

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

**Diagnóstico do Câncer de Mama Auxiliado por Computador - Um Protótipo
para Detecção de Microcalcificações**

por
Gilmário Barbosa dos Santos

Dissertação submetida à Univer-
sidade Federal de Santa Catarina
para obtenção do grau de Mestre
em Ciência da Computação.

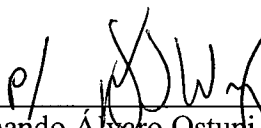
Prof. Dr. Aldo von Wangenheim, Dr.
Orientador

Florianópolis
1999

DIAGNÓSTICO DO CÂNCER DE MAMA AUXILIADO POR COMPUTADOR – UM PROTÓTIPO PARA DETECÇÃO DE MICROCALCIFICAÇÕES

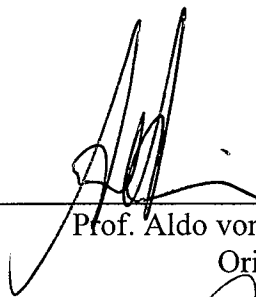
GILMÁRIO BARBOSA DOS SANTOS

ESTA DISSERTAÇÃO FOI JULGADA ADEQUADA PARA A OBTENÇÃO DO TÍTULO DE **MESTRE EM CIÊNCIAS DA COMPUTAÇÃO** E APROVADA EM SUA FORMA FINAL PELO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO



Prof. Fernando Alvaro Ostuni Gauthier, Dr.
Coordenador do Curso

Banca examinadora:



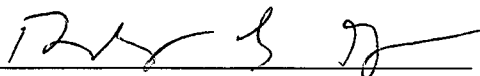
Prof. Aldo von Wangenheim, Dr.
Orientador



Prof. João Bosco da Mota Alves, Dr.



Prof. Luis Fernando Jacinto Maia, Dr.



Bráulio L. Fernandes, M.D.

Florianópolis, 22 de Dezembro de 1999

“...Bahia, ó quão dessemelhante...”

Caetano Veloso

Dedico este trabalho aos meus familiares, especialmente aos meus pais e avós.

AGRADECIMENTOS

À minha família.

À minha companheira Rúbia e seus familiares.

Ao Projeto Cyclops.

Aos professores e funcionários do PGCC, especialmente a Vera Lúcia (Verinha) e a Valdete.

À UFSC e à gratuidade do ensino superior.

À CAPES pela subvenção inicial imprescindível.

Aos amigos e colegas de curso.

Ao SUGAR - Smalltalk Users Group of Argentina, pela possibilidade de intercâmbio de informações.

Ao staff das clínicas DMI e SONITEC, especialmente ao Dr. Felipe, Dra Marcela, Dr. Evaldo, Cleci, Marli, Sandro.

Ao Dr. Braúlio (HU).

À todos aqueles que colaboraram com esse trabalho.

SUMÁRIO

LISTAS DE ILUSTRAÇÕES.....	viii
LISTA DE ABREVIATURAS.....	ix
RESUMO.....	x
ABSTRACT	xi
1. Introdução	1
1.1 Justificativa	2
1.2 Objetivos do Trabalho	2
1.3 Estrutura do Trabalho	3
2. O Câncer de Mama.....	5
2.1 Formação do Câncer	5
2.2 Câncer de Mama	6
2.3 Detecção Precoce	6
2.4 Exame Clínico	7
2.5 Técnicas Imagenológicas	7
2.6 Mamografia	7
2.7 Densidade Radiológica	8
2.8 Mamografia Digital	10
3. Detecção de Microcalcificações e CAD.....	12
3.1 Sinais Mamográficos do Câncer de Mama	12
3.2 Microcalcificações.....	12
3.2.1 Fatores Individuais	13
3.2.2 Fatores Grupais	13

3.2.2.1	Caraterísticas típicas	14
3.3	Detecção de Microcalcificações	15
3.4	CAD - Diagnóstico Auxiliado por Computador	16
3.4.1	Aquisição da Imagem	17
3.4.2	Pré-processamento	17
3.4.3	Detecção de Calcificações	17
3.4.4	Sugestão de Diagnóstico	17
3.5	Tecnologias para a Análise Computacional de Imagens Médicas	18
3.5.1	Domínio das Frequências	18
3.5.2	Domínio Espacial	18
3.6	Conceitos Básicos	19
3.6.1	Elemento Estrutural – Kernel	19
3.6.2	Convolução Discreta	19
3.6.3	Operação Threshold (limiar)	21
3.6.4	Segmentação	22
3.7	Morfologia matemática	22
3.7.1	Erosões e Dilatações sobre tons de cinza	26
3.8	Estado da Arte em CADs	28
4.	Implementação	41
4.1	Linguagens de Programação	41
4.2	Sistema Operacional	42
4.3	Implementação	43
4.3.1	Preparação	43
4.3.2	Técnica da Subtração de Imagens	49
4.4	Detecção	51
4.4.1	Threshold Binário Global	51
4.4.2	Threshold Binário Baseado em Histograma	51
4.5	Interface	53
4.6	Classes Utilizadas	53
4.6.1	Classes do Ambiente de Programação	54
4.6.2	Classes Nativas do Projeto Cyclops:	54

4.6.3 Classes Novas	54
4.6.3.1 MammaCanvasLauncher	54
4.6.3.2 MammaCanvas1	55
4.6.3.3 ProcImageLauncher	55
4.6.3.4 MammaMagnifyinGlass	56
4.6.3.5 CollumnKernel	56
4.6.3.6 MammaHistogramFactor	56
4.6.3.7 InterfaceMamasC	56
4.6.3.8 O modelo MVC	57
5. Testes	59
5.1 MIAS	59
5.2 Imagens Locais	60
5.3 Comparando Fontes de Dados	60
5.4 Preparação para Testes	61
5.5 Procedimentos de Avaliação	62
5.5.1 Avaliação Geral	64
5.5.2 Interface	65
6. Conclusões	68
6.1 Objetivos Alcançados	68
6.2 Dificuldades Encontradas	69
6.2.1 Geração da Mamografias	69
6.2.2 Digitalização Manual	69
6.2.3 Outras Dificuldades	70
6.3 Recomendações para Trabalhos Futuros	70
6.3.1 Inserção de Técnicas de IA	70
6.3.2 Conexão com Teleradiologia	70
6.3.3 Conexão ao DICOM	71
7. Referências Bibliográficas	72
Anexos	75

LISTA DE ILUSTRAÇÕES

Figura 2.1: Principais Vistas De Uma Mamografia. Os raios-X incidem atravessando um corpo que apresenta múltiplas densidades radiológicas. Adaptado de Meschan (1976, p. 2)	8
Figura 2.2: Os raios-X incidem atravessando um corpo que apresenta múltiplas densidades radiológicas. Adaptado de Meschan (1976, p. 2)	9
Figura 2.3: Comparação entre mamografias de mama com baixa densidade (a) e mama densa (b)	10
Figura 2.4: Diagrama de esquema de mamografo digital, retirado de FEIG e YAFFE (1995,p. 1205)11	11
Figura 3.5: Mamografia apresentando agrupamento microcalcificações	13
Figura 3.6: Classificação de microcalcificações segundo aspectos morfológicos e de distribuição segundo Moldes do ACR MAMMOGRAPHY LEXICON.	15
Figura 3.7: Um exemplo de CAD.	16
Figura 3.8: Convolução de uma imagem por um kernel resultando em uma imagem	20
Figura 3.9: Equivalente topográfico de uma imagem. HIGHNAM (1996, p.4)	21
Figura 3.10: Pixels de uma imagem com 500 dpi de resolução	23
Figura 3.11: Resultado da operação de erosão (figura A) e dilatação (figura B) de X por B.	25
Figura 3.12: Exemplos de operações de opening e closing, FILHO & NETO (1999, p. 191-2).....	26
Figura 3.13: Erosão em tons de cinza, o quadrado representa um kernel que erode os pixels de uma imagem, os pixels das filas inferiores representam a imagem resultante.	27
Figura 4.14: Sistema Smalltalk de execução de programas	41
Figura 4.15: Esquema em blocos da abordagem proposta.	43
Figura 4.16: Esquema de aquisição da imagem de trabalho.....	44
Figura 4.17: Threshold binário (esquerda) seguido de 4 dilatações	45
Figura 4.18: Aplicação do kernel coluna para definição de sentido (em X) de rotulação	47
Figura 4.19: Imagem anterior e posterior a separação da mama	49
Figura 4.20: Técnica da diferença entre imagens em um perfil hipotético	50
Figura 4.21: Histograma de uma imagem em tons de cinza	52
Figura 4.22: Especializações de classes do ambiente Visualworks 3	54
Figura 4.23: Interface que lança o protótipo	55
Figura 4.24: Principais classes implementadas e seus relacionamentos	56
Figura 4.25: Janela principal do protótipo	57
Figura 4.26: Principais classes que implementam o modelo MVC	58
Figura 5.27: Imagem normal (A), imagem após threshold com limiar = 10 (B) e limiar = 2 (C)	60
Figura 5.28: Comparação entre imagens do MIAS (esquerda) com imagens locais	61
Figura 5.29: Interface alterada para fins de teste de avaliação dos métodos de detecção	63
Figura 5.30: Percentual de acerto sobre o total de pixels detectados	66
Figura 5.31: Variação da acurácia e falsos positivos - Threshold por histograma	66
Figura 5.32: Variação da acurácia e falsos positivos - Threshold por estatística global	67
Figura 5.33: Defeito na imagem acarretando detecção de pixels indevidamente	67
Figura A1: Domínio de aplicação DICOM	78
Figura A2: Principais classes do sistema	79
Figura A3: (continuação da figura A1).....	80
Figura A4: Acesso a janela workspace para abertura do launcher da interface principal do sistema ..	83
Figura A5: Mammography Loader - lançador da janela principal	84
Figura A6: Mammography Viewer - janela principal e suas funções	85
Figura A7: Entrada de fator para cálculo de limiar por histograma	86
Figura A8: tela de magnificação	87
Tabela 1: Significado dos campos de registros na documentação MIAS	60
Tabela 2: Totais pixels detectados dentro da area suspeita (acertos) VERSUS total de pixels detectados fora da região suspeita (erros) - usando o MIAS Database	65

LISTA DE ABREVIATURAS

CAD	-	COMPUTER AIDED DIAGNOSIS.
DICOM	-	DIGITAL IMAGE COMMUNICATIONS IN MEDICINE.
DLL	-	DYNAMIC LINK LIBRARY.
PGM	-	PORTABLE GRAY MAP.
ROI	-	REGIONS OF INTEREST.
RX	-	RAIO X.
TAP	-	TRASFORMAÇÃO DE ÁREA EM PONTO.

RESUMO

Este trabalho apresenta o resultado das atividades realizadas no âmbito do Projeto Cyclops, no ano de 1999, sobre o diagnóstico precoce do câncer de mama auxiliado por computador, através do processamento de imagens oriundas de mamografias convencionais digitalizadas. Partindo-se da revisão do estado da arte e dos aspectos contextuais acerca do câncer de mama, construiu-se um protótipo com interface amigável. Tal protótipo contempla dois métodos de filtragem espacial de imagens: o *threshold* baseado em estatística global e o *threshold* baseado em histograma. Os testes finais foram feitos com imagens obtidas de mamografias convencionais digitalizadas, cedidas por clínicas especializadas locais, como também imagens obtidas no banco de imagens mamográficas mantido pelo MIAS (U.K.).

ABSTRACT

This thesis reports the results obtained after twelve months of work in the Cyclops Project; basically it is about breast cancer detection aided by computer. For this sake it has been used image processing techniques applied over images obtained from conventional mammographies. A prototype was built with two spatial filters: global threshold and threshold histogram based. In order to test the prototype, mammographies were used from local health institutions beside images from MIAS image database.

CAPÍTULO 1 - Introdução

O câncer de mama é considerado hoje uma epidemia, sua frequência especialmente acentuada sobre a população feminina caracteriza-o quase como exclusivo desta, não bastasse isso, o câncer de mama tem resistido a todos os esforços de pesquisa em se conseguir uma cura independente do seu estágio. Por outro lado, também não se encontram fatores de risco com o poder de predição adequado à implementação de políticas preventivas de saúde, mais eficientes. Este cenário tem conduzido os esforços dos setores competentes na direção da detecção precoce, dado saber-se que quanto mais cedo se detecta o câncer, maiores são as chances de uma cura ou sobrevida mais longa.

Dentre as muitas técnicas convencionais auxiliares à detecção precoce do câncer de mama, a mais popular e eficiente é a mamografia, ou seja, a radiografia da mama. A análise mamográfica é realizada por especialistas médicos, os quais procuram identificar sinais de anormalidades que possam revelar o desenvolvimento do câncer, dentre estes, um dos principais sinais procurados são as microcalcificações. Contudo, apesar da mamografia ser um exame altamente sensível às alterações na mama, há uma série de fatores que dificultam um diagnóstico mamográfico acurado, o que pode levar tanto a um diagnóstico errado, como à solicitação – por insegurança quanto ao diagnóstico – de exames extras mais decisivos que elevam o custo do tratamento, podendo vir a comprometer ou inviabilizar políticas públicas de prevenção.

Tanto sob o ponto de vista contábil quanto clínico, é premente melhorar a acurácia dos diagnósticos, portanto também é urgente a utilização de ferramentas de apoio ao especialista médico. Para atender esta demanda tem se investido muito no desenvolvimento de *softwares* para auxílio no diagnóstico precoce do câncer de mama, com vistas a (i) propiciar uma diminuição das taxas de mortalidade por essa doença e (ii) possibilitar a economia de recursos no que se refere a exames adicionais que são mais traumáticos para a paciente e que frequentemente são solicitados desnecessariamente. Essa categoria de *softwares* é conhecida como CAD ou *Computer Aided Diagnosis*, tais sistemas basicamente se utilizam de recursos da Ciência da Computação, principalmente métodos de processamento de imagens e visão computacional, acrescidos ou não de inteligência artificial e compreendem desde os que destacam ROI's - *Regions of Interest*, ou seja, regiões suspeitas de possuir sinais de câncer em uma

imagem de mamografia digital ou digitalizada, até aqueles mais “inteligentes” que de fato sugerem diagnósticos ao especialista.

Este trabalho foi desenvolvido no âmbito do projeto Cyclops, uma iniciativa de cooperação entre Brasil e Alemanha, tendo como coordenador no Brasil o prof. Dr. Aldo von Wangenheim (UFSC) e na Alemanha o prof. Dr. Michael M. Richter (Universidade de Kaiserslautern). O sistema Cyclops é um laboratório para a análise automatizada de imagens e tem como proposta desenvolver e implementar novos métodos, técnicas e ferramentas no campo da análise de imagens médicas, para tanto usa como fonte os conhecimentos oriundos da inteligência artificial e da área de visão computacional.

1.1. Justificativa

Pesquisar sobre contribuições da Ciência da Computação na área da saúde é um esforço plenamente justificável, pesquisar sobre tais contribuições no contexto de uma doença tão grave quanto o câncer de mama, mais ainda. Além da justificativa social envolvida, implica em travar contato com o que há de mais atual em processamento de imagens, visão computacional e inteligência artificial. Ao mesmo tempo significa construir um pensamento crítico a respeito da tecnologia utilizada, adequando-a a realidade brasileira, de forma a construir-se um *know-how* que, se continuamente incrementado, pode gerar um produto completo e por que não, competitivo internacionalmente.

1.2. Objetivos do Trabalho

1.2.1. Geral

Estudar o problema da detecção precoce do câncer de mama e em que aspectos a Ciência da Computação tem contribuído para a melhoria da acurácia dos diagnósticos.

1.2.2. Específicos

Implementar modelos de detecção de microcalcificações colhidos do levanta-



mento do estado da arte no que se refere à detecção de câncer de mama auxiliada por computador, de forma a encontrar ROIs.

Implementar uma interface amigável para os modelos implementados utilizando com o plataforma um PC, porém sem perder a perspectiva da portabilidade para outras plataformas.

Promover a apropriação do conhecimento tecnológico envolvido na problemática da detecção do câncer de mama, de forma a oferecer uma referência para estudos futuros.

1.3. Estrutura do Trabalho

Nos capítulos seguintes serão abordados os aspectos teóricos relacionados a investigação do tema:

No capítulo II aborda-se o problema do câncer de mama. Explicita-se quais as estratégias utilizadas para o diagnóstico precoce, bem como as ferramentas de diagnóstico através de imagem atualmente disponíveis para detecção dos sinais de câncer, coloca-se que a mamografia é o método atualmente mais utilizado, discute-se então a questão da densidade radiológica como fator de influência no grau de dificuldade na análise de uma mamografia.

No capítulo III faz-se uma descrição dos aspectos morfológicos das microcalcificações, descreve-se os tipos considerados típicos quanto à malignidade e benignidade, sendo citada uma classificação de microcalcificações e alguns aspectos que levam as microcalcificações a serem consideradas os sinais de câncer mais importantes em uma mamografia. Aprofunda-se também a questão dos problemas relativos a análise de mamografias, desta vez sob a perspectiva da busca por microcalcificações. Justifica-se, então, a necessidade de ferramentas CAD para apoio a diagnóstico, esclarece-se que tal conceito está associado a uma classe de *softwares* que aplicam técnicas da área de processamento de imagens e inteligência artificial. Coloca-se alguns conceitos básicos a respeito de processamento de imagens e relata-se o estado da arte no tema.

O capítulo IV é dedicado a descrição da implementação do protótipo, descreve-se os algoritmos implementados, o ambiente de programação, as linguagens de programação e o sistema operacional utilizado.

O capítulo V é dedicado ao material utilizado para os testes e o procedimento



adotado.

No capítulo VI faz-se as considerações finais e recomendações para trabalhos futuros.

Por último tem-se as referências bibliográficas e anexos.



CAPÍTULO 2 - O Câncer de Mama

Um dos principais causadores de óbitos entre mulheres na faixa etária dos 40 aos 50 anos de idade, segundo estudos realizados nos E.U.A, é o câncer de mama. De fato, na proporção de uma entre oito mulheres americanas desenvolve essa doença durante alguma fase da vida. No Brasil a realidade não é muito diferente, aqui, dentre os cânceres, o câncer de mama lidera as estatísticas no que se refere à mortalidade entre mulheres. O entendimento desse quadro conduz à compreensão do quão importante é a detecção precoce desta doença, tanto por ainda ser inviável prever a sua ocorrência, quanto pelo fato do diagnóstico nos estágios iniciais melhorar significativamente as chances de sobrevivência. Este diagnóstico precoce pode ser obtido basicamente por intermédio do exame clínico, ou pela realização de rastreamento por imagem da mama.

2.1. Formação do Câncer

As células se reproduzem naturalmente em função de diversas necessidades, tais como: para substituição de células mortas, para regenerar um ferimento ou devido ao próprio crescimento orgânico do corpo. Nesse processo, mecanismos celulares controladores são responsáveis por regular tanto a proliferação (quantidade de células novas produzidas) quanto a diferenciação celular (semelhança entre a nova célula e uma célula sadia). Esse mecanismo regulador que mantém o equilíbrio entre o crescimento celular e o bem estar do organismo nem sempre se mantém, pois eventualmente há uma perda de controle gerando tumores, os quais são classificados como:

- 1) **Malignos (câncer):** nesse caso as células recém criadas não guardam semelhança ou são pouco semelhantes às que lhes deram origem. Há um alto risco de vida envolvido, principalmente devido à possibilidade de metástase do câncer, ou seja, a sua dispersão pelo corpo em função do desprendimento de células cancerígenas das massas malignas, que ao alcançarem a corrente sanguínea ou linfática provocam a disseminação da doença, o que pode levar à irreversibilidade da mesma e por conseguinte à morte;
-

2) Benignos: apesar de terem a mesma origem dos malignos não desenvolvem metástase e raramente levam a risco de vida.

2.2. Câncer de Mama

O câncer de mama é o tipo de câncer que se origina nas mamas, sendo muito raro em homens especialmente devido a atrofia destes órgãos nos machos, já que estes não funcionam no ciclo reprodutivo como nutrízes. Nos EUA, assim como no Brasil, a incidência deste câncer em homens corresponde a aproximadamente 1% do total de todas as ocorrências de câncer de mama, portanto tal incidência é considerada estatisticamente desprezível.

Um tumor maligno na mama também apresenta uma reprodução desordenada, invasiva e rápida, tomando-se como referência os casos benignos. Deve-se destacar que a ocorrência de malignidade na mama é bem menor que a de tumores benignos ou alterações benignas.

2.3. Detecção Precoce

Como a ciência ainda não descobriu uma forma eficaz de prevenção do câncer de mama, muito menos uma cura absoluta (cura independente do estágio da doença) o que resta é detectar a sua ocorrência o mais cedo possível de forma a se ter um estadiamento inicial e com isso chances de cura ou prolongamento da vida (PRO-ONCO, 1994, p. 1). Sendo assim as políticas de saúde pública hoje têm apontado para a massificação de procedimentos ditos de rastreamento do carcinoma de mama, dentre eles estão o exame clínico e as técnicas de imagem da mama, todos visando principalmente a detecção precoce.



2.4. Exame Clínico

O exame clínico é levado a termo por profissional médico, sendo um procedimento periódico ao qual todas as mulheres devem se submeter segundo uma frequência que é função principalmente da idade.

Com vistas a melhorar ainda mais as chances de detecção do câncer em seu estágio impalpável, ou seja, do chamado "câncer mínimo", faz-se uso de técnicas de imagem da mama, as descobertas feitas sobre esses exames contribuem decisivamente com a eficácia do processo de rastreamento do câncer, pois complementam o próprio exame físico, colaborando para um diagnóstico mais correto e preciso das alterações na mama, podendo servir como guia para estudos mais aprofundados na investigação médica.

2.5. Técnicas Imagenológicas

Diversos são os recursos de imagem utilizados para auxiliar o diagnóstico do exame clínico, os mais frequentemente mencionados na literatura são a xeromamografia, ressonância magnética, ultra-sonografia, termografia e a mamografia. Destas a (atualmente) mais importante é a mamografia.

2.6. Mamografia

A mamografia, senografia ou mastografia, é a radiografia da mama. Normalmente a mamografia é bilateral, ou seja, é feita uma radiografia de cada mama, sendo que existem basicamente duas projeções: a crânio-caudal, onde o feixe de raios-X incide na direção do eixo da coluna vertebral e no sentido do crânio para os pés; e a médio-lateral, onde os raios-X incidem na direção perpendicular (ou oblíqua) à projeção crânio-caudal (Figura 2.1).

A utilização de várias projeções pode ser útil na determinação de *artefatos*¹, bem como confir-

1. Artefato é qualquer tipo de objeto detectado na mamografia mas que não faz parte de nenhum elemento da mama. Tais objetos são inseridos no resultado do exame durante o processo de execução do mesmo e/ou revelação do filme radiológico. Exemplos: fios de cabelo, cílios, impressões digitais, etc.



mação das evidências detectadas em uma, na outra projeção. Atualmente a mamografia é a técnica adotada como exame padrão no processo de rastreamento do "câncer mínimo", o INCA (1996, p.33).

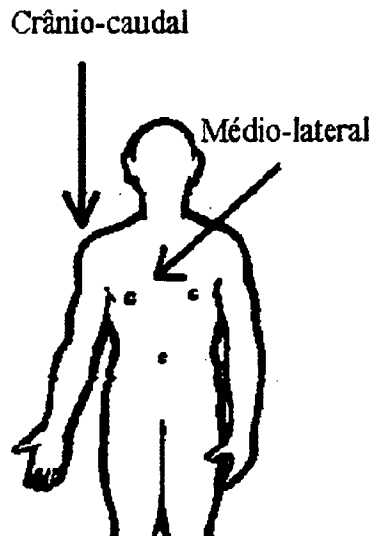


Figura 2.1: Principais vistas mamográficas

A principal vantagem da mamografia é a sua sensibilidade, uma mamografia consegue detectar o câncer de mama dois anos antes deste estar palpável, sendo um dos procedimentos mais confiáveis para o diagnóstico precoce, a aplicação de programas de rastreio baseado em mamografia permite diminuir em até 30% a mortalidade por câncer de mama em mulheres acima dos 50 anos de idade (NERCESSIAN, 1999, p.12). Por outro lado a sua desvantagem é a emissão de radiação ionizante, ela própria um risco para o desenvolvimento do câncer.

A sensibilidade da mamografia pode variar quando aplicada a mamas densas, para uma compreensão deste fenômeno faz-se necessário o entendimento de alguns fundamentos básicos de radiologia médica, o item seguinte se prestará a esse fim.

2.7. Densidade Radiológica

O termo densidade radiológica diz respeito a variação da capacidade de penetração dos RX (raios X) em função das diferenças entre os tecidos do corpo humano. Assim,



as partes de um órgão que têm a capacidade de absorver mais raios X têm uma aparência mais clara que as partes que absorvem menos os raios X, conforme figura abaixo (Figura 2.2).

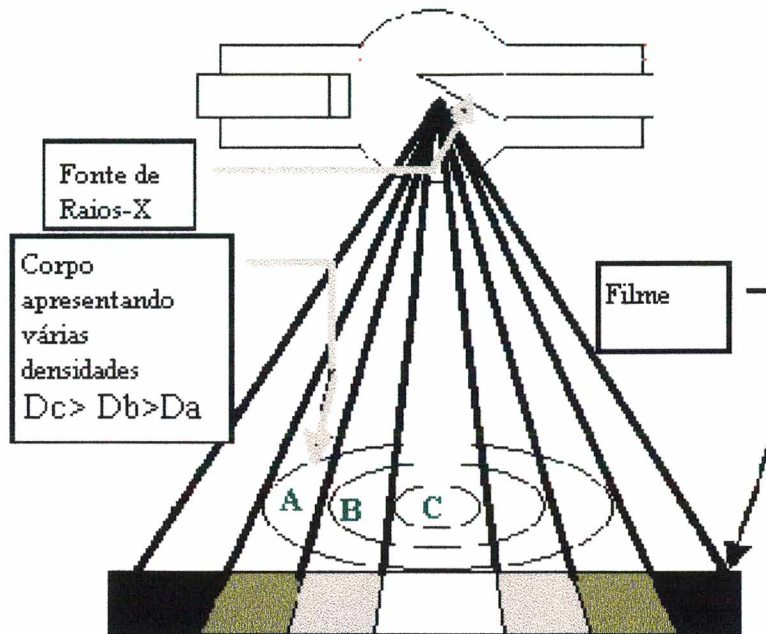


Figura 2.2: Raios-X atravessando um corpo com múltiplas densidades radiológicas. Adaptado de Meschan (1976, p. 2)

Na mamografia a variação da densidade radiológica pode ter efeitos negativos sobre a sensibilidade do exame, como pode ser visto comparando-se a mamografia de uma mama densa, como a da Figura 2.3-b, apresentando extensas áreas brancas, com a mamografia de uma mama com muita gordura e portanto menos densa, como a da Figura 2.3-a. Pode-se perceber o aspecto diferencial quanto a presença de áreas brancas, a maior extensão dessas áreas na mama densa se deve a uma maior absorção dos raios-X pelo seu estroma¹.

As microcalcificações, de forma semelhante ao estroma das mamas densas, também apresentam uma absorção relativa maior aos raios-X. Portanto, em mamas densas, as extensas áreas brancas verificadas devido à densidade maior do estroma, mascaram as microcalcificações, que também se apresentam na mamografia como pontos ou áreas brancas. Já nas mamas menos densas, há um destaque bem maior para as microcalcificações, devido ao contraste destas com o fundo mais escuro da mamografia. Dessa forma, a análise visual diferencial da mamografia de uma mama densa é menos sensível, portanto mais suscetível a enganos.

1. Tecido que forma a substância básica ou matriz de um órgão .



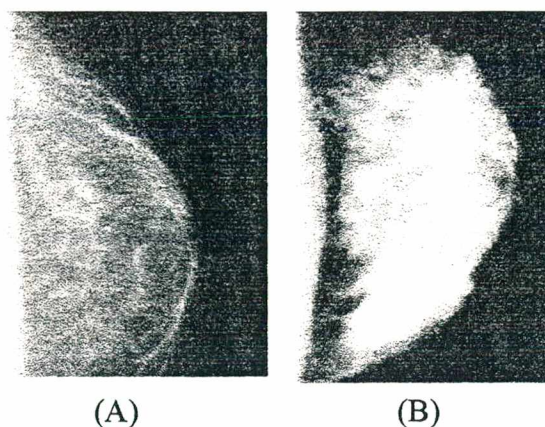


Figura 2.3: Comparação entre mamografias de mama com baixa densidade (a) e mama densa (b)

2.8. Mamografia Digital

A mamografia dita convencional é aquela cuja apresentação é feita sob a forma de películas (popularmente conhecida como “chapas de raio X”), constitui-se no resultado de revelação química de um filme que sofreu a incidência da radiação, sendo que a qualidade do resultado obtido é função de uma série de fatores tais como: a qualidade do produto químico utilizado, a perícia dos profissionais envolvidos, o grau de limpeza na câmara escura, etc, esses fatores exigem um controle contínuo sob pena de se comprometer o diagnóstico de um caso grave, por exemplo.

