

PONTIFÍCIA UNIVERSIDADE CATÓLICA  
DO RIO DE JANEIRO



**Vanessa de Oliveira Campos**

**Segmentação Multicritério para Detecção de Nódulos  
Pulmonares em Imagens de Tomografia Computadorizada**

**Tese de Doutorado**

Tese apresentada ao Programa de Pós-graduação  
em Engenharia Elétrica da PUC-Rio como requisito  
parcial para a obtenção do título de Doutor em  
Ciências em Engenharia Elétrica

Orientador: Raul Queiroz Feitosa  
Coorientador: Aristófanês Corrêa Silva

Rio de Janeiro, novembro de 2009



**Vanessa de Oliveira Campos**

**Segmentação Multicritério para Detecção de  
Nódulos Pulmonares em Imagens de  
Tomografia Computadorizada**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Prof. Raul Queiroz Feitosa**

Orientador

Departamento de Engenharia Elétrica – PUC-Rio

**Prof. Alexandre Street de Aguiar**

Departamento de Engenharia Elétrica – PUC-Rio

**Profa. Aura Conci**

UFF

**Prof. Marco Aurélio Cavalcanti Pacheco**

Departamento de Engenharia Elétrica – PUC-Rio

**Profa. Maria Luiza Fernandes Velloso**

UERJ

**Prof. Rodolfo Acatauassú Nunes**

UERJ

**Prof. José Eugenio Leal**

Coordenador Setorial do Centro Técnico Científico

Rio de Janeiro, 30 de novembro de 2009

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Vanessa de Oliveira Campos**

Graduou-se em Ciência da Computação na UFMT (Universidade Federal de Mato Grosso) em 2002 e recebeu o grau de Mestre em Ciências em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro em 2005.

#### Ficha Catalográfica

Campos, Vanessa de Oliveira

Segmentação multicritério para detecção de nódulos pulmonares em imagens de tomografia computadorizada / Vanessa de Oliveira Campos ; orientador: Raul Queiroz Feitosa ; coorientador: Aristófares Corrêa Silva. – 2009.

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

Tese (Doutorado em Engenharia Elétrica)–Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.

Inclui referências bibliográficas.

1. Engenharia elétrica – Teses. 2. Detecção auxiliada por computador (CAD). 3. Nódulo pulmonar. 4. Segmentação multicritério. 5. Tomografia computadorizada. I. Feitosa, Raul Queiroz. II. Silva, Aristófares Corrêa. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Aos maiores responsáveis pelas minhas conquistas:

Maria Carmem de Oliveira,  
minha mãezinha linda.

Melitina de Oliveira,  
minha vovozinha.

Antonio Graciano de Oliveira,  
meu vovozinho.

## Agradecimentos

A Deus que, com seu infinito amor e bondade, está sempre presente na minha vida.

Aos meus pais e familiares que sempre me deram apoio e carinho sempre com muito amor.

Ao meu orientador Dr. Raul Queiroz Feitosa pela instrução e paciência ao longo deste trabalho.

Ao meu coorientador Dr. Aristófanês Corrêa Silva que sempre esteve gentilmente disposto a ajudar.

Ao Dr. Rodolfo Acatauassu Nunes, pelo imenso auxílio com seus conhecimentos médicos no processo de interpretação de imagens de tomografia computadorizada, em especial, para detecção de nódulos pulmonares.

Ao Hospital Universitário Pedro Ernesto do Rio de Janeiro e ao grupo de desenvolvimento do LIDC pelo fornecimento do banco de dados.

À CAPES e à PUC-Rio, pelos auxílios financeiros concedidos.

Aos professores que participaram da Comissão examinadora.

A todos os professores e funcionários do Departamento pelos ensinamentos e pela ajuda.

A todos os meus amigos que, de uma forma ou de outra, me deram força.

## Resumo

Campos, Vanessa de Oliveira; Feitosa, Raul Queiroz; Silva, Aristófanés Corrêa. **Segmentação Multicritério para Detecção de Nódulos Pulmonares em Imagens de Tomografia Computadorizada.** Rio de Janeiro, 2009, 109 p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho propõe um novo algoritmo de segmentação baseado em crescimento de regiões para detecção de nódulos pulmonares em imagens de tomografia computadorizada. Para decidir, em cada iteração, se dois objetos adjacentes são fundidos em um único objeto, o algoritmo de segmentação calcula um índice de heterogeneidade baseada em múltiplos critérios. Entretanto, o algoritmo de segmentação depende de alguns parâmetros os quais foram encontrados utilizando algoritmo genético. Resultados experimentais mostraram que o método é robusto e promissor (chegando a uma sensibilidade de 80,9 % com 0,23 falsos positivos por exame). Além disso, indicam que o método proposto é capaz de fornecer um bom suporte para o diagnóstico do especialista.

## Palavras-chave

Detecção auxiliada por computador (CAD); nódulo pulmonar; segmentação multicritério; tomografia computadorizada.

