

DETECÇÃO DE AGRUPAMENTOS DE MICROCALCIFICAÇÕES EM IMAGENS MAMOGRAFICAS DIGITAIS UTILIZANDO ETAPAS DE SEGMENTAÇÃO DA MAMA E REALCE

VINICIUS RUELA PEREIRA BORGES¹, DENISE GULIATO²

RESUMO

Este artigo propõe o desenvolvimento de um novo método para a detecção de agrupamentos de microcalcificações em imagens mamográficas digitais. Este método é dividido em três etapas: segmentação da mama, realce de microcalcificações e detecção de agrupamentos de microcalcificações. A primeira etapa é composta por um procedimento que delimita a região da mama e outro que extrai o músculo peitoral, gerando uma nova imagem contendo apenas o tecido fibroglandular. A segunda etapa tem como objetivo realçar estas microcalcificações, já que estas não estão nítidas na imagem mamográfica. A última etapa segmenta as microcalcificações da imagem realçada e analisa se há ocorrência de agrupamentos, o que pode indicar ou não, um sinal precoce de câncer de mama. O método proposto foi avaliado na base de dados Digital Database for Screening Mammography (DDSM) e mostrou que os resultados obtidos são bastante promissores, o que pode contribuir positivamente com o aumento da sensibilidade nos exames de radiologia.

Palavras-chave: *diagnóstico auxiliado por computador, segmentação da mama, realce de microcalcificações, detecção de agrupamentos de microcalcificações.*

¹ Faculdade de Computação, Universidade Federal de Uberlândia, Campus Santa Mônica
Av. João Naves de Ávila, 2121, 38400-902 Uberlândia, MG, Brasil
viniciusrpb@gmail.com

² Orientadora, Faculdade de Computação, Universidade Federal de Uberlândia, Campus Santa Mônica
Av. João Naves de Ávila, 2121, 38400-902 Uberlândia, MG, Brasil
guliato@ufu.br

ABSTRACT

This paper proposes the development of a novel method for microcalcifications cluster detection in digital mammographic images. This method is divided in three steps: breast segmentation, microcalcifications enhance and microcalcifications clusters detection. The first step is composed by a procedure that delimits the mama region and other that extracts the pectoral muscle, creating a new image containing just the fibroglandular tissue. The second step aims to enhance these microcalcifications, since they are not clear in image. The last step segments the microcalcifications of the enhanced image and verify whether there is occurrence of clusters, which may indicate or not an earlier signal of breast cancer. The proposed method was tested in Digital Database for Screening Mammography (DDSM) medical image database and show that the results are very promising, which can contribute positively to the increased sensitivity of radiological examinations.

Keywords: *computer-aided diagnosis, breast segmentation, microcalcifications enhancement, detection of microcalcifications clusters.*

1. INTRODUÇÃO

O câncer de mama é o segundo tipo de câncer mais freqüente no mundo e o primeiro entre as mulheres. De acordo com o Instituto Nacional de Câncer - Brasil, o número estimado de casos em 2008 foi de aproximadamente 50.000 casos com um risco estimado de 51 casos a cada 100 mil mulheres (INCA - Instituto Nacional de Câncer, 2008). Apesar de ser considerado um câncer de bom prognóstico, quando diagnosticado e tratado precocemente, as taxas de mortalidade por câncer de mama continuam elevadas no Brasil, devido, provavelmente, ao diagnóstico tardio.

A mamografia é ainda o melhor exame para detecção precoce de sinais de câncer de mama e tem um papel importante nas decisões terapêuticas a serem adotadas. A mamografia pode revelar evidência de anormalidades como nódulos e calcificações, bem como sinais sutis como assimetria bilateral e distorção arquitetural. No entanto, a interpretação de mamografias não é uma tarefa fácil. Entre 10% e 30% das mulheres com presença da doença têm diagnóstico falso-negativo, ou seja, o radiologista interpretou o exame erroneamente como normal. Em aproximadamente 67% destes falsos-negativos o radiologista falhou na detecção

do câncer que era evidente em uma análise retrospectiva. Vários fatores podem contribuir para a falha no diagnóstico de câncer de mama em mamogramas: a baixa qualidade da imagem, o tamanho e a localização da lesão, a densidade do tecido mamário, a qualidade dos recursos técnicos e a habilidade de interpretação ou cansaço do radiologista.

Como alternativa para uma segunda “opinião” ou como uma estratégia de “pré-leitura”, Sistemas de Apoio ao Diagnóstico de Câncer de Mama – CAD têm sido propostos. A finalidade de um CAD é melhorar a eficácia do diagnóstico, assim como a interpretação da imagem radiológica, mediante o uso da resposta do computador como referência (Marques, 2001). Estudos recentes mostram que sistemas CAD podem melhorar a sensibilidade do radiologista sem um aumento significativo de taxa de *recall*.

Neste trabalho, é dado enfoque para a detecção de agrupamentos de microcalcificações, que são achados radiológicos em mamografias. Sua ocorrência é um sinal extremamente importante, principalmente porque podem constituir a única alteração radiológica visível na mamografia, além de poderem ser os sinais mais precoces de malignidade em exames mamográficos (Dengler, Behrens, Desaga, 2001).

No filme mamográfico, as regiões que contém microcalcificações costumam aparecer mais claras em relação aos demais tecidos da mama. Durante a digitalização do filme mamográfico, o brilho destas microcalcificações é perdido, as imagens contêm variação na distribuição de intensidade e incluem objetos que não fazem parte da estrutura da mama. Por isto, utilizar uma imagem sem correções de intensidade ou sem tratamentos adequados, pode trazer resultados insatisfatórios. Entre os ajustes a serem feitos na imagem, a retirada de regiões e objetos que interferem na detecção de microcalcificações e um realce na imagem mamográfica podem ser importantes para melhorar a eficácia do método.

Assim, este artigo tem como objetivo descrever um novo método para detectar agrupamentos de microcalcificações em imagens mamográficas digitais. Por se tratar de um problema complexo, decidiu-se dividir este método em três fases a fim de que as microcalcificações estejam suficientemente em condições de serem detectadas em imagens médicas com baixa qualidade. As três etapas são: *segmentação da mama*, *realce de microcalcificações* e *detecção de agrupamentos de microcalcificações*. A primeira etapa é importante para delimitar a área de interesse na imagem, enquanto que a segunda etapa melhora a nitidez e enfatiza detalhes na mamografia, e finalmente a terceira etapa consiste em se detectar os agrupamentos de microcalcificações. Espera-se que toda esta metodologia

contribua positivamente com o fornecimento de informações adicionais para que sensibilize o diagnóstico de sinais precoces de câncer de mama.

O artigo está organizado da seguinte maneira: a Seção 2 descreve os métodos utilizados na detecção de agrupamentos de microcalcificações, a Seção 3 analisa os resultados obtidos e finalmente a Seção 4 apresenta as considerações finais.

2. MATERIAL E MÉTODOS

Nesta Seção, são apresentadas as etapas necessárias desenvolvidos para a detecção de agrupamentos de microcalcificações. A Figura 1 esquematiza este método. A Seção 2.1, apresenta as etapas necessárias para a segmentação da mama: a detecção do contorno externo da mama e extração do músculo peitoral.

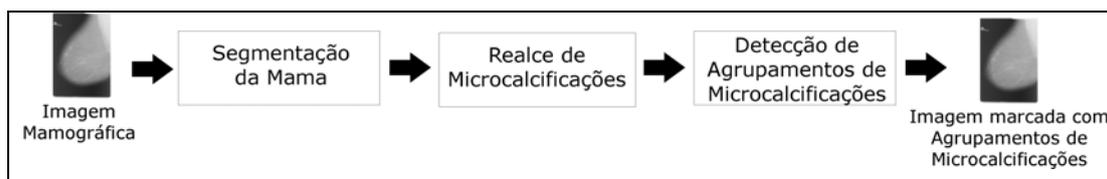


Figura 1: Fluxograma que ilustra as etapas do método de detecção de agrupamentos de microcalcificações.

A Seção 2.2 descreve um método de realce de microcalcificações. Finalmente, a Seção 2.3 descreve o processo de detecção de agrupamentos de microcalcificações.

2.1. Segmentação da Mama

A segmentação da mama é uma tarefa composta pela detecção do contorno externo da mama e pela extração do músculo peitoral. A Figura 2 ilustra como isto será feito.



Figura 2: Fluxograma que ilustra o funcionamento do método de segmentação da mama.