Com a crescente informatização dos hospitais e a necessidade de obter-se exames melhores, de forma mais eficiente e rápida, surgiu uma nova tecnologia que também tem como princípio a emissão de RX, mas que produz como resultado não mais uma película de filme, mas sim uma imagem mamográfica digitalizada (Figura 2.4), por digitalizado entenda-se como a informação na forma binária, portanto entendível pelo computador. Não por coincidência, essa característica possibilita diversas aplicações da ciência da computação no campo da informática médica, viabilizando: (i) o armazenamento dos exames mamográficos em bancos de dados informatizados, juntamente com informações sobre o paciente; (ii) a utilização **direta** da mamografia em sistemas de computadores em rede, com possibilidades enormes de aplicações que visem a melhoria dos diagnósticos; (iii) uma melhor padronização na geração das imagens, propiciando uma maior generalização dos métodos de processamento



de imagens voltados à imagenologia; (iv) a possibilidade de execução **direta** de processamentos sobre o resultado digital do exame, de forma a auxiliar-se o médico no seu trabalho; etc. Contudo, o mamografo digital ainda representa um investimento muito alto para a maioria dos hospitais e clínicas especializadas do Brasil.

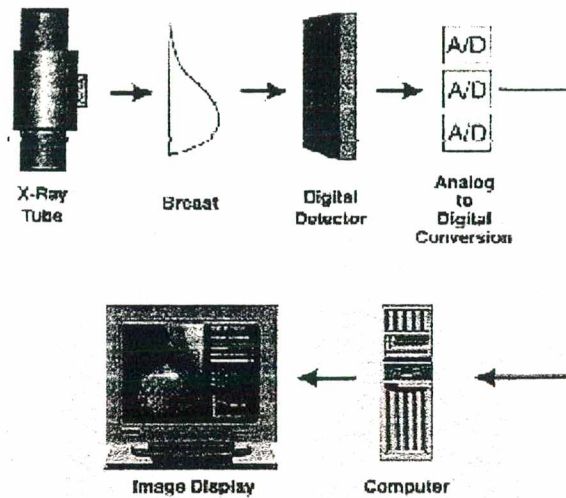


Figura 2.4: Diagrama mamografo digital, retirado de FEIG e YAFFE (1995,p. 1205)



CAPÍTULO 3 - Detecção de Microcalcificações e CAD

A mamografia tem atualmente um papel de destaque no processo de rastreamento do câncer em seu estágio inicial, porém sua sensibilidade pode ser afetada por fatores que dificultam a detecção de importantes sinais de câncer, como as microcalcificações. Elas podem apresentar-se sob diversas formas, tamanhos, densidades e quantidades, podendo ainda ocorrer agrupadas ou não. Para suplantar essa dificuldade e possibilitar uma maior exploração do potencial diagnóstico da mamografia, a ciência da computação, tem proposto alguns métodos como uma ferramenta de auxílio na detecção precoce do câncer. Neste capítulo aprofunda-se a análise destas questões.

3.1. Sinais Mamográficos do Câncer de Mama

Já foi devidamente esclarecido que a mamografia é a técnica de imagem de mama mais importante atualmente, mas, o que se procura em uma mamografia para fins de diagnóstico ? Dentre os diversos tipos de alterações identificáveis em uma mamografia, um dos mais importantes são as chamadas microcalcificações, elas assemelham-se a finos grãos de areia polvilhados (Figura 3.5), podem ocorrer em grupos ou solitárias, podendo estar associadas ou não a malignidade.

3.2. Microcalcificações

Microcalcificações são achados comuns na mamografia, a sua ocorrência é sinal extremamente importante, principalmente por que podem constituir a única alteração radiológica visível na mamografia, além de poderem ser os sinais mais precoces de malignidade em exames mamográficos (TABÁR, 1994, p. 16 e DUARTE, 1984, p. 13). Ademais, estima-se que de 20% a 30% dos casos que apresentam microcalcificações são diagnosticados

como malignos (OLIVEIRA C., 1997, p. 46).

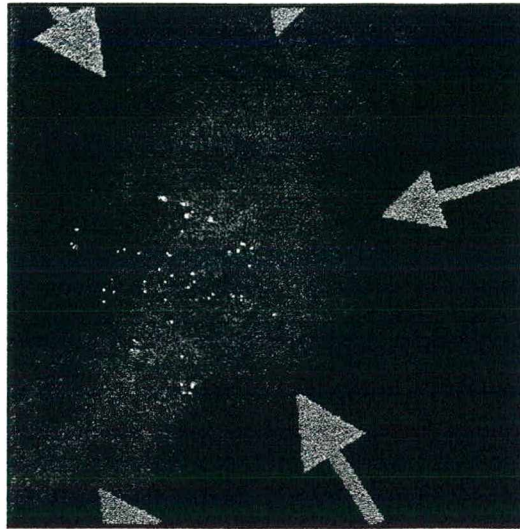


Figura 3.5: Mamografia apresentando agrupamento de microcalcificações

O diagnóstico mamográfico constitui em atividade de especialista e está longe de ser tarefa trivial, não apenas pela necessidade da habilitação profissional, mas também pela necessária habilidade na análise dos diversos fatores que se interrelacionam e devem ser considerados na definição do diagnóstico. A seguir lista-se um levantamento dos fatores encontrados na literatura médica considerados relevantes para fins de diagnóstico.

3.2.1. Fatores Individuais

Para cada microcalcificação concorrem os seguintes fatores:

1) Densidade: por serem resultantes de deposição de cálcio nos tecidos, as microcalcificações apresentam uma densidade relativamente alta, que pode variar entre exemplares de microcalcificações ou até mesmo no próprio exemplar, que pode apresentar variação de densidade em regiões diferentes na sua estrutura.

2) Forma/tamanho: as formas e tamanhos variam muito, há desde as formas de letras “Y”, “T” e “L”, formas lineares, puntiformes e ramificadas.

3.2.2. Fatores Grupais

Para os agrupamentos de microcalcificações concorrem os seguintes fatores:

1) Número: podem ocorrer numerosas ou não, quanto maior o seu número maior será a suspeita de malignidade.



2) Distribuição: é um aspecto de grande importância, em geral são unilaterais podendo estar agrupadas ou distribuídas em algum setor da mama.

3) Densidade do agrupamento: número de calcificações por cm³, em geral quanto maior a densidade maior o risco.

3.2.2.1. Características típicas

Na literatura calcificações **tipicamente malignas** possuem as seguintes características:

- a) São numerosas: quanto maior o seu número, maior a probabilidade de malignidade, quando o número é maior do que 30 as chances de malignidade são de 60% a 81% ;
- b) Apresentam-se em volume de cinco ou mais microcalcificações por cm³.
- c) São menores do que 0,5mm.
- d) São pleomórficas: têm uma alta capacidade de variação, quanto maior for esta variação (tamanho e forma) das microcalcificações entre si, maior o risco de malignidade.
- e) Possuem alta densidade, com variação entre elas e em si próprias.
- f) Distribuem-se de forma radiada a partir da papila.

Já as calcificações **tipicamente benignas**, que correspondem a maioria dos achados, são grandes e grosseiras, arredondadas com margens lisas ou não, podendo apresentar-se sob algumas formas peculiares tais como: anel, “casca de ovo”, “taça de chá”, “pipoca” (Figura 3.6).

Deve-se destacar que o exame definitivo para determinação da malignidade ou não de um caso, é o exame histológico, ou seja, o exame de tecido coletado na região suspeita. Tal procedimento só é aplicado em situações específicas, quando o especialista necessita de informação adicional que não é obtida nos convencionais.



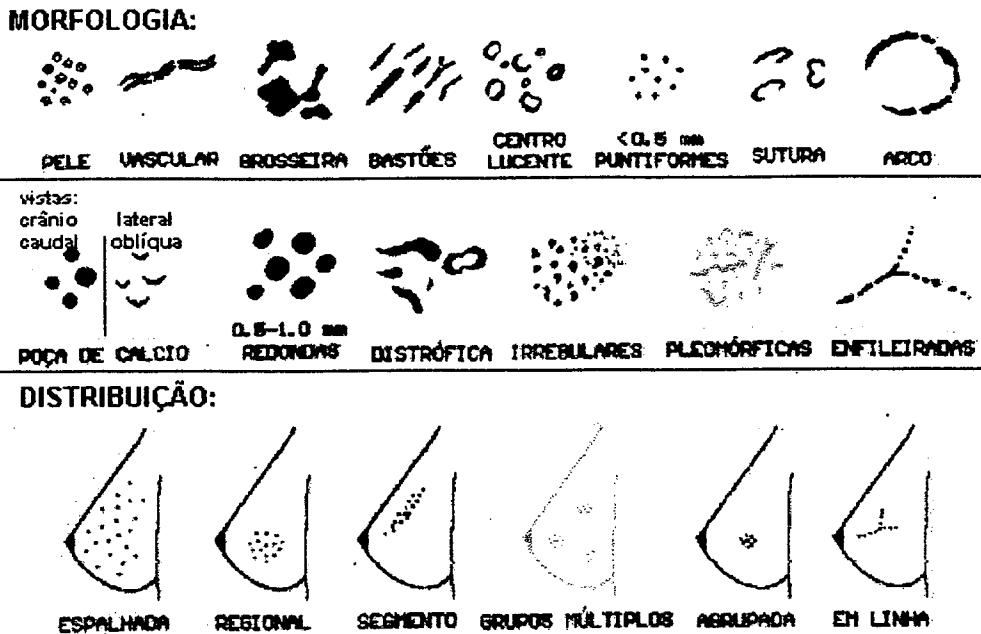


Figura 3.6: Microcalcificações segundo aspectos morfológicos e de distribuição segundo moldes do ACR MAMMOGRAPHY LEXICON.

3.3. Detecção de Microcalcificações

A sutileza com a qual as microcalcificações podem se apresentar aliada a fatores como a fadiga visual do médico ou a sua inexperiência, pode levar a detecção de casos falso positivos com conseqüente solicitação das biópsias desnecessárias. Vale ressaltar que uma biópsia apesar de ser uma pequena intervenção cirúrgica, implica em incremento de custo financeiro e do próprio stress psicológico que pode envolver a paciente. Segundo Heang-Ping CHAN (1988, p.664), muitas vezes os radiologistas têm que pesquisar na imagem de forma diligente e usando lentes de aumento para garantir uma alta taxa de detecção. Por causa da sutileza de muitas microcalcificações, o potencial para lapsos ou falha de está sempre presente. Essa atitude no mínimo implica na elevação dos custos de um tratamento. Segundo Atam P. DHAWAN (1993, p. 540-541) considerando-se a taxa de falso positivos de 0.88% (EUA), se todas as 50 milhões de mulheres americanas na faixa etária de risco para incidência de câncer de mama (ano de 1993) fossem examinadas, aproximadamente 520.000 delas, por ano, teriam recomendações para submeterem-se a exames de biópsia pelos seus médicos.



Ainda segundo o referido autor, a um custo de US\$ 1.600,00 por biópsia, isso implicaria em gastos da ordem de US\$ 832 milhões por ano. Porém, se a taxa de falso positivos baixasse para 0,66%; mantendo-se a taxa de verdadeiro positivos, seria possível uma economia de US\$ 184 milhões por ano.

Os dados de Atam Dhawan basearam-se nas solicitações de biópsia fundamentadas apenas na presença de microcalcificações, então, é bastante razoável dizer que a diminuições na taxa de falso positivos, sem perda de verdadeiros positivos, pode ser obtida através da utilização de alguma ferramenta de apoio a decisão atuando na análise de microcalcificações. Como em outras aplicações, tal tarefa pode ser desempenhada por algum software de computador desenvolvido para esse fim.

3.4. CAD - Diagnóstico Auxiliado por Computador

Nesse contexto surgiu a motivação para o desenvolvimento de ferramentas para auxílio na análise de mamografias, tais ferramentas comumente designadas como CAD, um acrônimo inglês para Computer Aided Diagnosis (diagnóstico auxiliado por computador) se caracterizam pela aplicação de técnicas da ciência da computação, principalmente processamento de imagens e inteligência artificial.

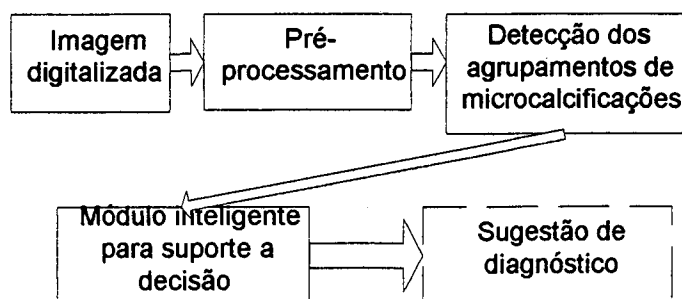


Figura 3.7: Um exemplo de CAD.

Como pôde ser visto na figura (Figura 3.7) acima, o esquema de um CAD para diagnóstico baseado em microcalcificações constituiu-se de quatro partes:



3.4.1. Aquisição da Imagem

Pode ocorrer de forma **direta**, pois alguns modernos equipamentos de mamografia já são capazes de produzir como resultado do exame uma imagem digital, ou **indireta**, nesse caso a aquisição pode ocorrer das seguintes formas: a) através da utilização de algum equipamento do tipo scanner, com o qual cria-se uma imagem digitalizada de uma mamografia comum, como fez Heang-Ping CHAN (1987, p.539) e outros; ou b) através da utilização de câmeras com capacidade de digitalizar a imagem fotografada ou filmada como Edward PATRICK (1991, p. 535), BETAL (1997, p. 903) e outros fizeram.

3.4.2. Pré-processamento

Lança-se mão de recursos da área processamento de imagens, tais recursos são aplicados com o objetivo de:

- 1) Eliminar de elementos espúrios que ocorrem na mamografia mas que não são microcalcificações (artefatos, fios de cabelo, cílios, impressões digitais, etc);
- 2) Realçar detalhes úteis que revelem microcalcificações.

3.4.3. Detecção de Calcificações

O objetivo da etapa de detecção de microcalcificações é obter uma imagem onde constem apenas microcalcificações, para tanto são também empregadas técnicas de processamento de imagens.

3.4.4. Sugestão de Diagnóstico

O módulo de sugestão de diagnóstico baseia-se em algum paradigma de inteligência artificial. Devido a diversidade de fatores que são considerados na definição do diagnóstico, bem como as possíveis correlações entre os mesmos, dotar o computador de inteligência o bastante para que esse efetivamente forneça um apoio confiável ao diagnóstico do câncer de mama, tem sido bastante desafiador. Apesar disso há muitas propostas baseadas principalmente no paradigma conexionista das redes neurais. Em tais sistemas uma rede neural artificial é treinada para identificar padrões de agrupamentos ou calcificação isoladas que caracterizem, por exemplo, malignidade.



3.5. Tecnologias para a Análise Computacional de Imagens Médicas

Para encontrar os achados patológicos individuais - microcalcificações - com base nos dados de imagem, pode-se lançar mão de uma série de técnicas da área de processamento de imagens. Tais técnicas pertencem basicamente a duas grandes áreas, na verdade duas formas distintas porém intercambiáveis, de abordar o processamento de imagens.

3.5.1. Domínio das Frequências

Por esta abstração não se faz referência a um ponto no espaço, como é natural para a percepção humana, mas sim à sua representação como sinal. O pressuposto é que as cores do espectro podem ser representadas por suas frequências de vibração, portanto uma imagem pode ser transformada em uma forma equivalente, composta de sinais caracterizados por suas amplitudes e fases. Dessa forma uma imagem convertida do domínio espacial ou Euclidiano para o domínio de frequências pode ser trabalhada através da utilização de filtros de frequências, filtros como o passa alta, passa baixa, passa faixa, etc. Após o devido tratamento, a imagem resultante pode ser reconvertida ao domínio espacial. Uma forma de conversão para o domínio de frequências é a aplicação da série de Fourier (sinais periódicos) e da transformada de Fourier (sinais com frequências arbitrárias). Existem também a transformada dos cossenos, a transformada de wavelets e outras.

3.5.2. Domínio Espacial

Relativo às coordenadas nos espaços n-dimensionais, onde os pontos pertencentes a certa dimensão n são representados por respectivas n-uplas, essa abstração torna-se útil na medida em que os *pixels*¹ de uma imagem podem ser tratados como pontos com pares de coordenadas, como pode-se ver na figura abaixo (Figura 3.10). Formalmente uma imagem no domínio espacial é considerada como um espaço bidimensional no qual há uma função que associa cada posição (*pixel*) da imagem a um mapa de cores ou tons de cinza:

$$f: U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^1$$

1. O termo pixel é um oriundo do inglês picture element, trata-se do menor fragmento de uma imagem (Figura 3.10).



Deve-se destacar que é comum em processamento de imagens no domínio espacial, definir-se a origem dos eixos coordenados como o canto superior esquerdo da imagem.

3.6. Conceitos Básicos

A essa altura faz-se necessária a colocação de alguns conceitos elementares no campo do processamento de imagens no domínio espacial. Reiterando-se que neste trabalho discutem-se e aplicam-se os métodos de processamento de imagens em tons de cinza.

3.6.1. Elemento Estrutural – *Kernel*

É o elemento chave em várias operações de processamento de imagens, consiste (usualmente) em uma forma plana matricial de dimensão quadrada e ímpar, em certas circunstâncias porém, pode-se lançar mão de *kernels* circulares. Um *kernel* ao operar sobre uma imagem original – O , deve percorrê-la no todo de forma que a combinação dos valores dos seus elementos, que representam valores de *pixels*, com os *pixels* da imagem origem O , através de certas operações gerem uma nova configuração de valores de *pixels* os quais (usualmente) são reproduzidos sobre os elementos de uma imagem destino- D .

3.6.2. Convolução Discreta

Processo pelo qual uma imagem original sofre uma operação (transformação) gerando-se uma imagem resultante transformada, porém de mesmas dimensões que as da imagem original. Neste processo a operação é efetuada a cada sub-região da imagem original delimitada por um elemento auxiliar, tal elemento possui dimensões bem menores que as da imagem original e pode ter várias designações, tais como: máscara de convolução, janela, elemento estrutural ou *kernel*.

Formalmente, seja O uma imagem. Seja W um *kernel* que corresponde a uma



matriz de dimensão 3×3 , com pesos W_1, \dots, W_9 :

W_1	W_2	W_3
W_4	W_5	W_6
W_7	W_8	W_9

E seja O' representada por uma matriz de pontos definida pela "sombra" de W sobre O . Portanto, O'_1, \dots, O'_9 são valores de *pixels* na imagem O :

O'_1	O'_2	O'_3
O'_4	O'_5	O'_6
O'_7	O'_8	O'_9

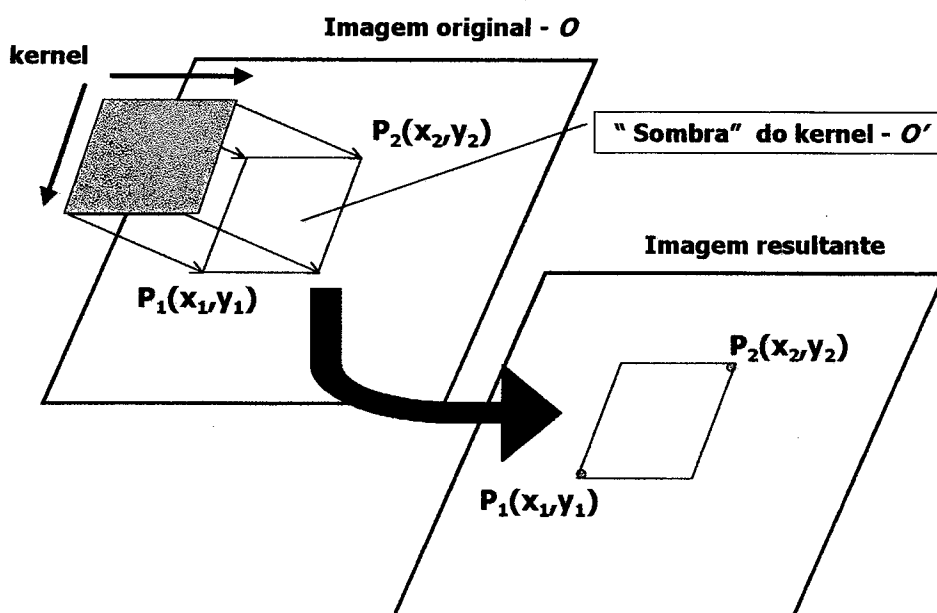


Figura 3.8: Convolução de uma imagem por um kernel resultando em outra imagem

O *kernel* percorrerá a imagem O no todo, deslocando-se a partir da sua origem, saltando linha a linha e coluna a coluna. A cada nova posição do *kernel* sobre a imagem original será efetuada a operação adequada (Figura 3.8). A operação pode consistir por exemplo em substituir cada *pixel* na imagem destino D , pelo valor $D(l,c)$ abaixo:

$$D(l,c) = (O' * W)(l,c) = O'(l,c) * W(l,c)$$



Onde $W(l',c')$ corresponde ao peso de uma célula da matriz W , $O'(l,c)$ corresponde ao tom de cinza ou cor do *pixel* correspondente à posição (l,c) em O' e $D(l,c)$ é o resultado do produto de cada *pixel* em O' pelo respectivo peso em W .

3.6.3. Operação *Threshold* (limiar)

Uma imagem, não necessariamente em tons de cinza, mas especialmente nesse mapa de cores, pode ser vista como uma superfície topográfica (Figura 3.9). Nesta forma de abstração, quanto mais branco o *pixel* mais alta é a sua representação topográfica, portanto os cumes das elevações equivalem aos *pixels* mais brilhantes, enquanto os mais escuros ocupam posições nos vales e depressões.

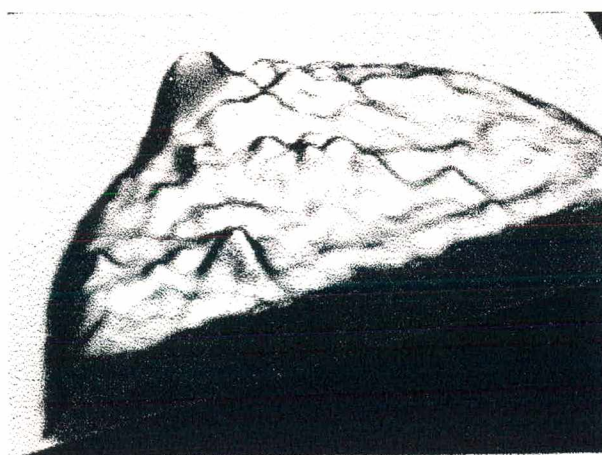


Figura 3.9: Equivalente topográfico de uma imagem. HIGHNAM (1996, p.4)

O *threshold* aplicado a uma imagem O , separa *pixels* cujos valores são menores ou iguais a um certo limiar, de *pixels* com valores superiores ao mesmo limiar, gerando uma segunda imagem binária D . Usualmente, os *pixels* superiores são reproduzidos em D com o valor máximo da escala utilizada (correspondente ao branco), enquanto os *pixels* que não conseguem “vencer” o limiar são reproduzidos em D com o valor mais baixo da escala (correspondente ao tom mais escuro), esse tipo de operação é chamada de *threshold* binário pois o resultado obtido é uma imagem com duas informações de cor, no caso do exemplo anterior, uma imagem em preto e branco. Não obstante, pode-se operar de forma diferente, por exemplo, preservando inalterados os *pixels* que vecem o limiar enquanto força-se os demais a receber o



valor de fundo da imagem.

Na analogia topográfica é como se houvesse uma grande inundação que em função do nível de precipitação das chuvas seria capaz de alagar vales e pequenas elevações, porém preservar montanhas.

Formalmente, seja O a imagem a sofrer *threshold*, D a imagem resultante da operação e L o limiar aplicado:

$$\forall (l,c) \in O, \begin{cases} \text{se } O(l,c) \leq L \Rightarrow D(l,c) = 0 \\ \text{se } O(l,c) > L \Rightarrow D(l,c) = 1 \text{ (i.e. máximo da escala)} \end{cases}$$

Onde zero corresponde ao valor mínimo na escala de cinza e 1 ao seu valor máximo.

Deve-se destacar que a operação de *threshold* usualmente prescinde de *kernel*.

3.6.4. Segmentação

Processo de divisão de uma imagem em segmentos bem definidos, de forma a isolar-se os componentes de interesse dentro da imagem, facilitando a sua análise. De fato o processo de *threshold* já implica em alguma segmentação de imagens, no entanto a área de segmentação abrange vários outros aspectos tais como: a detecção de linhas, pontos ou fronteiras de uma região.

3.7. Morfologia matemática

A área de conhecimento conhecida como Morfologia Matemática foi criada na França, mais precisamente em Fontainebleau na École Supérieure de Mines de Paris, tem como objetivo a análise das estruturas geométricas presentes em uma imagem, fundamenta-se na teoria de conjuntos e tem larga aplicação no processamento de imagens. A idéia por trás da Morfologia Matemática é que imagens são conjuntos de os *pixels* que formam agrupamentos em um espaço bidimensional portanto no domínio espacial (Figura 3.10).



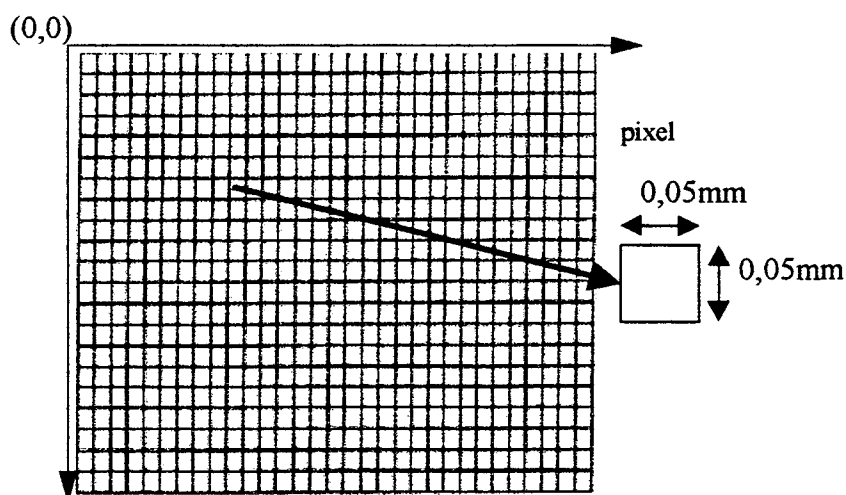


Figura 3.10: Pixels de uma imagem com 500 dpi de resolução

As operações básicas na Morfologia Matemática são a erosão e a dilatação, por aplicação dessas operações, *pixels* que se alinham com um determinado padrão sofrem alterações na cor (ou tom de cinza) que apresentam. Um padrão é definido em um elemento estrutural, este consiste de um subconjunto $B(x,y)$ do plano que interage com a imagem do objeto de forma a determinar propriedades topológicas e revelar informações estruturais do objeto (GOMES, 1994, p. 183).

Matematicamente, dado um conjunto B do plano com origem (base) em um ponto (x,y) e um subconjunto X do plano, as operações de erosão e dilatação são definidas da seguinte forma:

Erosão de X por um elemento estrutural $B(x,y)$:

$$X \ominus B(x,y) = \{(x,y) \mid B(x,y) \subset X\}$$

Dilatação de X por um elemento $B(x,y)$:



$$X \oplus B(x,y) = \{(x_1,y_1) = ((x_2+x_3),(y_2+y_3)) \mid (x_2,y_2) \in X \text{ e } (x_3,y_3) \in B \}$$

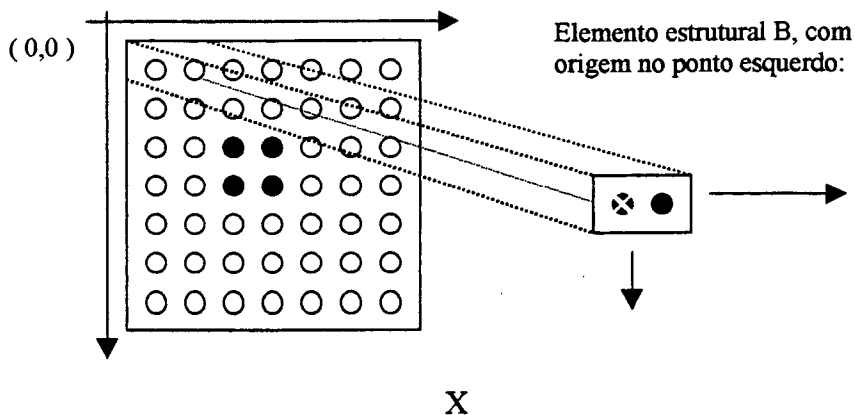
OU

$$X \oplus B(x,y) = C = A + B.$$

Exemplo de erosão e dilatação adaptado de PARKER (1997, p. 70-1):

Supondo um $X = \{(3,3);(3,4);(4,3);(4,4)\}$ e um elemento estrutural

$$B = \{(0,0);(0,1)\}$$



Aplicando-se B sobre o espaço, ao encontrar os pontos de X, de forma que a origem de B coincida com algum ponto deles, tem-se a seguinte erosão:

$$B(3,3) = \{(3,3), (4,3)\}; B(3,3) \subset X$$

Se (3,3) está contido em X então ele faz parte da erosão de X por B(x,y), o mesmo ocorre para B(3,4).

$$B(4,3) = \{(4,3), (5,3)\}; B(4,3) \not\subset X$$

Nesse caso o *pixel* no ponto de coordenada (4,3) não pertence ao conjunto erosão de X por B(x,y), o mesmo ocorre para B(4,4). Desta forma o conjunto erosão de X por B(x,y) é igual a $\{(3,3),(4,3)\}$, visualmente o resultado está na figura abaixo (Figura 3.11).

Para os mesmos conjuntos X e B, temos a seguinte dilatação, cujos resultados são apresentados na figura abaixo (Figura 3.11):



$B(3,3)$: $c_1 = ((3,3) + (0,0))$ e $c_2 = ((3,3) + (1,0)) \Rightarrow c_1 \cup c_2 = \{(3,3);(4,3)\}$, como $(c_1 \cup c_2) \subset (A + B)$,
então como $(c_1 \cup c_2) \subset X \oplus B$.

