

ARMAZENAMENTO DE MAPAS POLIGONAIS UTILIZANDO ESTRUTURAS QUADTREE**RICARDO LOMBA DE ARAUJO**

PETROBRÁS - Depex/Ditrex/Sepran - Av. Chile 65 Sala 1507 Rio de Janeiro

LUIS PAULO VIEIRA BRAGA

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO C.P. 68530 Rio de Janeiro CEP 21945

RESUMO

O presente trabalho tem como objetivo a definição e implementação de uma estrutura de dados voltada para o modelo vetorial . É tratada especificamente uma variante da "quadtree" para o arquivamento de informações do tipo linear (polígonos), de grande aplicação nas áreas de Cartografia Automatizada e Computação Gráfica em geral .

1 - ESTRUTURAS "QUADTREE"

As estruturas de dados hierárquicas têm grande importância e utilização nas áreas de computação gráfica, processamento de imagens e em sistemas de informações geográficas, devido ao fato de permitirem a focalização dos dados na forma de subconjuntos, conforme interesses específicos . Esta habilidade resulta em uma representação eficiente e na otimização do tempo de acesso e recuperação das informações, sendo portanto muito adequada para a realização de operações com conjuntos .

Em seu mais amplo contexto, o termo quadtree é usado para descrever uma classe de estruturas de dados hierárquicas que têm como propriedade comum o fato de se basearem no princípio da decomposição recursiva do espaço . Em duas dimensões, uma região quadrada é recursivamente dividida em quatro partes até que cada parte contenha apenas dados suficientemente simples para serem organizados por outra estrutura de dados . Os tipos de "quadtree" se diferenciam pelo tipo dos dados que elas representam (pontos isolados, retas, áreas, etc) e pelo critério de decomposição, que pode ser feita em partes de forma geométrica idêntica à da área original (regular) ou pode ser orientada pela ordem de entrada dos dados .

Dados em forma "raster", que não são o objetivo deste trabalho, podem ser representados em uma "region quadtree", que é construída a partir da subdivisão regular e recursiva de uma área quadrada que

envolve a região . Se um quadrante não for constituído inteiramente de "1s" ou inteiramente de "0s" ele é subdividido novamente até que se obtenham blocos uniformes .

Para dados puntiformes (vértices) pode-se utilizar a "point quadtree" na qual a região é dividida recursivamente em quatro retângulos cujo vértice comum é definido pelas coordenadas do próprio ponto pertencente à região . Neste caso , a decomposição é orientada pela ordem na qual os pontos são inseridos na estrutura .

Uma adaptação da "region quadtree" permite a representação de vértices através da "point region quadtree" . A PR "quadtree" armazena pontos em nós folhas e tem decomposição regular , que é aplicada até que nenhum quadrante contenha mais do que um vértice .

Para as regiões existem duas linhas principais de ação para sua representação : i) a que trata seu interior; ii) a que especifica as bordas . A "region quadtree" segue a primeira linha e as que visam representar polígonos seguem a segunda . Uma das variantes para representação de polígonos é a "MX-quadtree" (Hunter e Steiglitz) na qual os lados dos polígonos são representados por blocos de pixels alinhados .

Outra proposta é a "edge-quadtree" (Shneider) na qual uma região que contenha uma curva é sucessivamente dividida em subquadrantes até que cada um contenha uma feição que possa ser aproximada por um segmento de reta . Nenhuma destas estruturas entretanto é satisfatória para o arquivamento de mapas poligonais que contenham vértices ou lados isolados em conjunto com polígonos fechados, bem como regiões com áreas vazias em seus interiores . Deseja-se portanto uma estrutura de dados que armazene estes mapas sem a perda de informações resultante da digitalização, que não seja sensível ao posicionamento do mapa e que possa ser eficientemente manipulada .

2 - ESTRUTURAS PM-QUADTREE

Um mapa poligonal é um conjunto de polígonos que se interceptam apenas em vértices e que implicitamente definem uma coleção de regiões . A análise das estruturas a seguir levará em conta a realização de algumas operações sobre estes dados : localização de pontos, inserção de arestas, exclusão de arestas e superposição de mapas .

