

# Estruturas Tritree

## Compressão de imagens coloridas

Tiago Oliveira Cunha<sup>1</sup>, Vânia Cordeiro da Silva (orientadora)<sup>1</sup>  
<sup>1</sup>Departamento de Ciências Exatas e Tecnológicas -DCET  
Universidade Estadual de Santa Cruz (UESC) – Ilhéus – BA – Brazil  
tiago\_cic@yahoo.com.br, [vania@uesc.br](mailto:vania@uesc.br)

### Resumo.

Um novo método para segmentação/decomposição aplicada à compressão de imagens foi criado [5], o algoritmo de decomposição Tritree (TT). O objetivo é a segmentação da imagem em um conjunto de regiões triangulares com cores homogêneas, onde a diferença entre os valores de pixels não ultrapassem um determinado limite. Uma árvore é construída para representar a decomposição. Cada triângulo será um nó da TT. O triângulo inicial contendo toda a imagem é a raiz. Os últimos triângulos, que representam a imagem comprimida, são as folhas (nós terminais). Resultados experimentais com imagens em tons de cinza mostram que o método proposto possui resultados melhores do que a decomposição Quadtree (QT), em termos de PSNR e a taxa de compressão. É testado agora, o método desenvolvido para imagens coloridas.

### 1. Introdução

Muitas técnicas de compressão de imagens são baseadas em segmentação ou decomposição da imagem original em regiões homogêneas, o que pode ser representado por um reduzido número de bits. Este é o caso do conhecido método de decomposição *Quadtree* (QT), que tem sido empregado em diversos métodos híbridos de compressão de imagens encontrados na literatura [3, 4, 6, 7].

Este artigo apresenta a técnica de decomposição *Tritree* (TT), para representação de imagens reais bidimensionais, em tons de cinza, em diferentes níveis de resolução, desenvolvido em [5], e estende-a para imagens em cores espectrais. A decomposição TT é similar à decomposição QT. Contudo, enquanto QT trabalha subdividindo a imagem em regiões quadráticas progressivamente menores, a decomposição TT o faz para regiões triangulares. A meta é segmentar a imagem em um conjunto de regiões triangulares homogêneas, isto é, regiões onde a diferença dos valores de nível de cor dos *pixels* não exceda um determinado limiar. A estrutura de dados árvore de pesquisa é empregada para representar a decomposição. Cada elemento, ou triângulo, forma um nodo da árvore. O elemento inicial, que contém a imagem completa, é a raiz da árvore e suas folhas (nós terminais) guardam a

informação relevante da imagem comprimida, ou seja, a informação necessária para sua descompressão. Todas as árvores foram construídas por um procedimento *top-down*.

A aplicação mais imediata para este método seria em algoritmos híbridos de compressão de imagens. Entretanto, várias outras aplicações que fazem uso de decomposição ou segmentação, também seriam viáveis.

### 2. Imagens Coloridas

Em uma imagem digital colorida no sistema RGB, pode-se considerar um pixel como um vetor  $f(i,j) = (fr(i,j), fg(i,j), fb(i,j))$ , cujas componentes representam as intensidades de vermelho, verde e azul da cor. Pode-se considerar, portanto, que uma imagem colorida é a composição de três imagens monocromáticas,  $fr(i,j)$ ,  $fg(i,j)$ ,  $fb(i,j)$ , denominadas, respectivamente, de *banda vermelha* (ou *banda R*), *banda verde* (ou *banda G*), e *banda azul* (ou *banda B*) da imagem colorida, conforme ilustra a Figura 1.

O modelo RGB é o mais utilizado para representar cores em hardwares e aplicações que a suportam. É o modelo empregado nos dispositivos de visualização gráficos [2].