## Abstract

Campos, Vanessa de Oliveira; Feitosa, Raul Queiroz (Advisor); Silva, Aristófanés Corrêa (Co-advisor). **Multicriterion Segmentation for Lung Nodule Detection in Computed Tomography**. Rio de Janeiro, 2009. 109 p. D.Sc. Thesis – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This study proposes a novel segmentation algorithm for lung nodules detection in thorax computed tomography (CT). In order to decide, at each iteration, whether two adjacent objects should be merged or not, a region growing procedure calculates a heterogeneity growth based on multiple criteria. However, segmentation algorithm depends on some parameters which were found by genetic algorithm. Results produced by the proposed segmentation were closely consistent with the reference segments provided manually by an expert physician. The detection itself achieved 80,9% sensitivity with 0,23 false positive per slice which indicates that the proposed method is able to provide a good suggestion for the specialist. Results indicate the potential of proposed segmentation method and encourage a further investigation aiming at improving its accuracy.

## Keywords

Computer-aided detection (CAD); pulmonary nodule; multicriterion segmentation; computed tomography.

## Sumário

1. Introdução	16
1.1. Escopo e objetivos desta pesquisa	19
1.2. Organização da tese	19
2. Trabalhos Relacionados	21
2.1. Detecção de nódulos pulmonares	21
2.1.1. Delimitação do pulmão	22
2.1.2. Segmentação de nódulos	23
2.1.3. Identificação de nódulos	24
2.1.4. Redução de falsos positivos	25
2.2. Diagnóstico de nódulos pulmonares	25
3. Metodologia de Detecção de Nódulos Pulmonares	27
3.1. Passos fundamentais	27
3.2. Aquisição de imagens	28
3.3. Segmentação	29
3.3.1. Pré-segmentação	30
3.3.2. Segmentação Multicritério	32
3.3.3. Ajuste automático dos parâmetros de segmentação	35
3.3.3.1. Função de avaliação	38
3.4. Representação	40
3.4.1. Normalização	41
3.4.2. Seleção de variáveis	41
3.5. Reconhecimento	42
3.6. Indicadores de desempenho	43
4. Avaliação experimental	45
4.1. Base de dados e programa de implementação	45
4.1.1. Base de dados HUPE	45
4.1.2. Base de dados LIDC	46

4.2. Divisão da base de dados para treinamento e teste	47
4.2.1. Base de dados HUPE	48
4.2.2. Base de dados LIDC	48
4.3. Experimentos com o primeiro grupo de critérios para a segmentação	49
4.3.1. Critérios selecionados para segmentação de nódulos	49
4.3.2. Ajuste automático dos valores dos parâmetros de segmentação	51
4.3.2.1. Convergência do algoritmo genético	52
4.3.2.2. Análise de estabilidade dos valores dos parâmetros do modelo	56
4.3.3. Detecção de nódulos	58
4.4. Experimentos com o segundo conjunto de critérios para a segmentação	65
4.4.1. Critérios selecionados para segmentação de nódulos	66
4.4.2. Ajuste automático dos valores dos parâmetros de segmentação	67
4.4.2.1. Convergência do algoritmo genético	67
4.4.2.2. Análise de estabilidade dos valores dos parâmetros do modelo	70
4.4.3. Segmentação multicritério	71
4.4.4. Detecção de nódulos	80
4.5. Custo computacional	85
5. Conclusões	86
Trabalhos futuros	87
6. Referências Bibliográficas	88
Apêndice I – Limiarização Global pelo Algoritmo de Otsu	94
Apêndice II – Cálculo da Área Superficial e do Volume de um Objeto	96
Apêndice III – Medidas Descritivas dos Objetos	101

## Lista de figuras

Figura 1 – Exemplos de nódulos pulmonares	18
Figura 2 – Passos fundamentais para interpretação de imagens.	27
Figura 3 – Fatia de uma TC do tórax.	29
Figura 4 – Etapas da pré-segmentação representadas por uma fatia do volume todo.	31
Figura 5 – Segmentação multicritério (onde $H$ refere-se ao aumento de heterogeneidade provocado pela fusão de dois objetos adjacentes).	34
Figura 6 – Esquema geral de um algoritmo genético.	37
Figura 7 – Representação gráfica para a função de avaliação para casos onde existe um único conjunto de referência.	39
Figura 8 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o primeiro conjunto de critérios no banco de dados HUPE.	53
Figura 9 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o primeiro conjunto de critérios a partir do primeiro grupo de imagens do banco de dados LIDC.	54
Figura 10 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o primeiro conjunto de critérios a partir do segundo grupo de imagens do banco de dados LIDC.	54
Figura 11 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o primeiro conjunto de critérios a partir do terceiro grupo de imagens do banco de dados LIDC.	55
Figura 12 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o primeiro conjunto de critérios no banco de dados LIDC.	55

