

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**

Vilson Heck Junior

**Criação de uma Metodologia Computacional para a
Segmentação Bimodal de Patologias Timpânicas**

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação

Prof. Dr. rer.nat. Eros Comunello

Florianópolis, novembro de 2008.

CRIAÇÃO DE UMA METODOLOGIA COMPUTACIONAL PARA A SEGMENTAÇÃO BIMODAL DE PATOLOGIAS TIMPÂNICAS

Vilson Heck Junior

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação Área de Concentração Sistemas de Computação, e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

Prof. Dr. Frank Augusto Siqueira
Coordenador do Curso

Banca Examinadora

Prof. Dr. rer.nat. Eros Comunello
Orientador

Prof. Dr. Sady Selaimen da Costa

Prof^ª. Dr.^a Olga Regina Pereira Bellon

Prof^ª. Dr.^a Silvia Modesto Nassar

Aos meus pais e ao meu irmão, pelo afeto, apoio, carinho, educação e amor incondicionais, que serviram como forte alicerce de sustentação e força de impulsão para eu alcançar os meus objetivos.

À Greice pelo carinho e compreensão que me ajudaram a guiar esforços através deste trabalho.

À equipe de profissionais do Centro de Otite Média Crônica do Brasil, em especial aos doutores Sady e Cristina, pelo amparo técnico/médico e amizade.

Aos colegas e amigos do grupo Cyclops que estiveram sempre ao meu lado, em especial aos professores Eros e Aldo pela oportunidade de fazer parte desta equipe, onde tive a singular oportunidade de aprender.

SUMÁRIO

Capítulo 1	1
Introdução.....	1
1.1 Definição do Problema	2
1.2 Motivação	2
1.3 Objetivos do Trabalho	3
1.3.1 Objetivo Geral	3
1.3.2 Objetivos Específicos	3
1.4 Hipótese	3
1.5 Materiais e Métodos.....	4
1.6 Contribuições	5
1.7 Estudo do Trabalho	6
Capítulo 2	7
O Sistema Auditivo Humano.....	7
2.1 A Orelha Externa	7
2.2 A Orelha Média	7
2.2.1 A Membrana Timpânica.....	8
2.3 A Orelha Interna	9
2.4 Otite Média	9
2.4.1 Otite Média Crônica	9
2.5 Timpanoesclerose	10
2.6 Videoscopia Digital	10
2.7 Estado da Arte na Mensuração de Perfurações Timpânicas	11
Capítulo 3	13
Imagens Digitais.....	13
3.1 Elementos da Imagem.....	14

3.2	Imagens Binárias.....	15
3.3	Imagens Monocromáticas	16
3.4	Imagens Coloridas	16
3.4.1	A Cor.....	17
3.5	Modelos de Cor.....	17
3.5.1	Modelo de cor RGB	18
Capítulo 4		19
Processamento Digital de Imagens.....		19
4.1	Etapas do Processamento de Imagens.....	19
4.1.1	Aquisição de imagens digitais.....	20
4.1.2	Pré-processamento	21
4.1.2.1	Filtro Passa-Baixa.....	21
4.1.3	Segmentação.....	22
4.1.4	Representação e Descrição.....	22
4.1.5	Reconhecimento e Interpretação	23
4.2	Segmentação de Imagens	23
4.2.1	Detecção de Bordas	23
4.2.1.1	Operador de Sobel	24
4.2.1.2	Canny.....	25
4.2.2	Segmentação Orientada a Regiões	27
4.2.2.1	Mumford & Shah.....	27
4.2.2.2	Contornos Ativos.....	28
Capítulo 5		32
Metodologia.....		32
5.1	Pré-processamento	32
5.2	Segmentação Bimodal Semi-Automatizada	33

5.3	Definição da Área Timpânica	36
5.4	Validação	38
Capítulo 6	41
Resultados	41
6.1	Resultados da Validação	42
Capítulo 7	45
Discussões e Conclusões	45
7.1	Objetivos Alcançados	45
Referências	47
Anexos	50
A.	Medidas de Distância.....	50
A.1	Índice Rand	50
A.2	Correspondência de Grafos Bipartidos (BGM).....	51
B.	Tabela de Resultados	52

LISTA DE ILUSTRAÇÕES

Figura 1 - Sistema Auditivo Humano - [E] Orelha Externa, [M] Orelha Média e [I] Orelha Interna (Fonseca, Vasco, & Ferreira, 2002).....	7
Figura 2 - Orelha Média e seus principais componentes (Fonseca, Vasco, & Ferreira, 2002)	8
Figura 3 - Tímpano saudável – videotoscopia digital	8
Figura 4 - OMCNC - Tímpano perfurado	10
Figura 5 - Tímpano portador de timpanoesclerose.....	10
Figura 6 - Lenna colorida (Lenna, 1972) - Exemplo de imagem	13
Figura 7 - Representação de imagens digitais	13
Figura 8 - Espectro eletromagnético. Adaptado de (Pedrini & Schwartz, 2008)	14
Figura 9 – Lenna (Figura 6) – Imagem em versão binária	15
Figura 10 - Lenna (Figura 6) - Imagem em versão monocromática.....	16
Figura 11 - Faixa de comprimentos de ondas eletromagnéticas, em nanômetros, em relação à cor interpretada pelo olho humano (Gonzalez & Woods, 2000).....	17
Figura 12 - Cubo RGB	18
Figura 13 - Etapas do processamento de imagens. Adaptado de (Gonzalez & Woods, 2000) ...	20
Figura 14 - Digitalização de imagens.....	21
Figura 15 - Operador Sobel aplicado à Figura 6	24
Figura 16 – Detector de bordas Canny aplicado à Figura 6	25
Figura 17 – Resultado da aplicação do algoritmo de Mumford & Shah à Figura 6 com $\lambda = 1400$	28
Figura 18 - Fluxograma do algoritmo de Contornos Ativos segundo (Li, Xu, Gui, & Fox, 2005)	29
Figura 19 - Aplicação de Contornos Ativos - (a) Imagem processada com contorno inicial na cor verde; em (b), (c) e (d) evolução da curva após 600, 1100 e 1600 iterações, respectivamente.....	30
Figura 20 - Resultado da aplicação de 10 iterações do filtro passa-baixa. Imagens originais de (a) e (b): Figura 4 e Figura 5, respectivamente	32
Figura 21 - Histograma da Figura 4 (Imagem original)	33
Figura 22 - Histograma da Figura 20 (a) (Após aplicação do filtro).....	33
Figura 23 - Resultado da segmentação com o algoritmo de Mumford & Shah	34
Figura 24 - Exemplo de área inicial sendo desenhada por especialista em (a) e área inicial resultante em (b)	35
Figura 25 - Resultados dos processos de segmentação. Utilizando em (a) <i>Mumford & Shah</i> e em (b) <i>Contornos Ativos</i>	36
Figura 26 - Seleção elíptica do tímpano. Em (a) processo de seleção e em (b) imagem binária resultante	37
Figura 27 - Exemplos de <i>Ground Truths</i>	38

Figura 28 - Exemplos de Falso Negativo e Falso Positivo. Em (a) Segmentação timpânica e em (b) segmentação de patologia.....	39
Figura 29 - Resultado da mensuração do paciente 02 com 14,19% de perfuração. Em (a) imagem original, em (b) segmentação do tímpano e em (c) segmentação da patologia... 41	41
Figura 30 - Resultado da mensuração do paciente 03 com 23,90% de perfuração. Em (a) imagem original, em (b) segmentação do tímpano e em (c) segmentação da patologia... 41	41
Figura 31 - Correlação de Mensurações: resultados esperados em relação aos resultados obtidos	42
Figura 32 - Correlação de Timpanos: resultados esperados em relação aos resultados obtidos (em pixels)	43
Figura 33 - Correlação de Patologias: resultados esperados em relação aos resultados obtidos (em pixels)	44

LISTA DE TABELAS

Tabela 1 - Média e Desvio Padrão para o tamanho das patologias do conjunto avaliado.....	42
Tabela 2 - Avaliação da segmentação da área timpânica: Média e Desvio Padrão do índice Rand, índice BGM, percentual de Falso Negativo e percentual de Falso Positivo aplicados aos 39 casos.....	43
Tabela 3 - Avaliação da segmentação da área patológica: Média e Desvio Padrão do índice Rand, índice BGM, percentual de Falso Negativo e percentual de Falso Positivo aplicados aos 39 casos.....	44
Tabela 4- Resultados dos 39 casos utilizados para validação do método	52

ABREVIACOES

AOMC	Ambulatório de Otite Média Crônica
API	Application Programming Interface
BGM	<i>Bipartite Graph Matching</i> (Correspondência de Grafos Bipartidos)
CIE	Commission Internationale de L'Eclairage
CMY	Cyan, Magenta and Yellow
Com.BR	Centro de Otite Média do Brasil
GTs	<i>Ground Truths</i>
HCPA	Hospital de Clínicas de Porto Alegre
nm	Nanômetro
OMC	Otite Média Crônica
OMCC	Otite Média Crônica Colesteatomatosa
OMCNC	Otite Média Crônica Não Colesteatomatosa
PDI	Processamento Digital de Imagens
RGB	Red, Green and Blue
THZ	Terahertz

LISTA DE SIMBOLOS

θ	Ângulo
DX	Derivada no eixo X
DY	Derivada no eixo Y
G	Gradiente aproximado de uma imagem
G_x	Gradiente em X de uma imagem
G_y	Gradiente em Y de uma imagem
I	Imagem (Matriz de intensidades luminosas ou cores)
M	Máscara de convolução

RESUMO

Este trabalho mostra o desenvolvimento de uma metodologia computacional para fazer a segmentação bimodal de patologias timpânicas em imagens videotoscópicas digitais. Baseado nas necessidades apresentadas pelo Centro de Otite Média Crônica do Brasil, onde são efetuadas diariamente várias videoscopias digitais, verificou-se a necessidade de quantificar computacionalmente tais patologias.

As imagens videotoscópicas possuem baixo contraste e um reduzido conjunto de cores. Estes fatores tornam o desenvolvimento de uma metodologia semi-automatizada mais viável do que o desenvolvimento de uma metodologia totalmente automatizada. Depois de segmentada a imagem videotoscópica, é possível efetuar a mensuração das patologias segmentadas. Esta mensuração permite o acompanhamento da evolução da patologia e dos tratamentos aplicados à patologia. Para o processo de mensuração foi necessário criar um método de definição da área timpânica, a qual será comparada com a área da patologia para se obter o percentual afetado.

Os resultados obtidos com o uso da metodologia desenvolvida foram submetidos a processos de validação por similaridade conhecidos como Rand e BGM, os resultados foram também submetidos a cálculos para aferir as áreas percentuais obtidas que representam resultados falso-positivos ou falso-negativos.

Palavras-Chave: Segmentação de imagens; Mensuração de patologias; Otite Média Crônica.

ABSTRACT

This work presents the development of a computational methodology for the bimodal segmentation of tympanic pathologies in digital video-otoscopic images. Based on the needs exposed by the Brazilian Chronic Otitis Media Center, where several video-otoscopic images are produced daily, the need to computationally quantify such pathologies arose.

Video-otoscopic images possess a low contrast and a limited set of colors. These factors make the development of a semiautomated methodology more feasible than the development of a completely automated methodology. After the video-otoscopic image has been segmented, one can measure the pathologies that were segmented. This measuring allows the follow-up of the pathology and the treatment thereof. A method for the definition of the tympanic area was devised so that the ratio between the pathology area and the tympanic area can be computed.

The results obtained by using the methodology underwent two similarity-based validation processes, Rand and BGM. They were also subjected to calculations aiming at determining the percentage of the pathology areas that was misclassified and thus yielded false-positives or false-negatives.

Keywords: Image segmentation; Pathology measuring; Chronic otitis media.

Capítulo 1

Introdução

Sistemas computacionais provêm facilidades a todas as áreas da ciência. Nestes termos, a Ciência da Computação prove também recursos para o desenvolvimento da própria ciência em diversas áreas. Soluções genéricas para resolver diversos tipos de problemas diferentes são soluções complexas. Devido a esta complexidade a ciência da computação procura soluções específicas e menores, de forma a desenvolver, ou aprimorar, métodos que aos poucos amadurecem a ponto de crescer e resolver cada vez mais problemas.

A Visão Computacional é uma área da computação que utiliza diversos métodos e técnicas de processamento a fim de tornar imagens interpretáveis por computadores. O processo de interpretação pode ser definido como a transformação de um conjunto de dados digitais representando uma imagem (mono, bi, tri ou tetradimensional) em uma estrutura de dados que descreva uma semântica para um contexto (Wangenheim, 2007).

O Conjunto de transformações efetuados na visão computacional pode ser dividido em etapas, tais como filtragem, segmentação, representação entre outras. Para cada etapa, existe uma coleção diversificada de algoritmos. Cada algoritmo possui vantagens e desvantagens em relação a determinadas características encontradas nas imagens, por isso, para construir um sistema de visão computacional, é fundamental uma pesquisa completa em cada uma das etapas, combinando algoritmos de forma a atingir os resultados esperados.

A necessidade do desenvolvimento de tecnologias que provêm suporte a outras ciências é mostrada neste trabalho, onde uma parceria é estabelecida com o Centro de Otiite Média do Brasil (Com.BR – do Hospital de Clínicas de Porto Alegre / Universidade Federal do Rio Grande do Sul). Este centro demonstra a falta de ferramentas computacionais-otológicas que auxiliem no processo de diagnóstico e acompanhamento das patologias encontradas em seu banco de casos formado por mais de 700 pacientes com exames videotoscópicos bilaterais.

Este trabalho apresenta a pesquisa feita para o desenvolvimento de uma metodologia computacional, que obtêm bons resultados no processo de segmentação de patologias da membrana timpânica, em imagens de videotoscopia digital. A seguir é apresentada a definição do problema e questões que a permeiam.

1.1 Definição do Problema

A Otologia é a área da medicina que estuda o sistema auditivo humano, onde umas das patologias com maior prevalência mundial é a Otite Média Crônica. Sua principal característica é o rompimento da membrana timpânica, caracterizando a perfuração timpânica. Outra patologia muito comum é a timpanoesclerose.