Figura 1. Imagem colorida e suas bandas

### 3. Decomposição Tritree

Faz-se necessário algumas definições, observando o armazenamento das informações na estrutura de dados árvore. Cada item de dado armazenado numa árvore é chamado de nodo. O nodo inicial é a raiz. Sub-árvores podem originar de cada nodo, estabelecendo uma relação de pai-filho entre os nodos. Um nodo que não possui filhos é chamado nodo terminal ou folha. O número de filhos de um nodo determina seu grau. O número de gerações entre um nodo e a raiz da árvore determina o nível do nodo.

Todo o processo é realizado em cada uma das 3 bandas da imagem colorida. Decomposição TT começa pelo enquadramento da imagem em um triângulo equilátero, que é sempre sub-dividido em quatro novos triângulos equiláteros. O grau de homogeneidade é calculado para cada triângulo gerado e testado contra um valor de limiar dado empiricamente pelo usuário, para determinar se este deve ser novamente dividido ou não. Este processo é repetido um número fixo de vezes ou até todos os triângulos testados serem rotulados como homogêneos. Cada triângulo será um nodo da TT. O triângulo inicial, que contém a imagem toda, será a raiz. Os triângulos finais, representando a imagem decomposta, são as folhas.

A Figura 2 exemplifica a decomposição TT de uma imagem qualquer, resultando em uma árvore de grau 4. Uma TT com n níveis pode ter até  $4^{n-1}$  folhas. A Figura 3 mostra como as informações dos nodos da TT são armazenadas, utilizando um registro com 9 campos para cada nodo.

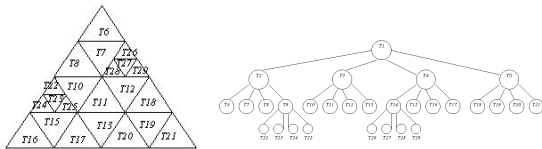


Figura 2: Informações de uma Estrutura Tritree

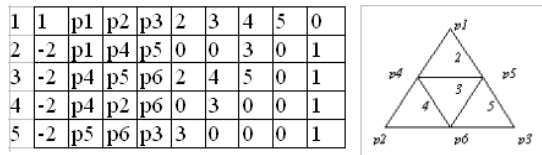


Figura 3: Informações em uma Estrutura Tritree

A lista dos vértices dos triângulos é organizada em uma árvore binária (grau 2). Registros com cinco campos são necessários para armazenar as informações de cada nodo da árvore binária: dois

campos para as coordenadas X-Y dos vértices, dois outros campos apontando para os vértices filhos, e o último guarda a informação do valor da cor do vértice. Quando um triângulo é dividido novos nodos são adicionados na árvore de vértices (a árvore binária). Estes nodos são os pontos do meio de cada aresta do triângulo, e correspondem aos vértices dos quatro novos triângulos criados pela decomposição. Os pontos são inseridos na árvore binária em ordem crescente, de acordo com os valores X-Y de suas coordenadas. Um ponto com coordenada (x,y) é definido como sendo menor do que outro ponto com coordenadas (u,v), respectivamente, se:

$$\begin{cases} x < u & \text{ou} \\ x = y & e \ y < v \end{cases}$$

A Figura 4 apresenta como a imagem é enquadrada no triângulo inicial. A largura e altura do triângulo são determinadas de acordo com as dimensões da imagem, garantindo, assim, o enquadramento perfeito. A altura  $a$  do triângulo raiz é dado por:  $a = l + c * \text{sen}60^\circ$  Onde  $l$  é o número de linhas da imagem e  $c$  é o número de colunas. A largura  $b$  do triângulo é calculada como se segue:  $b = c + 2 * l * \text{cos}60^\circ / \text{sen}60^\circ$  Em geral,  $a$  e  $b$  são valores fracionários, requerendo aproximação para o mais próximo, um dos valores da malha de coordenadas da imagem discreta. Como consequência, os cantos superiores da imagem, muito provavelmente, não tangenciam as arestas do triângulo inicial como ilustrado na Figura 4.

