

# Mapa Interativo das Estruturas da Mama Usando KD-Tree+

Vanessa Rodrigues Coelho Leite<sup>1</sup>, Edgar Moraes Diniz<sup>1</sup>, Aristófanés Correa Silva<sup>2</sup>,  
Anselmo Cardoso de Paiva<sup>1</sup>, Alexandre César Muniz de Oliveira<sup>1</sup>

<sup>1</sup>Departamento de Informática, Universidade Federal do Maranhão. <sup>2</sup>Departamento de Engenharia Elétrica, Universidade Federal do Maranhão  
São Luís – MA – Brasil

**Resumo.** *Este artigo apresenta uma nova técnica para a segmentação de mamogramas. Com base em características da imagem radiográfica, aplicou-se o KD-Tree+ para divisão da imagem em vários grupos, nos quais podemos identificar algumas estruturas importantes como nódulos e calcificações. A técnica proposta foi aplicada em imagens da base de dados Mini-MIAS e os resultados foram avaliados através de inspeção médica.*

## 1. Introdução

No Brasil, o câncer de mama é o câncer que mais causa mortes entre as mulheres. A incidência de câncer de mama no Brasil em 2006 foi de 48.930 casos, sendo o segundo mais incidente [INCA sd]. Estima-se que uma em cada oito mulheres desenvolverá câncer de mama no decorrer de sua vida [NCI sd].

A mamografia é o exame que permite identificar o nódulo antes que ele seja palpável ou se manifeste de outras formas e em 90% dos casos o câncer de mama é curável, desde que detectado precocemente, sendo este o método mais eficaz para o seu tratamento [IBCC sd].

Dessa forma, agrupar estruturas semelhantes se torna uma tarefa essencial para auxiliar na detecção precoce do câncer. Vários métodos já foram propostos com este mesmo fim. Em [Brás Jr *et. al.* 2006] a segmentação das estruturas é feita com base no K-means, em [Anguh e Silva 1997] a segmentação dessas estruturas utiliza como base o histograma, fazendo um mapeamento com pseudo-cores.

Apresentamos um algoritmo denominado KD-Tree+, que combina características do algoritmo KD-Tree Original [Bentley 1975] e do algoritmo Adaptive KD-Tree [Heredia *et. al.* 2003] de modo a gerar árvores balanceadas. Adicionalmente, apresentamos uma nova aplicação de KD-Tree's (árvores k dimensionais), de um modo geral: agrupamento e visualização de estruturas da mama em imagens de mamografia. Esta aplicação permite a identificação de lesões com maior facilidade, dentre elas calcificações e massas, auxiliando o especialista na análise de mamografias, que em geral é um processo repetitivo e cansativo: sujeito a erros, principalmente de identificação de estruturas após uma seqüência de análises.

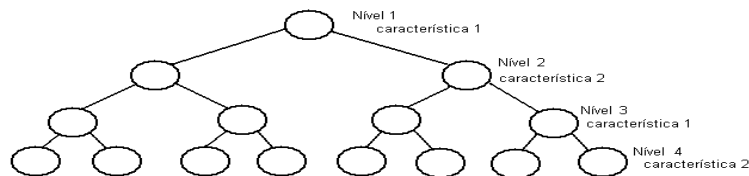
## 2. Métodos

### 2.1 KD-Tree Original e Adaptive KD-Tree

KD-Tree Original, ou simplesmente KD-Tree (*k dimensional tree*) foi proposta inicialmente por Bentley [Bentley 1975] para resolver o problema de busca baseada em

vários atributos, ou seja, encontrar os pontos que satisfazem mais de uma característica ao mesmo tempo. Consideramos, basicamente, dois tipos de KD-Tree: a KD-Tree Original e a Adaptive KD-Tree [Heredia *et. al.* 2003].

KD-Tree é uma estrutura de dados que representa uma árvore na qual cada característica utilizada na busca passa a corresponder a uma dimensão da árvore. Assim, não é preciso manter  $k$  árvores binárias de busca, uma para cada característica, pois a KD-Tree utiliza todos os atributos como indexadores, um a cada nível, ciclicamente. Chamamos de nível cada divisão horizontal, assim o nó raiz está no nível 1, os filhos deste estão no nível 2, e assim, sucessivamente. Supondo uma árvore construída com base em duas características, sua estrutura se apresentaria conforme Figura 1.