Estas patologias são diagnosticadas através de imagens geradas por um exame chamado de videotoscopia digital. Neste exame é feito uma filmagem da membrana timpânica do paciente, onde pode ser observada a perfuração timpânica e o remanescente da membrana não afetada.

A partir destas imagens, os médicos fazem a análise da área afetada pela patologia no tímpano do paciente. Esta análise pode ter objetivos diferentes, desde constatar a patologia, até mesmo acompanhar a evolução da patologia ou de seu tratamento, bem como a comparação de estados entre um pré e um pós-operatório.

O desenvolvimento de um método computacional para amparar este processo médico envolve uma área da Ciência da Computação, conhecida como Processamento Digital de Imagens, para quantificar as áreas das patologias encontradas nas imagens videotoscópicas. Estas imagens apresentam um conjunto de cores reduzido e gradientes tênues dividindo as diversas estruturas encontradas na imagem, isto dificulta o processo de segmentação das imagens que possibilitaria a quantificação de regiões / patologias.

1.2 Motivação

A ausência de métodos computacionais específicos para o suporte a esta subárea da medicina, provendo auxílio à mensuração de patologias, como forma de acompanhamento ao tratamento, é o primeiro fator motivamente do desenvolvimento deste trabalho. A possibilidade do acompanhamento quantitativo destas patologias

facilita as pesquisas nesta área da medicina, trazendo possíveis novas formas de tratamento, mais eficazes e/ou mais baratas.

Outro fator é que técnicas criadas para esta abordagem, não somente poderão servir para auxiliar a medicina no processo de evolução de patologias otológicas, mas como também pode trazer novas abordagens computacionais para o tratamento de outros problemas encontrados na área de processamento digital de imagens.

1.3 Objetivos do Trabalho

1.3.1 Objetivo Geral

Criar uma metodologia computacional de processamento de imagens capaz de efetuar a segmentação bimodal de perfurações timpânicas e de placas de timpanoesclerose que viabiliza o desenvolvimento de um processo computacional semi-automático de mensuração de patologias timpânicas.

1.3.2 Objetivos Específicos

1. Pesquisar o estado da arte no que tange soluções computacionais para doenças timpânicas;
2. Desenvolver um método de geração de *seed* inicial para um processo semi-automático de segmentação de imagens videotoscópicas digitais;
3. Compor um método para segmentar perfurações timpânicas ou placas de timpanoesclerose a partir do *seed* inicial fornecido.
4. Criar um método de definição da região timpânica que permita a correlação entre área da região timpânica e a área total afetada pela patologia segmentada.
5. Identificar métodos que permitam avaliar a similaridade dos resultados obtidos pela metodologia desenvolvida neste trabalho com imagens de referência (*Ground Truths*).

1.4 Hipótese

Através do uso de métodos de processamento digital de imagens é possível efetuar, com o auxílio de um especialista, a segmentação de imagens videotoscópicas digitais, e com isto obter a área afetada por patologias em membranas timpânicas, tornando possível a obtenção de informações quantitativas referentes à área patológica segmentada.

Com o auxílio de um processo computacional de mensuração, utilizando e adaptando técnicas computacionais já existentes, estima-se que é possível melhorar o processo de avaliação e de acompanhamento ao tratamento de patologias timpânicas.

1.5 Materiais e Métodos

Neste trabalho foi desenvolvida uma metodologia computacional para se fazer a segmentação de perfurações timpânicas e placas de timpanoesclerose, dentro de um processo de mensuração semi-automatizado. Isto foi feito através do uso de técnicas de processamento de imagens digital, área que foi amplamente pesquisada e adaptada conforme as necessidades encontradas. Técnicas matemáticas são, também, utilizadas, tanto no processamento digital de imagens, quanto num segundo processo também abrangido para tornar possível a mensuração das regiões, o processo de definição da região timpânica.

A base de materiais e conhecimento disponível para criação deste trabalho teve origem em duas grandes áreas diferentes: Medicina e Ciência da Computação. As informações são oriundas de várias fontes, mas entre elas destacam-se artigos científicos, livros e conhecimento intrínseco de especialistas de ambas as áreas.

A exposição do problema foi feita pela equipe de especialistas médicos do Centro de Otite Média Crônica do Brasil – Com.BR – do Hospital de Clínicas de Porto Alegre – HCPA. Esta equipe foi quem forneceu todas as informações inerentes ao conteúdo médico deste trabalho bem como todos os arquivos de videoscopias digitais, fundamentais ao processo de mensuração.

As videoscopias digitais (imagens com as patologias) foram fornecidas pelo grupo de pesquisa do HCPA, onde são tratados mais de 700 pacientes, com imagens de ambas as orelhas, totalizando mais de 1400 exames. Os pacientes podem ser

classificados em dois grupos de patologias diferentes: Otite Média Crônica e Timpanoesclerose.

Para fazer a mensuração das imagens foi necessário criar métodos para a segmentação de duas áreas, a área da patologia e a área total do tímpano. Na segmentação da patologia foi definido um algoritmo para segmentação que utiliza a geração de um *seed* inicial (área inicial para o processo de segmentação). A segmentação da área timpânica foi feita através da sua definição (seleção) com uma forma geométrica elíptica. Depois de segmentar as áreas, é calculado o percentual da membrana timpânica abrangida pela patologia.

Para finalizar, a metodologia foi validada através da avaliação de: 1) similaridade com os resultados esperados; 2) diferença entre os resultados esperados e os resultados obtidos; e 3) percentual de áreas falso negativo e de áreas falso positivo. Para isso, foram selecionadas aleatoriamente 39 imagens do banco de 1400 casos, e para cada uma das imagens selecionadas, foram criadas manualmente, por um especialista, as imagens de segmentação esperadas (os resultados esperados / padrão ouro / *ground-truth*).

A similaridade, entre os resultados obtidos e os resultados esperados, foi avaliada por métodos de verificação de similaridade de *clusters* conhecidos como Rand e BGM, além de ter sido verificada a diferença das áreas encontradas, diferenças essas que são chamadas de falso positivo e falso negativo.

1.6 Contribuições

Durante a realização deste trabalho, várias contribuições relevantes às áreas de aplicação da pesquisa foram desenvolvidas. Dentre elas:

- A criação de uma metodologia de segmentação de patologias mensuráveis por extensão do tímpano;
- A criação de um método matemático para a definição da região timpânica;
- A disponibilização para o meio médico de uma ferramenta para a mensuração de patologias timpânicas, principalmente perfurações e timpanoesclerose.

1.7 Estrutura do Trabalho

O presente trabalho está estruturado em sete capítulos, os quais compõem cinco grandes partes:

Parte I – Introdução

A contextualização é mostrada nesta parte do trabalho através da definição do problema, descrição da motivação, objetivos e da hipótese, bem como são descritos os materiais e métodos utilizados no desenvolvimento do trabalho. Esta parte compreende ao Capítulo 1.

Parte II – Fundamentação teórica

Nesta parte do trabalho são revisados todos os aspectos teóricos necessários à compreensão deste trabalho, que vão desde aspectos anatômicos e médicos até os conceitos de computacionais de processamento de imagens envolvidos. Esta parte compreende aos Capítulos 2, 3 e 4.

Parte III – Metodologia de desenvolvimento

Apresentação da metodologia desenvolvida através da exibição de cada passo executado na tentativa de criação da metodologia. Nesta parte também foi descrito o processo definido para a validação do método. Esta parte compreende ao Capítulo 5.

Parte IV – Validação, Resultados e Conclusões

Aqui são mostrados os resultados obtidos com o uso da metodologia desenvolvida, bem como são mostradas as informações obtidas pelo processo de validação quando aplicado aos resultados obtidos. Para encerrar, são feitas algumas conclusões sobre o trabalho, os resultados e objetivos alcançados. Esta parte compreende aos Capítulos 6 e 7.

Capítulo 2

O Sistema Auditivo Humano

O Sistema auditivo humano é responsável pela percepção sonora nos seres humanos, além de participar de outras funções como a do equilíbrio. Ele é composto por dois subsistemas auditivos – duas orelhas, esquerda e direita – os quais podem ser divididos em três grandes partes: Orelha Externa, Orelha Média e Orelha Interna.

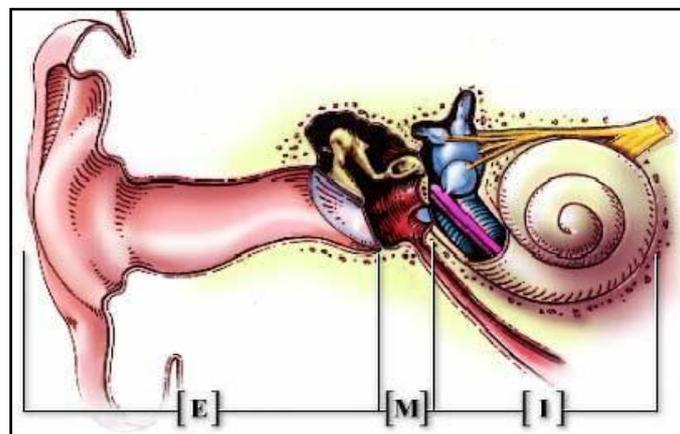


Figura 1 - Sistema Auditivo Humano - [E] Orelha Externa, [M] Orelha Média e [I] Orelha Interna (Fonseca, Vasco, & Ferreira, 2002)

2.1 A Orelha Externa

A orelha externa é composta pelo pavilhão auricular e conduto auditivo. Ela é responsável por captar os sinais sonoros externos e o conduzi-los até a orelha média.

Estabelecendo a comunicação entre o meio externo e a orelha média, a orelha externa tem cerca de três centímetros de comprimento e esta escavada no osso temporal. Internamente ela é revestida por pêlos e glândulas, que fabricam o cerume.

2.2 A Orelha Média

A orelha média é uma cavidade com ar, situada atrás da membrana timpânica. Ela é responsável por fazer a conversão dos sinais sonoros, captados pela orelha externa, em sinais mecânicos, e retransmiti-los a orelha interna.

Os principais componentes da orelha média são: Membrana Timpânica, Martelo, Bigorna, Estribo e Trompa de Eustáquio (também conhecida como Tuba Auditiva).

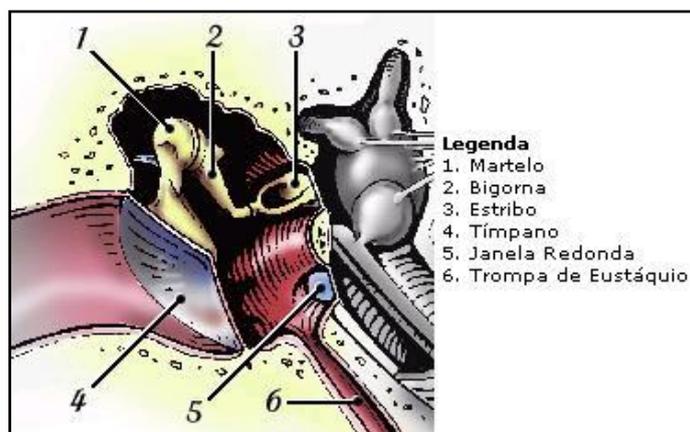


Figura 2 - Orelha Média e seus principais componentes (Fonseca, Vasco, & Ferreira, 2002)

2.2.1 A Membrana Timpânica

A Membrana timpânica se localiza na orelha média e faz a separação da orelha média com a orelha externa. Ela é responsável por duas principais tarefas: impedir a entrada de sujeira ou bactérias na orelha média e retransmitir o som que chega até a orelha média para a cadeia ossicular.



Figura 3 - Tímpano saudável – videotoscopia digital

2.3 A Orelha Interna

A conversão do sinal mecânico/sonoro em impulsos elétricos é feita na orelha interna. Esta conversão é um dos processos mais importantes do processo de audição, pois esses sinais elétricos são conduzidos até o cérebro, onde os mesmos são interpretados. A conversão é feita dentro da cóclea e a condução do sinal elétrico até o cérebro é feita pelo nervo auditivo.

Na orelha interna encontramos a cóclea, responsável pela audição e também uma parte posterior, relacionada com o equilíbrio e constituída pelo vestíbulo e pelos canais semicirculares.

2.4 Otite Média

A Otite Média é um processo inflamatório, infeccioso ou não, localizado focal ou generalizadamente na orelha média (Costa & Dornelles, 2006). É uma doença de alta prevalência mundial. A Otite Média pode se apresentar em diversas intensidades e formas, sendo a mais grave delas conhecida como Otite Média Crônica.

O processo inflamatório em si, muitas vezes pode extrapolar os limites do osso temporal, gerando complicações graves e mesmo fatais; ao comprometimento do estado geral do indivíduo e à perda auditiva funcional, reversível ou não (Costa & Dornelles, 2006).

2.4.1 Otite Média Crônica

Longos períodos sem tratamento ou reincidências da Otite Média Crônica, principalmente aguda, podem tornar a patologia crônica. A característica mais comum da cronificação é o rompimento da membrana timpânica (Otite Média Crônica não Colesteatomatosa - OMCNC). Outros casos, considerados como OMC, também podem ser caracterizados pela retração timpânica de média a severa (OMCNC), ou ainda, em graus muito elevados, a presença de colesteatomas (Otite Média Crônica Colesteatomatosa - OMCC).



Figura 4 - OMCNC - Tímpano perfurado

2.5 Timpanoesclerose

A timpanoesclerose é uma alteração histológica que ocorre na mucosa da orelha média. Pode evoluir com calcificação ou ossificação. Quando atinge a membrana timpânica, ela pode ser diagnosticada à otoscopia, tomando um aspecto de placas brancas com extensão variável, geralmente causa perda auditiva condutiva (Costa, Dornelles, Schmidt, Wangenheim, Comunello, & Heck Junior, 2008).