A detecção do contorno externo da mama tem como objetivo delimitar a região pertencente a mama na imagem. Além disso, também é possível remover qualquer artefato ou ruído presente na parte externa da mama, como a “identificação do paciente ou do filme” (região altamente densa) e demais estruturas, muito comuns nesta região, que podem interferir negativamente nos resultados de processamentos posteriores.

O músculo peitoral é uma região densa que está presente nas vistas médio laterais oblíquas (MLO) dos mamogramas e pode influenciar nos resultados de métodos que necessitem de informações como intensidade em níveis de cinza de uma imagem. Por exemplo, muitas das vezes, o músculo peitoral aparece com a mesma intensidade em níveis de cinza de lesões e calcificações presentes no tecido fibroglandular. Desta maneira, a detecção destes sinais na mama pode ser prejudicada pela presença do músculo peitoral na imagem mamográfica.

2.1.1. Detecção do Contorno Externo da Mama

Esta etapa que visa delimitar a região do tecido mamário, a imagem mamográfica digital passa por um processamento composto pelas seguintes sub-etapas: 1) Pré-processamento; 2) Limiarização global; 3) Filtragem da imagem resultante por um filtro morfológico abertura; 4) Detecção do contorno da mama e restauração do contorno para as dimensões originais. A Figura 3 ilustra estas etapas.

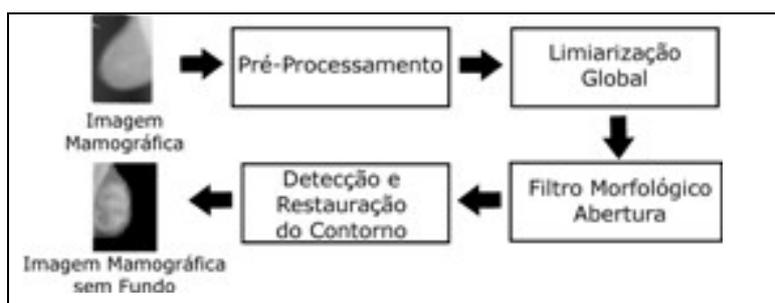


Figura 3: Fluxograma do Método de Detecção do Contorno Externo da Mama.

A idéia por trás deste método consiste em inicialmente remover e uniformizar as bordas da imagem mamográfica, eliminando ruídos e artefatos presentes nestas regiões, para depois separar o fundo do mamograma e o tecido mamário via limiarização global, utilizando-se um limiar baseado nas informações de intensidade dos *pixels* da imagem mamográfica.

2.1.1.1. Pré-processamento

Nesta sub-etapa, a imagem original (Figura 4(a)) é subamostrada em oito vezes para melhorar a performance do algoritmo. Note que as vistas da mama direita serão espelhadas com relação ao eixo vertical da imagem com o objetivo de padronizar o procedimento de análise e em virtude do posicionamento do músculo peitoral em vistas MLO.

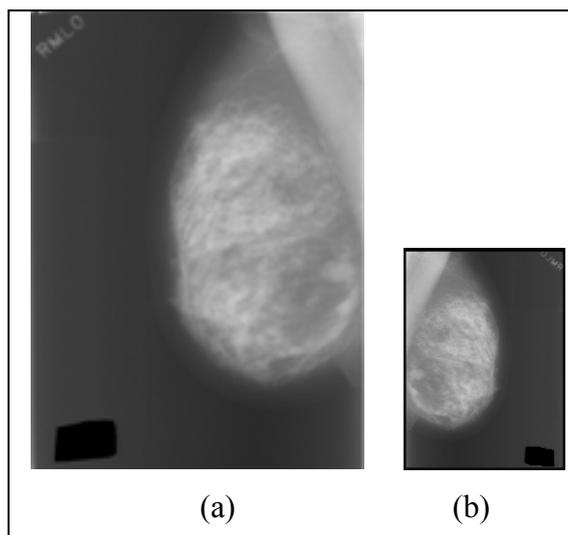


Figura 4: (a) Imagem mamográfica original; (b) Imagem subamostrada e espelhada da Figura (a).

Após a subamostragem e o espelhamento para mamas direita, uma máscara simétrica de dimensão $d \times d$ preenchida com valores iguais a zero é passada nas bordas da imagem de tal forma a eliminar artefatos e ruídos. O valor d é definido pelo usuário, exceto quando a máscara for utilizada na fronteira vertical esquerda, possuindo assim valor fixo igual a 5 ($d=5$). A coordenada inicial (x_0, y_0) da imagem tem seus valores reajustados, onde x_0 é igual a 6 e $y_0 = d+1$. A Figura 4(b) mostra a imagem resultante deste processamento.

2.1.1.2. Limiarização global

Uma limiarização global é uma maneira de dividir uma imagem em suas regiões constituintes particionando o seu histograma utilizando um único limiar T (Gonzalez, Wood, 2002).

O sucesso da limiarização depende do valor T , que é calculado baseado nos valores de *pixels* pertencentes ao fundo do mamograma. Uma máscara 15×15 , indicada por M_1 na

Figura 5(a), é sobreposta no canto direito, a 70% da altura da imagem sem considerar a região da borda que teve seus *pixels* alterados para zero na etapa anterior. Outra máscara M_2 é sobreposta no canto esquerdo da imagem a 40% da altura da imagem. O limiar T é calculado de acordo com

$$T = \begin{cases} \mu_1; & \text{se } \mu_1 \in [15,60] \\ \mu_2; & \text{caso contrário} \end{cases}$$

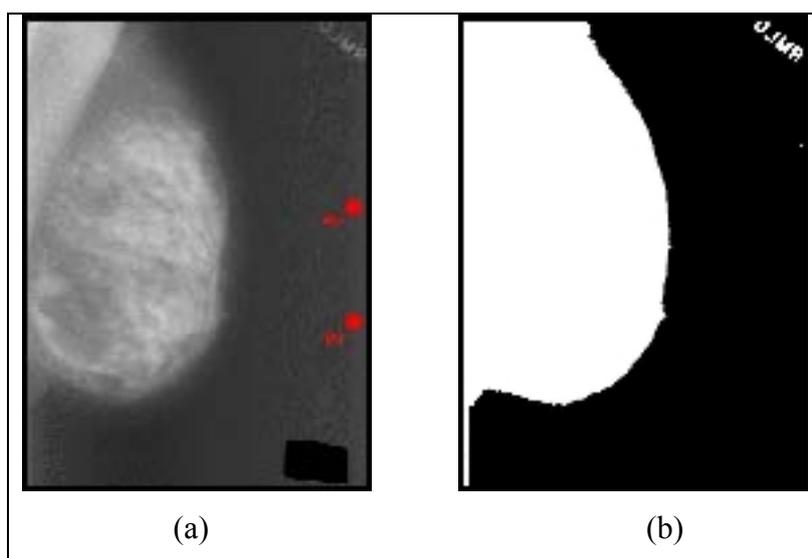


Figura 5: (a) Máscaras sobrepostas na imagem, destacadas em vermelho, para cálculo do limiar T ; (b) Imagem resultante da limiarização global com limiar $T = 76$ sobre Figura 4(b).

onde μ_1 é a média da máscara M_1 e μ_2 é a média da máscara M_2 . Estas condições evitam que se obtenham informações inconsistentes do fundo da imagem, já que podem existir ruídos nestas regiões. O resultado desta limiarização global é ilustrado na Figura 5(b).

2.1.3. Aplicação do filtro morfológico abertura

Nesta sub-etapa, um filtro morfológico de abertura para imagens binárias com elemento estruturante circular de raio igual a 25 *pixels*, é utilizado para eliminar ruídos existentes na imagem binária (Gonzalez, Woods, 2002). O raio deste elemento estruturante foi calculado com base na observação visual da remoção de artefatos e ruídos presentes no fundo de várias imagens binárias e também o quanto a borda da mama foi suavizada sem perder informações relevantes. O resultado é mostrado na Figura 6.