Como introdução aos critérios de decomposição convém lembrar que a "PR quadtree" usa o critério denominado C1 :

C1 : Um nó folha pode conter no máximo 1 vértice .

O limite superior para a profundidade da árvore em decorrência deste critério, denominando $DMIN(vv)$ como a menor distância entre dois vértices distintos, é dado por :

$$D1 = 1 + \log_2 \left(\frac{\sqrt{2}}{DMIN(vv)} \right)$$

Um outro critério de decomposição, denominado C2, onde define-se q-aresta como o trecho de aresta formado pela sua interseção com os limites dos nós, é dado por :

C2 : Um nó folha pode conter no máximo 1 q-aresta .

O problema de C2 é a possibilidade de se gerar árvores de profundidade infinita, que é resolvido substituindo-se C2 por C2' e C3 :

C2' : Se uma folha contém 1 vértice , então ela só pode conter q-arestas que incluam este vértice.

C3 : Se uma folha não contém vértices , então ela pode conter no máximo 1 (uma) q-aresta .

Considerando $DMIN(ev)$ como a menor distância entre uma aresta e um vértice não pertencente à esta aresta e $DMIN(ee)$ como a separação mínima entre duas q-arestas que não se interceptam , a profundidade gerada pela aplicação destes critérios é :

$$D2' = 1 + \log_2 \left(\frac{\sqrt{2}}{DMIN(ev)} \right)$$

$$D3 = 1 + \log_2 \left(\frac{\sqrt{2}}{DMIN(ee)} \right)$$

A quadtree que satisfaz C1, C2' e C3 é denominada "PM1 quadtree" . O inconveniente de C2 foi eliminado, entretanto, C3 é dependente da posição do mapa, o que não é desejável . Para isto estabelece-se o critério C3' :

C3' : Se uma folha não contém vértices , então ela pode conter apenas q-arestas que se encontram em um vértice comum exterior à ela .

Uma quadtree que satisfaz C1 , C2' e C3' é denominada "PM2 quadtree" . Deve-se ressaltar que a profundidade decorrente do critério C3' tem como limite superior $D2'$, por sua própria definição . Tem-se agora um critério para o qual a profundidade da árvore , em seu caso crítico , é menos sensível

O limite superior para a profundidade da árvore em decorrência deste critério, denominando $DMIN(vv)$ como a menor distância entre dois vértices distintos, é dado por :

$$D1 = 1 + \log_2 \left(\frac{\sqrt{2}}{DMIN(vv)} \right)$$

Um outro critério de decomposição, denominado C2, onde define-se q-aresta como o trecho de aresta formado pela sua interseção com os limites dos nós, é dado por :

C2 : Um nó folha pode conter no máximo 1 q-aresta .

O problema de C2 é a possibilidade de se gerar árvores de profundidade infinita, que é resolvido substituindo-se C2 por C2' e C3 :

C2' : Se uma folha contém 1 vértice , então ela só pode conter q-arestas que incluam este vértice.

C3 : Se uma folha não contém vértices , então ela pode conter no máximo 1 (uma) q-aresta .

Considerando $DMIN(ev)$ como a menor distância entre uma aresta e um vértice não pertencente à esta aresta e $DMIN(ee)$ como a separação mínima entre duas q-arestas que não se interceptam , a profundidade gerada pela aplicação destes critérios é :

$$D2' = 1 + \log_2 \left(\frac{\sqrt{2}}{DMIN(ev)} \right)$$

$$D3 = 1 + \log_2 \left(\frac{\sqrt{2}}{DMIN(ee)} \right)$$

A quadtree que satisfaz C1, C2' e C3 é denominada "PM1 quadtree" . O inconveniente de C2 foi eliminado, entretanto, C3 é dependente da posição do mapa, o que não é desejável . Para isto estabelece-se o critério C3' :

C3' : Se uma folha não contém vértices , então ela pode conter apenas q-arestas que se encontram em um vértice comum exterior à ela .

Uma quadtree que satisfaz C1 , C2' e C3' é denominada "PM2 quadtree" . Deve-se ressaltar que a profundidade decorrente do critério C3' tem como limite superior $D2'$, por sua própria definição . Tem-se agora um critério para o qual a profundidade da árvore , em seu caso crítico , é menos sensível

a rotações e translações do mapa poligonal . A questão passa a ser então se este critério pode ser melhorado .