Figura 5 - Tímpano portador de timpanoesclerose

2.6 Videotoscopia Digital

O processo de diagnóstico das patologias timpânicas envolve diversas variáveis e avaliações por parte do médico, porém, um dos exames mais dominantes é a Videotoscopia ou Otoscopia. Ambos os exames se caracterizam pela visualização da membrana timpânica, através da introdução na orelha de uma lente, geralmente em forma de fibra e iluminada.

A principal diferença entre a otoscopia e a videoscopia é que na otoscopia o médico observa o tímpano do paciente diretamente na lente do equipamento, enquanto na videoscopia uma câmera grava o exame em formato de vídeo e exhibe simultaneamente o mesmo em um dispositivo de visualização, como um monitor ou televisor.

A Videoscopia Digital é uma versão da Videoscopia onde o exame é gravado no formato digital, direto em um computador ou equipamento gravador digital.

Nas Figuras: Figura 3, Figura 4 e Figura 5; é possível observar imagens geradas a partir de exames de videoscopia digitais.

2.7 Estado da Arte na Mensuração de Perfurações Timpânicas

O processo de mensuração empregado tradicionalmente pelos especialistas é feito através da subjetividade do observador. Observando a imagem contendo a perfuração, o especialista cria, mentalmente, quatro divisões no tímpano, chamadas quadrantes. Baseado nesta orientação, o especialista então calcula mentalmente o percentual de cada um dos quadrantes afetados.

Depois de gerar os exames de videoscopia digital, eles são armazenados e gerenciados de diversas formas. No grupo colaborador do desenvolvimento desta pesquisa, o Com.BR do HCPA, o armazenamento é feito através da gravação dos arquivos de vídeo digital em CDs.

Uma abordagem mais atual é a utilização de sistemas online para o armazenamento e gerenciamento destes exames. Uma solução assim foi desenvolvida pelo Grupo Cyclops e fornecida ao Com.BR, e até a data de escrita deste trabalho encontrava-se em fase final de implementação (Cyclops, 2007).

Na pesquisa realizada foi encontrado apenas um método computacional desenvolvido para auxiliar no processo de mensuração de perfurações timpânicas. O método foi criado por (Hsu, Chen, Hwang, & Liu, 2004). Este método, apesar de computacional, é um processo completamente manual, onde o usuário deve desenhar

manualmente, com uso do mouse, as áreas referentes à perfuração do tímpano e tímpano total.

O método proposto por (Hsu, Chen, Hwang, & Liu, 2004) foi desenvolvido em Matlab® (Mathworks, 2007), uma ferramenta matemática/científica disponibilizada comercialmente.

Capítulo 3

Imagens Digitais

Segundo (Gonzalez & Woods, 2000), uma imagem pode ser definida como uma função de intensidade luminosa, denotada $f(x,y)$, cujo valor ou amplitude nas coordenadas espaciais (x,y) fornece a intensidade ou o brilho da imagem naquele ponto.



Figura 6 - Lenna colorida (Lenna, 1972) - Exemplo de imagem

Ou ainda conforme (Schalkoff, 1989), uma função bidimensional gerada através da visualização de uma cena é geralmente conhecida como imagem.

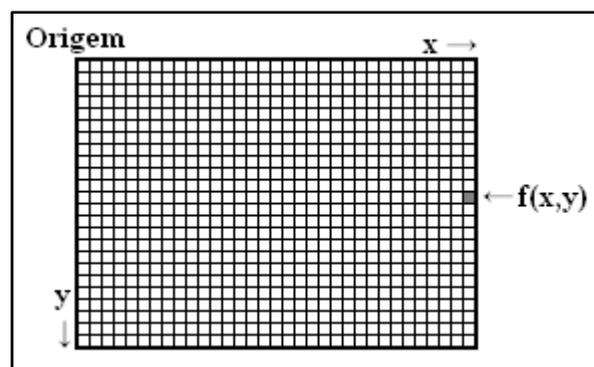


Figura 7 - Representação de imagens digitais

Na definição de (Mascarenhas & Velasco, 1989), matematicamente pode-se definir imagem como uma função $I(x,y)$, bidimensional, definida numa certa região. Para a maioria das imagens, a região de definição é um subconjunto limitado do plano, e os valores assumidos pela função são números reais limitados e não-negativos.

Tradicionalmente assume-se a forma retangular para imagens. As imagens são compostas por elementos de imagem que podem ser estruturas de diferentes modelos de cor ou podem ser simplesmente monocromáticos.

Uma imagem discreta pode ser armazenada em uma matriz, onde cada uma das posições armazena um elemento estruturante, definindo a cor ou intensidade de luminosidade naquele ponto.

Sinais podem ser capturados em uma dimensão (por exemplo: luminosidade), duas dimensões (por exemplo: imagens estáticas) ou três dimensões (por exemplo: tomografias computadorizadas). Ainda pode-se associar o fator tempo, a cada uma das formas de captura, compondo assim uma seqüência de sinais (por exemplo: Vídeos, seqüência de capturas de sinais bidimensionais). Imagens Digitais podem ser capturadas nas mais diferentes faixas do espectro do campo eletromagnético, possuindo as mais diferentes aplicações em cada faixa.

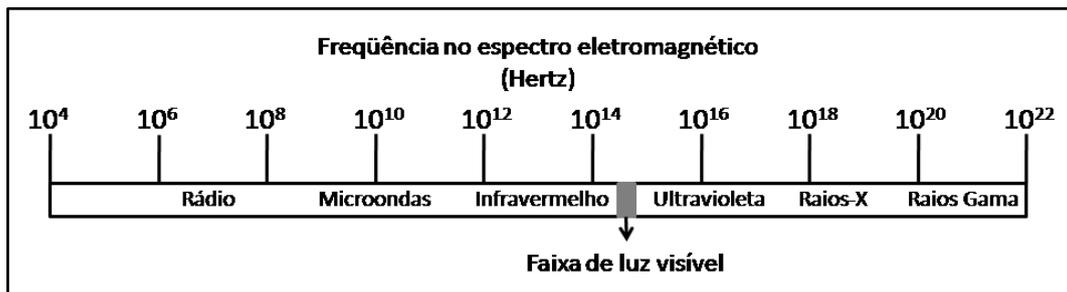


Figura 8 - Espectro eletromagnético. Adaptado de (Pedrini & Schwartz, 2008)

A aquisição de imagens digitais é comumente feita nas faixas de infravermelho, luz visível, ultravioleta e raios-x.

3.1 Elementos da Imagem

Os elementos de uma imagem são os valores de resposta da função imagem para cada coordenada. Estes elementos são estruturas que armazenam a intensidade de brilho, ou intensidade de cores, quando aplicável. Os elementos da imagem são comumente conhecidos como *pixel* (*Picture Element*).

O *pixel* é o menor elemento em um dispositivo de exibição como, por exemplo, em um monitor. Ele também é a menor unidade de imagem digital. Uma imagem digital

com tamanho de 640x480(pixels de largura x pixels de altura) possui 307200 pixels, outra imagem com tamanho de 1024x768 possui 786432 pixels. Nos casos de imagens tridimensionais, como em tomografias computadorizadas, este elemento é chamado *voxel*, e representa um pequeno cubo.

Para se armazenar o valor da intensidade de luminosidade, ou das cores, de cada pixel, é feita uma normalização da intensidade. Uma forma muito utilizada, por alguns modelos de cor e APIs, é a normalização entre zero e um. Sendo o valor “0” para ausência completa de luminosidade e o valor “1” para captura máxima de luminosidade. Outra forma, mais tradicional, é normalizar as intensidades em faixas de valores baseados em números potência de dois. A mais comum das normalizações utilizadas é entre os números 0 e 255, representando os 256 valores possíveis = 2^8 (oito bits por canal).

3.2 *Imagens Binárias*

As imagens binárias são matrizes onde os elementos da imagem podem assumir apenas dois valores, definidos conforme o contexto. Em geral imagens binárias são muito utilizadas na criação de mascaras, definição de regiões, descrição de bordas e/ou esqueletos de objetos, entre outros.



Figura 9 – Lenna (Figura 6) – Imagem em versão binária

Normalmente as imagens binárias são representadas graficamente por imagens em preto e branco, onde o a cor preta representa um dos dois valores enquanto o outro valor é representado pela cor branca.

3.3 Imagens Monocromáticas

Imagens Monocromáticas são imagens que apresentam apenas um canal de cor. Imagens em tons de cinza, televisões antigas e monitores antigos de fósforo verde são exemplos de imagens monocromáticas.



Figura 10 - Lenna (Figura 6) - Imagem em versão monocromática

As imagens monocromáticas são de grande utilização e importância. Uma grande área que utiliza muito este tipo de imagem é a medicina, onde são geradas imagens monocromáticas a partir de exames de raios-X, tomografia computadorizada, ressonâncias magnéticas, entre outros. No caso da medicina, as imagens monocromáticas servem para auxiliar aos médicos no processo de diagnóstico de diversas patologias.

3.4 Imagens Coloridas

O processo psicofisiológico de percepção de cores pelo cérebro humano não é ainda completamente compreendido pela ciência, mas sabe-se que os aspectos físicos da cor podem ser expressos por meio de uma base formal fundamentada por resultados teóricos e experimentais (Gonzalez & Woods, 2000).

As imagens coloridas têm por finalidade representar cenas com a maior parte possível do espectro visível de cores, aos olhos dos seres humanos.

Para isto existem diversos tipos de elementos estruturantes diferentes (Formatos de Pixel). Estas diferentes estruturas armazenam diferentes tipos de dados, em um padrão definido pelo modelo de cor utilizado.

Devido à existência de diferentes modelos de cor, a mesma cor pode ser representada por N formas diferentes, por modelos matemáticos abstratos, sendo N igual o número de modelos de cor em questão.

A Figura 6 exibe um exemplo de imagem colorida.

3.4.1 A Cor

A cor é uma percepção visual derivada do espectro da luz (distribuição da intensidade de luz versus o comprimento da onda de luz) (Gonzalez & Woods, 2000).

A luz é uma radiação eletromagnética, e toda radiação eletromagnética é caracterizada por um comprimento de onda (ou frequência). A faixa de comprimentos de onda visíveis aos olhos humanos (chamada de luz visível) é de aproximadamente 380nm até 780nm (788THZ até 384THZ).

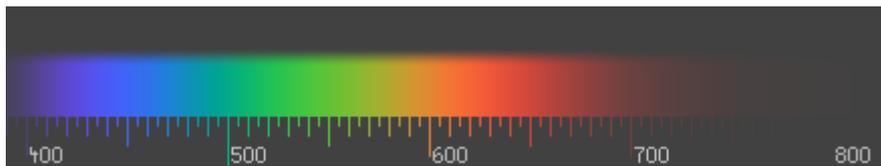


Figura 11 - Faixa de comprimentos de ondas eletromagnéticas, em nanômetros, em relação à cor interpretada pelo olho humano (Gonzalez & Woods, 2000)

3.5 Modelos de Cor

Um modelo de cor é um modelo matemático abstrato. Este modelo abstrato descreve um formato de representação para cores. Geralmente os modelos são compostos por tuplas de três ou quatro números.

Além da sua formação estrutural, os modelos de cor devem ter também uma função de mapeamento. Esta função de mapeamento é quem define a cor associada aos valores da tupla, ou seja, elas fazem a ligação entre os valores armazenados na tupla à cor específica no espectro de cores, bem como sua intensidade.

Cada modelo de cor existente é capaz de representar uma determinada *gamut*¹ de cores.

3.5.1 Modelo de cor RGB

RGB é uma abreviação na língua inglesa para três cores primárias de luz: *Red* (Vermelho), *Green* (Verde) and *Blue* (Azul). O RGB é o modelo de cor nativo dos computadores, e muito semelhante ao modelo de percepção humana, sendo que o modelo computacional foi inicialmente baseado no modelo humano (Pedrini & Schwartz, 2008).

O Modelo RGB é baseado no sistema de coordenadas cartesianas, em que o espaço de cores é um cubo. Cada uma das três cores primárias (vermelho, verde e azul) representa uma das dimensões neste cubo. As diferentes dimensões do cubo são sempre normalizadas em alguma faixa de valores, geralmente entre 0 e 1.

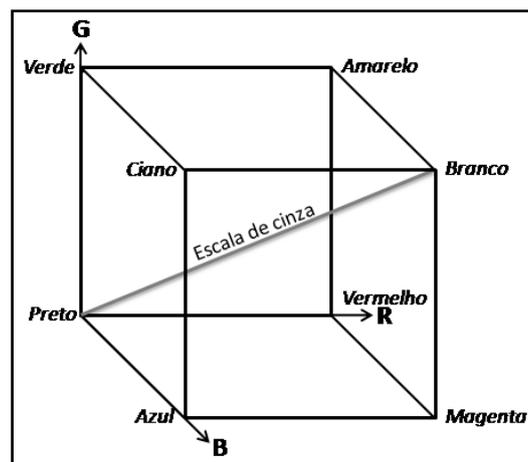


Figura 12 - Cubo RGB

Em uma das diagonais transversais do cubo (da cor preta até a cor branca), se encontra a escala de tons de cinza.

¹ “Um intervalo completo de possibilidades” (Longman, 2006), na computação, quando associado a cores, significa o conjunto de cores que podem ser representadas por uma estrutura de cor ou exibidas por um monitor de vídeo.

Capítulo 4

Processamento Digital de Imagens

O Processamento Digital de Imagens é um conceito que surgiu no início da década de 20. No principio o PDI foi usado para fazer pequenas restaurações de imagens digitalizadas para jornais, que eram enviadas por meio de cabo submarino de Londres para Nova York (Gonzalez & Woods, 2000).

Processar imagens é analisar e manipular imagens por computador. As finalidades podem ser diversas, mas as mais comuns são: extrair informações de imagens e transformar a imagem, para realçar elementos, por exemplo. (Mascarenhas & Velasco, 1989).

As informações obtidas através do processamento podem ser apresentadas de diversas formas, ou ainda, para as mais diversas aplicações. Geralmente o formato dos resultados se encontra em um formato pré-determinado, focando as necessidades posteriores.