**Figura 1. Modelo genérico de uma KD-Tree.**

Na KD-Tree cada nó possui dois descendentes, não importando o número  $k$  de dimensões. Em cada nó acessado é testado um valor chave. Esse valor chave está associado a um discriminador, que indica qual dimensão da KD-Tree foi utilizada naquele nível. Assim é possível determinar o valor chave que separa os nós restantes em duas sub-árvores. Os nós com os valores dessa dimensão maiores que o valor chave são armazenados na sub-árvore à direita e os restantes na sub-árvore à esquerda do nó acessado, entretanto, um dos problemas que podemos observar na KD-Tree é que ela depende da ordem em que os nós são inseridos.

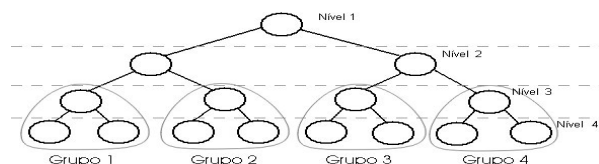
A Adaptive KD-Tree funciona semelhantemente à KD-Tree Original, a diferença entre essas estruturas é que enquanto na KD-Tree Original cada nó da árvore armazena as informações necessárias e o discriminador é escolhido com base na dimensão utilizada, na Adaptive KD-Tree, as informações são armazenadas apenas nos nós folhas (os nós intermediários armazenam um valor relacionado ao discriminador) e o discriminador é escolhido de maneira a dividir a árvore em duas sub-árvores, o mais balanceado possível.

## 2.2 KD-Tree+

Para a construção do KD-Tree+ propomos a utilização de características das duas árvores: escolha do melhor discriminador (característica da Adaptive KD-Tree), e armazenamento das informações em cada nó (característica da KD-Tree Original). Podemos escolher o melhor discriminador visto que os pontos que formarão a árvore são conhecidos. O discriminador que utilizamos é a mediana, uma medida de localização do centro da distribuição dos dados, ou seja, o valor que divide os pontos para que 50% sejam menores ou iguais a ele e 50% maiores ou iguais, garantindo assim o balanceamento da árvore. Para cada nível da árvore é feita esta política de escolha, com base na característica pertencente àquele nível. A partir disso constrói-se a árvore.

Assim, com a árvore construída, podemos gerar os agrupamentos das estruturas baseado nos níveis e nos grupos. Os grupos relacionam-se diretamente com os níveis, ou seja, quanto maior o valor do nível mais grupos teremos formados: seja  $n$  o número do

nível escolhido, para este nível teremos  $2^{n-1}$  grupos. Por exemplo, se o observamos o nível 3 teremos 4 grupos formados (Figura 2).



**Figura 2. Divisão dos grupos no nível 3.**

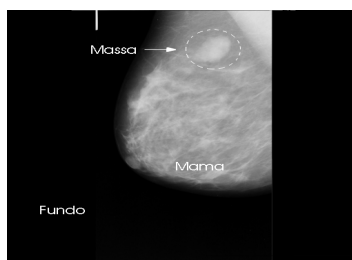
### **3. Aplicação de KD-Tree+ para mapear estruturas da mama**

O KD-Tree+ foi aplicado em imagens de mamografia para agrupar as estruturas com características semelhantes como tecidos e lesões, de forma a gerar um mapa dessas estruturas. E para tanto utilizamos uma característica da imagem: o pixel.

Com a árvore construída, geramos as imagens que representam o mapeamento. Para cada nível, é feita uma busca, que retorna os nós a partir do qual os grupos serão construídos. Cada nó retornado, juntamente com todos os seus filhos, formará um grupo ao qual será atribuída uma pseudo-cor, escolhida de forma aleatória. Quanto mais grupos formamos maior é a especificação de cada grupo. Este mapeamento permite a visualização das estruturas como massas, nódulos, calcificações dentre outras.