2.1.1.4. Remoção do Fundo do Mamograma

Nesta sub-etapa, o contorno da mama é extraído utilizando uma metodologia em que todos os *pixels* da imagem são analisados. Para cada *pixel* p da imagem binária, são analisados os seus oito vizinhos e uma nova imagem contendo apenas o contorno é gerada. O procedimento segue abaixo:

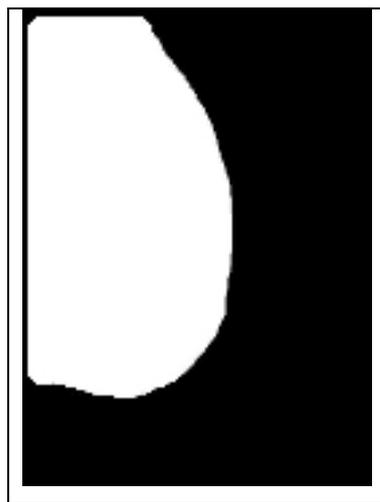


Figura 6: Filtro morfológico aplicado à Figura 5(b).

$$p' = \begin{cases} 0; & \text{se } p = 0 \text{ ou } p \text{ não possui vizinho } v \text{ onde } v = 0; \\ 1; & \text{se } p = 1 \text{ e } p \text{ tem algum vizinho } v \text{ onde } v = 0; \end{cases}$$

onde p' é um *pixel* da imagem contendo somente o contorno. Existindo apenas o contorno na imagem binária, utiliza-se a técnica *Chain-Code* (Gonzalez, Wood, 2002) para aloca-lo em uma estrutura adequada.

Caso ainda existam artefatos na imagem binária suavizada, o método pode retornar mais que um contorno. O contorno de interesse é escolhido seguindo o critério de que o contorno da mama será o maior, já que ocupa uma grande parte do mamograma. A Figura 6 ilustra esse procedimento e a Figura 7(a) mostra o contorno resultante sobreposto ao mamograma da Figura 3(b).

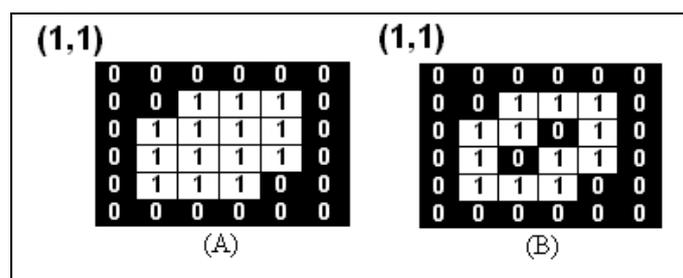


Figura 7: (a) Representação dos valores de níveis de cinza de uma imagem binária. (b) Contorno extraído da imagem binária: observa-se que os *pixels* (3,4) e (4,3) não pertencem ao contorno.

Finalmente, o contorno é restaurado para as condições e dimensões originais da imagem. Para se obter somente a mama sem o fundo da imagem mamográfica, faz-se uma operação AND entre a imagem original (Figura 3(a)) e a imagem binária contendo o contorno final restaurado e preenchido (Figura 8(b)), resultando na Figura 8(c).

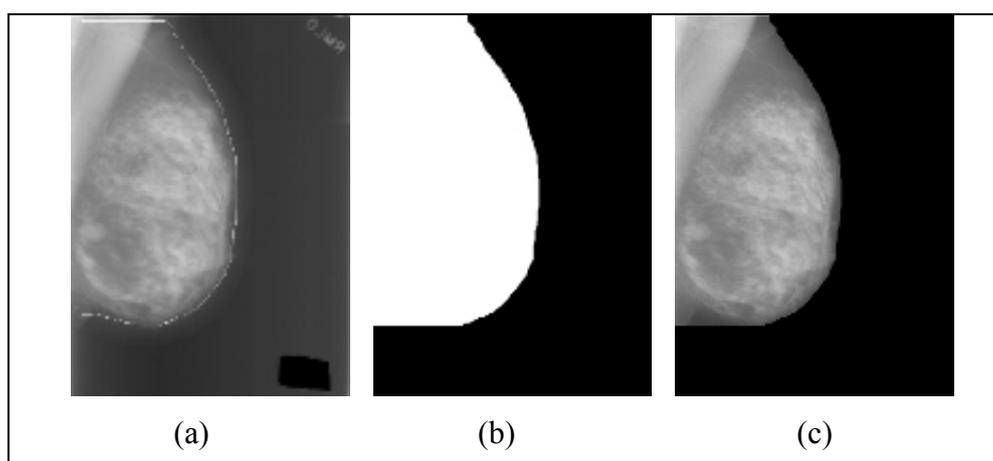


Figura 8. (a) Extração do contorno da Figura 6 sobreposto a imagem da Figura 4(a) espelhada; (b) Imagem binária com o contorno final restaurado preenchido internamente na posição original; (c) Operação AND entre a Figura 4(a) e a imagem do item (b), removendo o fundo do mamograma.

2.1.2. Extração do Músculo Peitoral

A extração do músculo peitoral recebe como entrada a imagem com o fundo da imagem mamográfica extraído e envolve as seguintes etapas: 1) delimitação da região de interesse (ROI); 2) cálculo de um limiar, que é utilizado na limiarização global na próxima

etapa; 3) processo iterativo, a fim de se obter a melhor segmentação do músculo peitoral; e 4) restauração do contorno. A Figura 9 ilustra estes passos.



Figura 9: Fluxograma que ilustra o processo de extração do músculo peitoral.

A idéia da extração do músculo peitoral em uma imagem mamográfica depende de um limiar (L) que leva em consideração a qualidade e a nitidez em termos de intensidade de nível de cinza do músculo peitoral. Em um processo iterativo, este limiar é alterado a cada iteração, até que se obtenha um valor ideal para que o músculo peitoral seja removido do restante da imagem.

A extração do músculo peitoral é detalhada nos tópicos seguintes.

2.1.2.1. Delimitação da Região de Interesse (ROI)

Esta sub-etapa consiste em delimitar a área contendo o músculo peitoral, que é a região de interesse (ROI) e se baseia na proposta de (Kwok, Chandrasekhar, Attikiouzel, 2001). Nas projeções MLO da mama esquerda (vale lembrar que as mamas direita foram espelhadas), a região do músculo peitoral está presente na parte superior esquerda da imagem mamográfica. Utilizam-se quatro pontos, obtidos automaticamente, para delimitar a região do músculo peitoral. Tais pontos são determinados como: N_1 - é o ponto superior esquerdo, ou, a coordenada inicial a partir da borda; N_2 - é o ponto superior que delimita a borda da mama ao lado direito; N_3 - é o ponto que está a 70% da altura da mama na ordenada X e na mesma ordenada Y do ponto N_1 ; N_4 - é o ponto que está na mesma ordenada X do ponto N_1 e a mesma ordenada Y do ponto N_2 . O procedimento abaixo ilustra como obter o ponto N_2 , onde *imagem* representa a imagem mamográfica sem o fundo, obtida da fase anterior.

```
Algoritmo determinaN2(**imagem)
```

1. $y_2 \leftarrow 1$;
2. largura \leftarrow obtém largura da imagem;
3. ENQUANTO imagem[6][y_2] \neq 0 E $y_2 \leq$ largura FAÇA
4. $y_2 \leftarrow y_2 + 1$;
5. $N_2 \leftarrow (6, y_2)$;

A ROI delimitada é apresentada conforme a Figura 10.

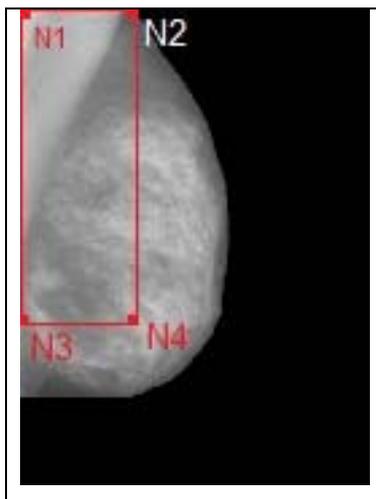


Figura 10: Retângulo formando a Região de Interesse (ROI).

2.1.2.3. Cálculo do Limiar (L)

A idéia desta etapa é calcular um limiar L , que estará diretamente envolvido com a etapa de limiarização global do processo iterativo (descrito na seção seguinte). O valor do limiar L será fixado dependendo da qualidade do músculo peitoral da ROI.



Figura 11: ROI com máscaras sobrepostas em cada região destacadas em vermelho.