$B(3,4)$: $c_3 = ((3,4) + (0,0))$ e $c_4 = ((3,4) + (1,0)) \Rightarrow c_3 \cup c_4 = \{(3,4);(4,4)\}$, como $(c_3 \cup c_4) \subset (A + B)$,
então como $(c_3 \cup c_4) \subset X \oplus B$.

$B(4,3)$: $c_5 = ((4,3) + (0,0))$ e $c_6 = ((4,3) + (1,0)) \Rightarrow c_5 \cup c_6 = \{(4,3);(5,3)\}$, como $(c_5 \cup c_6) \subset (A + B)$,
então como $(c_5 \cup c_6) \subset X \oplus B$.

$B(4,4)$: $c_7 = ((4,4) + (0,0))$ e $c_8 = ((4,4) + (1,0)) \Rightarrow c_7 \cup c_8 = \{(4,4);(5,4)\}$, como $(c_7 \cup c_8) \subset (A + B)$,
então como $(c_7 \cup c_8) \subset X \oplus B$.

Desta forma $X \oplus B(x,y) = \{(3,3);(4,3);(3,4);(4,4);(5,3);(5,4)\}$

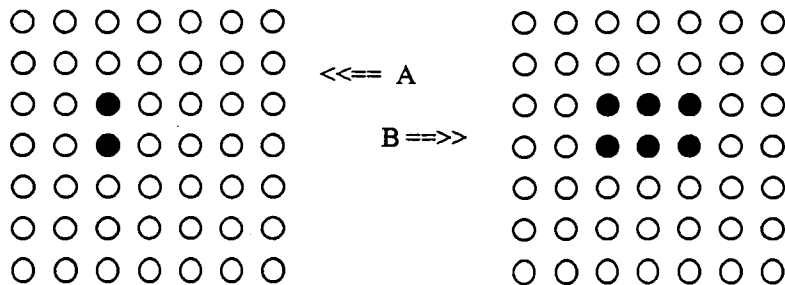


Figura 3.11: Resultado da operação de erosão (figura A) e dilatação (figura B) de X por B.

A erosão e a dilatação podem ser aplicadas através de outras transformações lineares, além disso pode-se combinar erosões e dilatações repetidamente dando origem a outros tipos de operações morfológicas tais como (PARKER, 1997, p. 84-5):

Opening: consiste em aplicar-se uma erosão seguida de dilatação, utilizando-se o mesmo elemento estrutural, a operação opening tem como objetivos principais: a) separar objetos muito próximos em uma imagem, ou seja, criar espaços (aberturas) entre objetos na imagem; b) eliminar ruídos (pixels negros aleatoriamente espalhados em toda a imagem).

Closing: consiste em aplicar-se uma dilatação seguida por erosão, utili-



zando-se o mesmo elemento estrutural, a operação closing tem como objetivos principais: a) eliminar espaços entre objetos em uma imagem, ou seja, criar espaços (aberturas) entre objetos na imagem; b) eliminar falhas dentro dos objetos da imagem (pixels brancos em um objeto negro, por exemplo); c) suavizar as bordas de um objeto na imagem.

Exemplos de closing e opening podem ser vistos na figura abaixo (Figura 3.12).

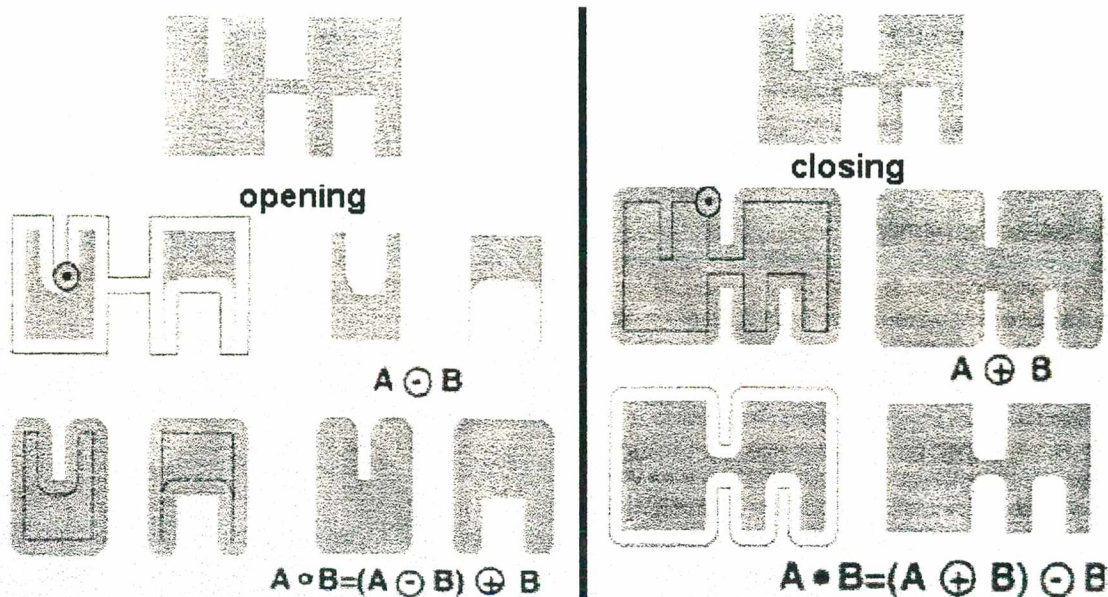


Figura 3.12: Exemplos de operações de opening e closing. retirado de FILHO & NETO (1999, p. 191-2)

3.7.1. Erosões e Dilatações sobre tons de cinza

As definições feitas acima são genéricas, ou seja, as operações que caracterizam uma erosão ou dilatação podem ser diversas desde que se atinja o objetivo de erodir ou dilatar, dessa forma, adotando-se outras operações, a erosão e dilatação podem ser extendidas à imagens em tons de cinza, imagens coloridas, etc. Por essa razão foi possível utilizar os a morfologia matemática neste trabalho que trata do processamento de imagens em tons de cinza com profundidade de 8 bits, ou seja, cujo mapa de tons corresponde a valores entre 0 (preto) e 255 (branco).

Uma erosão em tons de cinza consiste no mesmo princípio da erosão comum, exceto pela operação realizada na região definida pela vizinhança do *kernel* sobre a imagem a ser erodida. Nesse tipo de erosão a operação realizada é definida da seguinte forma: *cada pixel*



da imagem resultante recebe o valor do pixel central do kernel que convolve a imagem a ser erodida, esse valor de pixel, por sua vez, corresponde ao menor tom de cinza na vizinhança definida pelo kernel. Veja exemplo na figura abaixo (Figura 3.13).

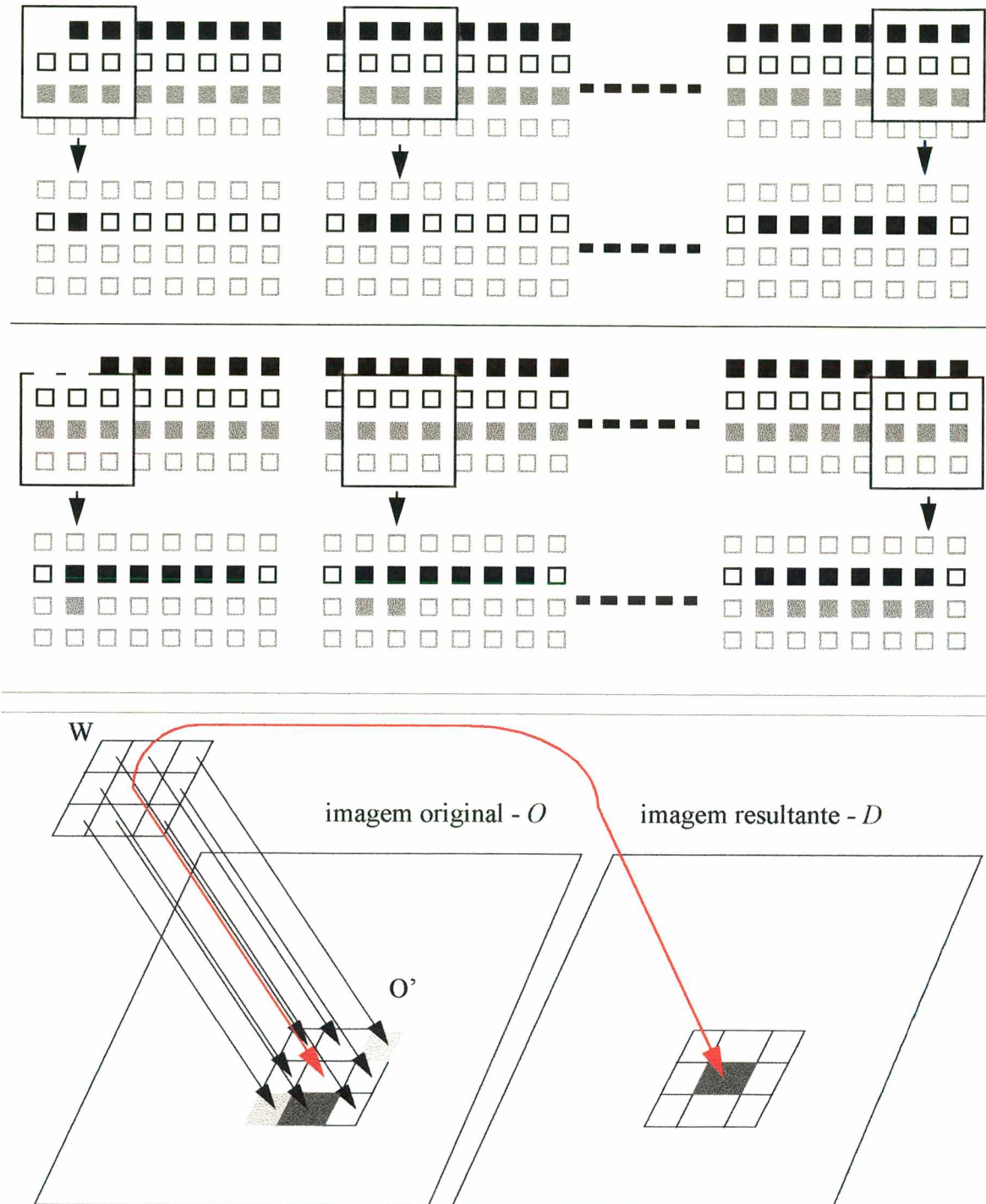


Figura 3.13: Erosão em tons de cinza, o quadrado representa um kernel que erode os pixels de uma imagem, os pixels das filas inferiores representam a imagem resultante.

Em uma dilatação em tons de cinza, modifica-se a operação para a seguinte:



cada pixel da imagem resultante recebe o valor do pixel central do kernel que convolve a imagem a ser erodida, esse valor de pixel, por sua vez, corresponde ao maior tom de cinza na vizinhança definida pelo kernel.

Os códigos em C para dilatação e erosão em tons de cinza são apresentados nos anexos a este trabalho.

3.8. Estado da Arte em CADs

Neste item faz-se um relato do estado da arte em relação a aspectos de sistemas da classe CAD. Para tanto foram selecionados sete textos que resumem as propostas atuais a respeito.

3) *Computer-aided detection of clustered microcalcifications: An improved method for grouping detected signals,*

Neste artigo, Robert M. NISHIKAWA (1993, p. 1661-6) e outros discutem um esquema para detecção automática de agrupamentos de microcalcificações a partir de imagens de mamografias, esse esquema é parte de um pacote CAD.

Técnicas empregadas para pré-processamento das imagens de mamografias com vistas a detecção de microcalcificações:

Um filtro linear (*kernel*) é empregado para reforçar o sinal da microcalcificação enquanto suprime a estrutura de fundo da mama. Um limiar (*threshold*) global de tons de cinza é aplicado com o objetivo de extrair possíveis microcalcificações. Aplica-se uma erosão morfológica para eliminar pequenos sinais causados por ruídos na imagem. O *threshold* local de tons de cinza é então aplicado para reduzir mais ainda o número de sinais falsos. Após isso aplica-se uma análise de textura (segmentação).

A aplicação do limiar local baseia-se em estatísticas tais como a média e desvio padrão locais dos valores de pixels (tons de cinza) em uma área de 5,1 X 5,1 mm na imagem. O nível do limiar = $K(\text{média} + \text{desvio padrão})$, onde K é um valor constante entre 3,0 e 4,2.

Transformação de agrupamentos:



Visa a criação de uma imagem de saída com agrupamentos de microcalcificações. Submete-se a imagem a um *kernel* de tamanho predefinido, de forma que se três ou mais sinais se apresentarem neste, o conteúdo do *kernel*, *pixel* por *pixel*, é reproduzido exatamente no *pixel* correspondente na imagem de saída. Se apenas um ou dois pixels se apresentarem no *kernel*, todos os pixels, na imagem de saída, que corresponderem à atual localização do *kernel* (na imagem de entrada), serão valorados em zero. Sinais desagrupados não são passados para a imagem de saída. O *kernel* utilizado foi escolhido entre vários tamanhos em uma faixa de 20 a 40 pixels. Como fase final do processo de agrupamento, os grupos de sinais dentro de um perímetro de 50 pixels um do outro, foram considerados pertencentes a um mesmo agrupamento.

Como procedimento auxiliar, utilizou-se uma técnica chamada transformação recursiva de área em ponto (TAP), onde a cada locação de um *kernel* 3 X 3 na imagem, o número de pixels acima da cor de fundo é contado, se houver mais do que dois, todo o *kernel* é valorado com cor de fundo, exceto o *pixel* inferior direito, que recebe valor máximo, caso haja menos que dois pixels com valores maiores que a cor de fundo nada é alterado nos pixels da atual locação do *kernel*. O *kernel* usado no TAP desloca-se de 2 em 2 pixels para a direita e de 2 em 2 linhas para baixo. A mesma imagem serve como entrada e saída dos resultados da operação de TAP. O procedimento de TAP tem complexidade de tempo menor que a oferecida pela técnica alternativa, que determina quais pixels pertenciam a cada sinal e mantém um mapa de quais sinais e pixels foram contados.

4) *Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network*

Neste texto, Heang-Pin CHAN (1995, p. 1555-67) e outros têm como meta a detecção automática de agrupamentos de microcalcificações em mamografias, e uso de um tipo de rede neural (CNN¹) aplicada com o fim de melhorar a exatidão do programa de detecção.

As técnicas empregadas nas imagens de mamografias com vistas a detecção de

1. Convolution Neural Network



microcalcificações compõem-se de três principais partes:

Pré-processamento:

Um detector de contorno detecta as bordas da mama e divide a imagem em duas regiões, interna e externa à mama. Um filtro de reforço de sinal (*kernel* 1X1) é empregado para reforçar as microcalcificações, e outro para supressão de sinais (filtro caixa de borda¹ com um *kernel* 8X8 de pesos constantes ao redor da borda e uma área central 4X4 com pesos nulos), para remover ou suprimir as microcalcificações e amenizar o ruído.

Subtraindo-se as duas imagens resulta em uma imagem com a relação sinal ruído² reforçada, na qual a estrutura de baixa frequência do fundo da mama é removida e o ruído de alta frequência é suprimido³. Quando o filtro de reforço de sinal e o filtro supressor de sinal são ambos lineares, esta técnica tem a equivalência de um filtro passa banda do domínio das frequências.

Segmentação:

Determina-se o histograma de tons de cinza da imagem dentro da região da mama. Uma técnica de limiar de tons de cinza é então usada para detectar locais de possíveis sinais acima de um limiar global. O limiar é mudado iterativamente até o número desses locais obtidos cair dentro de um intervalo com um máximo (4000) e um mínimo (3000).

Em cada local de possível sinal, um limiar de tons de cinza localmente adaptativo em combinação com uma técnica de crescimento de região, sendo aplicado para determinar o número de pixels conectados que estejam acima de um limiar local. Esse por sua vez, é calculado como o produto da raiz quadrada da média⁴ e um valor de limiar da relação sinal ruído. As características dos sinais a serem usados na etapa de classificação, tais como o tamanho, contraste máximo, relação sinal ruído e localização, são obtidas neste passo.

A técnica de aplicação de limiar local adaptativo é similar a extração de características do sinal descrita no passo anterior, exceto que o procedimento é feito na imagem

-
1. Box-rim filter
 2. SNR – Signal-to-Noise Ratio
 3. Técnica da diferença de imagem
 4. RMS – root mean square



que teve reforço da relação sinal ruído, ao invés de ser feito na imagem não processada.

Classificação:

Executam-se três testes para distinguir sinais, ruídos e artefatos, um limite inferior (dois pixels) é usado como referência de tamanho para excluir sinais menores que este e prováveis de serem ruído, um limite superior (80 pixels) é usado para excluir sinais maiores que este e que são prováveis de serem grandes calcificações benignas.

Um limite superior de contraste é também iniciado para excluir sinais potenciais que tenham um contraste mais alto do que um certo número (10), de pixels de desvios padrão¹, acima do contraste médio de todos os sinais potenciais encontrados com o limiar local. Esse critério exclui os sinais com contraste muito alto, que são prováveis artefatos e grandes calcificações benignas.

Um procedimento de agrupamento² por região é então aplicado para os sinais remanescentes, um sinal é mantido se o número de sinais encontrados dentro de uma vizinhança com um diâmetro fechado (1 cm) ao redor do sinal, é maior do que um certo valor mínimo. Os sinais remanescentes que não são encontrados na vizinhança de qualquer agrupamento potencial, será considerado ponto de ruído isolado ou calcificação e será excluído. Esse critério de agrupamento é útil para reduzir “falso positivos”, porque microcalcificações de interesse clínico sempre aparecem em agrupamentos.

Foi desenvolvido e aplicada uma técnicas baseada em redes neurais artificiais (CNN) para melhor acurar o processo de discriminação entre sinais falsos e verdadeiras microcalcificações.

Usam-se como parâmetros para alguns procedimentos: o RMS, SNR, média, desvio padrão. Os parâmetros específicos usados em cada passo podem ser encontrados nas referências Heang-Ping CHAN (1990, p. 1102-10),(1988, p. 664-71) e (1994, p. 1203-11).

1. SD – standard deviation
2. clustering



5) Malignant and Benign Clustered Microcalcifications: Automated Feature Analysis and Classification

Nesta referência, Yulei JIANG (1996, p. 671-8) e outros, expuseram o desenvolvimento de um método que extrai de uma imagem de mamografia, as características dos agrupamentos de microcalcificações, essas são então correlacionadas com aquelas que um radiologista humano leva em consideração para diagnóstico de agrupamentos como benignos ou malignos, produzindo com isso uma estimativa da probabilidade de malignidade.

Técnicas empregadas para a detecção de microcalcificações:

Oito características relacionadas a microcalcificações individuais e agrupamentos de microcalcificações foram extraídas de mamografias digitalizadas, essas características foram selecionadas com base na sua correlação qualitativa com a experiência dos radiologistas, e aparente utilidade para diagnóstico.

Utilizam-se técnicas como a segmentação de microcalcificações individuais através do uso de uma superfície polinomial de terceiro grau ajustada a uma região de interesse na mamografia digitalizada, bem como o cálculo do *thickness*¹, estimado pelo contraste do sinal (valor médio do *pixel* acima do fundo) da microcalcificação isolada.

6) *Computer-aided detection of clustered microcalcifications on digital mammograms*

Novamente R.M. NISHIKAWA (1995, p. 174-8) e outros, desta vez discutindo o desenvolvimento de um esquema CAD para auxílio a radiologistas no diagnóstico de agrupamentos de microcalcificações em mamografias.

Técnicas empregadas para pré-processamento das imagens de mamografias com vistas a detecção de microcalcificações:

Foram propostos três passos em sequência para detecção das microcalcificações:

Técnica da diferença de imagem:

Usada para reforçar os sinais das microcalcificações e suprimir a estrutura de

1. dimensão física ao longo de uma linha de projeção do raio X



fundo da mama, para isso usam-se dois filtros lineares (*kernels*), sendo a tarefa de supressão de sinais confiada a um filtro de borda¹. Após a aplicação desses filtros subtrai-se a imagem suprimida da imagem reforçada, obtendo-se a imagem diferença. Na prática os dois filtros são combinados em um único e um off-set de 512 é adicionado a cada *pixel* para evitar valores negativos.

Técnica da extração de sinal:

Procede-se uma operação de limiar global, baseado no histograma de tons de cinza da imagem total, de forma que 98% dos *pixels* na imagem diferença recebam valores de fundo (512 devido ao off-set). Procede-se uma erosão morfológica por meio de *kernels* com formatos em 'L' de 3 *pixels* e cadeias também de 3 *pixels*, essa operação elimina a presença de ruído aleatório. Finalmente, um limiar local de tons de cinza adaptativo é aplicado, baseado em estatísticas locais (média e desvio padrão) de valores de tons de cinza dentro de uma área de 5,1X5,1 mm na imagem processada.

Técnica de extração de características:

Consiste de dois passos: análise de textura e agrupamento de sinais. Análise de textura: uma região de 6,4X6,4 mm centrada em cada sinal que permanece na imagem após a etapa 2, é analisada. Inicialmente uma correção de fundo de baixa frequência com um polinômio de grau 3 é aplicada, então calcula-se um espectro de energia Fourier bidimensional, do qual calcula-se a média com a premissa de rotação simétrica, então o primeiro momento da média do espectro de força é calculado, sendo maior do que 3,00 ciclos mm⁻¹, o sinal é eliminado.

Agrupamento dos sinais:

Os sinais remanescentes do passo anterior são agrupados pela passagem de um quadro de dimensões 3,2X3,2 mm sobre a imagem, havendo 3 ou mais sinais no quadro, esses são repassados para a imagem de saída.

Maiores detalhes podem ser obtidos em:

1. box-rim filter



Katsuragawa, S., Doi, K. And MacMahon, H. (1988): "Image feature analysis and computer-aided diagnosis in digital radiography: Detection and characterization of interstitial lung disease in digital chest radiographs", *Med Phys.*, 15, pp 311-319.

Nishikawa, R.M., Doi, K., Giger, M.L., Vybomy, C.J., Schmidt, R.A. (1990b): "Automated detection of microcalcifications in mammograms: New feature-extraction technique using morphological filters", *Radiol.*, 177(P), p. 288 (abstract)

7) Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography

Neste artigo, Heang-Ping CHAN (1987, p. 538-48) e outros mostram o desenvolvimento de métodos baseados em computador para detecção de microcalcificações em mamografias digitais.

Técnicas empregadas para detecção de microcalcificações:

Filtragem através de reforço de sinal: foi empregado um filtro espacial que aproximadamente casou com o tamanho e variações de contraste de uma microcalcificação típica. Para esse filtro não foi levado em consideração o conteúdo de frequência do ruído correlacionado ao fundo da imagem, foi usado um *kernel* $n \times n$ com n ímpar. Baseado na análise de microcalcificações típicas, a variação do contraste das microcalcificações foi aproximada com os seguintes pesos para o *kernel* do filtro: quando $n=3$, os oito pesos mais externos (no *kernel*) terão peso 0,75 e o peso central será 1,00. Quando $n > 3$, os pesos mais externos serão 0,5 os segundos mais externos serão 0,75 e todos os outros serão 1,00.

Filtros para supressão de sinal (foram empregados dois tipos de filtro):

I. Filtro de mediana: filtro espacial não linear que substitui o valor de um dado *pixel* pelo valor do *pixel* de mediana dentro de um *kernel* centrado no *pixel* de interesse. Um filtro de mediana com tamanho e forma bem escolhidos pode virtualmente eliminar microcalcificações além de amortecer o ruído sem mudar o fundo da imagem de forma global. Usou-se aqui um *kernel* quadrado para os filtros de mediana.

II. Para reforçar a relação sinal ruído, enquanto remove a estrutura de fundo na imagem diferença, foi analisada uma segunda abordagem alternativa para supressão de sinal:



Um novo filtro chamado filtro de contraste reverso foi desenvolvido para reverter de forma seletiva o contraste das microcalcificações enquanto mantém a estrutura de baixa frequência do fundo e suprime o ruído de alta frequência. Essas funções podem ser levadas a termo por um filtro tendo fatores de transferência óticos¹ da forma : $F(u) = [2Fa(u)-1] Fb(u)$, onde $Fa(u)$ é um filtro de atenuação passa baixa e $Fb(u)$ é um filtro que suprime ruídos de alta frequência. Foram empregados filtros A e B com *kernels* quadrados no domínio espacial, o filtro A teve pesos similares a aqueles do filtro de casamento usados para reforçar sinais, o filtro B teve pesos uniformes independente do seu tamanho de *kernel*.

III. Diferença de Imagens:

É obtida pela subtração: (imagem_reforçada - imagem_suprimida), desde que o fundo é basicamente o mesmo nas duas imagens, esse componente será removido da imagem diferença. Para compensar a perda do nível de fundo, um valor constante de 512, é adicionado a todo *pixel* da imagem diferença para propósitos de exibição da imagem.

Limiar de tons de cinza

Com a finalidade de isolar microcalcificações do fundo remanescente, foram estudados dois tipos de limiar de tons de cinza:

A. Limiar global.

B. Limiar local: estatísticas locais são determinadas dentro de um *kernel* quadrado centrado no *pixel* de interesse; o valor do *pixel* é preservado apenas se ele é maior do que o valor do *pixel* médio subordinado a um múltiplo do desvio padrão. O tamanho do *kernel* deve ser suficientemente grande para dar uma boa estimativa da flutuação do ruído de fundo local e foi escolhido com as dimensões 51X51 *pixels*. A aplicação de um limiar de tons de cinza resulta em uma imagem na qual, grupos de *pixels* com valores acima do limiar são superpostos por um fundo uniforme.

Agrupamento de sinais

Detecta-se a região da mama na imagem não processada. Processa-se uma busca por sinais na região da mama, que passou por um filtro de limiar para determinar a localização, área e contraste de cada ponto. Um critério de limiar por área é executado nos sinais

1. OTF - optical transfer factor



detectados para eliminar pontos que tenham uma área menor do que um número pré-selecionado de *pixels*. Por fim, um critério de agrupamento é usado para identificar agrupamentos os quais contenham mais do que um número pre-selecionado de sinais dentro de uma região de certo diâmetro. Esses valores pre-selecionados foram determinados pela experiência dos radiologistas e empiricamente pelo processamento de um número de mamografias de teste com microcalcificações já conhecidas. O valor de limiar de área foi 2 ou 3 *pixels*, e um agrupamento foi deveria conter 3 ou mais sinais dentro de uma região de 1 a 1,5 cm de diâmetro. O último critério está em conformidade com a experiência clínica a qual requer que um mínimo de microcalcificações, usualmente entre 3 e 5, devem estar presentes antes do agrupamento ser considerado suspeito o bastante para ser necessária uma biópsia.

8) *Computer-Aided Detection of Microcalcifications in Mammograms. Methodology and Preliminary Clinical Study*

Neste Artigo, Heang-Ping CHAN (1988, p. 664-71) e outros, discutiram a viabilidade de métodos automáticos baseados em computadores para detecção de agrupamentos de microcalcificações em mamografias.

Técnicas empregadas para a detecção de microcalcificações:

Utiliza-se um filtro de borda¹ para supressão de sinais, este consiste basicamente em um filtro que calcula a média² de *pixels*, no qual os pesos na região central são nulos, excluindo-os do calculo da média. A aplicação deste filtro na imagem resulta na substituição dos valores atuais dos sinais, pela média dos valores dos *pixels* de fundo que o cercam, suavizando os ruídos da imagem. A seleção do tamanhos do *kernel* e da região central do filtro de borda é governado pelo tamanho do sinal a ser removido.

Utiliza-se de um filtro casado³ de tamanho 3X3. Uma imagem diferença é obtida diretamente pela utilização de uma combinação do filtro de borda e do filtro casado em um único filtro chamado filtro diferença.

Após a filtragem espacial, um limiar local adaptativo de tons de cinza é empre-

-
1. box-rim filter
 2. averaging filter
 3. matched filter



gado para extração de sinais da região da mama, para tanto, inicialmente aplica-se um limiar global de tons de cinza na imagem diferença, então o computador determina iterativamente o limite de tom de cinza, o qual situa aproximadamente 3.000 locais de sinais em potencial. O limiar local é então aplicado em cada um desses locais na imagem diferença ou na imagem não processada, no segundo caso, a variação de densidade do fundo em uma pequena região (51X51 *pixels*) centrada no sinal, é estimada por uma técnica de superfície de ajuste polinomial de duas dimensões, permitindo a atenuação do fundo. Após isso o valor do *pixel* médio e a raiz quadrada da média da flutuação de ruído do fundo local, são determinados em uma região de 51X51 *pixels*, excluindo o *pixel* central que pode conter uma sinal. Usando um limiar de relação sinal ruído pré-selecionado, definido com a razão do limiar de contraste do sinal e a raiz quadrada da média do ruído, o computador calcula o limiar local de tons de cinza como a soma do *pixel* médio do fundo local e o produto do limiar da relação sinal ruído e a raiz quadrada da média do ruído local. Uma técnica de crescimento de região é empregada para determinar os *pixels* conectados que têm valor de *pixel* acima do limiar local. O tamanho de um sinal detectado como o número desses *pixels* conectados, e o contraste é determinado como a diferença entre o valor do *pixel* máximo do sinal detectado e a média dos valores de *pixel* de fundo.

