

METODOLOGIAS DE INTEGRAÇÃO DE DADOS EM SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS

Carlos Alberto Felgueiras
Guaraci José Erthal
João Argemiro de Carvalho Paiva
Diógenes Salas Alves

Instituto de Pesquisas Espaciais-INPE
Departamento de Processamento de Imagens
Caixa Postal 515
12201 - São José dos Campos - SP

RESUMO

Este trabalho descreve e discute alguns aspectos importantes sobre o formato dos dados de entrada, os tipos de conversão entre formatos e as formas de manipulação e integração desses dados em Sistemas de Informações Geográficas. Para que o usuário possa utilizar de maneira eficiente as ferramentas fornecidas por um SIG, o trabalho apresenta, ainda, conceitos importantes sobre as metodologias de integração das informações contidas no sistema.

ABSTRACT

This work presents and analyses some important aspects related with the input format, conversion between formats, handling and integration of data stored and used in Geographical Information Systems - GIS. The data integration methodologies are also discussed in order to show to the GIS users how to work efficiently with the tools allowed by these systems.

1. INTRODUÇÃO

Para a solução de problemas complexos em áreas de aplicação de monitoramento ambiental, geologia, mapeamento, planejamento rural e urbano, e outras, muitas vezes a informações de sensoriamento remoto devem ser complementadas com outros tipos de informações, tais como, dados de mapeamento e de Modelos Numéricos de Terreno - MNT's. Os Sistemas de Informações Geográficas - SGI's - são banco de dados capazes de armazenar, recuperar e manipular informações digitais, georeferenciadas, provenientes de imagens, mapas e MNT's. Estes sistemas devem, além disso, fornecer ferramentas computacionais que permitam analisar e integrar essas informações com o objetivo de se obter soluções rápidas e precisas para problemas relacionados ao comportamento espacial dos dados contidos no sistema.

No ambiente de trabalho de um SIGA solução de um determinado problema, que é função de determinadas variáveis,

pode ser dividida segundo as seguintes tarefas:

1. Definição do problema e das variáveis que influenciam o comportamento do fenômeno que se quer estudar.

2. Definição do formato dos dados de entrada no SIG segundo critérios de precisão e de disponibilidade da informação.

3. Definição do formato de integração dos dados.

4. Definição da metodologia de integração das informações contidas na base de dados.

Este trabalho está organizado de forma a apresentar alguns conceitos sobre os formatos dos dados armazenados num SIG, os tipos de conversão possíveis entre esses dados, e as formas de integração dos mesmos, dando ênfase as metodologias de integração entre essas informações.

2. PROBLEMA GERAL ASSOCIADO A UM SIG

De uma forma simplificada, o problema a ser trabalhado dentro de um SIG está relacionado ao estudo de um fenômeno que ocorre numa região (R), definida por seus contornos em um determinado sistema de projeção. A solução do problema, ou seja a função que modela o fenômeno, depende de uma ou mais variáveis (vi's) dentro da região em questão.

Um SIG é um banco de dados, georeferenciado, que permite a entrada, o armazenamento e a manipulação e análise dos dados associados a uma região. Neste contexto a solução de um problema geral está associado, inicialmente, a entrada, no SIG, das variáveis relacionadas ao problema. Em uma segunda fase deve-se utilizar eficientemente as ferramentas de manipulação e análise, fornecidas pelo SIG, sobre as informações contidas na base de dados.

Em geral, as informações armazenadas no SIG estão distribuídas na forma de mapas, ou Planos de Informações (PI's), de acordo com o tipo de dados que elas representam ou segundo as variáveis que estão relacionadas ao problema a ser resolvido.

Algumas respostas fornecidas por um SIG dependem apenas de um tipo de informação, ou seja de um único PI. Dentro desta classe de problemas pode-se destacar os cálculos de áreas, volumes e as operações de reclassificação de um determinado PI. Problemas mais complexos exigem o uso de vários PI's e dependem fortemente da forma, ou metodologia, de integração entre as informações contidas nos PI's. O exemplo típico desse tipo de problema é a geração de mapas de aptidão que fornecem informações sobre as áreas que melhor se adaptam as restrições impostas pelas variáveis relacionadas ao problema.

3. FORMATO DOS DADOS

Os formatos dos dados dentro de um SIG estão associados as formas de representação digital, no computador, das informações contidas na determinada fonte de dado, que pode ser imagem, mapa ou outra.

