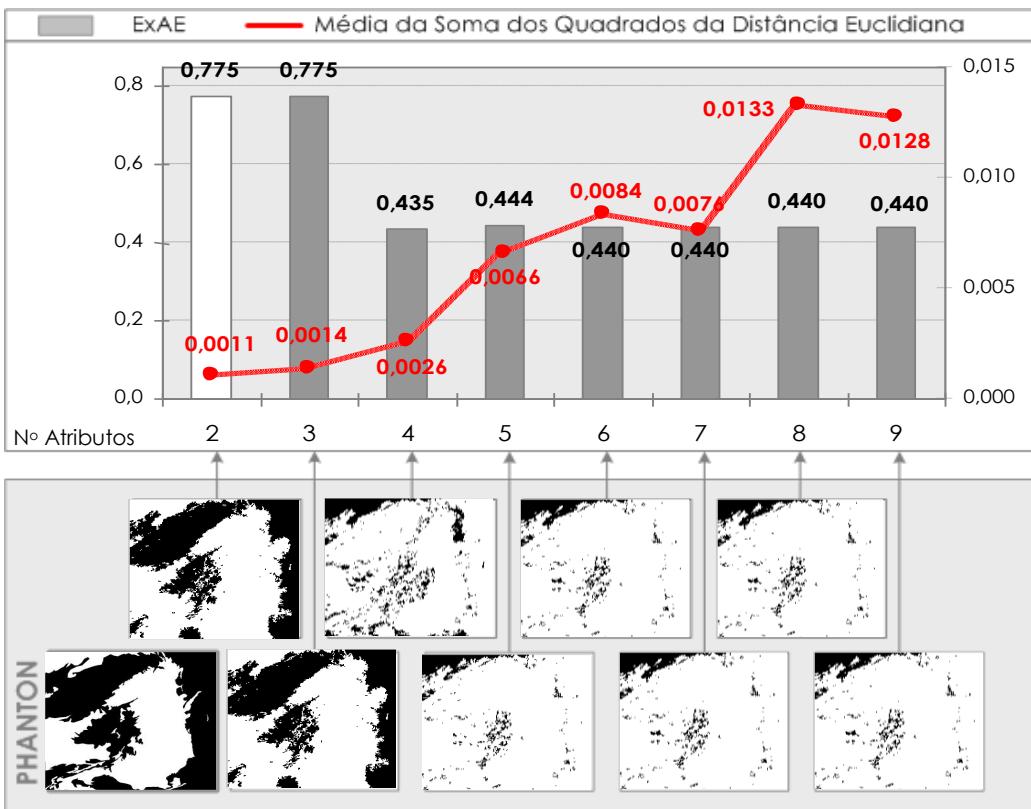


Figura 5.27 – Gráfico e espacialização das clusterizações realizadas para a Cena 7.

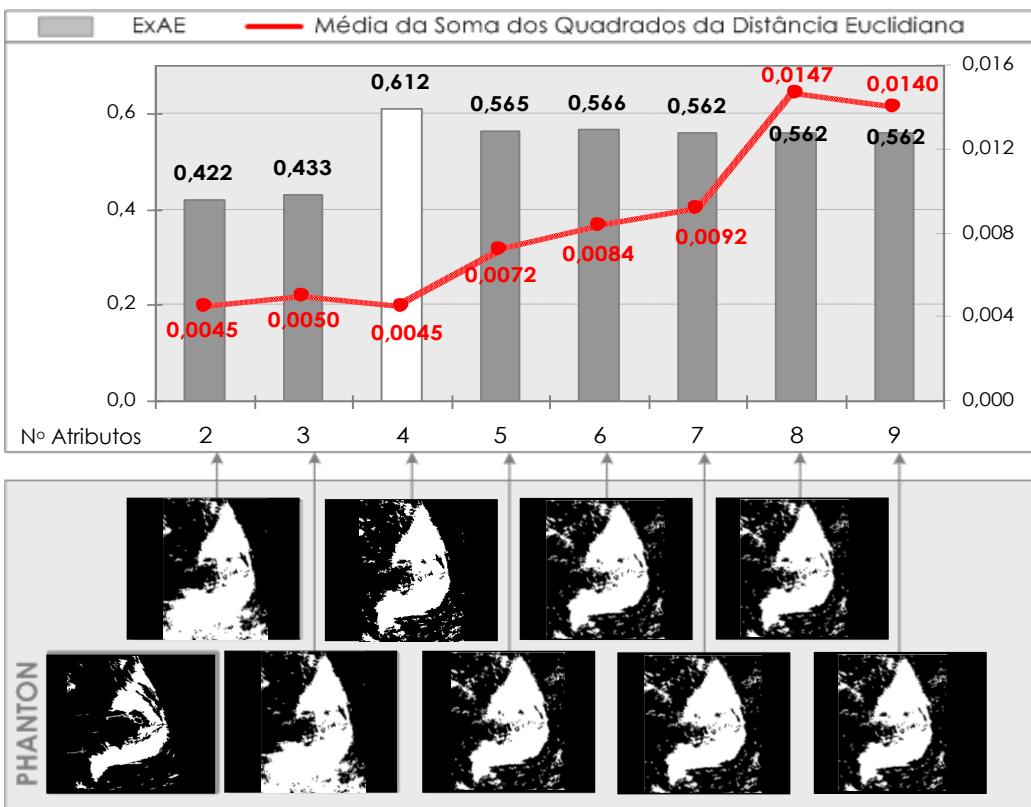


Figura 5.28 – Gráfico e espacialização das clusterizações realizadas para a Cena 8.

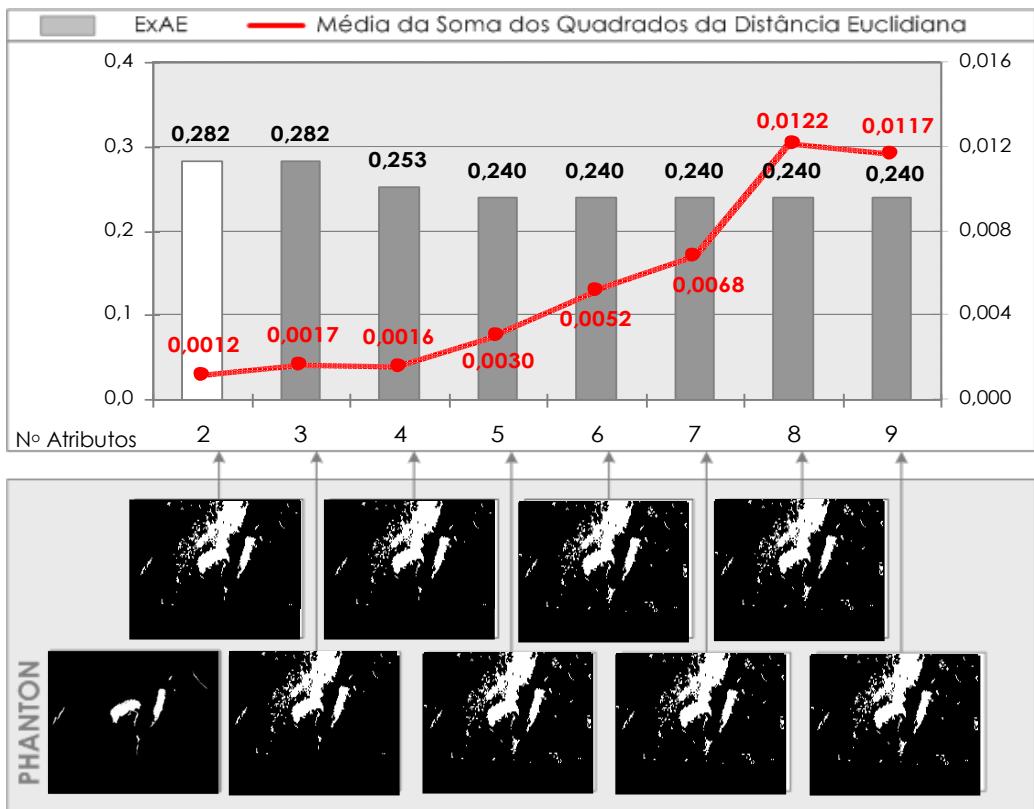


Figura 5.29 – Gráfico e espacialização das clusterizações realizadas para a Cena 9.

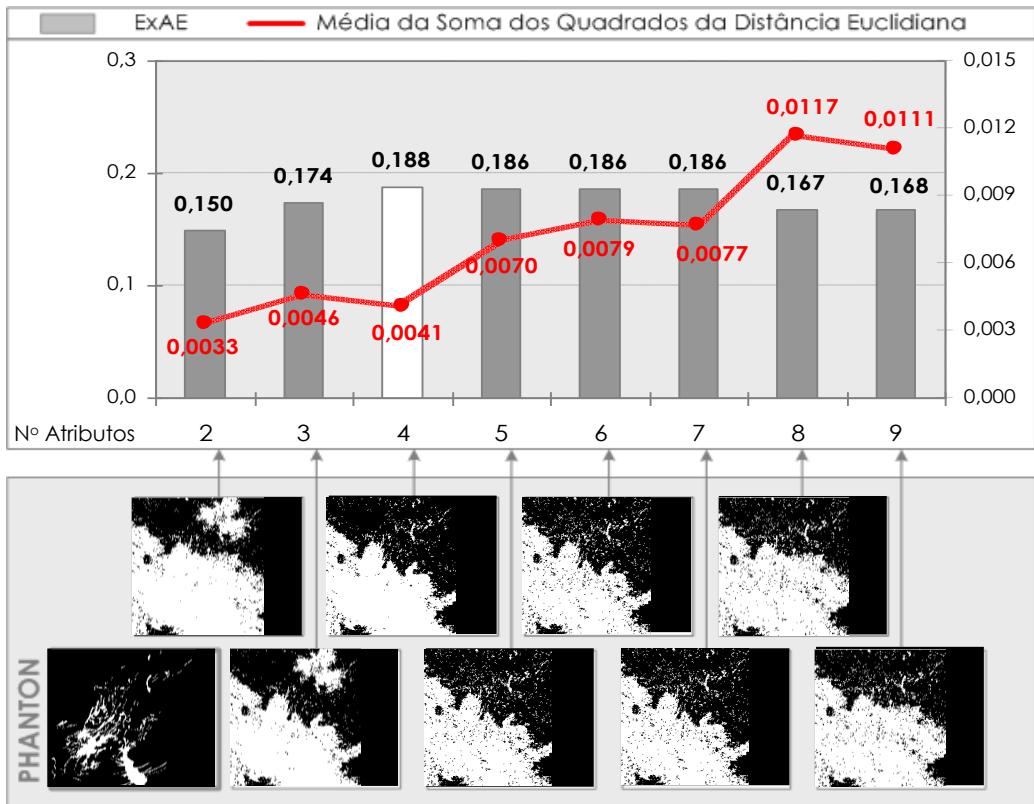


Figura 5.30 – Gráfico e espacialização das clusterizações realizadas para a Cena 10.

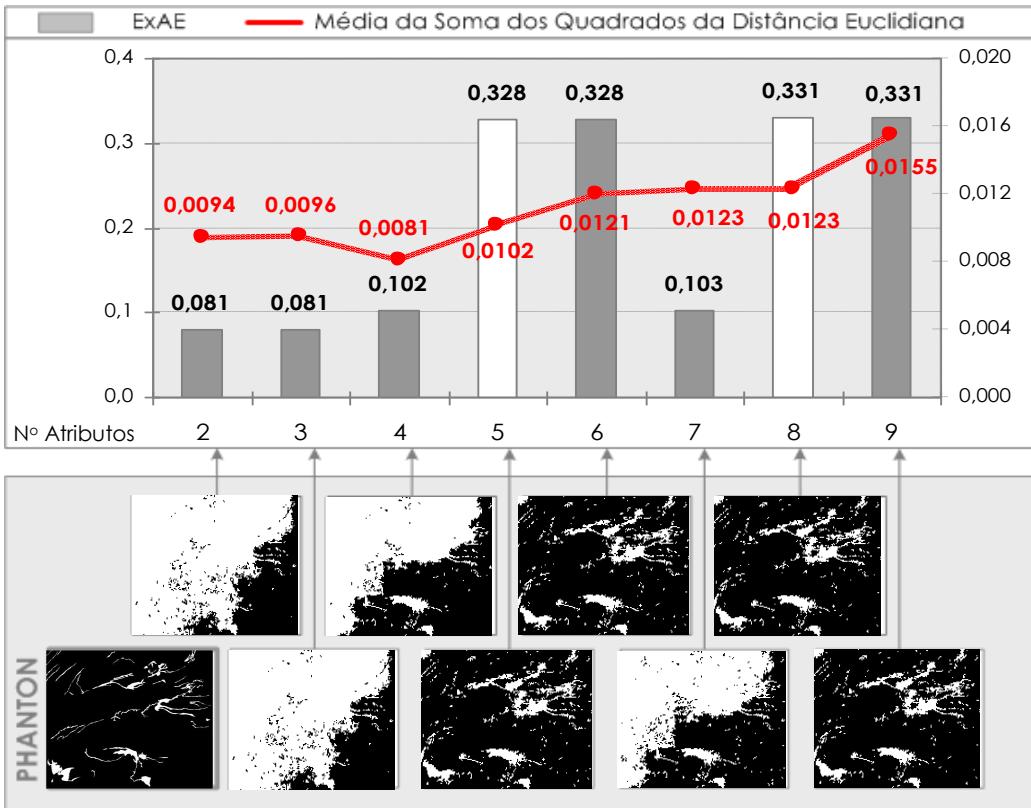


Figura 5.31 – Gráfico e espacialização das clusterizações realizadas para a Cena 11.

A *ME* tende a aumentar com o aumento do número de atributos, isto ocorre porque a cada nova clusterização o algoritmo procura convergir a melhor solução minimizando a *ME*, a inclusão de atributos tende a aumentar a *ME* devido ao aumento gradativo da dispersão dos *clusters*.

O critério utilizado para definir se o resultado de uma clusterização foi satisfatório, considerou uma *ExAE* maior ou igual a 0,5, equivalente a no mínimo 50% de representatividade em relação ao *Phantom*. Este critério foi definido interativamente pelo intérprete, a partir da comparação visual dos resultados obtidos pelo clusterizador em relação ao *Phantom*. Este critério poderia ser mais rigoroso, todavia, a imprecisão na definição exata dos limites de borda que separam um alvo escuro do oceano é inerente às imagens SAR, devido a inúmeros fatores descritos no Capítulo 2. Conforme comentado anteriormente, o próprio *Phantom* possui imprecisão associada ao delineamento manual dos limites de borda.

A cena 1 (Figura 5.21) convergiu para a melhor *ExAE* com 4 atributos e a menor *ME*. Neste caso, a utilização de um número menor de atributos (2 ou 3) prejudicou a performance da clusterização, com diferenças identificáveis na interpretação visual dos *clusters* produzidos. Quando comparados ao *Phantom*, os

resultados obtidos pela clusterização foram considerados não satisfatórios, sendo sua representatividade menor do que 50%, com índices de *ExAE* entre 0,32 e 0,40.

A inserção dos 9 atributos um a um na cena 2 (Figura 5.22) não alterou o índice de *ExAE*, e não provocou alterações visíveis na representatividade dos *clusters* obtidos. Neste caso, com apenas 2 atributos e a menor *ME* o algoritmo convergiu para a melhor *ExAE*. Os *clusters* obtidos pelo método automático foram considerados satisfatórios quando comparados visualmente com o *Phantom*, alcançando *ExAE* de 0,80, indicando 80% de representatividade.

Na cena 3 (Figura 5.23), não foi observado ganho significativo nos índices de *ExAE* com o aumento gradativo do número de atributos. Embora as maiores Exatidões tenham sido observadas com a utilização de 7, 8 e 9 atributos (0,349), a diferença entre estes e os índices obtidos com 2, 3 e 4 atributos (0,348) foi irrisória 0,001. Neste caso, com apenas 2 atributos foi possível convergir para a melhor solução, sendo que a inserção dos demais atributos não alterou os resultados obtidos. Nesta cena, a clusterização foi considerada não satisfatória com representatividade máxima de 35%.

A cena 4 (Figura 5.24) convergiu para a melhor *ExAE* com 9 atributos e com uma maior *ME*. Visualmente, os *clusters* resultantes não apresentaram mudanças discrepantes com a utilização de mais de 4 atributos, entretanto, a utilização de 2 ou 3 atributos prejudicou os resultados. Quando comparados ao *Phantom*, os resultados foram considerados não satisfatórios, sendo sua representatividade menor do que 50%, com índices de *ExAE* oscilando entre 0,08 e 0,215.

Na cena 5 (Figura 5.25) também é evidente que a utilização de 2 e 3 atributos não favoreceu o clusterizador. A inserção gradativa de atributos não alterou a representatividade visual dos clusters e, considerando a *ExAE*, os índices oscilaram de minimamente entre 0,242 e 0,248, representando um ganho de 0,006. Neste caso, o modelo mais simplificado seria utilizar 4 atributos com a menor *ME*. Os resultados também foram considerados não satisfatórios quando comparados com o *Phantom*, alcançando *ExAE* menores que 50%.

Na cena 6 (Figura 5.26) com apenas 2 atributos e a menor *ME* o algoritmo convergiu para a melhor *ExAE*, sendo este índice mantido estável com a inserção de até 4 atributos. De forma semelhante à cena 2, visualmente a diferença qualitativa entre os *clusters* obtidos com a inserção de até 9 atributos foi muito pequena. Neste caso, os *clusters* obtidos foram considerados satisfatórios quando comparados visualmente com o *Phantom*, alcançando 50% (0,51) de representatividade.

Um comportamento diferenciado foi observado na cena 7 (Figura 5.27), onde os melhores índices de *ExAE* foram obtidos com apenas 2 e 3 atributos, a partir dos quais a inserção gradativa de atributos prejudicou a detecção dos alvos escuros. Neste caso, seriam necessários pelo menos 2 atributos para atingir uma *ExAE* máxima. A performance do clusterizador automático foi considerada satisfatória, atingindo 77% (0,775) de representatividade em relação ao *Phantom*.

Para a cena 8 (Figura 5.28), a maior *ExAE* (0,612), juntamente com a menor *ME*, foi obtida com 4 atributos. Da mesma forma que para as cenas 1, 4 e 5, a utilização de 2 e 3 atributos prejudicou visualmente a performance do clusterizador na separação dos *clusters*. Neste caso, a solução que convergiu para a maior *ExAE* foi obtida com a utilização de 4 atributos. Os resultados obtidos pela clusterização automática foram considerados satisfatórios, sendo sua representatividade em relação ao *Phantom* maior do que 50%, com *ExAE* entre 0,56 (56%) e 0,61 (60%).

Na cena 9 (Figura 5.29) o maior índice de *ExAE* foi obtido com a inserção de 2 atributos, mantendo-se estável com até 3 atributos. Embora a *ExAE* tenha diminuído com a inserção gradativa de mais de 3 atributos, visualmente os *clusters* resultantes não apresentaram mudanças discrepantes. Neste caso, a padronização do número de atributos de entrada acima de 2 não alteraria visualmente a representatividade dos *clusters*. O desempenho do clusterizador foi considerado não satisfatório, atingindo exatidão máxima de 28%, menos que 50% em relação ao *Phantom*.

A melhor convergência do clusterizador para a cena 10 (Figura 5.30) atingiu *ExAE* máxima de 0,19 (19%) e foi obtida com a utilização de 4 atributos, acompanhada da menor *ME*. Este índice de *ExAE* manteve-se estabilizado até a inserção do 7º atributo, diminuindo cerca de 0,02 com a inserção do 8º e do 9º atributo. Embora a *ExAE* tenha diminuído, este comportamento não foi significativo em termos de representatividade em relação ao *Phantom*, uma vez que, os resultados foram considerados não satisfatórios, com representatividade menor que 50%, para todos os testes efetuados, atingindo *ExAE* máxima de 19% (0,186).

Para a cena 11 (Figura 5.31) a melhor *ExAE* (0,33) foi obtida com 5 atributos, mantendo-se estabilizada com a utilização de 6, 8 e 9 atributos. Visualmente, a utilização de um número menor do que 5 atributos prejudicou a qualidade dos *clusters* produzidos. Neste caso, os resultados obtidos pela clusterização automática também foram considerados não satisfatórios, sendo sua representatividade em relação ao *Phantom* menor do que 50%, com máximo percentual de *ExAE* de 33%.

Na maioria das cenas estudadas foi possível convergir para uma melhor solução com um menor número de atributos, oscilando entre 2 e 4 e com menores valores para a *ME*. No conjunto das 9 clusterizações realizadas para cada uma das 11 cenas, os maiores índices de *ExAE* foram obtidos com: a) apenas 2 atributos em 5 cenas; b) 4 atributos em 3 cenas; c) 5 atributos em 2 cenas, e; d) 9 atributos em 1 cena.

Nos casos em que a máxima *ExAE* foi obtida com um número maior ou igual a 4 atributos (6 cenas), a utilização de 2 a 3 atributos produziu *clusters* menos representativos em 6 das 11 cenas estudadas, sendo identificável tanto nos índices de *ExAE*, quanto visualmente. Estes resultados são interessantes porque reforçam que um número pequeno de atributos pode não ser suficiente para o clusterizador convergir *clusters* representativos, assim como um número exagerado de atributos também pode prejudicar os resultados.

As clusterizações também foram realizadas inserindo, seqüencialmente, um a um os 35 atributos (desconsiderando o atributo de Mínimo), considerando a ordem de importância indicada pelo chi-quadrado para cada cena (APÊNDICE V) conforme etapas indicadas no fluxograma metodológico da Figura 5.18. Este estudo foi realizado, para verificar se haveria diferença significativa na performance do clusterizador considerando o total de atributos e não apenas os 9 melhores do *ranking*.

Os resultados desta comparação são de extrema importância para validar se a automatização do sistema é factível a partir de um conjunto padrão mínimo de atributos, a ser utilizado a cada nova imagem SAR adquirida. A Tabela 5.21 apresenta a síntese dos resultados obtidos para as clusterizações efetuadas com os 9 melhores atributos do *ranking* e com os 35 atributos. Foram calculadas para cada cena a *ExAE* obtida com a clusterização simultânea dos 9 e dos 35 atributos e a Máxima *ExAE* obtida com a inserção gradual dos 9 e dos 35 atributos respeitando, respectivamente, a ordem de importância indicada na Tabela 5.19 e indicada pelo chi-quadrado (APÊNDICE V). O número de atributos necessário para retornar a máxima *ExAE* será variável de cena para cena.

O teste não paramétrico de Wilcox foi realizado com o objetivo de verificar se as Exatidões (*ExAE*) e Máximas Exatidões (*MExAE*) obtidas nas 4 situações indicadas na Tabela 5.21 eram estatisticamente diferentes entre si.

Tabela 5.21 - Síntese dos resultados obtidos considerando as clusterizações efetuadas com os 9 e 35 atributos.

Resultados Obtidos pela Clusterização		Cena 1	Cena 2	Cena 3	Cena 4	Cena 5	Cena 6	Cena 7	Cena 8	Cena 9	Cena 10	Cena 11	Média	Media	Máximo
Ranking dos 9 atributos pela análise efetuada	<i>ExAE: com todos os 9 Atributos</i>	0,351	0,799	0,349	0,215	0,248	0,335	0,440	0,562	0,240	0,168	0,331	0,367	0,335	0,799
Tab. 5.27	Nº de Atributos Necessários para atingir esta ExAE	9													
	Máxima ExAE obtida: inserção gradual dos 9 Atributos	0,392	0,799	0,348	0,215	0,248	0,507	0,775	0,612	0,282	0,186	0,331	0,427	0,348	0,799
	Nº de Atributos Necessários para atingir esta ExAE	4	2	4	9	7	2	3	4	2	5	9	5	4	9
Ranking dos 35 atributos fornecido pelo Chi-Quadrado	<i>ExAE: com todos os 35 Atributos</i>	0,190	0,002	0,062	0,054	0,020	0,228	0,476	0,257	0,240	0,191	0,337	0,187	0,191	0,476
	Nº de Atributos Necessários para atingir esta ExAE	35													
	Máxima ExAE obtida: inserção gradual dos 35 Atributos	0,392	0,802	0,349	0,215	0,284	0,547	0,841	0,612	0,282	0,194	0,328	0,441	0,349	0,841
	Nº de Atributos Necessários para atingir esta ExAE	4	12	4	16	12	18	2	12	3	5	5	8	5	18

Na maioria das cenas, os valores de *ExAE* obtidos considerando a clusterização simultânea dos 9 atributos foram maiores do que os obtidos para a clusterização simultânea de 35 atributos. Conforme esperado, o teste de Wilcox indicou que esta diferença nos valores de *ExAE* para 9 e 35 atributos foi estatisticamente significativa, com *p* valor de 0,028. Este resultado validou a hipótese de que, a partir de certo número de atributos, a inserção de novas variáveis ao modelo insere pouca ou nenhuma informação podendo inclusive prejudicar o desempenho do clusterizador (HUGHES, 1968). Desta forma, a seleção de um conjunto reduzido de atributos considerados significativos, é melhor do que a utilização de um amplo conjunto de dados que também inclui atributos não representativos, os quais podem prejudicar a performance do clusterizador.

Os Máximos valores de exatidão *MExAE* obtidos considerando a inserção gradativa dos 9 e dos 35 atributos, apresentaram comportamento semelhante quando comparados cena a cena. O teste de Wilcox confirmou, com um *p* valor de 0,068, que não há diferença estatisticamente significativa entre os valores de Máxima *MExAE* obtidos para a melhor combinação dentre os 9 ou dentre os 35 atributos. Esse resultado é de grande importância por validar a possibilidade de automatização do sistema a partir de um conjunto reduzido de atributos. A Figura 5.32 disponibiliza a espacialização das máximas *MExAE* obtidas considerando o melhor conjunto de dados para o universo de 9 e de 35 atributos, indicando o número necessário de atributos para alcançar estes índices.

A espacialização do resultado das clusterizações evidenciou que praticamente não há diferença em utilizar a melhor combinação de atributos obtida entre os 9 ou entre os 35 atributos, o que reforçou a possibilidade de automatização.

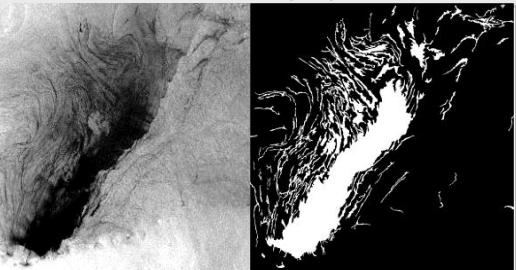
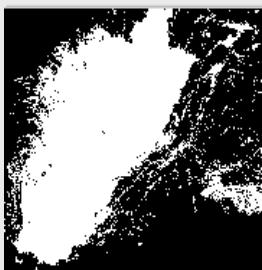
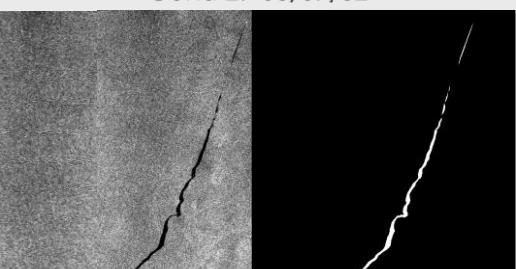
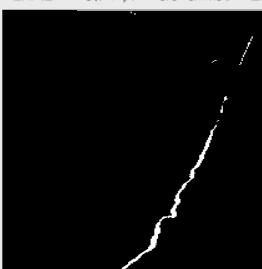
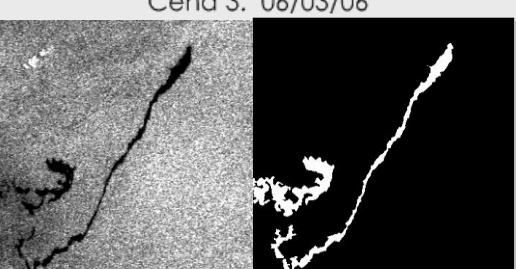
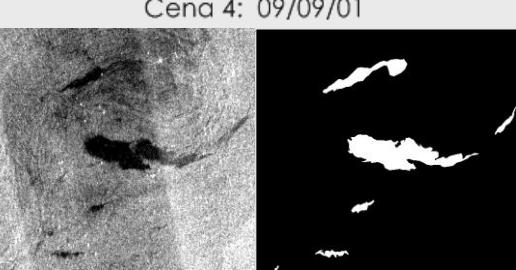
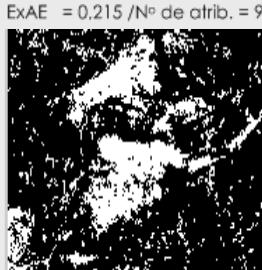
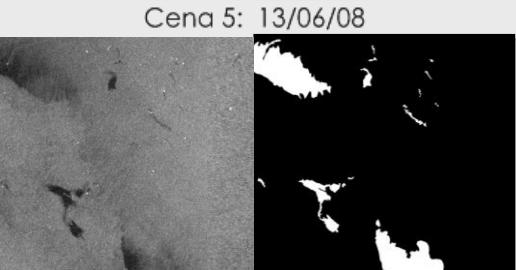
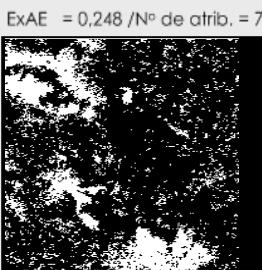
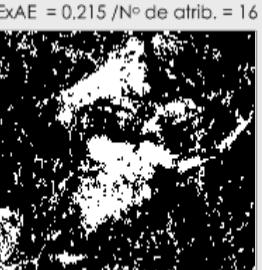
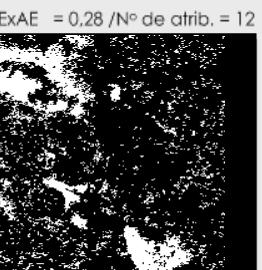
IMAGEM DE RADAR E PHANTON	MÁXIMA EXATIDÃO: 9 ATRIBUTOS	MÁXIMA EXATIDÃO: 35 ATRIBUTOS
Cena 1: 06/07/01		 ExAE = 0,39 /Nº de atrib. = 4
Cena 2: 06/07/02		 ExAE = 0,79 /Nº de atrib. = 2
Cena 3: 06/03/06		 ExAE = 0,348 /Nº de atrib. = 4
Cena 4: 09/09/01		 ExAE = 0,215 /Nº de atrib. = 9
Cena 5: 13/06/08		 ExAE = 0,248 /Nº de atrib. = 7
		 ExAE = 0,39 /Nº de atrib. = 4
		 ExAE = 0,80 /Nº de atrib. = 12
		 ExAE = 0,349 /Nº de atrib. = 4
		 ExAE = 0,215 /Nº de atrib. = 16
		 ExAE = 0,28 /Nº de atrib. = 12

Figura 5.32 - Espacialização dos melhores resultados da clusterização considerando o universo de 9 e 35 atributos

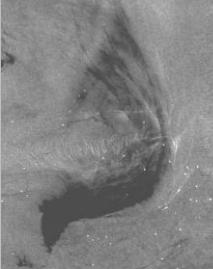
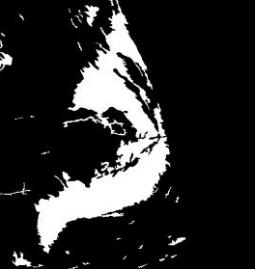
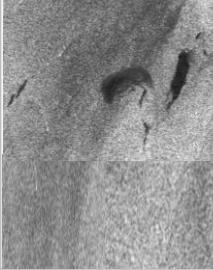
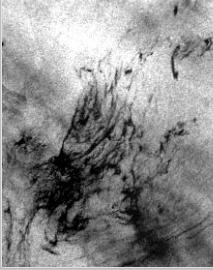
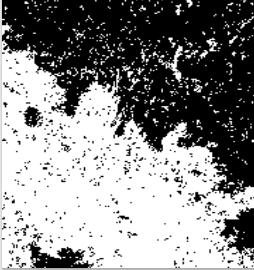
IMAGEM DE RADAR E PHANTON	MÁXIMA EXATIDÃO: 9 ATRIBUTOS	MÁXIMA EXATIDÃO: 35 ATRIBUTOS
Cena 6: 05/05/02	 	
	ExAE = 0,507 /Nº de atrib. = 2	ExAE = 0,547/Nº de atrib. = 18
Cena 7: 19/12/01	 	
	ExAE = 0,775 /Nº de atrib. = 3	ExAE = 0,841 /Nº de atrib. = 2
Cena 8: 03/04/02	 	
	ExAE = 0,61 /Nº de atrib. = 4	ExAE = 0,61/Nº de atrib. = 12
Cena 9: 04/07/01	 	
	ExAE = 0,282 /Nº de atrib. = 2	ExAE = 0,282 /Nº de atrib. = 3
Cena 10: 11/03/05	 	
	ExAE = 0,186/Nº de atrib. = 5	ExAE = 0,194 /Nº de atrib. = 5

Figura 5.32 - Continuação: Espacialização dos melhores resultados da clusterização considerando o universo de 9 e 35 atributos

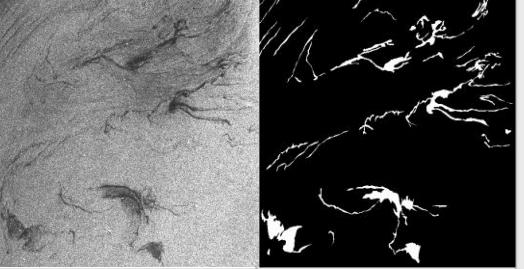
IMAGEM DE RADAR E PHANTON	MÁXIMA EXATIDÃO: 9 ATRIBUTOS	MÁXIMA EXATIDÃO: 35 ATRIBUTOS
Cena 11: 14/06/08 	ExAE = 0,337 /Nº de atrib. = 9	ExAE = 0,328 /Nº de atrib. = 5

Figura 5.32 - Continuação: Espacialização dos melhores resultados da clusterização considerando o universo de 9 e 35 atributos

Apesar da automatização considerando um número reduzido de atributos padrão ter sido indicada como factível, a análise global dos resultados evidenciou que, dentre as 11 cenas estudadas, o clusterizador foi capaz de convergir um conjunto de *clusters* considerados satisfatórios, com *ExAE* maiores ou iguais a 0,5, em apenas 4 cenas. Estes polígonos poderiam ser utilizados como entrada em um sistema de classificação, entretanto, os demais prejudicariam a performance do classificador uma vez que os atributos seriam calculados como base nestas geometrias.

Do ponto de vista operacional¹⁰, a detecção automática de um ou mais alvos escuros na imagem, mesmo que imprecisa, seria suficiente para acionar uma possível operação de contingenciamento, mesmo com *ExAE* inferiores a 0,5. As cenas 3, 9 e 11 enquadram-se nessa interpretação. A Figura 5.33 apresenta a sobreposição dos alvos escuros delineados como referência nos *Phantons* em vermelho, com os alvos escuros detectados automaticamente pelo clusterizador em branco.

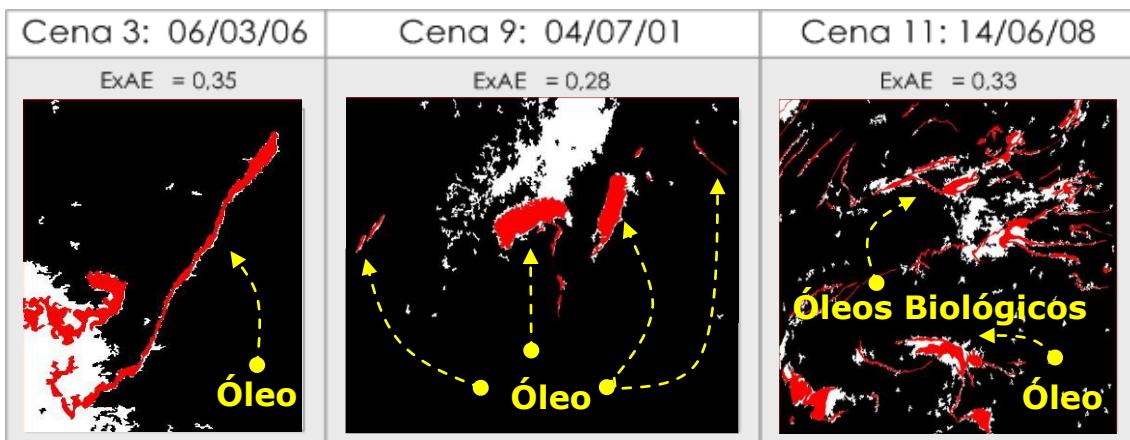


Figura 5.33 - Sobreposição dos *Phantons* com os resultados das clusterizações.

¹⁰ No contexto da tese a expressão “do ponto de vista operacional” remete à experiência que consolidei durante 6 anos de atuação junto à PETROBRAS na gerência de Controle de Contingências, tendo sido responsável pela atividade de Monitoramento Aéreo/Orbital.

A cena 11 constitui um exemplo interessante do ponto de vista operacional por mesclar, em um mesmo recorte de cena, óleo e assemelhados. Neste caso, mesmo que os resultados da classificação fossem duvidosos em função da imprecisão associada à geometria dos polígonos, uma verificação de campo seria necessária, pois não haveria garantia de que nenhum dos filetes detectados como alvos escuros poderia ser óleo.

Os resultados das clusterizações para as cenas 4, 5 e 10, foram muito mais ruidosos e não poderiam ser considerados satisfatórios. Para todos estes casos, as manchas de óleo foram identificadas e, operacionalmente, uma investigação sobre os eventos seria efetuada. Entretanto, como os resultados foram muito confusos, uma intervenção humana seria indispensável para indicar os polígonos de interesse. A Figura 5.34 apresenta a sobreposição dos *Phantons* em vermelho, com os alvos escuros clusterizados em branco.

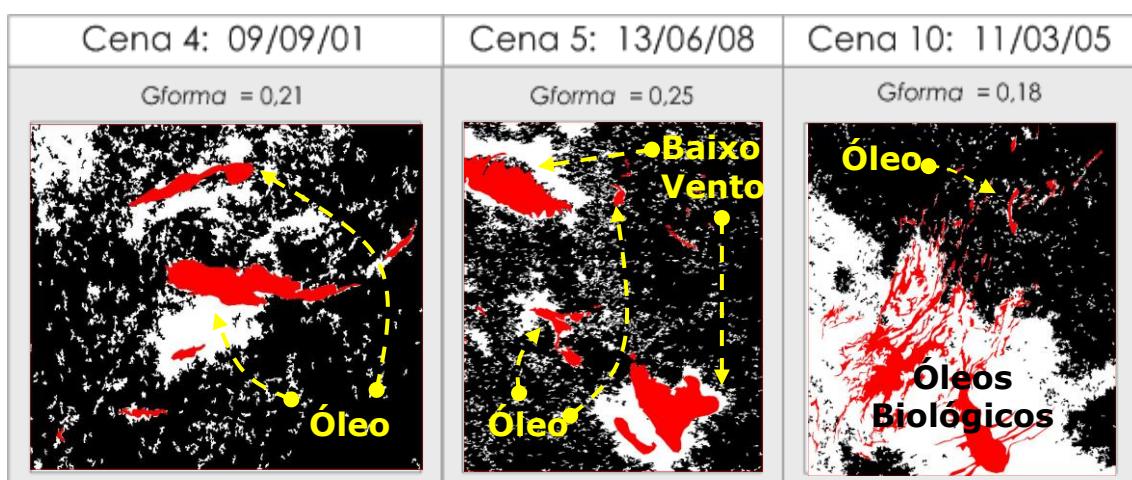


Figura 5.34 - Sobreposição dos *Phantons* com os resultados das clusterizações.

Portanto, mesmo com imprecisão, os alvos escuros detectados poderiam ser úteis para o acionamento das inspeções de campo. Entretanto, as geometrias não apresentaram acurácia suficiente em relação ao *Phantom* para serem utilizadas como dado de entrada em um sistema de classificação automática. Na pior hipótese, ao final da classificação, um resultado errôneo poderia ocasionar falhas no processo de combate a uma possível emergência envolvendo derramamento de óleo no mar. Também não seria aconselhável armazenar geometrias imprecisas em um Banco de Dados que tem por objetivo - além de registrar o número e o tipo dos eventos detectados - fornecer conhecimento pretérito para retroalimentar o sistema automático de detecção de óleo e assemelhados. O registro cumulativo de geometrias imprecisas e, consequentemente, de atributos imprecisos no Banco de Dados poderia levar à

ineficácia do sistema, confundindo o classificador na distinção dos alvos escuros em óleo ou assemelhados.

Com base nesta análise quantitativa e qualitativa dos resultados da clusterização, 5 cenas poderiam ser consideradas como satisfatórias sob a ótica do sistema automático, as cenas 2, 6, 7, 8 e 11, representando 45% dos recortes de cena. A primeira avaliação seria de que a performance do clusterizador não foi satisfatória, entretanto, é preciso salientar que, para se chegar a essa conclusão, seriam necessárias mais amostras, sempre considerando níveis de complexidade diferenciados. No âmbito da tese desenvolvida, seria impossível trabalhar com mais exemplos, considerando, a abrangência dos testes efetuados que avaliaram parâmetros envolvidos em todas as etapas, desde o pré-processamento das imagens.

É importante destacar que, especialmente quando os alvos escuros estavam associados a óleos biológicos, eventos difíceis de serem representados manualmente devido ao inúmero conjunto de estruturas filiformes, a imprecisão associada ao *Phantom* aumentou. Nestes casos, situações em que o clusterizador não incluiu exageradamente áreas oceânicas como alvos escuros, uma menor *ExAE* nos resultados da clusterização foi considerada satisfatória, como no exemplo da cena 11. Em contrapartida, exemplos como o da cena 10 não foram aceitos devido à inclusão exagerada de áreas oceânicas como parte dos alvos escuros.

A Figura 5.35 apresenta a *ExAE* dos resultados obtidos pelas clusterizações realizadas para as 5 cenas consideradas satisfatórias, destacando a espacialização: a) da intersecção dos alvos escuros: áreas detectadas corretamente como alvos escuros pela clusterização em relação ao *Phantom* (em branco); b) da Omissão: áreas de alvos escuro erroneamente clusterizadas como oceano (em vermelho); c) da Inclusão: áreas de oceano erroneamente clusterizadas como parte dos alvos escuro (em azul), e; d) áreas detectadas corretamente como oceano pela clusterização em relação ao *Phantom* (em preto).

Considerando que a contribuição relativa de cada atributo foi avaliada através de inúmeras clusterizações efetuadas no conjunto de 9 e de 35 atributos, os resultados obtidos representam o melhor desempenho do clusterizador em cada cena. A dificuldade associada à detecção automática de óleo em imagens SAR é destacada por inúmeros autores (TOPOUZELIS, 2008; TOPOUZELIS *et al.*, 2007a; CHANG *et al.* 2008, BREKKE & SOLBERG, 2005), sendo os fatores que influenciam na performance da detecção automática muito diversos. No âmbito dos experimentos realizados, as

observações a seguir buscam fornecer subsídios para enriquecer a interpretação dos resultados obtidos.

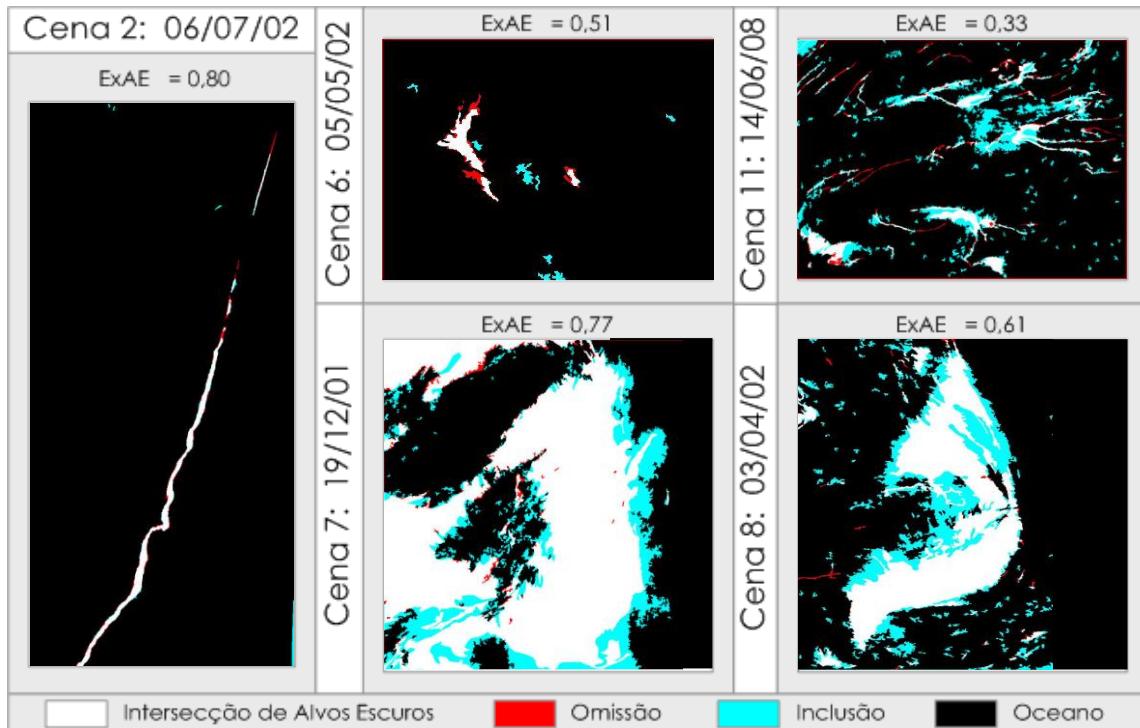


Figura 5.35 – Exatidão, Omissão e Inclusão nos resultados obtidos pelas clusterizações para as 5 cenas consideradas satisfatórias.

