

PONTIFÍCIA UNIVERSIDADE CATÓLICA  
DO RIO DE JANEIRO



**Pedro Marco Achancaray Diaz**

## **A Comparison of Segmentation Algorithms for Remote Sensing**

**Dissertação de Mestrado**

Dissertation presented to the Programa de Pós-Graduação em Engenharia Elétrica of the Departamento de Engenharia Elétrica, PUC-Rio as partial fulfillment of the requirements for the degree of Mestre em Engenharia Elétrica.

Advisor: Prof. Raul Queiroz Feitosa

Rio de Janeiro

May 2014



**Pedro Marco Achancaray Diaz**

## **A Comparison of Segmentation Algorithms for Remote Sensing**

Dissertation presented to the Programa de Pós-Graduação em Engenharia Elétrica of the Departamento de Engenharia Eletrica do Centro Técnico Científico da PUC-Rio, as partial fulfillment of the requirements for the degree of Mestre.

**Prof. Raul Queiroz Feitosa**

Advisor

Departamento de Engenharia Elétrica – PUC-Rio

**Prof. Sidnei Paciornik**

Departamento de Engenharia de Materiais – PUC-Rio

**Profa. Maria Luiza Velloso**

UERJ

**Prof. José Eugenio Leal**

Coordenador Setorial do Centro  
Técnico Científico

Rio de Janeiro, May 9th, 2014

All rights reserved.

**Pedro Marco Achancaray Diaz**

The author graduated in Mechanical and Electrical Engineering from Universidad Nacional de Ingenieria -UNI in 2010.

Bibliographic Data

Achancaray Diaz, Pedro Marco

A Comparison of Segmentation Algorithms for Remote Sensing / Pedro Marco Achancaray Diaz; advisor: Raul Queiroz Feitosa. – 2014.

84 f. : il. (color.) ; 30 cm

Dissertação (Mestrado em Engenharia Eletrica) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2014.

Incluí bibliografia.

1. Engenharia Elétrica – Teses. 2. Segmentação de Imagens. 3. Sensoriamento remoto. 4. Sintonização de Parâmetros. I. Feitosa, Raul Queiroz. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

For my grandparents, Maximo and Liduvina.  
For my parents, Pedro and Mercedes.  
For my brothers, Ever, Saul and David.  
For Karen, my love.

## Acknowledgments

I am truly grateful to my advisor, Prof. Raul Queiroz Feitosa, for the encouragement, his advices, stimulating talks and generous support throughout the development of my dissertation.

I thank my parents, Pedro and Mercedes, my brothers, Ever, Saul and David, for their support and unconditional love.

I want to thank all my colleagues from LVC, for the companionship and valuable scientific discussion, and all my friends I met during my studies.

I also gratefully acknowledge the financial support of CAPES and FAPERJ.

## Resumo

Achanccaray Diaz, Pedro Marco; Feitosa, Raul Queiroz. **Uma avaliação de Métodos de Segmentação para Aplicações em Sensoriamento Remoto.** Rio de Janeiro, 2014. 84p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação tem como objetivo avaliar algoritmos de segmentação para imagens de sensoriamento remoto. Quatro algoritmos de segmentação foram considerados neste estudo. Esses algoritmos têm abordagens diferentes tais como baseado em agrupamento, em crescimento de regiões, em modelos bayesianos e em grafos. Como cada algoritmo tem os seus próprios parâmetros, o processo de encontrar seus parâmetros ótimos foi feito usando um algoritmo de otimização, Nelder - Mead. O algoritmo Nelder - Mead procura os melhores parâmetros para cada algoritmo de segmentação, isto é, os parâmetros que proporcionam os resultados mais exatos com respeito a uma referência dada. A função objetivo foi definida a partir de sete métricas diferentes. Eles avaliam qualitativamente o resultado da segmentação baseadas na sua referência. Os experimentos foram realizados ao longo de três imagens de sensoriamento remoto de diferentes localidades do Brasil. Isso envolveu um total de 84 experimentos. Os resultados mostraram que as abordagens baseadas em grafos produzem os melhores resultados baseados em todas as métricas. As abordagens baseadas no crescimento de regiões e agrupamento apresentaram-se como boas opções para imagens de sensoriamento remoto.