Dois máscaras de dimensões 15×15 são utilizadas para analisarem *pixels* de duas regiões distintas. A primeira, indicada por M_1 na Figura 11, é colocada automaticamente na

parte superior esquerda da ROI, onde contém *pixels* referentes ao músculo peitoral, e a segunda, indicada por M_2 , é colocada automaticamente na parte inferior direita da ROI, que contém *pixels* referentes ao fundo, que é o tecido fibroglandular. Ambas as máscaras são preenchidas com os valores dos *pixels* das áreas sobrepostas. A obtenção do limiar é feita de acordo com o procedimento `encontraLimiar` a seguir:

```

Algoritmo encontraLimiar(**regiaoInteresse, altura, largura, *L);
1. Calcular  $\mu_p$  e  $\mu_m$  (médias da máscara  $M_1$  e  $M_2$ , respectivamente);
2. SE  $\mu_p - \mu_m > 10$ 
3.   ENTÃO  $i \leftarrow 1$ ;
4.   ENQUANTO regiaoInteresse[20][i] >  $\mu_m$  FAÇA
5.      $i \leftarrow i + 1$ ;
6.   RETORNA *T = regiaoInteresse[20][i-2];
7. SENÃO
8.   PARA  $i \leftarrow \lfloor \text{altura}/2 \rfloor$  ATÉ altura
9.     PARA  $j \leftarrow 1$  ATÉ largura
10.      regiaoInteresse[i][j]  $\leftarrow$  regiaoInteresse[i][j] - 60;
11.   RETORNA *L =  $\mu_p$ ;

```

Resumidamente, se a diferença entre μ_p e μ_m (linha 2) seja maior do que 10, será calculada um valor aproximado em nível de cinza da borda do músculo peitoral, pois há indícios que o músculo peitoral esteja bem definido. Desta maneira, utilizar este valor para o limiar L aumenta a possibilidade de se obter uma segmentação ideal do músculo peitoral. Na linha 4, foi escolhida aleatoriamente a linha 20 da ROI para encontrar o valor aproximado da borda do músculo peitoral.

Se a diferença entre μ_p e μ_m for menor do que 10, pode-se dizer que a imagem não está com boa qualidade ou que o músculo peitoral esteja com densidade menor que o tecido fibroglandular. Isto leva a realização de um procedimento para escurecer os *pixels* da metade inferior da ROI. Desta maneira, altera-se a intensidade da ROI e μ_p é o limiar L , o que aumenta as chances de segmentação do músculo peitoral nestas condições.

2.1.2.4. Processo iterativo para refinar o músculo peitoral

Neste próximo passo, será executado um processo iterativo sobre a ROI com o propósito de se obter o músculo peitoral ideal, como mostrado no fluxograma da Figura 12. A Figura 13 ilustra duas iterações que o processo executou para chegar a um contorno satisfatório.

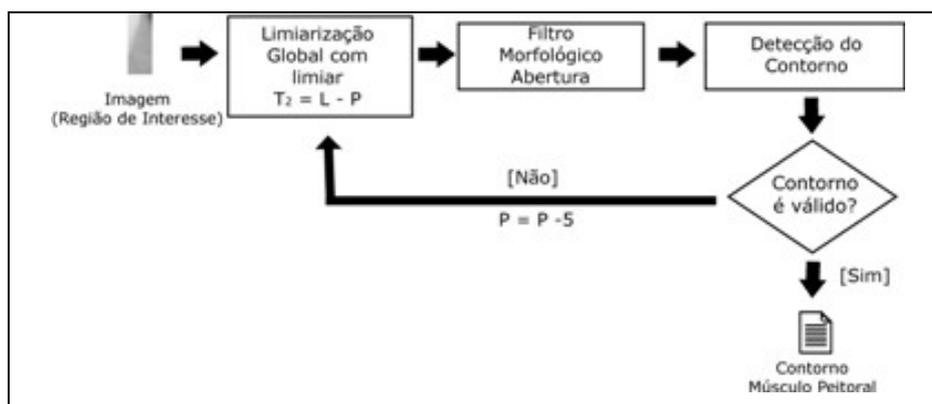


Figura 12: Diagrama que descreve o processo iterativo para refinar o músculo peitoral.

O objetivo deste processo iterativo é obter automaticamente um parâmetro P ideal e diferente para cada iteração executada, sendo P responsável por configurar o valor do limiar T_2 em uma dada iteração. Cada iteração executará sobre a ROI original as seguintes etapas: uma operação de limiarização global (Figuras 13(a) e 13(d)), filtragem morfológica abertura com elemento estruturante circular de raio 20 (Figuras 13(b) e 13(e)), extração de contorno (Figuras 13(c) e 13(f)) e verificação da validade do contorno (as condições para validade são discutidas no parágrafo a seguir). Caso o contorno seja inválido, o parâmetro P tem seu valor decrementado em cinco unidades e uma nova iteração é iniciada. Neste trabalho, o parâmetro P teve sua contagem iniciada em 25 unidades ($P = 25$) para casos onde o músculo peitoral está bem definido (isto é, se a diferença entre μ_p e μ_m for menor do que 10 unidades) e em 60 unidades ($P = 60$), caso contrário. Tais valores foram fixados ao parâmetro P observando-se os resultados.

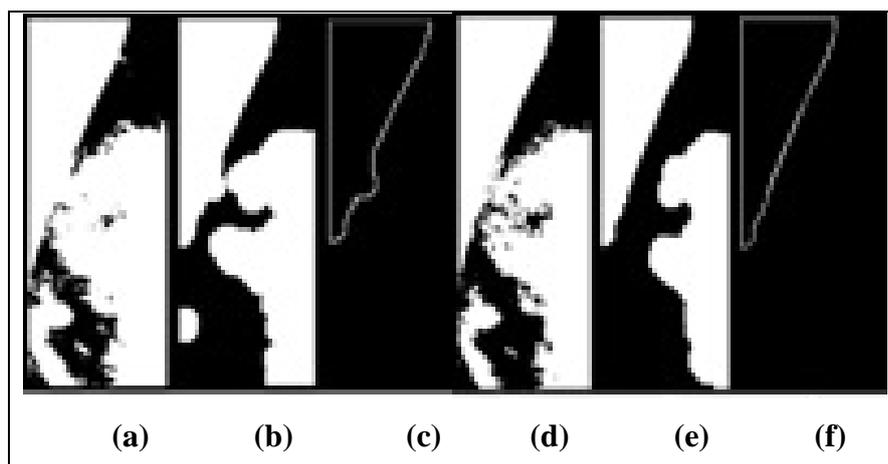


Figura 13: Etapas da execução da primeira iteração de (a) a (c) onde o contorno obtido não foi validado e da segunda iteração de (d) a (f) onde o contorno foi validado: (a) Limiarização global aplicada à Figura 12(a) limiar $T_2 = 136$; (b) filtro morfológico abertura; (c) extração do contorno; (d) Limiarização global aplicada à Figura 12(a) limiar $T_2 = 141$; (e) filtro morfológico abertura; (f) extração do contorno.

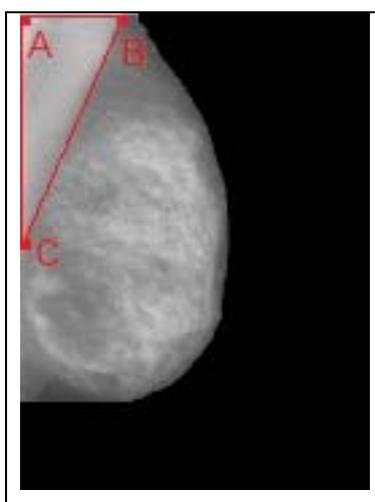


Figura 14: Músculo peitoral como se fosse um triângulo retângulo.

O músculo peitoral é considerado como se fosse um triângulo retângulo, representado na Figura 14. Essa consideração será importante para o estabelecimento de critérios para validação do contorno do músculo peitoral. Percorrendo-se o trecho do contorno obtido partindo do ponto B e chegando ao ponto C , caminha-se pelo contorno, *pixel a pixel*. O *pixel* a ser analisado deve aumentar ou manter o valor da ordenada X e reduzir ou manter o valor da ordenada Y em relação ao *pixel* analisado anteriormente. Todavia, algumas imperfeições são toleradas ao longo do caminhamento. Para imagens mamográficas em que o músculo peitoral está bem definido, tais restrições são impostas:

1. Aumento da ordenada Y mais do que quatro *pixels* seguidos (Figura 15(a));
2. Aumento da ordenada Y mais do que seis *pixels* não seguidos (Figura 15(b));
3. Devido à existência de falhas nas coordenadas mais próximas das linhas iniciais, a ordenada Y deve decrescer, no mínimo dez vezes até a metade da altura da ROI;

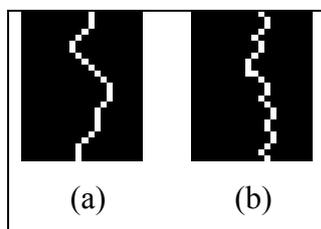


Figura 15: Exemplo dos critérios estabelecidos para validação do contorno do músculo peitoral: (a) trecho de contorno inválido pela condição 1; (b) trecho de contorno inválido pela condição 2.

