

VALIDAÇÃO DO CRITÉRIO DE CORREÇÃO DE UM SISTEMA DE QUADTREES LINEARES

VALIDATION OF THE CORRECTION CRITERIA OF A LINEAR QUADTREE SYSTEM

Emerson Raiman^{1,2}, Geison Luiz Pandiora¹, Joice Seleme Mota³, Maria Salete Marcon Gomes Vaz^{1,4}

^{1*} Autor para contato: Universidade Estadual de Ponta Grossa - UEPG, Departamento de Informática, Ponta Grossa, PR; e-mail: salete@uepg.br

² Prefeitura Municipal de Rio Negro - PR

³ UnC - Universidade do Contestado, Mafra, SC

⁴ UFPR - Universidade Federal do Paraná, Departamento de Informática, Curitiba, PR

Recebido para publicação em 15/05/2007

Aceito para publicação em 05/12/2007

RESUMO

Vários modelos conceituais de dados para SIG – Sistemas de Informação Geográfica – consideram classes e operações que permitem representar os aspectos espaciais e temporais das aplicações. Nos SIGs, esses aspectos são importantes, principalmente, para representar o histórico de dados georreferenciados. No nível do modelo interno dos sistemas atuais, as estruturas de dados armazenam e manipulam somente os aspectos espaciais dos dados geográficos, não contemplando os aspectos espaço-temporais propostos nos modelos conceituais. O objetivo deste trabalho foi comprovar o critério de correção de um Sistema de *Quadtrees* Lineares, onde foi implementada sua estrutura em linguagem visual, com base em algoritmos existentes, em um protótipo.

Palavras-chave: Informação Geográfica, *Quadtrees* Lineares

ABSTRACT

Several conceptual models of data for SIG – Geographic Information Systems – consider classes and operations that allow for representations of space and temporal aspects of the applications. In Geographic Information Systems these aspects are important mainly to represent the description of georeferentiated data. On the level of the internal models of the current systems, the data structures only store and manipulate spatial aspects of the geographic data, but do not contemplate

the space-temporal aspects proposed by the conceptual models. The aim of this work was to validate the correction criteria of a Linear Quadtree System whose structure was implemented in a visual language, on the basis of algorithms existing in a prototype.

Key words: Geographic Information, Linear *Quadtree*

1 Introdução

O termo *Quadtree* (*árvore quaternária*) é utilizado para descrever uma família de estruturas de dados, com propriedades baseadas na decomposição recursiva de formas geométricas, as quais são representadas, matricialmente, a partir de seus conteúdos (SAMET, 1989).

As estruturas *Quadtrees* decompõem dados espaciais representados em matrizes de pontos ou de células. Elas são utilizadas também como árvores de busca, possibilitando uma maior eficiência na execução de diversas operações espaciais.

A comprovação do critério de correção pode ser feita através de módulos que implementam as rotinas de armazenamento, dos aspectos espaciais de dados geográficos, mas não implementam as características temporais. Quando é necessário armazenar os aspectos temporais de dados geográficos com o histórico de imagens de uma determinada região, utilizando *quadtrees*, é armazenada a imagem completa a cada atualização.

A aplicação deve controlar a ordem das versões das imagens, já que o sistema de banco de dados ignora qualquer possível relacionamento entre duas ou mais *Quadtrees*. Tendo em vista que o espaço de memória está mais acessível, mas o espaço em disco ainda é um problema para bancos de dados geográficos, a transferência de dados entre diferentes computadores pode resultar em maior tempo de resposta.

Dentre os tipos de *Quadtrees*, existe uma classe chamada de *Quadtrees* Lineares (SAMET, 1990) e possuem vantagens para a representação e manipulação de dados espaços-temporais matriciais (CORDEIRO, 1999; MOTA, 1999), devido à facilidade de implementação, clareza conceitual, possibilidade de compactação e ao não armazenamento das células

vazias que compõem os espaços-temporais.

O objetivo deste artigo é apresentar a validação do critério de correção de um sistema de *quadtrees* lineares. Para tanto, está estruturado como segue. A Seção 2 é apresentado o embasamento teórico a respeito de gestão de dados geográficos, abordando Sistemas de Informação Geográfica e estruturas de dados para dados geográficos. Nessa seção são abordadas temporalidades de Dados e *Quadtrees*. A Seção 3 é apresentado uma análise do protótipo de validação do critério de correção de um sistema de *quadtrees* lineares. Na Seção 4, são feitas as considerações finais e sugestões de trabalhos futuros.