## Palavras-chave

Segmentação; Sensoriamento Remoto; Sintonização de Parâmetros.

## Abstract

Achanccaray Diaz, Pedro Marco; Feitosa, Raul Queiroz (Advisor). **A Comparison of Segmentation Algorithms for Remote Sensing**. Rio de Janeiro, 2014. 84p. Master Dissertation - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This dissertation aims to evaluate segmentation algorithms for remote sensing images. Four segmentation algorithms were considered in this study. These algorithms have different approaches such as clustering-based, region growing-based, bayesian-based and graph-based. As each algorithm has its own parameters, the process to find their optimum values was done using an optimization algorithm, Nelder – Mead. Nelder – Mead algorithm looks for the best parameters for each segmentation algorithm, i.e. the parameters that provide the most accurate results with respect to a given reference. The objective function was defined by seven different metrics. These metrics assess qualitatively the segmentation result based on its reference. The experiments were performed over three remote sensing images from different locations of Brazil. A total of 84 experiments have been performed. The results have shown that graph-based approaches produce the best results based on each metric. The region growing- and clustering-based approaches have shown to be good alternatives for remote sensing images.

## Keywords

Segmentation; Remote Sensing; Parameter Tuning

# Contents

1 INTRODUCTION	17
1.1. Objectives of the dissertation	18
1.2. Organization of the remainder parts	19
2 STATE OF THE ART	20
2.1. Image Segmentation in Remote Sensing	20
2.2. Segmentation Evaluation	25
2.3. Methodology	27
3 SEGMENTATION ALGORITHMS	29
3.1. Mean-Shift Segmentation (MS)	29
3.2. Graph-based Segmentation (Gb)	32
3.3. Region Merging – based Segmentation (Rm)	35
3.4. Conditional Random Fields – based Segmentation (CRFb)	38
3.4.1. Semantic Segmentation	38
3.4.2. Generative vs. Discriminative models	39
3.4.3. Conditional Random Fields (CRF)	40
4 NELDER – MEAD ALGORITHM	42
4.1. The Method	42
4.2. Simplex Initialization	43
4.3. Reflection	44
4.4. Expansion	45
4.5. Contraction	45
5 METRICS	47
5.1. Hoover metric (H)	47
5.2. Area Fit Index (AFI)	48
5.3. Shape Index (SI)	49
5.4. Rand Index (RI)	49

5.5. Precision-Recall (F)	50
5.6. Precision and Recall plots	52
5.7. Segmentation Covering (C)	53
5.8. Reference Bounded Segments Booster (RBSB)	54
6 RESULTS	56
6.1. Dataset	56
6.2. Experimental Procedure	59
6.3. Visual Quality Assessment	61
6.4. Comparison based on the selected metrics	69
6.5. Precision and Recall plots	71
6.6. Optimization algorithm	75
7 CONCLUSIONS	76
BIBLIOGRAPHIC REFERENCES	78