Crítérios de extração de sinal são então invocados, de forma que sinais verdadeiros possam ser diferenciados de ruídos e artefatos, primeiro, um limiar de área é aplicado eliminando todos os sinais detectados que tenham uma área menor do que um número pré-selecionado de *pixels*. Segundo, um limite superior de área e um limite superior de contraste são usados para excluir sinais com grandes áreas e altos contrastes, que podem ser causados por artefatos ou por grandes calcificações benignas.

Agrupamento de sinais

Para localização de agrupamentos que incluem um número mínimo pré-selecionado de sinais detectados, dentro de uma área com diâmetro pré-determinado (3 sinais detectados em um diâmetro de 1,2 cm), os procedimentos de detecção de grupos são realizados.

Em Segmentation and numerical analysis of microcalcifications on mammograms using mathematical morphology, os autores, D. BETAL (1997, p. 903-17) discutem



acerca da utilização dos algoritmos top-hat e watershed, da área de morfologia matemática, como ferramentas para detecção e segmentação automática de microcalcificações em mamografias digitalizadas.

Técnicas empregadas para pré-processamento das imagens de mamografias com vistas a detecção de microcalcificações

Detecção das microcalcificações:

Algoritmo top-hat¹: tomando-se um imagem 2D em tons de cinza como sendo uma superfície topográfica, onde as variações do domínio espacial se apresentam como um relevo em 3D. Uma cartola será colocada em cada parte da superfície, quando ocorrerem altos níveis de contraste (relevo muito alto) o topo da cartola será "rompido" e portanto detectados por esta técnica. Variando-se a altura e/ou largura da cartola pode-se identificar importantes características na imagem.

O *kernel* utilizado pode ser circular ou quadrado.

A aplicação do algoritmo da cartola consiste de dois passos:

Um opening é realizado, durante sua fase de erosão o *pixel* no centro do *kernel* recebe o menor valor da vizinhança de *pixels*. Na etapa de dilatação o *pixel* central do *kernel* é valorado com o máximo da vizinhança. Isso consiste em um opening usando algoritmo de reconstrução.

A segunda parte do algoritmo da cartola consiste em subtrair da imagem original a obtida no passo 1.

Segmentação das microcalcificações: Algoritmo da watershed²: esse algoritmo pode ser usado para segmentar características importantes em uma imagem baseada em tons de cinza, voltando à analogia do intensidade de sinais com um relevo topográfico, as fronteiras dos objetos podem ser representadas pelas bordas de bacias (hidrográficas) obtidas pelo "alagamento" de vales, que na imagem em tons de cinza correspondem a intervalos desde os valores mínimos de intensidade até um certo valor maior. Invertendo-se as intensidades dos sinais da imagem original, cada microcalcificação será representada por um mínimo local e então pode potencialmente ser segmentada pelo algoritmo watershed.

1. Em inglês cartola
2. Do inglês Watershed ou divisão de água.



Passos do algoritmo watershed (mais simples):

Localizar todos os mínimos locais dentro da imagem e rotulá-los.

Todos os *pixels* com valor na escala de cinza superior ao mínimo, são identificados e aqueles que sejam adjacentes ao mínimo local serão rotulados como pertencentes à bacia a qual corresponde o mínimo local. Aquelles que não são adjacentes, são colocados em uma fila e apropriadamente rotulados. Esse processo de inundação¹ é repetido sucessivamente para sucessivos valores na escala de tons até o máximo da intensidade do sinal na imagem inclusive.

Linhas divisórias são construídas nos pontos onde as bacias tendem a juntar-se.

Quando um único mínimo local identifica cada objeto de interesse, as linhas divisórias representam as fronteiras desses objetos. Esse é o watershed mais simples. Como ha sempre a possibilidade de ocorrência de mais mínimos do que características de interesse na imagem, essa sempre se apresenta super segmentada após um watershed convencional, para melhorar essa técnica dois incrementos foram feitos:

Foi aplicado o operador morfológico gradiente², para destacar as fronteiras das microcalcificações, o que significa uma superfície ideal para aplicar-se o watershed.

Para evitar a inundação de objetos que não correspondem a microcalcificações, a operação do algoritmo watershed foi restringida por uma função de apropriada. Um marcador interno único necessita ser identificado para cada microcalcificação, junto com um marcador externo. A inundação procede como descrito antes, porém dessa vez as bacias formadas, que não correspondam aos marcadores são inundadas pelo transbordamento das suas bacias vizinhas. As marcas internas são providas pelo algoritmo cartola, as marcas externas foram obtidas pela aplicação do algoritmo watershed sobre para a imagem invertida restringido pelos marcadores internos.

Após essas operações o watershed é aplicado para a imagem de gradiente morfológico restringido pelos marcadores internos e externos, as fronteiras das microcalcificações

1. Do inglês flooding

2. A imagem resultante desse operador corresponde a aplicar-se uma dilatação a uma imagem original, depois uma dilatação na mesma imagem original (ou vice versa) e encontrar-se a diferença entre a imagem obtida pela dilatação e a obtida pela erosão.



são então encontradas.

9) Expert Learning System Network for Diagnosis of Breast Calcifications

Edward A. PATRICK (1991, p. 534-9) e outros enfocam o desenvolvimento e avaliação de um sistema especialista para apoio a diagnóstico, pouco destaque é dado ao tratamento de imagens de mamografias.

Usou-se segmentação para extrair as calcificações e parâmetros básicos de geometria, chamados medidas de características relativos a área estatística de Reconhecimento de Padrões. Tais medidas incluem centro de gravidade, comprimento, largura, orientação, área, perímetro, momento, e perímetro convexo. As características são construídas a partir dessas medidas.



CAPÍTULO 4 - Implementação

Os estudos que substanciaram os capítulos anteriores bem como o contato com profissionais da área médica, serviram de subsídio para a implementação de um protótipo inicial de CAD, tal implementação é uma proposta básica e tem como objetivo servir como referência para trabalhos futuros, os quais devem estende-la com novas funções e técnicas aprimoradas que contribuam no processo de diagnóstico mamográfico.

4.1. Linguagens de Programação

O protótipo do sistema foi desenvolvido basicamente em Visualworks 3, portanto utiliza uma espécie de computador virtual sobre o qual o programa pode rodar. A existência de versões de máquinas virtuais para diversas plataformas é o que propicia a fácil portabilidade programas Smalltalk entre plataformas distintas (Figura 4.14). Em Smalltalk o computador virtual recebe o nome de Máquina Virtual, enquanto o programa Smalltalk é denominado Imagem Virtual (terminação *im*); o código fonte das classes e métodos desenvolvidos padrão do ambiente são guardados no arquivo *visual.sou*, enquanto o código fonte escrito pelo programador é armazenado em um arquivo que tem o mesmo nome do arquivo imagem, porém com terminação *cha*. No protótipo desenvolvido, o arquivo imagem ou imagem virtual recebeu o nome de *mammaViewer.im* enquanto o arquivo de mudanças recebeu o nome de *mammaViewer.cha*.

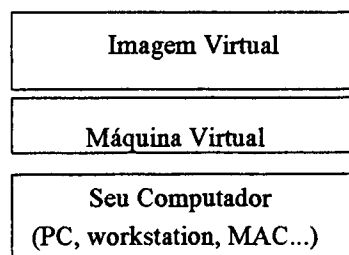


Figura 4.14: Sistema Smalltalk de execução de programas

Para implementação das funções de processamento de imagens foi escolhida a linguagem C padrão ANSI em ambiente Borland C++ 4.5, devido a exigências da plataforma

foi necessária a agregação das funções em C, dentro de uma biblioteca dinâmica (DLL) a qual chamou-se *mamas.dll*. Para que o Visualworks possa utilizar a *mamas.dll* faz-se necessário instalar a biblioteca *cw3230.dll* no diretório `c:\windows\system\`, uma das alternativas de obter tal biblioteca é o diretório `..\BC5\bin\`, ou seja, na própria instalação do ambiente Borland.

A escolha do Visualworks deveu-se ao uso deste como ambiente de programação do projeto Cyclops, no âmbito do qual foi desenvolvido este trabalho, além do que, o ambiente Visualworks tem suas virtudes, dentre elas a portabilidade. Mas, apesar de suas qualidades Smalltalk é uma linguagem interpretada e portanto relativamente lenta no que se refere a processamento de imagens, por conta disso, quando se necessitou de métodos voltados puramente ao processamento de imagens, optou-se por implementá-los em linguagem C. Para atenuar a perda de portabilidade implícita na aplicação final, optou-se pelo uso do ANSI C, visto que para portar-se a biblioteca de funções em C bastaria uma recompilação do código dado que o ANSI é um padrão contemplado por diversos compiladores em variadas plataformas.

Há de se relatar que a utilização de chamadas a uma biblioteca dinâmica de funções C (DLL) de dentro do código Smalltalk é algo que demanda cuidados especiais tendo em vista que trata-se de contemplar dois paradigmas bastante distintos (programação estruturada versus orientada a objetos) com formas diferentes de administrar a memória de trabalho, o Smalltalk com a sua *garbage collecting* enquanto o C deixa ao programador a administração do seu *heap* de memória.

4.2. Sistema Operacional

O protótipo foi implementado em ambiente Microsoft Windows, apesar da plataforma nativa do projeto Cyclops ser o UNIX, tal fato se deveu a um dos objetivos deste trabalho, ou seja, implementar a detecção em PCs. Tal escolha não implica na perda de utilidade do protótipo no contexto do projeto Cyclops, visto que, como foi explanado no item anterior, a portabilidade é importante característica do ambiente Smalltalk Visualworks. Para programação e testes foi utilizado um PC Pentium II, com relógio de 300 MHz e 64MB de RAM, recomenda-se uma configuração melhor, no mínimo um processador com relógio de 500 MHz e o dobro de memória de trabalho (RAM). Tais considerações são consideradas ideais para uma execução aceitável do programa em PC, por outro lado PCs com configurações como a sugere-



rida são cada vez mais acessíveis.

4.3. Implementação

Foram escolhidos dois métodos para detecção de sinais de microcalcificações, um baseado na aplicação de *threshold* global e outro na aplicação de *threshold* baseado no histograma da imagem. Na figura abaixo (Figura 4.15) mostra-se o esquema da abordagem que se propõe neste trabalho, a implementação de fato ocorre nos blocos marcados em linha cheia, ou seja, na preparação para detecção de sinais e na detecção propriamente dita

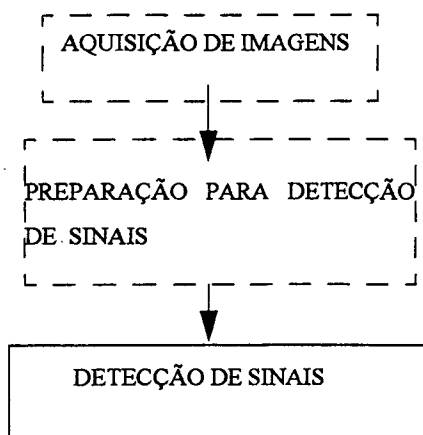


Figura 4.15: Esquema em blocos da abordagem proposta.

Entenda-se por "aquisição da imagem digitalizada" como a digitalização da mamografia. Durante o desenvolvimento deste trabalho não houve disponibilidade de equipamento de mamografia digital, portanto foram utilizadas imagens digitalizadas de mamografias convencionais. Tal processo foi realizado com *scanner* adequado, com produção de imagens em formato PGM (*Portable GreyMap*) com resolução de 300 dpi e/ou 600 dpi.

4.3.1. Preparação

A preparação para detecção de sinais consiste das etapas mostradas na figura

abaixo (Figura 4.16).

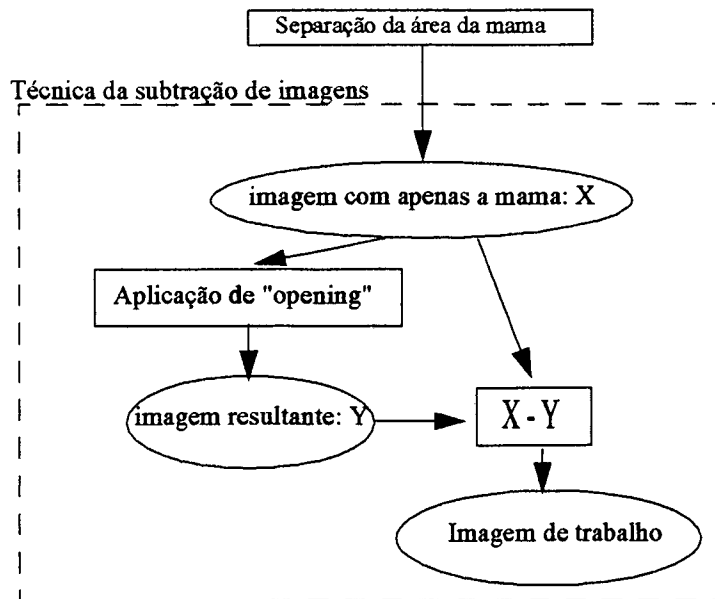


Figura 4.16: Esquema de aquisição da imagem de trabalho

A separação da área da mama consiste na segmentação da mesma do resto da imagem, através da aplicação de uma máscara, tal operação é necessária para obter-se uma imagem apenas com a área de interesse. O pressuposto básico deste algoritmo é que a maior área na imagem, que não corresponde ao fundo negro, é a mama. O algoritmo consiste nos seguintes passos:

- 1) A imagem original passa por um *opening* (erosão seguida de dilatação) para eliminação de pequenos ruídos na imagem;
- 2) Em seguida executa-se um *threshold* binário para obter-se uma imagem em preto e branco ideal para operação de rotulamento (Figura 4.17 esquerda);
- 3) Após a operação de *threshold* executam-se quatro dilatações com o objetivo de criar-se regiões contíguas fechando possíveis espaços que dificultariam a rotulação da imagem (Figura 4.17 direita);



Figura 4.17: *Threshold binário (esquerda) seguido de 4 dilatações*

- 4) Após as quatro dilatações executa-se um *closing* (dilatação seguida de erosão) para fechar pequenas fissuras dentro da imagem.
- 5) A área da imagem obtida de 4 (Figura 4.17 direita) será posteriormente rotulada, no entanto precisa-se determinar qual o sentido de convolução no eixo das abcissas (da esquerda para a direita ou o contrário) pois o algoritmo de rotulação é direcional, ou seja: no eixo horizontal, ele sempre tem que convolver a imagem da base da mama para o mamilo. Para definir isso usa-se um *kernel* coluna (Figura 4.18) com pesos distribuídos equilibradamente. O pressuposto é que na imagem binária a base da mama tem sempre uma coluna de *pixels* brancos maior que o mamilo, portanto a soma ponderada dos *pixels* encontrados pelo *kernel* coluna será zero quando se está na base da mama. Abaixo segue o trecho do código em Smalltalk que executa tal tarefa (Exemplo 4.1).



```

"-----
UFSC-CTC-Cyclops Project
    Esse método é usado para determinar a direção da convolução no eixo dos X durante a rotulação da
mamografia, através da detecção da base da mama. Tomando-se esta como referência a operação de rotulação
sempre se dá da base da mama para o mamilo.
Author-Gilmário
06/12/1999
"-----"
directionXaxeToBuildMaskHeap: arrayHeap and: anExtent
| arrayTemp acum kernel width height |
width := anExtent x.
height := anExtent y.
arrayTemp := ByteArray new:(ImageToBeProcessed image bits size) withAll: 0.
"array heap é uma área de memória administrada pelo C, é preciso convertê-lo para a memória administrada pelo
Visualworks, isso é feito nas próximas 5 linhas de código"
arrayHeap
    copyAt:0
    to: arrayTemp
    size: (arrayTemp size)
    startingAt: 1.

acum := 0.
"A priori: para cada coluna da matriz de pixels (imagem) é feita a seguinte soma ponderada:


$$\sum(pixel_i \times peso_i)$$


se esta soma for nula, testa-se 4 colunas adiante, se continuar nula trata-se de região da base da mama. Se a soma não for nula, testa-se 4 colu-
nas adiante, se continuar nula trata-se do mamilo da mama. Em ambas situações testa-se antes se há um número mínimo de pixels brancos
presentes na coluna"
1 to: width
do:[ : countCol |
    kernel := column refreshColumnUsing: countCol and: 1.
    ((column HowManyMaximumGrayTonePixelWith: arrayTemp maximumGrayTone: 255) >= ((height/3)asInteger) )
    ifTrue:[ kernel do: [:each| acum := ((arrayTemp at: (each at: 1)) * (each at: 2)) + acum. ].
        (acum ~= 0)
        ifTrue:[ kernel := column refreshColumnUsing: (countCol+4) and: 1.
            kernel do: [:each| acum := (arrayTemp at: (each at: 1)) * (each at: 2) + acum. ].
            (acum ~= 0)
            ifTrue:[ ^(#rightToLeft) ]
        ]
        ifFalse:[ kernel := column refreshColumnUsing: (countCol+4) and: 1.
            ((column HowManyMaximumGrayTonePixelWith: arrayTemp maximum GrayTone: 255) >= ((hei-
ght/3)asInteger))
            ifTrue:[ kernel do: [:each| acum := (arrayTemp at: (each at: 1)) * (each at: 2) + acum. ].
                (acum = 0)
                ifTrue:[ ^(#leftToRight) ].
            ]
        ]
    ].
].
].
].

```

Exemplo 4.1: Código em Smalltalk cuja função é detectar a base da mama



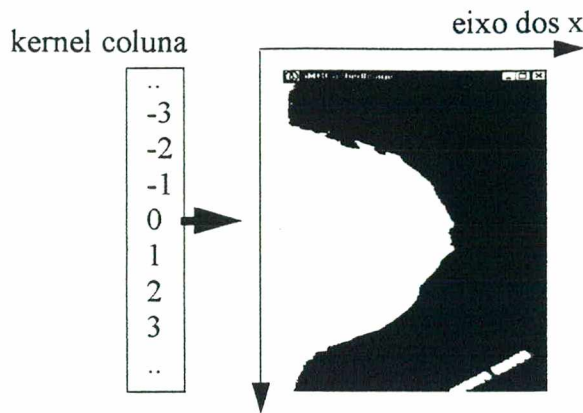


Figura 4.18: Aplicação do kernel coluna para definição de sentido (em X) de rotulação

- 6) Conhecida a direção de convolução nos eixo horizontal, a área da imagem obtida de 3 é então rotulada usando-se um *kernel* 3x3. Para tanto deve-se invocar a função em C adequada (convolução em X da esquerda-para-direita ou direita-para-esquerda). O pressuposto básico é que um *kernel* 3x3 só apontará para vizinhanças com nenhum *pixel* rotulado o ocorrerá no máximo um tipo de rótulo, nunca dois ou mais *pixels* com rótulos diferentes entre si. O algoritmo utilizado é o exibido no quadro abaixo (Exemplo 4.2), neste deve-se destacar que a operação de rotulamento é efetuada com convolução da esquerda para a direita no eixo das abcissas, a diferença deste algoritmo para o que executa a mesma convolução no sentido contrário está na linha de código que executa um *for* para contagem de colunas da imagem, trata-se de troca o *statement* que executa a operação da esquerda para a direita: `for(count_col=0; count_col < (im_w - k_dim); count_col++)`, pelo seguinte, que faz o contrário: `for(count_col=(im_w - (k_dim)); count_col >= 0; count_col--)`.
- 7) Verifica-se qual o rótulo modal na imagem rotulada;
- 8) Aplica-se um novo *threshold* sobre a imagem rotulada onde apenas os *pixels* correspondentes ao rótulo modal são "pintados" de branco, os demais são "pintados" de preto;
- 9) A máscara binária resultante do passo 8 é aplicada à imagem original, obtendo-se uma imagem apenas com a área da mama.




```

/*-----
UFSC-CTC- Cyclops Project
Esta função executa a rotulação em uma imagem, para tanto é usado um kernel 3x3 que convolve a matriz de
pixels (imagem). O sistema de coordenadas usa como origem o canto superior esquerdo da imagem e a
convolução no eixo das abcissas ocorre da esquerda para a direita.
Autor: Gilmário B. dos Santos
-----*/

#include "mamas.h"
__declspec(dllexport)
void left_to_right_build_mask( unsigned char * image_binary_array, unsigned int * byte_mask_array, unsigned int *
image_height, unsigned int * image_width, unsigned int * kernel_dimension)
{ unsigned int count_lin, count_col, count, im_h, im_w, k_dim, label, kernel_size, temp, pixel_value_in_binary_array;
  long int * kernel, array_size;
  unsigned char max_gray;
  label = 1;
  im_h = *image_height;
  im_w = *image_width;
  k_dim = *kernel_dimension;
  max_gray = (unsigned char) 255;
  array_size = (im_h*im_w);
  kernel_size = (k_dim*k_dim);
  byte_mask_array = integer_array_cleaner(byte_mask_array, array_size);
  if ( ( kernel =
(long int *) malloc(sizeof(long int)*kernel_size) ) == NULL )
  { exit(1); }
  kernel = long_integer_array_cleaner(kernel, kernel_size);
  /* count_lin representa o contador de linhas, count_col o contador de colunas */
  for(count_lin=0; count_lin < (im_h - (k_dim - 1)); count_lin++)
  { for(count_col=0; count_col < (im_w - (k_dim - 1)); count_col++)
    { kernel = refresh_squared_kernel_new(kernel, count_col, count_lin, im_w, k_dim); /*<== novo sítio do kernel*/
      if (kernel[kernel_size-1] <= array_size)
        if (temp = is_there_maximum_tone(image_binary_array, kernel, kernel_size, max_gray);
          if (temp == 1) /*<==== há pixels brancos no novo sítio ? */
            { temp = is_there_pixel_marked(kernel, kernel_size, byte_mask_array); /*<= pixels brancos já marcados ? */
              if (temp == 0)
                { label = label + 1; /*<==== há pixels brancos não marcados, crie novo label e marque-os com ele */
                  for(count = 0; count < kernel_size; count++)
                    { pixel_value_in_binary_array = image_binary_array[kernel[count]];
                      if (pixel_value_in_binary_array == 255)
                        byte_mask_array[kernel[count]] = label;
                    }
                }
              else /*há pixels brancos porém algum(ns) dele(s) já opssuem uma marca, marque os demais com a mesma */
                { for(count = 0; count < kernel_size; count++)
                    { pixel_value_in_binary_array = image_binary_array[kernel[count]];
                      if (pixel_value_in_binary_array == 255)
                        byte_mask_array[kernel[count]] = temp;
                    }
                }
            }
        }
    }
  }
  free(kernel);
}

```

Exemplo 4.2: Trecho de código em C onde se executa uma rotulação com deslocamento horizontal do kernel no sentido da esquerda para a direita



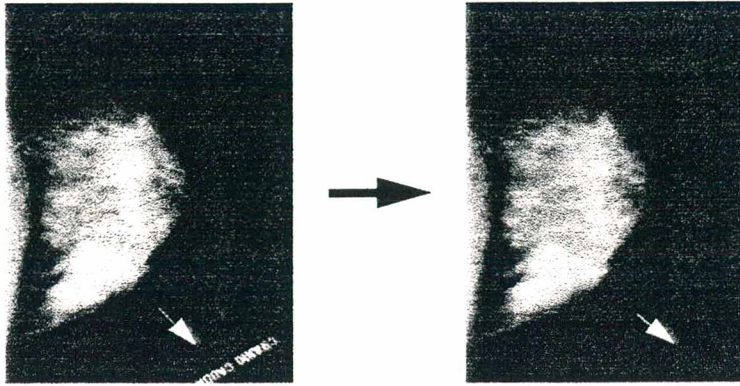


Figura 4.19: Imagem anterior e pós a separação da mama

4.3.2. Técnica da Subtração de Imagens

Uma vez obtida a imagem exclusivamente da mama (Figura 4.19 direita) pode-se aplicar a chamada "subtração de imagens" (BETAL, 1997, p. 903) que não consiste na simples operação de subtração entre duas imagens. Essa técnica assim como outras, também tira proveito da abstração de uma imagem como uma superfície topográfica e tem a finalidade de preparar a imagem para a aplicação dos métodos de detecção de sinais propriamente ditos. Agora o objetivo é atenuar o fundo da imagem (*background*) em relação aos sinais que são objeto de interesse, ou seja, as microcalcificações, essa portanto é a finalidade da técnica da diferença de imagens.

Na figura seguinte (Figura 4.20) têm-se exemplo dos resultados possíveis para a erosão e dilatação e subtração de imagens obtidas de um hipotético perfil topográfico de uma imagem (Figura 4.20 - a), deve-se destacar que o objetivo do *opening* realizado é obter-se (Figura 4.20 - c) uma imagem onde os *pixels* mais brilhantes sejam fortemente atenuados enquanto as estruturas de fundo (ou *background*) da imagem não sofram atenuação ou a sofram pouca, de forma que a imagem original ao ser subtraída desta (resultante do *opening*) praticamente elimine as estruturas de fundo enquanto preserva os *pixels* brancos deixando-os em destaque no resultado final do processo (Figura 4.20 - d).

Apesar de haver um reforço de sinal este é relativo, ou seja, os sinais de interesse ficam reforçados em relação ao fundo indesejado, porém toda a imagem sofre atenuação, por isso que a técnica da diferença não é suficiente, apenas prepara o "terreno" para o uso de

técnicas complementares.

É digno de nota que em BETAL (1997, p. 903) utiliza-se a denominação *top-hat* para a técnica acima descrita, porém após contato com textos de RUSS (1994, p. 253) notou-se que o conceito de *top-hat* é bem diferente do que é empregado por Betal, preferiu-se então utilizar a denominação "técnica da subtração de imagens", que parece ser bem mais apropriada. De qualquer sorte o *top-hat* de Russ foi testado durante a implementação do protótipo e não demonstrou apropriado para os objetivos propostos.

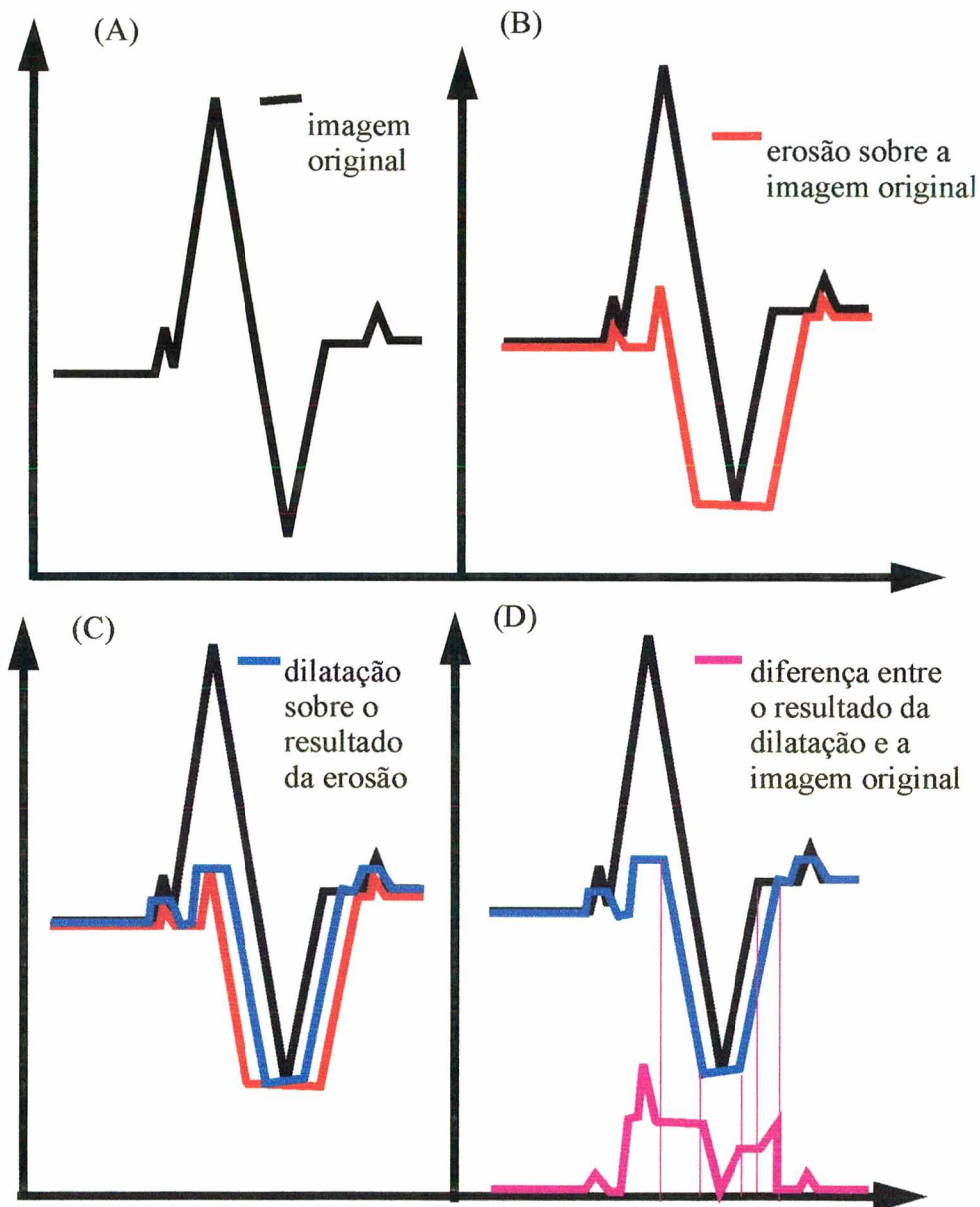


Figura 4.20 Técnica da diferença entre imagens em um perfil hipotético



Na implementação do protótipo utilizou-se *kernels* 7x7, tal dimensão é a utilizada em BETAL (1997), além disso foram feitos testes com *kernels* menores que demonstraram um desempenho inferior no que se refere à detecção de microcalcificações.