O maior ou menor potencial de detecção associado ao clusterizador esteve diretamente relacionado ao contraste existente entre os alvos escuros e o oceano. Entre as cenas escolhidas a intensidade deste contraste foi muito variada, sem um padrão referente à distribuição dos NC que fosse capaz de garantir que os mesmos atributos utilizados em cenas diferentes oferecessem um mesmo desempenho ao clusterizador.

Esta variação do contraste entre alvos escuros e oceano, dificultando o processo de clusterização, também esteve presente no contexto de uma mesma cena, sendo influenciada pela variação das condições meteo-oceanográficas na superfície do mar (cenas: 9, 10 e 11) e pela presença de falhas no imageamento (cena: 4). Os inúmeros fatores que influenciam esta ampla variação do contraste entre as cenas e nas cenas, os quais podem ocasionar o cálculo de atributos inexpressivos para o agrupamento das regiões, estão detalhados no Capítulo 2 sendo os principais: ângulo de incidência; presença do ruído *speckle*; parâmetros de largura da faixa, resolução espacial, freqüência do sinal e polarização associados ao modo de imageamento do(s) sensor(es); condições meteo-oceanográficas; quantidade e o tipo de óleo vazado; intervalo de tempo entre o vazamento e o momento da aquisição da imagem; grau de

intemperismo das manchas, dentre outros. As características do imageamento SAR em áreas oceânicas que dificultam a implementação de procedimentos automáticos encontram-se detalhadas no Capítulo 2 - item 2.2.

Além dos fatores citados, a alteração das propriedades estatísticas ocasionada pela compactação/descompactação das imagens no formato *MrSid* (Capítulo 4: item 4.1), também influenciou no desempenho das clusterizações, uma vez que as propriedades dos atributos calculados diferem dos calculados a partir de imagens originais. Para avaliar o quanto este formato de fato influenciou positivamente ou negativamente nos resultados, seria necessário efetuar clusterizações com os mesmos atributos calculados a partir de imagens no formato original. Entretanto, como o objetivo da tese estava focado no desenvolvimento de um método passível de ser implementado operacionalmente, imagens no formato original não foram utilizadas.

De certa forma, o que mais contribuiu para que os resultados da clusterização não fossem melhores foi o rigor utilizado na escolha dos recortes de cena estudados. Se as cenas tivessem sido recortadas sempre individualizando um único alvo com um bom contraste em relação ao oceano (priorizando eventos relacionados a óleo) - como frequentemente observado na literatura disponível – um maior número de cenas teria atingido o critério de Exatidão de Alvos Escuros (*ExAE*) necessário para que a clusterização fosse considerada satisfatória.

Uma reflexão importante é que a escolha dos exemplos a serem testados para validar um sistema automático pode tornar-se altamente tendenciosa. A importância da escolha das cenas teste é tal que, a conclusão de que um sistema proposto é automático e pode ser utilizado operacionalmente, pautada unicamente em exemplos de fácil reconhecimento, pode representar uma falácia.

Os recortes de cena consideraram desde exemplos comumente encontrados na literatura (cenas 2, 3 e 6), onde manchas de óleo com geometria bem definida e bom contraste em relação ao oceano foram selecionadas, até exemplos de maior complexidade em que ocorrências de óleo e assemelhados foram processadas simultaneamente (cenas 5, 10 e 11). Nesta escolha, a habilidade, tanto dos segmentadores como do clusterizador, em detectar alvos em diferentes escalas e com diferentes geometrias também foi avaliada. Alvos com geometria complexa associados a óleos biológicos e florescimento de algas também foram considerados isoladamente em alguns exemplos (recortes de cenas 1 e 8).

Esta dificuldade foi imposta ao clusterizador propositalmente, para que os resultados pudessem ser replicados e o método utilizado em outras imagens, justamente por ter sido definido considerando exemplos com nível diferenciado de dificuldade. Esta seleção buscou refletir ao máximo a realidade encontrada nas imagens SAR, incluindo até problemas no imageamento que podem ocorrer durante a aquisição das imagens pelo sensor ou no processamento das imagens, sendo visível a descontinuidade existente entre as faixas de imageamento na cena 4. Para esta cena, o clusterizador não conseguiu encontrar uma boa solução uma vez que não havia um comportamento padrão para os intervalos de NC, capaz de diferenciar os alvos escuros do oceano ao longo da mesma cena. Exemplo semelhante foi observado na cena 10, onde filetes de óleos biogênicos ocorreram simultaneamente sobrepostos a uma área com uma baixa intensidade de vento, reduzindo seu contraste na cena.

O aspecto negativo relacionado à escolha dos recortes de cena, foi que os resultados atingidos pelo clusterizador não foram satisfatórios para a maioria das cenas. O aspecto positivo foi o esforço realizado para interpretar os resultados, entender as propriedades dos atributos utilizados, evidenciando a necessidade da continuidade das pesquisas considerando estes exemplos que se aproximem da realidade encontrada nas imagens SAR. A busca de novos métodos de pré-processamento de imagens, que possam melhorar o contraste entre os alvos escuros e o oceano aprimorando a qualidade dos atributos calculados e, consequentemente, a performance do clusterizador, é uma linha de investigação interessante, considerando os inúmeros fatores que prejudicam a detecção dos alvos escuros em imagens SAR.

Para dar continuidade à pesquisa proposta foi preciso avançar para o próximo passo, verificar a possibilidade de fixar o número de atributos necessários para automatizar o sistema. Retornando aos resultados da Tabela 5.21, a Média indicou 5 atributos e a Mediana 4 atributos como necessários para se atingir a Máxima Exatidão MExAE, selecionando-os a partir do conjunto de 9 atributos indicados pela análise efetuada na Tabela 5.19. Estas indicações reforçaram a seleção de atributos efetuada pelo método de chi-quadrado, o qual também indicou que, de maneira geral, considerando os 36 atributos, as maiores importâncias estavam concentradas entre os 4 e 5 primeiros atributos, sendo reduzidas abruptamente nos atributos subseqüentes. Os gráficos com o *ranking* da importância obtida para os 36 atributos para cada cena, durante a etapa de seleção dos atributos, estão disponíveis para consulta no APÊNDICE V.

A fim de subsidiar a decisão sobre a utilização de 4 ou 5 atributos, o teste não paramétrico de Wilcox foi efetuado para verificar se a diferença entre os índices de *ExAE* obtidos para as 11 cenas com 4 e 5 atributos (Tabela 5.22) são estatisticamente significativas.

Tabela 5.22 – Índices de *ExAE* obtidos para as 11 cenas com 4 e 5 atributos.

<i>ExAE</i> Obtidas na Clusterização	Cena 1	Cena 2	Cena 3	Cena 4	Cena 5	Cena 6	Cena 7	Cena 8	Cena 9	Cena 10	Cena 11	Média
4 Atributos	0,39	0,80	0,35	0,15	0,24	0,51	0,44	0,61	0,25	0,19	0,10	0,37
5 Atributos	0,38	0,80	0,32	0,17	0,25	0,31	0,44	0,56	0,24	0,19	0,33	0,36

Com um *p*-valor de 0,68, o teste evidenciou que não há diferença estatisticamente significativa entre as *ExAE* obtidas com 4 ou 5 atributos, indicando que a escolha entre 4 ou 5 atributos não faria diferença na implementação automática de um sistema para detecção de alvos escuros em imagens SAR.

Apesar dos *clusters* obtidos considerando 4 ou 5 atributos serem visualmente semelhantes para a maioria das cenas (Figuras 5.21 a 5.30), na cena 11 (Figura 5.31) a utilização do 5º atributo melhorou significativamente o resultado da clusterização. Como não foram obtidos dados conclusivos, decidiu-se manter os 5 atributos para compor o método de detecção de alvos escuros em imagens SAR, considerando que para a cena 11 o 5º atributo foi essencial para o bom desempenho do clusterizador. Excluindo o atributo de Mínimo, o 5º atributo corresponde a uma métrica de textura, o GLCM da Média. Este atributo foi selecionado entre os 5 primeiros em 8 cenas e entre os 6 primeiros em 3 cenas (Capítulo 5: item 5.4.2.4). A Figura 5.36 destaca as diferenças encontradas entre os *clusters* utilizando 4 e 5 atributos para a cena 11.

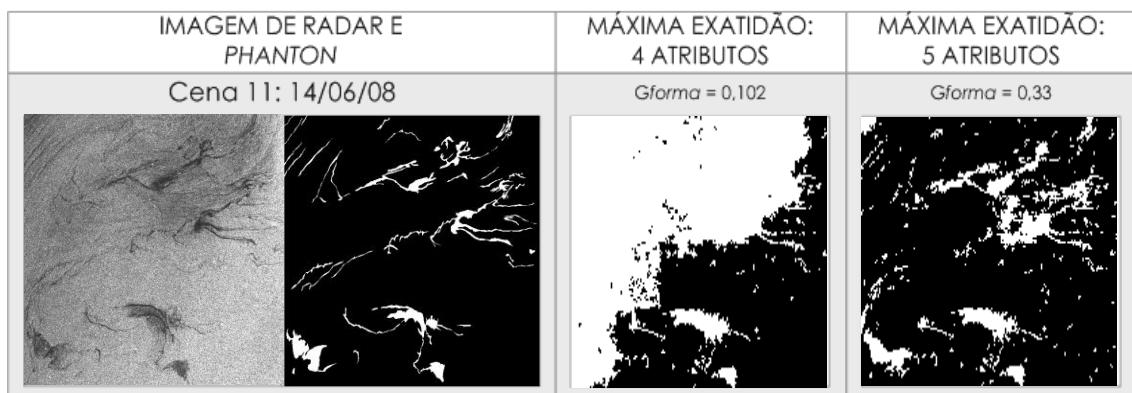


Figura 5.36 - Diferenças encontradas entre os *clusters* utilizando 4 e 5 atributos para a cena 11.

Ter considerado a contribuição relativa de cada atributo selecionado para detectar os diferentes tipos de alvos foi muito importante, pois forneceu subsídios para entender quais são as propriedades dominantes nas imagens e na geometria das

regiões que retornam os melhores contrastes nas cenas. Entretanto, para investigar a possibilidade de automatização do sistema, foi necessário utilizar os atributos mais recorrentes na seleção, sendo estes considerados os mais estáveis para serem utilizados como padrão. Neste contexto, os atributos estatísticos de primeira ordem Mediana, Média, Moda e Máximo, juntamente com o atributo Textural GLCM da Média, mostraram-se mais robustos por terem sido selecionados de forma recorrente em todas as cenas estudadas e validados com os resultados das clusterizações.

Importante destacar que apesar de existir multicolinearidade entre estes atributos (Item 5.4.2.4; APÊNDICE III), a sua utilização simultânea no *FP* foi o que gerou os melhores resultados para as clusterizações, validando esta seleção. Exemplos como o das cenas 4, 5, 8 e 11 evidenciaram que utilizando somente a Mediana e a Média os resultados não ficaram bons, ao inserir a Moda os *clusters* melhoraram significativamente (cenas 4, 5 e 8) e ao inserir o atributo de Máximo e GLCM da Média os *clusters* também melhoraram para a cena 11.

Com o objetivo de verificar se a seleção efetuada pelo chi-quadrado foi adequada, testes de seleção foram efetuados com mais 2 métodos implementados no software *Weka*, utilizando o método de avaliação por subconjuntos e dois métodos de busca, algoritmos genéticos (AG) e *Greedy Stepwise*¹¹ (GSTW). A busca efetuada pelo AG retornou 12 atributos sendo que todos os atributos estatísticos de primeira ordem multicorrelacionadas foram inclusos (Média, Mediana, Moda, Mínimo e Máximo) e a clusterização efetuada teve um bom resultado. A busca efetuada pelo GSTW retornou 6 atributos decorrelacionados, incluindo somente a Moda, a clusterização efetuada teve um resultado não satisfatório. Importante destacar que estes testes foram efetuados com a cena considerada mais fácil (cena 2), indicando que a exclusão dos atributos estatísticos de primeira ordem prejudica o reconhecimento de padrões.

A validação do método proposto e do conjunto de atributos selecionado foi realizada nos próximos dois itens nos quais os resultados obtidos automaticamente pelo *FP* no software *Clustan* foram comparados: a) com os resultados obtidos por métodos semi-automáticos aplicados às 11 cenas utilizadas ao longo da tese (item 5.4.4), e; b) com os resultados obtidos para os três novos recortes de cena escolhidos para validar o método proposto, utilizando os parâmetros definidos em cada uma das etapas da tese (item 5.5).

¹¹ *Greedy Stepwise*: combina a abordagem *forward* e *backward* durante a regressão.

5.4.4 – Comparação entre os Métodos Automático e Semi-automático

As classificações supervisionadas foram efetuadas com os algoritmos de Máxima Verossimilhança (Maxver) e Máquina de Suporte Vetorial (*Support Vector Machine* - SVM) para as 11 cenas, considerando os 5 melhores atributos. Os resultados provenientes da utilização de limiarização, processados para cada cena no Capítulo 5: item 5.3, também foram utilizados com o objetivo de proporcionar uma maior comparação do método automático com diferentes métodos supervisionados.

A limiarização foi realizada com uma maior intervenção humana em relação aos demais métodos supervisionados. A segmentação por crescimento de regiões foi realizada automaticamente, entretanto, a definição do limiar ideal para o agrupamento dos alvos escuros foi definido interativamente avaliando, através de interpretação visual, qual foi o melhor valor de corte que retornou *clusters* mais próximos aos indicados nos *Phantons*. Para os métodos Maxver e SVM, a intervenção humana ocorreu na seleção das amostras.

Para realizar as classificações com os algoritmos Maxver e SVM, um arquivo de imagem contendo 5 bandas foi gerado no *software* ENVI, onde cada banda correspondia a espacialização de 1 dos 5 atributos padronizados nas regiões resultantes das segmentações (Mediana, Média, Moda, Máximo e GLCM da Média). Amostras de treinamento a serem utilizadas pelos classificadores testados foram selecionadas para representar o comportamento estatístico dos alvos escuros e do oceano. Como o objetivo deste item não era realizar uma análise aprofundada sobre os métodos supervisionados para detecção de alvos escuros, mas apenas fornecer exemplos para comparação com o método não supervisionado, não foram investigados os melhores parâmetros a serem utilizados em cada algoritmo. O Maxver e o SVM foram processados considerando as mesmas amostras de treinamento e os mesmos parâmetros de entrada no *software* ENVI.

O Maxver foi escolhido por ser considerado um dos métodos mais comumente utilizados em classificação supervisionada. Apesar deste classificador ser utilizado para processamentos no nível do *pixel*, nos experimentos efetuados o algoritmo considerou como unidade de processamento as regiões produzidas pelas segmentações. O Maxver considera como referência os parâmetros estatísticos das amostras de treinamento, fornece uma probabilidade de cada região pertencer a um alvo escuro ou ao oceano e pondera as distâncias entre as médias dos atributos nas regiões (ENVI). O único parâmetro exigido pelo *software* para o processamento é

denominado fator de escala, sendo definido em função da resolução radiométrica da imagem. Como os dados utilizados estavam em 8 bits e o fator de escala é definido como $2^n - 1$, onde n representa o intervalo de NC correspondente aos bits da imagem, o fator de escala utilizado foi 255.

A classificação com SVM foi escolhida por ser um procedimento considerado eficiente para prover bons resultados a partir de dados complexos e ruidosos, como é o caso das imagens SAR (ENVI). Pesquisas realizadas por BENTZ (2006) indicaram o SVM como método de classificação automática eficaz para a separação de óleo e assemelhados. MERCIER & ARDHUIN (2005 a e b) também utilizaram SVM para a detecção de óleo e assemelhados. A classificação por SVM está fundamentada na teoria de aprendizagem estatística (VAPNICK, 1998), buscando otimizar a separação do conjunto de dados em classes distintas a partir de uma superfície de decisão que maximiza a distância existente entre as classes. Esta superfície é denominada hiperplano e os dados próximos ao hiperplano são denominados vetores de suporte. Mesmo que esta superfície apresente um comportamento não linear, os dados de entrada são transformados, utilizando uma série de funções matemáticas (funções de núcleo) para obter um novo conjunto de dados que possa ser linearmente separável (BURGHES, 1998).

Os parâmetros a serem definidos no *software* para o SVM são: a) o tipo da função *Kernel*: definida como a função de núcleo radial (RBF), indicada por geralmente apresentar bons resultados e por ter oferecido a melhor performance para a separação de óleo e assemelhados nos estudos realizados por BENTZ (2006); b) o critério de penalização, que aceita níveis de imprecisão associados à classificação definidos com o maior rigor como 100, não aceitando níveis de imprecisão, e; c) o parâmetro *Gamma* a ser utilizado na função *Kernel*, definido como 0,2 por corresponder ao inverso no número de bandas utilizadas (1/5). A Figura 5.37 disponibiliza os resultados das classificações automáticas e semi-automáticas realizadas para cada recorte de cena.

RESULTADOS DAS CLUSTERIZAÇÕES AUTOMÁTICA E SEMI-AUTOMÁTICA PARA AS DIFERENTES CENAS					
MÉTODO NÃO SUPERVISIONADO	MÉTODOS SUPERVISIONADOS: 5 ATRIBUTOS				
CLUSTAN	LIMIARIZAÇÃO	MAXVER	SVM		
Cena 1: 06/07/01					
Cena 2: 06/07/02					
Cena 3; 06/03/06					
Cena 4: 09/09/01					
Cena 5: 13/06/08					

Figura 5.37 – Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na deteção de alvos escuros em imagens SAR.

RESULTADOS DAS CLUSTERIZAÇÕES AUTOMÁTICA E SEMI-AUTOMÁTICA PARA AS DIFERENTES CENAS				
MÉTODO NÃO SUPERVISIONADO	MÉTODOS SUPERVISIONADOS: 5 ATRIBUTOS			
	CLUSTAN	LIMIARIZAÇÃO	MAXVER	SVM
Cena 6: 05/05/02				
Cena 7: 19/12/01				
Cena 8: 03/04/02				
Cena 9: 04/07/01				
Cena 10: 11/03/05				

Figura 5.37 – Continuação: Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na detecção de alvos escuros em imagens SAR.

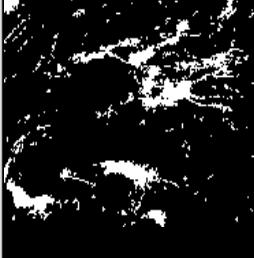
RESULTADOS DAS CLUSTERIZAÇÕES AUTOMÁTICA E SEMI-AUTOMÁTICA PARA AS DIFERENTES CENAS			
MÉTODO NÃO SUPERVISIONADO	MÉTODOS SUPERVISIONADOS: 5 ATRIBUTOS		
CLUSTAN	LIMIARIZAÇÃO	MAXVER	SVM
Cena 11: 14/06/08 			

Figura 5.37 – Continuação: Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na detecção de alvos escuros em imagens SAR.

A interpretação visual dos resultados indicou que, de maneira geral, os métodos semi-automáticos conseguiram retornar *clusters* mais representativos e, consequentemente, geometrias mais precisas para os alvos escuros, delineando com maior acurácia os limites que os separam do oceano. Este resultado era esperado à medida que estes métodos contam com o auxílio das amostras de treinamento no processo de reconhecimento de padrões.

Entretanto, é visualmente notável a boa performance automática do *Clustan* e dos atributos utilizados na identificação dos alvos escuros existentes nas cenas 2 e 6, sendo equivalente aos métodos supervisionados. Mesmo considerando uma menor precisão nas bordas, nas cenas 8 e 11 a performance automática também pôde ser considerada equivalente à dos métodos supervisionados, à medida que todos os métodos erraram incluindo ou omitindo regiões. Nas cenas 3 e 9 os semi-automáticos forneceram melhor resultado, entretanto, desconsiderando a confusão do clusterizador automático no canto dos recortes, os alvos escuros também foram detectados.

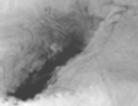
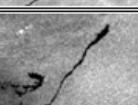
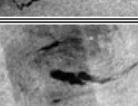
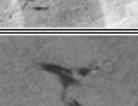
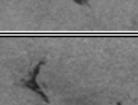
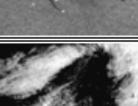
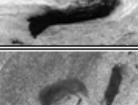
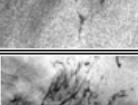
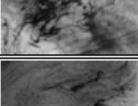
Para as cenas 4, 5 e 10 o método automático teve uma performance visualmente inferior aos semi-automáticos, detectou as manchas de óleo, mas apresentou ruído na região oceânica e nas bordas dos assemelhados. Na cena 1, todos os métodos apresentaram menor ruído no oceano, mas os alvos escuros foram detectados melhor pelos supervisionados Limiarização e Maxver.

A fim de subsidiar a interpretação dos resultados obtidos com dados quantitativos a Tabela 5.23 disponibiliza as métricas de avaliação calculadas para cada método utilizado e a Figura 5.38 apresenta a espacialização destas métricas.

Com exceção das cenas 2 e 6, contendo alvos escuros considerados fáceis de serem separados do oceano, todos os métodos - automático e supervisionados -

apresentaram erros de inclusão ou de omissão em maior ou menor escala. Estes erros foram perceptíveis visualmente nos recortes de cena em que foram considerados alvos muito filamentosos como os óleos biológicos (recortes 1, 10 e 11). Conforme comentado anteriormente para esse tipo de alvo a própria confecção manual dos *Phantons* possui uma maior imprecisão associada. Nestes casos, considerados complexos de serem solucionados pelos clusterizadores, observa-se que os diferentes métodos erraram incluindo ou omitindo regiões.

Tabela 5.23 - Métricas de avaliação calculadas para os resultados das clusterizações efetuadas pelos métodos automático e semi-automático

Cenas	Método Utilizado	Exatidão: AEscuros	Finters.: AEscuros	Omissão: AEscuros	Inclusão: AEscuros	Polig. Detectados
	Clustan	0,38	0,97	0,03	0,33	192
	Limiarização	0,61	0,79	0,21	0,29	131
	MaxVer	0,60	0,66	0,34	0,02	124
	SVM	0,51	0,92	0,08	0,17	169
	Clustan	0,80	0,88	0,12	0,00	4
	Limiarização	0,75	0,81	0,19	0,07	4
	MaxVer	0,51	0,88	0,12	0,01	4
	SVM	0,80	0,88	0,12	0,00	4
	Clustan	0,31	1,00	0,00	0,94	3
	Limiarização	0,50	0,52	0,48	0,04	3
	MaxVer	0,47	0,49	0,51	0,00	3
	SVM	0,52	0,54	0,46	0,00	3
	Clustan	0,17	0,73	0,27	0,22	6
	Limiarização	0,51	0,59	0,41	0,15	4
	MaxVer	0,38	0,71	0,29	0,05	5
	SVM	0,40	0,76	0,24	0,05	5
	Clustan	0,25	0,81	0,19	0,44	19
	Limiarização	0,61	0,76	0,24	0,24	11
	MaxVer	0,47	0,49	0,51	0,00	7
	SVM	0,63	0,80	0,20	0,03	14
	Clustan	0,31	0,61	0,39	0,01	3
	Limiarização	0,27	0,28	0,72	0,02	3
	MaxVer	0,44	0,61	0,39	0,01	3
	SVM	0,54	0,61	0,39	0,00	3
	Clustan	0,44	0,98	0,02	0,16	14
	Limiarização	0,86	0,90	0,10	0,05	8
	MaxVer	0,80	0,84	0,16	0,05	9
	SVM	0,85	0,92	0,08	0,08	14
	Clustan	0,56	0,98	0,02	0,19	33
	Limiarização	0,73	0,95	0,05	0,29	21
	MaxVer	0,68	0,70	0,30	0,01	3
	SVM	0,74	0,79	0,21	0,02	14
	Clustan	0,25	0,80	0,20	0,09	10
	Limiarização	0,67	0,74	0,26	0,10	9
	MaxVer	0,52	0,76	0,24	0,02	10
	SVM	0,76	0,86	0,14	0,01	9
	Clustan	0,18	0,97	0,03	0,58	92
	Limiarização	0,57	0,71	0,29	0,23	63
	MaxVer	0,46	0,72	0,28	0,06	53
	SVM	0,57	0,71	0,29	0,03	65
	Clustan	0,32	0,90	0,10	0,53	40
	Limiarização	0,32	0,40	0,60	0,23	31
	MaxVer	0,34	0,62	0,38	0,05	35
	SVM	0,38	0,54	0,46	0,03	26

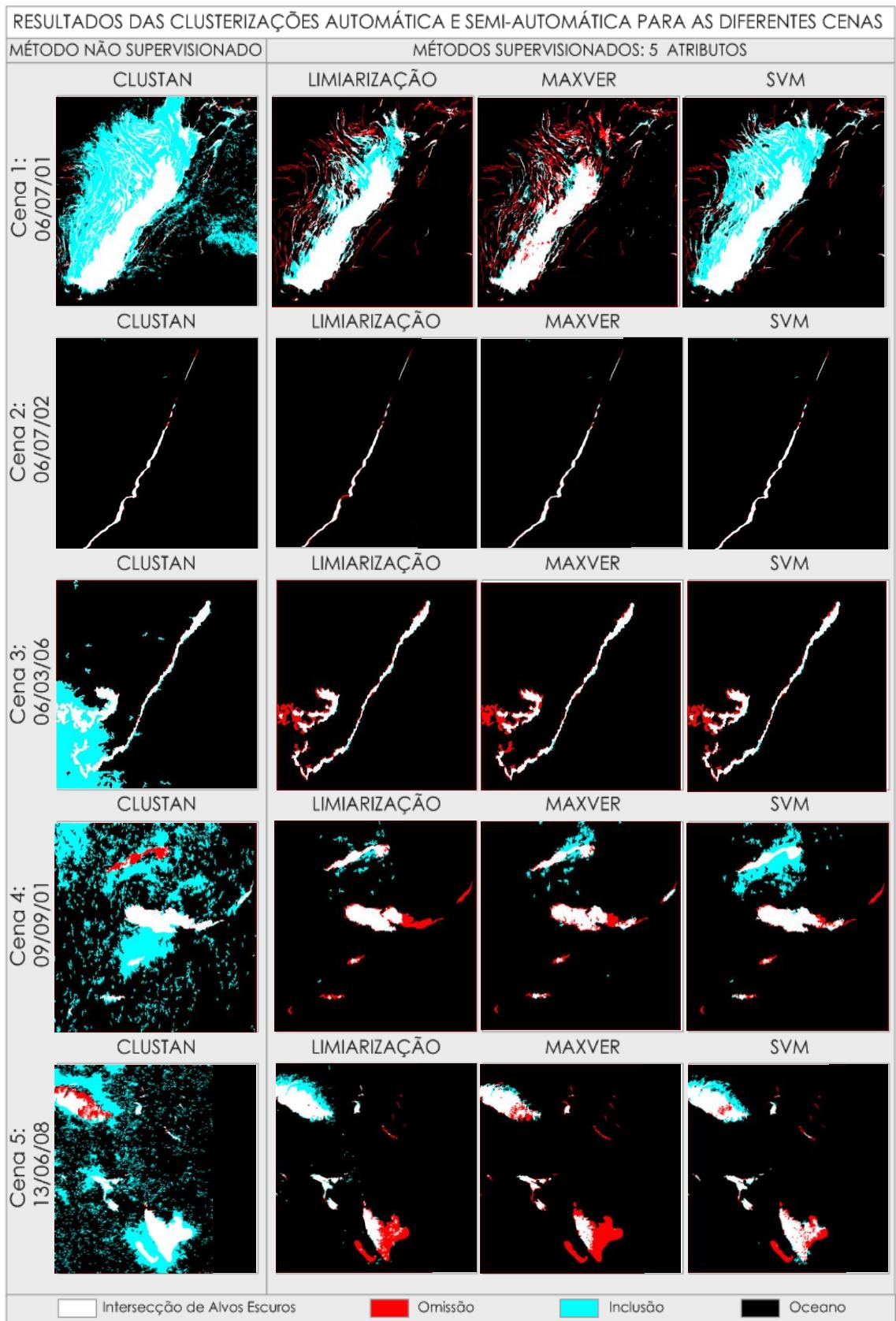


Figura 5.38 – Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na detecção de alvos escuros em imagens SAR.

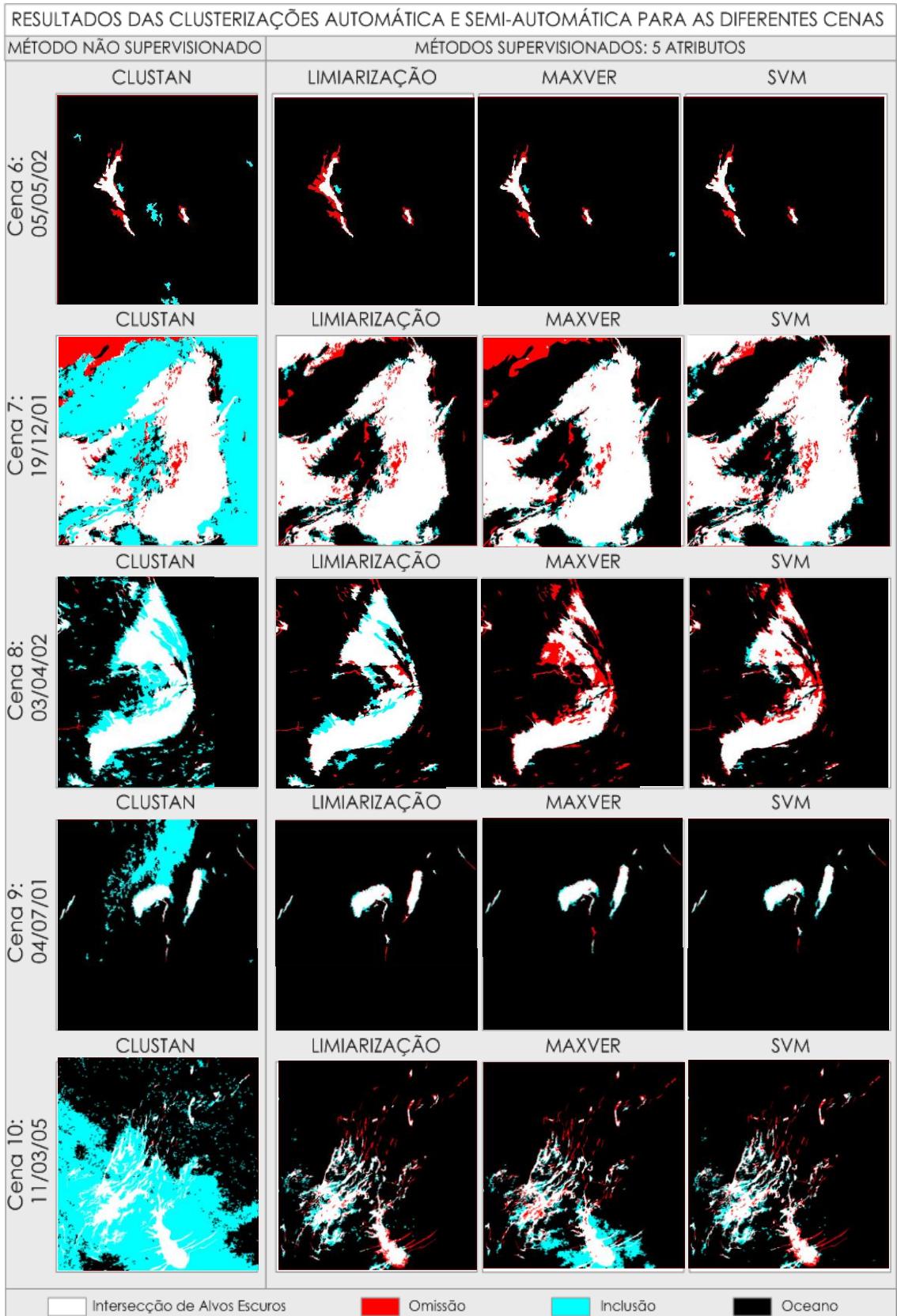


Figura 5.38 – Continuação: Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na detecção de alvos escuros em imagens SAR.

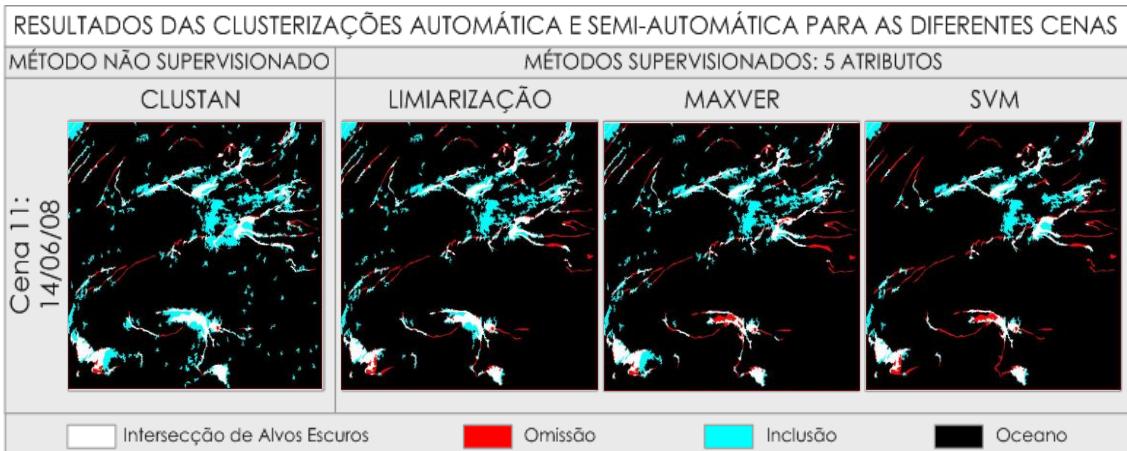


Figura 5.38 – Continuação: Resultados obtidos pela utilização de técnicas automáticas e semi-automáticas na detecção de alvos escuros em imagens SAR.

No caso do recorte de cena 1, apesar das melhores performances terem sido obtidas pelos métodos de Limiarização e Maxver, estes foram os que mais omitiram filamentos associados aos alvos escuros. O *Clustan* e o SVM erraram no sentido contrário, incluindo áreas de oceano como parte dos alvos escuros. De maneira geral, considerando a complexidade geométrica deste alvo o Maxver foi, visualmente (Figuras 5.37 e 5.38), o método que melhor conseguiu representar a estrutura filamentar.

No recorte de cena 10, as melhores *ExAE* foram obtidas pelos métodos de Limiarização e SVM. Os supervisionados omitiram mais áreas em relação ao *Phantom*, entretanto, o *Clustan* incluiu exageradamente regiões oceânicas como parte dos óleos biológicos. A presença de uma superfície menos rugosa, influenciada por uma baixa intensidade de vento ocorrendo simultaneamente no mesmo local dos óleos biológicos, prejudicou a definição dos *clusters*. Visualmente (Figuras 5.37 e 5.38) os melhores métodos foram a Limiarização e o SVM, com resultados similares.

De forma diferente das demais, a cena 11 pôde ser considerada bem clusterizada pelos 4 métodos avaliados (Figuras 5.37 e 5.38). O *Clustan* retornou a menor *ExAE*, por ter cometido mais erros de inclusão. Entretanto, todos os métodos cometeram erros em função da estrutura filamentar de baixo contraste em relação ao oceano, sendo que os supervisionados erraram mais omitindo regiões. A Figura 5.39 ilustra detalhes das cenas 1, 10 e 11, ampliando os trechos onde a densa malha de filamentos provenientes dos óleos biogênicos dificultou os processamentos. Os gráficos disponíveis na Figura 5.40 também auxiliam na interpretação quantitativa dos resultados.

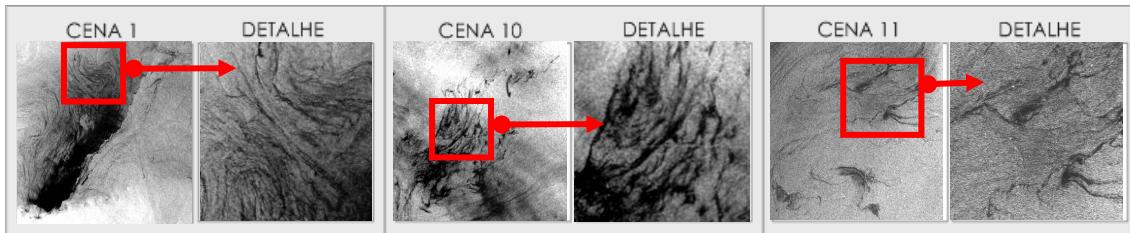


Figura 5.39 – Ampliação da malha de filamentos provenientes de óleos biogênicos existentes nas cenas 1, 10 e 11.

Erros exagerados de inclusão pelo software *Clustan* também foram observados nas cenas 4 e 5. A tendência obtida para a *ExAE* foi semelhante nas 2 cenas conforme indicado nos gráficos da Figura 5.40, sendo a Limiarização e o SVM os métodos que retornaram os melhores índices. Na cena 4 a descontinuidade na variação dos NC entre as faixas imageamento do RADAR prejudicou a separação automática das manchas de óleo do oceano. A utilização de amostras de treinamento possibilitou que os métodos supervisionados alcançassem melhor desempenho em relação ao *Clustan*. A Figura 5.41 disponibiliza detalhes da cena, ampliando os resultados das clusterizações na localização dos alvos escuros.

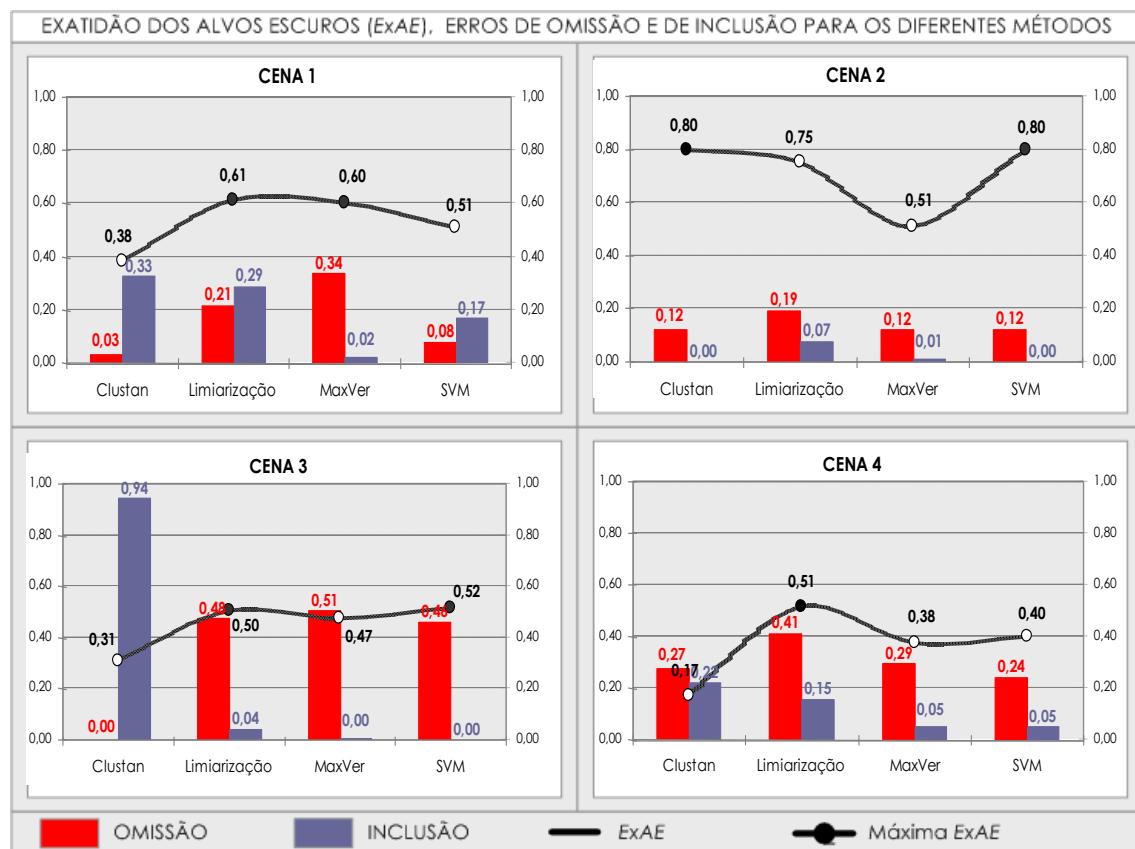


Figura 5.40 – Gráfico com os Erros de Omissão, de Inclusão e as *ExAE* obtidas para os métodos automático e semi-automáticos em cada recorte de cena.

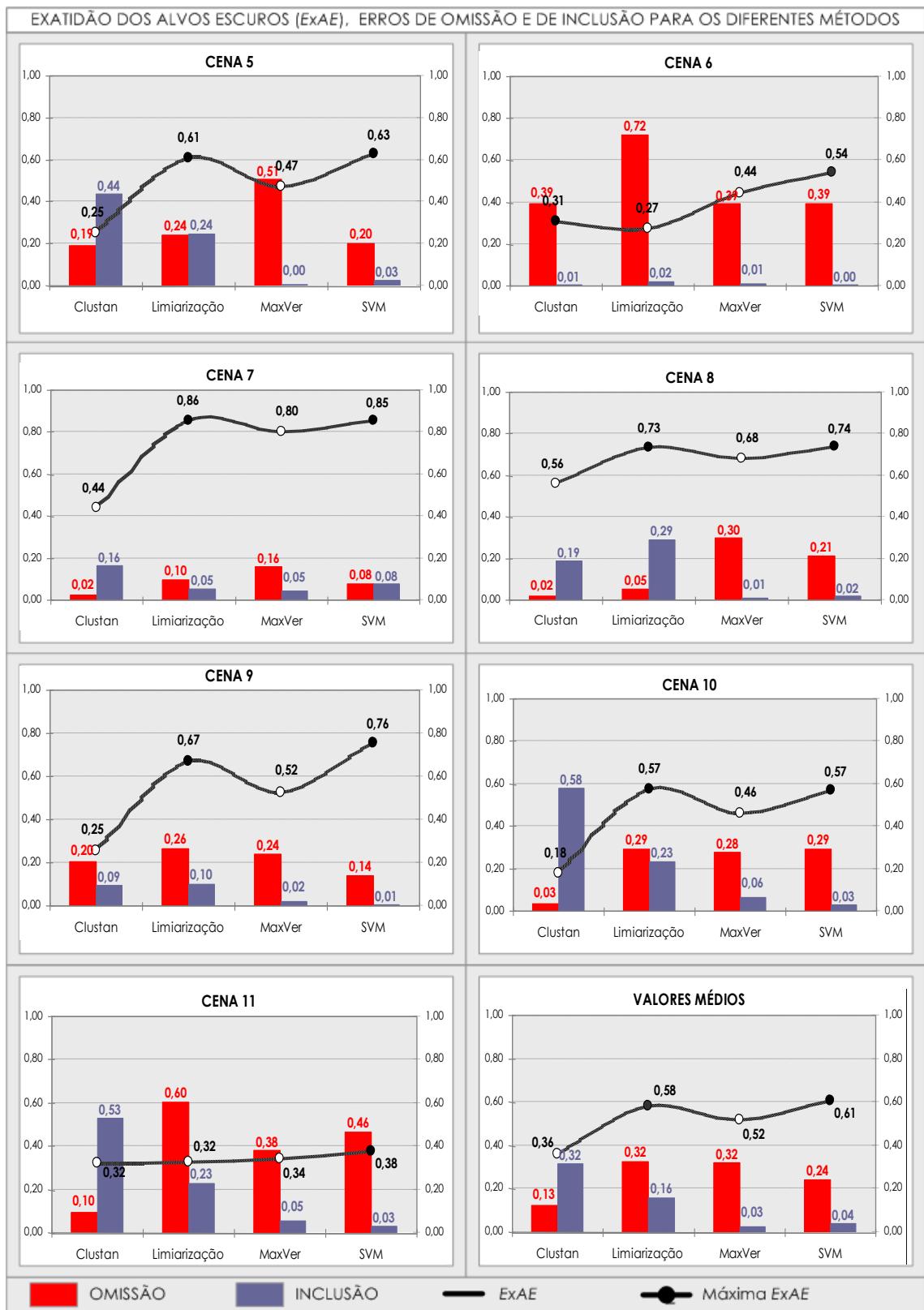


Figura 5.40 – Gráfico com os Erros de Omissão, de Inclusão e as *ExAE* obtidas para os métodos automático e semi-automáticos em cada recorte de cena.