## List of Figures

Figure 1: High level taxonomy of image segmentation algorithms (taken from Vantaram & Saber (2012)).	21
Figure 2: Classification of segmentation algorithms based on a low-level taxonomy. The blue rectangles represent common segmentation algorithms applied in remote sensing while the red ones represent the selected algorithms for this dissertation (modified from Vantaram & Saber (2012)).	22
Figure 3: Classification of methods for evaluation of segmentation quality according to Zhang (modified from Zhang (2001)).	25
Figure 4: Methodology followed for the development of this study.	27
Figure 5: Iterations of Mean Shift algorithm. An initial window is chosen and its center of mass is shifted until it converges (modified from Bradski & Kaehler (2008)).	31
Figure 6: Two possible merging regions, C1 (blue) and C2 (red), and the final result $C1 \cup C2$ (green).	37
Figure 7: Description of the Nelder – Mead algorithm.	43
Figure 8: Reflection Operation (modified from (Mathews & Fink, 2004)).	44
Figure 9: Expansion Operation (modified from (Mathews & Fink, 2004)).	45
Figure 10: Contraction Operation (modified from (Mathews & Fink, 2004)).	46
Figure 11: In case of a failed contraction, shrinking the triangle towards PI is done (modified from (Mathews & Fink, 2004)).	46
Figure 12: Precision and Recall definition. A segment of the segmentation outcome appears in red and its respective reference in green. There are four main regions, $t_p$ represents the true positives, $f_n$ the false negatives, $f_p$ the false positives and all pixels outside the union of the green and red regions are considered true negatives ( $t_n$ ).	51

Figure 13: Interpretation of areas (corners) in a Precision-Recall plot.	52
Figure 14: Explanation of over-segmentation and under-segmentation corners of a Precision-Recall plot. The result of the segmentation (S) and its ground truth (GT) are represented by the red and green contours respectively. In the first case (left), there are more $f_n$ than $f_p$ , which favor the over-segmentation. In the second case (right), the number of $f_p$ is greater than $f_n$ , which is related to an under-segmentation.	53
Figure 15: REDUC Image – Image 1 (top) and its ground truth (bottom).	57
Figure 16: Image of Maragogipe City – Image 2 (top) and its ground truth (bottom).	58
Figure 17: Image de Congonhas – Image 3 (top) and its ground truth (bottom).	59
Figure 18: Results of the experiments with Image 1, each row from top to bottom represents a metric: H, AFI, SI, RI, F, C and RBSB. From left to right, each column represents: GT, MS result, Gb result, Rm result and CRFb result.	62
Figure 19: Results of the experiments with Image 1, each row from top to bottom represents a metric: H, AFI, SI, RI, F, C and RBSB. From left to right, each column represents: GT, MS result, Gb result, Rm result and CRFb result.	63
Figure 20: Results of the experiments with Image 1, each row from top to bottom represents a metric: H, AFI, SI, RI, F, C and RBSB. From left to right, each column represents: GT, MS result, Gb result, Rm result and CRFb result.	64
Figure 21: Results of the experiments with Image 2, each row from top to bottom represents a metric: H, AFI, SI, RI, F, C and RBSB. From left to right, each column represents:GT, MS result, Gb result, Rm result and CRFb result.	65
Figure 22: Results of the experiments with Image 2, each row from top to bottom represents a metric: H, AFI, SI, RI, F, C and RBSB.	

From left to right, each column represents:GT, MS result, Gb result, Rm result and CRFb result.	66
Figure 23: Results of the experiments with Image 3, each row from top to bottom represents a metric: H, AFI, SI, RI, F, C and RBSB. From left to right, each column represents:GT, MS result, Gb result, Rm result and CRFb result.	67
Figure 24: Results of the experiments with Image 3, each row from top to bottom represents a metric: H, AFI, SI, RI, F, C and RBSB. From left to right, each column represents:GT, MS result, Gb result, Rm result and CRFb result.	68
Figure 25: Precision-Recall plot for Image 1. Each point represents an iteration of the optimization algorithm. Each color represents a segmentation algorithm (see legends).	72
Figure 26: Precision-Recall plot for Image 2. Each point represents an iteration of the optimization algorithm. Each color represents a segmentation algorithm (see legends).	73
Figure 27: Precision-Recall plot for Image 3. Each point represents an iteration of the optimization algorithm. Each color represents a segmentation algorithm (see legends).	73