4.4. Detecção

Foram implementados dois métodos para detecção de microcalcificações, o primeiro deles baseado em *threshold* global, cujo valor de limiar é obtido a partir do cálculo de estatísticas sobre a área da mama na imagem de trabalho (NISHIKAWA, 1993, p. 1661-6) e um segundo que efetua uma filtragem espacial com *threshold* cujo limiar é definido sobre o histograma de tons de cinza da imagem de trabalho (NISHIKAWA, 1995, p. 174-8).

4.4.1. *Threshold* Binário Global

Este *threshold* é igual ao *threshold* binário convencional exceto pelo fato do limiar ser calculado sobre estatísticas da imagem, no caso, a média e desvio padrão dos tons de cinza, para tanto utilizou-se as seguintes expressões:

$$m\u00e9dia = \frac{\sum_{i=1}^n X_i}{n}$$

$$desvio_padr\u00e3o = \sqrt{\frac{\sum_{i=1}^n (X_i - m\u00e9dia)^2}{n-1}}$$

Sendo o limiar calculado através da expressão abaixo obtida em NISHIKAWA (1993, p. 1662): $limiar = F (m\u00e9dia + desvio_padr\u00e3o)$, onde F é uma fator entre [3,2;4,0].

4.4.2. *Threshold* Binário Baseado em Histograma

Novamente trata-se de um *threshold* como outro qualquer exceto que o limiar é determinado sobre o histograma da área da mama na imagem de trabalho. O histograma for-



Ambos procedimentos de detecção (*threshold* global e baseado no histograma) fornecem como resultado uma coleção de *pixels* (coordenadas espaciais) que podem ser superpostos a imagem original da mama em cores que chamem a atenção, de fato, utiliza-se o amarelo para os sinais detectados pelo *threshold* global e vermelho para aqueles detectados pelo *threshold* baseado em histograma (Figura 4.25), com isso chega-se ao primeiro objetivo proposto para este trabalho, ou seja, a implementação de métodos para destacar regiões de interesse em mamografias convencionais digitalizadas, a seguir aborda-se a segunda parte da implementação, onde se coloca alguns pontos que merecem destaque no que diz respeito à construção da interface gráfica.

4.5. Interface

Nos itens anteriores foram discutidos os aspectos de implementação no que se refere a escolha dos métodos de processamento de imagens, como é objetivo deste trabalho a implementação de uma interface que sirva de "camada amigável" para estes métodos, passa-se à discussão de como foi montada essa interface.

Como já foi dito as linguagens de programação escolhidas para implementação foram o C e o Smalltalk, nos respectivos ambientes Borland 4.5 e Visualworks 3.0, quanto ao C pode-se dizer que cada uma das funções implementadas nessa linguagem serviu como um método do Visualworks dado que este ambiente provê tal facilidade. Deixa-se então os aspectos de modelagem das funções em C que compuseram a DLL referenciada pelo Smalltalk, para serem relatados na parte de documentos anexos (listagens de código documentadas). Quanto ao mais importante, as classes do Visualworks nas quais o sistema foi implementado, segue os diagramas de classes, maiores detalhes podem ser obtidos nos anexos a este trabalho.

4.6. Classes Utilizadas

Mais de 80% das classes utilizadas nesse trabalho foram implementadas para ele, as demais ou foram especializadas de classes nativas do ambiente ou foram especializadas de classes nativas do projeto Cyclops, a seguir tece-se algumas considerações a respeito.



4.6.1. Classes do Ambiente de Programação

Destas merecem destaque as classes Image e Depth8Image, as quais foram especializadas em ImageRXMamma e Depth8ImageRXMamma (Figura 4.22).

4.6.2. Classes Nativas do Projeto Cyclops:

Basicamente foi utilizada neste trabalho a classe MRCachedImage, que foi especializada na RXCachedImage (Figura 4.22);

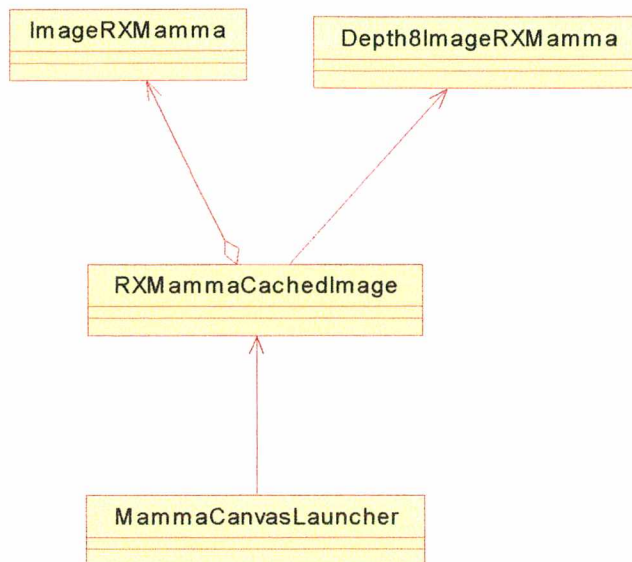


Figura 4.22: Especializações de classes do ambiente Visualworks 3

4.6.3. Classes Novas

Entenda-se como tal, as classes que foram desenvolvidas para implementar o protótipo (Figura 4.24).

4.6.3.1. MammaCanvasLauncher

Responsável por lançar as duas principais interfaces: a interface que lança sistema (Figura 4.23) e a interface MammaCanvas1 (Figura 4.25). Além disso, possibilita a comunicação entre as principais classes do sistema.



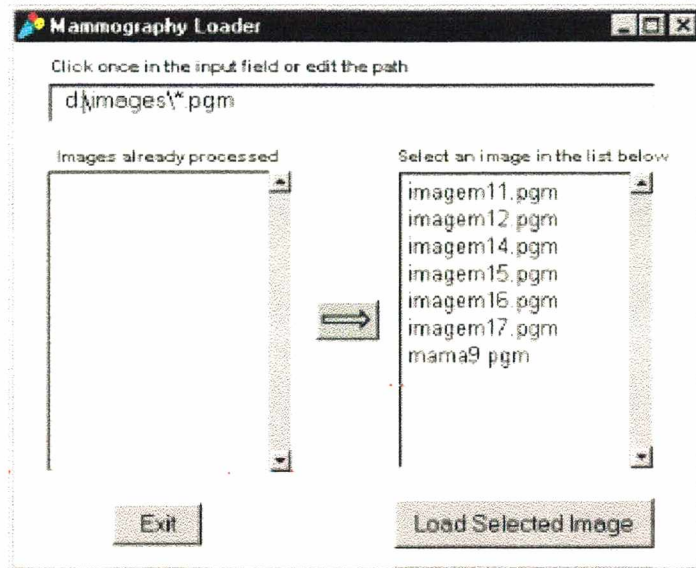


Figura 4.23: Interface que lança o protótipo

4.6.3.2. MammaCanvas1

É a principal interface do sistema (Figura 4.25), através dela o usuário tem acesso a todas as opções de detecção de sinais bem como a algumas outras utilidades, tais como: magnificação de trechos da imagem, inversão do mapa de cores, possibilidade de "riscar" sobre a imagem, capacidade de apagar riscos e marcação de sinais. Maiores detalhes sobre como o usuário tem acesso a esses utilitários e funções podem ser vistos no manual do usuário em anexo.

4.6.3.3. ProcImageLauncher

Classe que implementa os métodos de processamento de imagem implementados em Visualworks e que adequa os dados oriundos das chamadas a funções C, para dentro do Visualworks, bem como adequa os dados que devem ser passados às funções C. Entenda-se por adequação como sendo o processo de conversão de objetos Smalltalk em tipos C, alocação adequada de memória para áreas que o C trabalha, recuperação de memória alocada para o C e já em desuso, etc.

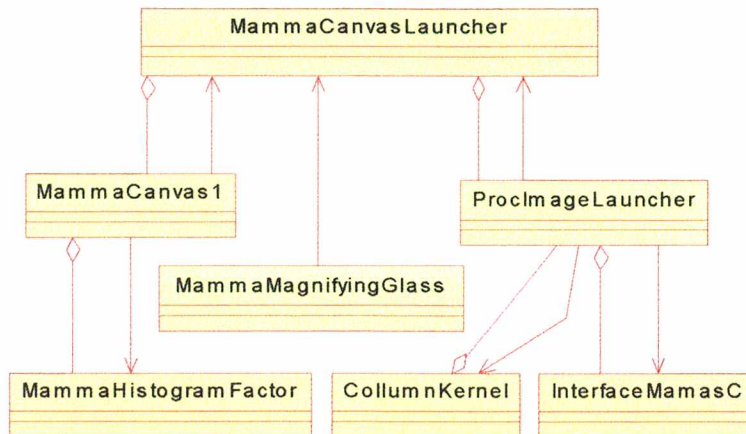


Figura 4.24: Principais classes implementadas e seus relacionamentos

4.6.3.4. MammaMagnifyinGlass:

Possui as especificações da janela que exhibe uma área magnificada selecionada na imagem, possui métodos que viabilizam vários graus de ampliação.

4.6.3.5. CollumnKernel

Emula um *kernel* coluna que é usado para definir o sentido de convolução no eixo do X, na operação de rotulação.

4.6.3.6. MammaHistogramFactor:

Esta classe possui as especificações de uma janela que permite ao usuário fornecer valores que serão usados para cálculo de limiar de *threshold* sobre um histograma.

4.6.3.7. InterfaceMamasC

Esta classe contempla um conjunto de métodos que na verdade são a interface entre o Visualworks e o C, tais métodos têm uma relação de um-para-um com as funções C presentes na biblioteca dinâmica (dll) que será invocada quando necessário. Maais uma vez, resalta-se que as funções C são exclusivas para processamento de imagens e se justificaram pela performance implícita que trouxeram ao sistema.



4.6.3.8. O modelo MVC

MVC é uma abreviatura de *Model-View-Controller* é o nome dado ao modelo utilizado para o desenvolvimento de aplicações no Smalltalk, esse modelo preconiza a divisão funcional como uma maneira mais prática de desenvolver aplicações interativas, possibilitando ao desenvolvedor da aplicação um melhor entendimento das partes e portanto facilitando modificações e extensões de cada uma delas. Há então três divisões da aplicação: 1. As partes que representam o modelo da base do domínio de aplicação (*Model*); 2. A forma como o modelo é apresentado para o usuário (*View*); e 3. A forma como o usuário interage com o modelo (*Controller*).

Todas as classes da figura (Figura 4.26), exceto a classe *MammaCanvas1*, implementam o modelo MVC. É digno de nota que algumas dessas classes, particularmente a *Sketch1*, *SketchView1* e *SketchController1*, são classes especializadas de *Sketch*, *SketchView* e *SketchController*, tais classes são nativas do Visualworks, porém fazem parte de exemplos do manual *Application Developer's Guide* do Visualworks 3 e portanto não são obrigatórias, muitas vezes torna-se necessária a digitação destas tomando como base o exemplo do manual.

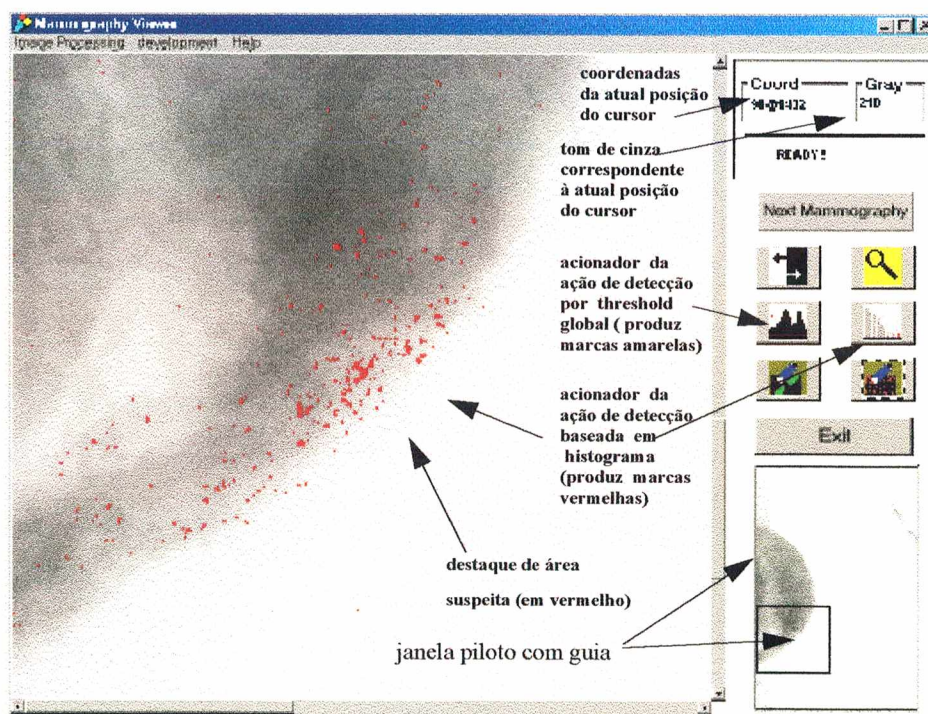


Figura 4.25: Janela principal do protótipo



As descrições feitas neste capítulo serão complementadas pelas informações presentes nas listagens documentadas dos códigos em C e Smalltalk bem como pelo manual do usuário do protótipo, ambos inclusos em anexo. Nestes documentos listam-se todas as classes Smalltalk e funções C, bem como os relacionamentos entre as classes Smalltalk.

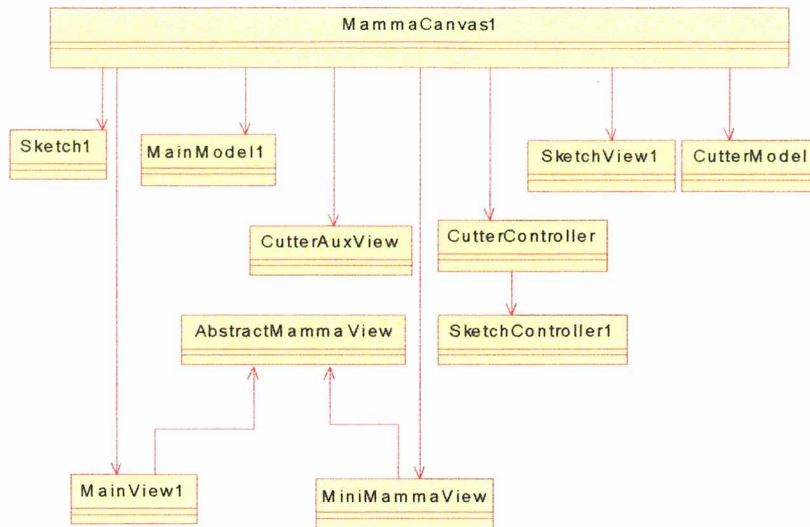


Figura 4.26: Principais classes que implementam o modelo MVC

Um melhor entendimento sobre o funcionamento da interface sob a ótica do usuário pode ser encontrada no manual do sistema em anexo.



CAPÍTULO 5 - Testes

Os testes foram realizados com mamografias coletadas de um banco de dados de imagens de mamografias provido pelo MIAS - Mammographic Image Analysis Society, através da internet. Para fins de comparação, também foram utilizadas algumas mamografias cedidas por uma importante clínica radiológica de Florianópolis. Por motivos que serão discutidos a seguir, os testes foram fundamentados no material do MIAS.

5.1. MIAS

MIAS é um consórcio de pesquisa inglês constituído de grupos multidisciplinares, voltados ao estudo do câncer de mama. Sua importância neste trabalho está no banco de imagens mamográficas que esta instituição mantém. Contudo deve-se destacar que muitas outras organizações de pesquisa contemplam satisfatoriamente esta necessidade (em http://min.ecn.purdue.edu/~sliu/mammo_db.html pode-se ter uma lista de tais instituições), no entanto, uma série de fatores privilegia o banco mantido pela MIAS, tais como: (i) a possibilidade de adquirir uma versão reduzida gratuita do banco; (ii) a boa documentação disponível a respeito dos casos; (iii) o formato gráfico usado para armazenar as imagens (PGM); e (iv) a ocorrência de citação (do MIAS) em DINTEN (1996, p. 351) com a mesma finalidade que se pretende empregá-lo neste trabalho, o que lhe confere credibilidade e confiabilidade.

O endereço na internet do "site" do MIAS é o seguinte: <http://s20c.smb.man.ac.uk/services/MIAS/MIAScom.html>, a partir deste pode-se dispor de *link* para o diretório de ftp, onde 335 casos estão disponíveis gratuitamente, desses fez-se *download* de vinte e cinco, escolhidos por serem citados na documentação como apresentando calcificação, mais três citados como normais, o que dá um total de vinte e oito casos. Quanto à documentação, no mesmo diretório que disponibiliza as imagens pode-se ter acesso a mesma. Seguem alguns exemplos de registros na documentação MIAS:

- (A) mdb015- G - CIRC - B - 595 - 684 - 68;
 - (B) mdb104 - D - ASYM - B - 357 - 365 - 50;
 - (C) mdb034 - D - NORM.
-

Tabela 1: Significado dos campos de registros na documentação MIAS

coluna 1	coluna 2 - tipo de tecido	coluna 3 - classe da anormalidade presente	coluna 4 - grau da anormalidade	colunas 5 e 6 - x, y	coluna 7
Número de referência no MIAS	F: Fatty G: Fatty-Glandular D: Dense	CALC- calcificação CIRC- lesão circunscrita SPIC- massa espiculada MISC - outras ARCH - distorção de arquitetura ASYM - assimetria NORM - normal	B - benigno M - maligno	coordenadas do centro da anormalidade	Raio aproximado do círculo que circunscreve a anormalidade.

5.2. Imagens Locais

Foram também utilizadas algumas imagens digitalizadas de mamografias convencionais, cedidas por clínica radiológica local.

5.3. Comparando Fontes de Dados

As imagens do MIAS mostraram-se mais brancas que as imagens locais, supõe-se que isso se deva a variações nos *settings* dos equipamentos geradores das mamografias, de qualquer forma a utilização destas imagens, pelo protótipo, exigiu um aumento no limiar de *threshold* para 10 no início do processo de separação da mama, para evitar distorções como as que podem ser vistas abaixo (Figura 5.27).

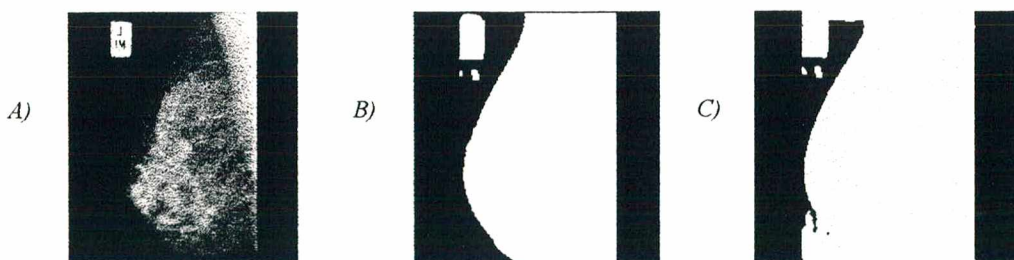


Figura 5.27: Imagem normal (A), imagem após *threshold* com limiar = 10 (B) e limiar = 2 (C)

A diferença entre as imagens locais do MIAS podem ser notada pela facilidade com que se vê o contorno da mama em uma imagem do MIAS e a impossibilidade prática de se identificar esse contorno a olho nú em uma imagem local na figura abaixo (Figura 5.28).

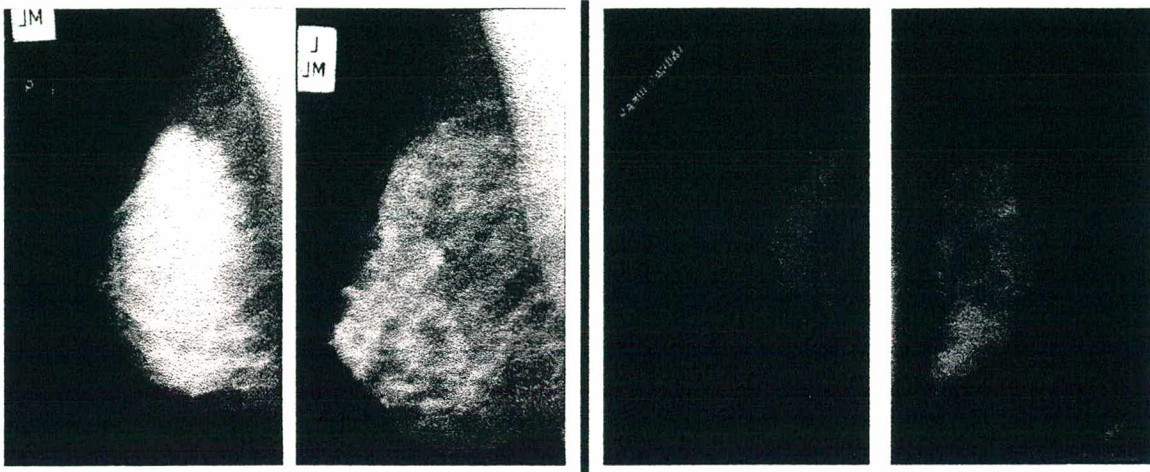


Figura 5.28: Comparação entre imagens do MIAS (esquerda) com imagens locais

5.4. Preparação para Testes

No que se refere as imagens oriundas de mamografias cedidas por clínicas locais, foi solicitado ao radiologista (especialista em laudos mamográficos) que as forneceu, uma marcação prévia das regiões suspeitas, dessa forma ao executar o protótipo com essas imagens basta conferir se os achados casavam com as marcas feitas pelo radiologista. Tais imagens foram digitalizadas em *scanner* especial com adaptador que possibilita a digitalização de elementos transparentes, no procedimento de digitalização não foram empregados quaisquer ganho de autocontraste ou brilho, a resolução utilizada foi de 300 dpi, parâmetro usualmente encontrado na literatura, com escala de 100%, por fim foi utilizado um mapa de tons de cinza com profundidade 8 bits, ou seja, que permite variações de cinza do zero (negro) a 255 (branco), o que também é usual na literatura. Tal configuração levou a imagens com tamanho médio de de 3 MegaBytes cada.

Como já foi dito, as mamografias obtidas do banco MIAS já foram adquiridas digitalizadas e com documentação, além disso as imagens do MIAS apresentam-se em dimensões 1024 Bytes por 1024 Bytes o que lhes confere tamanho total de 1 MegaByte, portanto bem inferior ao das imagens locais o que confere uma menor complexidade intrínseca aos programas.



Ainda em relação ao MIAS também deve-se destacar algumas peculiaridades, em primeiro lugar elas não estão em um formato PGM padrão, faz-se então necessária uma conversão para o PGM padrão, tal procedimento é extremamente simples, basta uma abertura e re-salvamento da imagem (*save as*) por intermédio de um editor de imagens que trabalhe com o formato PGM padrão. Em segundo lugar e extremamente importante de se ressaltar, as coordenadas da anormalidade citadas na documentação se referem a um sistema de coordenadas onde a origem dos eixos está no canto inferior esquerdo da imagem, diferentemente do usual em processamento de imagens onde a origem dos eixos localiza-se no extremo superior esquerdo. Mas, de fato, nenhuma dessas particularidades gera demérito para o banco MIAS.

Para execução do protótipo recomenda-se a utilização de sistema de vídeo de alta resolução, com monitor de no mínimo 17". O sistema foi montado pensando em uma clientela que dispõe de PCs com microprocessador de no mínimo 500 MHz e 131 Mb de memória RAM para que haja um bom desempenho.

5.5. Procedimentos de Avaliação

Inicialmente esclarece-se que os métodos de detecção aqui implementados são oriundos do estado da arte em processamento de imagens de mamografias digitalizadas colhidos na literatura, especialmente em BETAL (1997, p. 903) e NISHIKAWA (1995, p. 174-8, 1993, p. 1661-6). Nestas referências, publicadas em importantes meios de divulgação científica, pode-se encontrar relatos de testes de validação bastante criteriosos e rígidos, portanto trata-se de modelo de detecção já validado. Aqui o que fez foi testar a implementação do modelo frente a realidade local e as adaptações inseridas no decorrer do desenvolvimento. Tentou-se verificar o grau de acerto do protótipo para a casuística do *MIAS Database* (**referência de qualidade**).

Considera-se aqui os valores 0,999 para *Threshold por histograma* e 3,0 (no intervalo definido em 4.4.1) para *Threshold* por estatística global como valores *default*, visto que foram estes os definidos como tal na implementação.

Para cada imagem processada levantou-se:

a) o número de *pixels* detectados dentro da área delimitada como suspeita tendo-se como base a documentação do MIAS e;



b) o número total de *pixels* detectados na imagem.

Foram selecionadas vinte imagens do MIAS as quais foram submetidas aos dois métodos. Previamente, através da interface apresentada na Figura 5.29 (interface original devidamente alterada) foram informados:

- I. se a imagem apresenta calcificações ou não;
- II. as coordenadas do centro do círculo que delimita a anormalidade;
- III. o tamanho do raio do círculo que delimita a anormalidade.

Foi montado um arquivo onde cada registro tem a seguinte forma:

imagem_id método_usado (classe da anormalidade, centro da anormalidade, raio do círculo, pixels detectados dentro do círculo, total de pixels detectados). Os dois últimos dados foram calculados

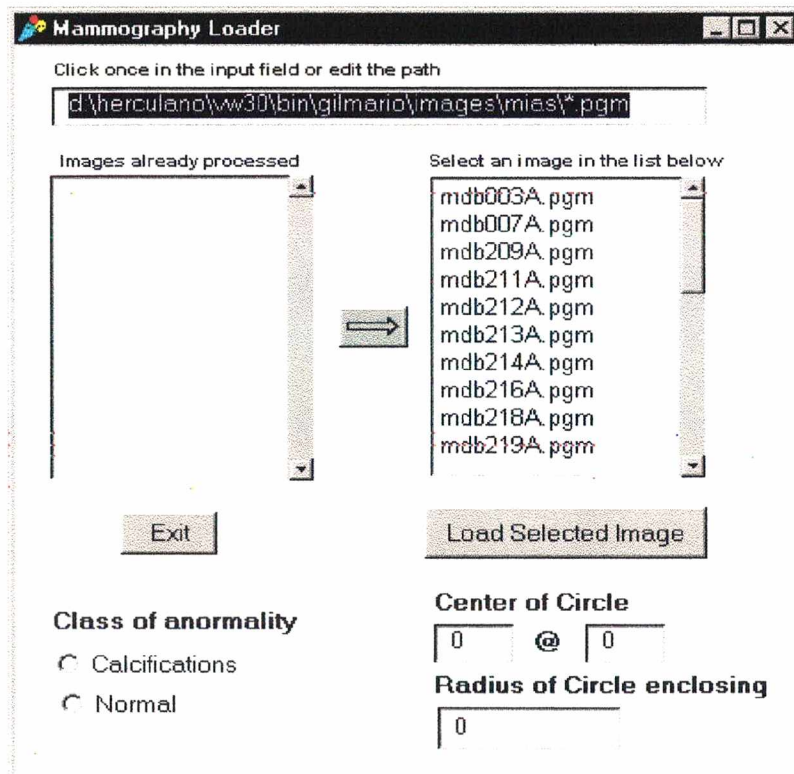


Figura 5.29: Interface alterada para fins de teste de avaliação dos métodos de detecção

tomando como base as informações previamente fornecidas através da interface acima (Figura 5.29). O conteúdo do arquivo de teste e o fragmento de código do método que monta o arquivo, encontram-se no anexo F. Foram utilizados os fatores 0,9;0,99;0,994;999;0,9996 para o *Threshold* por histograma e 1,5;2,0;2,6;3,0;3,6 para *Threshold* por estatística global.



5.5.1. Avaliação Geral

A) Como se esperava, pela análise dos gráficos na Figura 5.31 e na Figura 5.32 percebe-se que a sensibilidade dos métodos varia com os fatores utilizados para definir o "corte" sobre o histograma e o limiar global (ver 4.4.1 e 4.4.2). Percebe-se que quanto menos se "corta" o histograma ou quanto menor é o fator multiplicativo que define o limiar global, mais *pixels* são detectados dentro das áreas definidas como suspeitas na documentação, portanto maior é o grau de acurácia dos resultados obtidos;

B) Ainda pela análise dos gráficos na Figura 5.31 e na Figura 5.32 percebe-se que o aumento da contagem de *pixels* dentro das áreas suspeitas, pela diminuição dos fatores, implica também na proporcional elevação no total de *pixels* detectados na imagem (verdadeiro positivos e/ou falso positivos);

C) Através dos resultados apresentados pelo gráfico na Figura 5.30 e Tabela 2, percebe-se inicialmente um fraco grau de acerto devido a baixa percentagem apresentada de *pixels* dentro das áreas suspeitas. Deve-se esclarecer que tal fato ocorre por problemas nas imagens (ocorrência de artefatos por exemplo) tais problemas não são levados em consideração pelo analista humano mas interferem no processamento computacional. Como pode-se ver na Figura 5.33, devido a algum defeito na imagem, após seu processamento foi detectada uma região de *pixels* claramente falso-positiva, tal ocorrência incrementa expressivamente a contagem total de *pixels* e compromete o cálculo percentual do acerto. Tal cálculo é feito pela razão entre os *pixels* detectados corretamente e o total de *pixels* detectados, levando-se em consideração que a quantidade de *pixels* corretos está limitada por um círculo (documentação do MIAS) de pequenas dimensões e que na contagem total incidem também os *pixels* corretos, qualquer incremento demasiado desta segunda parcela implica em diminuição da percentagem de ocorrência de *pixels* corretos.