Isto é feito adotando-se unicamente o critério C1 e uma forma de organização das q-arestas dentro das folhas, sendo adotado sete dicionários aplicáveis conforme a classificação da q-aresta . A classe mais natural é a das q-arestas que se encontram em um vértice contido na folha . As demais q-arestas pertencentes à folha são classificadas de acordo com a direção com que cruzam a folha, gerando outras seis classes . A quadtree assim construída , "PM3 quadtree", é ilustrada na figura 1 e é a estrutura que requer menor quantidade de memória atendendo aos objetivos mencionados inicialmente .

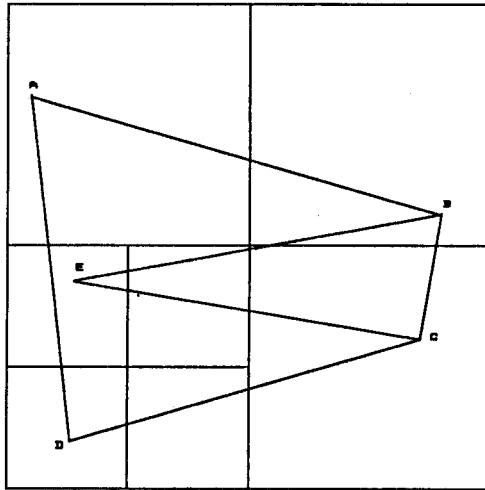


FIGURA 1 : PM3 QUADTREE

3 - SISTEMA PROTÓTIPO

3.1 - INSERÇÃO DE ARESTAS

Considerando uma árvore PM3 já existente , para inserir uma aresta AB , é feita a inserção das q-arestas que compõem AB em cada quadrante seccionado por AB . Em alguns destes quadrantes a inserção de uma q-aresta de AB pode causar a violação do critério de decomposição . Neste caso o

quadrante em questão deve ser subdividido e a inserção é então refeita . A subdivisão de um quadrante já existente na árvore , conforme citado acima , pode fazer com que as q-arestas nele existentes sejam redivididas gerando novas q-arestas . Assim, o custo final da inserção é freqüentemente devido a tantas inserções quantas q-arestas desta aresta forem posteriormente subdivididas para acomodar arestas subseqüentemente adicionadas .

Na análise de custo assume-se que o mapa está sendo construído dinamicamente , isto é , as informações sobre arestas futuras não são consideradas no momento que determinada aresta é inserida , e é um limite superior do custo real de construção do mapa quando somadas todas as inserções que perfazem o mesmo . Este custo é o produto do custo de inserir uma q-aresta pelo número de q-arestas que devem ser inseridas . O custo de inserir uma aresta é a soma da profundidade $D1$ da quadtree com a profundidade da estrutura do dicionário ($A1$) . Mais formalmente , um limite superior para o custo da construção de uma PM3 quadtree , inserindo uma aresta de cada vez , é dado por :

$$(3 \times E + L \times 2 \times 2^{D1}) \times (D1 + A1) \text{ ou } (5 \times L \times 2^{D1}) \times (D1 + A1)$$

Onde L é a soma dos comprimentos de todas as arestas que compõem o mapa e E é o número de arestas do mapa.

3.2 - EXCLUSÃO DE ARESTAS

Considerando uma árvore PM3 já existente , para excluir uma aresta AB , é feita a remoção das q-arestas que compõem AB em cada quadrante seccionado por AB . Em alguns casos a exclusão de uma q-aresta de AB pode fazer com que deixe de ser necessária , conforme o critério de decomposição , a existência de um conjunto de quatro nós filhos , o que leva à eliminação destes filhos e conseqüentemente ao grupamento das suas informações no nó pai . A junção de quatro nós filhos em um único nó (que deixa de ser cinza e passa a ser folha) , pode fazer com que algumas q-arestas neles existentes sejam grupadas , gerando uma única q-aresta .

Assim , o custo final da exclusão de uma aresta em uma árvore PM3 freqüentemente é devido ao grupamento de nós e q-arestas que não mais são necessários . A análise deste custo para a exclusão de uma aresta é similar à da inclusão , visto que são processos inversos mas que se baseiam nos mesmos princípios .