## List of Tables

Table 1: Metrics for segmentation evaluation selected for this study.	55
Table 2: Results obtained for each metric by the optimization algorithm for Image 1.	70
Table 3: Results obtained for each metric by the optimization algorithm for Image 2.	71
Table 4: Results obtained for each metric by the optimization algorithm for Image 3.	71
Table 5: Number of iterations needed by the optimization algorithm to find the best parameters for each segmentation algorithm.	74

## List of Symbols and Abbreviations

$MS$	Mean-Shift segmentation.
$Gb$	Graph-based segmentation.
$Rm$	Region Merging-based segmentation.
$CRFb$	Conditional Random Fields-based segmentation.
$NM$	Nelder – Mead algorithm.
$K$	Kernel function.
$d$	Represents the number of dimensions in a space.
$x$	$d$ – dimensional feature vector.
$x_i$	$d$ – dimensional feature vector of pixel $i$ .
$\mu$	Mean vector.
$\Sigma$	Covariance matrix.
$\sigma$	Standard deviation.
$\eta$	Number of Gaussians.
$P(x)$	Probability density function (pdf)
$\Omega$	Constant associated to Gaussian pdfs.
$\nabla P(x)$	Gradient of a function.
$\bar{x}$	Position of the center of mass.
$\Delta\bar{x}$	Mean-shift vector.
$G$	An undirected graph.
$G'$	A sub-graph.
$V$	Vertices of a graph $G$ .
$E$	Edges of a graph $G$ .
$w()$	Weight of an edge.
$S$	A segmentation outcome or machine segmentation.
$GT$	Ground truth.
$C$	A component of a segmentation $S$ , a segment.
$N$	Number of pixels in an image.
$N_S$	Number of segments in segmentation $S$ .
$N_{GT}$	Number of segments in ground truth $GT$ .
$O_i$	Number of pixels in the segment $C_i$ from $S$ .

$O_j$	Number of pixels in the segment $C_j$ from $GT$ .
$O_{ij}$	Number of pixels in both segments $C_i$ and $C_j$ .
$D$	A predicate defined to find boundaries.
$T$	A tree from a graph.
$ST$	A spanning tree.
$MST$	A minimum spanning tree.
$Int(C)$	Internal difference of a component $C$
$Dif(C_1, C_2)$	Difference between two components $C_1$ and $C_2$ .
$MInt(C_1, C_2)$	Minimum internal difference between $C_1$ and $C_2$ .
$\tau$	Threshold function.
$\theta$	Constant parameter related to graph-based segmentation.
$f$	Merging cost or degree of fitting.
$h_{color}, h_{shape}$	Spectral and shape components.
$\omega_{color}, \omega_L$	Spectral and band weights.
$A$	Area of a region or segment in pixels.
$L$	Spectral band.
$C_1 \cup C_2$	Resulting region after merging $C_1$ and $C_2$ .
$Sol$	Solidity.
$Comp$	Compactness.
$Abox$	Area of a bounding box.
$dmax$	Length of the major axis of an ellipse.
$E(\mathbf{x})$	Energy function.
$\psi_i(x_i)$	Unary potentials.
$\psi_{ij}(x_i, x_j)$	Pairwise cliques.
$N_i$	Neighborhood of pixel $i$ .
$\Gamma(i, j)$	Edge feature based on spectral difference of neighboring pixels.
$n$	Number of variables.
$P_0, P_1, \dots, P_n$	Points in $n$ – dimensional space defining a simplex.
$y_i$	Value of the objective function at $P_i$
$P_l, P_h$	Points where $y_i$ take its lowest and highest values respectively.
$P_m$	Centroid of the points without considering $P_h$ .
$\alpha$	Reflection constant.

$\beta$	Contraction constant.
$\gamma$	Expansion constant.
$P_r$	Point determined by reflection.
$P_e$	Point determined by expansion.
$P_c$	Point determined by contraction.
$CD$	Number of correct detections.
$\rho_k$	Perimeter of a segment $C_k$ .
$A_{l.i.}$	Area, in pixels, of the segment with largest intersection.