Figura 13 – Desempenho da detecção de nódulos pulmonares na base de dados HUPE.	61
Figura 14 – Falsos positivos por exame da detecção de nódulos pulmonares na base de dados HUPE.	61
Figura 15 – Desempenho da detecção de nódulos pulmonares na base de dados LIDC com o primeiro conjunto de critérios de segmentação.	64
Figura 16 – Falsos positivos por fatia da detecção de nódulos pulmonares na base de dados LIDC com o primeiro conjunto de critérios de segmentação.	64
Figura 17 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o segundo conjunto de critérios a partir do primeiro grupo de imagens do banco de dados LIDC.	68
Figura 18 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o segundo conjunto de critérios a partir do segundo grupo de imagens do banco de dados LIDC.	68
Figura 19 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o segundo conjunto de critérios a partir do terceiro grupo de imagens do banco de dados LIDC.	69
Figura 20 – Convergência de algoritmos genéticos para ajuste de valores dos parâmetros da segmentação com o segundo conjunto de critérios na base de dados LIDC.	69
Figura 21 – Nódulo pulmonar delimitado por quatro especialistas e pela segmentação multicritério.	73
Figura 22 – Duas visualizações de um mesmo nódulo sob diferentes ângulos.	74
Figura 23 – Nódulo pulmonar delimitado por quatro especialistas e pela segmentação multicritério.	75
Figura 24 – Duas visualizações de um mesmo nódulo sob diferentes ângulos.	76
Figura 25 – Nódulo pulmonar delimitado por quatro especialistas e	

pela segmentação multicritério.	78
Figura 26 – Duas visualizações de um mesmo nódulo sob diferentes ângulos.	78
Figura 27 – Sensibilidade da detecção de nódulos pulmonares em diferentes segmentações da mesma base de dados LIDC.	83
Figura 28 – Falsos positivos encontrados por fatia na detecção de nódulos pulmonares em diferentes segmentações da mesma base de dados.	83
Figura 29 – Dimensões de um voxel.	96
Figura 30 – Exemplo de GLCM em uma imagem 3D.	107

## Lista de tabelas

Tabela 1 – Matriz de resultados para detecção de nódulos.	43
Tabela 2 – Resultados do procedimento de parametrização da segmentação com o primeiro conjunto de critérios com base no banco de dados HUPE.	56
Tabela 3 – Resultados do procedimento de parametrização da segmentação com o primeiro conjunto de critérios com base no primeiro grupo de imagens do LIDC.	57
Tabela 4 – Resultados do procedimento de parametrização da segmentação com o primeiro conjunto de critérios com base no segundo grupo de imagens do LIDC.	57
Tabela 5 – Resultados do procedimento de parametrização da segmentação com o primeiro conjunto de critérios com base no terceiro grupo de imagens do LIDC.	58
Tabela 6 – Resultados da detecção de nódulos pulmonares em trabalhos encontrados na literatura.	59
Tabela 7 – Desempenho de trabalhos sobre detecção de nódulos utilizando o LIDC.	59
Tabela 8 – Resultados da detecção de nódulos pulmonares no banco de dados HUPE.	60
Tabela 9 – Resultados da detecção de nódulos pulmonares no banco de dados LIDC.	63
Tabela 10 – Resultados do procedimento de parametrização da segmentação com o segundo conjunto de critérios com base no primeiro grupo de imagens LIDC.	70
Tabela 11 – Resultados do procedimento de parametrização da segmentação com o segundo conjunto de critérios com base no segundo grupo de imagens LIDC.	70
Tabela 12 – Resultados do procedimento de parametrização da segmentação com o segundo conjunto de critérios com base no	

terceiro grupo de imagens.	71
Tabela 13 – Avaliação da segmentação multicritério com o segundo conjunto de critérios na base de dados LIDC.	80
Tabela 14 – Desempenho de diferentes segmentações avaliadas sobre o banco de dados LIDC.	82

## Lista de siglas

CAD	Detecção Auxiliada por Computador ( <i>Computer-aided Detection</i> )
CADx	Diagnóstico Auxiliado por Computador ( <i>Computer-aided Diagnosis</i> )
DICOM	Representação Digital para Comunicação em Medicina ( <i>Digital Imaging Communications in Medicine</i> )
FN	Falso Negativo
FP	Falso Positivo
GA	Algoritmos Genéticos ( <i>Genetic Algorithm</i> )
GLCM	Matriz de coocorrência de níveis de cinza ( <i>Gray-Level Co-occurrence Matrix</i> )
HU	Unidade Hounsfield
LIDC	Consórcio de Banco de Dados de Imagem de Pulmão ( <i>Lung Image Database Consortium</i> )
ROI	Região de Interesse ( <i>Region Of Interest</i> )
SVM	Máquina de Suporte Vetorial ( <i>Support Vector Machine</i> )
TC	Tomografia Computadorizada
XML	eXtensible Markup Language