3.3 - SUPERPOSIÇÃO DE MAPAS

A operação de superposição de mapas consiste em partindo de dois mapas da mesma região e armazenados em duas estruturas PM quadtree, obter uma terceira árvore PM resultante da superposição ou união das duas árvores originais. As duas árvores A e B que se deseja superpor são percorridas e analisadas em paralelo. Quando um nó de A é uma folha e o nó de B não é folha, o nó folha é transformado em um nó com quatro filhos (todos folhas e representando a mesma região) e este procedimento é aplicado recursivamente para estes filhos. Quando ambos os nós são folhas eles são agrupados formando uma folha da árvore final (C). É evidente que este agrupamento inclui a junção dos dicionários das folhas.

3.4 - DESCRIÇÃO DO SISTEMA

Algumas entidades se destacam naturalmente tais como Vértices, Arestas, Q-arestas e Nós. Da mesma forma se sobressaem alguns relacionamentos entre estas entidades, pois uma aresta é definida por seus vértices extremos, cada aresta é composta por várias q-arestas, que por sua vez estão contidas em um nó.

A entidade Vértices representa o conjunto de vértices do mapa poligonal definidos por suas coordenadas X e Y. Além disso ela também contém os vértices dos centros dos quadrados da estrutura celular.

A entidade Arestas representa o conjunto de arestas que compõem o mapa poligonal, onde cada aresta é definida por seus vértices extremos.

A entidade Q-aresta representa o conjunto de q-arestas do mapa e, mais do que isso, ela permite saber quais q-arestas pertencem a um determinado nó e também de qual aresta elas são oriundas.

A entidade Nó representa o conjunto de nós resultantes da decomposição em blocos. A definição unívoca de cada nó contém informações sobre o centro do quadrado que define sua região, se ele é um nó folha ou informações sobre seus nós filhos e também sobre as q-arestas nele contidas.

A entidade Quadrado representa o conjunto de quadrados que definem fisicamente os nós, contendo informações sobre o centro e o lado destes quadrados.

A análise destes atributos conduz quase que intuitivamente, no que se refere à implementação, ao trabalho com ponteiros em arquivos de acesso direto, evitando uma redundância de dados que a

princípio pareceria existir . Para melhor visualização destes arquivos e suas relações convém observar a figura 2 onde algumas ligações entre os diversos arquivos estão representadas .

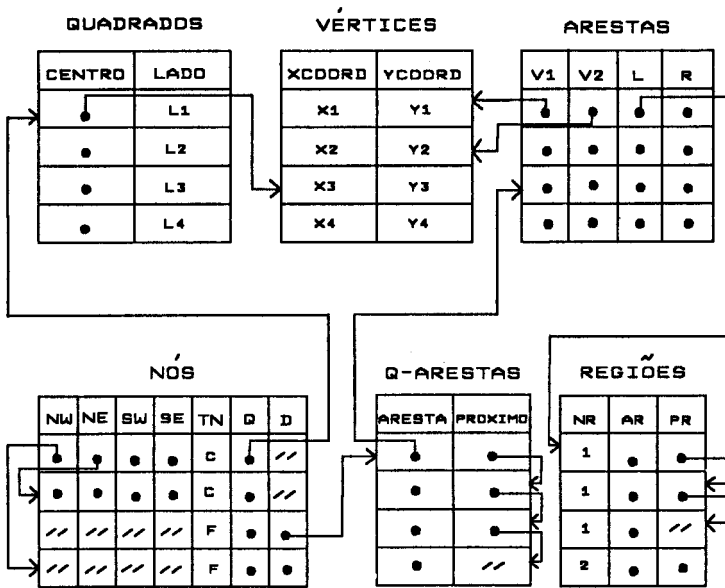


FIGURA 2 : ESQUEMA DOS ARQUIVOS

4 - IMPLEMENTAÇÃO

A implementação das rotinas de inclusão , exclusão , pesquisa e visualização de dados para uma estrutura PM3-Quadtree foi feita utilizando o compilador FORTRAN VS 2.3.0 (estruturado) , em um sistema operacional MVS/ESA 3.1.1 e um computador IBM 3090-600E/6VF . Na parte gráfica foi utilizado o editor gráfico GEOMAP , desenvolvido na PETROBRÁS , a estação gráfica IBM 5080 e a impressora LASER IBM 3800 .