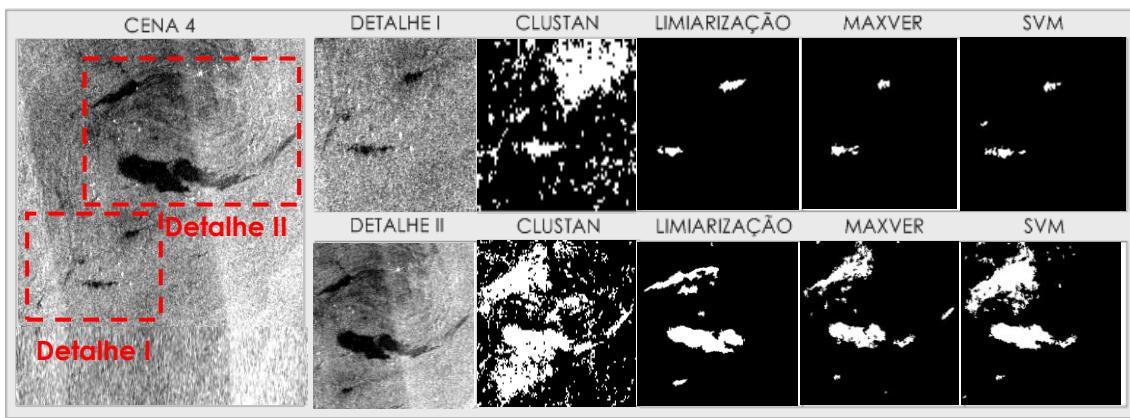


Figura 5.41 – Detalhes da cena 4, ampliando os resultados das clusterizações para os alvos escuros.

Na cena 5 foi perceptível a maior confusão e ruído nos resultados do *Clustan* para os alvos escuros relacionados a baixas intensidades de vento localizadas em pontos extremos do recorte de cena. A Figura 5.42 ilustra a clusterização destes alvos em detalhes para os 4 métodos avaliados, evidenciando que a geometria resultante varia muito entre eles. Esta diversidade de resultados reforça a dificuldade imposta aos clusterizadores por eventos de baixas intensidades de vento, os quais possuem bordas freqüentemente sem limites bem definidos em relação ao oceano.

As Figuras 5.43 e 5.44 ilustram detalhes da performance dos diferentes métodos para a detecção manchas de eventos de óleo com tamanho e contraste variados na cena 5. Foi perceptível que todos os métodos testados tiveram dificuldade na detecção das manchas menores com contraste menos intenso em relação ao oceano, errando incluindo (*Clustan*) ou excluindo regiões parcial ou completamente (Limiariação, Maxver e SVM).

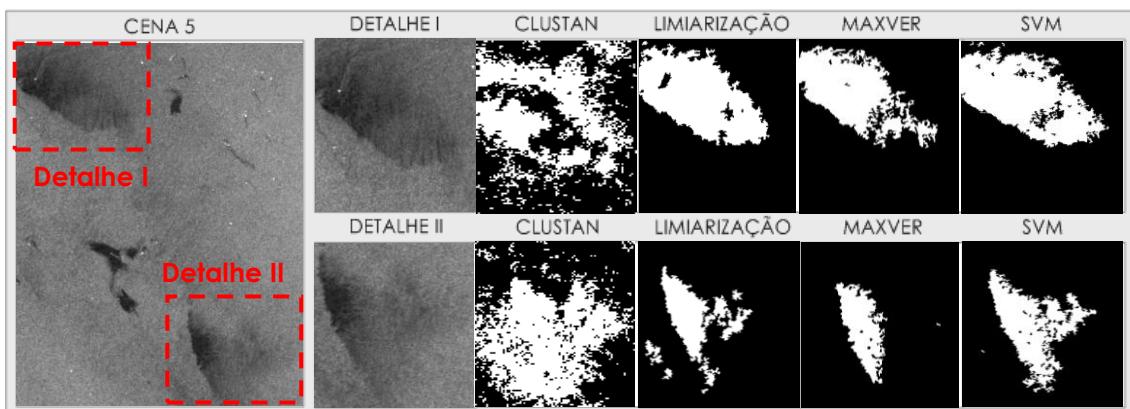


Figura 5.42 – Detalhes da cena 5, ampliando os resultados das clusterizações para os alvos escuros relacionados a baixas intensidades de vento.

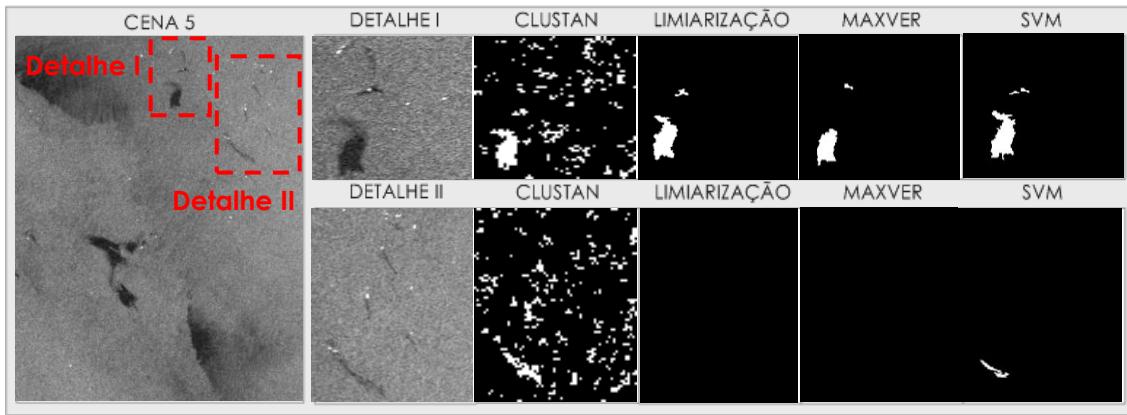


Figura 5.43 – Detalhes da cena 5, ampliando os resultados das clusterizações para os alvos escuros relacionados a óleo.

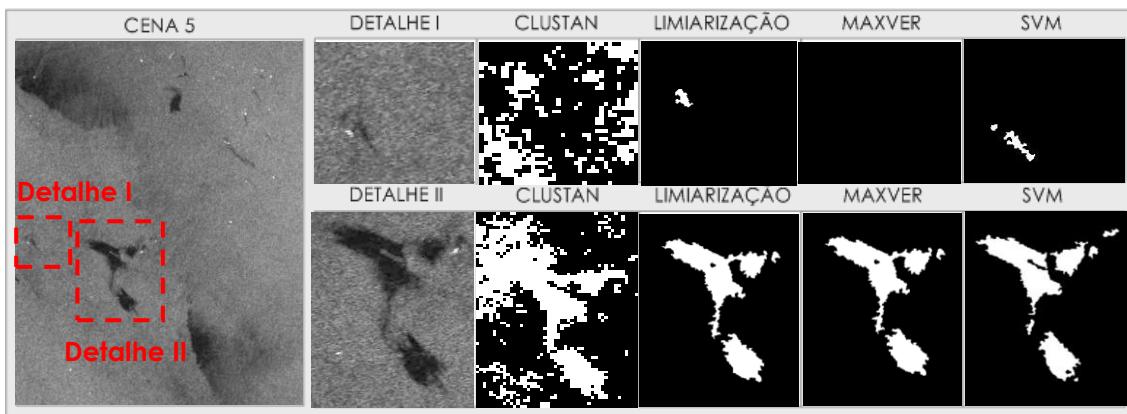


Figura 5.44 – Detalhes da cena 5, ampliando os resultados das clusterizações para os alvos escuros relacionados a óleo.

Considerando que a *ExAE* diminui com o aumento dos erros de inclusão e/ou de omissão, em 8 das 11 cenas estudadas, o *Clustan* teve um comportamento inverso do observado para os métodos supervisionados, cometendo mais erros de inclusão do que de omissão, conforme evidenciado no gráfico da Figura 5.40 (valores médios). Um exemplo interessante é o do recorte de cena 8, em que as melhores *ExAE* foram obtidas de forma semelhante para o *Clustan* e para o SVM, entretanto, os tipos de erros cometidos foram opostos.

Como, no contexto das imagens SAR, todo alvo escuro tem uma probabilidade de estar associado a óleo (KARATHANASSI *et al.*, 2006), sendo preferível errar aumentando as áreas estimadas como suspeita de óleo do que omitir um possível foco de vazamento que pode tornar-se uma emergência ambiental. Entretanto, embora estas geometrias imprecisas contribuam para o acionamento das inspeções de campo, no sistema automático prejudicariam a classificação. Num estágio posterior, uma auditoria no Banco de Dados seria necessária para indicar as geometrias que

precisariam ser refeitas a fim de evitar uma propagação de erros capaz de prejudicar a performance automática de sistemas baseados em conhecimento.

A média dos resultados (Figura 5.40) evidenciou que a melhor performance foi obtida pelo SVM seguida da Limiarização. A Limiarização foi comparada apenas para evidenciar como uma intervenção humana pode melhorar a qualidade dos resultados. Não seria viável automatizar este processo, pois os limiares foram definidos arbitrariamente, de forma interativa para cada cena, sendo impossível replicar operacionalmente a mesma eficácia dos resultados com novas aquisições. A técnica de Limiarização é freqüentemente utilizada na literatura (Capítulo 3), mas com uma abordagem automática, onde os limiares podem ser calculados de forma Global ou Local (Adaptativamente). Entretanto, inúmeros autores (KARANTZALOS & ARGIALAS, 2008; CHANG et al., 2008; TOPOUZELIS et al., 2007a; KARATHANASSI et al., 2006; SOLBER & SOLBERG, 1996) destacam a fragilidade destes métodos para agrupar *pixels* ou regiões (Capítulo: 3). A Figura 5.45 compara a performance do agrupamento realizado pelo *Clustan*, considerando as regiões da segmentação, e pelo *K-means* tradicional aplicado diretamente sobre os *pixels*, enfatizando o ganho associado aos processamentos efetuados a partir de regiões ao invés de *pixels*.

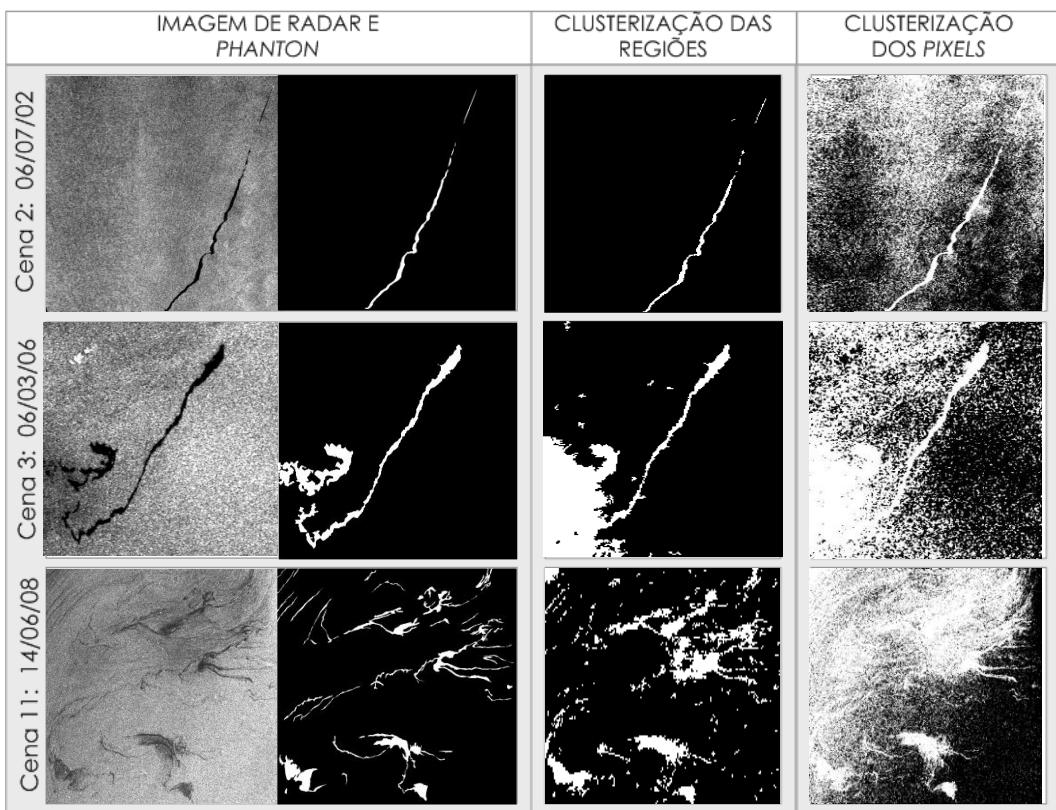


Figura 5.45 - Comparação entre o agrupamento realizado: a) pelo *Clustan*, considerando as regiões da segmentação, e; b) pelo *K-means* tradicional aplicado diretamente sobre os pixels.

Embora os resultados obtidos com a clusterização automática não tenham sido 100% eficazes, não seria justo afirmar que o procedimento não é passível de automatização. Uma breve comparação com os diferentes métodos e respectivos resultados disponíveis na literatura é interessante, buscando evidências do que os diversos autores consideraram como um bom resultado e um procedimento automático. Neste contexto, a Figura 5.46 disponibiliza as imagens originais e os resultados obtidos para a detecção de alvos escuros a partir de métodos considerados automáticos pelos referidos autores, comparando-os aos resultados obtidos na tese. Na figura, para cada autor foram identificados: a) se o mesmo trabalhou no nível do *pixel* ou das *regiões*; b) qual o tipo de segmentação ou classificação utilizada para separar os alvos escuros do oceano, e; c) se algum tipo de pós-processamento foi utilizado para agrupar os polígonos.

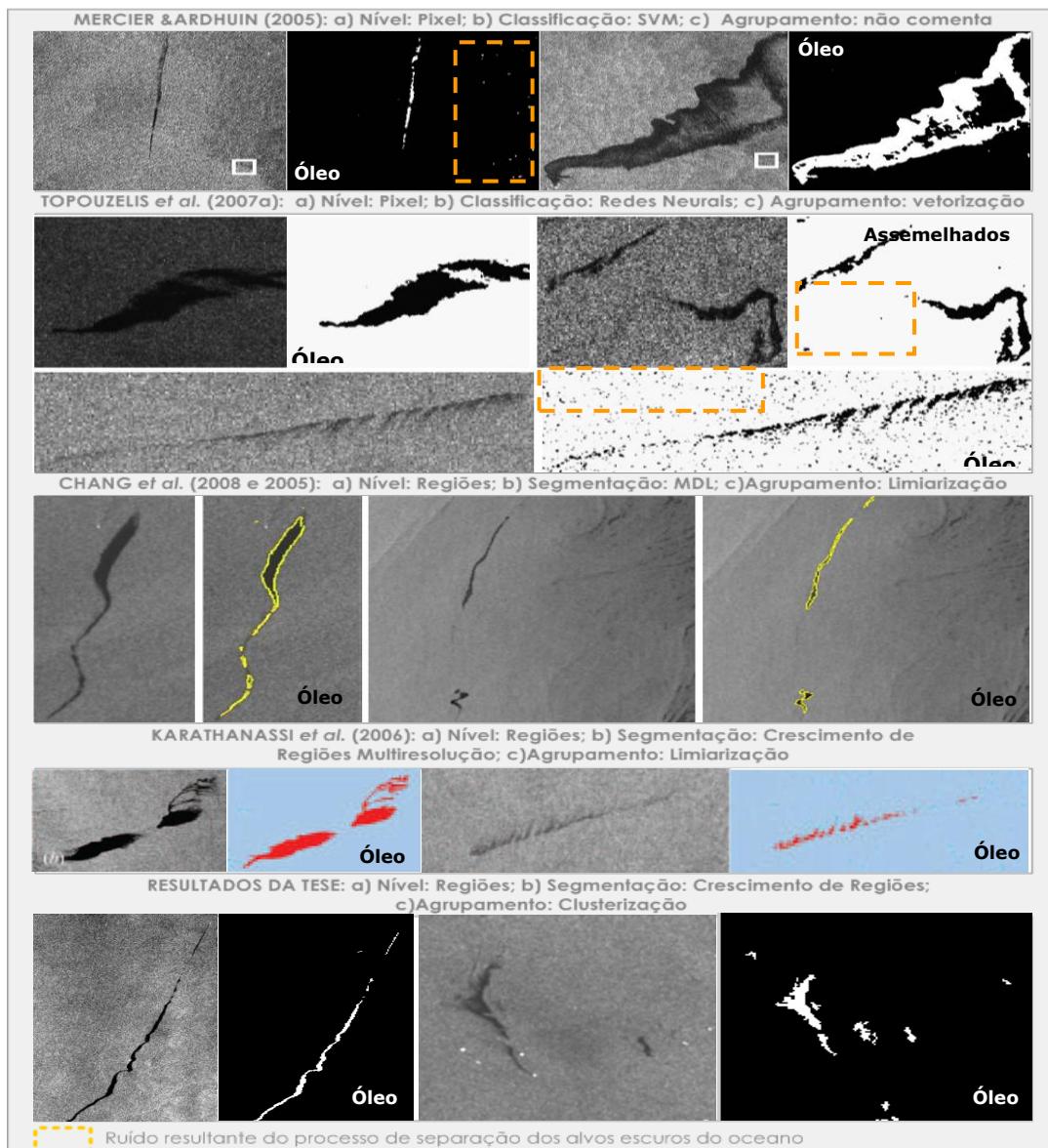


Figura 5.46 – Comparação entre resultados obtidos por diferentes métodos automáticos de detecção de alvos escuros disponíveis na literatura e os obtidos na pesquisa desenvolvida.

De acordo com os exemplos selecionados, os resultados obtidos por MERCIER & ARDHUIN (2005a) e TOPOUZELIS *et al.* (2007a) reforçaram as vantagens de se trabalhar com regiões ao invés de *pixels*. Apesar dos alvos escuros terem sido detectados o resultado é mais ruidoso na região oceânica (*background*), conforme indicado nas imagens. De forma contrária, os resultados produzidos nos estudos desenvolvidos na tese, por CHANG *et al.* (2005) e por KARATHANASSI *et al.* (2006), evidenciaram que o ruído, referente à indicação de áreas oceânicas como alvos escuros, é bem menor se os processamentos forem efetuados a partir de regiões e não de *pixels*. Outros autores como TOPOUZELIS *et al.* (2009), KERAMITSOGLOU *et al.* (2006) GALLAND *et al.* (2003 e 2004) também abordaram as vantagens de se trabalhar com *pixels* ao invés de regiões para a detecção de alvos escuros.

A maioria dos trabalhos que declararam suas metodologias como plenamente automáticas mostraram resultados considerando recortes de cena com alvos individuais isolados e geralmente com bom contraste em relação ao entorno, conforme exemplificado na Figura 5.46. A utilização de recortes de cena contendo mais de um alvo escuro em diferentes escalas é escassa na literatura, e de recortes de cena com ocorrências simultâneas de óleo e assemelhados mais escassos ainda. Isto é um ponto essencial a ser tratado, considerando que um sistema plenamente automático precisaria funcionar sem a intervenção humana, processando cenas inteiras lidando, não apenas com a diversidade e a escalabilidade dos alvos escuros, mas também com as variações dos NC que podem ocorrer em uma mesma cena.

Esta preocupação é recorrente na literatura onde, apesar de diversos autores (CHANG *et al.*, 2005 e 2008; TOPOUZELIS *et al.*, 2007a e 2004; KARATHANASSI *et al.*, 2006; MERCIER & ARDHUIN, 2005a e b) terem considerado seus métodos automáticos, indicaram a necessidade de efetuar mais testes com diferentes imagens, contendo diferentes tipos de alvos escuros com geometrias e texturas diferenciadas em diferentes condições meteo-oceanográficas. TOPOUZELIS *et al.* (2007a) ainda destacaram como perspectiva futura a necessidade de processar simultaneamente óleo e assemelhados em um mesmo recorte de cena.

Observando os exemplos disponíveis na Figura 5.46, considerados automáticos pelos respectivos autores, o método proposto na tese também poderia ser considerado automático. Entretanto, conforme comentado no item 5.4.3, esta conclusão é muito subjetiva e dependente dos tipos de alvos e da complexidade dos recortes de cena utilizados nos experimentos. A Figura 5.47 disponibiliza exemplos de metodologias que utilizaram recortes de cena, com alvos escuros considerados difíceis

de serem detectados automaticamente. Na figura, para cada autor foram identificados: a) se o mesmo trabalhou no nível do *pixel* ou das regiões; b) qual o tipo de segmentação ou classificação utilizada para separar os alvos escuros do oceano, e; c) se algum tipo de pós-processamento foi utilizado para agrupar os polígonos.

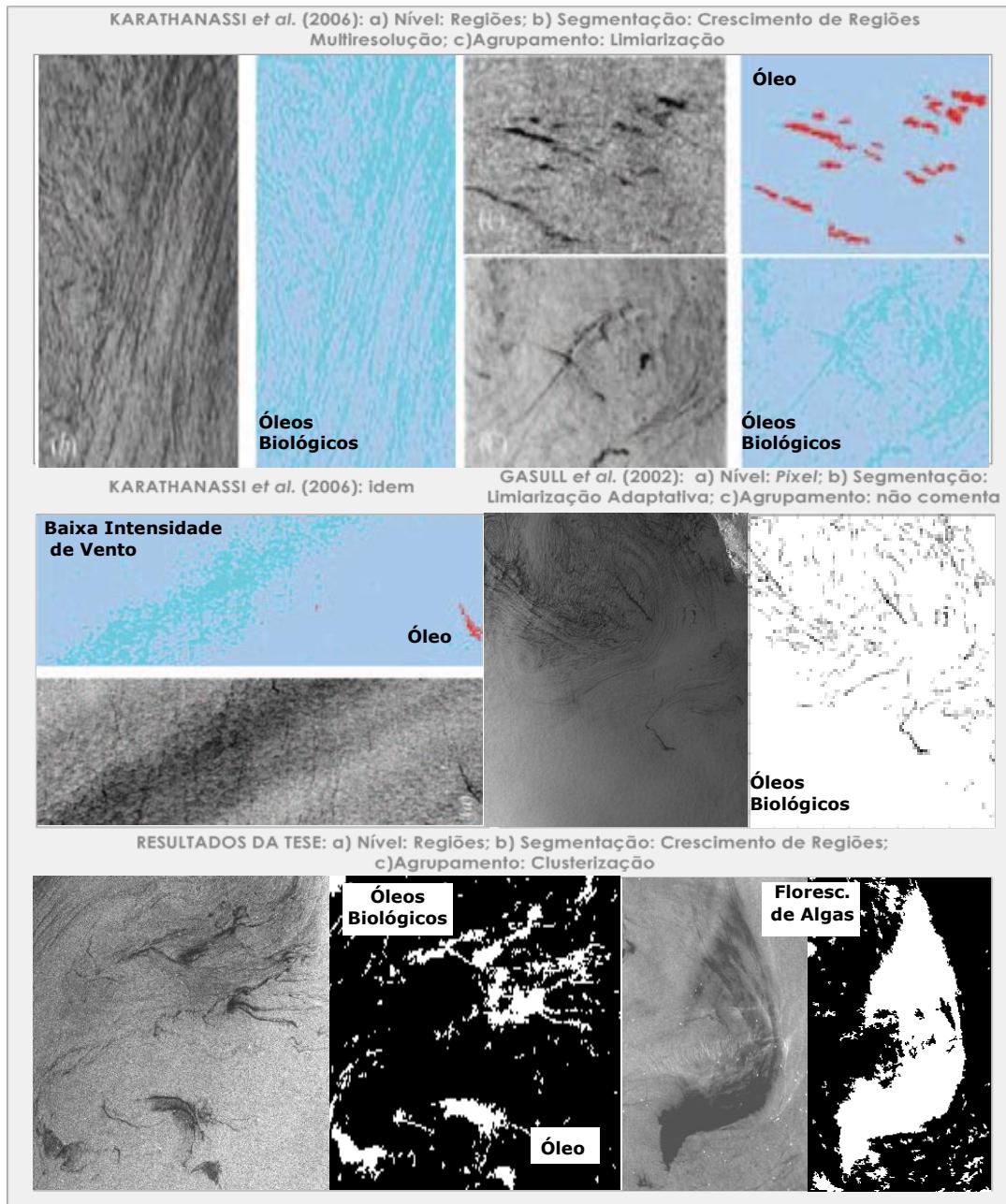


Figura 5.47 – Comparação entre resultados obtidos por diferentes métodos automáticos de detecção de alvos escuros disponíveis na literatura e os obtidos na pesquisa desenvolvida.

Os métodos propostos por KARATHANASSI *et al.* (2006) e por GASULL *et al.* (2002) detectaram alvos de geometria complexa difíceis de serem separados do oceano. Diferente dos demais autores KARATHANASSI analisou em uma mesma cena óleo e assemelhado, conseguindo um bom resultado na separação das regiões

com baixa intensidade de vento e com óleo do oceano. KARATHANASSI considerou seu método automático, mas ainda não operacional em função do custo computacional requerido, cerca de 2h. Uma fragilidade associada aos métodos propostos por KARATHANASSI e por GASULL é a utilização de limiarização para agrupar as regiões resultantes da segmentação em alvos escuros e oceano. Conforme comentado anteriormente, a limiarização automática pode não retornar bons resultados para todas as imagens, por ser sensível a variações nas condições meteo-oceanográficas predominantes na superfície do mar no momento da aquisição das imagens. De certa forma, todos os métodos sofrerão com a variabilidade dos intervalos de NC associados aos alvos escuros de cena para cena, por ser uma característica inerente aos dados SAR.

Os resultados obtidos automaticamente pelo *Clustan* nas cenas 8 e 11, contendo alvos igualmente difíceis de serem separados do oceano, também foram considerados satisfatórios. Na cena 11 foram detectados alvos escuros derivados da presença de assemelhados e de óleo. Considerando estes resultados o método proposto poderia ser considerado automático. Entretanto, conforme comentado anteriormente, os resultados obtidos pelo *Clustan* não foram aceitáveis para todas as complexidades de cenas e de tipos de alvos selecionados. Nestas circunstâncias, considerar que o método é automático e poderia ser reutilizado em qualquer nova imagem SAR, garantindo a mesma eficácia no reconhecimento de padrões associados a alvos escuros e oceano, não seria aceitável.

Especialmente nesta etapa da cadeia de processamentos, considerada a mais importante por definir a geometria dos alvos escuros que será utilizada em todas as etapas posteriores, tanto para o cálculo de atributos quanto para a classificação destes alvos em óleo e assemelhados, a acurácia é um fator a ser considerado. Uma alternativa seria criar um sistema híbrido onde a etapa de detecção dos alvos escuros seria feita com a intervenção humana, de forma supervisionada, para gerar geometrias mais precisas para a etapa de classificação automática. Este sistema poderia utilizar para a detecção de alvos escuros um classificador semi-automático como o SVM, a qual apresentou potencial nos testes efetuados.

ANGIULI et al. (2006), MONTALI et al (2006) e PELIZZARI & DIAS (2003) utilizaram métodos supervisionados para a detecção de alvos escuros em imagens SAR. Nestes sistemas a intervenção humana pode ocorrer em várias etapas do processo: a) indicando para o sistema quais são as áreas de interesse; b) definindo o limiar para separação dos alvos escuros do oceano interativamente, caso o cálculo

automático não tenha retornado resultados satisfatórios, e/ou; c) selecionando amostras de treinamento para a etapa de classificação em óleo e assemelhados.

Outra opção seria mesclar, já na etapa de detecção de alvos escuros, procedimentos automáticos e semi-automáticos, com o objetivo de aproveitar a performance automática do método proposto nas situações em que os resultados foram favoráveis. A sugestão seria criar uma regra de decisão a ser confrontada a cada nova imagem, onde seriam indicadas as cenas a serem processadas automaticamente e as cenas que necessitariam de supervisão. O item 5.5 buscou verificar a possibilidade de se utilizar uma regra de decisão válida para indicar o processamento automático ou semi-automático.

5.5 – Regra de Decisão e Validação do Modelo Proposto

Conforme evidenciado nos exemplos discutidos anteriormente, apesar da diversidade das pesquisas desenvolvidas e dos inúmeros métodos e algoritmos utilizados para a detecção de alvos escuros e classificação de óleo e assemelhados em imagens SAR (Capítulo 3 e APÊNDICE I), ainda não há um consenso sobre a utilização operacional de métodos plenamente automáticos. A proposta de elaboração de um sistema híbrido inteligente capaz de indicar, através de regras de decisão previamente estabelecidas, as imagens aptas para o processamento automático, seria uma alternativa interessante.

O método de Árvores de Classificação e Regressão - ACR (*Classification and Regression Trees* - CART), disponível no módulo *Data Mining* do Software *Statistica*, foi utilizado com o objetivo de estabelecer regras simples e compactas, capazes de indicar a aptidão das imagens para o processamento automático, considerando duas categorias: satisfatórias (S) e não satisfatórias (NS). As árvores são formadas a partir de uma coleção de regras estimadas com base no conjunto de variáveis preditoras. Estas regras são selecionadas com o objetivo de encontrar o número de partições e de valores para as variáveis preditoras, que melhor representem as classes indicadas pela variável dependente. O termo Árvores de Classificação é utilizado quando a variável dependente é categórica, sendo as Árvores de Regressão associadas à utilização de variáveis dependentes numéricas. O método utilizado é denominado Árvore de Classificação e Regressão por identificar automaticamente se as variáveis são categóricas ou numéricas, sendo capaz de efetuar os dois tipos de processamento (BREIMAN *et al.*, 1984).

As árvores são o resultado de partições recursivas onde: a) cada nó corresponde a uma das variáveis preditoras do conjunto de entrada; b) as ramificações estabelecem o critério de partição, considerando intervalos numéricos estabelecidos para os atributos de cada nó, e; c) as folhas contêm o número de indivíduos selecionados pela regra indicada, associados às classes estabelecidas pela variável dependente. Desta forma, as ACR constituem um conjunto de regras estabelecidas que se propagam da raiz em direção aos nós e folhas.

No modelo proposto, a variável dependente Y é categórica e está associada à performance do clusterizador, indicando se a mesma foi satisfatória (S) ou não (NS). O vetor X é composto pelo conjunto de 5 atributos de entrada utilizados como preditores, sendo $X = \{x_1, \dots, x_5\}$. Para buscar as regras de decisão através da ACR, 5 novos atributos foram calculados, considerando a distância entre os *clusters* formados. Estes atributos são resultantes da diferença entre as médias dos *clusters* referentes ao oceano e aos alvos escuros para cada atributo, sendo estes: a) $x_1 = \mu_{\text{Mediana Oceano}} - \mu_{\text{Mediana Alvos Escuros}}$; b) $x_2 = \mu_{\text{Média Oceano}} - \mu_{\text{Média Alvos Escuros}}$; c) $x_3 = \mu_{\text{Moda Oceano}} - \mu_{\text{Moda Alvos Escuros}}$; d) $x_4 = \mu_{\text{Máximo Oceano}} - \mu_{\text{Máximo Alvos Escuros}}$; e; e) $x_5 = \mu_{\text{GLCM Média Oceano}} - \mu_{\text{GLCM Média Alvos Escuros}}$. Os valores médios para cada atributo em cada *cluster* são calculados automaticamente pelo software *Clustan* ao final de cada clusterização. A Tabela 5.24 contém os atributos preditores, juntamente com a variável dependente que se refere a performance do clusterizador em cada recorte de cena.

Tabela 5.24 – Variável dependente e atributos preditores utilizados na ACR para as 11 cenas.

Recortes de Cena	$x_1 = \text{Dif. Mediana}$	$x_2 = \text{Dif. Media}$	$x_3 = \text{Dif. Moda}$	$x_4 = \text{Dif. Maximo}$	$x_5 = \text{Dif. GLCM Média}$	Performance do Clusterizador
Cena 1	0,23	0,20	0,21	0,12	-0,23	NS
Cena 2	0,33	0,35	0,31	0,22	-0,42	S
Cena 3	0,41	0,32	0,45	0,23	-0,34	NS
Cena 4	0,10	0,10	0,09	0,10	-0,40	NS
Cena 5	0,08	0,09	0,06	0,09	-0,43	NS
Cena 6	0,22	0,31	0,28	-0,10	0,43	S
Cena 7	-0,11	-0,12	-0,11	-0,15	0,37	NS
Cena 8	0,28	0,22	0,28	0,15	-0,27	S
Cena 9	0,29	0,16	0,27	0,14	-0,38	NS
Cena 10	0,20	0,15	0,25	0,11	-0,42	NS
Cena 11	0,36	0,32	0,35	0,14	-0,40	S

Os 11 recortes de cena estudados ao longo da tese foram utilizados na ACR para definir as regras. Com o objetivo de validar não apenas as regras formuladas pela ACR, mas todos os parâmetros padronizados ao longo da tese, 3 novas imagens foram escolhidas. A Figura 5.48 disponibiliza a metodologia utilizada para a criação e validação da regra de decisão para definir o processamento como automático ou semi-automático.

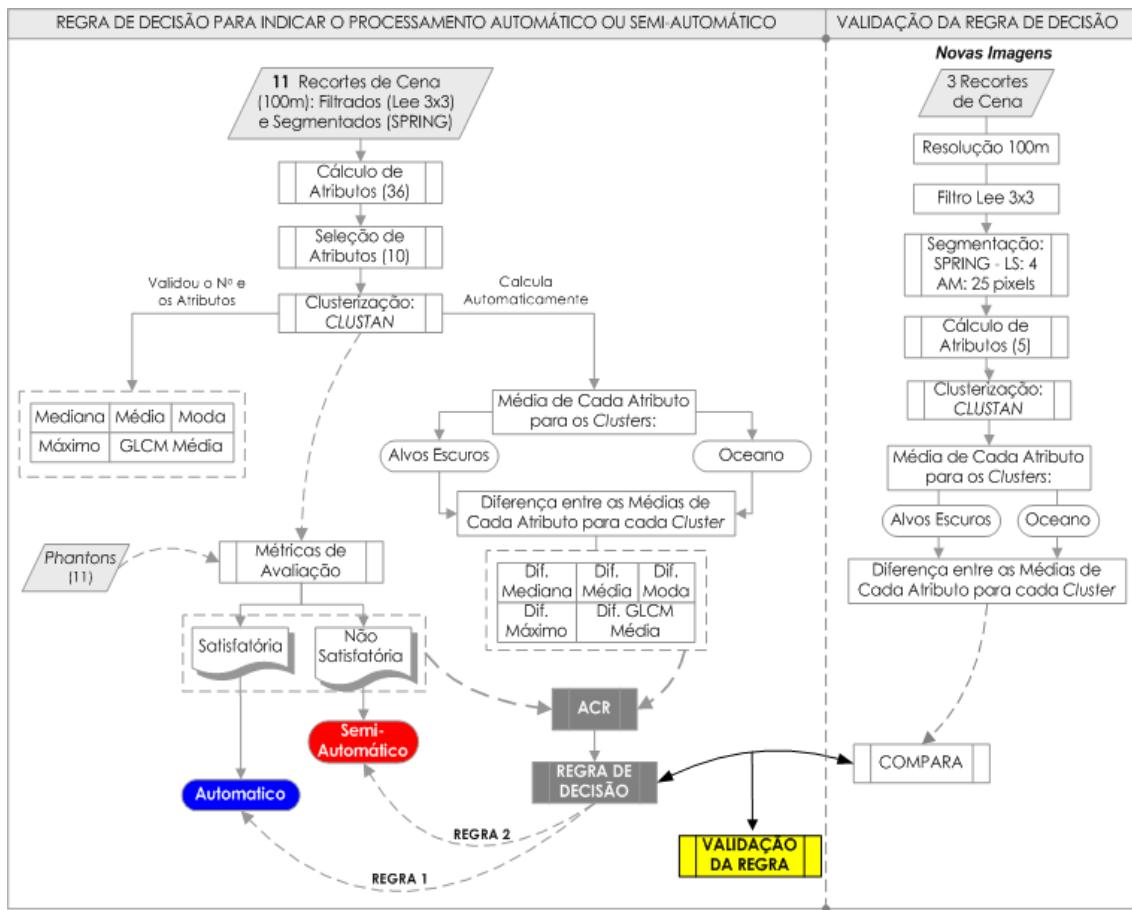


Figura 5.48 – Método para criação e validação da regra de decisão para definir o processamento como automático ou semi-automático.

A ACR retornou uma regra simples e compacta, considerando a diferença para a Moda como atributo de partição inicial associado ao nó raiz. A partir do nó raiz uma estrutura com 2 partições, resultando em 4 folhas e 3 nós terminais foi obtida. Na folha raiz o tamanho do nó é igual ao número de componentes de entrada, sendo igual ao número de cenas (11). A primeira regra de partição, baseada no atributo de Moda, associou 6 componentes ao ramo esquerdo e 5 ao ramo direito, sendo classificados respectivamente como NS e S. A segunda regra de partição, baseada no atributo de Máximo, subdividiu os 5 componentes restantes em NS (1) e S (4).

A ACR retornou um conjunto de regras simples e compactas que refletiu os resultados obtidos pelas clusterizações, indicando que apenas 4 cenas foram satisfatórias. A Figura 5.49 disponibiliza um fluxograma de decisão com as regras obtidas pela ACR a serem aplicadas às novas imagens com objetivo de indicar a aptidão das mesmas ao processamento automático ou semi-automático.

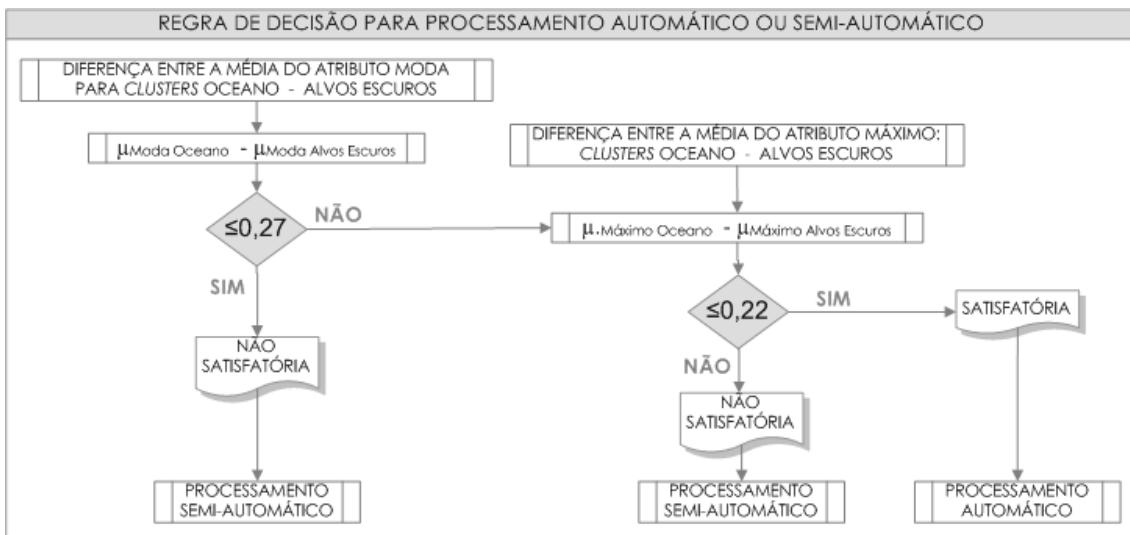


Figura 5.49 – Regra de decisão obtida pela ACR para definir o processamento como automático ou semi-automático.

Para a validação do método proposto e da regra de decisão fornecida pela ACR, 3 novas cenas RADARSAT foram recortadas contendo: a) recorte de cena A: alvo escuro derivado de vazamento de óleo, com bordas bem definidas e com bom contraste em relação ao oceano, e; b) recortes de cena B e C: alvos escuros em diferentes escalas, derivados da presença de assemelhados, sem bordas bem definidas e com maior heterogeneidade associada à variação dos NC no oceano, ocasionando um menor contraste em relação ao oceano para alguns alvos. Quanto mais evidente for o comportamento bimodal de um histograma, mais fácil será para o algoritmo identificar padrões que sejam eficazes para separar os alvos escuros do oceano. Nos recortes de cena definidos é perceptível a diminuição gradativa do comportamento bimodal dos histogramas, esta escolha teve como objetivo submeter o clusterizador a níveis crescentes de dificuldade. A Figura 5.50 ilustra os 3 novos recortes de cena escolhidos para validar o método proposto e a regra de decisão fornecida pela ACR, juntamente com seus respectivos histogramas.

Cada recorte de cena foi pré-processado com espaçamento de *pixel* de 100m, filtro Lee com janela 3x3, segmentado no software SPRING com área mínima de 25 pixels e limiar de similaridade 4, sendo este a média dos melhores limiares indicados no item 5.3 para cada uma das 11 cenas estudadas. Tendo como base as regiões claras e escuras resultantes da segmentação, o conjunto adequado de 5 atributos, indicado pelo processo de seleção e validado pelas análises de clusterização, foi calculado para os 3 novos recortes de cena. As clusterizações foram efetuadas automaticamente no software *Clustan* e, seqüencialmente, as métricas de avaliação foram calculadas tendo como referência os *Phantons* delineados manualmente para

os novos exemplos. A Tabela 5.25 disponibiliza as métricas de avaliação calculadas para as clusterizações automáticas efetuadas nos 3 novos recortes de cena para validação do método de detecção automática de alvos escuros proposto na tese, e a Figura 5.51 apresenta a espacialização destas métricas.

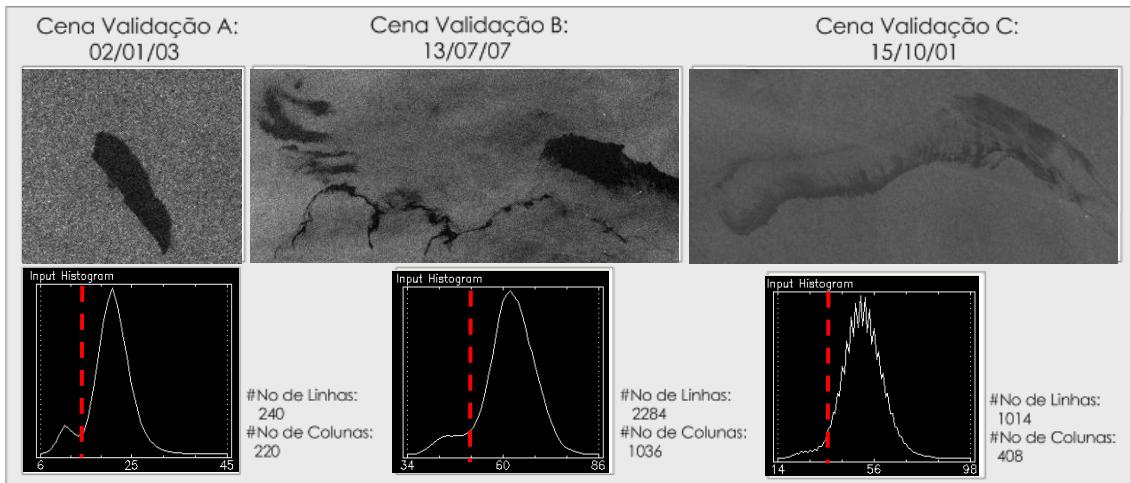


Figura 5.50 – Recortes de Cena escolhidos para validação do método proposto para a detecção automática de alvos escuros em imagens SAR.

Tabela 5.25 - Métricas de avaliação calculadas para os resultados das clusterizações efetuadas nos 3 novos recortes de cena para validação do método proposto na tese.

Cenas p/ Validação	Método Utilizado	<i>ExAE</i>	F.Interse- ção AE	Erro de Omissão	Erro de Inclusão	Nº de Alvos Escuros
Cena A	Clustan	0,90	0,98	0,02	0,01	1
Cena B	Clustan	0,46	0,93	0,07	0,10	23
Cena C	Clustan	0,49	0,88	0,12	0,08	21

A performance automática obtida pelo clusterizador no recorte de cena A foi considerada satisfatória, tendo evidenciado erros mínimos de omissão e de inclusão, atingindo uma *ExAE* de 0,89, resultado esperado em função do comportamento bimodal indicado no histograma da imagem. A clusterização para o recorte de cena B foi mais ruidosa em função da heterogeneidade dos NC no oceano associada a presença do ruído *speckle* e a redução localizada da intensidade do vento que ocorreu simultaneamente com os assemelhados. Neste exemplo, os erros de inclusão foram bem mais acentuados, e a *ExAE* foi de 0,46, performance considerada não satisfatória.