2 Gestão de Dados Geográficos

Nesta seção é apresentado o embasamento teórico inerente a gestão de dados geográficos, abordando Sistemas de Informação Geográfica e estruturas de dados para dados geográficos.

2.1 Sistemas de Informação Geográfica - SIGs

As aplicações geográficas armazenam e processam dados georreferenciados. Dados georreferenciados são dados referenciados em relação à superfície terrestre (LISBOA, 1997). Essas aplicações geográficas são implementadas com base em Sistemas de Informação Geográfica – SIG.

Os SIGs são sistemas automatizados, usados para armazenar, analisar, manipular e visualizar dados geográficos, ou seja, dados que representam fenômenos geográficos, cuja localização em relação à superfície terrestre é uma característica inerente à informação e indispensável para analisá-la (FARIA, 1998).

Os SIGs representam, em computador, fenômenos geográficos da natureza e aqueles criados pelo homem. Cada tipo específico de fenômeno geográfico pode ser representado de diversas formas. Na maioria dos modelos conceituais são representados como entidade ou objetos (LISBOA, 2000; FARIA, 1998; CARON, 1993; PEUQUET, 1998; BOTELHO, 1995; OLIVEIRA et al., 1997).

Um modelo conceitual permite representar como os objetos geográficos são e como se comportam perante a realidade que está sendo modelada, independentemente do sistema de gerência do banco de dados que irá armazená-los.

Os objetos geográficos apresentam características espaciais, descritivas e temporais. Estas últimas estão relacionadas ao histórico de estados assumidos pelos objetos ao longo do tempo. Alguns dos modelos conceituais propostos para representar objetos geográficos já tratam da temporalidade dos dados, ou seja, permitem que sejam modelados os aspectos temporais dos objetos. A face da terra se altera, dia a dia, pela intervenção do homem ou de fenômenos da natureza, o que aumenta a complexidade dos processos de armazenamento e manipulação de dados geográficos.

Tal complexidade deve ser tratada pelo sistema de gerência de banco de dados geográficos, que armazena os objetos geográficos para posterior manipulação pelos demais módulos do SIG. A maioria dos softwares de bancos de dados armazena somente as características espaciais e descritivas dos objetos geográficos. Novos estudos estão apresentando propostas para tratamento dos aspectos temporais de objetos geográficos (BOTELHO, 1995; FARIA, 1998).

Em SIG, o tempo é importante pela necessidade de tratamento do histórico das informações armazenadas e manipuladas. Por exemplo, em um estudo sobre a vegetação presente em uma determinada área, pode ser necessário saber se as matas e as florestas estão sendo preservadas ou desmatadas, ao longo de um período específico de tempo. Para isso seria necessário ter armazenado a sequência de estados pelos quais passou a vegetação da região ao longo do tempo. Cada um destes estados deveria, então, estar associado, no banco de dados, ao período de tempo em que foi válido.

2.2 Estruturas de dados para informações geográficas

Com base na modelagem conceitual, pode ser desenvolvido o modelo lógico que expressa os dados na visão da interface de definição do software de SIG. Neste momento, deve-se considerar se o modelo lógico implementado pelo SIG dá suporte aos aspectos temporais, o que facilita o mapeamento do modelo conceitual. Caso o modelo lógico não dê suporte a temporalidade dos dados geográficos, o tratamento dos aspectos temporais deve ficar sob a responsabilidade dos programas de aplicação.

Quando se trata do armazenamento dos dados, existem as estruturas de dados (CELES et al., 2004; SILVA, 2007) oferecidas pelo modelo físico do banco de dados implementado. No caso de dados geográficos no formato *raster* (por exemplo, imagens por satélite), existem várias estruturas de armazenamento possíveis. Dentre as mais utilizadas em SIGs, estão as *Quadtrees* (SAMET, 1989; SAMET, 1990), pela possibilidade de compactação no processo de armazenamento e manipulação de dados.