4.1 Etapas do Processamento de Imagens

O Processamento Digital de Imagens envolve muitas etapas distintas e cada uma dessas etapas pode ainda envolver muitos conceitos diferentes. As etapas são executadas seqüencialmente, devido à dependência de informação.

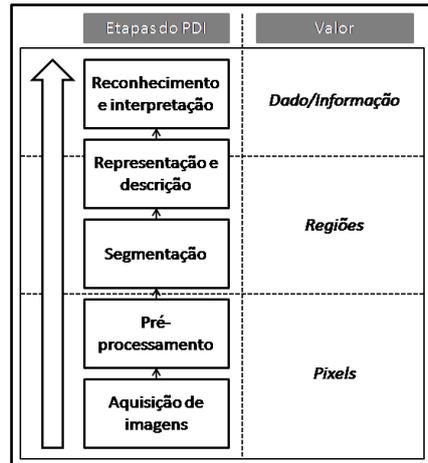


Figura 13 - Etapas do processamento de imagens. Adaptado de (Gonzalez & Woods, 2000)

Aquisição de imagens digitais, pré-processamento, segmentação, descrição e interpretação são etapas do processamento digital de imagens.

4.1.1 Aquisição de imagens digitais

No processo de aquisição de imagens digitais, também chamado de Digitalização, um dispositivo físico sensível a uma banda do espectro de energia eletromagnética (como raios-X, ultravioleta, luz visível ou infravermelha) recebe uma intensidade de sinal e produz um sinal elétrico de saída proporcional a um nível de energia percebida. Depois disto, um elemento digitalizador converte os sinais elétricos de sensoriamento para o formato digital (Gonzalez & Woods, 2000).

Como em toda forma de aquisição de imagens, para qualquer banda do espectro é necessário existir uma fonte geradora de energia na frequência da de interesse. Por exemplo, é fundamental a presença de luz solar, ou artificial (como lâmpadas), para a aquisição de imagens na banda de luz visível, assim como é necessário um gerador de ondas na banda de raios-X para gerar imagens de tomografias computadorizadas ou radiografias.

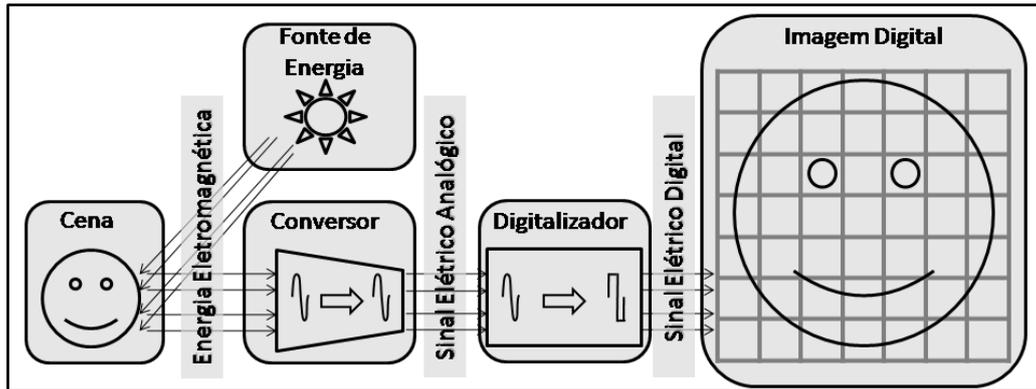


Figura 14 - Digitalização de imagens

Na Figura 14 é possível observar os elementos/estágios diferentes de um sistema digitalizador de imagens. Neste sistema encontra-se o dispositivo sensível a uma faixa do espectro eletromagnético e responsável pela geração de um sinal elétrico, representado pelo Conversor, e o dispositivo responsável pela conversão do sinal elétrico analógico para digital, o digitalizador.

4.1.2 Pré-processamento

O pré-processamento da imagem tem por função melhorar/preparar a imagem de forma a aumentar as chances de se obter sucesso nos processos seguintes. Nesta etapa geralmente são executados processos para atenuação de ruído, realce de contrastes, entre outros (Gonzalez & Woods, 2000).

Os ruídos ou distorções podem ser providos por diversas fontes. Cada uma das etapas do processo de aquisição está sujeita a interferências, imperfeições de equipamento ou similares, que acabam resultando em ruídos ou distorções.

Durante a etapa de pré-processamento são efetuadas correções nas imagens, de forma a suavizar a quantidade de distorções ou ruídos. As correções são feitas através de filtros específicos para determinados tipos de problemas. Um ou mais filtros podem ser aplicados na imagem, de forma a resolver mais do que um tipo de problema. O filtro passa-baixa é um dos filtros mais comumente usados no pré-processamento de imagens digitais.

4.1.2.1 Filtro Passa-Baixa

Uma imagem é formada pela aquisição de sinais elétricos, e por isso, está sujeita a ruídos adquiridos pelo processo de captura. O ruído encontrado em imagens digitais é de alta frequência, e é caracterizado pela alteração abrupta e difusa no contraste da imagem.

Em processamento de sinais, os filtros passa-baixa são utilizados para diminuição de sinais de alta frequência, e com isso, a diminuição do ruído de alta frequência. Em processamento de imagens os filtros passa-baixa são conhecidos também como filtro de mediana ou esmaecimento (*smooth*). Esse filtro é feito através da convolução de uma máscara na qual todos os valores são maiores que zero e geralmente cada ponto da máscara tem o mesmo valor em relação à distância do ponto central da máscara (*hot spot*). A Equação (1) mostra um exemplo simples de máscara para a filtragem de ruídos no formato passa-baixa.

$$I \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \div 9 \quad (1)$$

4.1.3 Segmentação

Na etapa de segmentação, a imagem é dividida em vários segmentos. O objetivo é que cada um destes segmentos corresponda a um objeto de interesse na cena apresentada pela imagem, ou partes de um objeto. Esta etapa é uma das mais difíceis de automatizar, principalmente porque dela dependem todas as próximas etapas, como reconhecer os objetos. A segmentação será abordada com maior ênfase à frente.

4.1.4 Representação e Descrição

Tipicamente a saída do processo anterior, do processo de segmentação, é constituída por dados em forma de pixels (“*raw pixel data*”), dados estes fazendo correspondência a regiões em imagens (Gonzalez & Woods, 2000) e em grande volume. O alto volume de dados em *raw pixel data* torna a comparação de padrões neste formato de altíssimo custo computacional.

Para diminuir a quantidade de informações sendo considerada na constituição dos padrões, e verificação de similaridades, são criados descritores para as imagens e seus segmentos. O processo de descrição de uma imagem também pode ser chamado de seleção de características. As características ou atributos dos segmentos ou da imagem podem ser informações quantitativas como tamanho de área ou informações estatísticas como coloração média, entre diversos outros tipos de atributos extraídos.

4.1.5 Reconhecimento e Interpretação

“Reconhecimento é o processo que atribui um rótulo a um objeto, baseado na informação fornecida pelo seu descritor. A interpretação envolve a atribuição de um significado a um conjunto de objetos reconhecidos.” (Gonzalez & Woods, 2000).

Em geral esta etapa compreende a aplicar técnicas de inteligência artificial ou estatísticas sobre os descritores encontrados na etapa anterior, a fim de identificar padrões esperados (objetos).

4.2 Segmentação de Imagens

A segmentação de imagens é um processo onde ocorre a divisão da imagem em regiões, ou seja, um particionamento da imagem. Esta divisão é muito importante, pois a partir dela, as regiões passaram a ser conhecidas por meio das suas propriedades geométricas e topológicas. Existem várias abordagens diferentes para fazer essa divisão, cada uma levando em consideração diferentes fatores, mas em essência todas buscam agrupar *pixels* que forem similares e se encontrem em uma mesma vizinhança. Essa busca é feita através da busca por discontinuidades ou similaridades no contraste da imagem.

4.2.1 Detecção de Bordas

O processo de detecção de bordas, também conhecido como detecção de discontinuidades, é um processo geralmente feito através de convoluções na imagem. Este processo objetiva detectar contornos de objetos em uma imagem. Para isso, geralmente são criadas máscaras que exerçam uma operação de identificação de

mudanças locais significativas no contraste da imagem. As mudanças podem ser descritas basicamente pelo conceito de derivada.

Muitas técnicas existem para a detecção de bordas em imagens. Algumas técnicas são clássicas, outras inovadoras, ou ainda outras bem difundidas, mas a verdade é que a idéia central em qualquer uma delas permanece a mesma. Todas as técnicas buscam por variações na cor, ou luminosidade, entre *pixels* em sua vizinhança.

Uma borda em uma imagem é a intensidade da variação de contraste em uma imagem, definindo contornos de objetos.

4.2.1.1 Operador de Sobel

O operador de Sobel é um popular operador diferencial discreto que consiste de duas pequenas máscaras. Assumindo a imagem I como entrada, a primeira máscara é utilizada para o cálculo de G_x (derivada da imagem I em x , (2), enquanto a segunda é utilizada para o cálculo de G_y (derivada da imagem I em y , (3) (Lindblad, 2005).



Figura 15 - Operador Sobel aplicado à Figura 6

Depois de encontrar as derivadas nas direções x e y da imagem, é calculado G (imagem estimada da magnitude do gradiente) utilizando a Equação (4). Também é possível estimar a direção do gradiente utilizando a Equação (5), obtendo assim a imagem Θ , conhecida como imagem de origem.

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * I \quad (2)$$

$$Gy = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I \quad (3)$$

$$G = \sqrt{(Gx)^2 + (Gy)^2} \quad (4)$$

$$\theta = \arctan\left(\frac{Gy}{Gx}\right) \quad (5)$$

4.2.1.2 Canny

O Detector de Bordas proposto por (Canny, 1986) é um tradicional e eficiente detector de bordas. O Detector foi concebido baseado em três principais necessidades identificadas pelo autor:

- Boa Detecção: encontrar todas as bordas existentes;
- Boa Localização: marcar as bordas no lugar mais próximo o possível da borda real;
- Resposta Mínima: cada borda deve ser marcada uma única vez.



Figura 16 – Detector de bordas Canny aplicado à Figura 6

O Algoritmo é composto por cinco passos: Filtragem, Cálculo de Gx e Gy , Cálculo da intensidade do Gradiente e Origem, Supressão não máxima e, por fim, Histerese.

A filtragem tem por objetivo retirar os ruídos de alta frequência da imagem. Geralmente é feita através do uso de filtros passa-baixa. O filtro mais comumente utilizado neste algoritmo é a convolução de uma máscara criada a partir da fórmula de distribuição de Gauss ((6). O tamanho da máscara e o valor para σ utilizado variam conforme parâmetros informados.

$$M(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-x^2}{2\sigma^2}} \quad (6)$$

O Cálculo de G_x e G_y é feito também através da convolução de máscaras, entretanto as máscaras são geradas a partir da primeira derivada da equação de distribuição de Gauss ((7).

$$M'(x) = \frac{-x}{\sqrt{2\pi\sigma^3}} e^{\frac{-x^2}{2\sigma^2}} \quad (7)$$

A Intensidade do Gradiente (G) a sua direção/origem (Θ) são calculados utilizando os valores encontrados para G_x e G_y . Estes servem para indicar uma correlação entre das bordas dos dois sentidos, em intensidade e sentido de direção.

$$G = \sqrt{G_x^2 + G_y^2} \quad (8)$$

$$\Theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (9)$$

Depois de encontrar a intensidade do gradiente, é possível visualizar os locais de resposta de bordas, mas cada borda ainda é representada mais de uma vez, para isto é então aplicado o processo de supressão não máxima.

No processo de supressão não máxima é feita uma busca, baseado na direção e intensidade do gradiente, por pontos de resposta local máxima. Este processo também é comumente chamado de afinamento de bordas, e resultará em uma única resposta para cada borda encontrada.

Após encontrar os pontos locais máximos, é executado o processo de traçar as bordas através da limiarização com histerese (um limiar adaptável), para obter uma imagem binária indicando as bordas da imagem (Figura 16).

4.2.2 Segmentação Orientada a Regiões

Existem duas formas básicas de segmentação orientada a regiões: Crescimento de regiões por agregação de pixels; e Divisão e união de regiões.

No crescimento de regiões por agregação de pixels, ocorre um agrupamento de pixels ou sub-regiões em regiões maiores (Gonzalez & Woods, 2000). A mais simples dessas abordagens começa com vários pontos “semente”, e a partir deles, cresce as regiões anexando a ela cada pixel vizinho que se considere similar ao inicial ou anterior, até que todos os pixels da imagem pertençam a alguma região.

Na divisão e união de regiões a imagem é dividida em um conjunto de regiões arbitrárias e disjuntas. Depois de dividir a imagem em diversas regiões, as consideradas similares são unidas, até não ser mais possível encontrar regiões similares (Pedrini & Schwartz, 2008).

4.2.2.1 Mumford & Shah

O Algoritmo de Mumford & Shah (Mumford & Shah, 1989) é um dos algoritmos mais utilizados para a segmentação de imagens através do crescimento de regiões. Este algoritmo é baseado em uma funcional de energia, descrita na Equação (10).

$$E(u, K) = \iint_{\Omega} \|u - g\|^2 dx dy + \iint_{\Omega/K} \|\nabla u\|^2 dx dy + \lambda * l(K) \quad (10)$$

(Mumford & Shah, 1989)

No algoritmo de Mumford & Shah as regiões (grupos de *pixels*) se comportam como se fossem pedaços de borracha com a elasticidade definida na funcional de energia. Esses pedaços de borracha são soltos sobre os pontos iniciais (*seeds*) e começam a ser esticados sobre as regiões. Quando as borrachas ainda possuem elasticidade para abranger mais *pixels*, mas estes já são abrangidos por outra borracha, pode acontecer a troca de borracha (região ou segmento) a qual os *pixels* estão associados, ou até mesmo a fusão das duas borrachas em uma só.