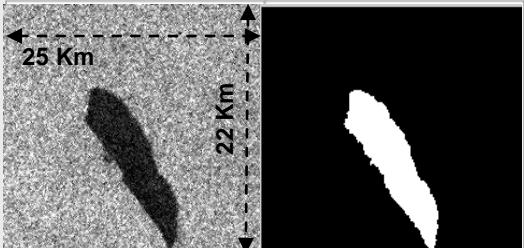
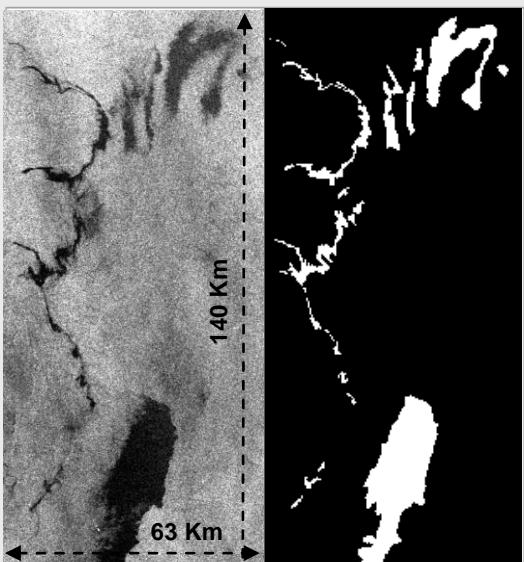
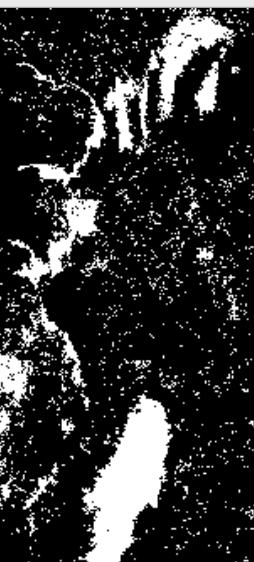
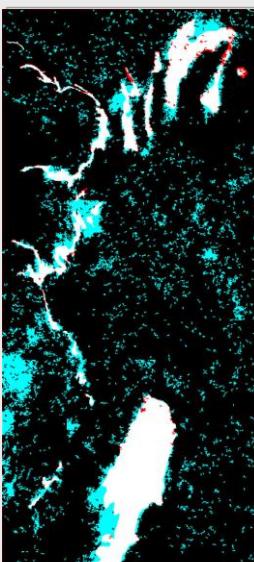
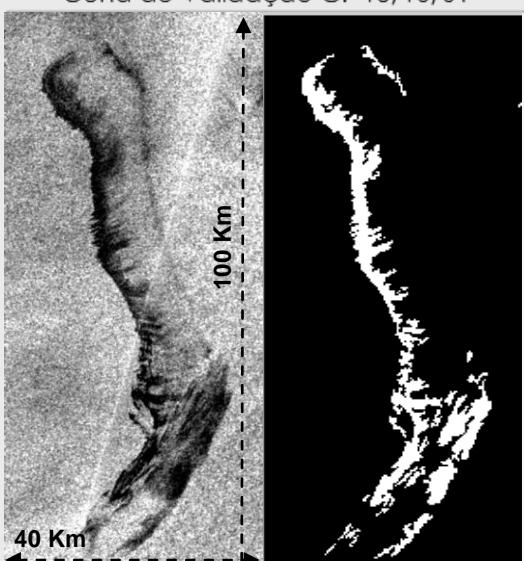
IMAGENS DE RADAR E PHANTONS PARA VALIDAÇÃO	CLUSTERIZAÇÃO: 5 ATRIBUTOS	EXATIDÃO E ERROS DE INCLUSÃO E OMISSÃO
Cena de Validação A: 02/01/03 	ExAE = 0.89 	Clusterização Satisfatória 
Cena de Validação B: 13/07/07 	ExAE = 0,46 	Clusterização Não Satisfatória 
Cena de Validação C: 15/10/01 	ExAE = 0,49 	Clusterização Satisfatória 

Figura 5.51 – Resultados obtidos para validação do método de detecção automática de alvos escuros proposto na tese.

Embora esta aparência ruidosa presente na região oceânica do recorte de cena B seja similar ao ruído produzido pelas clusterizações efetuadas no nível do pixel, esta associação é apenas visual e ocorre em função dos diferentes tamanhos de cena e

escala dos alvos escuros. Em cenas muito ruidosas, dotadas de maior heterogeneidade no oceano, a segmentação pode gerar regiões com o tamanho mínimo de até 25×25 pixels, parâmetro de área mínima definido no item 5.3 para ser utilizado nos segmentadores. A visualização da imagem em uma menor escala evidencia que não se tratam de pixels, mas de regiões delimitadas pelo clusterizador, uma clusterização efetuada no nível do pixel produziria resultados muito mais ruidosos. A Figura 5.52 disponibiliza: a) um mapa com as dimensões do recorte de cena 2; b) uma ampliação com área equivalente à do recorte de cena 1 para fins comparativos (detalhe 1), e; c) ampliações sucessivas (detalhes 2 e 3) evidenciando os pixels no interior das regiões escuras delimitadas pela segmentação.

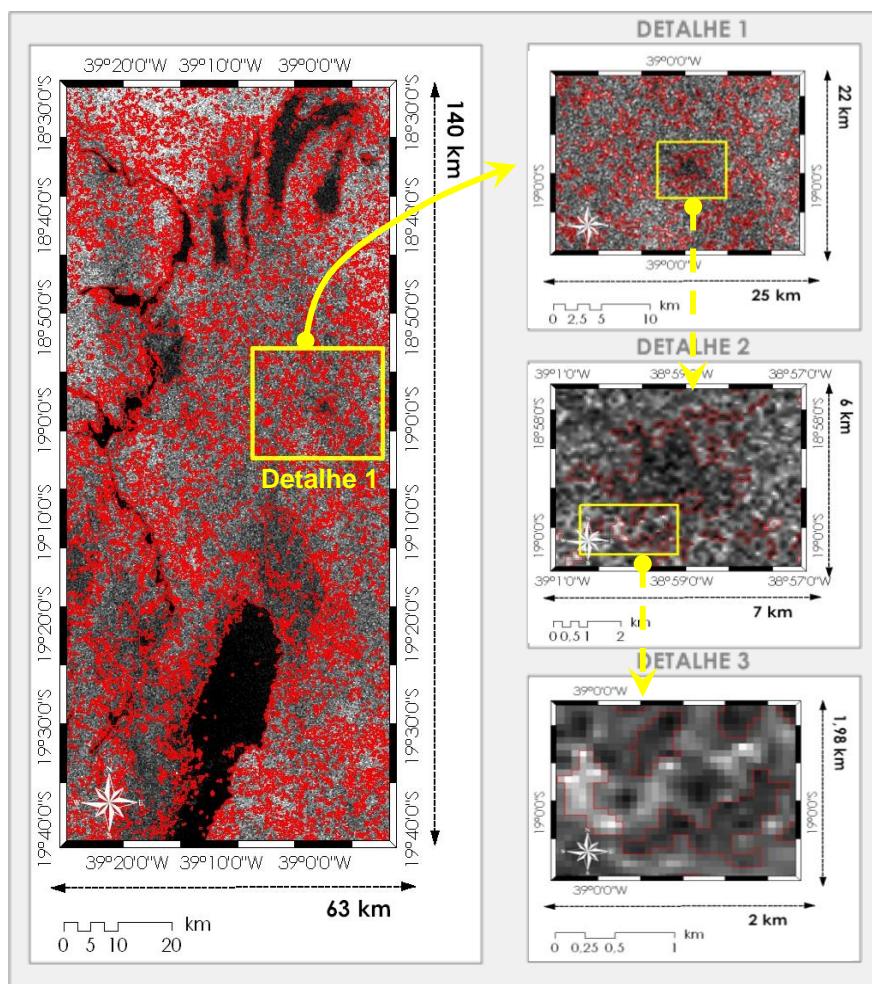


Figura 5.52 – Dimensões do recorte de cena B, evidenciando o conjunto de pixels nas regiões escuras delimitadas pela segmentação.

No recorte de cena C, onde era esperado que o clusterizador tivesse mais dificuldade para convergir *clusters* representativos, houve menos erro de inclusão no oceano. Este erro ficou concentrado ao redor do alvo, comportamento esperado uma vez que as bordas que separam o alvo do oceano não são bem definidas, especialmente no caso das imagens SAR. A ExAE (0,49) atingiu quase 50% de

representatividade em relação ao *Phantom* o que indicou a performance do clusterizador como satisfatória.

Para verificar se a regra de decisão formulada pela ACR seria válida para indicar as imagens aptas a serem processadas automaticamente, os seguintes atributos foram calculados: a) Média do atributo Moda para os *clusters* no oceano menos a Média do atributo Moda para os *clusters* dos alvos escuros, e; b) Média do atributo Máximo para os *clusters* no oceano menos a Média do atributo Máximo para os *clusters* dos alvos escuros. A Tabela 5.26 disponibiliza a diferença entre as médias dos *clusters* para os 5 atributos utilizados na clusterização, a *ExAE* e a classificação da performance do clusterizador como satisfatória (S) ou não satisfatória (NS).

Tabela 5.26 - Métricas de avaliação calculadas para os resultados das clusterizações automáticas efetuadas nos 3 novos recortes de cena para validação do método proposto.

Cenas de Validação	Dif. Mediana	Dif. Media	Dif. Moda	Dif. Máximo	Dif. GLCM Média	ExAE	Performance do Clusterizador
Cena A	0,91	0,88	0,79	0,45	0,07	0,89	Satisfatória
Cena B	0,13	0,13	0,26	0,10	-0,39	0,46	Não Satisfatória
Cena C	0,35	0,31	0,35	0,18	-0,50	0,49	Satisfatória

A Figura 5.53 apresenta os processamentos efetuados com os 3 novos recortes de cena, juntamente com os resultados da validação da regra de decisão indicada pela ACR, para definir o processamento das imagens como automático ou semi-automático. Ao final dos processamentos as geometrias produzidas com ou sem a intervenção humana são utilizadas como referência para o cálculo dos atributos em um sistema automático de classificação em óleo e assemelhados.

A regra de decisão (Figura 5.49) não foi validada com todos os exemplos utilizados. De acordo com a regra o processamento semi-automático seria indicado para os recortes de cena A e B. A regra foi válida para o recorte de cena B que obteve a menor *ExAE* e o maior erro de Inclusão (Tabela 5.25 e Figura 5.51). Entretanto, de forma contraditória, o procedimento supervisionado foi indicado para o recorte de cena A que apresentou a melhor performance automática. A cena C foi coerentemente indicada para o processamento automático, uma vez que os *clusters* obtidos foram representativos com *ExAE* de 0,49.

Desta forma, somente os recortes de cena B e C validaram a regra de decisão proposta pela ACR. Reprocessar supervisionadamente uma cena que teve resultados satisfatórios no modo automático significaria um custo computacional desnecessário. Entretanto, do ponto de vista operacional seria melhor do que indicar para o

processamento automático uma cena que não é favorável, produzindo resultados dotados de reduzida acurácia geométrica.

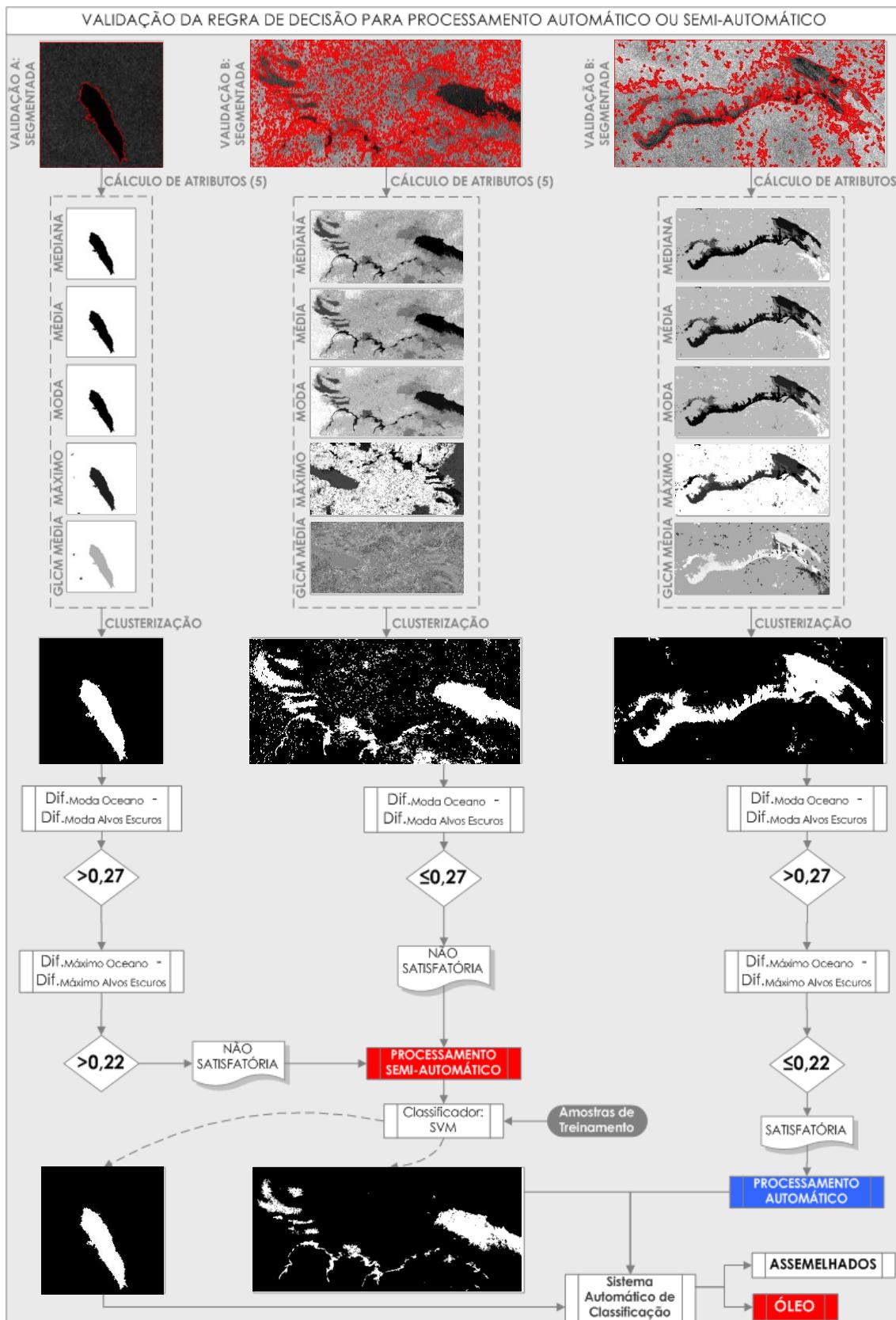


Figura 5.53 – Validação da regra de decisão para definir o processamento como automático ou semi-automático.

As 3 novas cenas processadas com os parâmetros padronizados ao longo da tese validaram as hipóteses relacionadas à possibilidade de automatização do método a partir de um conjunto padrão com número reduzido de atributos.

A utilização de um limiar único para a segmentação também foi validada, entretanto, neste quesito mais exemplos precisariam ser testados. Em recortes de cena menores com apenas um alvo, o limiar indicado resolve bem a separação dos alvos escuros do oceano. Nas cenas maiores, em que se têm alvos escuros de grande e pequena dimensão, o LS indicado pode super-segmentar alvos escuros de maior de dimensão, juntamente com o oceano (*background*), aumentando o erro de inclusão nas áreas oceânicas (exemplo da cena de validação C: Figura 5.51).

Os resultados evidenciaram o potencial da utilização de regras de decisão para auxiliar no processo de automatização, indicando situações em que a intervenção humana é necessária para produzir geometrias representativas para os alvos escuros. Entretanto, seria indicado aumentar o número de amostras para que a ACR possa retornar uma regra mais robusta, a ser aplicada a toda nova imagem SAR adquirida.

CAPÍTULO 6

CONCLUSÕES E RECOMENDAÇÕES

6 - Premissas e Hipóteses Adotadas no Desenvolvimento da Tese:

Os resultados obtidos e as discussões abordadas ao longo da tese forneceram contribuições significativas no âmbito dos sistemas automáticos para detecção de alvos escuros em imagens SAR. Estes resultados foram obtidos considerando a premissa de que o método proposto tivesse características que permitissem a sua futura implementação de forma automática e operacional, sendo investigadas a seguintes hipóteses:

- i) utilizar o formato compactado *MrSid* permitiria um rápido *download* e um processamento mais rápido por ser disponibilizado já com os pré-processamentos efetuados;
- ii) utilizar um espaçamento de *pixel* maior do que 50m reduziria o custo computacional;
- iii) utilizar filtros no pré-processamento das imagens amenizaria a influência do ruído *speckle* e, consequentemente, melhoraria a detecção de alvos escuros compensando a redução do espaçamento de *pixel*;
- iv) segmentar as imagens para efetuar os processamentos no nível das regiões e não dos *pixels*, produziria resultados menos ruidosos com um menor custo computacional;
- v) associar inúmeros atributos às regiões ampliaria as chances de encontrar uma combinação de atributos considerada adequada para agrupar as regiões em alvos escuros e oceano;
- vi) efetuar o agrupamento das regiões através de clusterização de dados, considerando que a segmentação não fornece um resultado definitivo em função da escalabilidade dos alvos, produziria geometrias mais representativas para os alvos escuros;
- vii) selecionar um subconjunto de atributos mais significativos melhoraria o desempenho da clusterização produzindo resultados mais satisfatórios;
- viii) efetuar uma seleção rigorosa dos recortes de cena a serem utilizados como teste permitiria obter resultados mais realistas, indicando a possibilidade de se automatizar ou não o sistema detecção de alvos escuros;
- ix) utilizar imagens de referência, denominadas *Phantoms*, permitiria uma justa avaliação dos resultados indicando quais os melhores softwares e parâmetros a serem utilizados no um sistema de detecção de alvos escuros;
- x) a acurácia obtida para a geometria dos alvos escuros é essencial para alimentar a etapa de classificação dos alvos nas diferentes classes de óleo e assemelhados;
- xi) reduzir a intervenção humana através da padronização dos parâmetros envolvidos na detecção, bem como utilizar um método para o agrupamento automático das regiões capaz de garantir a replicabilidade dos resultados, permitiria uma futura implementação automática e operacional do sistema;

Considerando a premissa e as hipóteses formuladas os seguintes resultados foram obtidos para as etapas investigadas ao longo da tese.

6.1 - Pré-processamento das imagens:

Espaçamento de Pixel:

- A avaliação da influência da utilização de diferentes espaçamentos de *pixel* (50, 100, 150 e 200m) no potencial de detecção de alvos escuros em imagens SAR, confirmou a hipótese (*ii*), indicando que: (a) o espaçamento de 100m, apesar de detectar um menor número de alvos escuros (7) quando comparado ao espaçamento de 50m (14), reduziu em 7 vezes o tempo de processamento, o que viabilizaria o uso operacional de cenas completas. (b) Esta escolha também foi influenciada pelas métricas de avaliação, as quais indicaram que, dentre os espaçamentos (100, 150 e 200m), o de 100m foi o que conseguiu definir melhor as bordas dos alvos escuros e detectar um maior número de pequenos alvos com um menor tempo de processamento, quando comparado com o espaçamento de 50m. (c) Embora a maioria dos autores utilize o espaçamento de 100m, não foram encontrados estudos avaliando comparativamente o custo/benefício de cada espaçamento de *pixel* no desempenho da detecção, o que torna interessante os resultados obtidos. (d) Outra evidência que reforçou a utilização do espaçamento de *pixel* de 100m foi a aquisição de imagens RADARSAT também no modo *ScanSAR Wide* (100m), sendo aconselhável, no âmbito da automatização do sistema padronizar o espaçamento de *pixel* das imagens de entrada.

Filtragem:

- A avaliação do potencial associado à detecção de alvos escuros foi realizada considerando imagens filtradas com os filtros Lee, Lee Enhanced, Frost, Frost Enhanced, Kuan, Gamma, Sigma e Mediana com 3 dimensões de janela (3x3, 5x5 e 7x7). Dentre os filtros avaliados o Lee, aplicado com janela 3x3, foi indicado como mais adequado para o pré-processamento das imagens, sendo também o mais utilizado na literatura. De acordo com as métricas de avaliação, apesar da aplicação deste filtro ter reduzido a *ExAE* e aumentado os erros de inclusão em relação ao recorte de cena sem filtragem (100m), aumentou o número de alvos escuros detectados de 7 para 10, confirmando a hipótese (*iii*). Esta avaliação foi importante para indicar a possibilidade de automatização, considerando a importância dos filtros para a suavização do ruído *speckle* inerente às imagens SAR.

- Uma avaliação mais robusta sobre a padronização ou não do filtro Lee 3x3 na etapa de pré-processamento foi realizada utilizando 11 cenas diferentes, segmentadas no MultiSeg e no SPRING (22 processamentos). Foi comprovado estatisticamente pelo teste de Wilcox que a aplicação de filtragem nas imagens interfere significativamente na detecção dos alvos escuros, melhorando os resultados na maioria das cenas (14). Com base nestas constatações, a utilização do filtro Lee aplicado com janela de dimensões 3x3, foi padronizada na etapa pré-processamento das imagens.

6.2 – Segmentação em Imagens SAR:

- Os segmentadores MultiSeg e SPRING foram aplicados em 11 cenas diferentes, com e sem filtragem, contendo diferentes tipos de alvos escuros em diferentes escalas. O resultado esperado era que, a segmentação de alvos escuros fosse melhor realizada pelo MultiSeg. Entretanto, embora o teste de Wilcox tenha indicado que não há diferença estaticamente significativa entre a qualidade dos resultados produzidos pelo SPRING ou pelo MultiSeg, o SPRING detectou um maior número de alvos escuros, tendo sido utilizado nas demais etapas da tese. Provavelmente, a alteração das propriedades estatísticas das imagens ocasionada pela compactação/descompactação dos dados a partir do formato *MrSid*, prejudicou o desempenho do MultiSeg, uma vez que as imagens foram processadas no modo específico para RADAR.

- A proposta de um método padrão para o cálculo do Limiar de Similaridade a partir de um *box plot* contendo as propriedades estatísticas dos alvos escuros e do oceano foi uma importante contribuição. Este método garantiu uma avaliação justa dos resultados obtidos, indicando o procedimento mais eficaz, considerando o melhor desempenho de cada um dos softwares testados. Outro aspecto interessante a ser comentado, refere-se à utilização do conhecimento dos especialistas para definir parâmetros coerentes para a automatização do método proposto, sendo o parâmetro de área mínima em *pixels* um exemplo. Este parâmetro foi estabelecido conjuntamente com os especialistas no combate à emergências ambientais por vazamento de óleo da PETROBRAS, os quais indicaram a mínima área de óleo identificável em campo e nas imagens de satélite.

- O método proposto para o cálculo do LS não sub-segmentou os recortes de cena, permitindo a detecção dos alvos escuros dotados de menor dimensão. A definição dos parâmetros de AM e LS adequados para serem aplicados em qualquer

imagem SAR adquirida, contribuíram para reduzir o nível de intervenção humana no processo, tornando factível uma futura automatização do sistema. Operacionalmente, a detecção de alvos escuros e oceano a partir de regiões e não de *pixels*, diminuiu a influência do ruído speckle e o custo computacional, além de permitir a associação de diferentes atributos às geometrias produzidas (hipóteses *iv* e *v*).

6.3 – Detecção dos Alvos Escuros por Reconhecimento de Padrões:

- O método utilizado para o cálculo do LS permitiu a detecção de alvos de pequena dimensão, entretanto, como os alvos de maior dimensão acabaram sendo particionados em um maior número de segmentos, o processo de agrupamento das regiões foi considerado indispensável para produzir geometrias mais representativas para os alvos escuros (hipótese *vi*).

- Neste contexto, considerando o estado da arte em sistemas para detecção de alvos escuros, os métodos propostos estão divididos em 2 etapas: a) segmentação das imagens, e; b) agrupamento dos *pixels* ou das regiões em alvos escuros e oceano. Pode-se considerar que o teor de inovação contido na tese esteve concentrado nesta etapa, sendo o método proposto dividido em 4 partes: a) segmentação das imagens; b) cálculo de diversos atributos para cada região originada pela segmentação; c) seleção do conjunto adequado de atributos, e; d) clusterização destes atributos em alvos escuros e oceano. Nos demais métodos o cálculo de atributos foi utilizado somente na etapa de classificação das imagens em óleo e assemelhados, bem como a utilização de técnicas para a seleção de atributos é escassa na literatura específica de detecção de óleo.

Análise Exploratória e Seleção de Atributos:

- O potencial associado aos 36 atributos calculados (8 estatísticos, 20 geométricos e 8 texturais) foi avaliado através dos procedimentos de análise exploratória e seleção de atributos efetuados sempre em relação ao atributo categórico Classe do *Phantom*. A avaliação dos atributos em relação a uma verdade previamente conhecida foi eficaz para identificar o conjunto de dados mais adequado para a clusterização.

- Os coeficientes de correlação de *Spearman* foram calculados para os atributos estatísticos, geométricos e texturais para cada uma das 11 cenas estudadas e utilizados para identificar os atributos mais relacionados ao *Phantom*, sendo estes os que teriam, provavelmente, o maior potencial para melhorar os resultados das

clusterizações. Os resultados evidenciaram que nenhuma das medidas de correlação obtidas (396) atingiu (r) acima de 0,7. Os atributos estatísticos foram os mais relacionados com o *Phantom*, seguidos dos texturas e dos geométricos com média global dos (r s de 0,37 (estatísticos), 0,20 (texturais) e 0,14 (geométricos). Apesar dos baixos valores de (r) obtidos para os três tipos de atributos, a maior parte das correlações obtidas foram consideradas estatisticamente significativas em relação ao *Phantom*, sendo 89% para os atributos estatísticos; 76% para os geométricos/poligonais; 72% para os geométricos/*skeletons*, e; 81% para os atributos texturais.

- A análise de componentes principais (ACP) indicou que conjunto mínimo de atributos necessário para manter 90% da representatividade do conjunto de dados inicial variou entre 4 e 10, sendo em média necessários 8 componentes.

- O método de seleção de atributos baseado no teste de chi-quadrado auxiliou no processo de redução da dimensionalidade do conjunto de dados, fornecendo um *ranking* com a importância dos 36 atributos, considerando a significância estatística (p valor) calculada para cada valor de chi-quadrado. Tendo como base os resultados da ACP, foram considerados na análise de seleção somente os 10 primeiros atributos mais importantes indicados pelo chi-quadrado em relação ao atributo Classe do *Phantom* em cada cena.

- Considerando que, para atender a premissa de automatização do sistema, é necessário padronizar os parâmetros envolvidos em cada etapa do fluxo de processamentos (hipótese x_1), um *ranking* considerando a freqüência de indicação acumulada dos 10 primeiros atributos para cada cena foi efetuado como objetivo de identificar os atributos recorrentes sendo estes, respectivamente: Mediana, Média, Moda, Mínimo, Máximo, GLCM da Média, Ajuste Retangular, GLCM da Homogeneidade, Assimetria dos Polígonos e Densidade. De forma coerente, os atributos indicados como menos correlacionados em relação *Phantom*, os geométricos e os texturais, foram os menos indicados entre os 10 primeiros do *ranking*.

- Com a espacialização dos 10 melhores atributos selecionados, foi interessante notar que, embora os atributos geométricos não tenham sido indicados como bons preditores em relação ao atributo Classe do *Phantom*, nas cenas em que manchas de óleo com geometrias bem definidas e alongadas estavam presentes, estes atributos realçaram estes alvos em relação ao oceano. Visualmente, o GLCM da Homogeneidade contribuiu para realçar os alvos em quase todas as cenas e o GLCM

da Média apresentou um comportamento irregular, sendo confuso em algumas cenas. Através da espacialização é notável que os atributos estatísticos de primeira ordem são, efetivamente, os que apresentaram maior potencial para auxiliar na detecção dos alvos escuros.

- Os atributos selecionados pelo chi-quadrado como mais importantes são multicorrelacionados entre si, entretanto, é notável que estes atributos foram os que apresentaram as melhores correlações com o *Phantom*, sendo os demais atributos muito pouco relacionados.

- A redução de 26 atributos do conjunto inicial de dados também representou uma contribuição para a operacionalização do sistema, reduzindo o custo computacional envolvido. A seleção realizada não foi considerada definitiva e foi utilizada para direcionar as inúmeras clusterizações efetuadas a partir da combinação dos atributos indicados (hipótese vii).

Clusterização das Regiões em Alvos Escuros e Oceano:

- Com o objetivo de validar o método de seleção utilizado, as clusterizações foram efetuadas inserindo os 10 melhores atributos um a um, na ordem indicada pela análise de freqüência acumulada (Tabela 5.19).

- Os resultados evidenciaram que a utilização do atributo de Mínimo isolado ou em conjunto com o atributo de Máximo piorou a detecção de alvos escuros, tendo sido indicada a sua eliminação do conjunto selecionado pelo chi-quadrado.

- Excluindo o atributo de mínimo, as clusterizações foram efetuadas considerando os 9 melhores atributos selecionados na ordem indicada na Tabela 5.27 e os 35 atributos na ordem indicada pelo chi-quadrado. Este estudo foi realizado, para verificar se haveria diferença na performance do clusterizador considerando o total de atributos e não apenas os 9 melhores do ranking. Para efetuar essa comparação foram consideradas: a) *ExAE* obtidas pela clusterização simultânea dos 9 e dos 35, e; b) a Máxima *ExAE* (*MExAE*) obtida considerando a inserção gradual dos 9 e dos 35 atributos.

- A média das *ExAEs* obtida com a clusterização simultânea de todos os atributos foi maior considerando os 9 atributos (0,37) do que considerando os 35 atributos (0,19), sendo esta diferença indicada como estatisticamente significativa pelo teste Wilcox com p valor de 0,028. Estes resultados validam a hipótese de que a seleção de um conjunto reduzido de atributos considerados significativos, é melhor do

que a utilização de um amplo conjunto de dados que inclua atributos não representativos.

- Os Máximos valores de exatidão *MExAE* obtidos considerando a inserção gradativa dos 9 e dos 35 atributos, apresentaram comportamento semelhante quando comparados cena a cena. A média das *MExAE* obtidas com a inserção gradativa de 9 atributos foi de 0,43, enquanto que a obtida com a inserção de 35 atributos foi de 0,44. Com um *p* valor de 0,068 o teste de Wilcox indicou que não há diferença estatisticamente significativa entre os valores de *MExAE* obtidos com a melhor combinação dentre os 9 ou dentre os 35 atributos.

- Considerando as hipóteses indicadas, esse resultado foi importante por validar a possibilidade de automatização do sistema de forma operacional a partir de um conjunto reduzido de atributos, minimizando também o custo computacional associado.

- Nas situações em que foi possível convergir para uma melhor *ExAE* com um menor número de atributos, oscilando entre 2 e 4, a inserção de mais atributos não prejudicou os resultados da clusterização. Entretanto, nos casos em que a máxima *ExAE* foi obtida com um número maior ou igual a 4 atributos, a redução do número de atributos para 2 ou 3 prejudicou a detecção dos alvos. Estes resultados indicaram que o conjunto de atributos a ser padronizado não poderia conter menos do que 4 atributos. O número de atributos a ser adotado poderia ser 4 ou 5, considerando que as *MExAE* foram obtidas na Média com 4 atributos e na Mediana com 5 atributos,. O teste Wilcox indicou que não há diferença estatisticamente significativa em se utilizar 4 ou 5 atributos. Neste contexto decidiu-se padronizar em 5 o número de atributos, considerando que a retirada do 5º atributo da cena 11 prejudicou significativamente os resultados e nas demais cenas não provocou alterações significativas.

- Importante destacar que embora as 5 variáveis indicadas sejam multicorrelacionadas, a sua utilização simultânea na clusterização foi o que gerou os melhores resultados, validando esta seleção. Estes atributos foram considerados robustos por terem sido selecionados de forma recorrente em todas as cenas. Testes foram efetuados considerando outros métodos de seleção e o resultado foi o mesmo, a clusterização só foi satisfatória para o conjunto de dados que continha as variáveis multicorrelacionadas.

- Apesar da automatização considerando um número reduzido de atributos padrão ter sido indicada como factível, a análise global dos resultados evidenciou que,

dentre as 11 cenas, apenas 5 obtiveram uma clusterização considerada satisfatória, sendo 4 com *ExAE* maiores ou iguais a 0,5 e 1 cena com *ExAE* menor do 0,5.

- Inúmeros fatores contribuíram para que os resultados não fossem considerados satisfatórios em todas as cenas sendo estes, direta ou indiretamente relacionados com o rigor adotado durante a seleção das cenas. A diversidade de cenas considerando diferentes tipos de alvos, com diferentes geometrias em diferentes escalas, ocorrendo isoladamente ou simultaneamente com outros eventos foi um ponto importante da pesquisa, enriquecendo os resultados e mostrando a dificuldade de se detectar alvos escuros automaticamente (hipótese *viii*). Um ponto a ser investigado é a influência do formato *MrSid* nesses resultados, considerando a alteração das propriedades estatísticas das imagens em função da compactação.

6.4 - Abordagem Automática e Semi-Automática:

- Considerando que ainda não há um consenso entre os pesquisadores sobre qual abordagem, manual, semi-automática ou automática, representa o melhor custo/benefício para operacionalizar a detecção de óleo e assemelhados em imagens SAR, uma comparação entre os métodos automático e semi-automático foi realizada para definir com maior rigor o procedimento mais eficaz. Foram utilizados os resultados do *Clustan* (considerado automático: não supervisionado) e os resultados obtidos pela Limiarização e pela aplicação dos classificadores Maxver e SVM (considerados semi-automáticos: supervisionados).

- A interpretação visual dos resultados indicou que os métodos semi-automáticos conseguiram retornar *clusters* mais representativos e, consequentemente, geometrias mais precisas para os alvos escuros, delineando com maior acurácia os limites que os separam do oceano. A Média das métricas de *ExAE* e de erros de Inclusão e de Omissão confirmou esta interpretação indicando que o melhor desempenho foi obtido pelo classificador SVM, seguido da Limiarização. Enquanto os métodos semi-automáticos erraram mais excluindo regiões que seriam alvos escuros, o método automático errou mais incluindo regiões oceânicas como alvos escuros.

- Uma comparação entre os resultados obtidos pela tese com resultados obtidos por diferentes autores que consideram seus métodos automáticos foi realizada. A comparação dos resultados efetuada entre alvos considerados fáceis e difíceis evidenciou que o método proposto também poderia ser considerado automático por ter conseguido resultados semelhantes ao de diversos autores.

Entretanto, o método proposto não obteve resultados considerados satisfatórios para todas as cenas estudadas.

- Do ponto de vista operacional, mesmo com imprecisão, os alvos escuros detectados pelo método de clusterização seriam úteis para o acionamento das inspeções de campo. Entretanto, geometrias com baixa exatidão não apresentariam acurácia suficiente em relação ao *Phantom* para serem utilizadas como dado de entrada em um sistema de classificação automática (hipótese x). Na pior hipótese, ao final da classificação um resultado errôneo poderia ocasionar falhas no processo de combate a uma possível emergência envolvendo derramamento de óleo no mar.

- Também não seria aconselhável armazenar geometrias imprecisas em um Banco de Dados que tem por objetivo fornecer conhecimento pretérito para retroalimentar o sistema automático de detecção de óleo e assemelhados. O registro cumulativo de geometrias imprecisas e, consequentemente, de atributos imprecisos no Banco de Dados poderia levar à ineficácia do sistema, confundindo o classificador na distinção dos alvos escuros em óleo ou assemelhados.

- Neste contexto, a elaboração de um sistema híbrido inteligente capaz de indicar através de regras de decisão as imagens aptas para o processamento automático ou semi-automático foi proposta. Para tal, regras de decisão foram estabelecidas e confrontadas a cada nova imagem SAR, indicando as cenas a serem processadas automaticamente e as cenas que necessitariam de supervisão. Esta proposta permitiu utilizar o processamento automático nas situações em que os resultados foram favoráveis.

6.5 - Validação e Síntese da Metodologia Proposta:

- Atributos que consideraram a distância entre as médias dos 5 atributos indicados como padrão, obtidas para os *clusters* dos alvos escuros e do oceano, foram calculados para as 11 cenas e processados pelo método de ACR para encontrar as regras de decisão. O método de ACR retornou uma regra compacta, considerando como atributo de partição inicial a diferença entre a Moda e como atributo de partição secundário a diferença entre o Máximo nos dois tipos de *clusters*.

- As 3 imagens de validação foram pré-processadas, segmentadas e agrupadas com todos os parâmetros padronizados ao longo da tese: a) Espaçamento de *Pixel*: 100m; b) Filtragem: Lee 3x3; c) Segmentação: SPRING com AM= 25 *pixels* e LS = 4, e; d) Clusterização: *Clustan*, considerando 2 *clusters*. Os atributos de

Diferença entre a Média dos *clusters* para o atributo de Moda e de Máximo foram calculados para serem confrontados com a regra de decisão encontrada.

- A regra não foi validada para todos os exemplos utilizados, onde 2 cenas deveriam ter sido indicadas para o processamento automático e apenas uma para o processamento semi-automático. A cena de validação A, que retornou o melhor índice de *ExAE*, foi indicada para o processamento semi-automático. Reprocessar supervisionadamente uma cena que teve resultados satisfatórios no modo automático significaria um custo computacional desnecessário. Entretanto, do ponto de vista operacional seria melhor do que indicar para o processamento automático uma cena não favorável, produzindo resultados dotados de reduzida acurácia geométrica.

- Os resultados evidenciaram o potencial destas regras para auxiliar no processo de automatização, sendo necessárias mais amostras para retornar regras mais robustas aplicáveis a toda nova imagem SAR adquirida.

- Desta forma, considerando a premissa, as hipóteses consideradas e os resultados obtidos ao longo da tese, o fluxograma da Figura 6.1 ilustra o método proposto para a detecção de alvos escuros em imagens SAR, adotando uma abordagem híbrida que integra métodos automáticos e semi-automáticos.

6.6 - Considerações Finais e Recomendações:

- Uma importante recomendação seria aumentar o número de exemplos obtidos com o mesmo método de detecção a fim de aprimorar a regra de decisão encontrada para que ela seja aplicável, com maior eficácia, a qualquer nova imagem SAR.

- Outra perspectiva seria investigar novos métodos para o pré-processamento das imagens com o objetivo de produzir novos atributos capazes de discriminar melhor os alvos escuros e o oceano, melhorando os resultados da clusterização. A possibilidade de utilizar diferentes técnicas de seleção de atributos e clusterização de dados também pode ser verificada.

- Efetuar testes comparativos utilizando os dados no formato CEOS e *MrSid*, seria interessante para verificar quantitativamente os benefícios e as perdas associados a utilização do formato *MrSid*.

MÉTODO PROPOSTO PARA A DETECÇÃO DE ALVOS E ESCUROS ACOPLADO COM O SISTEMA DE CLASSIFICAÇÃO AUTOMÁTICA:

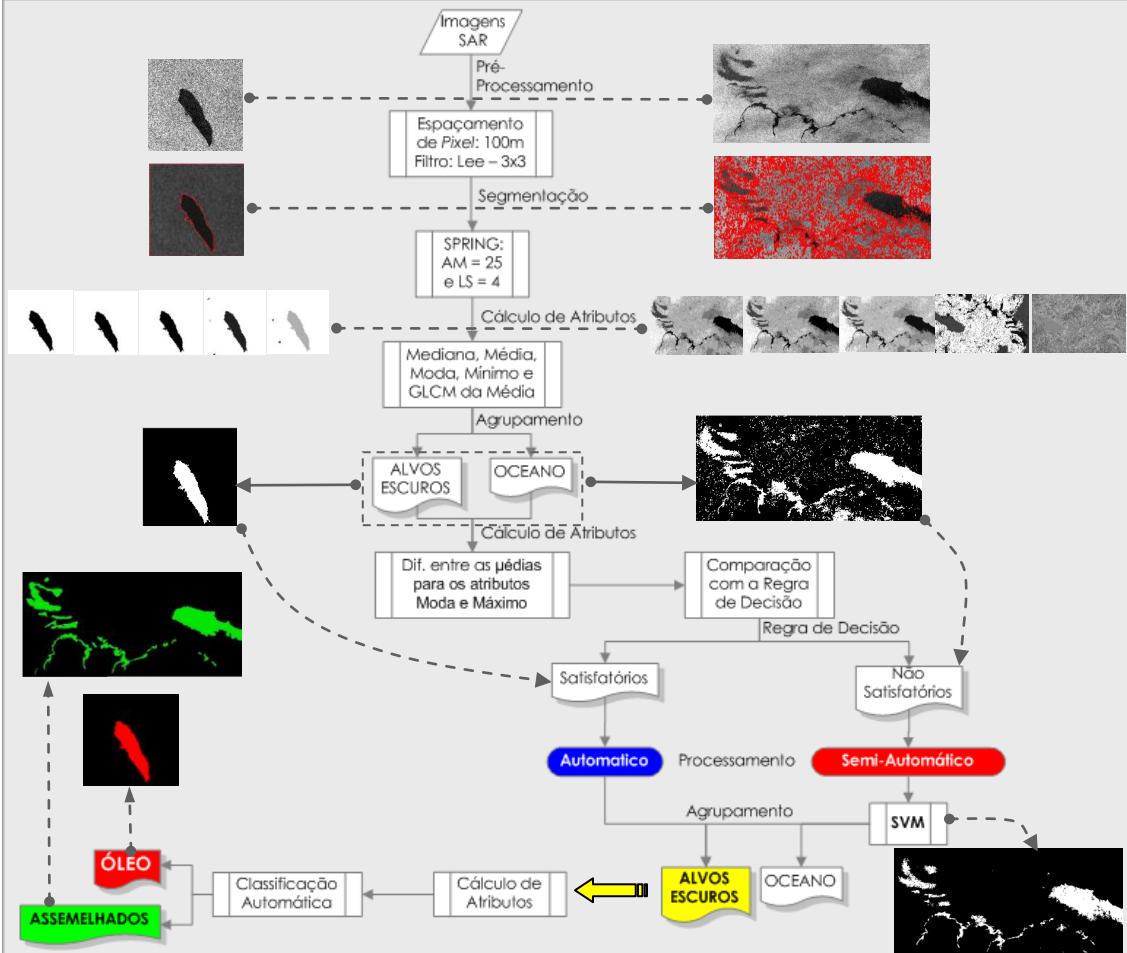


Figura 6.1 – Método proposto para a detecção automática de alvos escuros e classificação em óleo e assemelhados.

- Realizar experimentos utilizando o modo para segmentação de imagens óticas disponível no software MultiSeg, seria recomendado para confirmar se este produziria resultados melhores do que o SPRING.

- Verificar a possibilidade de utilizar dados contextuais referentes a intensidade da velocidade do vento em cada cena para auxiliar na separação dos alvos escuros do oceano, especialmente nas cenas em que alvos escuros referentes a óleo, algas, dentre outros ocorrem simultaneamente com eventos de baixa intensidade de vento.

- Testar o método proposto com imagens SAR adquiridas por outros satélites como o ENVISAT/ASAR, COSMO SKYMED, TERRASAR, dentro outros. Importante destacar que, embora no contexto da tese só tenham sido utilizadas imagens provenientes do satélite RADARSAT-1, o método é replicável para as imagens do RADARSAT-2. Isto é factível considerando que as configurações deste satélite são as mesmas do RADARSAT-1, sendo a única diferença a possibilidade de aquisição de imagens em outras polarizações que não somente a HH.

- Uma importante recomendação consiste em efetuar testes futuros acoplando as geometrias resultantes do método elaborado ao sistema automático de classificação proposto por BENTZ (2006). Este procedimento teria como objetivo verificar se as geometrias possibilitarão o cálculo de atributos eficientes para a classificação dos alvos escuros em óleo e assemelhados, considerando as diferentes sub-classes utilizadas por BENTZ (2006).

- Finalmente, considerando as dificuldades encontradas na detecção de alvos escuros em imagens SAR amplamente discutidas na literatura, seria essencial poder comparar os resultados dos diferentes métodos propostos para encontrar uma definição do que seria de fato um sistema automático e quais seriam os melhores métodos. Para que isso fosse possível seria recomendado que um conjunto de dados compilado, reunindo recortes de cena dotados de complexidade diferenciada, fosse disponibilizada para ser amplamente utilizado pelos pesquisadores e organizações interessados em desenvolver soluções para a detecção de óleo.