Para uma maior eficiência da utilização do espaço nos arquivos foi implementado um gerenciador de espaço . O primeiro campo do primeiro registro de cada arquivo aponta para o primeiro registro livre

do arquivo , que por sua vez , através de seu primeiro campo aponta para o próximo registro livre e assim sucessivamente até o último , que aponta para o registro 0 .

Para o procedimento de inclusão de arestas foram implementadas as seguintes rotinas : i) Clipa : para verificar se a aresta está contida em determinado nó; ii) Check : para verificar se a inclusão não violará o critério de decomposição; iii) Split : para dividir um nó folha em quatro nós filhos iv) Insert : para efetivar a inclusão.

Para o procedimento de exclusão de arestas foram implementadas as seguintes rotinas : i) Clipa : para verificar se a aresta está contida em determinado nó; ii) Trymerge : para agrupar quatro nós filhos em um nó pai, caso seja possível (após a exclusão de cada aresta; iv) Delete : para efetivar a exclusão .

Para o procedimento de pesquisa por área foi implementada a rotina Pesquisa e para a visualização foi implementada a rotina Desenha .

Após a implementação do sistema protótipo foi realizado um teste com dados oriundos da digitalização do contorno do estado do Rio de Janeiro (incluindo algumas ilhas e lagoas) de uma carta na escala de 1/1000000 que faz parte da Carta Internacional ao Milionésimo . Estes dados são compostos por 421 vértices que dão origem a 423 arestas .

Neste teste foi adotado o sistema de coordenadas da projeção Policônica , com as seguintes características : i) Meridiano Central = 54 graus W ; ii) Origem das coordenadas na interseção do Meridiano Central com o Equador ; iii) Falso Norte = 1000000 m e Falso Este = 500000 m .

Para a área correspondente à região do nó raiz foi adotado um quadrado com centro no ponto de coordenada este = 6412130.0 metros e coordenada norte = 7520270.0 metros e lado = 450000.0 metros .

A inserção das 423 arestas gera um conjunto de 1417 nós , sendo 354 nós cinzas e 1063 nós folhas, dos quais 472 são nós brancos (não contém q-aresta) e 591 são nós que contém q-arestas .

O total de vértices armazenados é de 1838 , sendo 1417 referentes aos centros dos nós . As 423 arestas são compostas por 846 q-arestas da classe 1 e 246 q-arestas da classe 2 , armazenadas nos arquivos correspondentes .

A árvore final da estrutura tem profundidade 8 , o que significa que o menor quadrado tem lado igual a 1757.8125 m . O resultado final com todas as arestas e quadrados , gerado pela rotina de visualização é mostrado pela figura 3 .