O termo *quadtree* é utilizado para descrever uma família de estruturas de dados em forma de árvores quaternárias com propriedade comum baseadas na decomposição recursiva de formas geométricas, as quais são representadas, matricialmente, a partir de seu conteúdo (SAMET, 1989; SAMET, 1990).

As estruturas *quadtrees* decompõem dados espaciais representados em matrizes de pontos ou de células. Atualmente, as *quadtrees* armazenam os aspectos espaciais de dados geográficos, não sendo capazes de armazenar suas características temporais.

Os tipos de dados a serem representados são pontos, áreas e curvas, entre outros. A decomposição é realizada com divisões quaternárias, ou ainda gerenciada pela inclusão de dados.

A classificação das estruturas *quadtrees* é apresentada pelas seguintes classes: *Quadtrees* de Pontos que servem para representar tipo de dado ponto e *Quadtrees* de Região servem para representar dados do tipo região. Cada classe de estruturas *quadtrees* possui um conjunto de operações para a sua manipulação, tais como: inserção, exclusão e busca de dados espaciais.

As principais vantagens da *quadtree* linear, em

relação às demais *quadrees*, são nos requisitos de espaço de armazenamento e tempo de execução nos processos de recuperação, que dependem unicamente da quantidade de nós; a eliminação da necessidade de armazenamento de ponteiros, de uso comum na implementação das estruturas em árvores dos demais tipos de *quadtree* e a facilidade de implementação da operação de compactação.

Outro aspecto a ser considerado é o fato das estruturas de *quadrees* lineares serem utilizadas para a economia de memória, inclusive por não utilizarem ponteiros. A não utilização de ponteiros também facilita a implementação. Por outro lado, por ser uma estrutura que não utiliza ponteiros, a sua principal desvantagem é não permitir o acesso randômico aos dados espaciais matriciais, sendo que o processo de recuperação de parte de um dado espacial matricial pode se tornar muito lento, por ser realizado sequencialmente.

3 Validação do Critério de Correção

O critério de correção refere-se ao fato da estrutura poder armazenar, manter e disponibilizar os dados corretamente. O critério de eficiência refere-se à economia de espaço de memória para a operação de armazenamento e ao tempo de resposta das operações de consulta.

Dentre os tipos de *quadtree* existentes, foi selecionada a classe *quadtree* linear para o desenvolvimento da aplicação. A seleção foi realizada de acordo com critérios de correção e eficiência. As *quadrees* lineares

possuem vantagens para a representação e manipulação de dados espaço-temporal matricial, devido à facilidade de implementação, clareza conceitual, possibilidade de compactação e ao não armazenamento das células vazias.

Essa classe manipula o histórico das imagens, armazenando a imagem original e apenas as diferenças das suas atualizações com os respectivos tempos de validade.

A *quadtree* linear utilizada é de tamanho fixo (SAMET,1990), que é a variação que trata a estrutura *quadtree* como uma coleção de nós “folhas” ou terminais. Cada nó é representado por um par de números (altura e código de localização). O código de localização é representado através de um número fixo de *bits*.

No método utilizado por Tzouramanis (TZOURAMANIS,1998), as imagens *rasters* são convertidas para representações *quadtree* linear de tamanho fixo e tem como resultado uma lista com os nós a serem armazenados.

Quando ocorrem alterações em imagens (Figura 1), são armazenadas as diferenças e quando ocorreram. Utilizando essa estrutura, com o histórico das imagens, ocorre economia de espaço de armazenamento, uma vez que apenas as diferenças estão sendo armazenadas a cada alteração.

A Figura 1 (a) corresponde a uma imagem (inicial) do ano de 1960. Já na Figura 1 (b), a mesma imagem em 1970, onde houve modificações (diferenças). Essas diferenças foram representadas e armazenadas no banco de dados, como aparece na Figura 1 (c).