Os termos da funcional de energia de Mumford & Shah são descritos a seguir:

$E(u, K)$	Energia funcional de Mumford & Shah
Ω	Domínio da Imagem
Ω_i	Região Sensata (<i>Piece-wise</i>)
$u(x, y)$	Função diferenciável dentro de cada região
$g(x, y)$	Intensidade de Luz
K	Fronteira entre regiões
λ	Parâmetro de escala
$l(K)$	Comprimento dos arcos das bordas

Na Figura 17 é possível visualizar um exemplo de resultado de segmentação obtido através do algoritmo de Mumford & Shah quando aplicado a imagem de Lenna e com $\lambda = 1400$. É possível observar boas definições de áreas, apesar de também serem encontrados alguns vazamentos de região.



Figura 17 – Resultado da aplicação do algoritmo de Mumford & Shah à Figura 6 com $\lambda = 1400$

4.2.2.2 Contornos Ativos

Um dos principais problemas encontrados quando se trabalha com detecção de bordas é a não continuidade das mesmas. Geralmente encontra-se dificuldade em determinar regiões fechadas em si, somente com o uso de algoritmos de detecção de bordas. Uma proposta de solução para este problema foi feita em (Kass, Witkin, & Terzopoulos, 1988) e chamada de *Snakes: Active contour models* (Modelos de Contornos Ativos).

Na abordagem de Contornos Ativos a região inicial (*seed*) é uma curva fechada em si que evolui sobre uma forma na imagem a fim de segmentá-la. Durante a evolução da curva diversas propriedades físicas são atribuídas à mesma, tais como elasticidade e

imagens de etapas anteriores. Os processos utilizados variam entre: Convolução, Magnitude e Operação.



Figura 19 - Aplicação de Contornos Ativos - (a) Imagem processada com contorno inicial na cor verde; em (b), (c) e (d) evolução da curva após 600, 1100 e 1600 iterações, respectivamente.

As *Convoluções* foram feitas com diferentes tipos de máscaras, cada máscara adequada à resposta esperada pelo algoritmo naquele momento. A maioria das convoluções é feita com operadores de derivação, descritos no algoritmo de Sobel nas equações: (2) e (3). A Operação de *Magnitude* do Gradiente utiliza duas imagens de derivadas, geradas por convolução (DX e DY), e é feita em todos os pontos da imagem. O cálculo da Magnitude também segue o modelo do cálculo de Magnitude do algoritmo de Sobel ((4)). Os processos de *Operação* fazem operações lógicas e/ou matemáticas, transformando mais de uma imagem de entrada em uma única imagem de saída.

Na Figura 19 é possível observar as diferentes etapas da segmentação de uma imagem através do uso de Contornos Ativos. Foi utilizada a Figura 6 – Lenna Colorida – para o processamento, e definida uma área inicial retangular ao redor da imagem da mulher na foto, a fim de conseguir a separação entre a mesma e o fundo da imagem. As

áreas delimitadas pela cor verde representam o segmento encontrado na imagem em uma determinada iteração pelo algoritmo, sendo que a subfigura (a) exibe a área inicial (iteração zero ou *startup seed*). As subfiguras (b) e (c) exibem respectivamente a curva durante a evolução nas iterações 600 e 1100. O resultado final do processamento, obtido após 1600 iterações é exibido na sub-figura (d). É possível observar, que em alguns pontos do resultado final, a curva incluiu algumas formas do fundo da imagem ao objeto de interesse, mas que em geral a imagem da mulher ao centro foi separada do fundo da imagem.

Capítulo 5

Metodologia

Para a criação da metodologia foram escolhidos e avaliados algoritmos para o pré-processamento e segmentação das imagens otológicas. No pré-processamento foi utilizado um filtro de passa-baixa como atenuador de ruídos. Para a etapa de segmentação foram testados e comparados os algoritmos de Mumford & Shah e Contornos Ativos.

5.1 Pré-processamento

Na etapa de pré-processamento foi utilizado um filtro passa-baixa. Este filtro tem por objetivo eliminar ruídos e atenuar bordas para uma detecção mais suave dos atributos da imagem.

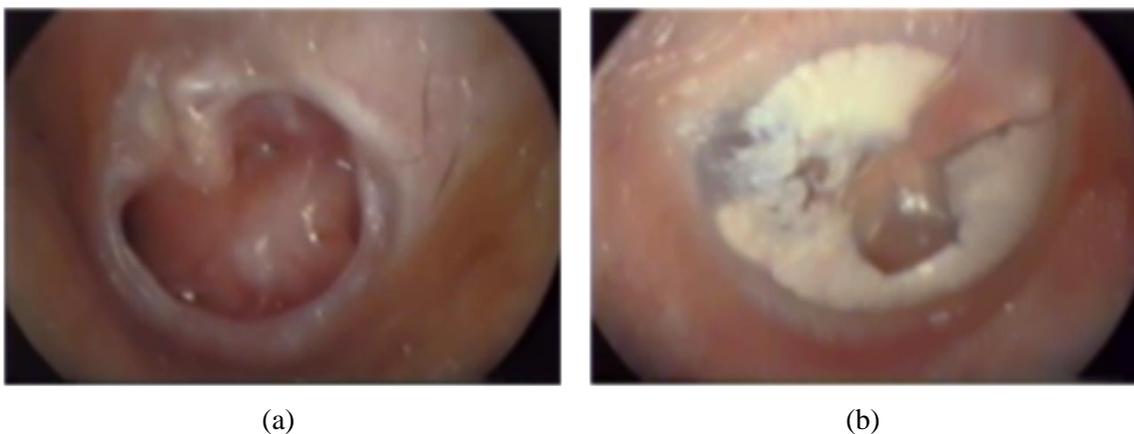


Figura 20 - Resultado da aplicação de 10 iterações do filtro passa-baixa. Imagens originais de (a) e (b): Figura 4 e Figura 5, respectivamente

É possível observar na Figura 20 os resultados obtidos com a aplicação de 10 iterações do filtro passa-baixa em imagens de videotoscopia digital. O número de 10 iterações foi escolhido após uma seqüência de testes laboratoriais, onde a aplicação deste número de iterações obteve resultados mais próximos ao esperado. Ocorre a suavização na alteração do contraste das imagens após a aplicação do filtro. O resultado visual é semelhante a um borrão, como é muitas vezes chamado o próprio algoritmo.

O resultado da filtragem passa-baixa pode ser observado em detalhes nas figuras: Figura 21 e Figura 22. Na primeira é exibido o histograma de uma imagem videotoscópica original e na segunda é mostrado o histograma da mesma imagem após o processo de filtragem. Ambas as imagens fazem a correlação da quantidade de pixels (Eixo Qtd) para cada tom de luminosidade (Eixo Lum) da imagem. Lembrando que o tom de luminosidade é um valor inteiro normalizado entre 0 e 255.

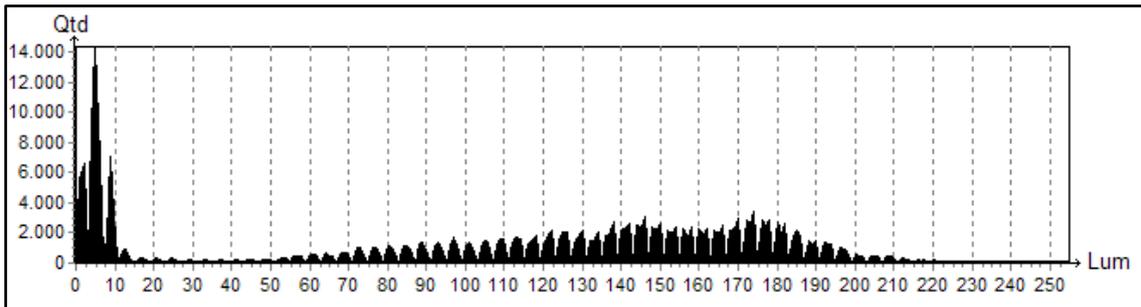


Figura 21 - Histograma da Figura 4 (Imagem original)

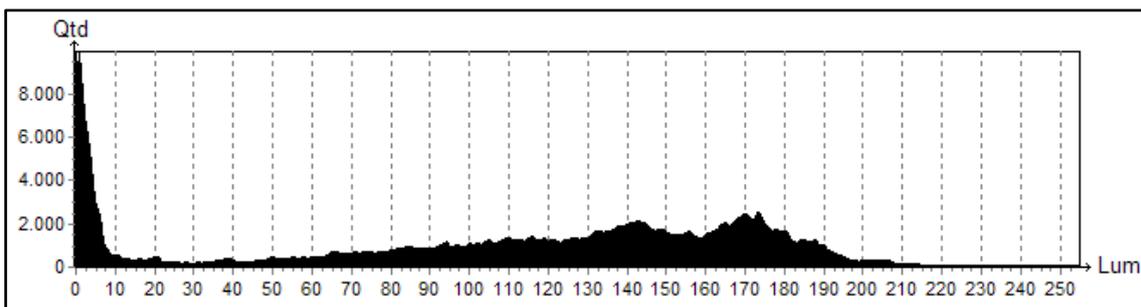


Figura 22 - Histograma da Figura 20 (a) (Após aplicação do filtro)

Através do histograma da Figura 21 percebe-se um *gamut* reduzido na câmera de aquisição, que é definido pela falta de resposta da câmera – a certos níveis de luminosidade ou cores – que é representada por lacunas no gráfico do histograma. Este *gamut* reduzido é responsável por imagens de baixa qualidade, o que dificulta o processo de segmentação.

5.2 Segmentação Bimodal Semi-Automatizada

Para a etapa de segmentação foram testados dois algoritmos de *estado da arte* apresentados anteriormente: *Mumford & Shah* e *Contornos Ativos*. O primeiro foi escolhido por obter bons resultados na segmentação em diversas formas de imagens e o segundo foi escolhido por obter resultados bimodais, parte do objetivo deste trabalho.

Nos testes com ambos os algoritmos, as imagens foram segmentadas com diferentes parâmetros e diferentes *seeds* iniciais. Os melhores parâmetros e formas de inicialização (*seeds*) foram escolhidos de forma subjetiva, conforme avaliação visual em relação aos resultados esperados.



Figura 23 – Imagem resultante da segmentação com o algoritmo de Mumford & Shah

O algoritmo de *Mumford & Shah* produz resultados em vários segmentos, não bimodal, ao contrário do algoritmo de *Contornos Ativos*. Por isto foi desenvolvido um método computacional que possibilita a interação do especialista médico com o resultado do algoritmo *Mumford & Shah*, de forma a facilitar o processo de segmentação bimodal das imagens de patologias timpânicas. O comportamento do modelo desenvolvido é descrito na Equação (11).

$$D = \bigcup_{i=1}^n R_i | (R_i \cap S \neq \emptyset) \quad (11)$$

No modelo desenvolvido são utilizados alguns elementos em forma de conjuntos, onde D é o conjunto de *pixels* que indica a região afetada pela patologia, R indica um conjunto de subconjuntos de *pixels* (resultantes do processo de segmentação *Mumford & Shah* - Figura 23), e S indica um conjunto de *pixels* que foram selecionados pelo especialista como *seed* inicial para a segmentação.

O conjunto S é um conjunto de *pixels* definido por uma linha, desenhada a mão livre pelo especialista sobre a imagem – como visto na Figura 24 (a). Após a linha ser

desenhada pelo especialista, as duas pontas da linha são ligadas, formando uma área fechada – exibida na Figura 24 (b).

O modelo faz a verificação de intersecção entre os subconjuntos de R com o conjunto S . Os subconjuntos de R que tiverem intersecção com o conjunto S (*Seed*) serão unidos em um novo conjunto chamado D , que representa o segmento da patologia – representado na Figura 25 (a).

Um ponto positivo neste modelo é a possibilidade de se criar tantas áreas iniciais quanto forem necessárias. Isto significa que mesmo em casos onde a primeira área inicial não intersectou segmentos suficientes para a total seleção da patologia, ainda é possível criar novas áreas iniciais para selecionar as áreas não tangidas previamente pelo processo, e incrementar o novo resultado ao resultado anterior. Quando isto for necessário, o usuário terá que fazer mais do que apenas uma interação com o processo.

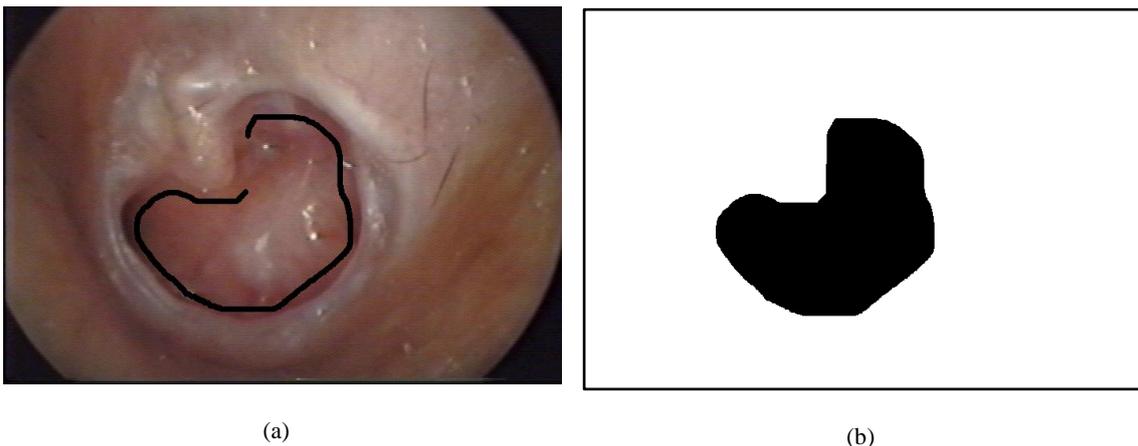


Figura 24 - Exemplo de área inicial sendo desenhada por especialista em (a) e área inicial resultante em (b)

A mesma abordagem utilizada para a criação do conjunto de *pixels* iniciais da metodologia de segmentação com *Mumford & Shah* foi utilizada para inicializar o processo de segmentação através de *Contornos Ativos* (Figura 24).