- O encontro SEASAR realizado em janeiro de 2010 na ESA, reunindo especialistas da área, focados no desenvolvimento científico e na aplicação operacional de sistemas para detecção de óleo obteve, como principal conclusão que: “Sistemas completamente automáticos para a detecção de vazamento de óleo ainda não constituem uma realidade a ser implementada operacionalmente. Entretanto, considerando a relevância do tema e o potencial dos dados SAR para a detecção destes eventos, a comunidade científica reforçou a necessidade da continuidade dos estudos nesta área.”

- Os resultados obtidos ao longo da tese refletem o estado da arte referente ao desenvolvimento de sistemas para a detecção automática de óleo e assemelhados. Neste contexto, a proposta de um sistema híbrido constitui uma alternativa interessante para as cenas dotadas de alvos mais difíceis de serem detectados, uma vez que a utilização de um método automático plenamente eficaz para todas as cenas ainda não foi encontrado.

REFERÊNCIAS BIBLIOGRÁFICAS

- AKKARTAL A.; F. SUNAR. The usage of radar images in oil spill detection, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B8. Beijing 2008, pp. 271-276, 2008.
- ALMEIDA, R. F.; MIRANDA F.P.; LORENZETTI, J.A.; PEDROSO, E.C.; BEISL, C.H.; LANDAU, L. RADARSAT-1 Images in Support of Petroleum Exploration: The Offshore Amazon River Mouth Example. Canadian Journal of Remote Sensing, Vol. 31, No 4, pp. 289-303, 2005.
- AYED, I. B.; HENNANE, N.; MITICHE, A. Unsupervised variational image segmentation/classification using a Weibull observations model. IEE Transections on Image Processing, vol 15, no 11, November, 2006.
- ALPERS, W.; HÜHNERFUSS, H. "Radar signatures of oil films floating on the sea surface and the Marangoni effect". Journal of Geophysics, Res., vol. 93(C4), pp. 3642–3648, 1988.
- ALPERS, W.; HÜHNERFUSS, H. The damping of ocean waves by surface films: a new look at on old problem. Journal of Geophysics, Res., vol. 94, no. C5, pp. 6251.6265, 1989.
- ALPERS, W.; WISMANN, V.; THEIS, R.; HUHNERFUSS, H.; BARTSCH, N.; MOREIRA, J.; Lyden, J. The damping of ocean surface waves by monomolecular sea slicks measured by airborne multi-frequency radars during the saxon-fpn experiment. Proceedings IGARSS'91, 1991.
- ALPERS, W.; ESPEDAL, H. A. Oils and Surfactants. Chapter 11 in Synthetic Aperture Radar Marine User`s Manual, Organized by Jackson C. R. & Apel J. R., National Oceanic and Atmospheric Administration – NOAA, 2004.
- ARVELYNA, Y.; OSHIMA, M.; KRISTIJONO, A.; GUNAWAN, I. Auto Segmentation of Oil Slick in RADARSAT SAR Image Data Around Rupat Island, Malaca Strait. Proceedings 22nd Asian Conference on Remote Sensing, Singapore, 2001.
- ANDRADE, L. P. Procedimento Interativo de Agrupamento de Dados [Rio de Janeiro], VIII, 193 p., COPPE/UFRJ, Dissertação de Mestrado, Engenharia Civil, 2004.

ANGIULI, E.; FRATE, F.; SALVATORI, L. Neural Networks for oil spill detection using ERS and ENVISAT imagery, 2006 [online] <http://earth.esa.int/workshops/seasar2006/proceedings/papers/s5_4_ang.pdf>, acesso novembro 2006.

GIRARD-ARDHuin, F.; G. MERCIER, R.; GARELLO. Oil slick detection by SAR imagery: potential and limitation. Marine Technology and Ocean Science Conf. (OCEANS'03), San Diego, USA, pp. 22-26 September ,2003.

ASSILZADEH, H.; MAGED, M.; MANSON, S.B.; MOHAMED, M.I. Application of trajectory model, remote sensing and geographic information systems (GIS) for oil spill emergency response in straits of Malacca. In: Proceedings of the 20th Asian Conference on Remote Sensing, Hong Kong, pp. 331–339, 1999.

BAATZ, M.; SCHÄPE, A., “Object-oriented and multi-scale image analysis in semantic networks”. In: *Proceedings of the 2nd International Symposium on Operationalization of Remote Sensing, ITC, Enschede, Netherlands*, Aug, 1999.

BAATZ, M.; SCHÄPE, A. “Multiresolution segmentation – An optimization approach for high quality multi-scale image segmentation”. In: *Angewandte Geographische Informationsverarbeitung XI. Beiträge zum AGIT – Symposium 1999*, Kalsruhe Herbert Wichmann Verlag, Salzburg, 2000.

BAATZ, M.; BENZ, U.; DEHGHANI, S.; HEYNEN, M.; HÖLTJE, A.; HOFMANN, P.; LINGENFELDER, I.; MIMLER, M.; SOHLBACH, M.; WEBER, M.; WILLHAUCK, G., *eCognition User Guide*, 2 ed, Definiens Imaging, München, 2003.

BAATZ, M.; BENZ, U.; DEHGHANI, S.; HEYNEN, M.; HÖLTJE, A.; HOFMANN, P.; LINGENFELDER, I.; MIMLER, M.; SOHLBACH, M.; WEBER, M.; WILLHAUCK, G. *e-Cognition Professional User Guide* 4. München, Germany: Definiens Imaging GmbH, 486 p, 2004.

BARNI, M.; BETTI, M.; MECOCCI, A. A fuzzy approach to oil spill detection on SAR images. Proc. IGARSS '95, vol. 1 (pp. 157– 159), 1995.

BENTZ C. M.; LORENZETTI J.A.; KAMPEL M.; POLITANO A.T.; GENOVEZ P.C., LUCCA A.V.D. Contribuição de dados ASTER, CBERS, R99/SIPAM e OrbiSAR-1 para o monitoramento oceânico – Resultados do projeto FITOSAT Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Abril de 2007a.

BENTZ C. M.; EBECKEN N. F.F.; POLITANO A. T. Identificação automática de eventos ambientais costeiros e oceânicos detectáveis em imagens de radares orbitais. Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Abril de 2007b.

BENTZ, C. M. Reconhecimento Automático de Eventos Ambientais Costeiros e Oceânicos em Imagens de Radares Orbitais [Rio de Janeiro], XIII, 115 p., COPPE/UFRJ, D.Sc., Engenharia Civil, 2006.

BENTZ, C.M.; POLITANO, A.T.; GENOVEZ, P.C; A.T. MIRANDA, R.S.P. "Offshore Oil Spill Detection and Monitoring in E&P Areas – An Operational Multi-Sensors Approach". *Anais*, Eighth International Conference On Health, Safety and Environment, Abu Dhabi - UAE, Abril de 2006.

BENTZ, C.; POLITANO, A.; SPINDOLA, R.; GENOVEZ, P. *A Multi-Sensor Approach for Oil Spill and Sea Surface Monitoring in Southeastern Brazil*, Halifax, Nova Scotia Canada. Proceedings of the 8th International Conference on Remote Sensing for Marine Environments, 2005a.

BENTZ, C.; POLITANO, A.; GENOVEZ, P, Monitoramento ambiental de áreas costeiras e oceânicas com múltiplos sensores orbitais, Goiânia, Brasil. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, 2005b.

BENTZ, C. M. Visita à Tromsø Satellite Station (TSS) e ao Nansen Environment and Remote Sensing Center (NERSC), Noruega. Relatório de Viagem Comunicação Técnica PETROBRAS, 2003.

BENTZ, C.M.; MIRANDA, F.P. Application of Remote Sensing Data for Oil Spill Monitoring in The Guanabara Bay, Rio De Janeiro, Brazil. Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'2001), Sydney, IEEE, Pisacataway, USA, 2001.

BEAULIEU, J. M.; TOUZI, R. Hierarchical segmentation of polarimetric SAR images. In: International Geoscience and Remote Sensing Symposium (IGARSS'02) and Canadian Symposium on Remote Sensing, 24. Proceedings Toronto, Canada: IGARSS, 2002. p. 2590–2592, 2002.

BEAUCHEMIN, M.; THOMSON. The evaluation of segmentation results and the overlapping area matrix. International Journal of Remote Sensing. v. 18, n. 18, p. 3895-3899, December, 1997.

BINS, L.; FONSECA, L.; ERTHAL, G. Satellite Imagery Segmentation: a region growing approach. in VIII Brazilian Symposium on Remote Sensing. São José dos Campos, BR: INPE, 1996.

BJERDE, K.W.; SOLBERG, A. H. S.; SOLBERG, R. Oil spill detection in SAR imagery. Proc. IGARSS'93, vol. 3, pp. 943–945, 1993.

BLUM, A.; Langley, P. *Selection of relevant features and examples in machine learning*. Artificial intelligence, 97, pp.245-271 , 1997.

BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 1984.

BREKKE, C.; SOLBERG, A. H. S. oil spill detection by satellite remote sensing – Review. Remote Sensing of Environment (95), pp 1-13, 2005.

BROWN C. E.; FINGAS M. F. New space-borne sensing for oil spill response, International Oil Spill Conference, 2001.

BURGES, C. J. C. “A tutorial on support vector machines for pattern recognition”, *Data Mining and Knowledge Discovery*, v. 2, n. 2, pp. 121–167, 1998.

CALABRESI, G.; DEL FRATE, F.; LICHTENEGGER, J.; PETROCCHI, A.; TRIVERO, P., “Neural networks for the oil spill detection using ERS-SAR data”. In: *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'99), Hamburg*, v. 1, IEEE, Pisacataway, USA, pp. 215 – 217, 1999.

CÂMARA, G., et al. SPRING: Integrating Remote Sensing and GIS with Object-Oriented Data Modelling. Computers and Graphics, 15(6): p. 13-22, 1996.

CANADA CENTRE FOR REMOTE SENSING, Natural Resources Canada (2006 - last updated 03 September). Tutorial: Fundamentals of Remote Sensing. from http://ccrs.nrcan.gc.ca/resource/tutor/fundam/index_e.php. Retrieved on June 11, 2007.

CAVES, R.; QUEGAN, S.; WHITE, R.G. Quantitative comparison of the performance of SAR segmentation algorithms. IEEE Transactions IP, 1996.

CAVES, R.; QUEGAN, S.; WHITE, R. G. Quantitative comparison of the performance of SAR segmentation algorithms. IEEE Transactions on Image Processing, v. 7, n. 11, p. 1534–1546, Nov. 1998.

CHANG, L.; CHENG, C. M.; TANG, Z. S. An automatic detection of oil spills in SAR images by using image segmentation approach. IEEE, 2005 [online] <<http://ieeexplore.ieee.org/iel5/10226/32596/01525287.pdf?arnumber=1525287>>, acesso novembro 2006.

CHANG, Y.; LEE, M.; SHIMADA, A.T.; SAKAIDA, F.; KAWAMURA, H.; CHAN, J.; W.-LU H. J. Wintertime high-resolution features of sea surface temperature and chlorophyll-a fields associated with oceanic fronts in the southern East China Sea, International Journal of Remote Sensing, v.29 n.21, p.6249-6261, November 2008 .

CHEN, M.; HAN, J.; YU, P. S. Data Mining: An Overview from a Database Perspective. IEEE Transections on Knownldge and data Engineering, Vol 8, No 6, December, 1996.

DELVES, L. M.; WILKINSON, R.; OLIVER, C. J.; WHITE, R. G. Comparing the performance of SAR segmentation algoritms. International Journal of Remote Sensing, v. 13, n.11, p. 21221-21249, 1992.

DEL FRATE, F.; PETROCCHI, A.; LICHTENEGGER, J.; CALABRESI, G. Neural Networks for oil spill detection using ERS-SAR data, IEEE Transc. on Geosc. and Remote Sensing, 3 (5), 2282-2287, 2000.

DEFINIENS PROFESSIONAL 5. Reference Book. Published by Definiens Ag, Trappentreus 1, München, Germany, 2006.

ENGELHARDT F. R. Remote Sensing for oil spill detection and response. Pure Appl, vol.71 (1), pp103-111, 1999.

ELACHI, C. *Introduction to the Physics and Techniques of Remote Sensing*, New York (NY) USA, John Wiley & Sons, 1987.

ESPEDAL, H.A.; WAHL, T. "Satellite SAR oil spill detection using wind history information", International Journal of Remote Sensing, 20 (1), 49-65, 1999.

ESPEDAL, H. A. Detection of oil spill and natural film in the marine environment by spaceborne synthetic aperture radar. PhD thesis, Department of Physics University of Bergen and Nansen Environment and Remote Sensing Center, Norway, 1998.

ESPINDOLA, G. M.; CÂMARA, G.; REIS, I. A.; BINS, L. S.; MONTEIRO, A. M. Spatial Autocorrelation Indicators for Evaluation of Remote Sensing Image Segmentation Algorithms. In: GIS and Spatial Analysis - 2005 Annual Conference of the International Association for Mathematical Geology, Toronto - Canada, 2005.

ESPINDOLA G. M. Ajuste de parâmetros em algoritmos de segmentação de imagens por crescimento de regiões. Dissertação de Mestrado em Sensoriamento Remoto, INPE, 2006.

EVANGELOU, I. E.; HADJIMITSIS, D.G.; LAZAKIDOU, A.; CLAYTON, C. I. Data mining and knowledge discovery in complex image data using artificial neural networks. In: International Conference on Logic Programming, 17, Workshop on Complex Reasoning on Geographical Data (ICLP'01), 2001.

ENVI 4.4 User's guide: The Environment For Visualizing Images. Version 4.4, Tutorial do software, consulta janeiro de 2010.

FERRARO, G.; BERNARDINI, A.; DAVID, M.; MEYER-ROUX S.; MUELLENHOFF, O.; PERKOVIC M.; TARCHI, D.; TOPOUZELIS, K. Towards an operational use of space imagery for oil pollution monitoring in the Mediterranean basin: A demonstration in the Adriatic Sea. Marine Pollution Bulletin, 54 403–422, 2007.

FINGAS, M.F.; BROWN, C.E. "Review of Oil Spill Remote Sensing", in *Proceedings of the Sixth International Conference on Remote Sensing for Marine and Coastal Environments*, Veridian ERIM International, Ann Arbor, MI, pp. I211-218, 2000a.

FINGAS, M.F.; BROWN, C.E. "Review of Oil Spill Remote Sensing", in *Proceedings of SPILLCON 2000*, Australian Marine Safety Authority, Sydney, Australia, www.meetingplanners.com.au/spillcon/, 2000b.

FINGAS, M.F.; BROWN, C.E. Review Of Oil Spill Remote Sensors. Presented at the Seventh International Conference on Remote Sensing for Marine and Coastal Environments, Miami, Florida, May, 2002.

FINGAS, M.F.; BROWN, C.E. Review Of Oil Spill Remote Sensing Presented at the Eighth International Conference on Remote Sensing for Marine and Coastal Environments, Halifax, Nova Scotia, May, 2005.

FISCELLA, B.; GIANCASPRO, A.; NIRCHIO, F.; PAVESE, P.; TRIVERO, P. Oil spill detection using marine SAR images. *International Journal of Remote Sensing* Vol. 21, No 18, 3561–3566, 2000.

FRERY, A. C.; CORREIA, A. H.; FREITAS, C. C. Classifying Multifrequency Fully Polarimetric Imagery with Multiple Sources of Statistical Evidence and Contextual Information. *IEEE Transactions on Geoscience and Remote Sensing*, v. 45, p. 3098-3109, 2007.

GADE, M.; SCHOLZ, J.; VON VIEBAHN, C. On the detectability of marine oil pollution in European marginal waters by means of ERS SAR imagery. *Proc. IGARSS 2000*, vol. 6 (pp. 2510– 2512), 2000.

GADE, M.; ALPERS, W.; HÜNERFUSS, H.; WISMANN, V. R.; LANGE, W.G. On the reduction of radar backscatter by oceanic surface films: Scatterometer measurements and their theoretical interpretation, *Remote Sens. of Env.* 66(1), 52-70, 1998.

GALLAND, F.; GERMAIN O.; BERTAUX, N.; RÉFRÉGIER, P. Partitioning into Homogeneous Regions Adapted to SAR Images, 2003 [online] <http://www.fresnel.fr/perso/galland/Galland04_ESA_EUSC_preprint.pdf>, acesso novembro, 2006.

GALLAND, F.; BERTAUX, N.; RÉFRÉGIER, P. Minimum description Length Synthetic Aperture RADAR. *Image Segmentation. Image Processing*, IEEE Transactions, 2003.

GALLAND, F.; RÉFRÉGIER, P.; GERMAIN, O. Synthetic Aperture Radar oil spill segmentation by stochastic complexity minimization. 2004, [online] <<http://ieeexplore.ieee.org/iel5/8859/28004/101109LGRS2004835346.pdf>>, acesso novembro, 2006.

GAMBLE, L. Comunicação pessoal efetuada através de email com o operador da estação de recepção de dados RADARSAT Gatineau. Contato efetuado em 2008.

GASULL, A.; FABREGAS, X.; JIMENEZ, J.; MARQUES, F.; MORENO, V.; HERRERO, M. Oil spills detection in SAR images using mathematical morphology. Proc. EUSIPCO'2002, Toulouse, France, vol. 1 (pp. 25–28), 2002.

GENOVEZ, P. C.; POLITANO, A. T.; BENTZ, C. M. Utilização de Imagens ENVISAT para a Detecção e Monitoramento de Acidentes Ambientais: Uma Abordagem Operacional. Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Abril de 2007.

GENOVEZ, P.; FREITAS, C. C.; BENTZ, C. M.; FREITAS, R. M.; DUTRA, L. V. Detecção de Óleo e Ocorrências Ambientais em Imagens SAR: Avaliação de Diferentes Parâmetros de Pré-processamento e Diferentes Segmentadores. Anais XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Abril de 2009.

GENOVEZ, P.; PALMEIRA, F. L. B.; BENTZ, C. M. Detecção e Monitoramento de Óleo através da Utilização Integrada de Imagens de Satélite, Verificações de Campo e Modelagem de Deriva de Óleo: Uma Abordagem Operacional. Anais XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Abril de 2009.

GOODMAN, R. H.; BROWN, C. E. Oil Detection Limits For A Number Of Remote Sensing Systems Presented at the Eighth International Conference on Remote Sensing for Marine and Coastal Environments, Halifax, Nova Scotia, May 2005.

GOODMAN, R. H., "Overview and Future Trends in Oil Spill Remote Sensing", *Spill Science and Technology Bulletin*, Vol. 1, No. 1, pp. 11-21, 1994.

GONZALES, R.C.; WINTZ, P. Digital image processing. New York: Addison-Wesley, 1987.

HANG, L. M.; DINH, D. N., Oil Spill Detection and Classification by ALOS PALSAR at Vietnam East Sea. 7th FIG Regional Conference Spatial Data Serving People: Land Gibvernance and the Environment – Building the Capacity . Hanoi, Vietnan, 19-22, October, 2009.

HAN, J.; KAMBER, M. Data Mining: Concepts and Techniques. Academic Press, United States, 2001.

HARALICK, R.; SHANMUGAN, K.; DINSTEIN, I., "Textural features for image classification", *IEEE Transactions on Systems, Man and Cybernetics*, v. 3, n. 1, pp. 610-621, 1973.

HARALICK, R.M. "Statistical and structural approaches to texture", Proceedings of the *IEEE*, v. 67, n. 5, pp. 786-804, 1979.

HARALICK, R.M.; SHAPIRO, L.G. Image segmentation techniques. Machine Vision International: Ann Arbor, 1984.

HOLT, B. SAR Imaging of Ocean Surface. Chapter 2 in Synthetic Aperture Radar Marine User's Manual, Organized by Jackson C. R. & Apel J. R., National Oceanic and Atmospheric Administration – NOAA, 2004.

HOVLAND, H.; JOHANNESSEN, J.A.; DIGRANES, G. Slick detection in SAR images. Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'94), IEEE, Pisacataway, USA, pp. 2038 - 2040, 1994.

HUGHES, G. On the Mean Accuracy of Statistical Pattern Recognizers. *IEEE TRANSACTIONS ON INFORMATION THEORY*, january, 1968

INDREGARD, M.; SOLBERG, A.; CLAYTON, P. D2-report on benchmarking oil spill recognition approaches and best practice. Tech. rep., Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177, 2004.

JAIN, A. K.; MURTY, M. N.; FLYN, P. J. Data Clustering: A Review. *ACM Computing Surveys*, Vol 31, No 3, September, 1999.

JAIN, A.. K.; DUBES, R. C. Algorithms for clustering data. Prentice-Hall, advanced reference series. Prentice-Hall Inc., Upper saddle River,NJ, 1988.

JOHNSON, R.A.; WICHERN, D.W. *Applied multivariate statistical analysis*. 3 ed. New Jersey, Prentice Hall, 1992.

KANAA, T.F.N.; TONYE, E.; MERCIER, G.; ONANA, V.P.; NGONO, J.M.; FRISON, P.L.; RUDANT, J.P.; GARELLO, R. Geoscience and Remote Sensing SymposiumIGARSS'03, 4 (21), 2750 – 2752, 2003.

KARANTZALOS, K.; ARGIALAS, D. Automatic detection and tracking of oil spills in SAR imagery with level set segmentation, International Journal of Remote Sensing, v.29 n.21, p.6281-6296, November, 2008.

KARATHANASSI, V.; TOPOUZELIS, K.; PAVLAKIS, P.; ROKOS, D. An object-oriented methodology to detect oil spills. *Int. J. Remote Sens.* 27, 5235-5251, 2006.

KERAMITSOGLOU, I.; CARTALIS, C.; KIRANOUDIS, C. Automatic Identification of oil spills on satellite images. Environmental Modelling & Software (21), pp 640-652, 2006.

KOTOVA, L. A.; ESPEDAL, H. A.; JOHANNESSEN, O. M. Oil spill detection using spaceborne SAR; a brief review. Proc. 27th ISRSE, Tromsb, Norway, 1998.

KUBAT, M.; HOLTE, C.; MATWIN, S. Machine learning for detection of oil spills in satellite Radar images, Machine Learning, vol.30, no 2-3, 195-215, 1998.

KURZ, L.; BENTEFTIFA, H. Analysis of variance in statistical image processing. Cambridge: Cambridge University Press, 1997.

LEHMANN, E.L.; JOSEPH, P.; ROMANO. *Testing Statistical Hypotheses* (3E ed.). New York: Springer, 2005.

LEE-LUENG, F.; HOLT, B. SeaSat Views Oceans and Sea Ice with Synthetic-Aperture Radar. National Aeronautics and Space Administration Jet Propulsion Laboratory (NASA), Califórnia -USA , 1982.

LEWIS, A. J.; HENDERSON, F. M. Radar Fundamentals: The Geoscience Perspective. Chapter 3 in Principles & Applications of Imaging RADAR – Manual of Remote Sensing, Volume 2 – edited by Henderson F. M. & Lewis A. J. John Wiley & Sons, 1998.

LE MOIGNE, J.; TILTON, J.C. Refining image segmentation by integration of edge and region data. IEEE Transactions on Geoscience and Remote Sensing, v. 33, n. 3, p. 605-615, May, 1995.

LIMA, J.A.M.; SCOFANNO, A.M., "New oceanographic challenges in the Brazilian deepwater oil fields". In: *Offshore Technology Conference – OTC*, Huston, Texas, May, 1999.

LIU, H.; SETIONO, R. "Chi2: Feature Selection and Discretization of Numeric Attributes," *Tools with Artificial Intelligence, IEEE International Conference on*, pp. 388, Seventh International Conference on Tools with Artificial Intelligence, 1995.

LOPEZ, L.; MOCTEZUMA, M.; PARMIGIANI, F. Oil Spill Detection Using GLCM and MRF. IEEE, 2005, [online] <<http://ieeexplore.ieee.org/iel5/10226/32597/01526349.pdf>> arnumber=1526349>, acesso dezembro 2006.

LUCCA, E. V. D. Avaliação e comparação de algoritmos de segmentação de imagens de radar abertura sintética. 207 p. Dissertação de Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 1998.

LUCCA, E. V. D.; FREITAS, C.; FRERY, A. C.; ANNA, S. J.S.S. Comparison of SAR Segmentation Algorithms. Anais IX Simpósio Brasileiro de Sensoriamento Remoto, Santos, 1998.

MACCANDLESSS, S. W.; JACKSON, C. R. Principles of Synthetic Aperture Radar. Chapter 2 in Synthetic Aperture Radar Marine User's Manual. Organized by Jackson C. R. & Apel J. R. National Oceanic and Atmospheric Administration – NOAA, 2004.

MANORE, M. J.; VACHON, P. W.; BJERKELUND, C.; EDEL, H. R.; RAMSAY, B. Operational use of RADARSAT SAR in the coastal zone: The Canadian experience. 27th *International Symposium on Remote Sensing of the Environment*, Tromsø, Norway, June 8–12 (pp. 115– 118), 1998.

MARANGONI, C. Sul principio della viscosità superficiale dei liquidi stabilito. *Nuevo Cimento*, ser. 2 (5/6), pp. 239.273, 1872.

MARGHANY, M. RADARSAT automatic algorithms for detecting coastal oil spill pollution, Intern. J. Appl. Earth Observation and Geoinformation, Vol. 3, N 2, pp. 191-196 , 2001.

MEINEL, G.; NEUBERT, M. A comparison of segmentation programs for high resolution remote sensing data. In: Proceedings International Society for Photogrammetry and Remote Sensing 29 Congress, XX, (ISPRS), Istanbul, Turkey - July – 2004.

MERCIER, G.; DERRODE, S.; PIECZYNSKI, W.; CAILLEC, J. -M. L.; GARELLO, R. Multiscale oil slick segmentation with Markov Chain Model. Proc. IGARSS'03, vol. 6 (pp. 3501– 3503), 2003.

MERCIER, G.; ARDHUIN, G. Oil Slick Detection by SAR Imagery Using Kernel Expansion, 2005a, [online] <<http://perso.enst-bretagne.fr/~mercierg/articles/2005/oceansSVM.pdf>>, acesso janeiro 2007.

MERCIER, G.; ARDHUIN, G. Unsupervised Oil Slick Detection by SAR Imagery Using Support Vector Machines, 2005b, [online] <<http://perso.enst-bretagne.fr/~mercierg/articles/2005/mercier.igarss05.pdf>>, acesso janeiro 2007.

MIGLIACCIO, M.; TRANFAGLIA, M. Oil Spill Observation by SAR: A Review, *Proc. US-Baltic International Symposium*, Klaipeda (Latvia), June 2004.

MIRANDA, F. P.; MENDOZA, A.; PEDROSO, E. C. ; BEISL, C. H. ; WELGAN, P. ; MORALES, L. M.. Analysis of RADARSAT-1 data for offshore monitoring activities in the Cantarell Complex, Gulf of Mexico, using the unsupervised semivariogram textural classifier (USTC). Canadian Journal of Remote Sensing, Ottawa, Ontario, Canada, v. 30, n. 3, p. 424-436, 2004.

MONTALI, A.; GIACINTO, G.; MIGLIACCIO, M.; GAMBARDELLA, A. Supervised Pattern Classification Techniques for Oil Spill Classification in SAR Images: Preliminary Results. Proceedings of SEASAR, Frascati, Italy, 2006.

NAVULUR, K. Multispectral image analysis using the object-oriented paradigm. CRC Press Taylor & Francis Group 6000 Broken Sound Parkway, London - New York, 2007.

NEUBERT, M.; HEROLD, H.; MEINEL, G. Evaluation of Remote Sensing Image Segmentation Quality- Further results and Concepts, 2006 – [online] <http://www.commission4.isprs.org/obia06/Papers/10_Adaption%20and%20further%20development%20II/OBIA2006_Neubert_Herold_Meinel.pdf>, acesso, março 2007.

NISBET, R. ; ELDER, J.; MINER, G. Handbook Of Statistical Analysis And Data Mining Applications. Academic Press an imprint for Elsevier, USA, San Diego, California, 2009.

PEKKARINEN, A. A method for the segmentation of very high spatial resolution images of forested landscapes. International Journal of Remote Sensing, v. 23, n. 14, p. 2817-2836, 2002.

PELIZZARI, S.; DIAS, J.M.B. Bayesian Adaptative Oil Spill Segmentation of SAR Images Via Graph Cuts, 2003 [online] <http://earth.esa.int/workshops/seasar2006/participants/256/paper_SoniaPelizzari_SEASAR2006_final.pdf>, acesso março 2007.

PINHO, C. M. D. Análise Orientada a Objetos de Imagens de Satélite de Alta Resolução Espacial Aplicada à Classificação de Cobertura do Solo no Espaço Intra-Urbano: O Caso De São José Dos Campos – SP. Dissertação de mestrado defendida no INPE, São José dos Campos, 2006.

PRATT, W. Digital Image Processing. New York Wiley-Interscience, 1991.

QUEGAN, S.; RYE, A.J.; HENDRY A.; SKINGLEY J.; ODDY C..J. Automatic interpretation strategies for synthetic aperture radar images. Philosophical Transactions of the Royal Society of London. A 324, p. 409-421, 1988.

RAMALHO, G.L.B.; MEDEIROS, F.N.S. Improving Reliability of Oil Spill Detection Systems Using Boosting for High-Level Feature Selection. M. Kamel and A. Campilho (Eds.): ICIAR 2007, LNCS 4633, pp. 1172–1181, 2007.

RANEY, K. R. Radar Fundamentals: Technical Perspective. Chapter 2 in Principles & Applications of Imaging RADAR – Manual of Remote Sensing, Volume 2 – edited by Henderson F. M. & Lewis A. J. John Wiley & Sons, 1998.

ROMÁN-ROLDÁN, R.; GÓMEZ-LOPERA, J. F.; ATAE-ALLAH, C; MARTÍNEZ-AROZA, J.; LUQUE-ESCAMILLA, P. L. A measure of quality for evaluating methods of segmentation and edge detection. Pattern Recognition., v.34, p:969-980, 2001.

SEASAR. Workshop, SEASAR 2010, entitled "Advances in SAR Oceanography from ENVISAT, ERS and ESA third party missions", which will be hosted at ESA ESRIN in Frascati, Italy, from 25 to 29 January 2010.

SCOFIELD, G. B.; SANT'ANNA, S. J. S.; FREITAS, C. C.; DUTRA, L. V.. Avaliação quantitativa do SegSar através de medidas de borda e regiões em imagens ópticas

sintéticas. Anais - XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Abril, 2007.

SCHMULLIUS, C.C.; EVANS, D. L. "Synthetic Aperture Radar (SAR) frequency and polarization requirements for applications in ecology, hydrology, and oceanography: A tabular status quo after SIR-C/X-SAR", *International Journal Remote Sensing*, v. 18, n. 13, pp. 2713-2722, 1997.

SIEGAL, S. Estatística não paramétrica para as ciências do comportamento. São Paulo, Mc Grow -Hill; 350p., 1975.

SILVA, M. P. S. Mineração de padrões de mudanças em imagens de sensoriamento remoto. 2006-02-03. 128 p. (INPE-14429-TDI/1130). Tese de Doutorado - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2006.

SILVA, M. P. S.; CÂMARA, G.; SOUZA, R. C. M.; VALERIANO, D. M. ; ESCADA, M. I. S. Mining Patterns of Change in Remote Sensing Image Databases. In *The Fifth IEEE International Conference on Data Mining*, Houston, 2005.

SKØELV, A. H. S.; WAHL, T. Oil spill detection using satellite based SAR, Phase 1B competition report. Technical report, Norwegian Defense Research Establishment, 1993.

SLOGGETT, D.; JORY, I., "Automatic satellite based oil slick detection and monitoring system". In: *First Workshop on ERS-1 Pilot Projects*, Toledo Spain, 1994.

SOLBERG, A. H. S; SOLBERG, R. A large-scale evaluation of features for automatic detection of oil spills in ERS SAR images, Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'97), Lincon, IEEE, Pisacataway, USA, pp.157-159, 1996.

SOLBERG, A. H. S; VOLDEN, E. Incorporation a prior knowledge in automatic classification of oil spills in ERS SAR images, Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'96), Lincon, IEEE, Pisacataway, USA, pp.1484-1486, 1997.

SOLBERG, A. H. S.; STORVIK, G.; SOLBERG, R.; VOLDEN, E. Automatic detection of oil spills in ERS SAR images, IEEE Trans. on Geosc. and Remote Sensing, 37 (4), July, 1999.

SOLBERG, A. H. S.; DOKKEN, S.T.; SOLBERG, R. Automatic detection of oil spills in Envisat, Radarsat and ERS SAR images, Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'2003), Toulouse, France, IEEE, Pisacataway, USA, 2003.

SOLBERG, A.; CLAYTON, P.; INDREGARD, M. D2 – Report on benchmarking oil spill recognition approaches and best practice. Kongsberg Satellite Services – Norway Archive No.: 04-10225-A-Doc, Issue/Revision : 2.1, 2005.

SOLBERG, A. Automatic Oil Apill Detection Based on ENVISAT, RADARSAT and ERS Images. Proceedings of the ENVISAT & ERS Symposium, Salzburg, Austria, ESA SP-572, 2005a.

SOLBERG, A. Automatic Detection and Estimating Confidence For Oil Spill Detection in SAR Images. 2005b.

SOH, L.K.; TSATSOULIS, C. Segmentation of satellite imagery of natural scenes using data mining. IEEE Transactions on Geoscience and Remote Sensing, v. 37, n. 2, p. 1086, 1999.

SOUSA Jr, M. A. Segmentação multi-níveis e multi-modelos para imagens de radar e ópticas. Tese de Doutorado em Computação Aplicada. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2005.

SOUSA Jr, M. A.; DUTRA, L. V.; FREITAS, C.C. Segmentação de Imagens JERS e TM/Landsat usando o Segmentador Icremental Multi-níveis SegSAR. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia - Brasil, 2005.

STAPLES, G.C.; HODGINS, D.O. RADARSAT-1 emergency response for oil spill monitoring. Proceedings of Fifth International Conference on Remote Sensing for Marine and Coastal Environments, CA./USA, October, 1998.

STATHAKIS, D.; TOPOUZELIS, K.; KARATHANASSI, V. Large-scale feature selection using evolved neural networks. In *Proceedings of SPIE, Image and Signal Processing for Remote Sensing XII*, Bruzzone, Lorenzo Ed.; SPIE Bellingham Wash. ETATS-UNIS: Stockholm, Sweden, September 13-14, p. 6365, 2006.

STATISTICA Data Analysis Software System, versão 8.0, Statsoft Inc.<www.statsoft.com>, 2007.

TELO, M.; LÓPEZ, M. C.; MALLORQUI, J. J. A Novel Algorithm for Automatic Ship and Oil Spill Detection Based on Time-Frequency Methods, 2006 – [online] <http://earth.esa.int/workshops/seasar2006/proceedings/papers/s5_16_tel.pdf>, acesso abril, 2007.

TILTON, J.; LAWRENCE, W. Interactive analysis of hierarchical image segmentation. In: International Geoscience and Remote Sensing Symposium IGARSS-2000, New York. IEEE Press, 2000.

TOPOUZELIS, K.; KARATHANASSI, V.; PAVLAKIS, P.; ROKOS, D. Oil Spill Detection: SAR Multiscale Segmentation and Objects features Evaluation. Proceedings of SPIE, Vol 4880, 2003.

TOPOUZELIS K.; KARATHANASSI, V.; PAVLAKIS, P.; ROKOS, D. Oil spill detection using RBF Neural Networks and SAR data, XXth ISPRS Congress, Istanbul, Turkey, July, 2004.

TOPOUZELIS K.; KARATHANASSI, V.; PAVLAKIS, P.; ROKOS, D. Detection and discrimination between oil spills and look-alike phenomena through neural networks, ISPRS J. Photogrammetry Rem. Sens., Vol. 62, N 9A, pp. 264-270, 2007a.

TOPOUZELIS K.; KARATHANASSI, V.; PAVLAKIS, P.; ROKOS, D. Oil Spill Detection Using RBF Neural Networks and SAR Data. 2004, [online] <<http://www.isprs.org/istanbul2004/comm7/papers/142.pdf>>, acesso março, 2007b.

TOPOUZELIS, K.N.. Oil spill detection by SAR images: Dark formation detection, feature extraction and classification algorithms, Sensors, Vol. 8, N 10, pp. 6642-6659, DOI: 10.3390/s8106642 , 2008.

TOPOUZELIS, K.; STATHAKIS, D.; KARATHANASSI, V. Investigation of genetic algorithms contribution to feature selection for oil spill detection, Intern. J. Rem. Sens., Vol. 30, N 3, pp. 611-625, 2009.

VACHON, P. W., THOMAS, S. J., CRANTON, J. A., BJERKELUND, C., DOBSON, F. W., OLSEN, R. B. Monitoring the coastal zone with the RADARSAT satellite. *Oceanology International* 98, UK, March 10–13, 10 pages, 1998.

VACHON, P. W.; WOLFE, J.; HAWKINS, R. K. The Impact of RADARSAT ScanSAR Image Quality on Ocean Wind Retrieval. SAR Workshop: CEOS Committee on Earth Observation Satellites; Working Group on Calibration and Validation, Proceedings of a Conference held 26-29 October 1999, Toulouse, France. Edited by Robert A. Harris and L. Ouwehand. Publisher: Paris: European Space Agency, ESA-SP vol. 450, ISBN: 9290926414, p.519, 2000.

VALENZUELA, G.R. "Theories for the interaction of electromagnetic and ocean waves – A review", *Boundary Layer Meteorology*, v. 13, pp. 61-85 , 1978.

VAPNIK, V. N., "An overview of statistical learning theory", *IEEE Transactions on Neural Networks*, v. 10, n. 5, pp. 988–999, 1999.

WEISTEEN, K.; SOLBERG, A.; SOLBERG, R. "Detection of oil spills in SAR images using a statistical classification scheme". In: *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'93)*, Tokyo, IEEE, Pisacataway, USA, pp. 943-945, 1993.

WISHART, D. Focal Point Clustering User Guide. Published by Clustan limited, Edinburgh Road, Scotland, Second Edition, 2004.

WISMANN, V.; GADE, M.; ALPERS, W.; HÜHNERFUSS, W H. "Radar signatures of marine mineral oil spills measured by an airborne multi-frequency radar", *International Journal of Remote Sensing*, v. 19, n. 18, pp. 3607-3623, 1998.

WU, S.; FLACH, P. A. Feature selection with labelled and unlabelled data. *ECML/PKDD'02 workshop on Integration and Collaboration Aspects of Data Mining, Decision Support and Meta-Learning*. M. Bohanec, B. Kasek, N. Lavrac, D. Mladenic, (eds.), pp. 156–167. August, 2002.

ZHANG, Y.J. A survey on evaluation methods for image segmentation. *Pattern Recognition*, v. 29, n. 8, p. 1335-1346, 1996.

ZHENG, Y.; DONG, H.; JIANG, Q.; Li, J. A modified FCM-Based algorithm for oil spill detection in SAR images. *Environmental Informatics Archives*, Vol 3, pp 346-351, 2005.

ZUCKER, S. W. Region growing: childhood and adolescence. *Computer Graphics and Image Processing*, 15, 382-399, 1976.

APÊNDICE I

Fluxograma: Métodos e Algoritmos utilizados para a Detecção de Óleo e Assemelhados

APÊNDICE II

Tabela 1: Métricas calculadas para o teste efetuado com todos os filtros testados na etapa de pré-processamento.