As imagens utilizadas foram obtidas do banco de dados mini-MIAS (*Mammographic Image Analysis Society*) [Suckling *et al* 1994]. Este banco de dados é composto por 322 imagens de mamografia (mama esquerda e direita) de 161 pacientes. As imagens foram digitalizadas na resolução 1024x1024 *pixels*, em padrão de cinza de 8 bits.

A Figura 3 apresenta uma imagem obtida do mini-MIAS.

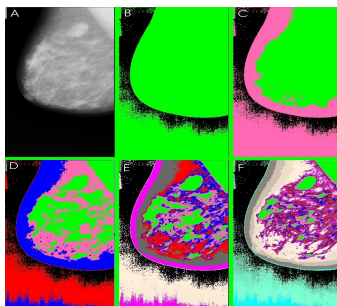


**Figura 3. Mamografia obtida do mini-MIAS, identificando os elementos da imagem.**

Para facilitar a visualização dos grupos gerados pela KD-Tree+, desenvolvemos uma ferramenta que é um módulo do CAMBI, que é um sistema de identificação e classificação de massas em mamas, que utilizamos para ler e abrir a imagem, inicializar a aplicação do KD-Tree+ (para construir a árvore), e gerar as imagens.

Após as imagens geradas, podemos fazer a visualização dos diversos grupos, de forma iterativa, assim fazemos um detalhamento das estruturas até a visualização da estrutura desejada pelo usuário. A Figura 4 representa a visualização da mamografia com massa em diversos níveis. A imagem A representa o nível 0 (zero), ou seja, a imagem original. A partir da imagem B até a imagem F temos os níveis de 1 a 5. Em B, por estarmos no nível 1, teremos a imagem completa como um único grupo, em D já podemos

perceber o refinamento dos grupos, e a melhor divisão das estruturas, em F temos o nódulo bem destacado.



**Figura 4. Imagens obtidas após aplicação do KD-Tree+. Mamografia com nódulo: A – Imagem original. B a F – Visualização dos níveis 1 a 5, respectivamente.**

#### 4. Conclusão

Visto que o câncer de mama tem sido a principal causa de morte entre as mulheres e que o diagnóstico precoce aumenta as chances de cura, propusemos neste artigo uma nova técnica, KD-Tree+, para auxiliar na segmentação das estruturas da mamografia, juntamente com uma ferramenta que auxilia a visualização das estruturas segmentadas de forma interativa. Facilitando assim a visualização das estruturas desejadas pelo especialista, mesmo aquelas que possivelmente passariam despercebidas, pois é possível destacá-las sobre a imagem original.

#### 5. Referências

- Anguh, M. M. and Silva, A. C. (1997) “Multiscale Segmentation and Enhancement in Mammograms”
- Bentley, J.L. (1975) “Multidimensional Binary Search Trees Used for Associative Searching.” *Communications of the ACM*, 18, 509-517.
- Brás Jr, G., Silva, A. C., Paiva, A. C. and Oliveira, A. C. M. de. (2006) “Identificação de Massas em Mamografias usando Textura, Geometria e Algoritmos de Agrupamento e Classificação”- In: VI Workshop de Informática Médica (WIM’2006), Vila Velha.
- INCA – Instituto Nacional do Câncer. “Estatísticas do Câncer”, [http://www.inca.gov.br/conteudo\\_view.asp?id=336](http://www.inca.gov.br/conteudo_view.asp?id=336)
- Heredia, L., Iochpe, C. and Comba, J. (2003) “Explorando a Multidimensionalidade da *Kd-Tree* para Suporte a Temporalidade em Dados Espaciais Vetoriais do Tipo Ponto.” *Proceedings of GeoInfo 2003*
- I.B.C.C. - Instituto Brasileiro de Combate ao Câncer. <http://www.ibcc.org.br>
- N.C.I. - National Cancer Institute. “Câncer” <http://www.cancer.gov>
- Suckling *et. al.* (1994) – Suckling, J., Parker, J., Dance, D. R., Astley, S., Hutt, I., Boggis, C. R. M., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S. L., Taylor, P., Betal, D. and Savage, J. “The Mammographic Image Analysis Society Digital Mammogram Database”. *Proc. Of The 2nd Intern. Workshop On Digital Mammography*, p.375-378.