A principal diferença entre a inicialização de ambos os modelos, é que quando utilizamos *Contornos Ativos*, não podemos fazer a segunda interação com o usuário, visto que o resultado do processo de segmentação é bimodal. Por tanto, neste processo não ocorre a seleção de segmentos, mas sim a evolução de uma curva delimitadora de uma única região a partir da área inicial.

A Figura 25 (b) mostra o resultado da segmentação bimodal obtida através do uso de *Contornos Ativos* com a mesma inicialização utilizada na geração do resultado através do uso de *Mumford & Shah*, exibido na Figura 25 (a). Para a segmentação através de *Contornos Ativos* foi utilizado o parâmetro de 1000 iterações (mesmo com um número maior de iterações o resultado não apresentou alterações).

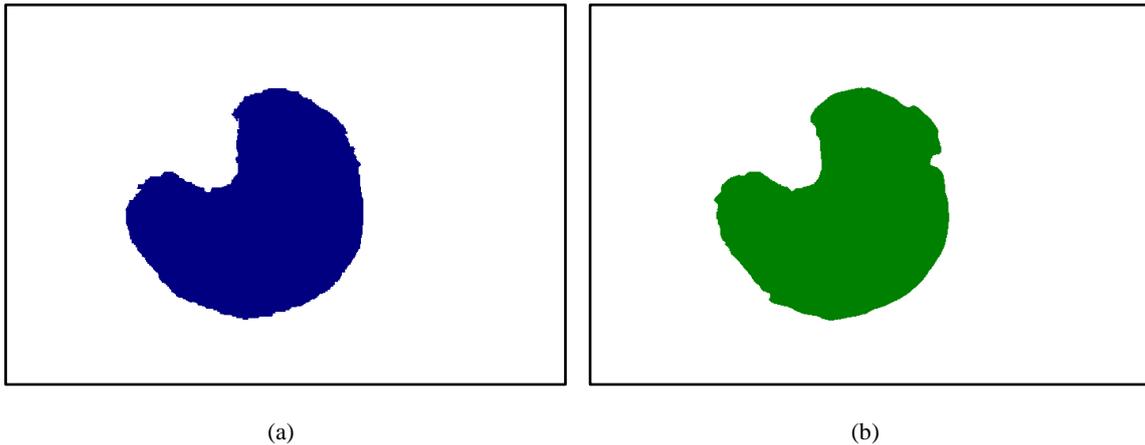


Figura 25 - Resultados dos processos de segmentação. Utilizando em (a) *Mumford & Shah* e em (b) *Contornos Ativos*

Ao comparar os resultados dos dois modelos (Figura 25) é possível observar que o modelo baseado em *Contornos Ativos* possui uma limitação de crescimento quando encontra variações elevadas no gradiente da imagem.

Outra limitação imposta pelo uso do algoritmo baseado em *Contornos Ativos* é a impossibilidade de desenhar novas áreas de seleção depois de segmentada a imagem. Isto ocorre por que o algoritmo funciona de forma iterativa e, por isso, toda e qualquer outra área previamente selecionada continuará evoluindo junto com a nova região, mesmo quando já estiver corretamente evoluída. Devido a estas limitações, escolheu-se utilizar apenas o modelo baseado na segmentação através do algoritmo de *Mumford & Shah*.

5.3 Definição da Área Timpânica

A quantificação das patologias timpânicas é uma das principais aplicações da segmentação das patologias timpânicas, mas para quantificar áreas em imagens digitais (mensurá-las), é necessária alguma associação (relação) por *pixel*. Esta associação pode ser métrica ou comparativa.

Para fazer uma associação métrica, o dispositivo de captura das imagens deve fornecer junto com a imagem uma informação métrica em relação a cada *pixel*, como ocorre, por exemplo, em imagens de tomografias digitais. Já na informação quantitativa associativa é possível fazer comparações de tamanho entre diferentes estruturas e estimar o tamanho da área segmentada, ou simplesmente pode-se obter um percentual de correspondência entre áreas.

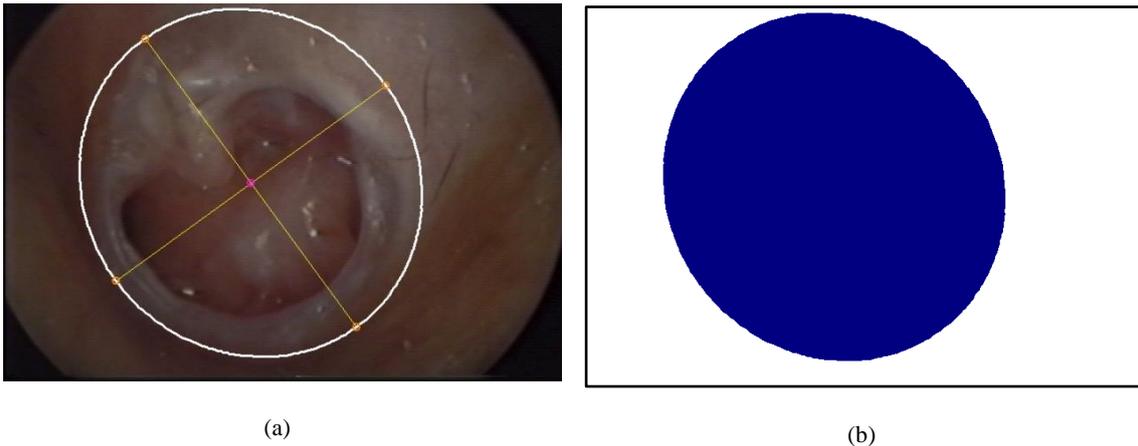


Figura 26 - Seleção elíptica do tímpano. Em (a) processo de seleção e em (b) imagem binária resultante

Infelizmente na aquisição de imagens através do exame de videotoscopia digital não é possível obter a associação métrica por pixel da imagem, por isto é necessário fazer uma associação comparativa entre a patologia segmentada e alguma outra estrutura do tímpano. A região a ser associada escolhida neste trabalho é a área total da membrana timpânica, devido à possibilidade direta de verificação do seu percentual afetado pela patologia.

Para obter valores percentuais de correspondência entre as áreas afetadas pelas patologias, foi criada a seleção de região timpânica elíptica, exibida na Figura 26 (a). Nesta técnica assume-se que o tímpano possui um formato elíptico. A partir disto, defini-se uma elipse sobre a imagem, esta elipse será transladada, rotada e redimensionada até que suas bordas estejam sobre os limites da membrana timpânica.

No fim do processo de sobreposição da elipse sobre a membrana timpânica, é obtida uma imagem binária representando a área timpânica, exibida na Figura 26 (b). Esta área serve como referência para se fazer a correlação entre área timpânica e área afetada pela patologia em percentual afetado.

5.4 Validação

Para aferir a acurácia dos métodos de segmentação, de ambas as estruturas (patologia e anatomia timpânica), foi empregada uma metodologia baseada em *ground truths* (GTs) alinhada com (Yoo, Ackerman, & Vannier, 2000).

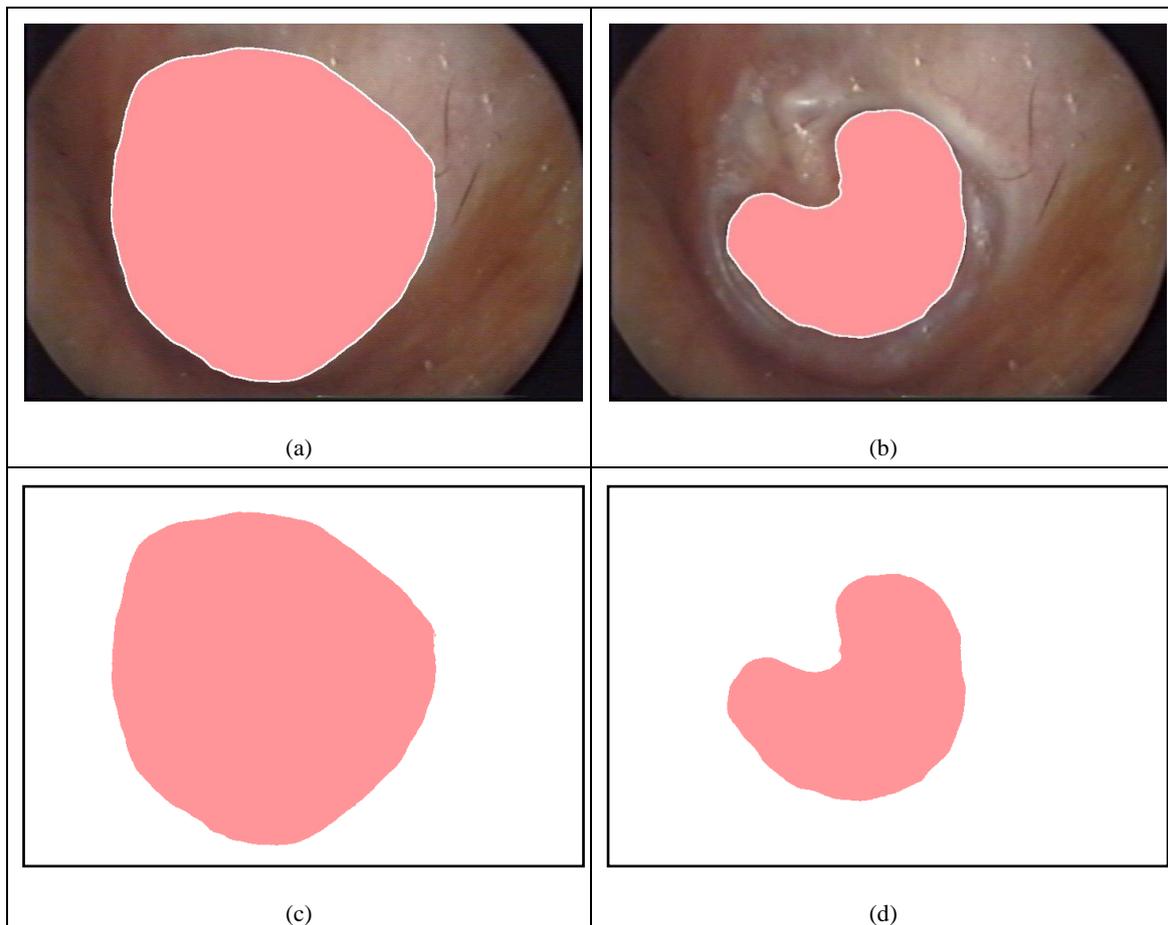


Figura 27 - Exemplos de *Ground Truths*

Neste caso, os *ground truths* (GTs) são imagens segmentadas manualmente por um especialista, o qual usa um dispositivo de entrada especial conhecido como *digitizing tablet* (mesa digitalizadora) para desenhar uma imagem que representa a segmentação bimodal esperada (da patologia ou do tímpano). Como exemplo, imagens de GT podem ser vistas na Figura 27. Em (a) desenho do especialista para selecionar o tímpano, em (b) desenho do especialista para selecionar a patologia, em (c) área binária desenhada em (a) e em (d) imagem binária desenhada em (b). O objetivo da geração dos GTs é prover um padrão ouro morfológico para a acurácia da segmentação. Por este motivo foi:

- a) Escolhido um conjunto de imagens de patologias timpânicas de 39 pacientes;
- b) Seguido a metodologia de geração de GTs proposta por (Martin, Fowlkes, Tal, & Malik, 2001).

Os 39 pacientes foram selecionados aleatoriamente a partir da base de dados, contendo mais de 1.000 pacientes com filmagem bilateral, do Hospital de Clínicas de Porto Alegre. Foram analisados 39 casos devido ao tempo disponibilizado pelos especialistas, visto que o processo de geração de GTs possui um alto consumo de tempo. As imagens foram submetidas a dois processos distintos:

- Processadas com o método computacional desenvolvido neste trabalho; e
- Geração manual dos GTs por usuários utilizando *digitizing tablets*.

Os resultados do método computacional desenvolvido foram comparados com os GTs por técnicas conhecidas por Rand (Rand, 1971) e BGM (Jiang, Marti, Irniger, & Bunke, 2006). Ambas as técnicas, Rand e BGM, geram índices no intervalo de 0 a 1, de forma que 0 significa uma correspondência perfeita entre as segmentações e 1 significa que não existe similaridade entre elas. O fato de ambas os índices estarem no mesmo intervalo facilita a comparação entre algoritmos com diferentes abordagens e características. Detalhes sobre estas duas medidas de avaliação podem ser observadas no Anexo A.

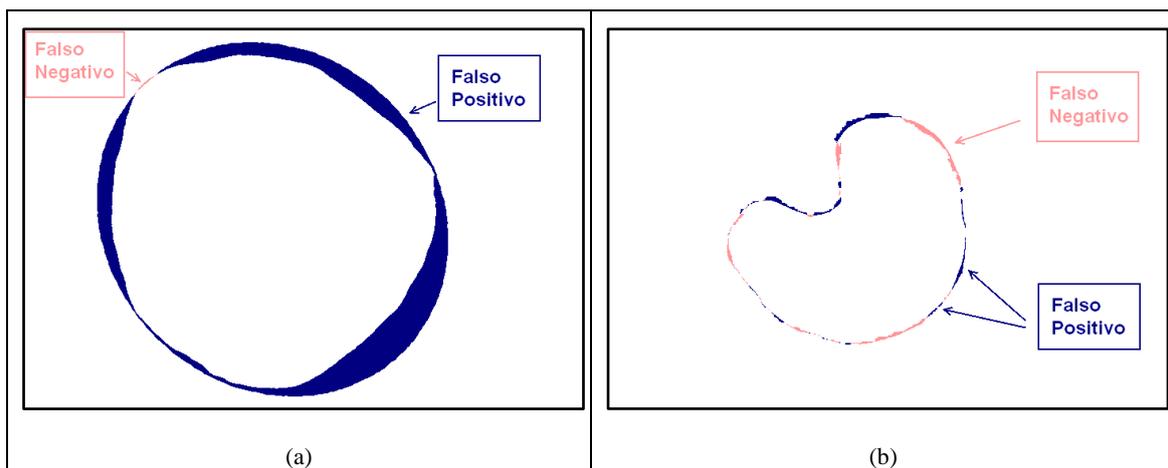


Figura 28 - Exemplos de Falso Negativo e Falso Positivo. Em (a) Segmentação timpânica e em (b) segmentação de patologia

Outra comparação feita foi a verificação do percentual da área segmentada que não corresponde com a área definida no GT. Áreas que não são selecionadas na

segmentação, mas são selecionadas no GT são chamadas de Falso Negativo (percentual de área não segmentada em relação a área total do GT). Áreas que são selecionadas na segmentação, mas não são selecionadas no GT são chamadas de Falso Positivo (percentual de área segmentada em relação a área total do GT). Exemplos de Falso Negativo e Falso Positivo, tanto na definição da área timpânica quanto na área patológica, são exibidas nas Figura 28 (a) e (b).