Dentre os formatos mais usados podemos destacar:

- Formato vetorial: conjunto de pontos definidos no espaço 2-D (x,y) que representam entidades tais como: rios, estradas, isolinhas, divisões

entre propriedades, lineamentos, etc. A cada uma dessas entidades pode-se relacionar atributos tais como uma classe ou um valor de cota.

- Formato imagem: matriz de pontos distribuídos regularmente no espaço 2-D (x,y) que representam uma área da região terrestre. Para cada ponto da imagem, pixel, existe um valor de nível de cinza (NC) que pode representar uma refletância, um nível de declividade, um teor geoquímico, ou outra informação associada ao pixel.

- Formato amostras-3D: são pontos definidos no espaço 3-D (x,y,z) que servem como amostras do comportamento de uma variável z dentro de uma região definida no espaço (x,y). Esta representação serve principalmente para armazenar informações obtidas através de levantamentos em campo de fenômenos pontuais associados a região de interesse. Nestas classes incluem-se as amostras de teores geoquímicos e geofísicos, de toxicidade do solo, de batimetria, etc....

- Formato grade: conjunto de pontos definidos no espaço 3-D (x,y,z) distribuídos no plano (x,y) segundo uma figura geométrica poligonal, geralmente triangular ou retangular. Os valores de cota z desses pontos servem para representar os mesmos tipos de dados que as amostras-3D

4. CONVERSÃO ENTRE FORMATOS

Em geral a integração entre os diversos PIs de uma região é feita num determinado formato, ou seja, todos os PIs que serão integrados devem estar num determinado formato. Como os SIG's aceitam como entrada informações em qualquer um dos formatos descritos no item anterior, estes sistemas devem prover ferramentas de conversão entre formatos para que as informações possam ser combinadas.

Abaixo estão relacionados algumas conversões entre formatos que podem ser manipuladas dentro de um SIG.

- Vetorial para Imagem: em geral, consiste em preencher as áreas pertencentes aos polígonos definidos pelas linhas vetoriais. A cada ponto da imagem é associado o valor da classe do polígono a que ele pertence.

- Imagem para vetorial: consiste na extração dos contornos definidos pelas classes da imagem. Neste caso o sistema deve criar, ainda, as relações

topológicas entre os polígonos vetoriais extraídos.

- Grade para vetorial: consiste em extração de contornos, isolinhas, segundo a definição de cotas de corte da superfície representada pela grade.

- Amostras-3D para grade: a grade obtida através das amostras-3D pode ser obtida por um processo de interpolação das amostras (grade retangular) ou por um processo de conexão entre as mesmas para definir áreas afins segundo critérios de proximidade (grade triangular). Uma grade pode ser gerada a partir de informações no formato vetorial, isolinhas, desde que a essas linhas estejam associados valores de cota.

- Grade para imagem: consiste na quantização dos limites dos valores de cota de uma grade em níveis de cinza da imagem, geralmente 255 valores. Nesta transformação é comum obter-se uma imagem com resolução melhor do que a grade. Neste caso os valores de NC dos pontos da imagem são obtidos através de interpolação dos valores de cota da grade e posterior quantização para nível de cinza.

5. INTEGRAÇÃO ENTRE PLANOS DE INFORMAÇÃO

Considerando-se o formato de representação pode-se dividir as formas de integração de dados dentro de um SIG em 3 categorias principais: integração no formato vetorial, integração no formato grade e integração no formato imagem.

Os algoritmos computacionais de "overlay" de mapas vetoriais, geralmente operações booleanas entre classes dos PI's envolvidos na integração, são bastante complexos pois exigem basicamente 3 fases de processamento: em primeiro lugar deve-se realizar uma **busca das intersecções** dos contornos dos PI's; a fase seguinte **edição das linhas** ou seja a geração de novos nós de intersecção e a separação das linhas cortadas por esses nós. A terceira fase é a **reconstrução da nova estrutura topológica** e seus atributos segundo as operações de integração definidas pelo usuário. Nesta fase inclui-se também a extração das linhas que separam polígonos de mesma classe. Apesar de complexos, estes algoritmos fornecem resultados bastantes precisos.