Nº Polígonos do <i>Phantom</i> : 34	Filtros Avaliados	Métricas no Nível do Pixel				Métricas no Nível das Regiões						
		Tempo de Processamento	Exatidão: AEscuros	Intersecção: AEscuros	Omissão: AEscuros	Inclusão: AEscuros	Intersecção de Polígonos	Taxa de Int. de Polígonos	G forma	Fitxy	Fiti	Fitn
Le 3x3	0,83	0,75	0,94	0,06	0,24	10	0,29	0,92	0,82	0,81	0,96	0,28
Medn 3x3	0,76	0,76	0,94	0,06	0,24	10	0,29	0,91	0,82	0,81	0,97	0,28
Sig 3x3	0,89	0,76	0,94	0,06	0,24	10	0,29	0,88	0,82	0,82	0,92	0,29
Kua 3x3	0,75	0,76	0,93	0,07	0,24	10	0,29	0,90	0,80	0,82	0,91	0,30
Sem Filtro	1,36	0,80	0,88	0,12	0,09	7	0,21	0,77	0,93	0,90	0,33	0,72
Medn 7x7	0,34	0,79	0,91	0,09	0,16	6	0,18	0,92	0,78	0,74	0,11	0,96
Sig 5x5	0,45	0,77	0,92	0,08	0,19	6	0,18	0,93	0,78	0,76	0,10	0,96
Kua 5x5	0,38	0,76	0,92	0,08	0,20	6	0,18	0,92	0,78	0,74	0,11	0,96
Kua 7x7	0,60	0,77	0,91	0,09	0,19	6	0,18	0,92	0,78	0,75	0,11	0,96
G 3x3 1L	0,42	0,76	0,91	0,09	0,21	6	0,18	0,93	0,78	0,76	0,09	0,97
G 3x3 3L	0,42	0,75	0,91	0,09	0,22	6	0,18	0,93	0,78	0,76	0,09	0,97
G 5x5 1L	0,37	0,77	0,91	0,09	0,18	6	0,18	0,93	0,78	0,75	0,10	0,97
ELe 3x3	0,42	0,75	0,91	0,09	0,21	6	0,18	0,93	0,78	0,76	0,09	0,97
ELe 5x5	0,38	0,77	0,92	0,08	0,19	6	0,18	0,92	0,78	0,75	0,11	0,96
Medn 5x5	0,35	0,77	0,91	0,09	0,19	5	0,15	0,93	0,77	0,73	0,09	0,98
Le 5x5	0,43	0,75	0,92	0,08	0,23	5	0,15	0,94	0,77	0,77	0,08	0,98
Le 7x7	0,41	0,76	0,91	0,09	0,20	5	0,15	0,93	0,77	0,77	0,09	0,97
Sig 7x7	0,45	0,77	0,90	0,10	0,18	5	0,15	0,94	0,77	0,76	0,08	0,98
G 5x5 3L	0,38	0,78	0,92	0,08	0,17	5	0,15	0,93	0,77	0,75	0,09	0,97
G 7x7 1L	0,35	0,77	0,90	0,10	0,16	5	0,15	0,93	0,77	0,75	0,09	0,97
G 7x7 3L	0,36	0,77	0,91	0,09	0,17	5	0,15	0,93	0,77	0,75	0,09	0,97
Fr 3x3	0,45	0,75	0,91	0,09	0,21	5	0,15	0,93	0,77	0,76	0,09	0,97
Fr 5x5	0,47	0,75	0,90	0,10	0,20	5	0,15	0,94	0,77	0,74	0,08	0,98
Fr 7x7	0,50	0,77	0,91	0,09	0,18	5	0,15	0,93	0,78	0,73	0,10	0,97
EFr 3x3	0,50	0,75	0,92	0,08	0,22	5	0,15	0,93	0,78	0,73	0,10	0,97
EFr 7x7	0,46	0,77	0,90	0,10	0,17	5	0,15	0,94	0,77	0,74	0,08	0,99
ELe 7x7	0,35	0,78	0,90	0,10	0,15	4	0,12	0,94	0,77	0,75	0,08	0,99
EFr 5x5	0,44	0,75	0,91	0,09	0,21	4	0,12	0,93	0,77	0,76	0,09	0,97

APÊNDICE III

Análise de Matriz de Correlação dos Atributos em Relação ao *Phantom*

(A): Atributos Espectrais - Estatísticos

Tabela 1 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 1

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 1: 06/07/01) Correlações Significativas com o <i>Phantom</i> ($p < .05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	-0.1013*	0.9991*	0.9784*	0.9814*	0.2746*	0.08652*	0.9952*	-0.6456*	
Desvio Padrão	-0.1013*	1,000	-0.1015*	0.03937*	-0.2277*	0.03312*	0.1168*	-0.1054*	0.2632*	
Mediana	0.9991*	-0.1015*	1,000	0.9762*	0.9795*	0.2564*	0.08399*	0.9964*	-0.6443*	
Maximo	0.9784*	0.03937*	0.9762*	1,000	0.9341*	0.3304*	0.1601*	0.9708*	-0.6098*	
Minimo	0.9814*	-0.2277*	0.9795*	0.9341*	1,000	0.3183*	0.02307	0.9747*	-0.6508*	
Assimetria	0.2746*	0.03312*	0.2564*	0.3304*	0.3183*	1,000	0.1568*	0.2312*	-0.08167*	
Curtosi	0.08652*	0.1168*	0.08399*	0.1601*	0.02307	0.1568*	1,000	0.08587*	-0.07862*	
Moda	0.9952*	-0.1054*	0.9964*	0.9708*	0.9747*	0.2312*	0.08587*	1,000	-0.6419*	
Classe Phantom	-0.6456*	0.2632*	-0.6443*	-0.6098*	-0.6508*	0.08167*	0.07862*	-0.6419*	1,000	
Ranking	2	6	3	5	1	7	8	4		

Tabela 2 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 2

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 2: 06/07/02) Correlações Significativas com o <i>Phantom</i> ($p < .05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	-0.1301*	0.9715*	0.7848*	0.8078*	0.234*	0.1892*	0.8733*	-0.5175*	
Desvio Padrão	-0.1301*	1,000	-0.178*	0.2371*	-0.4721*	0.005999	-0.1392*	-0.2698*	0.3629*	
Mediana	0.9715*	-0.178*	1,000	0.7333*	0.7997*	0.1595*	0.1631*	0.8905*	-0.5247*	
Maximo	0.7848*	0.2371*	0.7333*	1,000	0.5117*	0.4878*	0.4223*	0.6225*	-0.5139*	
Minimo	0.8078*	-0.4721*	0.7997*	0.5117*	1,000	0.373*	0.1565*	0.7514*	-0.5259*	
Assimetria	0.234*	0.005999	0.1595*	0.4878*	0.373*	1,000	0.6246*	0.1013	-0.2605*	
Curtosi	0.1892*	-0.1392*	0.1631*	0.4223*	0.1565*	0.6246*	1,000	0.2058*	-0.2638*	
Moda	0.8733*	-0.2698*	0.8905*	0.6225*	0.7514*	0.1013	0.2058*	1,000	-0.5222*	
Classe Phantom	-0.5175*	0.3629*	-0.5247*	-0.5139*	-0.5259*	-0.2605*	-0.2638*	-0.5222*	1,000	
Ranking	4	6	2	5	1	8	7	3		

Tabela 3 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 3

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 3: 06/03/06) Correlações Significativas com o <i>Phantom</i> ($p < .05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.1274	0.9924*	0.8994*	0.9368*	0.3891*	0.266*	0.9592*	-0.6928*	
Desvio Padrão	0.1274	1,000	0.09029	0.3909*	-0.09064	0.2407*	0.1946*	0.05178	0.2068*	
Mediana	0.9924*	0.09029	1,000	0.8717*	0.9386*	0.3332*	0.2295*	0.9681*	-0.6938*	
Maximo	0.8994*	0.3909*	0.8717*	1,000	0.775*	0.5273*	0.4472*	0.8391*	-0.6375*	
Minimo	0.9368*	-0.09064	0.9386*	0.775*	1,000	0.4167*	1,000	0.5741*	0.9321*	-0.6857*
Assimetria	0.3891*	0.2407*	0.3332*	0.5273*	0.4167*	1,000	0.2994*	0.2994*	-0.198*	
Curtosi	0.266*	0.1946*	0.2295*	0.4472*	0.1624	0.5741*	1,000	0.2037*	-0.2726*	
Moda	0.9592*	0.05178	0.9681*	0.8391*	0.9321*	0.2994*	0.2037*	1,000	-0.6838*	
Classe Phantom	-0.6928*	0.2068*	-0.6938*	-0.6375*	-0.6857*	-0.198*	-0.2726*	-0.6838*	1,000	
Ranking	2	7	1	5	3	8	6	4		

Tabela 4 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 4

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 4: 09/09/01) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.3607*	0.9933*	0.9447*	0.9595*	0.4687*	0.1029*	0.9795*	-0.3308*	
Desvio Padrão	0.3607*	1,000	0.3346*	0.5402*	0.1769*	0.342*	0.2457*	0.3147*	-0.003535	
Mediana	0.9933*	0.3346*	1,000	0.9293*	0.9651*	0.4309*	0.07546*	0.9893*	-0.3323*	
Maximo	0.9447*	0.5402*	0.9293*	1,000	0.861*	0.5884*	0.2506*	0.9115*	-0.2883*	
Minimo	0.9595*	0.1769*	0.9651*	0.861*	1,000	0.4736*	-0.01146	0.9568*	-0.3324*	
Assimetria	0.4687*	0.342*	0.4309*	0.5884*	0.4736*	1,000	0.2972*	0.384*	0.009045	
Curtosi	0.1029*	0.2457*	0.07546*	0.2506*	-0.01146	0.2972*	1,000	0.07154*	-0.05319*	
Moda	0.9795*	0.3147*	0.9893*	0.9115*	0.9568*	0.384*	0.07154*	1,000	-0.3294*	
Classe Phantom	-0.3308*	0.003535	-0.3323*	-0.2883*	-0.3324*	0.009045	0.05319*	-0.3294*	1,000	
Ranking	3	8	2	5	1	7	6	4		

Tabela 5 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 5

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 5: 13/06/08) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.6294*	0.9975*	0.9493*	0.9442*	0.4211*	0.1119*	0.9821*	-0.3326*	
Desvio Padrão	0.6294*	1,000	0.6171*	0.7898*	0.4359*	0.3634*	0.2436*	0.5948*	-0.1637*	
Mediana	0.9975*	0.6171*	1,000	0.9398*	0.9409*	0.3853*	0.102*	0.9876*	-0.3342*	
Maximo	0.9493*	0.7698*	0.9398*	1,000	0.8439*	0.542*	0.2732*	0.9198*	-0.3054*	
Minimo	0.9442*	0.4359*	0.9409*	0.8439*	1,000	0.4747*	0.005298	0.9228*	-0.3246*	
Assimetria	0.4211*	0.3634*	0.3863*	0.542*	0.4747*	1,000	0.3799*	0.3419*	-0.02625	
Curtosi	0.1119*	0.2436*	0.102*	0.2732*	0.005298	0.3799*	1,000	0.1087*	-0.04689*	
Moda	0.9821*	0.5948*	0.9876*	0.9198*	0.9228*	0.3419*	0.1087*	1,000	-0.3369*	
Classe Phantom	-0.3326*	-0.1637*	-0.3342*	-0.3054*	-0.3246*	-0.02625	0.04689*	-0.3369*	1,000	
Ranking	3	6	2	5	4	8	7	1		

Tabela 6 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 6

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 6: 05/05/02) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	-0.08127	0.9958*	0.9301*	0.9592*	0.4887*	0.136	0.9826*	-0.6103*	
Desvio Padrão	-0.08127	1,000	-0.08815	0.09926	-0.2204	0.291*	0.09831	-0.1179	0.5661*	
Mediana	0.9958*	-0.08815	1,000	0.9266*	0.9533*	0.4657*	0.1346	0.9851*	-0.6066*	
Maximo	0.9301*	0.09926	0.9266*	1,000	0.8252*	0.6071*	0.2563*	0.9047*	-0.548*	
Minimo	0.9592*	-0.2204	0.9533*	0.8252*	1,000	0.4114*	0.007613	0.9518*	-0.6171*	
Assimetria	0.4887*	0.291*	0.4657*	0.6071*	0.4114*	1,000	0.3139*	0.4381*	-0.08852	
Curtosi	0.136	0.09831	0.1346	0.2563*	0.007613	0.3139*	1,000	0.1271	-0.2073	
Moda	0.9826*	-0.1179	0.9851*	0.9047*	0.9518*	0.4381*	0.1271	1,000	-0.6095*	
Classe Phantom	-0.6103*	0.5661*	-0.6066*	-0.548*	-0.6171*	-0.08852	-0.2073	-0.6095*	1,000	
Ranking	2	5	4	6	1	8	7	3		

Tabela 7 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 7

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 7: 19/12/01) Correlações Significativas com o Phantom ($p < ,05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.1075*	0.9912*	0.9056*	0.8225*	-0.02559	0.2809*	0.8901*	-0.05068	
Desvio Padrão	0.1075*	1,000	0.09658*	0.3499*	-0.2476*	0.08962*	-0.02615	0.02256	0.1218*	
Mediana	0.9912*	0.09658*	1,000	0.8838*	0.7963*	-0.104*	0.265*	0.9125*	-0.1006*	
Maximo	0.9056*	0.3499*	0.8838*	1,000	0.6272*	0.1533*	0.3728*	0.7741*	-0.01522	
Minimo	0.8225*	-0.2476*	0.7963*	0.6272*	1,000	0.1065*	0.1287*	0.6939*	0.07788	
Assimetria	-0.02559	0.08962*	-0.104*	0.1533*	0.1065*	1,000	0.1858*	-0.2158*	0.2887*	
Curtosi	0.2809*	-0.02615	0.265*	0.3728*	0.1287*	0.1858*	1,000	0.28*	-0.09038*	
Moda	0.8901*	0.02256	0.9125*	0.7741*	0.6939*	-0.2158*	0.28*	1,000	-0.2171*	
Classe Phantom	-0.05068	0.1218*	-0.1006*	-0.01522	0.07788	0.2887*	0.09038*	-0.2171*	1,000	
Ranking	7	3	4	8	6	1	5	2		

Tabela 8 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 8

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 8: 03/04/02) Correlações Significativas com o Phantom ($p < ,05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.0764*	0.9956*	0.9284*	0.9705*	0.3096*	0.1375*	0.9886*	0.6792*	
Desvio Padrão	0.0764*	1,000	0.05788	0.2896*	-0.1067*	0.211*	0.2083*	0.04085	-0.08605*	
Mediana	0.9956*	0.05788	1,000	0.9154*	0.9736*	0.2802*	0.1195*	0.9947*	0.6847*	
Maximo	0.9284*	0.2896*	0.9154*	1,000	0.8476*	0.4469*	0.2876*	0.9041*	0.5971*	
Minimo	0.9705*	-0.1067*	0.9736*	0.8476*	1,000	0.3036*	0.04058	0.9695*	0.6751*	
Assimetria	0.3096*	0.211*	0.2802*	0.4469*	0.3036*	1,000	0.252*	0.2496*	-0.004693	
Curtosi	0.1375*	0.2083*	0.1195*	0.2876*	0.04058	0.252*	1,000	0.1174*	0.09213*	
Moda	0.9886*	0.04085	0.9947*	0.9041*	0.9695*	0.2496*	0.1174*	1,000	0.6834*	
Classe Phantom	0.6792*	0.08605*	0.6847*	0.5971*	0.6751*	0.004693	0.09213*	0.6834*	1,000	
Ranking	3	7	1	5	4	8	6	2		

Tabela 9 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 9

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 9: 04/07/01) Correlações Significativas com o Phantom ($p < ,05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.2008*	0.9941*	0.93*	0.9417*	0.5394*	0.1079	0.9732*	-0.5079*	
Desvio Padrão	0.2008*	1,000	0.1773*	0.4212*	0.005209	0.3693*	0.1424*	0.1589*	0.3258*	
Mediana	0.9941*	0.1773*	1,000	0.9179*	0.9429*	0.5109*	0.09632	0.9791*	-0.5118*	
Maximo	0.93*	0.4212*	0.9179*	1,000	0.8209*	0.645*	0.2423*	0.8969*	-0.4329*	
Minimo	0.9417*	0.005209	0.9429*	0.8209*	1,000	0.5438*	0.006644	0.9277*	-0.5085*	
Assimetria	0.5394*	0.3693*	0.5109*	0.645*	0.5438*	1,000	0.3006*	0.4768*	-0.01858	
Curtosi	0.1079	0.1424*	0.09632	0.2423*	0.006644	0.3006*	1,000	0.08672	-0.1593*	
Moda	0.9732*	0.1589*	0.9791*	0.8969*	0.9277*	0.4768*	0.08672	1,000	-0.5126*	
Classe Phantom	-0.5079*	0.3258*	-0.5118*	-0.4329*	-0.5085*	-0.01858	-0.1593*	-0.5126*	1,000	
Ranking	4	6	2	5	3	8	7	1		

Tabela 10 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 10

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 10: 11/03/05) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.7293*	0.9941*	0.93*	0.9417*	0.5394*	0.1079	0.9732*	-0.5079*	
Desvio Padrão	0.7293*	1,000	0.1773*	0.4212*	0.005209	0.3693*	0.1424*	0.1589*	0.3258*	
Mediana	0.9972*	0.721*	1,000	0.9112*	0.9183*	0.2721*	0.1719*	0.9853*	-0.5185*	
Maximo	0.9217*	0.8632*	0.9112*	1,000	0.7885*	0.5004*	0.4212*	0.8949*	-0.4573*	
Minimo	0.9245*	0.5185*	0.9183*	0.7885*	1,000	0.3826*	0.09103*	0.8975*	-0.4941*	
Assimetria	0.3117*	0.3814*	0.2721*	0.5004*	0.3826*	1,000	0.6042*	0.2414*	-0.0741*	
Curtosi	0.1903*	0.3245*	0.1719*	0.4212*	0.09103*	0.6042*	1,000	0.1733*	-0.1455*	
Moda	0.9806*	0.706*	0.9853*	0.8949*	0.8975*	0.2414*	0.1733*	1,000	-0.5166*	
Classe Phantom	-0.5154*	-0.2751*	-0.5185*	-0.4573*	-0.4941*	-0.0741*	-0.1455*	-0.5166*	1,000	
Ranking	3	6	1	5	4	8	7	2		

Tabela 11 – Matriz de Correlação dos Atributos Estatísticos em relação ao atributo da Classe do Phantom: Recorte de Cena 11

Matriz de Correlação dos Atributos Estatísticos em Relação a Classe do Phantom (Cena 11: 14/06/08) Correlações Significativas com o Phantom ($p < .05$) estão										
Atributos	Media	Desvio Padrão	Mediana	Maximo	Minimo	Assimetria	Curtosi	Moda	Classe Phantom	
Media	1,000	0.7261*	0.9972*	0.9217*	0.9245*	0.3117*	0.1903*	0.9806*	-0.5154*	
Desvio Padrão	0.7261*	1,000	0.721*	0.8632*	0.5185*	0.3814*	0.3245*	0.706*	-0.2751*	
Mediana	0.9948*	0.7031*	1,000	0.8808*	0.9138*	0.2756*	0.1816*	0.9883*	-0.557*	
Maximo	0.8978*	0.8761*	0.8808*	1,000	0.7532*	0.5506*	0.466*	0.8693*	-0.5037*	
Minimo	0.9187*	0.527*	0.9138*	0.7532*	1,000	0.3643*	0.09824*	0.8997*	-0.55*	
Assimetria	0.3166*	0.4687*	0.2756*	0.5506*	0.3643*	1,000	0.6835*	0.2599*	-0.2556*	
Curtosi	0.2005*	0.3681*	0.1816*	0.466*	0.09824*	0.6835*	1,000	0.19*	-0.2536*	
Moda	0.981*	0.6848*	0.9883*	0.8693*	0.8997*	0.2599*	0.19*	1,000	-0.5528*	
Classe Phantom	-0.5585*	-0.3267*	-0.557*	-0.5037*	-0.55*	-0.2556*	-0.2536*	-0.5528*	1,000	
Ranking	1	6	2	5	4	7	8	3		

(B): Atributos de Forma - Genéricos e Poligonais

Tabela 1 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 1

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 1: 06/07/01) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	0.0489*	0.9313*	0.5392*	0.1273*	-0.4234*	0.8782*	-0.4093*	0.6888*	0.8711*	0.5078*	-0.6586*	0.5976*	0.1009*
Assimetria	0.0489*	1,000	0.1538*	0.1304*	-0.7479*	-0.1036*	0.4075*	-0.1379*	0.2412*	-0.2839*	0.1493*	-0.2471*	0.2254*	0.2282*
Comprimento de Borda	0.9313*	0.1538*	1,000	0.7458*	-0.1223*	-0.644*	0.9359*	-0.6415*	0.8955*	0.6469*	0.4036*	-0.8626*	0.5731*	0.116*
Compação	0.5392*	0.1304*	0.7458*	1,000	-0.4203*	-0.8697*	0.6784*	-0.9001*	0.8732*	0.6581*	0.1375*	-0.8581*	0.3346*	0.08913*
Densidade	0.1273*	-0.7479*	-0.1223*	-0.4203*	1,000	0.4815*	-0.2602*	0.5187*	-0.4074*	0.2414*	0.03963*	0.4264*	-0.0933*	-0.1979*
Ajuste Elíptico	-0.4234*	-0.1036*	-0.644*	-0.8697*	0.4815*	1,000	-0.5332*	0.9682*	-0.813*	-0.5608*	-0.0768*	0.8171*	-0.2606*	-0.03811*
Comprimento	0.8782*	0.4075*	0.9359*	0.6784*	-0.2602*	-0.5332*	1,000	-0.5471*	0.8301*	0.6711*	0.4732*	-0.8066*	0.6341*	0.1914*
Ajuste Retangular	-0.4093*	-0.1379*	-0.6415*	-0.9001*	0.5187*	0.9582*	-0.5471*	1,000	-0.8253*	-0.5413*	-0.07122*	0.8285*	-0.2629*	-0.04711*
Índice de Forma	0.6888*	0.2412*	0.8955*	0.8732*	-0.4074*	-0.813*	0.8301*	-0.8253*	1,000	0.6739*	0.2236*	-0.9717*	0.446*	0.1088*
Largura	0.8711*	-0.2839*	0.8469*	0.6581*	0.2414*	-0.5608*	0.6711*	-0.5413*	0.6739*	1,000	0.3528*	-0.6463*	0.4407*	-0.003691
μ dos Comprimentos de Borda	0.5078*	0.1493*	0.4036*	0.1375*	0.03963*	-0.0768*	0.4732*	-0.07122*	0.2236*	0.3528*	1,000	-0.193*	0.7436*	0.1795*
Compactação (Polígono)	0.6586*	-0.2471*	-0.8628*	-0.8881*	0.4264*	0.8171*	-0.8065*	0.8285*	-0.9717*	-0.6463*	-0.193*	1,000	-0.4376*	-0.1219*
Comprimento da maior Borda	0.5976*	0.2254*	0.5731*	0.3346*	-0.0933*	-0.2606*	0.6341*	-0.2629*	0.446*	0.4407*	0.7436*	-0.4376*	1,000	0.2087*
Classe Phantom	0.1009*	0.2282*	0.116*	0.08913*	-0.1979*	-0.03811*	0.1914*	-0.04711*	0.1088*	-0.003691	0.1795*	-0.1219*	1,000	
Ranking	9	1	7	10	3	12	4	11	8	13	5	6	2	

Tabela 2 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 2

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 2: 06.07.02)															
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho															
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Comprimento da menor Borda	Classe Phantom
Área	1,000	0,07007	0,8766*	0,4526*	0,08253	-0,3606*	0,7867*	-0,3609*	0,6021*	0,7677*	0,3442*	-0,5909*	0,4766*	0,1424*	
Assimetria	0,07007	1,000	0,2747*	0,2397*	-0,7629*	-0,1493*	0,5415*	-0,2053*	0,4002*	-0,364*	0,2536*	-0,3789*	0,3427*	0,483*	
Comprimento de Borda	0,8766*	0,2747*	1,000	0,6991*	-0,2807*	-0,6008*	0,9157*	-0,6187*	0,8996*	0,6591*	0,3447*	-0,877*	0,5819*	0,2921*	
Compactação	0,4526*	0,2397*	0,6991*	1,000	-0,5363*	-0,8623*	0,6106*	-0,8893*	0,8097*	0,5208*	0,1172*	-0,7963*	0,3591*	0,2658*	
Densidade	0,08253	-0,7629*	-0,2807*	-0,5363*	1,000	0,5163*	-0,4088*	0,5511*	-0,5681*	0,27*	-0,1465*	0,5648*	-0,3024*	-0,4578*	
Ajuste Elíptico	0,3606*	-0,1493*	-0,6008*	-0,8623*	0,5163*	1,000	-0,4474*	0,9541*	-0,7249*	-0,5009*	-0,07125	0,7298*	-0,2649*	-0,1598*	
Comprimento	0,7867*	0,5415*	0,9157*	0,6106*	0,4088*	-0,4474*	1,000	-0,4868*	0,8301*	0,4004*	0,4035*	-0,8059*	0,6172*	0,4121*	
Ajuste Retangular	0,3609*	-0,2053*	-0,6187*	-0,8983*	0,5511*	0,9541*	-0,4868*	1,000	-0,7543*	-0,4782*	-0,08031	0,7562*	-0,2837*	-0,2125*	
Índice de Forma	0,6021*	0,4002*	0,8998*	0,8097*	-0,5681*	-0,7249*	0,8301*	-0,7543*	1,000	0,4433*	0,2632*	-0,9752*	0,5321*	0,3667*	
Largura	0,7677*	-0,364*	0,6591*	0,5208*	0,27*	-0,5009*	0,4004*	-0,4782*	0,4433*	1,000	0,1457*	-0,4395*	0,2541*	-0,1528*	
μ dos Comprimentos de Borda	0,3442*	0,2536*	0,3447*	0,1172*	-0,1465*	-0,07125	0,4035*	-0,08031	0,2632*	0,1457*	1,000	-0,2253*	0,6708*	0,2756*	
Compactação (Polígono)	0,5909*	-0,3789*	-0,877*	-0,7963*	0,5648*	0,7298*	-0,8059*	0,7562*	-0,9752*	-0,4395*	-0,2253*	1,000	-0,5121*	-0,3555*	
Comprimento da maior Borda	0,4766*	0,3427*	0,5819*	0,3591*	-0,3024*	-0,2649*	0,6172*	-0,2837*	0,5321*	0,2541*	0,6708*	-0,5121*	1,000	0,3452*	
Classe Phantom	0,1424*	0,483*	0,2921*	0,2658*	-0,4578*	-0,1598*	0,4121*	-0,2125*	0,3667*	-0,1528*	0,3452*	1,000			
Ranking	13	1	7	8	2	11	3	10	4	12	9	5	6		

Tabela 3 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 3

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 3: 06.03.06)															
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho															
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Comprimento da menor Borda	Classe Phantom
Área	1,000	-0,04648	0,9102*	0,4082*	0,1913*	-0,2681*	0,8373*	-0,264*	0,6525*	0,8579*	0,4957*	-0,5967*	0,5597*	0,384*	
Assimetria	0,04648	1,000	0,08911	0,2418*	-0,7401*	-0,1698*	0,3603*	-0,23*	0,2*	-0,3658*	-0,01011	-0,2143*	0,1017	0,2658*	
Comprimento de Borda	0,9102*	0,08911	1,000	0,6569*	-0,1014	0,5355*	0,9184*	-0,5506*	0,8936*	0,8082*	0,3686*	-0,85*	0,5642*	0,4166*	
Compactação	0,4082*	0,2418*	0,6569*	1,000	-0,5574*	-0,8219*	0,6172*	-0,9019*	0,8033*	0,4595*	0,05144	-0,8034*	0,2899*	0,3376*	
Densidade	0,1913*	-0,7401*	-0,1014	-0,5574*	1,000	0,5952*	-0,2306*	0,6442*	-0,3962*	0,2944*	0,1627	0,4312*	-0,05	-0,2718*	
Ajuste Elíptico	0,2681*	-0,1698*	-0,5355*	-0,8219*	0,5952*	1,000	-0,4224*	1,000	-0,4699*	0,8215*	0,594*	0,4156*	-0,7689*	0,5993*	0,4611*
Comprimento	0,8373*	0,3603*	0,9184*	0,6172*	-0,2306*	-0,4224*	1,000	-0,4699*	1,000	-0,7418*	-0,3525*	0,01101	0,7615*	-0,2409*	-0,2789*
Ajuste Retangular	-0,264*	-0,23*	-0,5506*	-0,9019*	0,6442*	0,944*	-0,4699*	1,000	-0,7418*	-0,4718*	1,000	0,608*	-0,9786*	0,4725*	0,3272*
Índice de Forma	0,6525*	0,2*	0,8936*	0,8033*	-0,3962*	-0,7187*	0,8215*	-0,7418*	1,000	0,608*	0,1807	-0,9786*	0,4725*	0,3272*	
Largura	0,8579*	-0,3658*	0,8082*	0,4595*	0,2944*	-0,3699*	0,594*	-0,3525*	0,608*	1,000	0,3815*	-0,5618*	0,4294*	0,2656*	
μ dos Comprimentos de Borda	0,4957*	-0,01011	0,3686*	0,05144	0,1627	-0,008808	0,4156*	0,01101	0,1807	0,3815*	1,000	-0,1627	0,7059*	0,3569*	
Compactação (Polígono)	0,5967*	-0,2143*	-0,85*	-0,8034*	0,4312*	0,732*	-0,7889*	0,7615*	-0,9786*	-0,5618*	-0,1627	1,000	-0,4813*	-0,3247*	
Comprimento da maior Borda	0,5597*	0,1017	0,5642*	0,2899*	-0,05	-0,2569*	0,5993*	-0,2409*	0,4725*	0,4294*	0,7059*	-0,4813*	1,000	0,3658*	
Classe Phantom	0,384*	0,2658*	0,4166*	0,3376*	-0,2718*	-0,2678*	0,4611*	-0,2789*	0,3272*	0,2656*	0,3569*	-0,3247*	0,3658*	1,000	
Ranking	3	12	2	6	10	11	1	9	7	13	5	8	4		

Tabela 4 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 4

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 4 09.09.01)															
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho															
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Comprimento da menor Borda	Classe Phantom
Área	1,000	-0,04703	0,9332*	0,5357*	0,2561*	-0,4052*	0,8818*	-0,3969*	0,712*	0,8916*	0,5101*	-0,6759*	0,5718*	0,1128*	
Assimetria	0,04703	1,000	0,03733	0,6648*	-0,7005*	-0,07968*	0,2951*	-0,1063*	0,1181*	-0,3295*	0,0572*	-0,1291*	0,127*	0,08577*	
Comprimento de Borda	0,9332*	0,03733	1,000	0,733*	1,000	-0,3579*	-0,8473*	0,678*	-0,8777*	0,8846*	0,6618*	0,13*	-0,8621*	0,3146*	0,08723*
Compactação	0,5357*	0,6648*	0,733*	1,000	-0,3579*	1,000	0,473*	-0,1131*	0,5012*	-0,2576*	0,3091*	0,1439*	0,2916*	0,02005	-0,0498
Densidade	0,2561*	-0,7005*	0,0159	-0,3579*	1,000	0,473*	-0,1131*	0,5012*	-0,2576*	-0,5397*	-0,08298*	0,7935*	-0,244*	-0,06752*	
Ajuste Elíptico	0,4052*	-0,07968*	-0,6196*	-0,8473*	0,473*	1,000	-0,5196*	0,9564*	-0,777*	-0,5397*	-0,08298*	0,7935*	-0,244*	-0,06752*	
Comprimento	0,8818*	0,2951*	0,9318*	0,678*	-0,1131*	-0,5196*	1,000	-0,5344*	1,000	-0,7909*	-0,526*	-0,0886*	0,8085*	-0,2309*	0,5849*
Ajuste Retangular	0,3969*	-0,1063*	-0,6209*	-0,8777*	0,5012*	0,9554*	-0,5344*	1,000	-0,7909*	1,000	0,7218*	0,2035*	-0,9734*	0,4111*	0,1272*
Índice de Forma	0,712*	0,1181*	0,9076*	0,8546*	-0,2576*	-0,777*	0,834*	-0,7909*	1,000	0,7218*	0,2035*	-0,9734*	0,4111*	0,1272*	
Largura	0,8916*	-0,3295*	0,8743*	0,6618*	0,3091*	-0,5397*	0,7114*	-0,526*	0,7218*	1,000	0,3841*	-0,6954*	0,4511*	0,09374*	
μ dos Comprimentos de Borda	0,5101*	0,0572*	0,3907*	0,13*	0,1439*	-0,08298*	0,4453*	-0,0886*	0,2035*	0,3841*	1,000	-0,1649*	1,000	-0,3905*	-0,1171*
Compactação (Polígono)	0,6759*	-0,1291*	-0,8718*	-0,8521*	0,2916*	0,7935*	-0,8053*	0,8085*	-0,9734*	-0,6954*	-0,1649*	1,000	-0,3905*	-0,1171*	
Comprimento da maior Borda	0,5718*	0,127*	0,5343*	0,3146*	0,02005	-0,244*	0,5849*	-0,2309*	0,4111*	0,4511*	0,7243*	-0,3905*	1,000	0,1122*	
Classe Phantom	0,1128*	0,06577*	0,127*	0,08723*	-0,0498*	-0,06752*	0,1328*	0,06516*	0,1272*	0,09374*	0,1075*	-0,1171*	1,000	0,1122*	
Ranking	5	11	3	9	13	10	1	12	2	8	7	4	6		

Tabela 5 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 5

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 5: 13/06/08)														
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	-0.0651*	0.9301*	0.5512*	0.1968*	-0.4534*	0.8506*	-0.4428*	0.6963*	0.8679*	0.4679*	-0.6605*	0.5573*	0.08485*
Assimetria	-0.0651*	1,000	0.03523*	0.06816*	-0.7506*	-0.04931*	0.3394*	-0.07793*	0.1387*	-0.3936*	0.1016*	-0.1483*	0.1593*	-0.003589
Comprimento de Borda	0.9301*	0.03523*	1,000	0.7483*	0.04667*	-0.6559*	0.9104*	-0.6589*	0.9018*	0.8417*	0.3643*	-0.8639*	0.5383*	0.08095*
Compactação	0.5512*	0.06816*	0.7483*	1,000	-0.3601*	-0.8265*	0.6835*	-0.6884*	0.8619*	0.6497*	0.1297*	-0.8432*	0.324*	0.05563*
Densidade	0.1968*	-0.7506*	-0.04667*	-0.3801*	1,000	0.427*	-0.2268*	0.4618*	-0.3242*	0.3191*	0.05317*	0.3499*	0.07125*	0.01575
Ajuste Elíptico	0.4534*	-0.04931*	-0.6559*	-0.8285*	0.427*	1,000	-0.531*	0.9515*	-0.7966*	-0.5641*	-0.1049*	0.8043*	-0.2778*	-0.03095*
Comprimento	0.8506*	0.3394*	0.9104*	0.6883*	-0.2268*	-0.531*	1,000	-0.5514*	0.6208*	0.6181*	0.442*	-0.7934*	0.6096*	0.07964*
Ajuste Retangular	0.4428*	-0.07793*	-0.6589*	-0.8884*	0.4618*	0.9515*	-0.5514*	1,000	-0.8148*	-0.5488*	-0.09446*	0.8218*	-0.2724*	-0.03125*
Índice de Forma	0.8963*	0.1387*	0.9018*	0.8619*	-0.3242*	-0.7966*	0.8208*	-0.8149*	1,000	0.673*	0.2006*	-0.9702*	0.4338*	0.06309*
Largura	0.8679*	-0.3936*	0.8417*	0.3191*	-0.5641*	0.6181*	-0.5488*	0.673*	-0.3049*	1,000	-0.1712*	0.7072*	0.07541*	
μ dos Comprimentos de Borda	0.4679*	0.1018*	0.3643*	0.1297*	0.05317*	-0.1049*	0.442*	-0.09446*	0.2006*	0.3049*	1,000	-0.1712*	0.7072*	0.07541*
Compactação (Polígono)	0.6605*	-0.1483*	-0.8639*	-0.8432*	0.3499*	0.8043*	-0.7934*	0.8218*	-0.9702*	-0.6389*	-0.1712*	1,000	-0.4312*	-0.05268*
Comprimento da maior Borda	0.5573*	0.1593*	0.5383*	0.324*	0.07125*	-0.2778*	0.6096*	-0.2724*	0.4338*	0.3853*	0.7072*	-0.4312*	1,000	0.03827*
Classe Phantom	0.08485	0.003589	0.08095*	0.05583*	0.01575	-0.03095*	0.07964*	0.03125*	0.06309*	0.08946*	0.07541*	-0.05258*	0.03827*	1,00
Ranking	2	13	3	7	12	11	4	10	6	1	5	8	9	

Tabela 6 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 6

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 6: 05/05/02)														
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	0.0787	0.9247*	0.4712*	0.109	-0.3712*	0.8518*	-0.389*	0.7174*	0.8663*	0.3843*	-0.7018*	0.5095*	0.317*
Assimetria	0.0787	1,000	0.1873	0.2517*	-0.8121*	-0.2774*	0.4203*	-0.2689*	0.2467*	-0.2708*	0.1495	-0.2656*	0.227	0.4962*
Comprimento de Borda	0.9247*	0.1873	1,000	0.6651*	-0.1213	-0.6026*	0.9324*	-0.6126*	0.913*	0.8218*	0.3394*	-0.8879*	0.4986*	0.3354*
Compactação	0.4712*	0.2517*	0.6551*	1,000	-0.4904*	-0.8725*	0.6361*	-0.5465*	0.9597*	-0.7786*	-0.4741*	-0.1483	0.7805*	-0.2566*
Densidade	0.109	-0.8121*	-0.1213	-0.4904*	1,000	0.5459*	-0.2726*	0.5402*	-0.3438*	0.2373	-0.07106	0.3667*	-0.1213	-0.4566*
Ajuste Elíptico	0.3712*	-0.2774*	-0.6026*	-0.8725*	0.5459*	1,000	-0.5465*	0.9597*	-0.7786*	-0.4741*	-0.1483	0.7805*	-0.2566*	
Comprimento	0.8518*	0.4203*	0.9324*	0.6361*	-0.2726*	-0.5465*	1,000	-0.5562*	0.8896*	0.6395*	0.3635*	-0.8413*	0.5645*	0.4102*
Ajuste Retangular	-0.389*	-0.2689*	-0.6128*	-0.8982*	0.5402*	0.9597*	-0.5652*	1,000	-0.7845*	-0.4854*	-0.1632	0.7816*	-0.29*	-0.2562*
Índice de Forma	0.7174*	0.2467*	0.913*	0.7726*	-0.3438*	-0.7786*	0.8598*	-0.7845*	1,000	0.69*	0.2554*	-0.9644*	0.3946*	0.3064*
Largura	0.8663*	-0.2708*	0.8218*	0.5729*	0.2373	-0.4741*	0.6395*	-0.4854*	0.69*	1,000	0.2887*	-0.6689*	0.3322*	0.1958
μ dos Comprimentos de Borda	0.3843*	0.1495	0.3394*	0.08141	0.07106	-0.1483	0.3635*	-0.1632	0.2554*	0.2887*	1,000	-0.2359	0.6942*	0.2399
Compactação (Polígono)	-0.7018*	-0.2656*	-0.8879*	-0.7724*	0.3667*	0.7805*	-0.8413*	0.7816*	-0.9644*	-0.6689*	-0.2359	1,000	-0.4457*	0.316*
Comprimento da maior Borda	0.5095*	0.227	0.4986*	0.2095	-0.1213	-0.2584*	0.5645*	-0.29*	0.3946*	0.3322*	0.6942*	-0.4457*	1,000	0.0316*
Classe Phantom	0.317*	0.4962*	0.3354*	0.3401*	-0.4566*	-0.2656*	0.4102*	-0.2562*	0.3064*	0.1958	0.2399	-0.3401*	0.316*	1,00
Ranking	7	1	6	4	2	10	3	11	9	13	12	5	8	

Tabela 7 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 7

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 7 19/12/01)														
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	0.004198	0.95*	0.5735*	0.2366*	-0.4374*	0.8994*	-0.4202*	0.7476*	0.8937*	0.4337*	-0.7115*	0.576*	-0.04268
Assimetria	0.004198	1,000	0.05976	0.02229	-0.7037*	0.01669	0.3219*	-0.007881	0.1238*	-0.3046*	0.21*	-0.1236*	0.2221*	0.1202*
Comprimento de Borda	0.95*	0.05976	1,000	0.7357*	-0.4023*	-0.6105*	0.9347*	-0.6048*	0.9098*	0.8803*	0.3583*	-0.8715*	0.5506*	-0.01117
Compactação	0.5735*	0.02229	0.7357*	1,000	-0.2675*	-0.8341*	0.6712*	-0.8678*	0.8504*	0.685*	0.1025*	-0.8366*	0.3143*	0.01804
Densidade	0.2366*	-0.7037*	0.04327	-0.2675*	1,000	0.3492*	-0.1046*	0.3899*	-0.2227*	0.3477*	0.02233	0.2348*	-0.04025	-0.1403*
Ajuste Elíptico	0.4374*	0.01669	-0.6105*	-0.8341*	0.3492*	1,000	-0.4895*	0.9432*	-0.7634*	-0.5682*	-0.04625	0.7654*	-0.2463*	-0.01955
Comprimento	0.8994*	0.3219*	0.9347*	0.6712*	-0.1046*	-0.4895*	1,000	-0.4981*	0.8382*	0.7179*	0.4423*	-0.8052*	0.6187*	0.01563
Ajuste Retangular	0.4202*	-0.007881	-0.6048*	-0.8678*	0.3899*	0.9432*	-0.4961*	1,000	-0.7733*	-0.5465*	-0.02298	0.7717*	-0.2229*	-0.0227
Índice de Forma	0.7476*	0.1238*	0.9098*	0.8504*	-0.2227*	-0.7634*	0.8382*	-0.7733*	1,000	0.7373*	0.2277*	-0.9671*	0.4511*	0.0355
Largura	0.8937*	-0.3046*	0.8803*	0.685*	0.3477*	-0.5682*	0.7179*	-0.5465*	0.7373*	1,000	0.2807*	-0.7105*	0.4333*	-0.08723*
μ dos Comprimentos de Borda	0.4337*	0.21*	0.3583*	0.1025*	0.02233	-0.04625	0.4423*	-0.02298	0.2277*	0.2807*	1,000	-0.1813*	0.7208*	-0.02499
Compactação (Polígono)	0.7115*	-0.1236*	-0.8715*	-0.8366*	0.2348*	0.7654*	-0.8052*	0.7717*	-0.9671*	-0.7105*	-0.1813*	1,000	-0.4343*	-0.03775
Comprimento da maior Borda	0.576*	0.2221*	0.5506*	0.3143*	-0.04025	-0.2463*	0.6187*	-0.2229*	0.4511*	0.4333*	0.7208*	-0.4343*	1,000	0.0111
Classe Phantom	0.04266	0.1202*	-0.01117	0.01804	-0.1403*	-0.01955	0.01563	-0.0227	0.0355	-0.08723*	-0.02499	-0.03775	0.0111	1,00
Ranking	4	2	12	10	1	9	11	8	6	3	7	5	13	

Tabela 8 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 8

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 8 03/04/02)														
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	0.0895*	0.9363*	0.5253*	0.1526*	-0.3909*	0.8853*	-0.3699*	0.7189*	0.8717*	0.5463*	-0.6842*	0.6469*	-0.07063
Assimetria	0.08955*	1,000	0.1918*	0.1825*	-0.7525*	-0.1406*	0.4331*	-0.1677*	0.2738*	-0.2508*	0.2022*	-0.2942*	0.2675*	-0.1704*
Comprimento de Borda	0.9363*	0.1918*	1,000	0.7279*	0.08038	-0.603*	0.9425*	-0.5935*	0.9075*	0.8412*	0.4543*	-0.8735*	0.6291*	-0.09499*
Compactação	0.5253*	0.1825*	0.7279*	1,000	-0.43*	-0.8643*	0.6639*	-0.89*	0.8647*	0.6082*	0.154*	-0.8494*	0.3489*	-0.09526*
Densidade	0.1526*	-0.7525*	-0.08038*	-0.43*	1,000	0.4877*	-0.2143*	0.5253*	-0.3454*	0.291*	0.05353	0.375*	-0.05316	0.1356*
Ajuste Elíptico	0.3909*	-0.1406*	-0.603*	-0.8843*	0.4877*	1,000	-0.4998*	0.9635*	-0.774*	-0.5011*	-0.07117	0.782*	-0.24*	0.05924
Comprimento	0.8853*	0.4331*	0.9425*	0.6639*	-0.2143*	-0.4998*	1,000	-0.5029*	0.8504*	0.6723*	0.5365*	-0.8276*	0.6985*	-0.1379*
Ajuste Retangular	0.3699*	-0.1677*	-0.5935*	-0.89*	0.5253*	-0.5029*	1,000	-0.7803*	-0.4766*	-0.06571	0.786*	-0.2356*	0.08007*	
Índice de Forma	0.7189*	0.2738*	0.9075*	0.8647*	-0.3454*	-0.774*	0.8504*	-0.7803*	1,000	0.68*	0.2895*	-0.9713*	0.5126*	-0.09583*
Largura	0.8717*	-0.2508*	0.8412*	0.6082*	0.291*	-0.5011*	0.6723*	-0.4766*	0.68*	1,000	0.3732*	-0.6454*	0.4711*	-0.003394
μ dos Comprimentos de Borda	0.5463*	0.2022*	0.4543*	0.154*	0.05353	-0.07117	0.5365*	-0.06571	0.2895*	0.3732*	1,000	-0.261*	0.7856*	-0.1086*
Compactação (Polígono)	0.6842*	-0.2942*	0.8735*	-0.8494*	0.375*	0.782*	-0.8276*	0.786*	-0.9713*	0.6454*	-0.261*	1,000	0.5058*	0.08639*
Comprimento da maior Borda	0.6469*	0.2675*	0.6291*	0.3489*	-0.05316	-0.24*	0.6985*	-0.2356*	0.5126*	0.4711*	0.7856*	-0.5058*	1,000	-0.071
Classe Phantom	0.07063	-0.1704*	0.08499*	0.09526*	0.1366*	0.05924	-0.1379*	0.08007*	-0.09583*	0.003394	-0.1086*	0.08639*	-0.071	1,00
Ranking	11	1	7	6	3	12	2	9	5	13	4	8	10	

Tabela 9 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 9

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 9 04/07/01)														
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	0.04983	0.9394*	0.4358*	0.2521*	-0.3235*	0.8804*	-0.3005*	0.7035*	0.8769*	0.537*	-0.6929*	0.5631*	0.1641*
Assimetria	0.04983	1,000	0.1187*	0.116*	-0.634*	-0.05848	0.3929*	-0.08625	0.1637*	-0.2631*	0.1722*	-0.1609*	0.1724*	0.2952*
Comprimento de Borda	0.9394*	0.1187*	1,000	0.64*	0.02306	-0.5299*	0.9269*	-0.5204*	0.8919*	0.866*	0.4593*	-0.8781*	0.5683*	0.1791*
Compactação	0.4358*	0.116*	0.64*	1,000	-0.4424*	-0.8086*	0.5665*	-0.845*	0.8089*	0.5576*	0.1416*	-0.7854*	0.3173*	0.0573
Densidade	0.2521*	-0.634*	0.02306	-0.4442*	1,000	0.4954*	-0.09964	0.5398*	-0.2786*	0.3189*	0.126*	0.2703*	0.04314	-0.1298*
Ajuste Elíptico	0.3235*	-0.05848	-0.5299*	-0.8086*	0.4954*	1,000	-0.4009*	0.9566*	-0.7331*	-0.4686*	-0.05156	0.7288*	-0.2006*	0.03583
Comprimento	0.8804*	0.3929*	0.9269*	0.5865*	-0.09964	-0.4009*	1,000	-0.4075*	0.811*	0.6818*	0.5185*	-0.7995*	0.613*	0.2608*
Ajuste Retangular	0.3005*	-0.08625	-0.5204*	-0.845*	0.5398*	0.9566*	-0.4075*	1,000	-0.7408*	-0.4406*	-0.04562	0.731*	-0.1933*	0.01542
Índice de Forma	0.7035*	0.1637*	0.8919*	0.8089*	-0.2786*	-0.7331*	0.811*	-0.7408*	1,000	0.7042*	0.2763*	-0.9696*	0.4559*	0.1056
Largura	0.8769*	-0.2631*	0.866*	0.5576*	0.3189*	-0.4686*	0.6818*	-0.4406*	0.7042*	1,000	0.415*	-0.7004*	0.4813*	0.02072
μ dos Comprimentos de Borda	0.537*	0.1722*	0.4593*	0.1416*	0.126*	-0.05156	0.5185*	-0.04662	0.2763*	0.415*	1,000	-0.2539*	0.6958*	0.2069*
Compactação (Polígono)	0.6929*	-0.1609*	-0.8781*	0.7854*	0.2703*	0.7288*	-0.7995*	0.731*	-0.9696*	-0.7004*	-0.2539*	1,000	-0.4548*	-0.1221*
Comprimento da maior Borda	0.5631*	0.1724*	0.5683*	0.3173*	0.04314	-0.2006*	0.613*	-0.1933*	0.4569*	0.4813*	0.6958*	-0.4548*	1,000	0.2068*
Classe Phantom	0.1641*	0.2952*	0.1791*	0.0573	-0.1298*	0.03583	0.2608*	0.01542	0.1056	0.02072	0.2069*	-0.1221*	0.2068*	1,00
Ranking	6	1	5	10	7	11	2	13	9	12	3	8	4	