Para casos onde o músculo peitoral não está bem definido, exemplificados nas Figuras 16(a) e 16(b), a ordenada Y não deve aumentar ao caminhar por cinco *pixels* descontínuos. Se algum dos critérios acima for violado, o contorno é considerado inválido. Caso contrário, se o contorno for aceito, o contorno será restaurado, descrito na Seção seguinte.

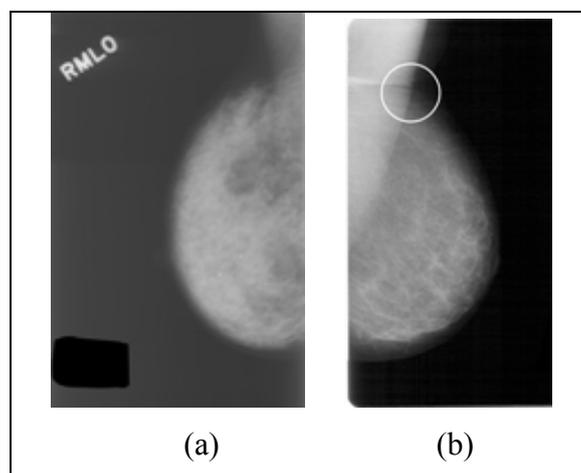


Figura 16: Exemplo de dois mamogramas onde o músculo peitoral está mal definido: (a) caso onde há baixa nitidez; (b) caso onde há falha (circulado em branco) que compromete a fase de detecção de contorno.

2.1.2.5. Restauração do Contorno

Nesta sub-etapa, o músculo peitoral é extraído fazendo-se uma operação lógica AND entre a imagem sem o fundo do mamograma (Figura 8(c)) e a imagem binária contendo o contorno final preenchido (Figura 8(b)), com os valores dos *pixels* internos ao mesmo iguais a 1, resultando na Figura 20(c), que foi espelhada.

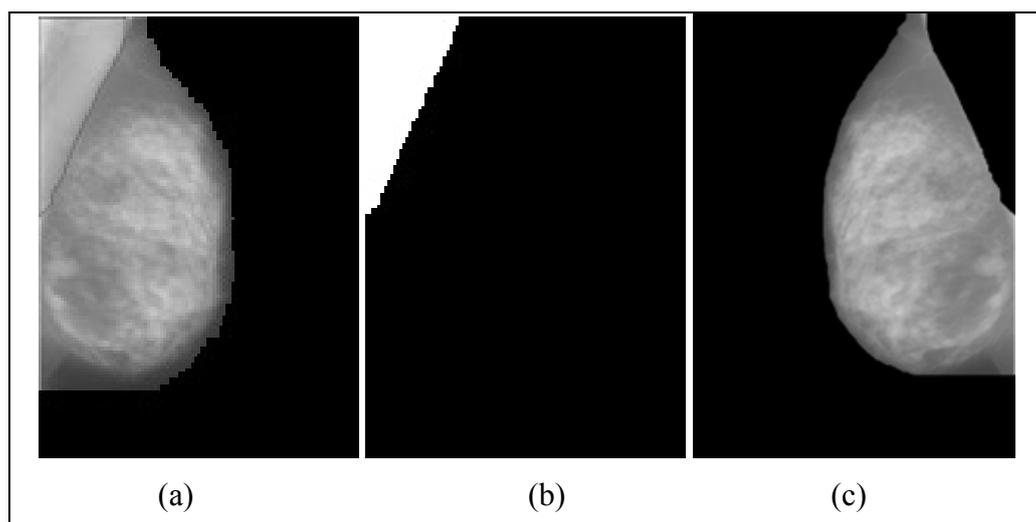


Figura 20: (a) Contorno final do músculo peitoral; (b) Imagem binária do contorno internamente preenchido; (c) Imagem com o fundo e o músculo peitoral removidos e espelhada.

2.2. Realce de Microcalcificações

O objetivo desta etapa do processo de detecção de microcalcificações é aumentar o contraste da imagem, reduzindo a densidade do tecido fibroglandular e realçando as microcalcificações, de forma a torná-las mais nítidas. Desta maneira, há uma redução na detecção de sinais *falsos-positivos* (*FP*) e maior eficácia em se detectar *verdadeiros-positivos* (*VP*).

O método de realce proposto recebe como parâmetro a imagem mamográfica segmentada (somente com o tecido fibroglandular). A Figura 21 ilustra as etapas para realçar as microcalcificações.

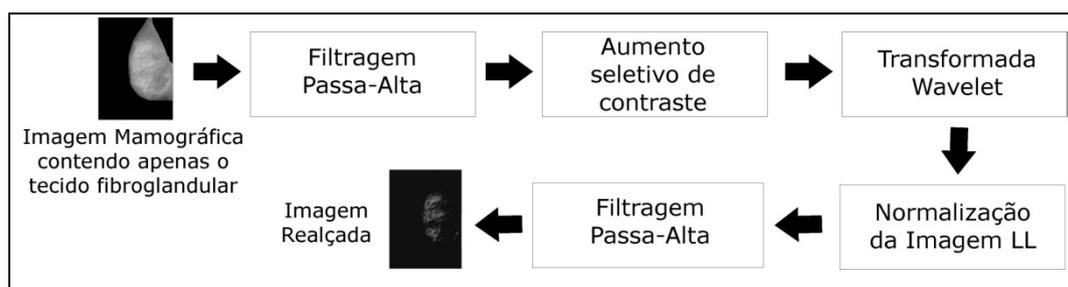


Figura 21: Diagrama que ilustra as etapas do método de Realce de Microcalcificações.

A seguir, cada uma das etapas será descrita em detalhes. Por conveniência, serão apresentados nos resultados de cada fase apenas as Regiões de Interesse (ROI's), ou seja, a área da imagem mamográfica que contém o agrupamento de microcalcificações.

2.2.1. Filtragem Passa-Alta

Nesta etapa, a imagem contendo apenas o tecido fibroglandular, é filtrada com o propósito de enfatizar detalhes finos, como as microcalcificações, pois podem estar com densidades menores que o tecido fibroglandular.

Este problema pode ser solucionado utilizando-se um filtro passa-alta (Gonzalez, Woods, 2002), que realça as transições entre regiões diferentes da mama e mantém informações do tecido fibroglandular. A soma algébrica dos coeficientes deste filtro é diferente de zero, o que permite manter informações sobre a imagem original.

Os coeficientes deste filtro foram obtidos observando-se o quanto as transições entre as bordas das microcalcificações e o tecido fibroglandular eram realçadas. A Equação (1) mostra os coeficientes deste filtro.

$$\text{Filtro Linear Passa-Alta} = [-0.6 \ -0.4 \ 3 \ -1] \quad (1)$$

O resultado da aplicação deste filtro sobre a imagem da Figura 20(b), representada pela ROI na Figura 22(a) é apresentado na Figura 22(b).

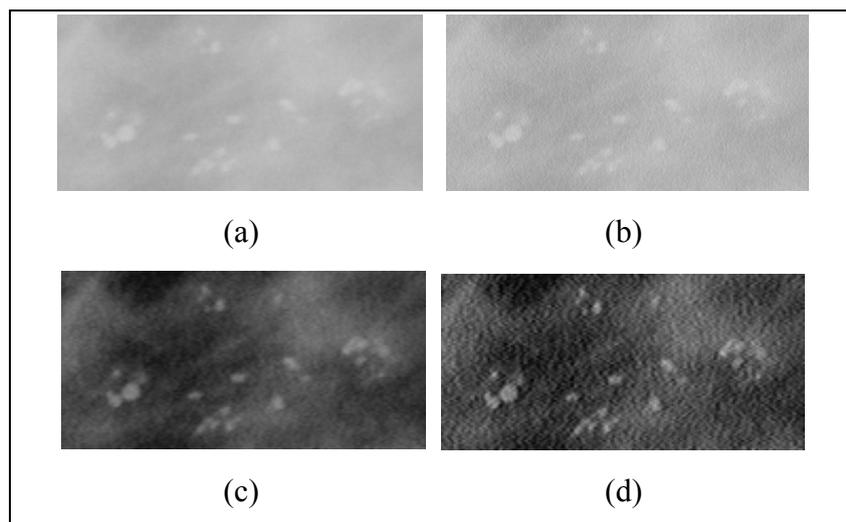


Figura 22: (a) ROI contendo microcalcificações da Figura 20(b); (b) ROI contendo microcalcificações da imagem filtrada com filtro passa-alta; (c) ROI contendo microcalcificações da imagem resultante do aumento seletivo de contraste sobre (b); (d) ROI contendo microcalcificações da imagem resultante da segunda filtragem passa-alta sobre a componente *LL*

2.2.2. Aumento seletivo de contraste

Com o objetivo de manter apenas as regiões mais claras, foi aplicada à imagem representada pela ROI da Figura 22(b) a transformação mostrada na Figura 23.