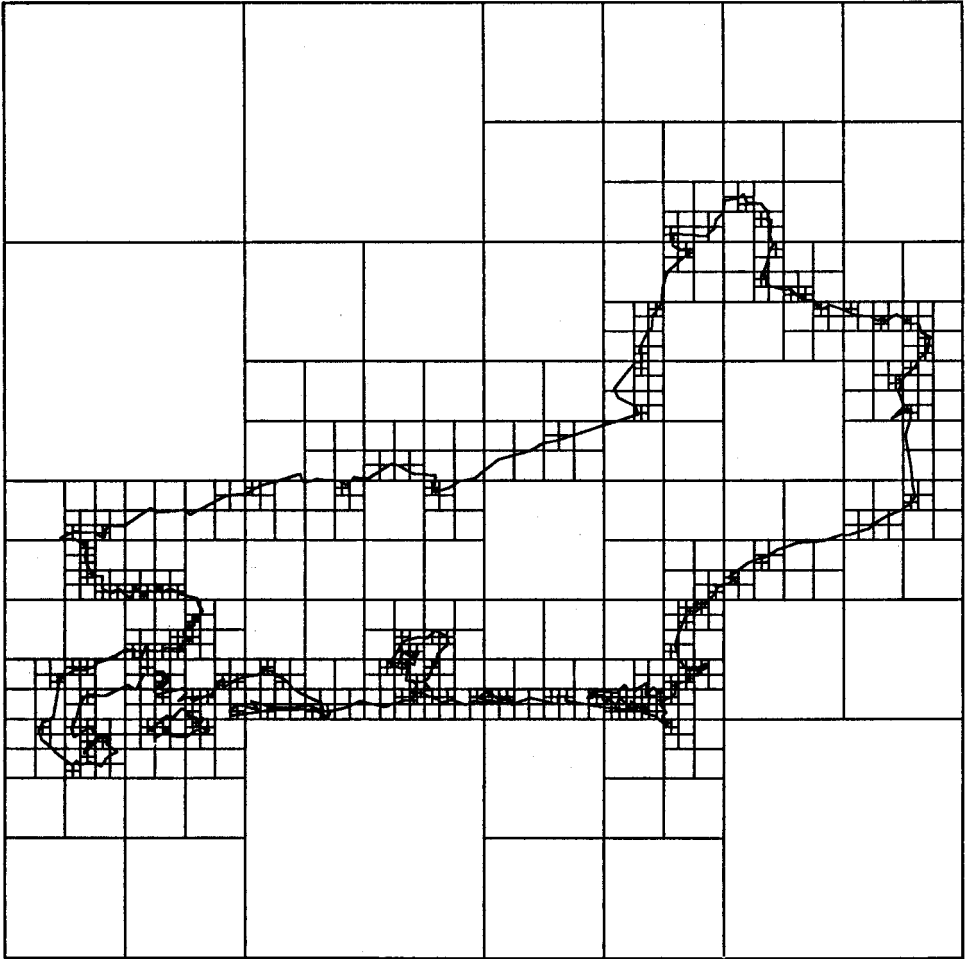


FIGURA 3 : RIO DE JANEIRO

5 - ANÁLISE DOS RESULTADOS

Quanto à quantidade de q-arestas geradas é conveniente ressaltar a quantidade média de q-arestas por nó / folha : 1.43 q-arestas tipo 1 e 0.42 q-arestas tipo 2 e também o número máximo de q-arestas por folha : 4 do tipo 1 e 3 do tipo 2 . Isto permite concluir que a utilização de uma lista encadeada para organizar estes dois arquivos é plenamente satisfatória .

Outra consideração é quanto à escolha da região do nó raiz e , conseqüentemente , quanto à sensibilidade da estrutura a operações de translação no que se refere à quantidade de memória. Para uma melhor avaliação desta influência utilizou-se a mesma massa de dados do teste e alterou-se a área do nó raiz de forma a se comparar os resultados. A diferença encontrada entre os valores extremos foi de 9% em termos utilização de memória . Estes números mostram que não ocorre um aumento excessivo da memória requerida como em outras estruturas .

6 - CONCLUSÕES

Como vértices isolados não constituem nenhum tipo de problema , a PM3 quadtree permite o armazenamento de pontos , linhas e regiões , que são os principais elementos de um mapa poligonal. Os resultados encontrados , ainda que tenham sido realizados poucos testes , permitem afirmar que a estrutura PM3-Quadtree pode ser utilizada para armazenamento de dados tipicamente cartográficos tais como contorno de estados e municípios , sistemas viários e fluviais e também para dados de mapas temáticos como delimitação de bacias e campos produtores de petróleo , reservas florestais e mapas de uso do solo em geral .

BIBLIOGRAFIA

1. G. M. Hunter e K. Steiglitz , Linear Transformation of Pictures Represented by Quadtrees ,Computer Graphics and Image Processing, **10**,289,(1979).
2. H. Samet e R. E. Webber , Using Quadtrees to Represent Polygonal Maps,Computer Vision and Pattern Recognition(Washington),IEEE,127, (1983).
3. H. Samet e R. E. Webber , Storing a Collection of Polygons Using Quadtrees,ACM Transactions on Graphics, **4** , 182,(1985).
4. H. Samet , Using Quadtrees to Represent Spatial Data,NATO ASI Series Computer Architectures for Spatially Distributed Data, **F18** ,229,(1985).