Das características levantadas nos itens *A*, *B*, *C* e **levando-se em consideração o contexto do MIAS**, deduz-se que:

A qualidade das imagens **gratuitas** do MIAS *Database* não têm a qualidade que se esperava. De fato recomenda-se fortemente que, para novas aplicações de bancos de imagens, seja adquirida uma versão completa com todas as suas garantias de qualidade ou montada uma casuística com imagens locais, também dentro de parâmetros rígidos.

De qualquer forma, para uma melhoria do grau de acerto faz-se necessário acrescentar uma etapa de tratamento de artefatos e/ou outros defeitos que possam gerar o



baixo grau de acerto. Por outro lado pode-se optar pela utilização de métodos complementares aos implementados, métodos que tenham como base o processamento local, de forma a tirar proveito da segmentação inicial propiciada pelos métodos (globais) aqui testados. Neste caso, o *top-hat*, segundo a definição dada por RUSS (1994, p. 253), parece ser uma boa alternativa, tal método baseia-se em parâmetros em vizinhança de *pixels* de interesse e deve ser aplicado em complemento a processamentos globais de forma a fornecer resultados melhores que a simples aplicação de *threshold* baseado em parâmetros globais.

5.5.2. Interface

A interface foi considerada bem construída e amigável, foram utilizados ícones com boa semântica associada e botões bem posicionados, além de bem distribuídos. A quantidade de funções disponíveis se mostrou ainda limitada, mas tratando-se dos objetivos propostos é aceitável, ficando outras extensões para trabalhos futuros.

Tabela 2: Totais pixels detectados dentro da area suspeita (acertos) *VERSUS* total de pixels detectados fora da região suspeita (erros) - usando o MIAS Database

	total de acertos	total de erros	total geral
threshold por estatística global, fator = 1,5	1441	74517	75958
percentual	1,9%	98,1%	100%
threshold por histograma, fator = 0,9	561	40903	41464
percentual	1,4%	98,6%	100%



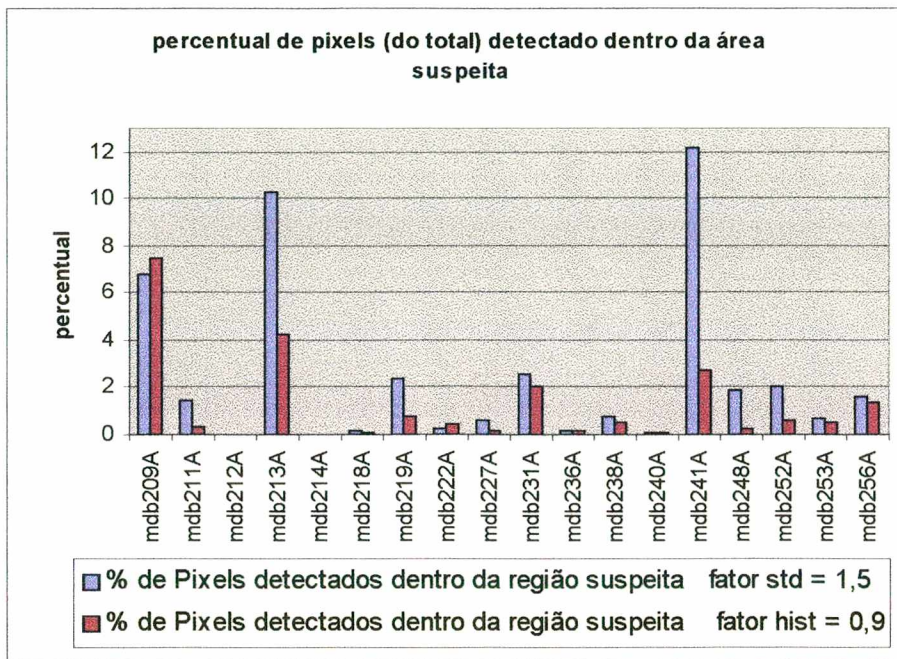


Figura 5.30: Percentual de acerto sobre o total de pixels detectados

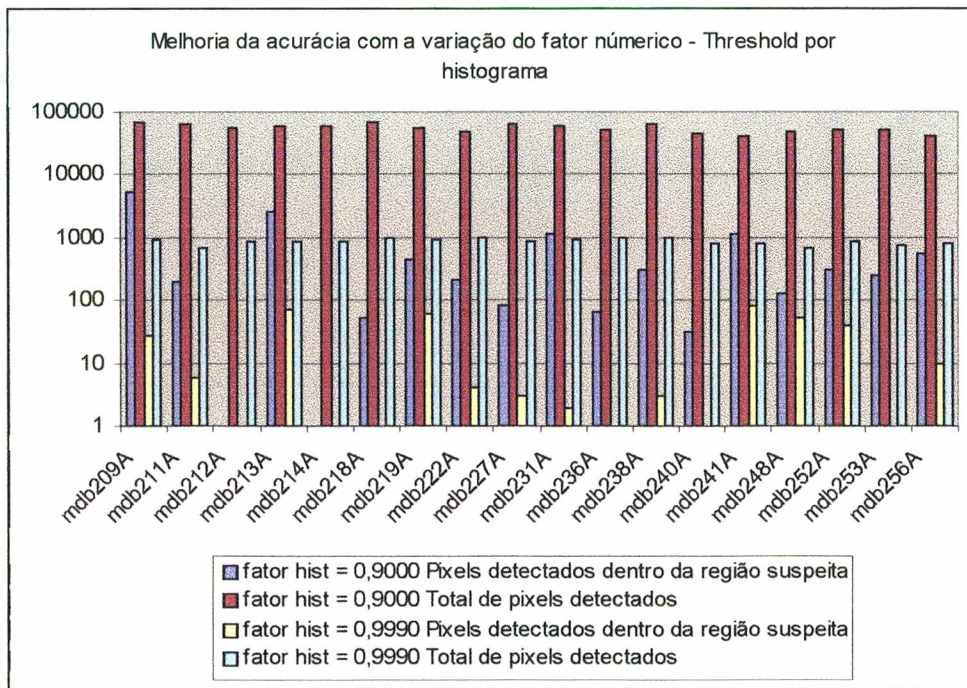


Figura 5.31: Variação da acurácia e falsos positivos - Threshold por histograma



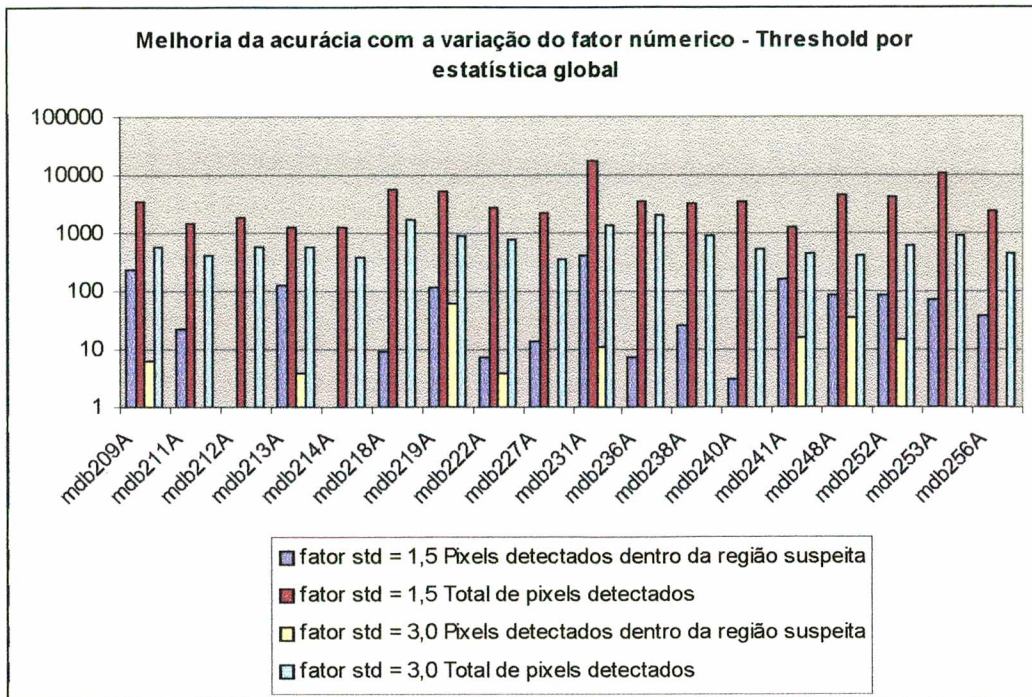


Figura 5.32: Variação da acurácia e falsos positivos - Threshold por estatística global

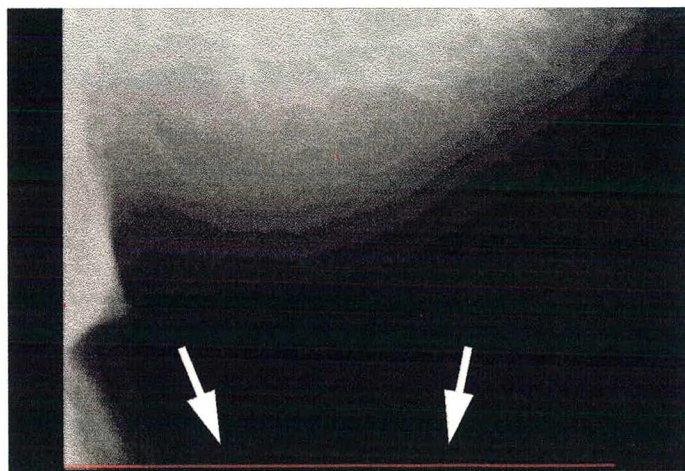


Figura 5.33: Defeito na imagem acarretando detecção de pixels indevidamente



CAPÍTULO 6 - Conclusões

Esse trabalho nasceu de uma primeira proposta de pesquisa na área de mineração de dados. Pretendia-se montar uma base de conhecimento acerca de agrupamentos de microcalcificações através da análise de padrões, como tal projeto não progrediu, optou-se por uma outra proposta a altura, a construção de um protótipo de CAD voltado à detecção de calcificações, um projeto no qual haveria o aproveitamento de parte do que já havia sido estudado. Apesar do esforços empreendidos, este novo desafio se mostrou muitas vezes pouco reconhecido, talvez porque pesquisar sobre ferramentas CAD, no terceiro mundo, seja uma atitude quase quixotesca, dado o desnível tecnológico brutal entre o contexto nacional/regional e o contexto dos países ditos do primeiro mundo, onde a pesquisa no campo da detecção do câncer de mama está bastante evoluída. Apesar disso, é com satisfação e orgulho do dever cumprido, que se conclui este trabalho.

6.1. Objetivos Alcançados

Os objetivos colocados no início deste trabalho foram alcançados a contento:

- A) A implementação de métodos de detecção: dois métodos de detecção foram colhidos do estado da arte e implementados, um baseado em *threshold* global e outro em *threshold* baseado em histograma;
 - B) O encapsulamento destes em uma interface amigável: foram implementadas algumas interfaces, uma principal onde concentram-se todos os recursos e duas outras secundárias, auxiliares;
 - C) A construção de um sistema portátil: o requisito de portabilidade foi buscado ao máximo, no entanto o desempenho do Visualworks em processamento de imagens forçou a utilização de chamadas a funções C implementadas em DLL, isso quebrou parcialmente a portabilidade do sistema em virtude desta ser uma característica do Visualworks, no entanto todas as implementações em C foram feitas em ANSI C, como trata-se de padrão bastante comum em diversas plataformas, a adoção da proposta ANSI deve atenuar a perda de portabilidade provocada.
 - D) Prover uma referência para estudos futuros: na medida em que todos os objetivos anteriores foram alcançados, alcançou-se também mais este.
-

6.2. Dificuldades Encontradas

As maiores dificuldades para execução deste trabalho dizem respeito a disponibilidade de imagens mais adequadas à análise computacional. Tem-se uma expectativa de que, com o advento do mamógrafo digital sejam produzidas imagens mais próprias para o processamento computacional e com isso se obtenha melhores resultados na detecção de calcificações e outras lesões da mama, ao invés de empreender-se tempo em desenvolvimento e em ciclos de CPU, com procedimentos de preparação que garantam uma qualidade mínima das imagens, como se vê hoje.

A seguir tecem-se comentários mais específicos acerca das principais dificuldades encontradas.

6.2.1. Geração da Mamografias

Sabe-se que os aparelhos de RX sofrem ajuste, havendo normas e margens legais para tais configurações, ocorre que esses ajustes não objetivam gerar imagens que serão processadas por computador, mas sim para serem analisadas pelo olho humano. Tal premissa aplicada na prática, produz imagens muito escuras, como as obtidas das clínicas locais. Como exemplo de imagens mais adequadas para o processamento computacional, tem-se as imagens colhidas do banco MIAS, através da Figura 5.28 já citada, pode-se ter um comparativo entre as imagens das duas fontes.

A vantagem das imagens mais claras está na maior facilidade com que se determina um *threshold* ideal, ou seja, um limiar que produz uma imagem bastante compacta e ideal para a rotulação proposta aqui.

6.2.2. Digitalização Manual

O segundo fator diz respeito à necessidade de digitalização manual, neste caso faz-se referência apenas às imagens obtidas das clínicas locais.

A repetibilidade da tarefa, a lentidão do processo devido ao não uso de redução das imagens, aliados a fatores como a sujeira no anteparo transparente do scanner, o posicionamento da mamografia no momento da operação de digitalização e a própria qualidade da mamografia, produz imagens nem sempre adequadas ao processamento.



6.2.3. Outras Dificuldades

Como dificuldades secundárias pode-se citar o desconhecimento inicial do ambiente de programação Visualworks cujo domínio demandou algum tempo e a carência de computadores mais adequados ao desafio proposto.

6.3. Recomendações para Trabalhos Futuros

A proposta básica deste trabalho foi montar um protótipo inicial que sirva como referência para trabalhos futuros, o que se oferece neste momento é justamente isso, um estudo teórico e uma implementação com os quais gradativamente será possível construir-se um *know-how* próprio que possibilite o desenvolvimento de novas propostas com abordagens inovadoras e mais próximas da realidade nacional. Segue então as seguintes sugestões para trabalhos futuros:

6.3.1. Inserção de Técnicas de IA

O maior desafio dos softwares CAD é o efetivo apoio a decisão, para tanto a aplicação de técnicas de inteligência artificial são imprescindíveis, dada a complexidade e quantidade de fatores envolvidos no processo de diagnóstico. Para contemplar tal desafio tem-se usado principalmente o modelo das redes neurais artificiais para convolução de imagens. A inserção de tal ferramenta no protótipo sem dúvida será de grande valia.

6.3.2. Conexão com Teleradiologia

O aproveitamento deste futuro CAD no contexto da teleradiologia será bastante proveitoso para médicos e pacientes que, devido a distância aos grandes centros, não dispõem de contato com profissionais mais qualificados. Já existem estudos no âmbito do projeto Cyclops com vistas a implementação de recursos de teleradiologia, no futuro poder-se-a agregar o protótipo nesta rede.



6.3.3. Conexão ao DICOM

No capítulo II quando justificou-se o surgimento da mamografia digital, fez-se alguns comentários sobre a emergente a informatização do setor de saúde, na onda desta mesma informatização surgiu o DICOM, que entre outras coisas é um protocolo de comunicação entre equipamentos médicos (vide anexos para uma descrição melhor) que permite ao seu usuário o acesso a informações geradas por diversos equipamentos médicos (tomógrafo, mamografo digital, etc), por informações entenda-se por exemplo: toda a ficha do paciente e todos os exames já feitos por ele. Fica a sugestão de no futuro implementar-se o acoplamento ao DICOM, do protótipo desenvolvido.



Referências Bibliográficas

- BETAL, D.; ROBERTS, N.; WHITEHOUSE, G. H. Segmentation and numerical analysis of microcalcifications on mammograms using mathematical morphology. **The British Journal of Radiology**. v. 70, n. 837, p. 903-917, sep., 1997.
- CHAN, H.P., et al. Image feature analysis and computer-aided diagnosis in digital radiography. I. automated detection of microcalcifications in mammography. **Medical Physics**, [s.l.], v.14, n. 4, p. 538-48, jul./aug., 1987.
- CHAN, H.P., et al. Computer-aided detection of microcalcifications in mammograms. Methodology and preliminary clinical study. **Invest Radiol**, [s.l.], v. 23, n. 9, p. 664-71, sep., 1988.
- CHAN, H. P.; et al. Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network. **Medical Physics**, [s.l.], v. 22, n. 10, p. 1555-67, oct., 1995.
- CHAN, H.P.; et al. Improvement in radiologist's detections of clustered microcalcifications on mammograms. The potential of computer aided-diagnosis. **Invest. Radiol.**, [s.l.], v. 25, p. 1102-1110, 1990.
- CHAN, H.P., Digitization requirements in mammography: effects in computer-aided detection of microcalcifications, **Medical Physics**. [s.l.] v. 21, p.1203-1211, 1994.
- D'ORSI, Carl J.; DEBOR, Mary Dale. Communication Issues in Breast Imaging, **Radiologic Clinics of North America**, [s.l.] v. 33, n. 6, p. 1231-45, nov., 1995.
- DHAWAN, Atam P., "Computerized mammographic image analysis for reducing false positive rate for biopsy recommendation". Proceedings SPIE Biomedical Image Processing and Biomedical Visualization, San Diego, vol. 1905 (2), pp. 540-542, 1993.
- DINTEN, Jean-Marc; DARBOUX, Michel; NICOLAS, Eric. Features Extraction for a Precise Characterization of Microcalcifications in Mammograms, **IEEE**, [s.l.], p. 351-4, 1996.
- DUARTE, Dakir L. Mamografia. In: MONTORO, Antonio Franco. **Mastologia**. São Paulo, Sarvier, 1984.
- FEIG, Stephen A., YAFFE, Martin J. Digital Mammography, Computer-Aided
-

- Diagnosis and TeleMammography. **Radiologic Clinics of North America**, [s.l.], v. 33, n. 6, p. 1205-30, nov., 1995.
- FILHO, Ogê Marques; NETO, Hugo vieira. **Processamento Digital de Imagens**. Rio de Janeiro, Brasport, 1999.
- GOMES, Jonas; VELHO, Luiz. **Computação Gráfica: imagem**. Rio de Janeiro, IMPA/SBM, 1994.
- HIGHNAM, Ralph et al., A representation for mammographic image processing, **Medical Image Analysis**, [s.l.], v. 1, n. 1, p. 1-18, 1996.
- INCA, Instituto Nacional de Câncer - Ministério da Saúde. **Falando sobre doenças da mama**. Rio de Janeiro, INCA - Instituto Nacional do Câncer, 1996.
- JIANG, Y.L. et al. Malignant and benign clustered microcalcifications - Automated feature analysis and classification. **Radiology**, [s.l.], v. 198, n. 3, p. 671-8, mar., 1996
- MESCHAN, Isadore. **Synopsis of analysis of roentgen signs in general radiology**. [s.l.], W. B. Saunders Company, 1976.
- NERSESSIAN, Ana Carolina., et al. Microcalcificações Agrupadas: correlação anátomo-radiológica, **Radiologia Brasileira**, São Paulo, n. 2, p. 63-7, 1999.
- NISHIKAWA, R. M., et al. Computer-aided detection of clustered microcalcifications: An improved method for grouping detected signals. **Medical Physics**. [s.l.], v. 20, n. 6, p. 1661-6, nov./dec., 1993
- NISHIKAWA, R.M., et al. Computer-aided detection of clustered microcalcifications on digital mammograms. **Medical & Biological Engineering & Computing**. [s.l.], v. 33, n. 2, p.174-8, mar., 1995.
- OLIVEIRA, C.; at al., Câncer de mama – diagnóstico por imagem – mamografia. In: Franco, Josélio M. **Mastologia – formação de especialistas**. São Paulo, Atheneu, 1997.
- PARKER, J. R., **Algorithms for Image Processing and Computer Vision**. [s.l.], John Wiley & Sons, Inc., 1997.
- PATRICK, Edward A., et al. Expert Learning system Network for Diagnosis of Breast Calcifications. **Invest. Radiol.**, [s.l.], vol. 26, n. 6, p. 534-539, jun, 1991.
- PRO-ONCO, Coordenação de Programas de Controle de Câncer - Ministério



da Saúde - Instituto Nacional de Câncer. **Câncer de mama - Estadiamento clínico - TNM/1987**. [s.l.], 1994. Cartaz informativo.

RUSS, John C. **The Image Processing Handbook**. [s.l.], Editora CRC Press, 1994.

TABÁR, László; DEAN, Peter B. **Atlas de mamografia**. Rio de Janeiro, Revinter, 1994.



ANEXO A - Bibliografia Consultada

Além da bibliografia consultada e referenciada há a seguinte lista de material que foi consultado porém não referenciado neste trabalho:

ACS, American Cancer Society. **What's Cancer ?** [on line], Documento disponível na internet via URL: http://www3.cancer.org/cancerinfo/gen_main_cont.asp?st=wi, out., 1998a.

ACS, American Cancer Society. **Breast Cancer.** [on line], Documento disponível na internet via URL: http://www3.cancer.org/cancerinfo/main_cont.asp?ct=5&st=wi, out., 1998b.

BARROS, Alfredo C. S. D.; PINOTTI, José A., Bases da carcinogênese mamária. In: DIAS, Ezio N. et al. **Controvérsias em Mastologia.** Rio de Janeiro, Revinter, 1998.

BENNINGTON, James L. **Saunders Dictionary & Encyclopedia of Laboratory Medicine and Technology.** [s.l.], Press of W. B. Saunders Company, 1984.

BERNARDELLO, Edgardo T. L.; URIBURU, Júlio V. Patologia mamária no homem. In: MONTORO, Antonio Franco. **Mastologia.** São Paulo, Sarvier, 1984.

BURNS, George W.; BOTTINO, Paul J., **Genética,** Rio de Janeiro, Guanabara Koogan S.A., 1991.

CALEFFI, Maira., Genética e câncer de mama. In: DIAS, Ezio N. et al. **Controvérsias em Mastologia.** Rio de Janeiro, Revinter, 1998.

CHAVES, Júlio Edi. Detecção precoce do carcinoma de mama - uma visão epidemiologia. **Femina - Revista da Federação Brasileira das Sociedades de Ginecologia e Obstetrícia,** Rio de Janeiro, v. 14, n. 1, p. 64-68, jan., 1986.

D'ORSI, Carl J.; DEBOR, Mary Dale. Communication Issues in Breast Imaging, **Radiologic Clinics of North America,** [s.l.] v. 33, n. 6, p. 1231-45, nov., 1995.

DANGELO, José Geraldo; FANTTINI, Carlo Américo. **Anatomia Básica dos Sistemas Orgânicos - Com a descrição dos ossos, juntas, músculos, vasos e nervos.** [s.l.] Editora Atheneu, 1995.

FILHO, Antonio F., RIBEIRO, Lair B. C. Exame Clínico. In: MONTORO, Antonio Franco. **Mastologia.** São Paulo, Sarvier, 1984.

HARDY, Ellen E.; PINOTTI, J. A.; FAUNDES, Anibal. Variações internacionais da incidência e da mortalidade por câncer de mama. **Femina - Revista da Federação Brasileira das Sociedades de Ginecologia e Obstetrícia,** Rio de Janeiro, v. 11, n. 2, p. 106-15, fev., 1983.

INCA, Instituto Nacional de Câncer - Ministério da Saúde. **O que é câncer ?** [on line], Documento disponível na internet via URL: <http://www.inca.org.br/cancer/cancer.html>. out., 1998a.

INCA, Instituto Nacional de Câncer - Ministério da Saúde. **Câncer de mama.** [on line], Documento disponível na internet via URL: <http://www.inca.org.br/cancer/tipos/mama.html>,

out., 1998b.

INCA, Instituto Nacional de Câncer - Ministério da Saúde. **Câncer de pulmão**. [on line], Documento disponível na internet via URL: <http://www.inca.org.br/cancer/tipos/pulmao.html>, out., 1998c.

LEIS JÚNIOR, Henry P., Fatores de risco e epidemiologia do carcinoma de mama. In: MONTORO, Antonio Franco. **Mastologia**. São Paulo, Sarvier, 1984.

LINS, Ronaldo M.; SILVA, Henrique M. S., Doppler colorido na propedêutica da mama. In: DIAS, Ezio N. et al. **Controvérsias em Mastologia**. Rio de Janeiro, Revinter, 1998.

LUNA, Marconi M. de, Significado das microcalcificações em mamografia. In: DIAS, Ezio N. et al. **Controvérsias em Mastologia**. Rio de Janeiro, Revinter, 1998

MENKE, Carlos H., et al. Falsos negativos da mamografia. **RBGO - Revista Brasileira de Ginecologia e Obstetrícia**, Rio de Janeiro, n. 1, p. 1-3, 1989.

MONTORO, Antonio F., Termografia Mamária. In: MONTORO, Antonio Franco. **Mastologia**. São Paulo, Sarvier, 1984.

NORA, James J.; FRASER, F. Clarke, **Genética Médica**. Rio de Janeiro, Editora Guanabara Koogan S.A., 1991.

OLIVEIRA, Laerte J. de; URBANETZ, Almir A.; HATSCHBACH, Sérgio B. B.; REHME, Marta F. B. Prevenção e detecção do câncer de mama. **Femina - Revista da Federação Brasileira das Sociedades de Ginecologia e Obstetrícia**, Rio de Janeiro, v. 19, n. 3, p. 203-8, mar., 1991.

PINOTTI, J. A.; PACHECO, J. S.; COLLIER, A. M. Termografia de placa. **Femina - Revista da Federação Brasileira das Sociedades de Ginecologia e Obstetrícia**, Rio de Janeiro, v. 14, n. 10, p. 985-91, nov., 1986.

PINHEIRO, José C. L. Câncer de mama feminina - Algumas considerações. **Femina - Revista da Federação Brasileira das Sociedades de Ginecologia e Obstetrícia**. Rio de Janeiro, v. 14, n. 1, p. 41-48, jan., 1986.

THOMPSON, James S.; THOMPSON, Margareth W. **Genética Médica**. Rio de Janeiro, Editora Guanabara, 1986.



ANEXO B - DICOM

A informatização da medicina levou ao desenvolvimento de padrões e protocolos específicos, dentre eles o DICOM ou *Digital Image Communications in Medicine* (Figura A1). O DICOM contempla uma série de requisitos necessários à avaliação mais acurada de uma imagem radiológica, diferentemente de uma imagem em formato convencional, uma imagem DICOM trás consigo informações adicionais que podem ser tanto *parâmetros de aquisição*, ou seja, parâmetros de regulagem do equipamento radiológico, como dados sobre o prontuário do paciente. Além disso o DICOM permite que essa imagem e suas informações associadas possam ser intercambiadas entre certos equipamentos radiológicos, computadores e hospitais. Nesse sentido o DICOM estabelece uma linguagem comum através da qual pode haver uma interoperabilidade entre equipamentos diversos de diferentes fabricantes.

O DICOM vem sendo desenvolvido desde meados da década de oitenta, por representantes de companhias que lidam com imagens médicas (NEMA - National Electrical Manufacturers Association) e importantes entidades que representam a comunidade médica, principalmente dos EUA, tais entidades foram: ACC - American College of Cardiology, ACR - American College of Radiology, ASE - Society of Echocardiography, ESC - European Society of Cardiology, ASNC - and American Society of Nuclear Cardiology. O padrão inicialmente desenvolvido foi denominado ACR/NEMA e só mais tarde recebeu a denominação DICOM.

DICOM contempla basicamente os aspectos seguintes:

I - Um protocolo de comunicação entre equipamentos radiológicos e softwares hospitalares;

II - Um formato de dados diferente para cada modalidade de imagem médica como tomografia computadorizada, raio X, ressonância magnética, etc;

III - Um protocolo de serviços, como por exemplo transmitir dados de exame ou procurar exames de paciente.

DICOM encontra-se hoje na versão 3.0 e é um padrão aberto, podendo ser estendido para englobar novos serviços ou tipos de imagens ou dados adicionais em um tipo já

existente. A utilização de DICOM não está sujeita a nenhum tipo de restrição de direitos ou royalties. A documentação do padrão pode ser encontrada no site do ACR, da NEMA ou RSNA ou adquirida em CD. Para um equipamento ou um produto de software ser considerado DICOM-compatível é necessária a certificação-DICOM, realizada anualmente.