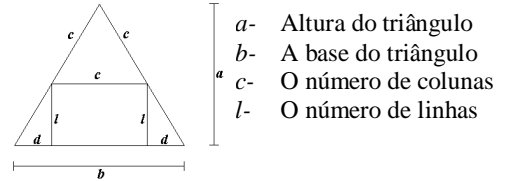


Figura 4: Imagem Enquadrada em Triângulo Inicial

Internamente, a estrutura de dados utilizada para representar imagens é uma matriz bi-dimensional [4].

### 4. Reconstrução

A reconstrução da imagem é realizada através da reconstrução dos triângulos produzidos pela decomposição TT. Dois diferentes algoritmos de reconstrução foram pesquisados: replicação da média e interpolação linear [3]. O primeiro é representado pela média dos valores de níveis de cor do triângulo, a reconstrução consiste da replicação deste valor médio dentro dos limites físicos dos triângulos. A interpolação linear é executada em cada triângulo-folha (nodo terminal) da TT, de acordo com a equação:  $ax + by + cz = d$ .

Abaixo, na Figura 5, as imagens reconstruídas utilizando o método Tritree com a reconstrução feita por interpolação linear.



Figura 5: Imagens reconstruídas

## 5. Codificação

O procedimento de codificação inclui codificar a informação da estrutura da árvore, bem como codificar os níveis de cor dos vértices dos triângulos. Para a codificação da estrutura da árvore o valor '1' é utilizado para rotular os nodos pais e '0' é reservado para rotular os nodos folhas [5][6]. Todos os nodos do último nível de decomposição são folhas, portanto, o processo de codificação da estrutura da árvore pode parar no penúltimo nível. O código da raiz, único nodo do primeiro nível, será sempre '1', podendo ser suprimida, também, esta informação. Por exemplo, o código resultante para a árvore TT apresentada na Figura 2 é 1111 - 0001 - 0000 - 1000 - 0000.

A informação do nível de cor de uma imagem reconstruída por interpolação linear entre os vértices dos triângulos é armazenada utilizando-se codificação preditiva sem perdas. Somente os vértices dos nodos folhas são armazenados. Um ponto de vértice, o primeiro a ser calculado, é escolhido para ser o preditor e o único a ser codificado com 8 bits (para o caso de imagens com 256 níveis de cor). Para os outros pontos de vértice, somente a diferença entre estes e o preditor é codificada, com o mesmo número de bits do limiar. Isto acontece por que triângulos adjacentes tem, no mínimo, um vértice em comum [2][3].

## 6. Teste de Homogeneidade

Foram utilizados dois testes de homogeneidade, no primeiro, descobrem-se os pixels de maior e menor valor em seguida a diferença entre os dois é com parada com o limiar de decisão informado anteriormente (foram utilizados dois tipos de limiares, o primeiro é passado de forma empírica, o

segundo utiliza um valor passado empiricamente para calcular um novo valor de limiar para cada nível da árvore TT é utilizado em conjunto com o teste de homogeneidade desvio da média), caso seja maior há divisão, senão o triângulo é considerado terminal, o segundo é o desvio da média [7], a decisão ou não da divisão é feita através da comparação do valor absoluto da diferença dos valores de níveis de cor de cada pixel do triângulo com o valor médio dos níveis de cor dos pixels do triângulo, essa diferença é comparada com o valor do limiar de divisão caso seja maior há a divisão.

## 7. Medidas de Avaliação

A eficiência da compressão é medida pela taxa de compressão, expressa como a média (Tx) do número de bits requeridos por pixel para codificar a imagem comprimida, dado por:

$$Tx \text{ bpp} = \frac{\text{Número de bits para codificar a imagem comprimida}}{\text{Total de pixels na imagem original}}$$