No formato grade e imagem as operações de integração são realizadas entre os valores de cota z dos pontos

correspondentes, ou seja com mesma localização espacial, de cada PI. Quando essas operações são do tipo aritmética, tais como: somas, multiplicação, exponenciação, funções trigonométricas e outras, convém utilizar-se os PI's no formato grade. Como os valores de cota dos pontos das grades são valores reais os resultados dos cálculos aritméticos não acumulam erros de quantização que ocorrem quando se usam imagens para essas operações. Quando se quer um resultado no formato imagem segundo classes pode-se transformar a grade resultado em imagem e então classificá-la segundo algum critério de aglutinação dos valores dos pixels. Quando a função de integração está definida para as classes de cada PI a combinação deve ser feita no formato imagem. Neste caso obtém-se um PI resultado classificado segundo as restrições impostas nas regras de integração entre as classes dos PI's operandos.

Os algoritmos computacionais de integração entre PI's nos formatos grades e imagens são de implementação simples e muito rápidos em relação ao tempo de processamento. Porém o usuário deve se preocupar com a resolução dos dados de entrada para obter resultados satisfatórios.

6. METODOLOGIAS DE INTEGRAÇÃO

A definição de uma metodologias de integração de dados num SIG envolve o conhecimento das relações entre as variáveis importantes na solução de um determinado problema, ou seja, no estudo do comportamento de um determinado fenômeno que ocorre numa região. O usuário, além de conhecer as variáveis que afetam um determinado fenômeno em estudo, deve saber qual a metodologia apropriada para o estudo do problema em questão. Isto significa que o usuário deve ter a sensibilidade para saber como se relacionam as variáveis que ele elegeu como importantes para a solução do problema. Dentro deste contexto pode-se distinguir dois tipos de metodologias a serem utilizadas num SIG: as metodologias já equacionadas e as metodologias a serem equacionadas. O equacionamento de uma metodologia compreende a definição de uma função de relacionamento entre as variáveis envolvidas no fenômeno e os testes necessários para se garantir o funcionamento satisfatório da metodologia.

As metodologias já equacionadas, e portanto testadas, podem ser repetidas para regiões com características semelhantes à região utilizada para a

definição e teste do equacionamento. Em sistemas que utilizam conceitos de inteligência artificial, essas metodologias podem ser armazenadas em bases de conhecimento. Neste caso essas informações podem então ser recuperadas e a metodologia pode ser repetida automaticamente pelo sistema sempre que se quiser utilizá-la para uma nova região.

As metodologias não equacionadas devem ser definidas pelo usuário do sistema. Neste caso o usuário utiliza uma série de informações predefinidas e uma função de relacionamento entre essas informações que ele acha que pode vir a representar o comportamento do fenômeno em estudo para a região. Assim várias metodologias podem vir a ser testadas, variando-se as informações ou a função de relacionamento, até que se encontre, **empiricamente**, um resultado que condiz com a realidade do comportamento do fenômeno na região. Este tipo de procedimento pode ser bastante explorado principalmente quando se possui ferramentas computacionais que permitem a repetição desses processos rapidamente.

7. CONCLUSÕES

Neste trabalho procurou-se discutir alguns aspectos importantes relacionados com a integração de informações utilizado-se ferramentas computacionais disponíveis num SIG. A maioria dos tópicos aqui descritos não se apresentam como novidade às pessoas que já trabalham ou tem um certo conhecimento na área. Porém o que se observa atualmente é uma falta de conhecimento relativamente grande desse tipo de tecnologia no país e é exatamente neste contexto que se encaixa os propósitos deste trabalho. É importante salientar que numa época dominada pela informática, as informações tem um valor cada vez maior principalmente quando atualizadas e confiáveis. O advento dos Sistemas de Informações baseados em computadores, incluindo aí as informações manipuladas pelos SIG's, são os mecanismos que agilizam e melhoram a confiabilidade das informações utilizadas no mundo atual e sobre as quais se baseiam as decisões a serem tomadas em diversas áreas de atuação.

8 - REFERÊNCIAS BIBLIOGRÁFICAS

Burrough, P. A. Principles of Geographical Information Systems for Land Resource Assessment. Clarendon Press - Oxford, 1986.

Sistema de Informações Geográficas - SGI V2.0. Manual do usuário. Engespaço Indústria e Comércio Ltda. São José dos Campos SP. 1989.