Capítulo 6

Resultados

O método computacional desenvolvido foi implementado em um protótipo de *software* para propósitos de testes. Este *software* foi chamado de *Auris* (Heck Junior, Wangenheim, Abdala, Comunello, Dornelles, & Costa, 2007). Os primeiros resultados quantitativos obtidos com o uso deste *software* foram feitos pelos especialistas do Hospital de Clínicas de Porto Alegre (HCPA) em um trabalho de pesquisa. Os dados da análise clínica foram publicados em periódicos da medicina tal como em: (Costa, Dornelles, Schmidt, Wangenheim, Comunello, & Heck Junior, 2008) e em (Dornelles, Costa, Wangenheim, Rosito, Comunello, & Heck Junior, 2007).

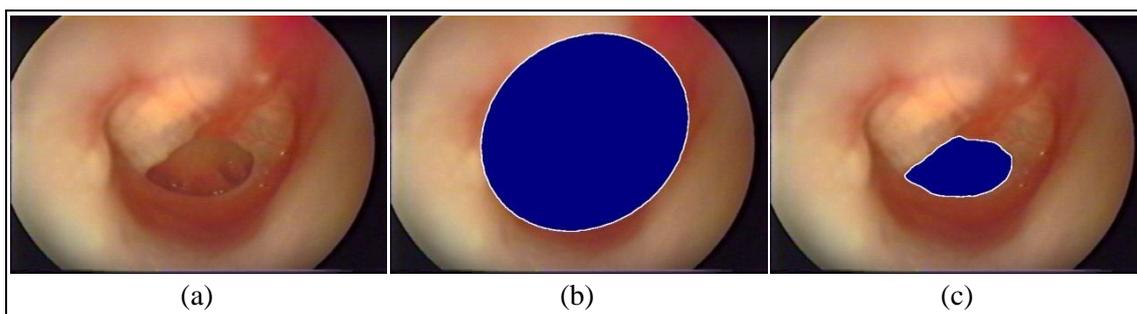


Figura 29 - Resultado da mensuração do paciente 02 com 14,19% de perfuração. Em (a) imagem original, em (b) segmentação do tímpano e em (c) segmentação da patologia.

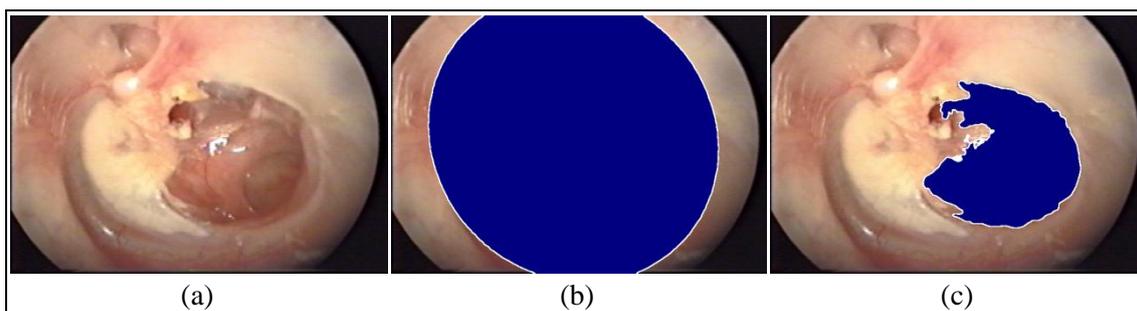


Figura 30 - Resultado da mensuração do paciente 03 com 23,90% de perfuração. Em (a) imagem original, em (b) segmentação do tímpano e em (c) segmentação da patologia.

Alguns outros exemplos, além do referente ao paciente 01 - já exibido na etapa de desenvolvimento na Figura 25 (a) e Figura 26 (b) com uma patologia que abrange 36,71% da membrana timpânica -, são exibidos na Figura 29 e Figura 30. O conjunto completo das 39 imagens pode ser visualizado no endereço web:

<http://www.lapix.ufsc.br/auris/>. Informações sobre as mensurações do conjunto completo de casos selecionados podem ser vistas também na Tabela 4, no Anexo B.

6.1 Resultados da Validação

Os resultados do processamento dos 39 casos selecionados para a validação foram comparados com os seus respectivos GTs. Desta comparação foi possível extrair diversas informações sobre: O percentual do tímpano afetado pela patologia - tanto pelo o método computacional desenvolvido, quanto pelos GTs -, os índices Rand e BGM, bem como os valores percentuais dos resultados de Falso Positivo e de Falso Negativo. Também foi calculada a média e o desvio padrão para cada valor percentual ou índice.

Informação	GT	Auris	Diferença
<i>Média</i>	23,65%	21,94%	3,10%
<i>Desvio Padrão</i>	14,23%	12,97%	3,73%

Tabela 1 - Média e Desvio Padrão para o percentual da área do tímpano afetada pelas patologias do conjunto de 39 casos avaliados

Na análise quantitativa da patologia, foi calculada média da diferença entre os resultados do método e os valores obtidos com os GTs. A média de percentual do tímpano afetado pela patologia no conjunto de pacientes selecionado, bem como o respectivo desvio padrão, é exibida na Tabela 1. Na Figura 31 é possível observar o diagrama de dispersão, a linha de tendência e o coeficiente de correlação da mensuração dos resultados obtidos em relação a mensuração dos resultados esperados.

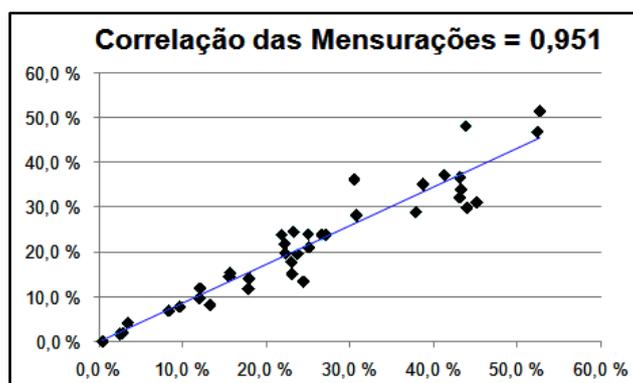


Figura 31 - Correlação de Mensurações: resultados esperados em relação aos resultados obtidos

O resultado da mensuração para cada um dos 39 casos, tanto utilizando imagens segmentadas pelo método computacional desenvolvido quanto pelos GTs, bem como os índices de avaliação Rand e BGM e os percentuais de Falso Negativo e Falso Positivo são exibidos na Tabela 4, no Anexo B.

Informação	Índice Rand	Índice BGM	Falso Negativo	Falso Positivo
<i>Média</i>	0,119	0,063	2,388%	14,360%
<i>Desvio Padrão</i>	0,069	0,063	5,450%	10,359%

Tabela 2 - Avaliação da segmentação da área timpânica: Média e Desvio Padrão do índice Rand, índice BGM, percentual de Falso Negativo e percentual de Falso Positivo aplicados aos 39 casos

Resumos da Tabela 4 são exibidos nas tabelas Tabela 2 e Tabela 3. Na primeira são exibidas as informações médias e de desvio padrão para cada um dos índices e percentuais aferidos na segmentação de áreas timpânicas, enquanto a segunda representa a mesma aferição para os resultados da segmentação de patologias timpânicas.

As Figuras 32 e 33 mostram os diagramas de dispersão, linhas de tendência e os coeficientes de correlação nas comparações entre os resultados obtidos pelos processos de definição da região timpânica e da segmentação de patologias em relação aos seus respectivos resultados esperados.

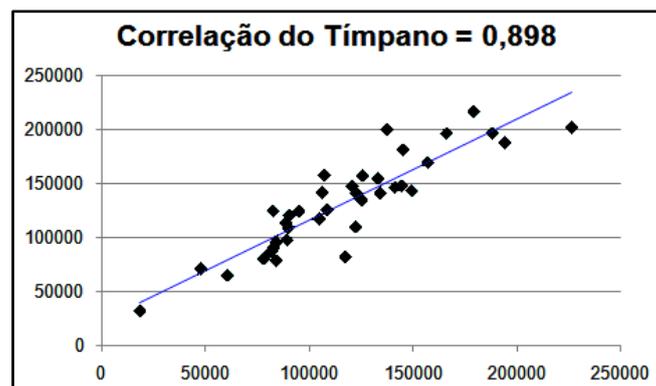


Figura 32 - Correlação de Timpanos: resultados esperados em relação aos resultados obtidos (em pixels)

É possível observar através da comparação entre as tabelas, Tabela 2 e Tabela 3, que na avaliação dos resultados de segmentação de patologias, as mesmas apresentaram índices médios menores do que os resultados da segmentação de tímpanos. Os índices obtidos na avaliação da segmentação da área patológica foram próximos a zero,

representando alta similaridade, à medida que os percentuais de valores falsos (positivo e negativo) mantiveram-se abaixo de 4,5%.

Informação	Índice Rand	Índice BGM	Falso Negativo	Falso Positivo
<i>Média</i>	0,0102	0,0052	4,433%	2,264%
<i>Desvio Padrão</i>	0,0158	0,0082	5,178%	4,643%

Tabela 3 - Avaliação da segmentação da área patológica: Média e Desvio Padrão do índice Rand, índice BGM, percentual de Falso Negativo e percentual de Falso Positivo aplicados aos 39 casos

O processo de segmentação de tímpanos obteve índices maiores, representando menor similaridade, devido ao uso da estrutura geométrica elíptica para estimar a forma anatômica do tímpano. O valor de Falso Positivo também foi maior na segmentação de membranas timpânicas, isto ocorreu também devido ao uso da estrutura elíptica, pois ao se assegurar que toda a membrana estava coberta pela elipse, algumas áreas não pertencentes ao tímpano foram incluídas na área segmentada.

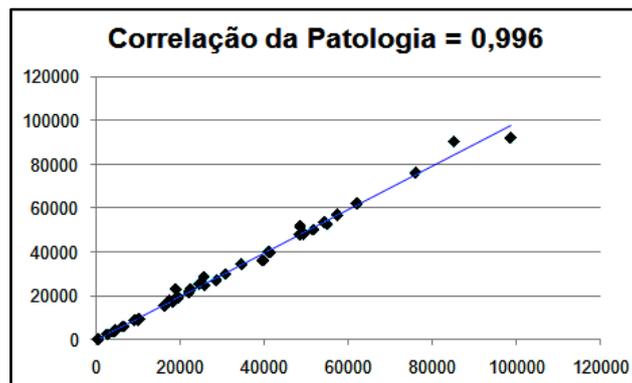


Figura 33 - Correlação de Patologias: resultados esperados em relação aos resultados obtidos (em pixels)

Um dos maiores casos de falso positivo encontrado, na definição da área timpânica, foi no caso 04 (ver Tabela 4 – Anexo B). Este caso ilustra também uma das maiores diferenças de percentual aferido entre os resultados obtido e esperado. Isto também ocorre nos casos 05, 20, 21, 31 e 32.

O caso número 28 também chama atenção por apresentar um valor muito acima da média em área de falso negativo no resultado de segmentação da patologia, porém, seus índices Rand e BGM são também os menores encontrados. Isto se deve ao tamanho da patologia, visto que a área da patologia é muito pequena.

Capítulo 7

Discussões e Conclusões

A metodologia desenvolvida, e exposta neste trabalho, obteve resultados satisfatórios, segundo avaliação feita durante o processo de validação. Os resultados da validação do método de segmentação bimodal da patologia apresentam valores que indicam alto índice de similaridade entre o resultado obtido e o resultado esperado.

Os resultados da validação do método de definição da área timpânica apontaram um bom grau de similaridade entre o resultado obtido e o resultado esperado. Este método obteve índices menores do que o método de segmentação da patologia. Um dos principais motivos constatados é a tendência à obtenção de valores falso-positivos pelo método, que se deve à natureza elíptica do método criado para selecionar a estrutura anatômica.

Os valores de falso-positivo e de falso-negativo são variações nos resultados inerentes a métodos, mas nem todas as aplicações são totalmente tolerantes a variações. Na maioria das vezes uma faixa de variação é aceita, por se tornar insignificante, mas nem sempre. Por isto, cabe uma análise *a priori* dos índices de variação encontrados na metodologia para julgar a aplicabilidade desta em cada contexto.

7.1 Objetivos Alcançados

Na pesquisa realizada não foram encontrados trabalhos com a mesma abrangência deste trabalho, apesar de ter sido encontrado um trabalho correlato que é identificado e comentado na seção 2.7.

O método para a geração de um *seed* inicial, utilizado pelo método de segmentação bimodal da patologia, bem como o próprio método de segmentação bimodal foram desenvolvidos. O desenvolvimento destes métodos é apresentado na seção 5.2.

Para poder completar o processo de mensuração, através de uma área de referência, foi criado o método de definição da região timpânica. Este modelo permite uma seleção elíptica da estrutura timpânica e é descrito na seção 5.3.

Por fim, foram encontrados e utilizados quatro métodos como forma de estimar o grau de precisão da metodologia. Estes métodos foram aplicados aos resultados obtidos com a metodologia desenvolvida em relação a imagens criadas de *Ground-Truths*. Os valores resultantes destes processos mostraram alto grau de similaridade entre os resultados obtidos e os resultados esperados (*ground-truths*). Esta validação está descrita na seção 6.1.