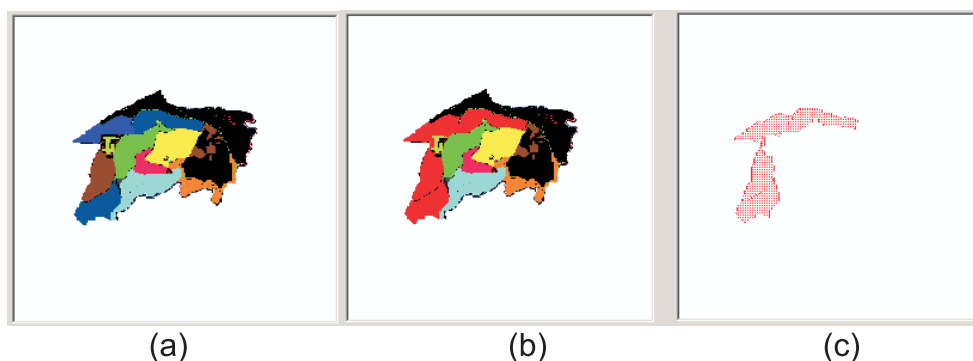


Figura 1 - Imagens Geográficas e Armazenamento em Banco de Dados

A Figura 2 (a) corresponde a uma imagem geográfica do ano de 1980, onde é representada no banco de dados as modificações ocorridas (Figura 2 (b)).

Já na Figura 2 (c) houve modificações em relação 1980, em 1990, e são representadas na Figura 2 (d).

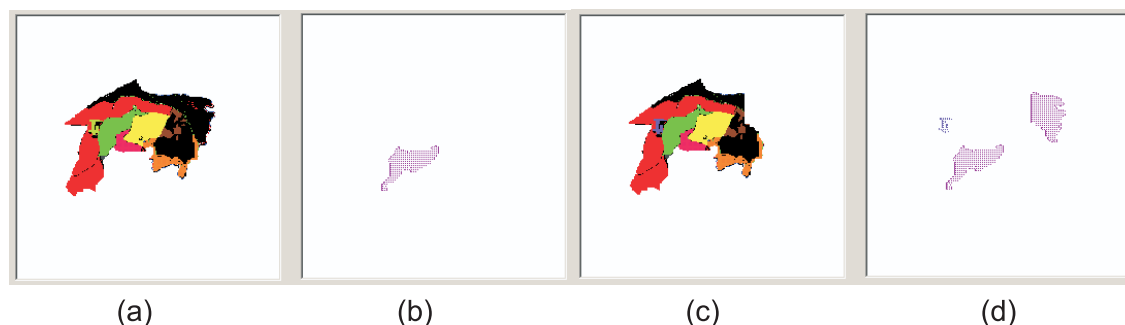


Figura 2 - Histórico de Imagens

O Protótipo do Sistema de *Quadrees* Lineares Temporais foi desenvolvido na Linguagem Pascal 7.0, e possui módulos que implementam operações de inserção, atualização e consulta. Os dados de entrada correspondem a um espaço-temporal com data e tuplas com (coordenadas, cor) das imagens. Inicialmente, é realizada a leitura desses dados e a conversão para uma *quadtree* linear.

Nas Seções 3.1 a 3.3 são descritos os processos de inserção, atualização e recuperação de informações do sistema de *quadtree* linear.

3.1 Inserção no Espaço-Temporal

As rotinas que compõem o módulo da inserção, estado inicial, no espaço-temporal são:

- *Conversão*: lê os dados de entrada e converte-os para números binários;
- *Colorir*: lê as cores das tuplas;
- *Codifica*: gera a lista linear, com o código de localização e a cor;
- *Ordena*: ordena a lista linear;
- *Compacta*: compacta a lista já ordenada; e
- *Grava_índice*: grava a data inicial da lista linear.

Para a verificação do funcionamento do protótipo foram realizadas as operações, para um conjunto de imagens representadas por um espaço-temporal. As Figuras 3 apresenta a matriz de uma imagem, bem como a lista linear correspondente gerada pelo protótipo.

	0	1	2	3	4	5	6	7
0								
1		A						
2		A	A	A				
3		A	A	A	A			
4			V	V	V	L	L	R
5			V	V	V	L	L	R
6					V	Vr	Vr	R
7					V	Vr	Vr	R

(a)

1985
 003 A 0 - 021 A 0 - 023 A 0 - 030 A 3 -
 122 A 0 - 210 V 3 - 300 V 0 - 301 L 0 -
 302 V 0 - 303 L 0 - 310 L 0 - 311 R 0 - 312
 L 0 - 313 R 0 - 320 V 0 - 321 Vr 0 - 322 V 0
 - 323 Vr 0 - 330 Vr 0 - 331 R 0 - 332 Vr 0
 333 R 0.