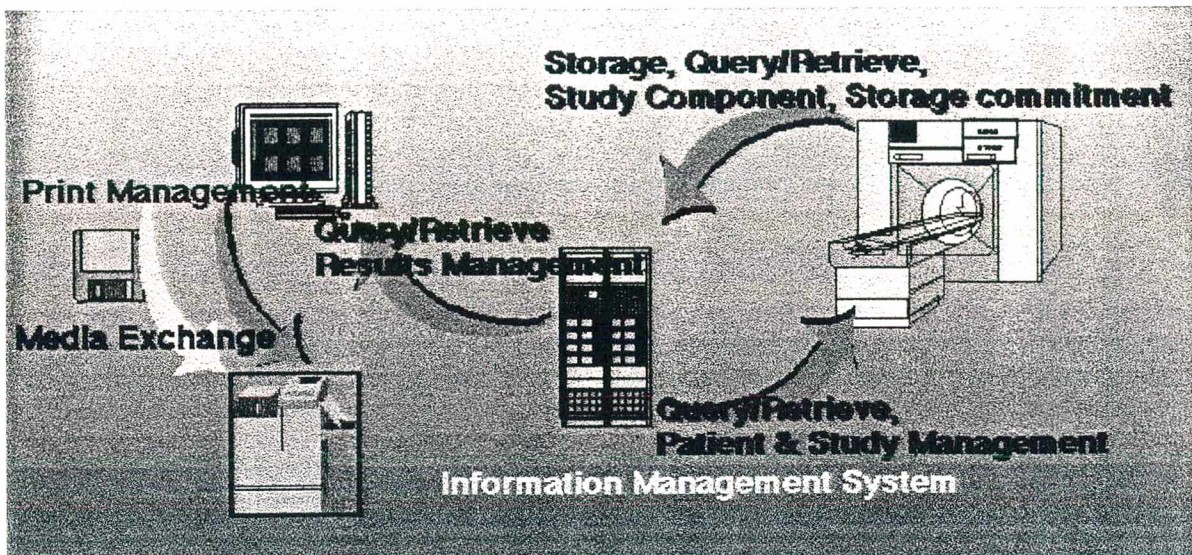


Figura A1: Domínio de aplicação DICOM

ANEXO C - DIAGRAMAS DE CLASSES

Segue nas figuras abaixo (Figura A2 e Figura A3) os diagramas das principais classes implementadas:

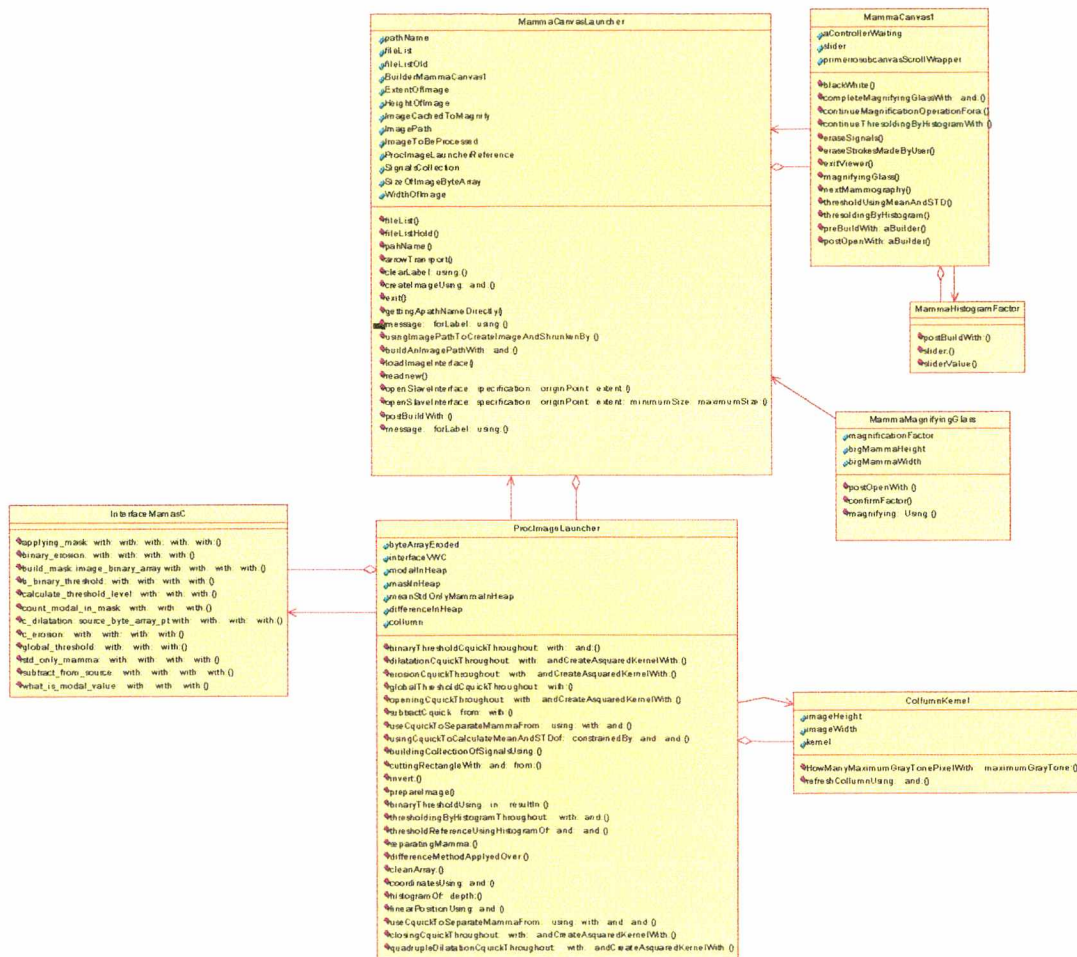


Figura A2: Principais classes do sistema

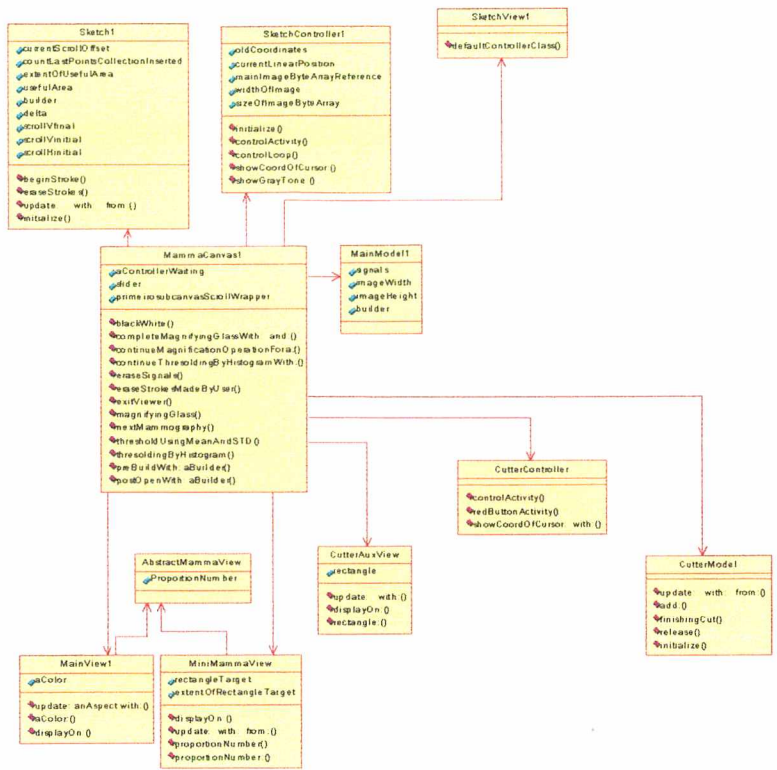


Figura A3: (continuação da figura A1)



ANEXO D - MANUAL DO USUÁRIO

1) Instalação

Por ainda se tratar de um protótipo a instalação do detector de microcalcificações exige algum conhecimento sobre o ambiente Visualworks, porém aqui não se entrará em detalhes quanto à instalação do ambiente nem configuração da imagem virtual, pois ao usuário cabe utilizar a interface não instalar o ambiente.

2) Partida do Sistema

O protótipo do sistema foi desenvolvido como uma imagem virtual do Visualworks 3, assim a execução do mesmo fica vinculada a este ambiente. Além disso deve existir mais três itens obrigatórios: 1) a biblioteca dinâmica (DLL) a qual chamou-se *mamas.dll*, onde estão todas as funções C de processamento de imagens; 2) o arquivo de prototipos das funções C, a qual chamou-se *mamas.h*, ambos arquivos devem estar obrigatoriamente no mesmo diretório que o arquivo imagem do sistema. Qualquer alteração na localização destes implica em mudança nos caminhos fornecidos no *builder* do *External Interface Finder* do ambiente Visualworks; e 3) para que o Visualworks possa utilizar a *mamas.dll* faz-se necessário instalar a biblioteca *cw3230.dll* no diretório *c:\windows\system*, uma das alternativas de obter tal biblioteca é o diretório *..\BC5\bin*, ou seja, na própria instalação do ambiente Borland.

Partindo-se do pressuposto que o Visualworks 3 está instalado corretamente, e que dos arquivos *mammaViewer.im* e *mammaViewer.cha*, arquivo imagem e arquivo de código fonte das classes e métodos da imagem, estejam também corretamente configurados (*settings* da imagem virtual). A partir da execução da imagem virtual, deve-se abrir a janela *workspace* (Figura A4 - A), na janela *workspace* deve-se digitar o nome da classe que lança a interface principal do protótipo do sistema (MammaCanvasLauncher) seguida pela mensagem *open*, conforme pode ser visto na figura abaixo, marca-se essa linha de comando (Figura A4 - B), com o botão direito do mouse clica-se na opção *do it* (Figura A4 - C), tal operação dará partida na interface que de fato lança a interface principal do sistema.

Inicialmente a interface de lançamento da janela principal se mostra como na

figura D abaixo (Figura A4 - D), ou seja, apenas um caminho default no disco (local) é exibido (este caminho é editável), para que sejam exibidos os arquivos imagem faz-se necessário que o usuário, sobre o campo onde se exibe o caminho no disco: clique uma vez com o botão esquerdo ou direito do mouse. Caso precise editar o caminho o usuário, para que a lista de arquivos se atualize, o usuário deve utilizar apenas um clique com o botão direito. Exibida a lista de arquivos obtém-se a aparência da Figura A5, para ser disparada a interface principal *Mammography Viewer*, basta selecionar um arquivo imagem na lista direita, com um clique no botão direito do mouse, e acionar o botão *load selected image*.

Na janela *Mammography Viewer* (Figura A6) efetivamente o usuário terá acesso aos métodos de detecção bem como algumas facilidades típicas de ferramentas para procesamento de imagens, tais como: (ii) a possibilidade de recortar trechos da imagem para magnificação, bastando para tanto clicar no ícone de uma lupa e recortar o trecho da imagem desejado; (ii) execução de riscos na tela de forma a marca-la quando conveniente, através do deslocamento do *mouse* com o botão esquerdo pressionado delimitando-se a área retangular pretendida; (iii) facilidades para apagamento das marcações feitas pelo usuário bem como dos sinais detectados pelos métodos de detecção; (iv) campos informativos sobre coordenadas do cursor e tom de cinza associado.



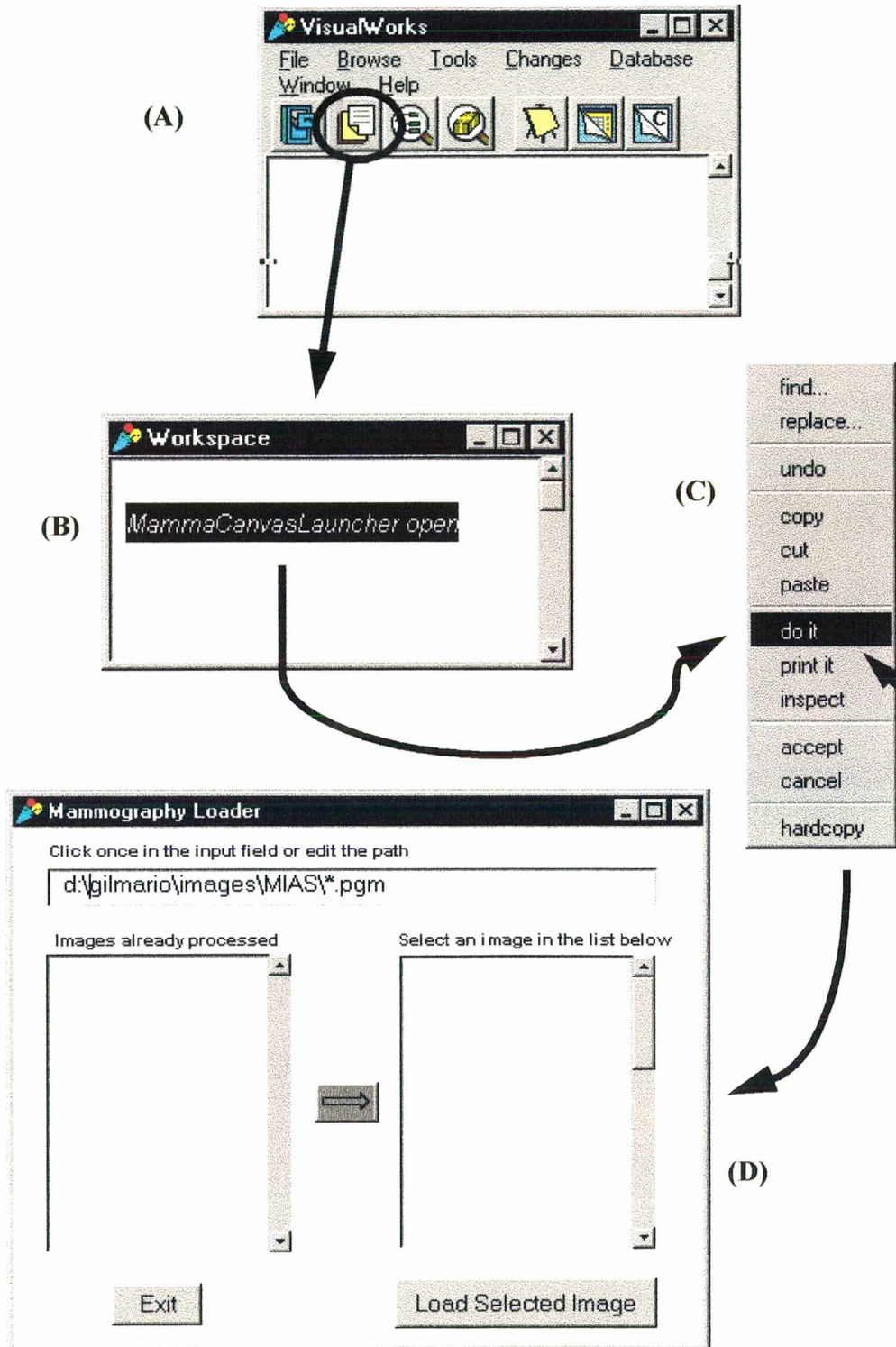


Figura A4: Acesso a janela workspace para abertura do launcher da interface principal do sistema

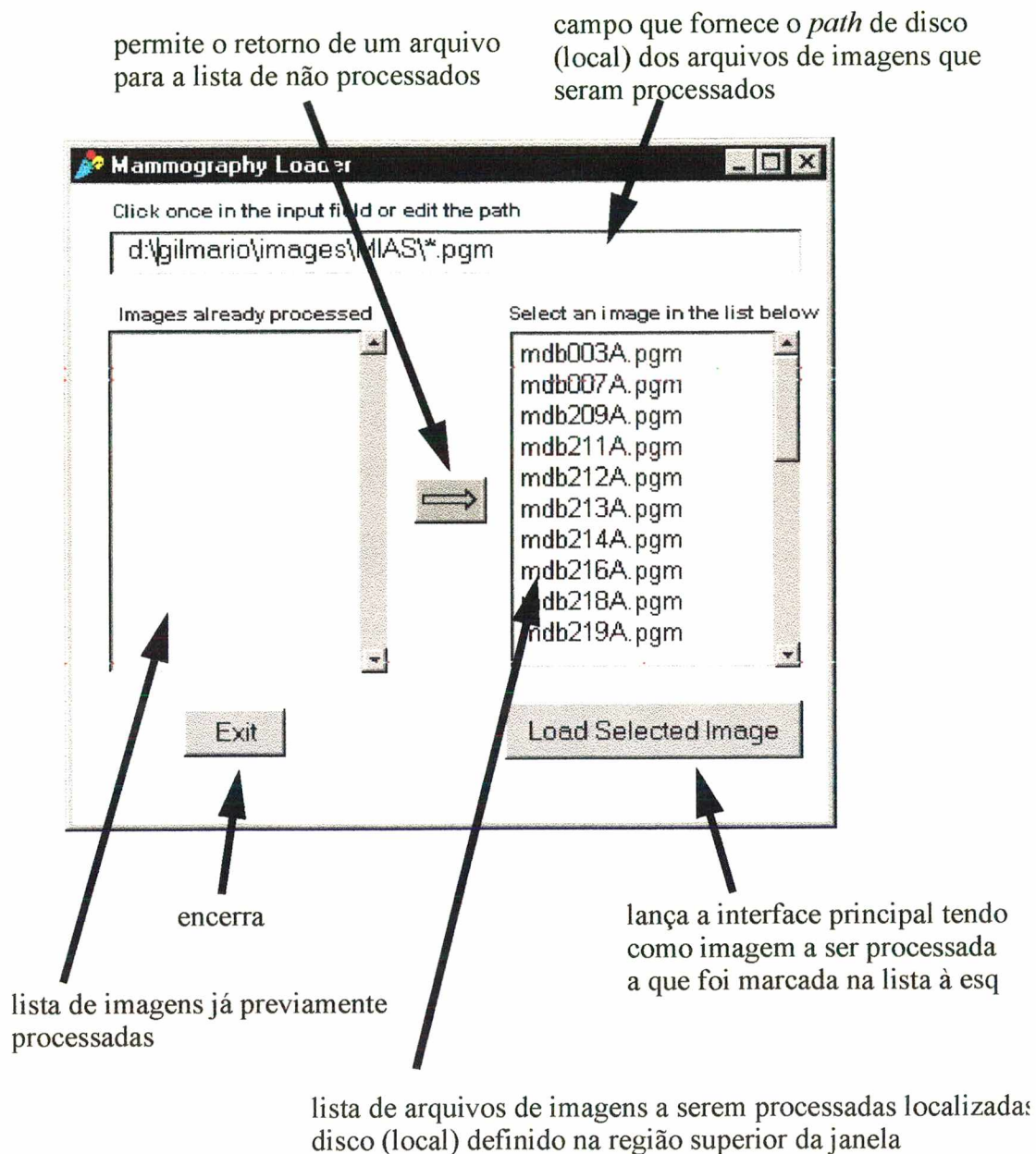
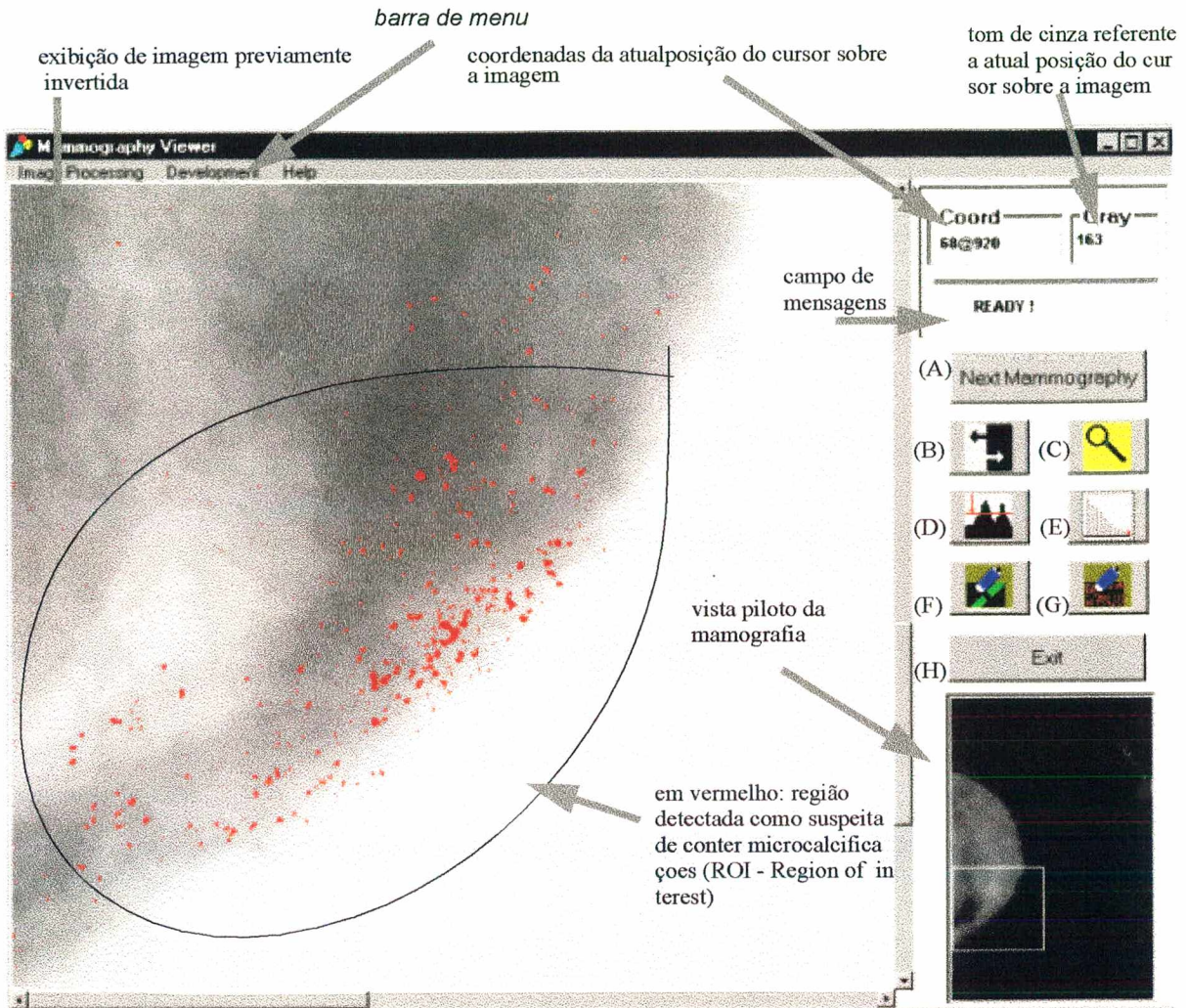


Figura A5: Mammography Loader - lançador da janela principal





descrição das funções dos botões:

- (A) Fecha janela atual e abre janela de lançamento para escolha de nova imagem a ser processada;
- (B) Inverte pixels da imagem que está sendo trabalhada;
- (C) Permite a magnificação de um trecho retangular recortado da imagem;
- (D) Executa o threshold global e destaca os sinais suspeitos em regiões amarelas;
- (E) Executa o Thrshold baseado em histograma, destaca os sinais detectados em vermelho;
- (F) Apaga rabiscos feios pelo usuário com o mouse;
- (G) Apaga marcas de regiões detectadas por D e E;
- (H) Fecha esta janela.

Figura A6: Mammography Viewer - janela principal e suas funções



Abaixo seguem duas janelas auxiliares: (i) a janela que permite ao usuário fornecer um valor que servirá como percentual a ser aplicado ao tamanho da imagem (número de pixels) de forma a encontrar no histograma de frequências o tom de cinza que servirá de limiar para o *threshold* baseado em histograma (Figura A7); (ii) a janela que exibe uma área recortada na imagem (pelo usuário) e que pode ser magnificada (Figura A8).

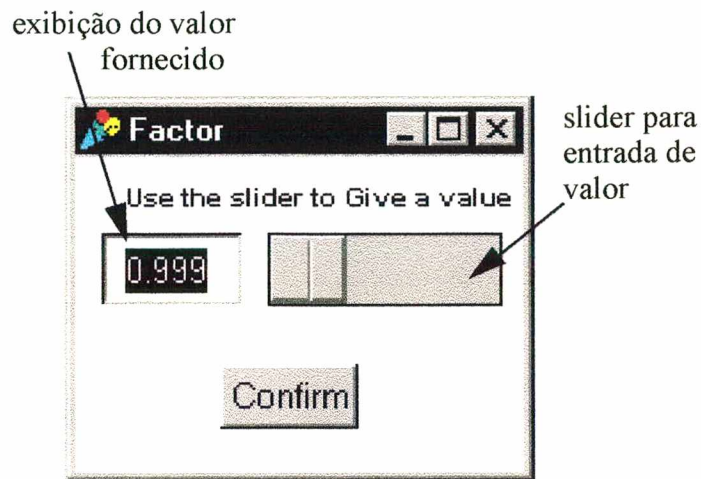


Figura A7: Entrada de fator para calculo de limiar por histograma



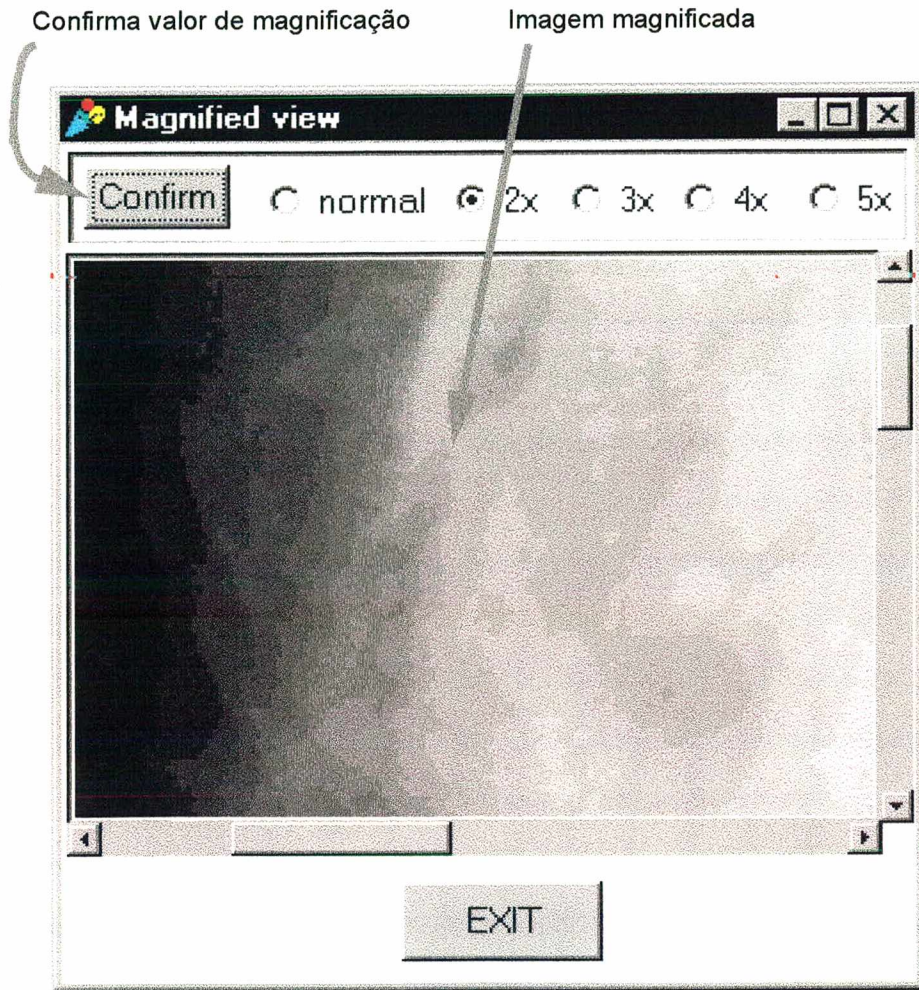


Figura A8: Tela de magnificação

ANEXO E - FRAGMENTOS DE CÓDIGO EM C

```
/*-----
UFSC - CTC - Projeto Cyclops
Função utilizada para erodir uma imagem em tons de cinza com profundidade 8 bits.

Autor: Gilmar B. dos Santos
-----*/
#include <stdlib.h>
#include <stdio.h>

__declspec(dllexport)
void erosion(unsigned char * source_byte_array_pt, unsigned char * result_byte_array_pt, unsigned int *imageWidth,
unsigned int *imageHeight, unsigned int *kernelDimension)
{
    unsigned int    kernel_size, k_dim, im_h, im_w, cont_lin, cont_col, cont;
    unsigned char    minimum_gray, current_gray_tone_image_A;
    long int         *kernel_pt, image_array_size, result_array_size, target_position_at_image_B,
current_kernel_origin_over_image_A, pixel_position_image_A;

    im_h = *imageHeight;
    im_w = *imageWidth;
    k_dim = *kernelDimension;
    kernel_size = (k_dim*k_dim);
    image_array_size = im_h*im_w;
    result_byte_array_pt = (unsigned char *) byte_array_cleaner(result_byte_array_pt, image_array_size);
    if ( ( kernel_pt = (long int *) malloc(sizeof(long int)*kernel_size) ) == NULL )
    { exit(1); }
    kernel_pt = long_integer_array_cleaner(kernel_pt, kernel_size);
    for(cont_lin=0; cont_lin < (im_h - (k_dim - 1)); cont_lin++)
    {
        for(cont_col=0; cont_col < im_w; cont_col++)
        {
            kernel_pt = refresh_squared_kernel_new(kernel_pt, cont_col, cont_lin, im_w, k_dim);

            if (kernel_pt[kernel_size-1] <= image_array_size)
            { for(cont=0; cont < kernel_size; cont++)
            {
                pixel_position_image_A = kernel_pt[cont];
                current_gray_tone_image_A = (unsigned char)source_byte_array_pt[pixel_position_image_A];
                if(cont == 0)
                    { minimum_gray = current_gray_tone_image_A; }
                else
                    if ( minimum_gray > current_gray_tone_image_A )
                        { minimum_gray = current_gray_tone_image_A; }
            }
            current_kernel_origin_over_image_A = linearPosition(cont_col, cont_lin, im_w);
            target_position_at_image_B = kernelCenterPositionAtImage( (k_dim/2), (k_dim/2),
current_kernel_origin_over_image_A, im_w);
            result_byte_array_pt[target_position_at_image_B] = (unsigned char) minimum_gray;
        }
    }
}
free(kernel_pt);
```