7.2 Trabalhos Futuros

Dois principais trabalhos futuros podem ser desenvolvidos e são recomendados:

1 – Melhoria do método de definição da área timpânica – isto traria menores áreas de falso positivo à esta etapa e, por consequência, resultados mais consistentes;

2 – Automatização do processo de segmentação – o primeiro esforço já foi dado no sentido de segmentar estas patologias, a automatização traria mais facilidades e simplicidade para o processo, do ponto de vista do especialista.

Referências

- Canny, J. F. (1986). A Computational Approach To Edge Detection. *IEEE Trans. Pattern Analysis and Machine Intelligence* , 679-714.
- Caselles, V., & Coll, T. (1996). Snakes in movement. *SIAM Journal on Numerical Analysis* , 2445-2456.
- Cohen, L. (1991). On active contour models and balloons. *CVGIP: Image Understanding* , 211-218.
- Costa, S. S., & Dornelles, C. d. (2006). *Otorrinolaringologia – Princípios e Prática, 2ª edição*. Artmed.
- Costa, S. S., Dornelles, C., Schmidt, L. P., Wangenheim, A. v., Comunello, E., & Heck Junior, V. (2008). *Correlação do percentual de timpanoesclerose na membrana timpânica com a perda auditiva*. Belo Horizonte: Revista Brasileira de Otorrinolaringologia.
- Cyclops. (12 de Dezembro de 2007). *Portal de Telemedicina*. Acesso em 2007, disponível em Portal de Otologia: <http://www3.telemedicina.ufsc.br/teleoto/>
- Dornelles, C., Costa, S. S., Wangenheim, A. v., Rosito, L., Comunello, E., & Heck Junior, V. (2007). Análise quantitativa de patologias da membrana timpânica: uma técnica computacional semi-automática para acompanhamento e mensuração. *Revista Brasileira de Otorrinolaringologia (Online)* .
- Fonseca, H., Vasco, S., & Ferreira, A. (2002). *Constituição do Sistema Auditivo Humano*. Acesso em 20 de Janeiro de 2008, disponível em <http://telecom.inescn.pt>
- Gonzalez, R. C., & Woods, R. E. (2000). *Processamento de Imagens Digitais*. São Paulo: Edgard Blücher.
- Heck Junior, V., Wangenheim, A. v., Abdala, D. D., Comunello, E., Dornelles, C. d., & Costa, S. S. (2007). Computational Techniques for Accompaniment and Measuring of Otology Pathologies. *20th IEEE International Symposium on Computer-Based Medical Systems* (pp. 53-58). Maribor: IEEE Computer Society.

- Hsu, C. Y., Chen, Y. S., Hwang, J. H., & Liu, T. C. (2004). A computer program to calculate the size of tympanic membrane perforations. *Clinical Otolaryngology* , pp. 340–342.
- Jiang, X., Marti, C., Irniger, C., & Bunke, H. (2006). Distance measures for image segmentation evaluation. *EURASIP Journal on Applied Signal Processing* , 1-10.
- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: active contour models. *Int. Journal of Computer Vision* , 321-331.
- Lenna. (1972). *Lenna*. Acesso em 20 de Dezembro de 2007, disponível em Wikipedia: <http://en.wikipedia.org/wiki/Lenna>
- Li, C., Xu, C., Gui, C., & Fox, M. D. (2005). Level set evolution without re-initialization: A new variational formulation. *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 430-436). Washington: IEEE Computer Society.
- Lindblad, T. (2005). *Image Processing Using Pulse Coupled Neural Networks*. Springer.
- Longman. (2006). Longman Dictionary of Contemporary English.
- Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. *8th Int'l Conf. Computer Vision*, (pp. 416-423).
- Mascarenhas, N. D., & Velasco, F. R. (1989). *Processamento Digital de Imagens*. Buenos Aires: Kapelusz.
- Mathworks. (2007). *MATLAB and Simulink for Technical Computing*. Acesso em 07 de Dezembro de 2007, disponível em The MathWorks: <http://www.mathworks.com/>
- Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications On Pure & Applied Mathematics* , 557-684.

Pedrini, H., & Schwartz, W. R. (2008). *Análise de Imagens Digitais*. São Paulo: Thomson Learning.

Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*. Vol. 66. , 846-850.

Schalkoff, R. J. (1989). *Digital Image Processing and Computer Vision*. John Wiley.

Wangenheim, A. v. (2007). *Visão Computacional*. Acesso em 10 de Dezembro de 2007, disponível em Seminário Introdução à Visão Computacional: <http://www.inf.ufsc.br/~visao/>

Yoo, T. S., Ackerman, M. J., & Vannier, M. (2000). Toward a Common Validation Methodology for Segmentation and Registration Algorithms. *Medical Image Computing and Computer-Assisted Intervention*. Pittsburgh: Springer Berlin / Heidelberg.

Anexos

A. Medidas de Distância

A.1 Índice Rand

O índice Rand, divulgado pela primeira vez em (Rand, 1971) e revisado em (Jiang, Marti, Irniger, & Bunke, 2006), é uma medida de similaridade desenvolvida especialmente para verificar a qualidade de algoritmos de *clustering* (aglomeramento). A verificação é feita através da comparação com outros resultados de *clustering* ou com um padrão ouro (*ground truth*). Para comparar dois resultados de *clustering* $C1 = \{c1_1, c1_2, \dots, c1_N\}$ e $C2 = \{c2_1, c2_2, \dots, c2_M\}$ na mesma imagem $P = \{p_1, p_2, \dots, p_K\}$ onde cada elemento de $C1$ ou $C2$ é um subconjunto de P e $c1_j = \{p1_j, p2_j, \dots, pL_j\}$, as seguintes quantidades são calculadas:

- N_{11} – o número de *pixels* convergentes de ambos os *clusters* $C1$ e $C2$.
- N_{00} – o número de *pixels* divergentes de ambos os *clusters* $C1$ e $C2$.

O índice Rand é definido pela Equação (12).

$$R(C1, C2) = 1 - \frac{N_{11} + N_{00}}{\frac{n(n-1)}{2}} \quad (12)$$

Para computar as quantidades de N_{11} e N_{00} , um deles deve iterar sobre toda a imagem, por cada *pixel*, para poder avaliar as condições definidas anteriormente, gerando assim um algoritmo de ordem $O(n^4)$. Uma solução mais prática é usar o método descrito em (Jiang, Marti, Irniger, & Bunke, 2006), onde uma matriz de correspondência é usada para diminuir as ocorrências dos *pixels* na classe respectiva. A matriz de correspondência é construída alocando cada *cluster* de $C1$ para uma linha e cada *cluster* de $C2$ para uma coluna. As células da matriz são então definidas como uma intersecção dos *clusters* especificando cada linha e coluna. Se a matriz de correspondência tem um tamanho $k \times l$, cada célula pode ser definida como $m_{ij} = |c_i \cap c_j|$, $c_i \in C1$, $c_j \in C2$.

As quantidades de N_{11} e N_{00} podem ser computadas pelas condições da matriz de correspondência como segue:

$$N_{11} = \frac{1}{2} \left(\sum_{i=1}^k \sum_{j=1}^l m_{ij}^2 - n \right) \quad (13)$$

$$N_{00} = \frac{1}{2} \left(n^2 - \sum_{i=1}^k n_i^2 - \sum_{j=1}^l n_j^2 + \sum_{i=1}^k \sum_{j=1}^l m_{ij}^2 \right) \quad (14)$$

onde n é a cardinalidade de P e n_i e n_j são a cardinalidade dos *clusters* $c1_i$ e $c2_j$.

A.2 Correspondência de Grafos Bipartidos (BGM)

O índice BGM (Jiang, Marti, Irniger, & Bunke, 2006) calcula a correlação de um para um entre os *clusters* e ao mesmo tempo tenta maximizar a sua correlação. O BGM considera cada *cluster* de $C1$ e $C2$ como vértices de um grafo bipartido. Bordas são adicionadas entre cada vértice das duas partições e elas são estimadas como $|c1_i \cap c2_j|$, um valor que pode ser extraído diretamente da matriz de correspondência. Então o grafo bipartido com peso máximo é definido como o sub-grafo $\{(c1_{i1}, c2_{j1}), \dots, (c1_{ir}, c2_{jr})\}$, onde apenas as bordas de $c1_i$ a $c2_j$ com peso máximo estão presentes. Depois que todas as bordas com valores máximo foram encontradas, o peso geral do grafo é calculado pela soma de todos os pesos de bordas restantes.

B. Tabela de Resultados

Identificadores		Forma de Mensuração			Avaliação do Tímpano				Avaliação da Patologia			
Paciente	Nome arq.	%GT	%AURIS	DIF %	Rand	BGM	% FN	% FP	Rand	BGM	% FN	% FP
01	aomc009e	43,002	36,709	6,293	0,124	0,066	0,025	14,360	0,014	0,007	2,362	2,023
02	aomc0107d	17,905	14,194	3,711	0,119	0,063	0,891	18,828	0,005	0,003	4,424	1,254
03	aomc029e	27,005	23,902	3,103	0,260	0,272	0,014	17,221	0,080	0,042	11,154	16,895
04	aomc032e	45,052	31,122	13,930	0,383	0,351	0,000	31,374	0,019	0,010	2,283	2,926
05	aomc034e	43,144	34,062	9,083	0,172	0,095	0,048	20,383	0,016	0,008	2,974	2,105
06	aomc035d	2,412	1,876	0,536	0,248	0,145	5,614	20,225	0,002	0,001	8,957	1,084
07	aomc040d	41,180	37,268	3,912	0,121	0,065	4,893	11,722	0,016	0,008	3,863	1,400
08	aomc043e	52,579	51,502	1,077	0,096	0,050	4,522	7,089	0,021	0,011	2,068	2,709
09	aomc048e	11,918	9,847	2,071	0,090	0,047	4,760	13,261	0,007	0,004	10,760	1,634
10	aomc049e	9,554	7,913	1,641	0,099	0,052	2,388	12,969	0,005	0,003	8,158	1,130
11	aomc051d	15,592	15,532	0,060	0,150	0,082	0,000	18,651	0,040	0,020	6,980	24,035
12	aomc063e	26,566	23,924	2,642	0,093	0,049	2,573	14,938	0,010	0,005	2,046	5,033
13	aomc064d	30,631	28,213	2,417	0,075	0,039	2,270	7,219	0,014	0,007	4,433	1,499
14	aomc064e	30,718	28,279	2,439	0,086	0,045	0,762	8,211	0,012	0,006	2,411	1,952
15	aomc077d2	43,743	48,173	4,430	0,087	0,046	5,613	2,427	0,055	0,028	2,308	8,297
16	aomc089e	8,233	7,015	1,218	0,114	0,061	1,284	15,153	0,019	0,010	6,829	6,019
17	aomc089e	23,647	19,698	3,949	0,114	0,061	1,284	15,153	0,007	0,003	5,277	2,264
18	aomc093d	24,998	21,164	3,834	0,186	0,104	6,554	19,101	0,009	0,005	3,605	1,432
19	aomc097d	2,807	2,098	0,709	0,148	0,080	0,949	23,096	0,001	0,001	6,841	3,219
20	aomc098e	37,781	29,004	8,778	0,218	0,124	1,967	21,584	0,016	0,008	4,453	0,442
21	aomc110e	22,975	15,176	7,798	0,220	0,126	0,000	33,961	0,011	0,005	4,739	4,764
22	aomc114e	15,499	14,733	0,767	0,081	0,042	3,067	6,712	0,006	0,003	2,897	1,686
23	aomc116d	52,287	46,844	5,444	0,119	0,064	3,446	7,615	0,049	0,025	7,457	1,163
24	aomc120d	22,924	17,891	5,032	0,190	0,106	0,000	25,299	0,014	0,007	2,642	6,817
25	aomc135d	21,736	23,935	2,200	0,172	0,095	12,444	1,949	0,007	0,003	4,151	2,527
26	aomc255d	22,214	19,797	2,417	0,055	0,029	2,496	8,712	0,008	0,004	6,095	1,350
27	aomc337e	30,431	36,346	5,914	0,076	0,040	10,960	5,354	0,021	0,011	0,878	11,783
28	aomc338e	0,352	0,224	0,128	0,074	0,038	6,854	13,440	0,000	0,000	32,864	2,055
29	aomc346e	38,671	35,198	3,473	0,070	0,036	1,841	10,846	0,010	0,005	2,275	2,481
30	aomc367e	13,188	8,374	4,813	0,150	0,082	4,010	36,063	0,002	0,001	5,859	1,253
31	aomc402e	24,338	13,512	10,827	0,081	0,042	0,000	44,281	0,002	0,001	4,572	4,224
32	aomc415d	43,872	29,970	13,902	0,167	0,092	0,103	25,602	0,020	0,010	8,492	0,237
33	aomc423d	43,008	32,253	10,756	0,160	0,088	0,000	23,813	0,015	0,008	3,976	2,446
34	aomc424d	17,753	11,944	5,809	0,256	0,150	0,000	32,167	0,007	0,004	3,690	2,894
35	aomc443ere	12,006	12,024	0,019	0,108	0,057	8,515	4,628	0,010	0,005	6,885	3,079
36	aomc505e	3,344	4,313	0,969	0,186	0,104	29,742	0,147	0,003	0,001	10,143	0,983
37	aomc527e	23,202	24,603	1,401	0,073	0,038	10,320	0,226	0,011	0,005	5,524	0,875
38	aomc531d	24,921	24,125	0,796	0,084	0,044	7,613	11,122	0,006	0,003	2,200	2,811
39	aomc531d	22,139	21,943	0,196	0,084	0,044	7,613	11,122	0,010	0,005	3,509	6,344
Média		23,647	21,943	3,103	0,119	0,063	2,388	14,360	0,010	0,005	4,433	2,264
Desvio		14,232	12,972	3,726	0,069	0,063	5,450	10,359	0,016	0,008	5,178	4,643
Erro		2,279	2,077	0,597	0,011	0,010	0,873	1,659	0,003	0,001	0,829	0,744

Tabela 4- Resultados dos 39 casos utilizados para validação do método