(b)

Figura 3 - Imagem representada por matriz – Primeira inserção

Na Figura 3 (b), as listas lineares são representadas por seu tempo de início de validade. A Figura 3 (a) é a representação de uma imagem de 1985, a qual é representada pela lista linear da Figura 3 (b), no protótipo implementado.

Os três campos que aparecem separados por traço correspondem a código de localização, cor da célula na matriz e marca de compactação, respectivamente. O código de localização é um código quaternário que corresponde às coordenadas da matriz. A marca de compactação indica se existe compactação de células.

Tabela 1 - Campos da lista linear.

Ano da Imagem	Célula	Cor	Compactação
1985	003	A	0
	230	V	4

Na Tabela 1 é mostrado um exemplo. O ano da imagem é 1985. A célula é a 003, onde a cor é azul e 0 (zero) indica que não houve compactação. Já para a célula 230, significa que existem mais células seqüências da mesma cor. A cor verde aparece nas células 230, 231, 232 e 233, visto que o campo de compactação indica 4.

3.2 Atualização dos Dados de Imagens

Na operação de atualização, os dados de en-

trada passam pelo mesmo processo da inserção, até o processo de ordenação. Após, é recuperada a lista com data inferior a lista de entrada, onde é feita a comparação entre as listas e gerada a lista com as diferenças. Esse processo é executado em cascata até atingir a lista com estado inicial e posterior armazenamento.

A estratégia utilizada no protótipo foi de armazenar cada atualização em um arquivo. Não foi prevista, na implementação do protótipo, a situação de possuir data de validade inferior ao estado final do Sistema de *Quadrtrees* Lineares Temporais. As rotinas para o módulo de atualização são as que seguem.

- *Conversão*: lê os dados de entrada e converte para números binários;
- *Colorir*: lê as cores das células;
- *Codifica*: gera a lista linear, com o código de localização e cor;
- *Ordena*: ordena a lista linear;
- *Trazer*: carrega os dados com data inferior, para serem comparados com a lista de entrada;
- *Descompacta*: descompacta os dados para compará-los com a lista de entrada;
- *Atualiza*: compara as listas, gerando uma nova lista com a diferença (esse processo é realizado em cascata até se chegar a lista com estado inicial);
- *Ordena*: ordena a lista com a diferença;
- *Compacta*: compacta a lista já ordenada; e
- *Grava_índice*: grava a data da lista.

	0	1	2	3	4	5	6	7
0								
1								
2		A	A	A				
3		A	A	A	A			
4			V	V	V	L	L	L
5			V	V	V	L	L	L
6			Am	Am	V	VR	VR	R
7			Am	Am	V	VR	VR	R

(a)

1990
003 – 230 Am 4 - 311 L 0 313 L 0

(b)

Figura 4 - Imagem representada por matriz – Primeira atualização

Na Figura 4 (a) constatou-se que aparecem pontos amarelos nas células 230 a 233, conforme re-

presentação implementada, na primeira atualização em 1990, na Figura 4 (b).

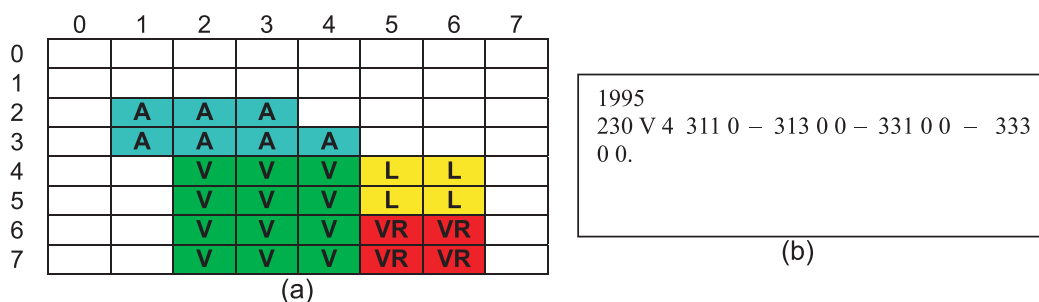


Figura 5 - Imagem representada por matriz – Segunda atualização.