Código em C para erosão em tons de cinza


```

-----
/*
UFSC - CTC - Projeto Cyclops
Esta função devolve o label modal de uma imagem rotulada previamente

Autor: Gilmario B. dos Santos
-----
-*/
#include <stdio.h>
#include <stdlib.h>
__declspec(dllexport) void what_is_modal_value(unsigned int * modal_pt, unsigned int * array_labeled, unsigned int * image_width, unsigned int * image_height)
{
/* this function answers the modal label that is the label which is found more times in array_labeled. Valid labels are greater than zero*/
long int *label_counting_list, current_label_at_array_labeled, frequency_current_label, modal_label_frequency, count, array_labeled_size;
unsigned int label_quantity, i_wid, i_heigh, modal_label;
/* label_counting_list has one cell for each label, the first important assumption is that the number of different labels is <= 65353, in fact this number is sizeof(unsigned int). ->For save memory space this number was reduced to 1000.

The second assumption is about the maximum frequency (counting) of occurrences of any label, here this value is sizeof(long int)
*/
label_quantity = 1001; /*<- we suppose that the maximum number of different labels is 1000 zero is not considered a valid label because black gray tone is not a valid label*/

i_wid = *image_width;
i_heigh = *image_height;
array_labeled_size = (i_wid*i_heigh);
modal_label = 0; /* setting initial value for modal_label */
label_counting_list = (long int *) malloc(sizeof(long int)* label_quantity);
if ( label_counting_list == NULL )
{ exit(1); }
label_counting_list = long_integer_array_cleaner(label_counting_list, label_quantity);

for(count = 0; count < array_labeled_size; count++)
{ current_label_at_array_labeled = array_labeled[count];
/* count == 0 is used to start the looping */
if((current_label_at_array_labeled > 0)|| (count == 0))
{ frequency_current_label = label_counting_list[current_label_at_array_labeled];
frequency_current_label = frequency_current_label + 1;
label_counting_list[current_label_at_array_labeled] = frequency_current_label;

modal_label_frequency = label_counting_list[modal_label];
if(frequency_current_label > modal_label_frequency)
{ modal_label = current_label_at_array_labeled;
}
}
}
}
*modal_pt = modal_label;
free(label_counting_list);
}

```

Definição do label modal em uma imagem rotulada



```

/*-----
UFSC - CTC- Projeto Cyclops
Função que executa uma dilatação em uma imagem em tons de cinza
Autor: Gilmar B. dos Santos
-----*/

#include <stdlib.h>
#include <stdio.h>
__declspec(dllexport)
void dilatation(unsigned char * source_byte_array_pt, unsigned char * result_byte_array_pt, unsigned int
*imageWidth,unsignedint *imageHeight, unsigned int *kernelDimension)
{ unsigned int kernel_size, k_dim,im_h,im_w,
cont_lin, cont_col, cont;
unsigned char maximum_gray, current_gray_tone_image_A;
long int *kernel_pt, image_array_size,
target_position_at_image_B, current_kernel_origin_over_image_A, pixel_position_image_A;
im_h = *imageHeight;
im_w = *imageWidth;
k_dim = *kernelDimension;
kernel_size = (k_dim*k_dim);
image_array_size = im_h*im_w;
result_byte_array_pt = byte_array_cleaner(result_byte_array_pt,image_array_size);
if ( ( kernel_pt = (long int *) malloc(sizeof(long int)*kernel_size) ) == NULL )
{ exit(1); }
kernel_pt = long_integer_array_cleaner(kernel_pt,kernel_size);
for(cont_lin=0;cont_lin < (im_h -(k_dim -1));cont_lin++)
{
for(cont_col=0; cont_col < im_w; cont_col++)
{ kernel_pt = refresh_squared_kernel_new(kernel_pt,cont_col, cont_lin, im_w, k_dim);
if(kernel_pt[kernel_size-1] <= image_array_size)
{
for(cont=0; cont < kernel_size; cont++)
{
if(cont == 0) { maximum_gray = '\0'; }
pixel_position_image_A = kernel_pt[cont];
current_gray_tone_image_A = (unsigned char) source_byte_array_pt[pixel_position_image_A];
if(cont == 0)
{maximum_gray = current_gray_tone_image_A;}
else
if ( maximum_gray < current_gray_tone_image_A)
{ maximum_gray = current_gray_tone_image_A; }
}
current_kernel_origin_over_image_A = linearPosition(cont_col,cont_lin,im_w);
target_position_at_image_B = kernelCenterPositionAtImage( (k_dim/2),(k_dim/2),
current_kernel_origin_over_image_A,im_w);
result_byte_array_pt[target_position_at_image_B]= (unsigned char) maximum_gray;
}
}
}
free(kernel_pt);
}

```

Dilatação em tons de cinza



```

/*-----
-
UFSC - CTC - Projeto Cyclops
Função que executa um threshold global, baseado na média e desvio padrão da região
da mama.

Autor: Gilmário B. dos Santos
-----*
/
#include <stdio.h>
#include <stdlib.h>
__declspec(dllexport) void global_threshold(unsigned char * source_byte_array, unsigned
char * result_byte_array_pt, long int * source_byte_array_size,
                mean_std_only_mamma * mean_std)

{ unsigned int temp, temp1;
  long int count;
  float mean_gray, top_hat_height, std_gray;
  mean_gray = (mean_std->mean_gray);
  std_gray = (mean_std->std_gray);
  top_hat_height = 3.0*(mean_gray + std_gray);

for(count = 0; count < (*source_byte_array_size); count++)
{ temp = source_byte_array[count];
  temp1 = ((unsigned int)top_hat_height);
  if(temp > temp1 )
    result_byte_array_pt[count] = (unsigned char) 255;
  else
    result_byte_array_pt[count] = '\0';
}
}

```

Threshold global



ANEXO F - Validação

Conteúdo do arquivo de teste:

formato dos registros:

vector (class abnormality, center abnormality, radius, pixels detected into circle enclosing abnormality, total pixels detected)

Os fatores considerados, *fator hist* e *fator std*, são valores numéricos utilizados para definir respectivamente: o percentual de "corte" do histograma e o multiplicador da expressão que determina o limiar global, conforme explicado em 4.4.1 e 4.4.2.

Threshold por histograma

fator hist = 0,9000

mdb003A.pgm	ThresholdHist	(NORMAL, '', '', '', 66296)
mdb007A.pgm	ThresholdHist	(NORMAL, '', '', '', 48292)
mdb209A.pgm	ThresholdHist	(CALCIFICATIONS, 647@503, 87, 5193, 69343)
mdb211A.pgm	ThresholdHist	(CALCIFICATIONS, 680@327, 13, 197, 62910)
mdb212A.pgm	ThresholdHist	(CALCIFICATIONS, 687@882, 3, 0, 56407)
mdb213A.pgm	ThresholdHist	(CALCIFICATIONS, 547@520, 45, 2510, 58804)
mdb214A.pgm	ThresholdHist	(CALCIFICATIONS, 582@916, 11, 0, 59474)
mdb218A.pgm	ThresholdHist	(CALCIFICATIONS, 519@629, 8, 53, 68047)
mdb219A.pgm	ThresholdHist	(CALCIFICATIONS, 546@756, 29, 444, 55651)
mdb222A.pgm	ThresholdHist	(CALCIFICATIONS, 398@427, 17, 211, 48961)
mdb227A.pgm	ThresholdHist	(CALCIFICATIONS, 504@467, 9, 83, 62814)
mdb231A.pgm	ThresholdHist	(CALCIFICATIONS, 603@538, 44, 1165, 57880)
mdb236A.pgm	ThresholdHist	(CALCIFICATIONS, 276@824, 14, 66, 51229)
mdb238A.pgm	ThresholdHist	(CALCIFICATIONS, 522@553, 17, 306, 64396)
mdb240A.pgm	ThresholdHist	(CALCIFICATIONS, 643@614, 23, 31, 44360)
mdb241A.pgm	ThresholdHist	(CALCIFICATIONS, 453@678, 38, 1110, 40327)
mdb248A.pgm	ThresholdHist	(CALCIFICATIONS, 378@601, 10, 127, 47101)
mdb252A.pgm	ThresholdHist	(CALCIFICATIONS, 439@367, 23, 299, 53005)
mdb253A.pgm	ThresholdHist	(CALCIFICATIONS, 733@564, 28, 246, 50445)
mdb256A.pgm	ThresholdHist	(CALCIFICATIONS, 400@484, 37, 561, 41164)

fator hist = 0.9990

mdb003A.pgm	ThresholdHist	(NORMAL, '', '', '', 614)
mdb007A.pgm	ThresholdHist	(NORMAL, '', '', '', 673)
mdb209A.pgm	ThresholdHist	(CALCIFICATIONS, 647@503, 87, 28, 887)
mdb211A.pgm	ThresholdHist	(CALCIFICATIONS, 680@327, 13, 6, 681)
mdb212A.pgm	ThresholdHist	(CALCIFICATIONS, 687@882, 3, 0, 852)
mdb213A.pgm	ThresholdHist	(CALCIFICATIONS, 547@520, 45, 70, 850)
mdb214A.pgm	ThresholdHist	(CALCIFICATIONS, 582@916, 11, 0, 876)
mdb218A.pgm	ThresholdHist	(CALCIFICATIONS, 519@629, 8, 0, 997)
mdb219A.pgm	ThresholdHist	(CALCIFICATIONS, 546@756, 29, 62, 928)
mdb222A.pgm	ThresholdHist	(CALCIFICATIONS, 398@427, 17, 4, 964)
mdb227A.pgm	ThresholdHist	(CALCIFICATIONS, 504@467, 9, 3, 835)
mdb231A.pgm	ThresholdHist	(CALCIFICATIONS, 603@538, 44, 2, 883)
mdb236A.pgm	ThresholdHist	(CALCIFICATIONS, 276@824, 14, 0, 1017)
mdb238A.pgm	ThresholdHist	(CALCIFICATIONS, 522@553, 17, 3, 998)
mdb240A.pgm	ThresholdHist	(CALCIFICATIONS, 643@614, 23, 0, 782)
mdb241A.pgm	ThresholdHist	(CALCIFICATIONS, 453@678, 38, 82, 776)
mdb248A.pgm	ThresholdHist	(CALCIFICATIONS, 378@601, 10, 52, 706)
mdb252A.pgm	ThresholdHist	(CALCIFICATIONS, 439@367, 23, 38, 875)
mdb253A.pgm	ThresholdHist	(CALCIFICATIONS, 733@564, 28, 0, 746)
mdb256A.pgm	ThresholdHist	(CALCIFICATIONS, 400@484, 37, 10, 811)

fator hist = 0.9996

mdb003A.pgm	ThresholdHist	(NORMAL, '', '', '', 326)
mdb007A.pgm	ThresholdHist	(NORMAL, '', '', '', 371)
mdb209A.pgm	ThresholdHist	(CALCIFICATIONS, 647@503,87,0,412)
mdb211A.pgm	ThresholdHist	(CALCIFICATIONS, 680@327,13,0,416)
mdb212A.pgm	ThresholdHist	(CALCIFICATIONS, 687@882,3,0,414)
mdb213A.pgm	ThresholdHist	(CALCIFICATIONS, 547@520,45,0,418)
mdb214A.pgm	ThresholdHist	(CALCIFICATIONS, 582@916,11,0,397)
mdb218A.pgm	ThresholdHist	(CALCIFICATIONS, 519@629,8,0,402)
mdb219A.pgm	ThresholdHist	(CALCIFICATIONS, 546@756,29,23,406)
mdb222A.pgm	ThresholdHist	(CALCIFICATIONS, 398@427,17,0,413)
mdb227A.pgm	ThresholdHist	(CALCIFICATIONS, 504@467,9,0,368)
mdb231A.pgm	ThresholdHist	(CALCIFICATIONS, 603@538,44,0,369)
mdb236A.pgm	ThresholdHist	(CALCIFICATIONS, 276@824,14,0,389)
mdb238A.pgm	ThresholdHist	(CALCIFICATIONS, 522@553,17,0,393)
mdb240A.pgm	ThresholdHist	(CALCIFICATIONS, 643@614,23,0,398)
mdb241A.pgm	ThresholdHist	(CALCIFICATIONS, 453@678,38,6,404)
mdb248A.pgm	ThresholdHist	(CALCIFICATIONS, 378@601,10,33,398)
mdb252A.pgm	ThresholdHist	(CALCIFICATIONS, 439@367,23,0,415)
mdb253A.pgm	ThresholdHist	(CALCIFICATIONS, 733@564,28,0,403)
mdb256A.pgm	ThresholdHist	(CALCIFICATIONS, 400@484,37,0,418)

fator hist = 0,9940

mdb003A.pgm	ThresholdHist	(NORMAL, '', '', '', 2710)
mdb007A.pgm	ThresholdHist	(NORMAL, '', '', '', 3495)
mdb209A.pgm	ThresholdHist	(CALCIFICATIONS, 647@503,87,239,3501)
mdb211A.pgm	ThresholdHist	(CALCIFICATIONS, 680@327,13,31,3599)
mdb212A.pgm	ThresholdHist	(CALCIFICATIONS, 687@882,3,0,3692)
mdb213A.pgm	ThresholdHist	(CALCIFICATIONS, 547@520,45,331,3309)
mdb214A.pgm	ThresholdHist	(CALCIFICATIONS, 582@916,11,0,3008)
mdb218A.pgm	ThresholdHist	(CALCIFICATIONS, 519@629,8,7,4655)
mdb219A.pgm	ThresholdHist	(CALCIFICATIONS, 546@756,29,114,3959)
mdb222A.pgm	ThresholdHist	(CALCIFICATIONS, 398@427,17,11,3983)
mdb227A.pgm	ThresholdHist	(CALCIFICATIONS, 504@467,9,19,3232)
mdb231A.pgm	ThresholdHist	(CALCIFICATIONS, 603@538,44,101,4211)
mdb236A.pgm	ThresholdHist	(CALCIFICATIONS, 276@824,14,15,5088)
mdb238A.pgm	ThresholdHist	(CALCIFICATIONS, 522@553,17,25,3244)
mdb240A.pgm	ThresholdHist	(CALCIFICATIONS, 643@614,23,3,3466)
mdb241A.pgm	ThresholdHist	(CALCIFICATIONS, 453@678,38,294,3265)
mdb248A.pgm	ThresholdHist	(CALCIFICATIONS, 378@601,10,78,2894)
mdb252A.pgm	ThresholdHist	(CALCIFICATIONS, 439@367,23,77,2934)
mdb253A.pgm	ThresholdHist	(CALCIFICATIONS, 733@564,28,22,3172)
mdb256A.pgm	ThresholdHist	(CALCIFICATIONS, 400@484,37,62,3434)

fator hist = 0,9900

mdb003A.pgm	ThresholdHist	(NORMAL, '', '', '', 6646)
mdb007A.pgm	ThresholdHist	(NORMAL, '', '', '', 5298)
mdb209A.pgm	ThresholdHist	(CALCIFICATIONS, 647@503,87,490,6996)
mdb211A.pgm	ThresholdHist	(CALCIFICATIONS, 680@327,13,46,5760)
mdb212A.pgm	ThresholdHist	(CALCIFICATIONS, 687@882,3,0,5458)
mdb213A.pgm	ThresholdHist	(CALCIFICATIONS, 547@520,45,593,6829)
mdb214A.pgm	ThresholdHist	(CALCIFICATIONS, 582@916,11,0,6335)
mdb218A.pgm	ThresholdHist	(CALCIFICATIONS, 519@629,8,11,6883)
mdb219A.pgm	ThresholdHist	(CALCIFICATIONS, 546@756,29,133,7067)
mdb222A.pgm	ThresholdHist	(CALCIFICATIONS, 398@427,17,21,5736)
mdb227A.pgm	ThresholdHist	(CALCIFICATIONS, 504@467,9,26,4734)
mdb231A.pgm	ThresholdHist	(CALCIFICATIONS, 603@538,44,154,5771)
mdb236A.pgm	ThresholdHist	(CALCIFICATIONS, 276@824,14,17,7824)
mdb238A.pgm	ThresholdHist	(CALCIFICATIONS, 522@553,17,37,4687)
mdb240A.pgm	ThresholdHist	(CALCIFICATIONS, 643@614,23,3,4950)
mdb241A.pgm	ThresholdHist	(CALCIFICATIONS, 453@678,38,393,5030)
mdb248A.pgm	ThresholdHist	(CALCIFICATIONS, 378@601,10,83,4372)
mdb252A.pgm	ThresholdHist	(CALCIFICATIONS, 439@367,23,98,6316)
mdb253A.pgm	ThresholdHist	(CALCIFICATIONS, 733@564,28,39,4700)
mdb256A.pgm	ThresholdHist	(CALCIFICATIONS, 400@484,37,84,5163)



Threshold or estatísticas globais**fator std = 3,00**

mdb003A.pgm	ThresholdSTD:	(NORMAL, '', '', '', 257)
mdb007A.pgm	ThresholdSTD:	(NORMAL, '', '', '', 353)
mdb209A.pgm	ThresholdSTD:	(CALCIFICATIONS, 647@503, 87, 6, 582)
mdb211A.pgm	ThresholdSTD:	(CALCIFICATIONS, 680@327, 13, 0, 420)
mdb212A.pgm	ThresholdSTD:	(CALCIFICATIONS, 687@882, 3, 0, 567)
mdb213A.pgm	ThresholdSTD:	(CALCIFICATIONS, 547@520, 45, 4, 583)
mdb214A.pgm	ThresholdSTD:	(CALCIFICATIONS, 582@916, 11, 0, 396)
mdb218A.pgm	ThresholdSTD:	(CALCIFICATIONS, 519@629, 8, 0, 1703)
mdb219A.pgm	ThresholdSTD:	(CALCIFICATIONS, 546@756, 29, 62, 928)
mdb222A.pgm	ThresholdSTD:	(CALCIFICATIONS, 398@427, 17, 4, 768)
mdb227A.pgm	ThresholdSTD:	(CALCIFICATIONS, 504@467, 9, 0, 367)
mdb231A.pgm	ThresholdSTD:	(CALCIFICATIONS, 603@538, 44, 11, 1365)
mdb236A.pgm	ThresholdSTD:	(CALCIFICATIONS, 276@824, 14, 0, 1999)
mdb238A.pgm	ThresholdSTD:	(CALCIFICATIONS, 522@553, 17, 0, 917)
mdb240A.pgm	ThresholdSTD:	(CALCIFICATIONS, 643@614, 23, 0, 520)
mdb241A.pgm	ThresholdSTD:	(CALCIFICATIONS, 453@678, 38, 16, 458)
mdb248A.pgm	ThresholdSTD:	(CALCIFICATIONS, 378@601, 10, 37, 434)
mdb252A.pgm	ThresholdSTD:	(CALCIFICATIONS, 439@367, 23, 15, 610)
mdb253A.pgm	ThresholdSTD:	(CALCIFICATIONS, 733@564, 28, 0, 923)
mdb256A.pgm	ThresholdSTD:	(CALCIFICATIONS, 400@484, 37, 1, 454)

fator std = 2.6

mdb003A.pgm	ThresholdSTD:	(NORMAL, '', '', '', 285)
mdb007A.pgm	ThresholdSTD:	(NORMAL, '', '', '', 439)
mdb209A.pgm	ThresholdSTD:	(CALCIFICATIONS, 647@503, 87, 10, 662)
mdb211A.pgm	ThresholdSTD:	(CALCIFICATIONS, 680@327, 13, 0, 432)
mdb212A.pgm	ThresholdSTD:	(CALCIFICATIONS, 687@882, 3, 0, 634)
mdb213A.pgm	ThresholdSTD:	(CALCIFICATIONS, 547@520, 45, 10, 612)
mdb214A.pgm	ThresholdSTD:	(CALCIFICATIONS, 582@916, 11, 0, 483)
mdb218A.pgm	ThresholdSTD:	(CALCIFICATIONS, 519@629, 8, 0, 2077)
mdb219A.pgm	ThresholdSTD:	(CALCIFICATIONS, 546@756, 29, 72, 1184)
mdb222A.pgm	ThresholdSTD:	(CALCIFICATIONS, 398@427, 17, 4, 903)
mdb227A.pgm	ThresholdSTD:	(CALCIFICATIONS, 504@467, 9, 0, 448)
mdb231A.pgm	ThresholdSTD:	(CALCIFICATIONS, 603@538, 44, 28, 2019)
mdb236A.pgm	ThresholdSTD:	(CALCIFICATIONS, 276@824, 14, 0, 2090)
mdb238A.pgm	ThresholdSTD:	(CALCIFICATIONS, 522@553, 17, 3, 998)
mdb240A.pgm	ThresholdSTD:	(CALCIFICATIONS, 643@614, 23, 0, 606)
mdb241A.pgm	ThresholdSTD:	(CALCIFICATIONS, 453@678, 38, 25, 499)
mdb248A.pgm	ThresholdSTD:	(CALCIFICATIONS, 378@601, 10, 46, 604)
mdb252A.pgm	ThresholdSTD:	(CALCIFICATIONS, 439@367, 23, 30, 745)
mdb253A.pgm	ThresholdSTD:	(CALCIFICATIONS, 733@564, 28, 8, 1560)
mdb256A.pgm	ThresholdSTD:	(CALCIFICATIONS, 400@484, 37, 3, 498)

fator std = 3,60

mdb003A.pgm	ThresholdSTD:	(NORMAL, '', '', '', 242)
mdb007A.pgm	ThresholdSTD:	(NORMAL, '', '', '', 326)
mdb209A.pgm	ThresholdSTD:	(CALCIFICATIONS, 647@503, 87, 0, 546)
mdb211A.pgm	ThresholdSTD:	(CALCIFICATIONS, 680@327, 13, 0, 406)
mdb212A.pgm	ThresholdSTD:	(CALCIFICATIONS, 687@882, 3, 0, 487)
mdb213A.pgm	ThresholdSTD:	(CALCIFICATIONS, 547@520, 45, 1, 552)
mdb214A.pgm	ThresholdSTD:	(CALCIFICATIONS, 582@916, 11, 0, 335)
mdb218A.pgm	ThresholdSTD:	(CALCIFICATIONS, 519@629, 8, 0, 1231)
mdb219A.pgm	ThresholdSTD:	(CALCIFICATIONS, 546@756, 29, 53, 760)
mdb222A.pgm	ThresholdSTD:	(CALCIFICATIONS, 398@427, 4, 0, 637)
mdb227A.pgm	ThresholdSTD:	(CALCIFICATIONS, 504@467, 9, 0, 312)
mdb231A.pgm	ThresholdSTD:	(CALCIFICATIONS, 603@538, 44, 2, 790)
mdb236A.pgm	ThresholdSTD:	(CALCIFICATIONS, 276@824, 14, 0, 1892)
mdb238A.pgm	ThresholdSTD:	(CALCIFICATIONS, 522@553, 17, 0, 779)
mdb240A.pgm	ThresholdSTD:	(CALCIFICATIONS, 643@614, 23, 0, 491)
mdb241A.pgm	ThresholdSTD:	(CALCIFICATIONS, 453@678, 38, 8, 415)
mdb248A.pgm	ThresholdSTD:	(CALCIFICATIONS, 378@601, 10, 26, 362)
mdb252A.pgm	ThresholdSTD:	(CALCIFICATIONS, 439@367, 23, 7, 535)
mdb253A.pgm	ThresholdSTD:	(CALCIFICATIONS, 733@564, 28, 0, 565)
mdb256A.pgm	ThresholdSTD:	(CALCIFICATIONS, 400@484, 37, 0, 427)



fator std = 2,00

```

mdb003A.pgm      ThresholdSTD: (NCRMAL, '', '', '', 614)
mdb007A.pgm      ThresholdSTD: (NORMAL, '', '', '', 1133)
mdb209A.pgm      ThresholdSTD: (CALCIFICATIONS, 647@503, 87, 41, 1231)
mdb211A.pgm      ThresholdSTD: (CALCIFICATIONS, 680@327, 13, 1, 523)
mdb212A.pgm      ThresholdSTD: (CALCIFICATIONS, 687@882, 3, 0, 852)
mdb213A.pgm      ThresholdSTD: (CALCIFICATIONS, 547@520, 45, 45, 742)
mdb214A.pgm      ThresholdSTD: (CALCIFICATIONS, 582@916, 11, 0, 696)
mdb218A.pgm      ThresholdSTD: (CALCIFICATIONS, 519@629, 8, 2, 3244)
mdb219A.pgm      ThresholdSTD: (CALCIFICATIONS, 546@756, 29, 101, 2425)
mdb222A.pgm      ThresholdSTD: (CALCIFICATIONS, 398@427, 17, 5, 1428)
mdb227A.pgm      ThresholdSTD: (CALCIFICATIONS, 504@467, 9, 3, 834)
mdb231A.pgm      ThresholdSTD: (CALCIFICATIONS, 603@538, 44, 154, 5771)
mdb236A.pgm      ThresholdSTD: (CALCIFICATIONS, 276@824, 14, 0, 2563)
mdb238A.pgm      ThresholdSTD: (CALCIFICATIONS, 522@553, 17, 9, 1304)
mdb240A.pgm      ThresholdSTD: (CALCIFICATIONS, 643@614, 23, 2, 1114)
mdb241A.pgm      ThresholdSTD: (CALCIFICATIONS, 453@678, 38, 65, 690)
mdb248A.pgm      ThresholdSTD: (CALCIFICATIONS, 378@601, 10, 65, 1421)
mdb252A.pgm      ThresholdSTD: (CALCIFICATIONS, 439@367, 23, 50, 1279)
mdb253A.pgm      ThresholdSTD: (CALCIFICATIONS, 733@564, 28, 39, 4700)
mdb256A.pgm      ThresholdSTD: (CALCIFICATIONS, 400@484, 37, 10, 812)

```

fator std = 1,5

```

mdb003A.pgm      ThresholdSTD: (NORMAL, '', '', '', 2710)
mdb007A.pgm      ThresholdSTD: (NORMAL, '', '', '', 3495)
mdb209A.pgm      ThresholdSTD: (CALCIFICATIONS, 647@503, 87, 239, 3501)
mdb211A.pgm      ThresholdSTD: (CALCIFICATIONS, 680@327, 13, 22, 1535)
mdb212A.pgm      ThresholdSTD: (CALCIFICATIONS, 687@882, 3, 0, 1865)
mdb213A.pgm      ThresholdSTD: (CALCIFICATIONS, 547@520, 45, 127, 1232)
mdb214A.pgm      ThresholdSTD: (CALCIFICATIONS, 582@916, 11, 0, 1290)
mdb218A.pgm      ThresholdSTD: (CALCIFICATIONS, 519@629, 8, 9, 5529)
mdb219A.pgm      ThresholdSTD: (CALCIFICATIONS, 546@756, 29, 121, 5157)
mdb222A.pgm      ThresholdSTD: (CALCIFICATIONS, 398@427, 17, 7, 2898)
mdb227A.pgm      ThresholdSTD: (CALCIFICATIONS, 504@467, 9, 14, 2287)
mdb231A.pgm      ThresholdSTD: (CALCIFICATIONS, 603@538, 44, 431, 16954)
mdb236A.pgm      ThresholdSTD: (CALCIFICATIONS, 276@824, 14, 7, 3558)
mdb238A.pgm      ThresholdSTD: (CALCIFICATIONS, 522@553, 17, 25, 3244)
mdb240A.pgm      ThresholdSTD: (CALCIFICATIONS, 643@614, 23, 3, 3466)
mdb241A.pgm      ThresholdSTD: (CALCIFICATIONS, 453@678, 38, 156, 1288)
mdb248A.pgm      ThresholdSTD: (CALCIFICATIONS, 378@601, 10, 83, 4372)
mdb252A.pgm      ThresholdSTD: (CALCIFICATIONS, 439@367, 23, 86, 4263)
mdb253A.pgm      ThresholdSTD: (CALCIFICATIONS, 733@564, 28, 73, 11171)
mdb256A.pgm      ThresholdSTD: (CALCIFICATIONS, 400@484, 37, 38, 2348)

```

Fragmento de código do método implementado para criação do arquivo de teste acima:

```

.....
.....
aux := super builder window application.
mainModell := self modelOfComponentNumberTwoOn: #primeirosbancas.
mainModell1 := self modelOfComponentNumberThreeOn: #primeirosbancas.
total :=      mainModell signals size.
circleCenter := (aux xCenterAbnormal value) @ (aux yCenterAbnormal value).
circleRadius := aux radiusAbnormal value.

circleCenter ~= (0@0)
    ifTrue: [      howMuchIntoTheCircle := mainModell howMuchIntoACircleWith:
circleRadius and: circleCenter.

```



```

        howMuchIntoTheCircle11 := mainModel11 howMuchIntoACircleWith:
circleRadius and: circleCenter.

```

```

]

```

```

ifFalse: [
    howMuchIntoTheCircle := '' .
    howMuchIntoTheCircle11 := '' .
    circleCenter := '' .
    circleRadius := '' ].

```

```

total11 := mainModel11 signals size.

```

```

(aux classAbnormal value = 0)

```

```

    ifTrue: [ class := 'NORMAL' ]

```

```

    ifFalse: [class := 'CALCIFICATIONS'].

```

```

LogFile

```

```

    nextPutAll:'ThresholdSTD: ';nextPutAll:'('; nextPutAll: class;nextPutAll:',';

```

```

nextPutAll: circleCenter printString;nextPutAll:',';

```

```

nextPutAll: circleRadius printString;nextPutAll:','; nextPutAll:howMuchIntoTheCir-
cle;

```

```

printString;nextPutAll:',';      nextPutAll: total printString; nextPutAll: ')';

```

```

cr;

```

```

    nextPutAll:'ThresholdHist  ';nextPutAll:'(';nextPutAll: class;nextPutAll:',';

```

```

nextPutAll: circleCenter printString;nextPutAll:',';nextPutAll:

```

```

circleRadius printString;nextPutAll:',';nextPutAll: howMuchIntoTheCircle11 print-
String;nextPutAll:',';nextPutAll: total11 printString; nextPutAll: ')'; cr.

```