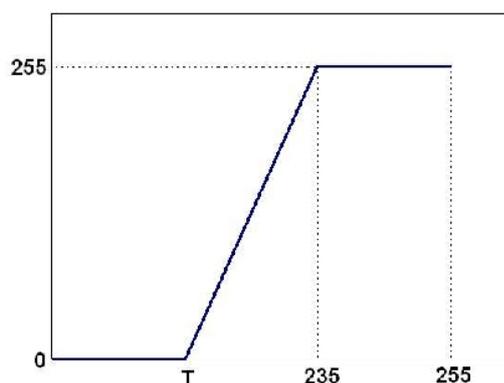


Figura 23: Transformação que amplia o intervalo dinâmico da imagem original no intervalo $[T, 235]$ para $[0, 255]$.

O resultado da aplicação desta função de transformação é ilustrado na Figura 22(c). O valor de T é fornecido pelo usuário e o valor 235 foi fixado observando-se os valores de

níveis de cinza das microcalcificações em diversos mamogramas, pois os mesmos estão no intervalo $[T,235]$.

2.2.3. Transformada Wavelet

A Transformada *Wavelet* permite descrever a imagem em termos de suas frequências e de localização destas frequências (Pedrini, Schwartz, 2007). Neste trabalho, a Transformada *Wavelet* foi decomposta em apenas um nível. A Figura 24 ilustra este esquema de decomposição. Aqui, as variáveis M e N representam a largura e a altura da imagem mamográfica, respectivamente. Os filtros LF_x e HF_x são descritos como filtro passa-baixa e filtro passa-alta, respectivamente aplicados sobre as linhas da imagem, enquanto que os filtros LF_y e HF_y se referem ao filtro passa-baixa e passa-alta aplicado sobre as colunas da imagem. O símbolo $\downarrow 2$ se refere à subamostragem, indicando que a cada duas linhas da imagem, uma deve ser removida.

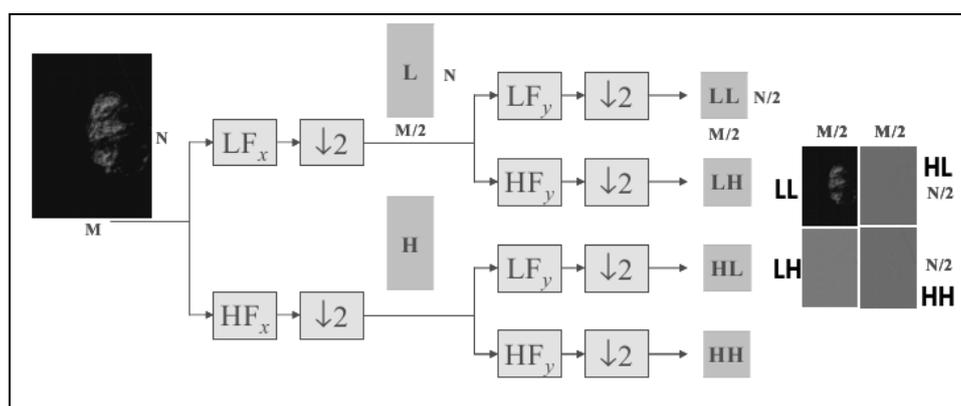


Figura 24: Esquema Geral da Transformada *Wavelet*.

A componente LL possui informações de baixa frequências nas orientações vertical e horizontal, enquanto que a componente HH possui informações de alta frequências nas orientações vertical e horizontal. A componente LH possui informações de baixa frequências na orientação horizontal e alta frequência na orientação vertical, enquanto que a componente HL possui informações de baixa frequências na orientação vertical e alta frequência na orientação horizontal.

Foi utilizada a família de *Wavelets* de *Daubechies* de ordem 4 ($db4$), o que significa que, os filtros desta transformada possuem ordem 4, ou seja, quatro coeficientes. Esta família define os coeficientes dos filtros passa-baixa e passa-alta como mostra as Equações (2) e (3).

$$\text{Filtro de Daubechies Passa-Alta} = \left[\frac{1-\sqrt{3}}{4\sqrt{2}} - \frac{3-\sqrt{3}}{4\sqrt{2}} \frac{3+\sqrt{3}}{4\sqrt{2}} - \frac{1+\sqrt{3}}{4\sqrt{2}} \right] \quad (2)$$

$$\text{Filtro de Daubechies Passa-Baixa} = \left[\frac{1+\sqrt{3}}{4\sqrt{2}} \frac{3+\sqrt{3}}{4\sqrt{2}} \frac{3-\sqrt{3}}{4\sqrt{2}} \frac{1-\sqrt{3}}{4\sqrt{2}} \right] \quad (3)$$

2.2.4. Normalização da componente *LL*

A componente resultante da Transformada *Wavelet* a ser utilizada para esta etapa é *LL*, pois a intensidade de níveis de cinza do tecido fibroglandular está mais suprimida e menos suscetível a detecção de ruídos. Porém, observa-se que valores de níveis de cinza dos *pixels* desta imagem estão fora do intervalo realizável [0,255]. A operação que re-escalona os *pixels* da imagem no intervalo [0,255] é definida pela Equação (4):

$$\text{out}(i, j) = \left(\frac{\text{img}(i, j) - \min(\text{img})}{\max(\text{img}) - \min(\text{img})} \right) \quad (4)$$

onde *img* é a imagem a ser normalizada, *min* e *max* são os menor e o maior valor de *img*, respectivamente.

Mais detalhes sobre a operação de normalização podem ser encontrados em (Gonzalez, R. C. and Woods, R. E., 2000).

2.2.5. Segunda Filtragem Passa-Alta

O objetivo principal desta segunda passagem do filtro na imagem é realçar informações perdidas na componente *LL* resultante da Transformada *Wavelet*. A imagem originada desta última etapa conterá as microcalcificações de forma mais atenuadas.

O filtro utilizado é o mesmo usado na Seção 2.2.1. Nesta segunda filtragem, observa-se que as microcalcificações estão mais visíveis do que na imagem resultante da primeira filtragem. A ROI resultante é ilustrada na Figura 22(d).

Finalmente, obtém-se a imagem mamográfica realçada. Repare que, comparando-se esta imagem com a imagem original, as microcalcificações da imagem realçada estão mais nítidas.

2.3. Detecção de Agrupamento de Microcalcificações

A detecção de agrupamentos de microcalcificações é uma etapa importante para o diagnóstico de câncer de mama, já que, para fins de um diagnóstico correto e seguro, é mais interessante encontrar agrupamentos de microcalcificações do que microcalcificações isoladas, pois as primeiras estão relacionadas à casos de malignidade.

Inicialmente, a imagem mamográfica obtida no processo anterior é segmentada via limiarização global com a finalidade de isolar as microcalcificações do fundo do tecido fibroglandular. Cada sinal resultante da segmentação na imagem é transformado em um único ponto com o objetivo de facilitar a contagem de sinais na fase de agrupamento. Considera-se um agrupamento toda ocorrência de três ou mais sinais em uma área de $3,2 \text{ mm}^2$ na imagem (Nishikawa et. al., 1993). Caso esta condição seja satisfeita, os sinais pertencentes a esta área permanecerão na imagem final.

A Figura 25 ilustra os passos do método de detecção de agrupamentos de microcalcificações utilizados neste trabalho.

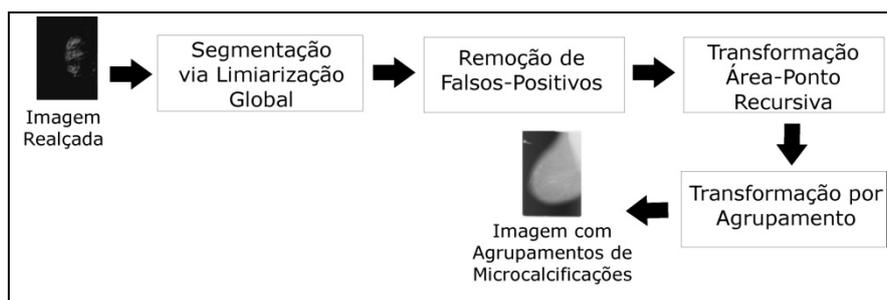


Figura 25: Fluxograma completo que ilustra o método de Detecção de Agrupamentos de Microcalcificações.