Na Figura 5 (a) constatou-se que aparecem pontos amarelos nas células 230 a 233, conforme repre-

sentação implementada, na segunda atualização em 1995, na Figura 5 (b).

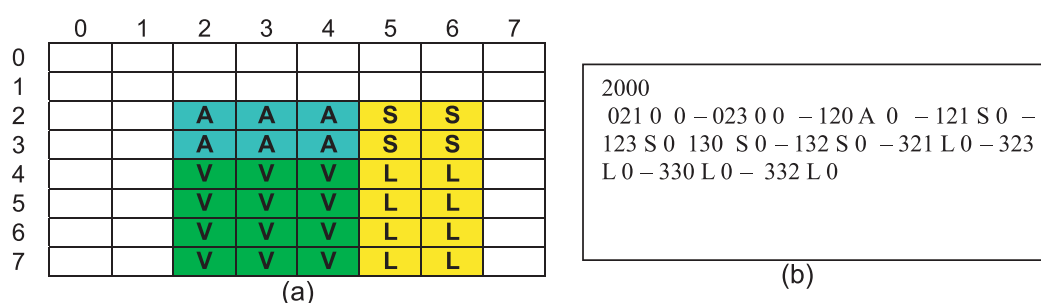


Figura 6 - Imagens representadas por matrizes – terceira atualização.

Todas as imagens que correspondem a Figura 3, que representa a primeira Inserção no ano de 1985, foram feitas mais 3 inserções (Figura 6), 1ª Atualização, em 1990, 2ª Atualização, em 1995 e 3ª Atualização, em 2000, são comparadas com suas anteriores e o resultado dessa comparação é o que fica armazenado em seu histórico.

3.3 Consulta e Recuperação da Informação do Sistema de Quadrees

A operação de consulta é realizada de acordo com a data solicitada pelo usuário. Não foi prevista, no desenvolvimento do protótipo, a situação de solici-

tação de uma consulta com data inferior ao estado inicial do Sistema de *Quadrees* Lineares Temporais. As rotinas para o módulo de consulta são as que seguem.

- *Trazer*: recupera os dados de acordo com a data solicitada e o arquivo com data anterior a essa;
- *Descompacta*: descompacta os arquivos para recuperação;
- *Consulta*: gera a lista a ser recuperada; e
- *Ordena*: ordena a lista gerada.

As operações de consulta realizadas no protótipo, dependendo da data de validade são apresentadas nas Figuras 7 a 10.

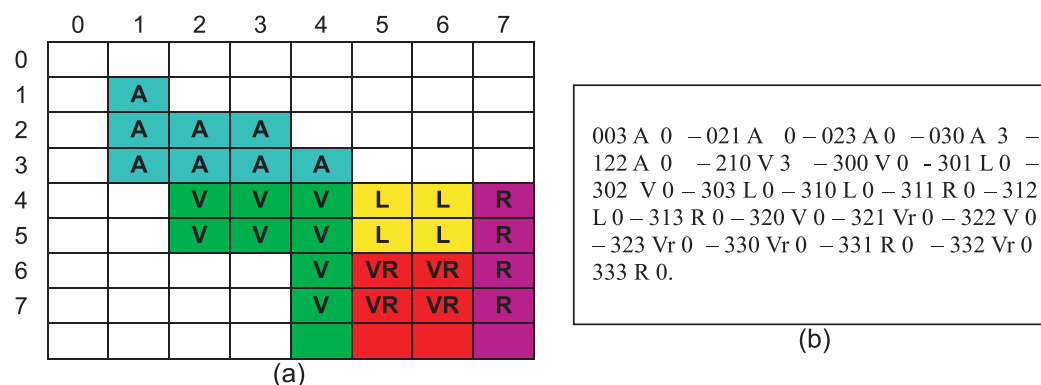


Figura 7 - Consulta e Recuperação da Imagem válida em 1987.