A qualidade da imagem reconstruída é mensurada pela relação sinal-ruído de pico, ou PSNR, do inglês “peak signal-to-noise ratio”, calculada como:

$$PSNR = 10 \cdot \log_{10} \left\{ \frac{A^2}{MSE} \right\} dB$$

Nesta equação, A representa o valor máximo de nível de cor que os pixels da imagem testada podem alcançar. Nos testes realizados foram utilizadas apenas imagens com 8 bits/pixel, ou seja, o valor de A foi sempre igual a 255. A variável MSE é o erro médio quadrático (Mean Square Error) da diferença, ou erro, entre o sinal da imagem original (I(i,j)) e a imagem comprimida (R(i,j)). Sendo L o número de linhas das imagens e C o número de colunas, o MSE é calculado como se segue:

$$MSE = \frac{1}{LC} \left\{ \sum_{i=0}^L \sum_{j=0}^C [I(i,j) - R(i,j)]^2 \right\}$$

## 8. Resultados Experimentais

Serão apresentados aqui os resultados relativos à imagem Lena por se tratar da imagem mais utilizada em testes de algoritmos de compressão de imagens.

Nos testes foram utilizados os métodos Tritree e Quadtree, sendo que no primeiro foram testadas duas estratégias de reconstrução, replicação da média das cores e interpolação linear das cores e dois testes de homogeneidade, diferença entre maior e menor pixel e desvio da média, no segundo somente a replicação da média e diferença entre maior e menor pixel.

Os cálculos foram feitos para as bandas da imagem logo cada medida avaliativa vem acompanhada da banda a qual faz referência, também foram calculados o número de vértices e o

número partições (triângulos ou quadrado, de acordo com o método) gerados por cada método, os resultados obtidos para o método Tritree estão expostos nas Tabelas 1 (reconstrução por replicação da média), Tabela 2 (reconstrução por interpolação linear) e Tabela 4 (reconstrução por replicação e desvio da média), os resultados referentes ao método Quadtree são expostos na Tabela 3.

Tabela 1:Tritree: reconstrução por replicação da média.

Limiar	Vértices	Partições	MSE(R)	MSE(G)	MSE(B)	PSNR(R)	PSNR(G)	PSNR(B)
5	188028	415645	106,00	105,00	80,00	27,8777	27,9189	29,0999
10	129612	277079	153,00	121,00	94,00	26,2838	27,3029	28,3995
20	78109	159949	200,00	153,00	125,00	25,1205	26,2838	27,1617
30	55921	111527	234,00	184,00	152,00	24,4386	25,4826	26,3123
50	33989	66106	307,00	253,00	215,00	23,2594	24,0995	24,8064

Tabela 2:Tritree: reconstrução por interpolação linear.

Limiar	Vértices	Partições	MSE(R)	MSE(G)	MSE(B)	PSNR(R)	PSNR(G)	PSNR(B)
5	188028	415645	59,00	43,00	41,00	30,4222	31,7961	32,0029
10	129612	277079	107,00	59,00	55,00	27,8369	30,4222	30,7271
20	78109	159949	154,00	92,00	88,00	26,2555	28,4929	28,6859
30	55921	111527	191,00	123,00	116,00	25,3204	27,2317	27,4862
50	33989	66106	264,00	197,00	180,00	23,9147	25,1861	25,5780

Tabela 3:Quadtree original.

Limiar	Vértices	Partições	MSE(R)	MSE(G)	MSE(B)	PSNR(R)	PSNR(G)	PSNR(B)
5	555928	890084	180,00	107,00	90,00	25,5780	27,8369	28,5883
10	353523	546892	184,00	111,00	94,00	25,4826	27,6775	28,3995
20	191293	282384	197,000	125,00	106,00	25,1861	27,1617	27,8777
30	126421	180024	216,00	144,00	125,00	24,7862	26,5471	27,1617
50	66384	90676	272,00	194,00	181,00	23,7851	25,2527	25,5540

Tabela 4:Tritree: reconstrução por interpolação linear, desvio da média.