2.3.1. Segmentação via Limiarização Global

Esta etapa recebe a imagem contendo apenas o tecido fibroglandular com as microcalcificações realçadas e usando o limiar T_A , definido na Equação (6), a imagem é

transformada em uma imagem binária, restando apenas os *pixels* com valores de níveis de cinza maiores que T_A .

$$T_A = T - 10 \quad (6)$$

Na Equação (6), T é o limiar fornecido pelo usuário na Seção 2.2.2. A segmentação utilizando o limiar T_A gera uma imagem binária mantendo a maioria dos *pixels* correspondentes às microcalcificações, como também sinais mais claros da imagem, onde parte deles serão eliminados na fase seguinte. A Figura 26(b) ilustra esta operação sobre a figura representada pela ROI da Figura 26(a).

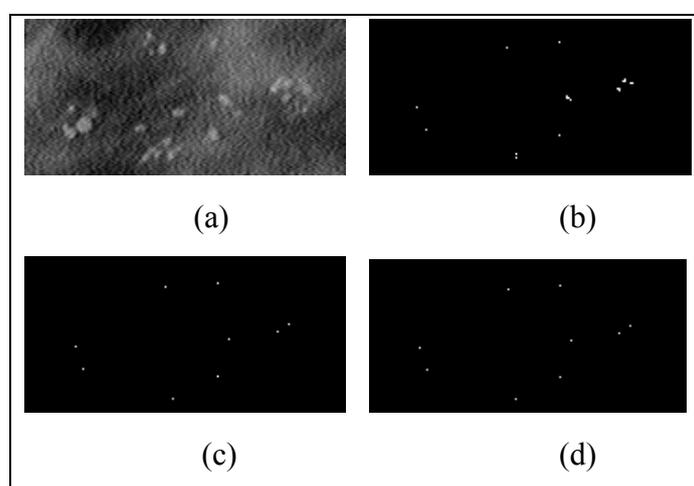


Figura 26: Visualização dos resultados na ROI da execução de cada etapa do método de Detecção de Agrupamentos de Microcalcificações: (a) ROI realçada; (b) ROI segmentada; (c) ROI após a aplicação da Transformação Área-Ponto Recursiva em (b); (d) ROI após a aplicação da Transformação por Agrupamento em (c).

2.3.2. Remoção de Falsos-Positivos

A imagem binária resultante da segmentação pode conter sinais que não apresentam características de microcalcificações. Sabe-se também que objetos como macrocalcificações e nódulos podem permanecer na imagem, pois possuem altos valores de níveis de cinza. Esta etapa tem como objetivo eliminar os sinais falsos-positivos.

As microcalcificações são sinais que possuem um diâmetro menor que $0,5 \text{ mm}$, o que equivale a uma imagem da base de dados *DDSM*, a 10 pixels . Desta maneira, sinais que

possuírem diâmetro maior que 10 *pixels* são eliminados da imagem binária. Para se obter o tamanho do sinal em número de *pixels*, calcula-se a diferença entre dois pontos extremos que o compõe.

2.3.3. Transformação Área-Ponto Recursiva

Com base na idéia da transformação área-ponto, apresentada por *Nishikawa et al.* (Nishikawa et. al., 1993) foi desenvolvido um procedimento com o objetivo de transformar cada sinal presente na imagem segmentada em um único *pixel*. A transformação é realizada da seguinte maneira: uma máscara de 3×3 *pixels* é passada sobre a imagem binária começando do canto superior esquerdo. Para cada posição, é contabilizado o número de *pixels* com níveis de cinza superior ao do fundo da imagem. Havendo dois ou mais *pixels* com essa característica, todos os *pixels* sob a máscara são ajustados para um valor igual ao do fundo. Neste caso o *pixel* do canto inferior direito do núcleo, tem seu valor fixado para 255. O núcleo da máscara é movido 2 *pixels* para a direita e o processo é repetido até processar todas as colunas da imagem. O processo salta de 2 em 2 linhas até atingir o fim da imagem. A finalidade desse processo é facilitar a contagem das estruturas para a identificação dos possíveis agrupamentos da imagem. A Figura 27 ilustra esta tarefa.

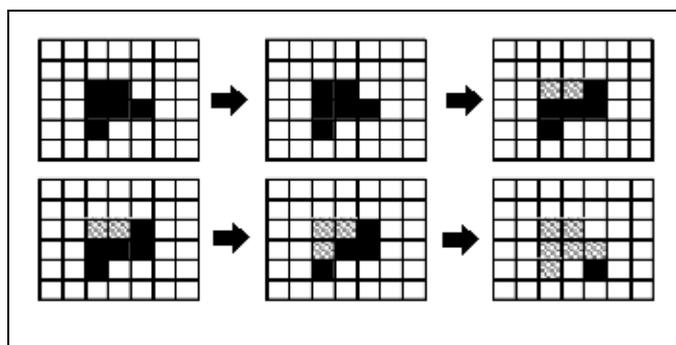


Figura 27: Exemplo de utilização da Transformação Área-Ponto Recursiva proposto por *Nishikawa et. al.*

A Figura 26(c) ilustra esta operação sobre a Figura 26(b), que representa uma imagem com as microcalcificações segmentadas.

2.3.4. Transformação por Agrupamento

A detecção de microcalcificações agrupadas é o processo que efetivamente identifica os agrupamentos existentes na imagem, indicando também a sua localização e a quantidade de microcalcificações neles contidas.

Este método consiste em posicionar máscaras seqüencialmente ao longo da imagem binária resultante da *Transformação Área-Ponto* e verificar em cada posicionamento a quantidade de sinais existentes. Se forem contados mais do que três sinais, um agrupamento é identificado. Quanto maior a restrição sobre o número de sinais na máscara, menor é a detecção de falsos-positivos, no entanto diminui a sensibilidade do método. Inicialmente a máscara é posicionada na imagem na coordenada inicial da imagem e, para cada posicionamento, verifica se há uma quantidade igual ou maior do que três sinais. Caso esta condição seja satisfeita, o contador de agrupamento é incrementado, a posição de cada sinal é armazenada e os sinais localizados sob a máscara são removidos para que não sejam contados novamente. Ocorre, então, o deslocamento da máscara e o processo se repete até que toda a imagem tenha sido percorrida.

O tamanho da máscara deve estar de acordo com o tamanho do *pixel*. Neste método, foi utilizada uma máscara de 66 *pixels*, pois um *pixel* da base de dados *DDSM* possui tamanho $0,05 \times 0,05 \text{ mm}$. Nishikawa et. al trabalhou com mamografias onde cada *pixel* tinha $0,1 \times 0,1 \text{ mm}$. Logo, a máscara utilizada em seu método possuía dimensões $32 \times 32 \text{ pixels}$.

A Figura 26(d) ilustra a utilização deste método na Figura 26(c), que representa uma imagem resultante da *Transformação Área-Ponto Recursiva*. A Figura 28 ilustra o funcionamento da operação de *Transformação por Agrupamento*.

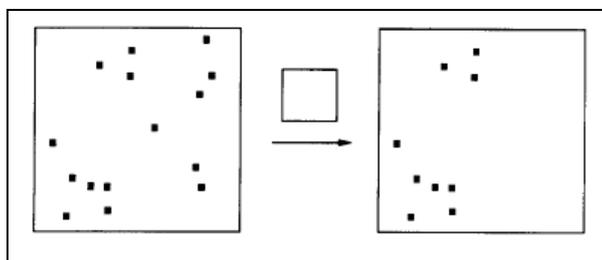


Figura 28: Exemplo de utilização da *Transformação por Agrupamento* por *Nishikawa et. al*: Imagem de entrada à esquerda e imagem de saída à direita.

Finalmente, a imagem com os sinais que identificam as microcalcificações é reamostrada para seu tamanho original e esses sinais são sobrepostos à imagem original. Desta maneira, obtém-se uma imagem mamográfica com os agrupamentos de microcalcificações detectados, que podem indicar regiões suspeitas. A Figura 29 apresenta o resultado final.