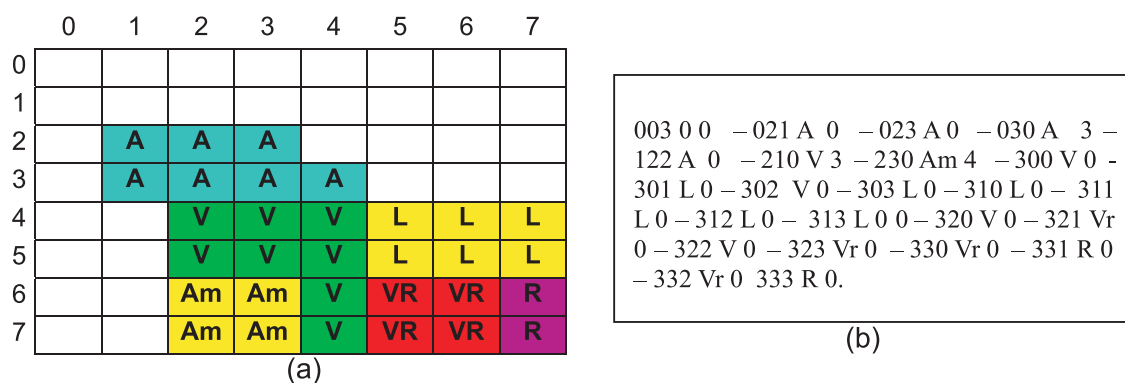


Figura 8 - Consulta e Recuperação da Imagem válida entre 1991 e 1994.

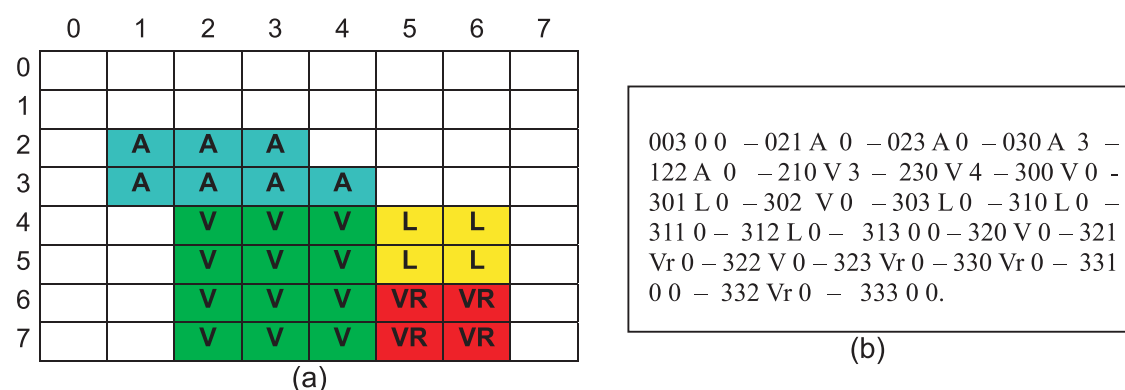


Figura 9 - Consulta e Recuperação da imagem válida em 1998.

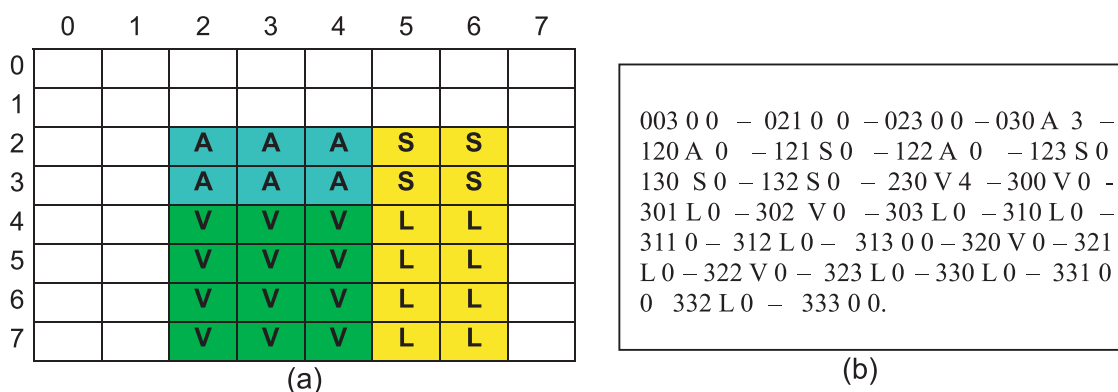


Figura 10 - Consulta e Recuperação da imagem válida em 2000.

Na operação de consulta, como ilustram as Figura 7 a 10, a imagem representada é recuperada (1987 – 1991/1994 – 1998-2000), através dos códigos representados. A construção da imagem se faz pela comparação da imagem que se deseja consultar com as

imagens com datas anteriores a sua.