Tabela 10 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 10

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 10: 11/03/05)														
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento de Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	0.02349	0.9163*	0.3466*	0.3555*	-0.2021*	0.8682*	-0.1737*	0.6336*	0.8718*	0.5991*	-0.5639*	0.546*	0.1324*
Assimetria	0.02349	1,000	0.1039*	0.1272*	-0.6523*	-0.1401*	0.385*	-0.1823*	0.1702*	-0.267*	0.04877*	-0.17*	0.1333*	0.1626*
Comprimento de Borda	0.9163*	0.1039*	1,000	0.5866*	0.09343*	-0.4757*	0.9088*	-0.4583*	0.8748*	0.8564*	0.4128*	-0.8107*	0.48*	0.08267*
Compactação	0.3466*	0.1272*	0.5866*	1,000	-0.4103*	-0.7993*	0.5382*	-0.8403*	0.7681*	0.516*	0.01765	-0.7611*	0.1922*	-0.01478
Densidade	0.3555*	-0.6523*	0.09343*	-0.4103*	1,000	0.5485*	-0.02994	0.5958*	-0.234*	0.3955*	0.3284*	-0.2746*	0.1189*	-0.0274
Ajuste Elíptico	0.2021*	-0.1401*	-0.4757*	-0.7993*	0.5485*	1,000	-0.3612*	0.9397*	-0.7172*	-0.359*	0.0916*	0.7274*	-0.1019*	0.03869*
Comprimento	0.8682*	0.3857*	0.9088*	0.5382*	-0.02994	-0.3612*	1,000	-0.3687*	0.7587*	0.6755*	0.494*	-0.7014*	0.5408*	0.1597*
Ajuste Retangular	0.1737*	-0.1823*	-0.4583*	-0.8403*	0.5958*	0.9397*	-0.3687*	1,000	-0.7167*	-0.3263*	0.1126*	0.7299*	0.08892*	0.04059*
Índice de Forma	0.6336*	0.1702*	0.8748*	0.7681*	-0.234*	-0.7172*	0.7587*	-0.7167*	1,000	0.6624*	0.1362*	-0.958*	0.3215*	-0.001485
Largura	0.8718*	-0.287*	0.8564*	0.516*	0.3955*	-0.359*	0.6755*	-0.3263*	0.6624*	1,000	0.4518*	-0.6065*	0.4354*	0.03959*
μ dos Comprimentos de Borda	0.5991*	0.04877*	0.4128*	0.01765	0.3284*	0.0916*	0.494*	0.1126*	0.1362*	0.4518*	1,000	-0.05898*	0.7006*	0.2004*
Compactação (Polígono)	-0.5639*	-0.17*	-0.8107*	-0.7681*	0.2746*	0.7274*	-0.7014*	0.7299*	-0.958*	-0.6065*	-0.05898*	1,000	-0.2706*	-0.01016
Comprimento da maior Borda	0.546*	0.1333*	0.48*	0.1922*	0.1189*	-0.1019*	0.5408*	0.08892*	0.3215*	0.4354*	0.7006*	-0.2706*	1,000	0.1717*
Classe Phantom	0.1324*	0.1626*	0.08267*	-0.01478	0.03869*	0.1597*	0.04059*	0.001485	0.03959*	0.02004*	-0.01016	0.1717*	1,00	
Ranking	5	3	6	11	10	9	4	7	13	8	1	12	2	

Tabela 11 – Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em relação ao atributo da Classe do Phantom: Recorte de Cena 11

Matriz de Correlação dos Atributos de Forma (Genéricos e Poligonais) em Relação a Classe do Phantom (Cena 11: 14/06/08) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho														
Atributos	Área	Assimetria	Comprimento da Borda	Compação	Densidade	Ajuste Elíptico	Comprimento	Ajuste Retangular	Índice de Forma	Largura	μ dos Comprimentos de Borda	Compactação (Polígono)	Comprimento da maior Borda	Classe Phantom
Área	1,000	0.1102*	0.9236*	0.3797*	0.2376*	-0.2221*	0.8715*	-0.1974*	0.6598*	0.8284*	0.5832*	-0.6082*	0.5518*	0.1068*
Assimetria	0.1102*	1,000	0.1892*	0.1637*	-0.7095*	-0.1842*	0.468*	-0.2031*	0.2532*	-0.2607*	0.1164*	-0.2611*	0.212*	0.2921*
Comprimento da Borda	0.9236*	0.1892*	1,000	0.5932*	0.091	-0.4757*	0.9079*	-0.4564*	0.8829*	0.8101*	0.4241*	-0.8327*	0.4861*	0.07704*
Compactação	0.3797*	0.1637*	0.5932*	1,000	-0.406*	-0.7791*	0.5338*	-0.8362*	0.7528*	0.5365*	0.09341*	-0.7489*	0.2219*	0.02249
Densidade	0.2378*	-0.7095*	0,091	-0.406*	1,000	0.5344*	-0.1421*	0.5719*	-0.291*	0.3663*	0.2087*	0.3245*	0.03563	-0.2053*
Ajuste Elíptico	-0.2231*	-0.1842*	-0.4757*	-0.7791*	0.5344*	1,000	-0.36*	0.9341*	-0.6982*	-0.3642*	0.04353	0.708*	-0.1069*	0.01801
Comprimento	0.8715*	0.468*	0.9079*	0.5338*	-0.1421*	-0.36*	1,000	-0.3617*	0.7746*	0.5995*	0.5058*	-0.7397*	0.5626*	0.2047*
Ajuste Retangular	0.1974*	-0.2031*	-0.4564*	-0.8362*	0.5719*	0.9341*	-0.3617*	1,000	-0.6964*	-0.3459*	0.04605	0.7106*	-0.09532	0.009502
Índice de Forma	0.6598*	0.2532*	0.8829*	0.7528*	-0.291*	-0.8982*	0.7746*	-0.6964*	1,000	0.6333*	0.1819*	-0.9618*	0.3335*	0.03435
Largura	0.8284*	-0.2807*	0.8101*	0.5365*	0.3663*	-0.3642*	0.5995*	-0.3459*	0.6333*	1,000	0.4255*	-0.592*	0.4028*	-0.0513
μ dos Comprimentos de Borda	0.5832*	0.1164*	0.4241*	0.90341*	0.2087*	0.04533	0.5058*	0.04605	0.1819*	0.4255*	1,000	-0.1244*	0.7434*	0.1987*
Compactação (Polígono)	0.6082*	-0.2611*	-0.8327*	-0.7489*	0.3245*	0.708*	-0.7397*	0.7106*	-0.9618*	-0.592*	-0.1244*	1,000	-0.3037*	-0.0456
Comprimento da maior Borda	0.5518*	0.212*	0.4861*	0.2219*	0.03563	-0.1069*	0.5626*	0.09532*	0.3335*	0.4028*	0.7434*	-0.3037*	1,000	0.2114*
Classe Phantom	0.1068*	0.2921*	0.07704*	0.02249	-0.2053*	0.01801	0.2047*	0.009502	0.03435	-0.0513	0.1987*	-0.0456	0.2114*	1,00
Ranking	6	1	7	11	3	12	4	13	10	8	5	9	2	

(C): Atributos Forma – Skeletons

Tabela 1 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 1

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 1: 06/07/01) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Comprimento da Linha Principal I	Comprimento da Linha Principal II	Comprimento/Largura dos Ramos	Máximo Comprimento dos Ramos	Desv. Padrão do Comprimento das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.5897*	0.5852*	0.5789*	0.9025*	0.1025*	0.09765*	0.02956*	
Comprimento da Linha Principal I	0.5897*	1,000	0.9959*	0.9104*	0.6791*	0.3118*	0.1839*	0.1312*	
Comprimento da Linha Principal II	0.5852*	0.9959*	1,000	0.9061*	0.6817*	0.309*	0.18*	0.1327*	
Comprimeno/Largura dos Ramos	0.5789*	0.9104*	0.9061*	1,000	0.6188*	0.1644*	-0.1735*	0.09948*	
Máximo Comprimento dos Ramos	0.9025*	0.6791*	0.6817*	0.6188*	1,000	0.1829*	0.2123*	0.05811*	
Desv. Padrão do Comprimento das Bordas	0.1025*	0.3118*	0.309*	0.1644*	0.1829*	1,000	0.4031*	0.2279*	
Largura dos Ramos	0.09765*	0.1839*	0.18*	-0.1735*	0.2123*	0.4031*	1,000	0.07899*	
Classe Phantom	0.02956*	0.1312*	0.1327*	0.09948*	0.05811*	0.2279*	0.07899*	1,000	
Ranking	7	3	2	4	6	1	5		

Tabela 2 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 2

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 2: 06/07/02) Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Comprimento da Linha Principal I	Comprimento da Linha Principal II	Comprimento/Largura dos Ramos	Máximo Comprimento dos Ramos	Desv. Padrão do Comprimento das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.3782*	0.3782*	0.4132*	0.9128*	-0.02458	-0.06062	0.04258	
Comprimento da Linha Principal I	0.3782*	1,000	1*	0.9005*	0.4869*	0.3441*	-0.01182	0.3348*	
Comprimento da Linha Principal II	0.3782*	1*	1,000	0.9005*	0.4869*	0.3441*	-0.01182	0.3348*	
Comprimeno/Largura dos Ramos	0.4132*	0.9005*	0.9005*	1,000	0.4638*	0.2136*	-0.373*	0.3188*	
Máximo Comprimento dos Ramos	0.9128*	0.4869*	0.4869*	0.4638*	1,000	0.04658	1,000	0.1953*	
Desv. Padrão da Comprimento das Bordas	-0.02458	0.3441*	0.3441*	0.2136*	0.04658	1,000	0.1953*	0.3762*	
Largura dos Ramos	-0.06062	-0.01182	-0.01182	-0.373*	0.05373	0.1953*	1,000	-0.03204	
Classe Phantom	0.04258	0.3348*	0.3348*	0.3188*	0.06628	0.3762*	-0.03204	1,000	
Ranking	6	2	3	4	5	1	7		

Tabela 3 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 3

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 3: 06/03/06)									
Correlações Significativas com o Phantom ($p < ,05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.5206*	0.5206*	0.5007*	0.8167*	0.04671	0.06397	0.106	
Compr./to da Linha Principal I	0.5206*	1,000	1*	0.9066*	0.686*	0.3476*	0.113	0.4015*	
Compr./to da Linha Principal II	0.5206*	1*	1,000	0.9066*	0.686*	0.3476*	0.113	0.4015*	
Comprimeno/Largura dos Ramos	0.5007*	0.9066*	0.9066*	1,000	0.5826*	0.24*	-0.2286*	0.337*	
Máximo Compr./to dos Ramos	0.8167*	0.686*	0.686*	0.5826*	1,000	0.2164*	0.2222*	0.3013*	
Desv. Padrão do Compr./to das Bordas	0.04671	0.3476*	0.3476*	0.24*	0.2164*	1,000	0.3674*	0.3338*	
Largura dos Ramos	0.06397	0.113	0.113	-0.2286*	0.2222*	0.3674*	1,000	0.1234	
Classe Phantom	0.106	0.4015*	0.4015*	0.337*	0.3013*	0.3338*	0.1234	1,000	
Ranking	7	1	2	3	5	4	6		

Tabela 4 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 4

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 4 09/09/01)									
Correlações Significativas com o Phantom ($p < ,05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.588*	0.5881*	0.5785*	0.8829*	0.05253	0.07881*	0.07272*	
Compr./to da Linha Principal I	0.588*	1,000	1*	0.9018*	0.6985*	0.2597*	0.1758*	0.1195*	
Compr./to da Linha Principal II	0.5881*	1*	1,000	0.9015*	0.6985*	0.2597*	0.1759*	0.1196*	
Comprimeno/Largura dos Ramos	0.5785*	0.9018*	0.9015*	1,000	0.6414*	0.112*	-0.1967*	0.1019*	
Máximo Compr./to dos Ramos	0.8829*	0.6985*	0.6985*	0.6414*	1,000	0.1265*	0.1848*	0.1034*	
Desv. Padrão do Compr./to das Bordas	0.05253	0.2597*	0.2597*	0.112*	0.1265*	1,000	0.3728*	0.09485*	
Largura dos Ramos	0.07881*	0.1758*	0.1759*	-0.1967*	0.1848*	0.3728*	1,000	0.06598*	
Classe Phantom	0.07272*	0.1195*	0.1196*	0.1019*	0.1034*	0.09485*	0.06598*	1,000	
Ranking	6	2	1	4	3	5	7		

Tabela 5 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 5

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 5: 13/06/08)									
Correlações Significativas com o Phantom ($p < ,05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.5534*	0.55*	0.5531*	0.8907*	0.01452	0.01843	0.04102*	
Compr./to da Linha Principal I	0.5534*	1,000	0.997*	0.9065*	0.6502*	0.2726*	0.1279*	0.08049*	
Compr./to da Linha Principal II	0.55*	0.997*	1,000	0.9035*	0.6521*	0.2708*	0.1248*	0.08104*	
Comprimeno/Largura dos Ramos	0.5531*	0.9065*	0.9035*	1,000	0.5891*	0.112*	-0.2393*	0.0599*	
Máximo Compr./to dos Ramos	0.8907*	0.6502*	0.6521*	0.5891*	1,000	0.09835*	0.1504*	0.05487*	
Desv. Padrão do Compr./to das Bordas	0.01452	0.2726*	0.2708*	0.112*	0.09835*	1,000	0.3804*	0.01429	
Largura dos Ramos	0.01843	0.1279*	0.1248*	-0.2393*	0.1504*	0.3804*	1,000	0.06685*	
Classe Phantom	0.04102*	0.08049*	0.08104*	0.0599*	0.05487*	0.01429	0.06685*	1,000	
Ranking	6,00	2,00	1,00	4,00	5,00	7,00	3,00		

Tabela 6 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 6

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 6: 05/05/02)									
Correlações Significativas com o Phantom ($p < ,05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.5685*	0.5685*	0.5457*	0.9377*	0.04998	0.1317	0.2247	
Compr./to da Linha Principal I	0.5685*	1,000	1*	0.9081*	0.6588*	0.2799*	0.1321	0.3518*	
Compr./to da Linha Principal II	0.5685*	1*	1,000	0.9075*	0.6587*	0.2798*	0.1321	0.3518*	
Comprimeno/Largura dos Ramos	0.5457*	0.9081*	0.9075*	1,000	0.6063*	0.1676	-0.2258	0.3704*	
Máximo Compr./to dos Ramos	0.9377*	0.6588*	0.6587*	0.6063*	1,000	0.1265	0.1971	0.2774*	
Desv. Padrão do Compr./to das Bordas	0.04998	0.2799*	0.2798*	0.1676	0.1265	1,000	0.2116	0.2632*	
Largura dos Ramos	0.1317	0.1321	0.1321	-0.2258	0.1971	0.2116	1,000	0.05125	
Classe Phantom	0.2247	0.3518*	0.3518*	0.3704*	0.2774*	0.2632*	0.05125	1,000	
Ranking	6	2	3	1	4	5	7		

Tabela 7 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 7

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 7 19/12/01)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.6267*	0.6262*	0.6113*	0.9217*	0.08639	0.207*	-0.05519	
Compr./to da Linha Principal I	0.6267*	1,000	0.9999*	0.9195*	0.7222*	0.2898*	0.3136*	0.003264	
Compr./to da Linha Principal II	0.6262*	0.9999*	1,000	0.9194*	0.722*	0.29*	0.3131*	0.003505	
Comprimeno/Largura dos Ramos	0.6113*	0.9195*	0.9194*	1,000	0.6725*	0.1524*	-0.01326	0.02469	
Máximo Compr./to dos Ramos	0.9217*	0.7222*	0.722*	0.6725*	1,000	0.1593*	0.3031*	-0.07464	
Desv. Padrão do Compr./to das Bordas	0.08639	0.2898*	0.29*	0.1524*	0.1593*	1,000	0.4406*	0.02576	
Largura dos Ramos	0.207*	0.3136*	0.3131*	-0.01326	0.3031*	0.4406*	1,000	-0.07542	
Classe Phantom	-0.05519	0.003264	0.003505	0.02469	-0.07464	0.02576	-0.07542	1,000	
Ranking	3	7	6	5	2	4	1		

Tabela 8 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 8

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 8 03/04/02)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.6512*	0.6511*	0.6157*	0.9023*	0.1613*	0.2288*	-0.003393	
Compr./to da Linha Principal I	0.6512*	1,000	0.9999*	0.9139*	0.7176*	0.4207*	0.289*	-0.11*	
Compr./to da Linha Principal II	0.6511*	0.9999*	1,000	0.9137*	0.7174*	0.4208*	0.289*	-0.11*	
Comprimeno/Largura dos Ramos	0.6157*	0.9139*	0.9137*	1,000	0.6329*	0.264*	-0.0518	-0.09488*	
Máximo Compr./to dos Ramos	0.9023*	0.7176*	0.7174*	0.6329*	1,000	0.2601*	0.3414*	-0.0056	
Desv. Padrão do Compr./to das Bordas	0.1613*	0.4207*	0.4208*	0.264*	0.2601*	1,000	0.4627*	-0.09107*	
Largura dos Ramos	0.2288*	0.289*	0.289*	-0.0518	0.3414*	0.4627*	1,000	-0.02162	
Classe Phantom	-0.003393	-0.11*	-0.11*	-0.09488*	-0.0056	-0.09107*	-0.02162	1,000	
Ranking	7	1	2	3	6	4	5		

Tabela 9 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 9

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 9 04/07/01)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.6201*	0.6201*	0.5659*	0.8819*	0.1606*	0.223*	-0.004729	
Compr./to da Linha Principal I	0.6201*	1,000	1*	0.9105*	0.6827*	0.3367*	0.2544*	0.2123*	
Compr./to da Linha Principal II	0.6201*	1*	1,000	0.9104*	0.6826*	0.3367*	0.2544*	0.2127*	
Comprimeno/Largura dos Ramos	0.5659*	0.9105*	0.9104*	1,000	0.5785*	0.2052*	-0.1035	0.1244*	
Máximo Compr./to dos Ramos	0.8819*	0.6827*	0.6826*	0.5785*	1,000	0.2245*	0.3216*	0.01427	
Desv. Padrão do Compr./to das Bordas	0.1606*	0.3367*	0.3367*	0.2052*	0.2245*	1,000	0.353*	0.2354*	
Largura dos Ramos	0.223*	0.2544*	0.2544*	-0.1035	0.3216*	0.353*	1,000	0.1548*	
Classe Phantom	-0.004729	0.2123*	0.2127*	0.1244*	0.01427	0.2354*	0.1548*	1,000	
Ranking	7	3	2	5	6	1	4		

Tabela 10 – Matriz de Correlação dos Atributos de Forma (Skeletons) em relação ao atributo da Classe do Phantom: Recorte de Cena 10

Matriz de Correlação dos Atributos de Forma (Skeletons) em Relação a Classe do Phantom (Cena 10: 11/03/05)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.4683*	0.4661*	0.4313*	0.7163*	0.003819	0.137*	-0.01082	
Compr./to da Linha Principal I	0.4683*	1,000	0.9982*	0.8464*	0.5914*	0.2275*	0.2164*	0.1078*	
Compr./to da Linha Principal II	0.4661*	0.9982*	1,000	0.8445*	0.593*	0.226*	0.2145*	0.1082*	
Comprimeno/Largura dos Ramos	0.4313*	0.8464*	0.8445*	1,000	0.4333*	0.03765*	-0.2597*	0.01872	
Máximo Compr./to dos Ramos	0.7163*	0.5914*	0.593*	0.4333*	1,000	0.1325*	0.3293*	0.05999*	
Desv. Padrão do Compr./to das Bordas	0.003819	0.2275*	0.226*	0.03765*	0.1325*	1,000	0.3965*	0.1752*	
Largura dos Ramos	0.137*	0.2164*	0.2145*	-0.2597*	0.3293*	0.3965*	1,000	0.1772*	
Classe Phantom	-0.01082	0.1078*	0.1082*	0.01872	0.05999*	0.1752*	0.1772*	1,000	
Ranking	7	4	3	6	5	2	1		

Tabela 11 – Matriz de Correlação dos Atributos de Forma (*Skeletons*) em relação ao atributo da Classe do Phantom: Recorte de Cena 11

Matriz de Correlação dos Atributos de Forma (<i>Skeletons</i>) em Relação a Classe do Phantom (Cena 11: 14/06/08)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	Grau de Ramificação do Skeleton	Compr./to da Linha Principal I	Compr./to da Linha Principal II	Comprime no/Largura dos Ramos	Máximo Compr./to dos Ramos	Desv. Padrão do Compr./to das Bordas	Largura dos Ramos	Classe Phantom	
Grau de Ramificação do Skeleton	1,000	0.4843*	0.4755*	0.4468*	0.7479*	0.03697	0.1951*	-0.01684	
Compr./to da Linha Principal I	0.4843*	1,000	0.9928*	0.8706*	0.5849*	0.2897*	0.2347*	0.1383*	
Compr./to da Linha Principal II	0.4755*	0.9928*	1,000	0.8639*	0.591*	0.2827*	0.2274*	0.1402*	
Comprimeno/Largura dos Ramos	0.4468*	0.8706*	0.8639*	1,000	0.4569*	0.09875*	-0.1886*	0.06308	
Máximo Compr./to dos Ramos	0.7479*	0.5849*	0.591*	0.4569*	1,000	0.1394*	0.3364*	0.02108	
Desv. Padrão do Compr./to das Bordas	0.03697	0.2897*	0.2827*	0.09875*	0.1394*	1,000	0.4464*	0.2158*	
Largura dos Ramos	0.1951*	0.2347*	0.2274*	-0.1886*	0.3364*	0.4464*	1,000	0.1578*	
Classe Phantom	-0.01684	0.1383*	0.1402*	0.06308	0.02108	0.2158*	0.1578*	1,000	
Ranking	7	4	3	5	6	1	2		

(D): Atributos Texturais - Haralick

Tabela 1 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 1

Matriz de Correlação dos Atributos Texturais (Haralick) em Relação a Classe do Phantom (Cena 1: 06/07/01)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.5834*	-0.6106*	0.5745*	-0.9898*	0.7159*	-0.03134*	0.4884*	-0.2721*
GLCM Contraste	0.5834*	1,000	-0.9568*	0.9944*	-0.6316*	0.11*	-0.006844	0.9266*	-0.1219*
GLCM Correlação	-0.6106*	-0.9568*	1,000	-0.9433*	0.652*	-0.1438*	0.04275*	-0.7872*	0.1849*
GLCM Dissimilaridade	0.5745*	0.9944*	-0.9433*	1,000	-0.6289*	0.07702*	-0.007358	0.932*	-0.1234*
GLCM Entropia	-0.9898*	-0.6316*	0.652*	-0.6289*	1,000	-0.6767*	0.02928	-0.5405*	0.2695*
GLCM Homogeneidade	0.7159*	0.11*	-0.1438*	0.07702*	-0.6767*	1,000	0.02173	0.05772*	-0.1877*
GLCM Média	-0.03134*	-0.006844	0.04275*	-0.007358	0.02928	0.02173	1,000	0.04142*	0.2628*
GLCM Desvio Padrão	0.4884*	0.9266*	-0.7872*	0.932*	-0.5405*	0.05772*	0.04142*	1,000	-0.02967*
Classe Phantom	-0.2721*	-0.1219*	0.1849*	-0.1234*	0.2695*	-0.1877*	0.2628*	-0.02967*	1,00
Ranking	1	7	5	6	2	4	3	8	

Tabela 2 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 2

Matriz de Correlação dos Atributos Texturais (Haralick) em Relação a Classe do Phantom (Cena 2: 06/07/02)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.5043*	-0.5797*	0.4583*	-0.9743*	0.5378*	-0.5936*	0.3459*	-0.3781*
GLCM Contraste	0.5043*	1,000	-0.9087*	0.9875*	-0.5712*	-0.06404	-0.6475*	0.8686*	-0.005071
GLCM Correlação	-0.5797*	-0.9087*	1,000	-0.8772*	0.6393*	-0.03702	0.7318*	-0.6113*	0.2725*
GLCM Dissimilaridade	0.4583*	0.9875*	-0.8772*	1,000	-0.5426*	-0.1391*	-0.628*	0.8837*	0.02077
GLCM Entropia	-0.9743*	-0.5712*	0.6393*	-0.5426*	1,000	-0.4599*	0.665*	-0.4128*	0.3939*
GLCM Homogeneidade	0.5378*	-0.06404	-0.03702	-0.1391*	-0.4599*	1,000	-0.1198	-0.1539*	-0.2474*
GLCM Média	-0.5936*	-0.6475*	0.7318*	-0.628*	0.665*	-0.1198	1,000	-0.4543*	0.5177*
GLCM Desvio Padrão	0.3459*	0.8686*	-0.6113*	0.8837*	-0.4128*	-0.1539*	-0.4543*	1,000	0.2521*
Classe Phantom	-0.3781*	-0.005071	0.2725*	0.02077	0.3939*	-0.2474*	0.5177*	0.2621*	1,00
Ranking	3	8	4	7	2	6	1	5	

Tabela 3 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 3

Matriz de Correlação dos Atributos Texturais (Haralick) em Relação a Classe do Phantom (Cena 3: 06/03/06)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.7164*	-0.6816*	0.7017*	-0.9803*	0.3506*	-0.2045*	0.5443*	-0.2358*
GLCM Contraste	0.7164*	1,000	-0.8836*	0.9902*	-0.7805*	-0.1311	-0.4518*	0.8224*	-0.3809*
GLCM Correlação	-0.6816*	-0.8836*	1,000	-0.843*	0.7355*	0.1116	0.5387*	-0.4962*	0.5736*
GLCM Dissimilaridade	0.7017*	0.9902*	-0.843*	1,000	-0.7708*	-0.1524	-0.4268*	0.8533*	-0.3374*
GLCM Entropia	-0.9803*	-0.7805*	0.7355*	-0.7708*	1,000	-0.2569*	0.2554*	-0.611*	0.287*
GLCM Homogeneidade	0.3506*	-0.1311	0.1116	-0.1524	-0.2569*	1,000	0.3239*	-0.1284	0.1889*
GLCM Média	-0.2045*	-0.4518*	0.5387*	-0.4268*	0.2554*	0.3239*	1,000	-0.2488*	0.51*
GLCM Desvio Padrão	0.5443*	0.8224*	-0.4962*	0.8533*	-0.611*	-0.1284	-0.2488*	1,000	0.01254
Classe Phantom	-0.2358*	-0.3809*	0.5736*	-0.3374*	0.287*	0.1889*	0.51*	0.01254	1,00
Ranking	6	3	1	4	5	7	2	8	

Tabela 4 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 4

Matriz de Correlação dos Atributos Texturais (Haralick) em Relação a Classe do Phantom (Cena 4 09/09/01)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.7*	-0.6719*	0.684*	-0.984*	0.5607*	0.1158*	0.6407*	0.01343
GLCM Contraste	0.7*	1,000	-0.9474*	0.9945*	-0.757*	0.1004*	-0.0592*	0.9274*	-0.09717*
GLCM Correlação	-0.6719*	-0.9474*	1,000	-0.9345*	0.7179*	-0.0788*	0.0941*	-0.7706*	0.1688*
GLCM Dissimilaridade	0.684*	0.9945*	-0.9345*	1,000	-0.7488*	0.05335*	-0.05792*	0.9319*	-0.09441*
GLCM Entropia	-0.984*	-0.757*	0.7179*	-0.7488*	1,000	-0.5091*	-0.1074*	-0.7042*	-0.006817
GLCM Homogeneidade	0.5607*	0.1004*	-0.0788*	0.05335*	-0.5091*	1,000	0.2005*	0.1013*	0.1272*
GLCM Média	0.1158*	-0.0592*	0.0941*	-0.05792*	-0.1074*	0.2005*	1,000	-0.01031	0.0667*
GLCM Desvio Padrão	0.6407*	0.9274*	-0.7706*	0.9319*	-0.7042*	0.1013*	-0.01031	1,000	0.009623
Classe Phantom	0.01343	-0.09717*	0.1688*	-0.09441*	-0.006817	0.1272*	0.0667*	0.009623	1,00
Ranking	6	3	1	4	8	2	5	7	

Tabela 5 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 5

Matriz de Correlação dos Atributos Texturais (Haralick) em Relação a Classe do Phantom (Cena 5: 13/06/08)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.4781*	-0.3621*	0.482*	-0.9872*	0.5537*	0.1291*	0.5422*	0.1559*
GLCM Contraste	0.4781*	1,000	-0.9302*	0.9847*	-0.5321*	-0.08001*	-0.09476*	0.8866*	-0.1267*
GLCM Correlação	-0.3621*	-0.9302*	1,000	-0.8895*	0.4033*	0.1385*	0.1395*	-0.6688*	0.1926*
GLCM Dissimilaridade	0.482*	0.9847*	-0.8895*	1,000	-0.547*	-0.1166*	-0.1018*	0.9079*	-0.1046*
GLCM Entropia	-0.9872*	-0.5321*	0.4033*	-0.547*	1,000	-0.4881*	-0.1127*	-0.6026*	-0.1471*
GLCM Homogeneidade	0.5537*	-0.08001*	0.1385*	-0.1166*	-0.4881*	1,000	0.2264*	0.0104	0.2*
GLCM Média	0.1291*	-0.09476*	0.1395*	-0.1018*	-0.1127*	0.2264*	1,000	-0.01647	0.06359*
GLCM Desvio Padrão	0.5422*	0.8866*	-0.6688*	0.9079*	-0.6026*	0.0104	-0.01647	1,000	-0,07
Classe Phantom	0.1559*	-0.1267*	0.1926*	-0.1046*	-0.1471*	0.2*	0.06359*	-0,068	1,00
Ranking	3	5	2	6	4	1	7	8	

Tabela 6 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 6

Matriz de Correlação dos Atributos Texturais (<i>Haralick</i>) em Relação a Classe do Phantom (Cena 6: 05/05/02)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.7714*	-0.8172*	0.7531*	-0.9887*	0.6111*	-0.1792	0.5855*	-0.601*
GLCM Contraste	0.7714*	1,000	-0.9596*	0.9921*	-0.8125*	0.2657*	-0.4196*	0.8912*	-0.5241*
GLCM Correlação	-0.8172*	-0.9596*	1,000	-0.9464*	0.841*	-0.3423*	0.4018*	-0.7465*	0.5521*
GLCM Dissimilaridade	0.7531*	0.9921*	-0.9464*	1,000	-0.8007*	0.2194	-0.4413*	0.8957*	-0.5078*
GLCM Entropia	-0.9887*	-0.8125*	0.841*	-0.8007*	1,000	-0.5871*	0.2183	-0.6372*	0.601*
GLCM Homogeneidade	0.6111*	0.2657*	-0.3423*	0.2194	-0.5871*	1,000	0.06941	0.1108	-0.431*
GLCM Média	-0.1792	-0.4196*	0.4018*	-0.4413*	0.2183	0.06941	1,000	-0.3804*	0.3052*
GLCM Desvio Padrão	0.5855*	0.8912*	-0.7465*	0.8957*	-0.6372*	0.1108	-0.3804*	1,000	-0.3168*
Classe Phantom	-0.601*	-0.5241*	0.5521*	-0.5078*	0.801*	-0.431*	0.3052*	-0.3168*	1,00
<i>Ranking</i>	1	4	3	5	2	6	8	7	

Tabela 7 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 7

Matriz de Correlação dos Atributos Texturais (<i>Haralick</i>) em Relação a Classe do Phantom (Cena 7 19/12/01)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.699*	-0.6762*	0.6785*	-0.9915*	0.06059	-0.05912	0.6696*	-0.06447
GLCM Contraste	0.699*	1,000	-0.9689*	0.9924*	-0.7186*	-0.3063*	-0.06245	0.9411*	0.08551
GLCM Correlação	-0.6762*	-0.9689*	1,000	-0.9685*	0.6982*	0.2964*	0.02639	-0.8338*	-0.06998
GLCM Dissimilaridade	0.6785*	0.9924*	-0.9685*	1,000	-0.7046*	-0.3538*	-0.08329	0.9387*	0.08744*
GLCM Entropia	-0.9915*	-0.7186*	0.6982*	-0.7046*	1,000	0.005604	0.08314	-0.6843*	0.05463
GLCM Homogeneidade	0.06059	-0.3063*	0.2964*	-0.3538*	0.005604	1,000	0.05731	-0.2875*	-0.1805*
GLCM Média	-0.05912	-0.06245	0.02639	-0.08329	0.08314	0.05731	1,000	-0.1117*	-0.4932*
GLCM Desvio Padrão	0.6696*	0.9411*	-0.8338*	0.9387*	-0.6843*	-0.2875*	-0.1117*	1,000	0.1039*
Classe Phantom	-0.06447	0.08551	-0.06998	0.08744*	0.05463	-0.1805*	-0.4932*	0.1039*	1,00
<i>Ranking</i>	7	5	6	4	8	2	1	3	

Tabela 8 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 8

Matriz de Correlação dos Atributos Texturais (<i>Haralick</i>) em Relação a Classe do Phantom (Cena 8 03/04/02)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.6019*	-0.6308*	0.5951*	-0.9904*	0.6449*	0.1396*	0.4846*	0.0917*
GLCM Contraste	0.6019*	1,000	-0.9513*	0.9953*	-0.6547*	0.01201	-0.02212	0.9136*	0.07448*
GLCM Correlação	-0.6308*	-0.9513*	1,000	-0.9404*	0.6758*	-0.04356	0.06602	-0.759*	-0.147*
GLCM Dissimilaridade	0.5951*	0.9953*	-0.9404*	1,000	-0.653*	-0.01529	-0.01729	0.9175*	0.07101
GLCM Entropia	-0.9904*	-0.6547*	0.6758*	-0.653*	1,000	-0.5957*	-0.1341*	-0.5385*	-0.09353*
GLCM Homogeneidade	0.6449*	0.01201	-0.04356	-0.01529	-0.5957*	1,000	0.193*	-0.03089	0.01345
GLCM Média	0.1396*	-0.02212	0.06602	-0.01729	-0.1341*	0.193*	1,000	0.04527	-0.199*
GLCM Desvio Padrão	0.4846*	0.9136*	-0.759*	0.9175*	-0.5385*	-0.03089	0.04527	1,000	-0.03931
Classe Phantom	0.0917*	0.07448*	-0.147*	0.07101	-0.09353*	0.01345	-0.199*	-0.03931	1,000
<i>Ranking</i>	4	5	2	6	3	8	1	7	

Tabela 9 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 9

Matriz de Correlação dos Atributos Texturais (<i>Haralick</i>) em Relação a Classe do Phantom (Cena 9: 04/07/01)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.6634*	-0.7057*	0.6383*	-0.9804*	0.5825*	0.04814	0.5274*	-0.3193*
GLCM Contraste	0.6634*	1,000	-0.9525*	0.993*	-0.7263*	0.05179	-0.135*	0.9162*	-0.1801*
GLCM Correlação	-0.7057*	-0.9525*	1,000	-0.9331*	0.7551*	-0.1271*	0.1975*	-0.7651*	0.2757*
GLCM Dissimilaridade	0.6383*	0.993*	-0.9331*	1,000	-0.7105*	-0.0008367	-0.1299*	0.9272*	-0.1708*
GLCM Entropia	-0.9804*	-0.7263*	0.7551*	-0.7105*	1,000	-0.5206*	-0.0222	-0.6013*	0.3209*
GLCM Homogeneidade	0.5825*	0.05179	-0.1271*	-0.0008367	-0.5206*	1,000	0.1972*	-0.06258	-0.1882*
GLCM Média	0.04814	-0.135*	0.1975*	-0.1299*	-0.0222	0.1972*	1,000	-0.04718	0.2119*
GLCM Desvio Padrão	0.5274*	0.9162*	-0.7651*	0.9272*	-0.6013*	-0.06258	-0.04718	1,000	-0.02395
Classe Phantom	-0.3193*	-0.1801*	0.2757*	-0.1708*	0.3209*	-0.1882*	0.2119*	-0.02395	1,00
Ranking	2	6	3	7	1	5	4	8	

Tabela 10 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 10

Matriz de Correlação dos Atributos Texturais (<i>Haralick</i>) em Relação a Classe do Phantom (Cena 10: 11/03/05)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.549*	0.01519	0.5501*	-0.9895*	0.3445*	0.1442*	0.5796*	0.1292*
GLCM Contraste	0.549*	1,000	-0.1422*	0.9887*	-0.5902*	-0.1201*	0.06251*	0.8524*	-0.1348*
GLCM Correlação	0.01519	-0.1422*	1,000	-0.1407*	-0.002766	0.1774*	0.04743*	0.02648	0.297*
GLCM Dissimilaridade	0.5501*	0.9887*	-0.1407*	1,000	-0.5951*	-0.128*	0.07888*	0.8676*	-0.1137*
GLCM Entropia	-0.9895*	-0.5902*	-0.002766	-0.5951*	1,000	-0.2814*	-0.1335*	-0.6095*	-0.1125*
GLCM Homogeneidade	0.3445*	-0.1201*	0.1774*	-0.128*	-0.2814*	1,000	0.2199*	0.05519*	0.2615*
GLCM Média	0.1442*	0.06251*	0.04743*	0.07888*	-0.1335*	0.2199*	1,000	0.2412*	0.1913*
GLCM Desvio Padrão	0.5796*	0.8524*	0.02648	0.8676*	-0.6095*	0.05519*	0.2412*	1,000	0.1486*
Classe Phantom	0.1292*	-0.1348*	0.297*	-0.1137*	-0.1125*	0.2615*	0.1913*	0.1486*	1,00
Ranking	6	5	1	7	8	2	3	4	

Tabela 11 – Matriz de Correlação dos Atributos Texturais (*Haralick*) em relação ao atributo da Classe do Phantom: Recorte de Cena 11

Matriz de Correlação dos Atributos Texturais (<i>Haralick</i>) em Relação a Classe do Phantom (Cena 11: 14/06/08)									
Correlações Significativas com o Phantom ($p < .05$) estão assinaladas em vermelho									
Atributos	GLCM 2o Momento Angular	GLCM Contraste	GLCM Correlação	GLCM Dissimilaridade	GLCM Entropia	GLCM Homogeneidade	GLCM Média	GLCM Desvio Padrão	Classe Phantom
GLCM 2o Momento Angular	1,000	0.5535*	-0.07307*	0.5364*	-0.9869*	0.3849*	0.1956*	0.5327*	0.1222*
GLCM Contraste	0.5535*	1,000	-0.2145*	0.98*	-0.6*	-0.07912*	-0.02815	0.8149*	-0.01359
GLCM Correlação	-0.07307*	-0.2145*	1,000	-0.1865*	0.07493*	0.2102*	0.187*	0.07727*	0.3378*
GLCM Dissimilaridade	0.5364*	0.98*	-0.1865*	1,000	-0.593*	-0.1099*	0.009605	0.8483*	0.0308
GLCM Entropia	-0.9859*	-0.6*	0.07493*	-0.593*	1,000	-0.3181*	1,000	-0.1782*	-0.5759*
GLCM Homogeneidade	0.3849*	-0.07912*	0.2102*	-0.1099*	-0.3181*	1,000	0.3743*	0.07717*	0.2243*
GLCM Média	0.1956*	-0.02815	0.187*	0.009605	-0.1782*	0.3743*	1,000	0.2605*	0.4238*
GLCM Desvio Padrão	0.5327*	0.8149*	0.07727*	0.8483*	-0.5759*	0.07717*	0.2605*	1,000	0.3214*
Classe Phantom	0.1222*	-0.01359	0.3378*	0.0308	-0.1214*	0.2243*	0.4238*	0.3214*	1,00
Ranking	5	8	2	7	6	4	1	3	

APÊNDICE IV:

Análise de Componentes Principais

Tabela 1 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 1

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 01: 06/07/2001 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	9,407	26,131	9,407	26,131	73,869
2	5,882	16,338	15,289	42,469	57,531
3	5,352	14,866	20,641	57,335	42,665
4	3,505	9,735	24,145	67,070	32,930
5	3,103	8,620	27,248	75,690	24,310
6	1,791	4,976	29,040	80,666	19,334
7	1,435	3,986	30,475	84,652	15,348
8	0,950	2,639	31,425	87,291	12,709
9	0,856	2,379	32,281	89,669	10,331
10	0,712	1,978	32,993	91,647	8,353
11	0,499	1,386	33,492	93,033	6,967
12	0,416	1,156	33,908	94,189	5,811
13	0,390	1,083	34,298	95,272	4,728
14	0,357	0,991	34,655	96,263	3,737
15	0,288	0,800	34,943	97,063	2,937
16	0,211	0,587	35,154	97,650	2,350
17	0,186	0,516	35,340	98,167	1,833
18	0,129	0,359	35,469	98,525	1,475
19	0,109	0,301	35,578	98,827	1,173
20	0,092	0,257	35,670	99,083	0,917
21	0,066	0,182	35,736	99,265	0,735
22	0,063	0,174	35,798	99,439	0,561
23	0,049	0,136	35,847	99,575	0,425
24	0,042	0,117	35,889	99,692	0,308
25	0,032	0,089	35,921	99,781	0,219
26	0,015	0,041	35,936	99,822	0,178
27	0,014	0,038	35,950	99,860	0,140
28	0,012	0,033	35,961	99,893	0,107
29	0,009	0,026	35,971	99,919	0,081
30	0,008	0,022	35,979	99,940	0,060
31	0,006	0,017	35,985	99,957	0,043
32	0,005	0,014	35,990	99,971	0,029
33	0,004	0,012	35,994	99,983	0,017
34	0,003	0,009	35,997	99,992	0,008
35	0,003	0,007	36,000	99,999	0,001
36	0,000	0,001	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 2 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 2

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 02: 06/07/2002 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	11,376	31,600	11,376	31,600	68,400
2	8,222	22,839	19,598	54,439	45,561
3	4,464	12,399	24,062	66,838	33,162
4	3,793	10,537	27,855	77,374	22,626
5	2,875	7,987	30,730	85,361	14,639
6	1,424	3,955	32,154	89,316	10,684
7	0,720	2,000	32,874	91,317	8,683
8	0,687	1,908	33,561	93,224	6,776
9	0,511	1,419	34,071	94,643	5,357
10	0,413	1,148	34,485	95,791	4,209
11	0,327	0,907	34,811	96,698	3,302
12	0,235	0,654	35,047	97,352	2,648
13	0,212	0,590	35,259	97,942	2,058
14	0,172	0,477	35,431	98,418	1,582
15	0,152	0,421	35,582	98,840	1,160
16	0,109	0,302	35,691	99,141	0,859
17	0,088	0,244	35,779	99,385	0,615
18	0,053	0,148	35,832	99,533	0,467
19	0,044	0,122	35,876	99,655	0,345
20	0,035	0,098	35,911	99,753	0,247
21	0,030	0,083	35,941	99,836	0,164
22	0,016	0,045	35,957	99,880	0,120
23	0,014	0,038	35,971	99,919	0,081
24	0,007	0,019	35,978	99,938	0,062
25	0,006	0,017	35,984	99,955	0,045
26	0,005	0,013	35,988	99,968	0,032
27	0,004	0,011	35,992	99,979	0,021
28	0,003	0,008	35,995	99,987	0,013
29	0,002	0,006	35,997	99,993	0,007
30	0,001	0,003	35,998	99,995	0,005
31	0,001	0,002	35,999	99,997	0,003
32	0,001	0,002	36,000	99,999	0,001
33	0,000	0,001	36,000	100,000	0,000
34	0,000	0,000	36,000	100,000	0,000
35	0,000	0,000	36,000	100,000	0,000
36	0,000	0,000	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 3 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 3

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 03: 06/03/2006 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	11,110	30,860	11,110	30,860	69,140
2	6,616	18,377	17,725	49,237	50,763
3	4,324	12,010	22,049	61,247	38,753
4	3,653	10,147	25,702	71,394	28,606
5	3,044	8,455	28,746	79,849	20,151
6	1,421	3,948	30,167	83,796	16,204
7	1,321	3,669	31,487	87,465	12,535
8	1,022	2,838	32,509	90,303	9,897
9	0,810	2,249	33,319	92,552	7,448
10	0,616	1,710	33,934	94,262	5,738
11	0,445	1,235	34,379	95,497	4,503
12	0,240	0,665	34,618	96,162	3,838
13	0,227	0,631	34,846	96,793	3,207
14	0,209	0,580	35,055	97,374	2,626
15	0,178	0,493	35,232	97,867	2,133
16	0,158	0,439	35,390	98,306	1,694
17	0,142	0,395	35,533	98,702	1,298
18	0,102	0,284	35,635	98,986	1,014
19	0,088	0,246	35,723	99,231	0,769
20	0,049	0,137	35,772	99,368	0,632
21	0,043	0,120	35,815	99,487	0,513
22	0,037	0,103	35,852	99,590	0,410
23	0,030	0,083	35,882	99,673	0,327
24	0,027	0,074	35,909	99,747	0,253
25	0,021	0,059	35,930	99,806	0,194
26	0,015	0,043	35,945	99,849	0,151
27	0,015	0,042	35,960	99,890	0,110
28	0,009	0,026	35,970	99,916	0,084
29	0,008	0,023	35,978	99,939	0,061
30	0,007	0,019	35,985	99,959	0,041
31	0,005	0,014	35,990	99,972	0,028
32	0,004	0,012	35,994	99,984	0,016
33	0,003	0,007	35,997	99,992	0,008
34	0,001	0,004	35,999	99,996	0,004
35	0,001	0,003	36,000	99,999	0,001
36	0,000	0,001	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 4 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 4