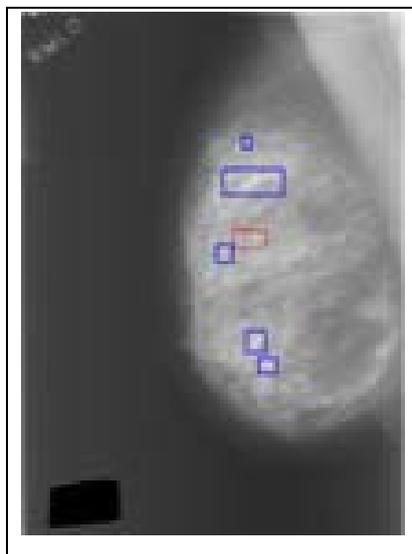


Figura 29: Imagem resultante final do método de Detecção de Agrupamentos de Microcalcificações: seis agrupamentos foram detectados. O agrupamento correto está marcado em vermelho.

3. RESULTADOS

Os testes foram realizados em um computador *Pentium 4 2.6 Ghz*, com *1024 MB* de memória RAM e controlado pelo Sistema Operacional *Linux*. O software utilizado foi o *Mathworks Matlab R2007a*.

As imagens mamográficas utilizadas nos testes pertencem à base de dados *DDSM* “*Digital Database for Screening Mammography*” da *University of South Florida*. Esta base é pública e possui grande diversidade de casos compostos por duas vistas mamográficas de cada mama (*CC* e *MLO*). As imagens estão amostradas com $200 \mu\text{m}$ e 16 bits de resolução (Guliatto, Barcelos, Dias, 2008).

No método aqui proposto, as imagens foram processadas mantendo-se as dimensões originais das mamografias digitalizadas.

3.1. Resultados do Método de Detecção do Contorno Externo da Mama

Um total de 76 imagens da base de dados *DDSM*, selecionadas aleatoriamente, foram processadas. Em 70 imagens, os contornos da mama foram detectados corretamente, delimitando de forma esperada a região ocupada pelo tecido mamário na imagem mamográfica.

Em quatro imagens, o método funcionou parcialmente, já que áreas não pertencentes à mama foram contornadas, pois o fundo da imagem possuía níveis de cinza mais elevados. No entanto, esses quatro casos não foram considerados falha total, pois não houve perda de informações relevantes da mama, como lesões, tumores, microcalcificações etc. Em duas imagens o método falhou completamente, pois apenas uma região parcial do tecido mamário foi detectada devido ao baixo contraste da mama em relação ao fundo do mamograma, o que comprometeria resultados de processamentos posteriores.

3.2. Resultados do Método de Extração do Músculo Peitoral

Um total de 55 imagens MLO da base de dados *DDSM*, selecionadas aleatoriamente, foram processadas. Para 49 imagens, o músculo peitoral foi detectado satisfatoriamente. Para sete imagens, o método funcionou parcialmente e em apenas um mamograma o músculo peitoral não foi detectado por conter artefatos e ruídos que comprometeram sua identificação, ou devido ao músculo peitoral não estar bem definido na imagem mamográfica, ou por causa de regiões densas pertencentes ao tecido mamário estarem bem próximos ao músculo peitoral.

O tempo de processamento médio por imagem variou entre 10 e 25 segundos. Esta diferença mostra que a quantidade de iterações influi diretamente no tempo e varia de imagem para imagem. Assim, as imagens mamográficas que possuírem o músculo peitoral bem definido e com a densidade contrastada em relação ao tecido fibroglandular, gastarão menos iterações no processo iterativo para refinar o músculo peitoral, pois será mais fácil segmentá-lo do que em uma imagem em condições adversas, por requerer mais tentativas para encontrar o limiar ideal nesta fase.

3.3 Resultados do Método de Detecção de Agrupamentos de Microcalcificações

Um total de 19 imagens da referida base de dados foram analisadas. Dentre elas, 17 continham microcalcificações e duas não continham. Para todos os casos, as imagens tiveram o fundo e o músculo peitoral removidos, restando apenas o tecido fibroglandular na imagem. Os resultados foram analisados por um médico especialista e foram separados nas seguintes classes, conforme Tabela 1.

I - Detectou corretamente o(s) agrupamento(s) de microcalcificações;

II - Detectou o(s) agrupamento(s) de microcalcificações e outros objetos (macrocalcificações, artefatos, ruídos e lesões);

III - Não detectou nenhum agrupamento de microcalcificações ou nenhum objeto;

IV - Não detectou nenhum agrupamento de microcalcificações, mas detectou outros objetos;

V - O mamograma não continha agrupamento de microcalcificações e nenhum objeto foi detectado;

VI - O mamograma não continha microcalcificações e outros objetos foram detectados.

Tabela 1: Classificação dos resultados nas classes descritas

Classe	Quantidade	%
I	3	15,8%
II	10	52,7%
III	0	0,0%
IV	4	21,0%
V	0	0,0%
VI	2	10,5%
TOTAL	19	100%

O método detectou algum agrupamento de microcalcificações em aproximadamente 68,5% dos mamogramas analisados, ou seja, são os casos pertencentes às classes I e II. Nestas imagens, as microcalcificações estão originalmente mais enfatizadas comparadas com microcalcificações de imagens pertencentes a outras classes.

Nos casos onde o método não funcionou corretamente, classes IV e VI, as microcalcificações estão mal definidas e com baixa intensidade de níveis de cinza. Assim, outros objetos maior intensidade de nitidez são realçados e ocorre a detecção de falsos positivos.

6. CONSIDERAÇÕES FINAIS

Este trabalho apresentou um novo método que é composto por uma fase de segmentação, uma fase de realce de microcalcificações e a fase final de detecção dos agrupamentos de microcalcificações utilizando-se de técnicas simples de processamento de imagens. Apesar de todas as dificuldades encontradas relativas à qualidade das imagens mamográficas, o método apresentou resultados promissores, dando ênfase para a fase de segmentação.

O realce de microcalcificações contribuiu para melhorar a qualidade da imagem. Durante a digitalização da mamografia, o brilho característico destes sinais foi perdido e os mesmos passaram a ter valores de níveis de cinza muito baixos. Esta etapa recuperou parte desta alta intensidade de níveis de cinza que as microcalcificações originalmente deveriam possuir.

O método de detecção de agrupamentos de microcalcificações identificou todos os sinais mais evidentes na imagem realçada. A alta detecção de falsos positivos é inevitável nos casos em que a imagem possui baixo contraste.

Por fim, o desenvolvimento um sistema que forneça um diagnóstico seguro e preciso para a detecção precoce do câncer de mama não é tarefa simples. Espera-se, como resultado, que este trabalho contribua com o aumento da sensibilidade nos exames de radiologia de mama, com a redução de biopsias e conseqüentemente com a redução de riscos para a mulher e com a redução de custos com a saúde.

7. REFERÊNCIAS BIBLIOGRÁFICAS

DENGLER J., BEHRENS S., DESAGA J. F. Segmentation of Microcalcifications in Mammograms, IEEE Transactions on Medical Imaging, Vol. 12, n. 4, p. 634-642, 1993.

FERRARI R. J., RANGAYYAN R. M., DESAUTELS J. E. L., FRÉRE A. F. Segmentação de mamogramas: Identificação da borda da mama, músculo peitoral e disco glandular, Anais CBEB, p. 1255-1261, 2000.

GONZALEZ, R. C., WOODS, R. E., 2002. Digital Image Processing, Prentice Hall, 793 p., 2002.

GULIATO D., BARCELOS C. A. Z. e DIAS W. B. O uso de Equações de Difusão no Processo de Detecção de Regiões Suspeitas em Mamografias, In: VIII Workshop de Informática Médica, 2008, p. 1-10, 2008.

INCA - Instituto Nacional de Câncer, 2008, Incidência de câncer no Brasil, Ministério da Saúde. Disponível em: <http://www.inca.gov.br>: Acesso em: Março de 2010.

KWOK S. M., CHANDRASEKHAR R., ATTIKIIOUZEL Y. Automatic Pectoral Muscle Segmentation and Cliff Detection, Seventh Australian and New Zealand Intelligent Information Systems Conference, p. 67-72, 2001.

MARQUES P. M. A., 2001, “Diagnóstico Auxiliado por Computador na Radiologia”, In: Radiol Brás, p. 295-293, 2001.

NISHIKAWA R. M., GIGER M. L., Doi K., VYBORNÝ C. J. e SCHMIDT R. A., 1993, Computer-aided detection of clustered microcalcifications: An improved method for grouping detected signals, Med. Phys., p. 1661-1666, 1993.

PEDRINI H. e SCHWARTZ W. R. Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações, Thomson Pioneira, 528 p., 2007.