Complementar a este trabalho foi desenvolvido o aperfeiçoamento de um sistema de quadrees lineares (PASDIORA, 2004; RAIMAN, 2004) que envolve o projeto e implementação do sistema.

Considerações Finais e Perspectivas de Trabalhos Futuros

Esse trabalho apresentou comprovação do critério de correção de um Sistema de *Quadrees* Lineares, através do desenvolvimento de um protótipo. A representação gráfica foi desenvolvida através de matrizes. O critério de correção se deu através da estrutura de armazenamento e disponibilização dos dados. Dentre os tipos de *quadtree* existentes, foi selecionada a classe *quadtree* linear para o desenvolvimento da aplicação.

Como perspectiva de trabalho futuro sugere-se a continuidade deste protótipo, criando funções que não estão implementadas neste protótipo, como a operação de exclusão e uma segunda maneira de inserir as imagens. Neste protótipo as imagens têm uma sequência de armazenamento, começando pela antiga, até a atual. Essa nova inserção poderia ser armazenada em qualquer momento do histórico. Por exemplo, a imagem atual é a de 2000, e apareça uma imagem de 1999.

REFERÊNCIAS

1. BOTELHO, M.A. **Incorporação de facilidades espaço-temporais em bancos de dados orientados a objetos**. São Paulo: IMECC-UNICAMP, 1995.
2. CELES, W.; CERQUEIRA, R.; RANGEL, J. L. **Introdução a Estruturas de Dados**. Rio de Janeiro: Elsevier, 2004.
3. CARON, C.; BÉDARD, Y. **Extending the individual formalism for a more complete modeling of urban spatially referenced data**. Computers, Environment and Urban Systems, [S.l.], v.17, p.337-346, 1993.
4. CORDEIRO, M. J. **Estruturas de dados aplicadas a cartografia automatizada**. Instituto Militar de Engenharia, 1999. Tese.
5. FARIA, G. **Um banco de dados espaço-temporal para desenvolvimento de aplicações em sistemas de informação geográfica**. Campinas: Instituto de Computação da UNICAMP, 1998.
6. LISBOA, J. **Modelos conceituais de dados para Sistemas de Informações Geográficas**. Porto Alegre: CPGCC da UFRGS, 1997. 119p. (EQ-12).
7. LISBOA, J. **Projeto conceitual de banco de dados geográficos através da reutilização de esquemas, utilizando padrões de análise e um framework conceitual**. Porto Alegre: PPGC da UFRGS, 2000. Tese de doutorado.
8. MOTA, J.S. **Modelos de dados temporais: um estudo comparativo**. Porto Alegre: PPGC da UFRGS, 1999. (TI-904).
9. OLIVEIRA, J.; PIRES, F.; MEDEIROS, C.B. Na Environment for integrated modelling and analysis of geographic information. **GeoInformatica**, Boston, n.1, p.29-58, 1997.
10. PASDIORA, G. L. **Armazenamento de dados georreferenciados: implementando um Sistema de Quadrees Lineares**. Monografia de Especialização em Administração de Banco de Dados. Universidade Estadual de Ponta Grossa, 2004.
11. PEUQUET, D. et al. **Delineating operations for visualization and analysis of space-time data in GIS**. Pennsylvania: Department of Geography. The Pennsylvania State University, 1998.
12. RAIMAN, E. **Armazenamento de dados Georreferenciados: aperfeiçoando um Sistema de Quadrees Lineares**. Monografia de Especialização em Administração de Banco de Dados. Universidade Estadual de Ponta Grossa, 2004.
13. SAMET, H. **The desing and analysis of spatial data structures**. Addison Wesley, 1989.
14. SAMET, H. **Applications of Spatial Data Structures**. Addison Wesley, 1990.
15. SILVA, O. Q. **Estrutura de dados e algoritmos usando C – fundamentos e aplicações**. Rio de Janeiro: Editora Ciência Moderna, 2007.
16. TZOURAMANIS, T. Overlapping linear *Quadrees*: a spatio-temporal access method. In: ACM-GIS, 1998. **Proceedings...**, Bethesda, MD, November 1998, p. 1-7.