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 04: 09/09/2001 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	11,110	30,860	11,110	30,860	69,140
2	6,616	18,377	17,725	49,237	50,763
3	4,324	12,010	22,049	61,247	38,753
4	3,653	10,147	25,702	71,394	28,606
5	3,044	8,455	28,746	79,849	20,151
6	1,421	3,948	30,167	83,796	16,204
7	1,321	3,669	31,487	87,465	12,535
8	1,022	2,838	32,509	90,303	9,697
9	0,810	2,249	33,319	92,552	7,448
10	0,616	1,710	33,934	94,262	5,738
11	0,445	1,235	34,379	95,497	4,503
12	0,240	0,665	34,618	96,162	3,838
13	0,227	0,631	34,846	96,793	3,207
14	0,209	0,580	35,055	97,374	2,626
15	0,178	0,493	35,232	97,867	2,133
16	0,158	0,439	35,390	98,306	1,694
17	0,142	0,395	35,533	98,702	1,298
18	0,102	0,284	35,635	98,986	1,014
19	0,088	0,246	35,723	99,231	0,769
20	0,049	0,137	35,772	99,368	0,632
21	0,043	0,120	35,815	99,487	0,513
22	0,037	0,103	35,852	99,590	0,410
23	0,030	0,083	35,882	99,673	0,327
24	0,027	0,074	35,909	99,747	0,253
25	0,021	0,059	35,930	99,806	0,194
26	0,015	0,043	35,945	99,849	0,151
27	0,015	0,042	35,960	99,890	0,110
28	0,009	0,026	35,970	99,916	0,084
29	0,008	0,023	35,978	99,939	0,061
30	0,007	0,019	35,985	99,959	0,041
31	0,005	0,014	35,990	99,972	0,028
32	0,004	0,012	35,994	99,984	0,016
33	0,003	0,007	35,997	99,992	0,008
34	0,001	0,004	35,999	99,996	0,004
35	0,001	0,003	36,000	99,999	0,001
36	0,000	0,001	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 5 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 5

Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	9,411	26,141	9,411	26,141	73,859
2	6,494	18,039	15,905	44,180	55,820
3	4,083	11,341	19,988	55,521	44,479
4	3,763	10,453	23,751	65,974	34,026
5	2,974	8,260	26,724	74,235	25,765
6	1,659	4,610	28,384	78,844	21,156
7	1,351	3,753	29,735	82,597	17,403
8	1,080	2,999	30,815	85,596	14,404
9	0,981	2,726	31,796	88,322	11,678
10	0,875	2,430	32,671	90,753	9,247
11	0,622	1,727	33,293	92,480	7,520
12	0,502	1,395	33,795	93,876	6,124
13	0,419	1,165	34,215	95,041	4,959
14	0,349	0,969	34,564	96,010	3,990
15	0,256	0,711	34,820	96,721	3,279
16	0,237	0,657	35,056	97,379	2,621
17	0,218	0,606	35,274	97,984	2,016
18	0,150	0,418	35,425	98,402	1,598
19	0,116	0,322	35,541	98,724	1,276
20	0,103	0,285	35,643	99,009	0,991
21	0,059	0,164	35,702	99,173	0,827
22	0,054	0,151	35,757	99,324	0,676
23	0,051	0,143	35,808	99,467	0,533
24	0,048	0,132	35,856	99,599	0,401
25	0,033	0,092	35,889	99,691	0,309
26	0,026	0,073	35,915	99,764	0,236
27	0,020	0,057	35,936	99,821	0,179
28	0,018	0,051	35,954	99,872	0,128
29	0,015	0,040	35,969	99,913	0,087
30	0,009	0,025	35,978	99,938	0,062
31	0,007	0,020	35,985	99,958	0,042
32	0,007	0,019	35,992	99,977	0,023
33	0,004	0,011	35,996	99,988	0,012
34	0,002	0,006	35,998	99,994	0,006
35	0,002	0,004	35,999	99,999	0,001
36	0,001	0,001	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 6 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 6

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 06: 05/05/2006 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	13,380	37,166	13,380	37,166	62,834
2	8,720	24,221	22,099	61,387	38,613
3	6,068	16,855	28,167	78,243	21,757
4	2,908	8,077	31,075	86,320	13,680
5	1,701	4,724	32,776	91,044	8,956
6	0,821	2,282	33,597	93,326	6,674
7	0,672	1,867	34,269	95,193	4,807
8	0,455	1,265	34,725	96,458	3,542
9	0,313	0,870	35,038	97,327	2,673
10	0,225	0,626	35,263	97,953	2,047
11	0,192	0,533	35,455	98,486	1,514
12	0,153	0,426	35,608	98,912	1,088
13	0,132	0,365	35,740	99,278	0,722
14	0,066	0,183	35,806	99,461	0,539
15	0,057	0,158	35,863	99,619	0,381
16	0,045	0,126	35,908	99,745	0,255
17	0,025	0,071	35,934	99,816	0,184
18	0,019	0,053	35,953	99,869	0,131
19	0,015	0,042	35,968	99,911	0,089
20	0,012	0,033	35,980	99,944	0,056
21	0,007	0,018	35,986	99,962	0,038
22	0,005	0,013	35,991	99,975	0,025
23	0,003	0,007	35,993	99,982	0,018
24	0,001	0,004	35,995	99,986	0,014
25	0,001	0,004	35,996	99,989	0,011
26	0,001	0,003	35,997	99,992	0,008
27	0,001	0,002	35,998	99,994	0,006
28	0,001	0,002	35,999	99,996	0,004
29	0,001	0,002	35,999	99,998	0,002
30	0,000	0,001	36,000	99,999	0,001
31	0,000	0,000	36,000	99,999	0,001
32	0,000	0,000	36,000	100,000	0,000
33	0,000	0,000	36,000	100,000	0,000
34	0,000	0,000	36,000	100,000	0,000
35	0,000	0,000	36,000	100,000	0,000
36	0,000	0,000	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 7 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 7

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 07: 19/12/2001 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	12,087	33,575	12,087	33,575	66,425
2	5,679	15,774	17,765	49,349	50,651
3	5,058	14,049	22,823	63,397	36,603
4	3,108	8,634	25,931	72,032	27,968
5	2,107	5,852	28,038	77,884	22,116
6	1,557	4,324	29,595	82,208	17,792
7	1,248	3,467	30,843	85,676	14,324
8	0,935	2,597	31,778	88,272	11,728
9	0,830	2,307	32,609	90,579	9,421
10	0,775	2,152	33,383	92,732	7,268
11	0,631	1,753	34,014	94,485	5,515
12	0,374	1,038	34,388	95,522	4,478
13	0,302	0,840	34,690	96,362	3,638
14	0,201	0,559	34,891	96,920	3,080
15	0,180	0,501	35,072	97,422	2,578
16	0,165	0,457	35,236	97,879	2,121
17	0,150	0,416	35,386	98,295	1,705
18	0,114	0,316	35,500	98,610	1,390
19	0,104	0,289	35,604	98,899	1,101
20	0,077	0,215	35,681	99,114	0,886
21	0,061	0,170	35,742	99,284	0,716
22	0,052	0,145	35,794	99,429	0,571
23	0,049	0,137	35,844	99,566	0,434
24	0,048	0,135	35,892	99,701	0,299
25	0,026	0,073	35,919	99,774	0,226
26	0,024	0,066	35,942	99,840	0,160
27	0,015	0,042	35,957	99,882	0,118
28	0,014	0,038	35,971	99,919	0,081
29	0,009	0,026	35,980	99,946	0,054
30	0,008	0,023	35,989	99,969	0,031
31	0,005	0,013	35,993	99,982	0,018
32	0,003	0,009	35,997	99,991	0,009
33	0,002	0,005	35,998	99,996	0,004
34	0,001	0,003	36,000	99,999	0,001
35	0,000	0,001	36,000	100,000	0,000
36	0,000	0,000	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 8 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 8

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 08: 03/04/2002 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	10,779	29,941	10,779	29,941	70,059
2	6,996	19,432	17,774	49,373	50,627
3	5,035	13,986	22,809	63,360	36,640
4	3,256	9,044	26,065	72,404	27,596
5	2,685	7,459	28,751	79,863	20,137
6	1,594	4,427	30,344	84,289	15,711
7	1,381	3,837	31,725	88,126	11,874
8	0,922	2,560	32,647	90,686	9,314
9	0,828	2,299	33,475	92,985	7,015
10	0,637	1,768	34,111	94,754	5,246
11	0,353	0,980	34,464	95,733	4,267
12	0,331	0,921	34,795	96,654	3,346
13	0,236	0,657	35,032	97,311	2,689
14	0,170	0,472	35,202	97,783	2,217
15	0,134	0,372	35,336	98,156	1,844
16	0,118	0,328	35,454	98,484	1,516
17	0,110	0,304	35,564	98,788	1,212
18	0,081	0,226	35,645	99,014	0,986
19	0,062	0,172	35,707	99,186	0,814
20	0,054	0,151	35,761	99,337	0,683
21	0,049	0,136	35,810	99,474	0,526
22	0,045	0,124	35,855	99,598	0,402
23	0,042	0,116	35,897	99,714	0,286
24	0,035	0,099	35,933	99,813	0,187
25	0,015	0,040	35,947	99,853	0,147
26	0,011	0,030	35,958	99,883	0,117
27	0,009	0,024	35,966	99,907	0,093
28	0,007	0,020	35,973	99,926	0,074
29	0,007	0,019	35,980	99,946	0,054
30	0,005	0,014	35,985	99,959	0,041
31	0,004	0,011	35,989	99,971	0,029
32	0,003	0,009	35,993	99,980	0,020
33	0,003	0,007	35,995	99,987	0,013
34	0,002	0,006	35,998	99,993	0,007
35	0,002	0,005	35,999	99,998	0,002
36	0,001	0,002	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 9 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 9

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 09: 04/07/2001 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	11,570	32,138	11,570	32,138	67,862
2	6,946	19,296	18,516	51,434	48,566
3	5,104	14,177	23,620	65,611	34,389
4	3,138	8,717	26,758	74,328	25,672
5	2,418	6,716	29,176	81,044	18,956
6	1,702	4,728	30,878	85,772	14,228
7	1,147	3,185	32,024	88,957	11,043
8	1,052	2,921	33,076	91,878	8,122
9	0,767	2,130	33,843	94,008	5,992
10	0,577	1,602	34,420	95,610	4,390
11	0,349	0,970	34,769	96,580	3,420
12	0,266	0,738	35,034	97,317	2,683
13	0,227	0,630	35,261	97,947	2,053
14	0,160	0,444	35,421	98,391	1,609
15	0,126	0,349	35,547	98,740	1,260
16	0,108	0,301	35,655	99,042	0,958
17	0,077	0,214	35,732	99,255	0,745
18	0,068	0,189	35,800	99,444	0,556
19	0,050	0,138	35,850	99,583	0,417
20	0,033	0,091	35,883	99,674	0,326
21	0,031	0,085	35,913	99,759	0,241
22	0,022	0,062	35,936	99,821	0,179
23	0,019	0,053	35,955	99,874	0,126
24	0,011	0,029	35,965	99,904	0,096
25	0,008	0,023	35,974	99,927	0,073
26	0,007	0,020	35,981	99,947	0,053
27	0,005	0,015	35,986	99,962	0,038
28	0,005	0,013	35,991	99,975	0,025
29	0,003	0,008	35,994	99,983	0,017
30	0,002	0,006	35,996	99,989	0,011
31	0,001	0,004	35,997	99,993	0,007
32	0,001	0,003	35,998	99,995	0,005
33	0,001	0,002	35,999	99,998	0,002
34	0,001	0,001	36,000	99,999	0,001
35	0,000	0,000	36,000	100,000	0,000
36	0,000	0,000	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 10 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 10

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 10: 11/03/2005 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	9,518	26,440	9,518	26,440	73,560
2	6,546	18,183	16,064	44,623	55,377
3	4,533	12,591	20,597	57,214	42,786
4	3,322	9,228	23,919	66,442	33,558
5	3,065	8,513	26,984	74,955	25,045
6	1,821	5,060	28,805	80,015	19,985
7	1,177	3,270	29,983	83,285	16,715
8	1,011	2,807	30,993	86,092	13,908
9	0,848	2,356	31,841	88,448	11,552
10	0,805	2,236	32,646	90,685	9,315
11	0,723	2,008	33,369	92,693	7,307
12	0,445	1,237	33,815	93,930	6,070
13	0,358	0,995	34,173	94,925	5,075
14	0,298	0,829	34,471	95,754	4,246
15	0,288	0,800	34,759	96,554	3,446
16	0,216	0,600	34,975	97,153	2,847
17	0,179	0,498	35,154	97,651	2,349
18	0,171	0,475	35,326	98,126	1,874
19	0,132	0,368	35,458	98,495	1,505
20	0,102	0,283	35,560	98,777	1,223
21	0,086	0,238	35,645	99,015	0,985
22	0,068	0,188	35,713	99,203	0,797
23	0,055	0,153	35,768	99,355	0,645
24	0,052	0,144	35,820	99,500	0,500
25	0,048	0,132	35,867	99,632	0,368
26	0,033	0,091	35,900	99,723	0,277
27	0,023	0,063	35,923	99,786	0,214
28	0,019	0,054	35,942	99,840	0,160
29	0,015	0,040	35,957	99,880	0,120
30	0,012	0,033	35,969	99,914	0,086
31	0,010	0,028	35,979	99,942	0,058
32	0,008	0,023	35,987	99,965	0,035
33	0,007	0,019	35,994	99,984	0,016
34	0,003	0,009	35,997	99,993	0,007
35	0,001	0,004	35,999	99,997	0,003
36	0,001	0,003	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

Tabela 11 - Autovalores, Número de Autovalores, Variância Total e Acumulada em Porcentagem, Autovalores Acumulados e o Percentual de Ganho referente a Inserção das Variáveis no Modelo: Recorte de Cena 11

Matriz de Correlação dos Autovalores e Estatísticas Relacionadas para os Atributos da Cena 11: 14/06/2008 (Atributos Normalizados)					
Componentes	Autovalores	Variância Total %	Autovalores Acumulados	Variância Acumulada %	Percentual de Ganho
1	10,164	28,235	10,164	28,235	71,765
2	6,714	18,651	16,879	46,885	53,115
3	4,330	12,028	21,209	58,913	41,087
4	3,364	9,344	24,573	68,257	31,743
5	3,021	8,391	27,593	76,648	23,352
6	1,760	4,890	29,354	81,538	18,462
7	1,118	3,106	30,472	84,643	15,357
8	0,932	2,588	31,403	87,231	12,769
9	0,807	2,241	32,210	89,473	10,527
10	0,659	1,830	32,869	91,303	8,697
11	0,629	1,748	33,498	93,050	6,950
12	0,362	1,007	33,861	94,057	5,943
13	0,318	0,884	34,179	94,941	5,059
14	0,290	0,806	34,469	95,746	4,254
15	0,265	0,735	34,733	96,482	3,518
16	0,217	0,603	34,950	97,085	2,915
17	0,201	0,557	35,151	97,642	2,358
18	0,173	0,479	35,324	98,121	1,879
19	0,145	0,402	35,468	98,523	1,477
20	0,103	0,287	35,572	98,810	1,190
21	0,083	0,230	35,654	99,040	0,960
22	0,075	0,207	35,729	99,247	0,753
23	0,070	0,194	35,799	99,441	0,559
24	0,052	0,145	35,851	99,586	0,414
25	0,047	0,131	35,898	99,717	0,283
26	0,026	0,073	35,924	99,790	0,210
27	0,021	0,057	35,945	99,847	0,153
28	0,014	0,039	35,959	99,887	0,113
29	0,011	0,030	35,970	99,917	0,083
30	0,010	0,029	35,981	99,946	0,054
31	0,007	0,020	35,988	99,966	0,034
32	0,005	0,015	35,993	99,980	0,020
33	0,003	0,010	35,996	99,990	0,010
34	0,002	0,005	35,998	99,995	0,005
35	0,001	0,003	35,999	99,998	0,002
36	0,001	0,002	36,000	100,000	0,000

* Estão assinalados em vermelho o número de atributos que representa um ganho de pelo menos 20% na representatividade dos dados

APÊNDICE V:

Seleção de Atributos pelo Método Chi-Quadrado

Tabela 1 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phanton*: Recorte de Cena 1

<i>Ranking</i> dos Atributos Independentes	Chi-Quadrado	%	p - value	% Acumulado	% de Representatividade
Média	2875,313	0,152	0,000	0,15	
Mediana	2834,644	0,150	0,000	0,30	
Moda	2802,638	0,148	0,000	0,45	
Mínimo	2708,007	0,143	0,000	0,59	
Máximo	2436,396	0,129	0,000	0,72	
GLCM Média	846,676	0,045	0,000	0,77	
GLCM 2º Momento Angular	530,665	0,028	0,000	0,79	
Desvio Padrão	510,779	0,027	0,000	0,82	
Assimetria dos Polígonos	395,382	0,021	0,000	0,84	
GLCM Homogeneidade	375,551	0,020	0,000	0,86	
Desv. Padrão do Compr./to das Bordas	299,331	0,016	0,000	0,88	
Média do Compr./to das Bordas	235,619	0,012	0,000	0,89	
GLCM Correlação	165,573	0,009	0,000	0,90	
Compr./to do Trecho de Borda mais Longo	157,284	0,008	0,000	0,91	
Comprimento	144,907	0,008	0,000	0,91	
Comprimento/Largura	142,613	0,008	0,000	0,92	
Índice de Forma	133,597	0,007	0,000	0,93	
Compactação (Polígono)	121,792	0,006	0,000	0,94	
GLCM Entropia	119,768	0,006	0,000	0,94	
Ajuste Retangular	119,059	0,006	0,000	0,95	
Ajuste Elíptico	115,750	0,006	0,000	0,95	
GLCM Dissimilaridade	113,142	0,006	0,000	0,96	
GLCM Contraste	104,479	0,006	0,000	0,97	
Densidade	83,277	0,004	0,000	0,97	
Largura (Linha Principal)	76,798	0,004	0,000	0,97	
Compr./to de Linha Principal (no cycles)	76,331	0,004	0,000	0,98	
Máximo Compr./to dos Ramos	70,502	0,004	0,000	0,98	
Largura	68,806	0,004	0,000	0,99	
Compr./to de Linha Principal (regarding cycles)	55,438	0,003	0,000	0,99	
Largura da Borda	44,753	0,002	0,000	0,99	
Grau de Ramificação do Skeleton	39,678	0,002	0,000	0,99	
Assimetria	38,282	0,002	0,012	1,00	
Compactação	29,824	0,002	0,000	1,00	
Curtosi	27,534	0,001	0,025	1,00	
GLCM Desvio Padrão	18,769	0,001	0,764	1,00	
Área	12,771	0,001	0,386	1,00	

Tabela 2 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 2

<i>Ranking</i> dos Atributos Independentes	Chi-Quadrado	%	<i>p</i> -value	% Acumulado	% de Representatividade
Densidade	339,115	0,105	0,000	0,105	até 50%
Máximo	282,120	0,088	0,000	0,193	
Mínimo	278,178	0,086	0,000	0,279	
GLCM Média	272,361	0,085	0,000	0,364	
Assimetria dos Polígonos	175,932	0,055	0,000	0,418	
Comprimento/Largura	147,335	0,046	0,000	0,464	
GLCM 2º Momento Angular	141,761	0,044	0,000	0,508	
GLCM Entropia	127,809	0,040	0,000	0,548	
Ajuste Retangular	107,043	0,033	0,000	0,581	
Média	104,307	0,032	0,000	0,613	
GLCM Homogeneidade	98,989	0,031	0,000	0,644	até 75%
Desvio Padrão	96,684	0,030	0,000	0,674	
GLCM Desvio Padrão	95,912	0,030	0,000	0,704	
Índice de Forma	95,725	0,030	0,000	0,733	
Desv. Padrão do Compr./to das Bordas	73,291	0,023	0,000	0,756	
Comprimento	71,576	0,022	0,000	0,778	até 90%
Mediana	69,092	0,021	0,000	0,800	
GLCM Contraste	66,623	0,021	0,000	0,820	
Compactação (Polígono)	64,154	0,020	0,000	0,840	
Compactação	61,132	0,019	0,000	0,859	
Média do Compr./to das Bordas	55,611	0,017	0,000	0,877	
Assimetria	49,759	0,015	0,005	0,892	
CurtoSI	45,316	0,014	0,002	0,906	
GLCM Dissimilaridade	40,562	0,013	0,045	0,919	até 95%
Ajuste Elíptico	37,531	0,012	0,039	0,930	
GLCM Correlação	35,069	0,011	0,137	0,941	
Compr./to de Linha Principal (regarding cycles)	34,235	0,011	0,000	0,952	
Máximo Compr./to dos Ramos	28,439	0,009	0,019	0,961	até 100%
Compr./to de Linha Principal (no cycles)	23,966	0,007	0,001	0,968	
Largura da Borda	21,750	0,007	0,010	0,975	
Compr./to do Trecho de Borda mais Longo	21,750	0,007	0,001	0,982	
Largura (Linha Principal)	21,643	0,007	0,001	0,988	
Largura	18,411	0,006	0,104	0,994	
Grau de Ramificação do Skeleton	11,251	0,003	0,259	0,998	
Moda	7,814	0,002	0,553	1,000	
Área	0,121	0,000	0,989	1,000	

Tabela 3 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 3

Ranking dos Atributos Independentes	Chi-Quadrado	%	p-value	% Acumulado	% de Representatividade
Mediana	150,776	0,093	0,000	0,093	até 50%
Média	137,741	0,085	0,000	0,178	
Mínimo	134,087	0,083	0,000	0,260	
Moda	126,810	0,078	0,000	0,338	
Máximo	107,783	0,066	0,000	0,404	
GLCM Média	101,869	0,063	0,000	0,467	
Ajuste Retangular	77,685	0,048	0,000	0,515	
GLCM Correlação	72,225	0,044	0,000	0,559	
Densidade	67,179	0,041	0,000	0,601	
Índice de Forma	52,836	0,033	0,000	0,633	
GLCM Homogeneidade	46,357	0,029	0,000	0,662	até 75%
Comprimento	43,023	0,026	0,000	0,688	
Assimetria	42,114	0,026	0,004	0,714	
GLCM Dissimilaridade	39,919	0,025	0,008	0,739	
Comprimento/Largura	39,812	0,024	0,002	0,763	
Assimetria dos Polígonos	34,781	0,021	0,144	0,784	até 90%
Compr./to de Linha Principal (no cycles)	34,646	0,021	0,003	0,806	
Compactação (Polígono)	33,222	0,020	0,099	0,826	
GLCM 2º Momento Angular	31,859	0,020	0,000	0,846	
GLCM Desvio Padrão	27,518	0,017	0,154	0,863	
Largura da Borda	27,027	0,017	0,001	0,879	
Grau de Ramificação do Skeleton	24,852	0,015	0,016	0,895	
GLCM Contraste	24,824	0,015	0,052	0,910	
Ajuste Elíptico	24,368	0,015	0,276	0,925	
Largura (Linha Principal)	21,592	0,013	0,010	0,938	
Desvio Padrão	20,731	0,013	0,146	0,951	
Largura	12,280	0,008	0,423	0,959	até 100%
Curtosi	11,746	0,007	0,946	0,966	
Compr./to de Linha Principal (regarding cycles)	10,856	0,007	0,093	0,972	
Máximo Compr./to dos Ramos	10,281	0,006	0,591	0,979	
Desv. Padrão do Compr./to das Bordas	8,133	0,005	0,521	0,984	
Média do Compr./to das Bordas	8,133	0,005	0,521	0,989	
GLCM Entropia	7,868	0,005	0,248	0,994	
Compactação	7,607	0,005	0,055	0,998	
Compr./to do Trecho de Borda mais Longo	2,136	0,001	0,907	1,000	
Área	0,562	0,000	0,997	1,000	

Tabela 4 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 4

<i>Ranking</i> dos Atributos Independentes	Chi-Quadrado	%	p-value	% Acumulado	% de Representatividade
Moda	754,217	0,162	0,000	0,162	
Mediana	738,074	0,158	0,000	0,320	
Média	601,811	0,129	0,000	0,449	
Mínimo	579,527	0,124	0,000	0,573	
Máximo	240,376	0,051	0,000	0,624	
GLCM Média	166,516	0,036	0,000	0,660	
Comprimento	129,629	0,028	0,000	0,688	
Ajuste Retangular	93,257	0,020	0,000	0,708	
Densidade	88,927	0,019	0,000	0,727	
Média do Compr./to das Bordas	87,205	0,019	0,000	0,745	
Assimetria dos Polígonos	83,165	0,018	0,000	0,763	
Compr./to de Linha Principal (no cycles)	69,421	0,015	0,000	0,778	
Compactação	67,892	0,015	0,000	0,793	
GLCM Correlação	67,096	0,014	0,000	0,807	
Largura (Linha Principal)	60,609	0,013	0,000	0,820	
Índice de Forma	60,317	0,013	0,000	0,833	
Compactação (Polígono)	59,532	0,013	0,000	0,846	
Comprimento/Largura	50,735	0,011	0,000	0,856	
Desv. Padrão do Compr./to das Bordas	50,254	0,011	0,000	0,867	
GLCM Contraste	47,628	0,010	0,008	0,877	
Máximo Compr./to dos Ramos	46,093	0,010	0,000	0,887	
GLCM Dissimilaridade	46,024	0,010	0,013	0,897	
GLCM Entropia	45,991	0,010	0,013	0,907	
GLCM Desvio Padrão	43,740	0,009	0,022	0,916	
GLCM 2º Momento Angular	42,646	0,009	0,011	0,925	
GLCM Homogeneidade	42,156	0,009	0,032	0,934	
Assimetria	41,134	0,009	0,040	0,943	
Curtozi	39,226	0,008	0,000	0,952	
Ajuste Elíptico	36,638	0,008	0,048	0,960	
Largura	36,345	0,008	0,002	0,967	
Largura da Borda	32,595	0,007	0,001	0,974	
Área	31,588	0,007	0,002	0,981	
Compr./to do Trecho de Borda mais Longo	24,202	0,005	0,019	0,986	
Desvio Padrão	22,949	0,005	0,085	0,991	
Compr./to de Linha Principal (regarding cycles)	21,694	0,005	0,041	0,996	
Grau de Ramificação do Skeleton	19,453	0,004	0,078	1,000	

Tabela 5 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 5

Ranking dos Atributos Independentes	Chi-Quadrado	%	p - value	% Acumulado	% de Representatividade
Mediana	1932,603	0,142	0,000	0,142	até 50%
Moda	1904,259	0,139	0,000	0,281	
Média	1893,474	0,139	0,000	0,420	
Mínimo	1703,969	0,125	0,000	0,544	até 75%
Máximo	1178,008	0,086	0,000	0,631	
GLCM Média	539,936	0,040	0,000	0,670	
Ajuste Retangular	457,677	0,034	0,000	0,704	
GLCM Homogeneidade	362,612	0,027	0,000	0,730	
Compactação	306,594	0,022	0,000	0,753	
Densidade	292,742	0,021	0,000	0,774	
GLCM Entropia	262,595	0,019	0,000	0,793	até 90%
GLCM 2º Momento Angular	240,899	0,018	0,000	0,811	
GLCM Contraste	237,807	0,017	0,000	0,828	
GLCM Dissimilaridade	224,140	0,016	0,000	0,845	
Compr./to de Linha Principal (regarding cycles)	214,345	0,016	0,000	0,861	
Compr./to de Linha Principal (no cycles)	206,885	0,015	0,000	0,876	
Comprimento	157,282	0,012	0,000	0,887	
GLCM Correlação	157,071	0,012	0,000	0,899	até 95%
Índice de Forma	154,745	0,011	0,000	0,910	
Desvio Padrão	143,578	0,011	0,000	0,921	
Comprimento/Largura	140,419	0,010	0,000	0,931	
Largura (Linha Principal)	137,724	0,010	0,000	0,941	
Largura	130,070	0,010	0,000	0,950	
Compactação (Polígono)	93,066	0,007	0,000	0,957	até 100%
Área	80,587	0,006	0,000	0,963	
Máximo Compr./to dos Ramos	74,140	0,005	0,000	0,969	
Média do Compr./to das Bordas	67,745	0,005	0,000	0,974	
Largura da Borda	63,051	0,005	0,000	0,978	
Ajuste Elíptico	57,520	0,004	0,000	0,982	
Grau de Ramificação do Skeleton	55,910	0,004	0,000	0,987	
GLCM Desvio Padrão	45,550	0,003	0,000	0,990	
Compr./to do Trecho de Borda mais Longo	36,433	0,003	0,002	0,993	
Assimetria dos Polígonos	35,253	0,003	0,133	0,995	
Assimetria	28,217	0,002	0,400	0,997	
Desv. Padrão do Compr./to das Bordas	21,531	0,002	0,253	0,999	
Curtosi	17,208	0,001	0,142	1,000	

Tabela 6 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 6

Ranking dos Atributos Independentes	Chi-Quadrado	%	p-value	% Acumulado	% de Representatividade
Mínimo	49,149	0,069	0,000	0,069	até 50%
Média	45,601	0,064	0,000	0,134	
Mediana	45,059	0,064	0,000	0,197	
Desvio Padrão	40,105	0,057	0,000	0,254	
Moda	39,480	0,056	0,000	0,309	
GLCM Dissimilaridade	36,210	0,051	0,000	0,360	
GLCM Correlação	35,169	0,050	0,000	0,410	
GLCM Média	34,968	0,049	0,000	0,459	
GLCM Contraste	31,308	0,044	0,000	0,503	
Densidade	29,143	0,041	0,000	0,544	
Máximo	27,381	0,039	0,000	0,583	até 75%
GLCM Homogeneidade	25,304	0,036	0,001	0,619	
GLCM 2º Momento Angular	23,509	0,033	0,000	0,652	
Assimetria dos Polígonos	22,813	0,032	0,004	0,684	
Comprimento	22,244	0,031	0,000	0,715	
Comprimento/Largura	22,237	0,031	0,002	0,747	
Compactação (Polígono)	20,923	0,029	0,013	0,776	
Índice de Forma	20,582	0,029	0,001	0,805	até 90%
Ajuste Elíptico	17,310	0,024	0,027	0,830	
GLCM Desvio Padrão	16,833	0,024	0,010	0,853	
GLCM Entropia	16,817	0,024	0,005	0,877	
Ajuste Retangular	16,633	0,023	0,002	0,900	
Máximo Compr./dos Ramos	16,159	0,023	0,001	0,923	
Largura da Borda	7,819	0,011	0,050	0,934	até 95%
Compr./to de Linha Principal (no cycles)	7,429	0,010	0,024	0,945	
Grau de Ramificação do Skeleton	6,986	0,010	0,137	0,955	
Desv. Padrão do Compr./to das Bordas	5,727	0,008	0,126	0,963	
Média do Compr./to das Bordas	5,162	0,007	0,076	0,970	até 100%
Compactação	4,986	0,007	0,026	0,977	
Largura	3,868	0,005	0,145	0,982	
Assimetria	3,677	0,005	0,885	0,988	
Compr./to do Trecho de Borda mais Longo	3,521	0,005	0,172	0,992	
Curtosi	3,020	0,004	0,555	0,997	
Largura (Linha Principal)	1,186	0,002	0,553	0,998	
Compr./to de Linha Principal (regarding cycles)	0,705	0,001	0,703	0,999	
Área	0,420	0,001	0,517	1,000	

Tabela 7 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 7

Ranking dos Atributos Independentes	Chi-Quadrado	%	p - value	% Acumulado	% de Representatividade
GLCM Média	188,660	0,118	0,000	0,118	até 50%
Mínimo	163,080	0,102	0,000	0,220	
Média	106,929	0,067	0,000	0,287	
Mediana	104,691	0,066	0,000	0,353	
Moda	99,169	0,062	0,000	0,415	
Máximo	67,621	0,042	0,000	0,457	
Assimetria	66,719	0,042	0,000	0,499	
Compactação	52,757	0,033	0,001	0,532	até 75%
GLCM Homogeneidade	49,110	0,031	0,000	0,563	
Índice de Forma	47,469	0,030	0,000	0,593	
Ajuste Elíptico	47,362	0,030	0,003	0,622	
Assimetria dos Polígonos	42,721	0,027	0,028	0,649	
GLCM Entropia	42,662	0,027	0,028	0,676	
Ajuste Retangular	39,651	0,025	0,055	0,701	
Densidade	37,472	0,023	0,087	0,724	até 90%
GLCM Contraste	34,527	0,022	0,151	0,746	
Comprimento/Largura	32,370	0,020	0,020	0,766	
GLCM Dissimilaridade	31,632	0,020	0,246	0,786	
Grau de Ramificação do Skeleton	31,308	0,020	0,008	0,805	
Desvio Padrão	30,624	0,019	0,080	0,825	
GLCM 2º Momento Angular	29,952	0,019	0,038	0,843	
Comprimento	28,370	0,018	0,019	0,861	até 95%
GLCM Correlação	28,258	0,018	0,398	0,879	
Compr./to de Linha Principal (no cycles)	27,324	0,017	0,026	0,896	
Largura	24,068	0,015	0,064	0,911	
Máximo Compr./to dos Ramos	23,312	0,015	0,078	0,926	
Média do Compr./to das Bordas	21,784	0,014	0,592	0,939	
GLCM Desvio Padrão	20,111	0,013	0,826	0,952	
Desv. Padrão do Compr./to das Bordas	17,336	0,011	0,500	0,963	até 100%
Compactação (Polígono)	12,113	0,008	0,207	0,970	
Largura (Linha Principal)	10,720	0,007	0,772	0,977	
Largura da Borda	10,619	0,007	0,562	0,984	
Curtosi	9,934	0,006	0,622	0,990	
Compr./to do Trecho de Borda mais Longo	7,817	0,005	0,799	0,995	
Compr./to de Linha Principal (regarding cycles)	4,143	0,003	0,902	0,997	
Área	4,071	0,003	0,907	1,000	

Tabela 8 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phanton*: Recorte de Cena 8

<i>Ranking</i> dos Atributos Independentes	Chi-Quadrado	%	p-value	% Acumulado	% de Representatividade
Mediana	587,071	0,154	0,000	0,154	até 50%
Moda	585,717	0,154	0,000	0,308	
Média	571,305	0,150	0,000	0,457	
Mínimo	533,220	0,140	0,000	0,597	
Máximo	419,099	0,110	0,000	0,707	
GLCM Média	106,452	0,028	0,000	0,735	
Ajuste Retangular	88,142	0,023	0,000	0,758	até 90%
GLCM Homogeneidade	73,322	0,019	0,000	0,777	
Ajuste Elíptico	68,499	0,018	0,000	0,795	
Densidade	56,996	0,015	0,000	0,810	
GLCM 2º Momento Angular	54,888	0,014	0,000	0,825	
Índice de Forma	48,502	0,013	0,000	0,837	
Assimetria dos Polígonos	47,313	0,012	0,009	0,850	
GLCM Correlação	44,944	0,012	0,016	0,862	
Compr./to de Linha Principal (no cycles)	43,919	0,012	0,000	0,873	
Desvio Padrão	40,866	0,011	0,002	0,884	
GLCM Dissimilaridade	40,654	0,011	0,044	0,894	
Comprimento	38,948	0,010	0,001	0,905	
GLCM Contraste	36,704	0,010	0,047	0,914	até 95%
GLCM Entropia	32,887	0,009	0,017	0,923	
Comprimento/Largura	31,775	0,008	0,023	0,931	
Assimetria	31,506	0,008	0,140	0,939	
Largura (Linha Principal)	30,722	0,008	0,031	0,948	
Média do Compr./to das Bordas	29,417	0,008	0,044	0,955	
Largura	28,518	0,007	0,019	0,963	até 100%
Compr./to de Linha Principal (regarding cycles)	23,884	0,006	0,067	0,969	
Compactação (Polígono)	18,053	0,005	0,035	0,974	
Curtosi	15,402	0,004	0,220	0,978	
Largura da Borda	13,313	0,003	0,347	0,981	
Grau de Ramificação do Skeleton	12,791	0,003	0,384	0,985	
Área	11,923	0,003	0,452	0,988	
Máximo Compr./to dos Ramos	11,319	0,003	0,730	0,991	
Compactação	10,193	0,003	0,117	0,993	
Compr./to do Trecho de Borda mais Longo	9,833	0,003	0,830	0,996	
GLCM Desvio Padrão	7,819	0,002	0,996	0,998	
Desv. Padrão do Compr./to das Bordas	7,539	0,002	0,941	1,000	

Tabela 9 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 9

<i>Ranking</i> dos Atributos Independentes	Chi-Quadrado	%	<i>p</i> -value	% Acumulado	% de Representatividade
Mediana	344,113	0,121	0,000	0,121	até 50%
Mínimo	311,964	0,110	0,000	0,231	
Moda	305,486	0,108	0,000	0,339	
Média	188,418	0,066	0,000	0,406	
Máximo	175,845	0,062	0,000	0,468	
Densidade	123,439	0,044	0,000	0,511	
GLCM Média	123,146	0,043	0,000	0,555	até 75%
Ajuste Retangular	115,356	0,041	0,000	0,595	
Assimetria dos Polígonos	110,236	0,039	0,000	0,634	
Desvio Padrão	96,863	0,034	0,000	0,668	
GLCM 2º Momento Angular	78,972	0,028	0,000	0,696	
GLCM Homogeneidade	78,469	0,028	0,000	0,724	
Média do Comprimento das Bordas	73,467	0,026	0,000	0,750	até 90%
GLCM Correlação	68,916	0,024	0,000	0,774	
Comprimento	65,483	0,023	0,000	0,797	
Comprimento/Largura	64,816	0,023	0,000	0,820	
GLCM Dissimilaridade	51,789	0,018	0,003	0,838	
Comprimento de Linha Principal (no cycles)	45,650	0,016	0,000	0,854	
Largura	39,040	0,014	0,001	0,868	
Ajuste Elíptico	37,757	0,013	0,037	0,881	até 95%
Compactação (Polígono)	36,102	0,013	0,000	0,894	
Índice de Forma	32,679	0,012	0,018	0,906	
GLCM Contraste	31,884	0,011	0,060	0,917	
Largura (Linha Principal)	31,238	0,011	0,008	0,928	
Desv. Padrão do Comprimento das Bordas	28,539	0,010	0,005	0,938	
Largura da Borda	24,444	0,009	0,018	0,947	até 100%
GLCM Desvio Padrão	23,802	0,008	0,473	0,955	
Assimetria	20,494	0,007	0,809	0,962	
Máximo Comprimento dos Ramos	20,399	0,007	0,157	0,969	
GLCM Entropia	20,295	0,007	0,161	0,977	
Compactação	16,467	0,006	0,001	0,982	
Curtozi	13,679	0,005	0,550	0,987	
Grau de Ramificação do Skeleton	10,254	0,004	0,803	0,991	
Comprimento do Trecho de Borda mais Longo	9,227	0,003	0,683	0,994	
Comprimento de Linha Principal (regarding cycles)	8,776	0,003	0,187	0,997	
Área	7,784	0,003	0,556	1,000	

Tabela 10 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 10

Ranking dos Atributos Independentes	Chi-Quadrado	%	p-value	% Acumulado	% de Representatividade
Moda	1480,862	0,148	0,000	0,148	até 50%
Mediana	1470,773	0,147	0,000	0,296	
Média	1433,100	0,143	0,000	0,439	
Mínimo	1216,873	0,122	0,000	0,561	até 75%
Máximo	962,869	0,096	0,000	0,657	
GLCM Correlação	508,367	0,051	0,000	0,708	
GLCM Média	401,517	0,040	0,000	0,748	
GLCM Homogeneidade	365,963	0,037	0,000	0,785	
Desvio Padrão	302,776	0,030	0,000	0,815	
Média do Compr./to das Bordas	184,090	0,018	0,000	0,834	
GLCM Contraste	152,664	0,015	0,000	0,849	
Assimetria dos Polígonos	136,320	0,014	0,000	0,863	até 90%
Desv. Padrão do Compr./to das Bordas	134,686	0,013	0,000	0,876	
Comprimento	120,242	0,012	0,000	0,888	
GLCM 2º Momento Angular	105,524	0,011	0,000	0,899	
GLCM Dissimilaridade	103,045	0,010	0,000	0,909	
GLCM Entropia	97,867	0,010	0,000	0,919	
Compactação	82,586	0,008	0,000	0,927	
Curtozi	72,291	0,007	0,000	0,934	até 95%
Índice de Forma	67,113	0,007	0,000	0,941	
Comprimento/Largura	64,971	0,007	0,000	0,948	
Compr./to de Linha Principal (no cycles)	63,462	0,006	0,000	0,954	
Largura (Linha Principal)	58,924	0,006	0,000	0,960	
Compr./to do Trecho de Borda mais Longo	57,467	0,006	0,000	0,966	
Largura	52,996	0,005	0,000	0,971	
Ajuste Retangular	46,891	0,005	0,003	0,976	
Ajuste Elíptico	43,070	0,004	0,026	0,980	
Compactação (Polígono)	39,881	0,004	0,022	0,984	
Máximo Compr./to dos Ramos	32,064	0,003	0,006	0,987	até 100%
Compr./to de Linha Principal (regarding cycles)	30,827	0,003	0,002	0,990	
Assimetria	22,554	0,002	0,208	0,992	
GLCM Desvio Padrão	16,792	0,002	0,857	0,994	
Área	16,022	0,002	0,190	0,996	
Densidade	15,984	0,002	0,001	0,997	
Grau de Ramificação do Skeleton	13,202	0,001	0,355	0,999	
Largura da Borda	13,136	0,001	0,359	1,000	

Tabela 11 – *Ranking* dos atributos indicados pelo *Chi-Quadrado* como mais representativos em relação à Classe do *Phantom*: Recorte de Cena 11

Ranking dos Atributos Independentes	Chi-Quadrado	%	P-value	% Acumulado	% de Representatividade
Mediana	545,869	0,133	0,000	0,133	até 50%
Média	506,087	0,123	0,000	0,256	
Moda	457,196	0,111	0,000	0,367	
Mínimo	370,374	0,090	0,000	0,457	
Máximo	326,401	0,079	0,000	0,536	
GLCM Média	240,214	0,058	0,000	0,594	até 75%
GLCM Correlação	212,705	0,052	0,000	0,646	
Assimetria dos Polígonos	153,471	0,037	0,000	0,683	
Desvio Padrão	107,885	0,026	0,000	0,709	
Assimetria	93,063	0,023	0,000	0,732	
GLCM Homogeneidade	92,732	0,023	0,000	0,754	
Média do Compr./to das Bordas	69,102	0,017	0,000	0,771	até 90%
Densidade	65,519	0,016	0,000	0,787	
GLCM Desvio Padrão	63,352	0,015	0,000	0,802	
Curtosi	62,571	0,015	0,000	0,818	
GLCM Dissimilaridade	56,239	0,014	0,001	0,831	
Ajuste Elíptico	55,017	0,013	0,001	0,845	
GLCM Contraste	48,545	0,012	0,007	0,856	
Comprimento	47,061	0,011	0,000	0,868	
GLCM 2º Momento Angular	47,060	0,011	0,001	0,879	
GLCM Entropia	46,543	0,011	0,011	0,891	
Largura	46,118	0,011	0,000	0,902	até 95%
Compactação	45,076	0,011	0,016	0,913	
Desv. Padrão do Compr./to das Bordas	44,583	0,011	0,000	0,924	
Ajuste Retangular	40,465	0,010	0,046	0,933	
Compr./to do Trecho de Borda mais Longo	40,009	0,010	0,000	0,943	
Compactação (Polígono)	38,408	0,009	0,003	0,953	até 100%
Índice de Forma	34,839	0,008	0,010	0,961	
Largura (Linha Principal)	34,646	0,008	0,003	0,969	
Largura da Borda	30,988	0,008	0,002	0,977	
Compr./to de Linha Principal (no cycles)	30,261	0,007	0,011	0,984	
Comprimento/Largura	24,490	0,006	0,270	0,990	
Máximo Compr./to dos Ramos	14,538	0,004	0,485	0,994	
Compr./to de Linha Principal (regarding cycles)	11,648	0,003	0,234	0,997	
Grau de Ramificação do Skeleton	9,626	0,002	0,649	0,999	
Área	4,455	0,001	0,974	1,000	

Figura 1; Ranking da importância indicado pelo Chi-Quadrado para os 36 atributos, efetuado para as 11 cenas.