Limiar	Vértices	Partições	MSE(R)	MSE(G)	MSE(B)	PSNR(R)	PSNR(G)	PSNR(B)
5	270222	508860	83,00	48,00	48,00	28,9400	31,3183	31,3183
10	147478	278422	131,00	76,00	73,00	26,9580	29,3226	29,4975
20	72849	137042	188,00	135,00	116,00	25,3892	26,8274	27,4862
30	45733	86136	234,00	185,00	165,00	24,4386	25,4590	25,9559
50	23125	44862	344,00	296,00	273,00	22,7652	23,4178	23,7691

Observando-se a coluna vértices e partições das tabelas, pode-se observar que o número de vértices e partições gerados pela Quadtree é sempre maior que o número de vértices e partições gerado pela Tritree. Este fato é demonstrado matematicamente por Mayer em [6], que conclui no final de sua pesquisa: "Dada uma imagem I, particionada em um dado número de regiões de mesma área A, o particionamento triangular requer cerca de metade dos vértices requeridos para o particionamento em quadrados". Para o particionamento em regiões de diferentes áreas, esta afirmação pode ser verificada empiricamente através dos experimentos, tanto neste trabalho como no de Mayer [6]. Isto implica em uma quantidade menor de informação passada ao descompressor, ou seja, mais compressão.

Analisando as colunas que contém as PSNRs, pode-se verificar a vantagem obtida pelo método Tritree com reconstrução por interpolação e diferença entre o maior e o menor pixel, sobre as outras três combinações e visto que o método Tritree ainda se encontra em desenvolvimento pode-se esperar por ganhos ainda maiores em relação ao método Quadtree, que na concorrência com o método Tritree sempre perdeu em termos de PSNR, MSE, vértices e partições com mesma imagem e limiar de divisão.

## 9. Conclusão

Os resultados parciais confirmam a viabilidade do algoritmo TT para compressão de imagens coloridas, por produzir, simultaneamente, PSNR superior e número de vértices e partições inferior aos do algoritmo QT, um dos mais populares esquemas de codificação baseada em segmentação aplicado em compressão de imagens.

A utilização da decomposição TT nos mesmos métodos híbridos de compressão de imagens que fazem uso da QT geraria resultados superiores. As aplicações que fazem uso de decomposição ou segmentação de imagens, também poderiam utilizar a TT, como foi observado em imagens em níveis de cinza. Isto posta, a gama de possíveis aplicações da TT é consideravelmente ampliada.

## 10. Referências

- [1] Silva, V.C. e Carvalho, J.M. Image Compression by Tritree Decomposition. *Scientia - Unisinos*, 13(2):133–144, Julho/Dezembro 2002.
- [2] Bárbara, R.M.B e Alcaim, A. Compressão de Imagens Baseada em Decomposição Quadtree com Limiares Variáveis, Alocação de bits Adaptativa e DCT. SBT volume 38(3), March 1999.
- [3] Sullivan, G.J. e Baker, R.L. Efficient Quadtree Coding of Images and Video. *IEEE Transaction on Image Processing*, 3(3), May 1994.
- [4] Wille, S.O. Structured Tri-Tree Search Method for Generation of Optimal Unstructured Finite Element Grids in Two and Three Dimensions. *International Journal for Numerical Methods in Fluids*, 14:861–881, 1992.
- [5] Silva, V.C. Decomposição Triangular de Imagens: uma Aplicação em Compressão. Tese de doutorado,
- [6] Mayer, J. *Blending Models for Image Enhancement and Coding*. Tese de doutorado, Universidade da Califórnia, Santa Cruz, Dezembro 1999.
- [7] Cunha, T.O. e Silva, V.C. Decomposição Triangular de Imagens Coloridas: Uma Aplicação em Compressão. IV Semanainfo, Faculdade Metodista Granbery, Juiz de fora, MG, maio de 2007.
- [8] Cunha, T. O. ; Silva, V. C. . Compressão de Imagens Coloridas Utilizando Estruturas Tritree. In: VIII ERBASE - Escola Regional de Computação Bahia - Alagoas - Sergipe, 2008, Salvador/BA. Anais do VIII ERBASE - Escola Regional de Computação dos Estados da Bahia, Alagoas e Sergipe, 2008.