

Mineração Visual de Multiatributos Sísmicos para Classificação de Multifácies

Hélio Lopes, Alex Bordignon, Rener de Castro, Thomas Lewiner, Geovan Tavares - PUC-Rio. Rogério Santos e Amin Murad - Petrobras

Copyright 2007, SBGF - Sociedade Brasileira de Geofísica

This paper was prepared for presentation at the 10th International Congress of The Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 19-22 November 2007.

Contents of this paper were reviewed by the Technical Committee of the 10th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGF, its officers or members. Electronic reproduction, or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Abstract

The understanding of large data clustering is a relevant problem in several areas of science and engineering. Seismic signal are in general represented by large data sets. For petroleum geophysicists the choice of areas of interest is a demanding task. In this paper we introduce data mining visualization techniques, stellar coordinates and transfer functions, which ease the task of dealing visually with data clustering in an interactive and intuitive way allowing the user to make meaningful choices of 3D regions of interest to work with.

Introdução

A exploração e análise de dados multiatributos é um importante assunto em diferentes áreas como simulação de fluidos, imagens médicas e geofísicas. Pesquisadores da área de visualização da informação desenvolveram inúmeras técnicas para explorar e analisar interativamente dados multivariados. Milhões de amostras conduzem a uma dificuldade de seleção de poucas amostras que conduza a um interesse econômico ou geocientífico, num determinado bancos de dados, que geralmente são muito grandes e complexos. Uma solução efetiva é alcançada ao se colocar o conhecimento humano e a interação nesse processo de exploração.

Objetivo

Multiatributos de um determinado objeto, ou imagem, em análise é um tema de crescente importância em diversas áreas da geofísica de petróleo, medicina e engenharia. Descrever e selecionar um conjunto de atributos focalizando um determinado interesse de análise, por vezes representadas por anomalia de dados, são tarefas que podem ser entregues a processos computacionais de inteligência artificial. Entretanto, dependendo das características dessas anomalias, mesmo os mais sofisticados algoritmos matemáticos não conseguem ter sucesso. Não conseguem superar, nem mesmo substituir, a experiência, a percepção e a inteligência humanas. Para esses casos torna-se crítico a identificação, ou discriminação, por percepção visual, de aglomerações, ou conjunção, de dados que se arranjam em estruturas que sigam alguma lei de formação espacial, e que possam ser marcadas como possíveis candidatas a conjuntos anômalos, destacadas no interior de uma grande massa de dados.

Esse trabalho trata de análises de bancos de dados multivariados estruturados geometrica e espacialmente. Inicialmente, supõe-se que cada elemento, ou um arranjo, desse dado (pixel ou voxel) possui vários atributos, considerando as estruturas formadas como objetos geométricos multivariados. Para esses objetos, considera-se as informações espaciais 2D e 3D como sendo também atributos. O principal objetivo é estabelecer uma ferramenta simples e interativa para facilitar a tarefa de exploração e análise desses objetos.

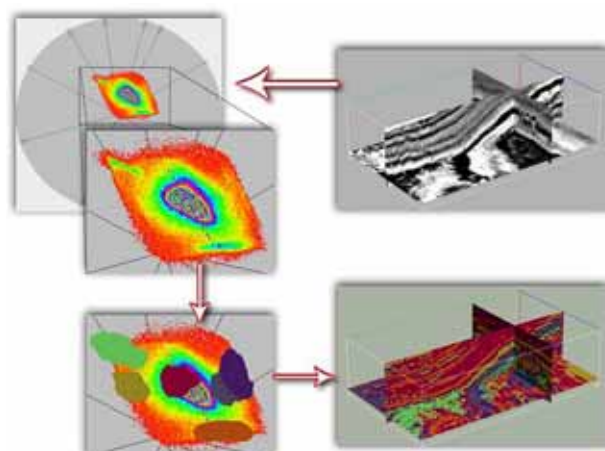


Figura 1 - Segmentando um objeto em dados sísmicos, usando 11 atributos, primeiro o dado é projetado para o plano usando coordenadas escolhidas pelo usuário, segundo marca-se interativamente os aglomerados no mapa de densidade e recebe-se o retorno visual da seleção.

Contribuições resumidas

Desenvolveu-se uma nova técnica que permite explorar visualmente e iterativamente os objetos geométricos multivariados. Uma interface gráfica simples e intuitiva é prototipada. Nessa interface, o usuário controla um eixo de visualização, que permite detectar relações e estruturas dentro do objeto. Assim, é possível se identificar uma característica útil em grande massa de dados, assinalar uma região em questão para uma função de transferência, que irá mostrar tal seleção usando cores e opacidade. No fim do processo, a região é renderizada no domínio do objeto, com o auxílio da função de transferência.

Coordenadas estelares

Coordenadas estelares generalizam o conceito de plote de espalhamento, ou nuvem de pontos, para dimensões mais altas do que 3. Apresenta-se aqui uma descrição formal para tal sistema de coordenadas. A idéia principal é colocar n eixos dentro de um círculo no plano.

Considere uma família de funções $F\{A,S\}$ que mapeia a caixa unitária de n dimensões definida por $B=[\min_1, \max_1] \times \dots \times [\min_n, \max_n]$ em R^n no plano R^2 Essa família de funções é parametrizada pelo conjunto

$A=\{a_1, \dots, a_n\}$ de n vetores unitários em R^2 e por um conjunto $S=\{s_1, \dots, s_n\}$ de números reais não negativos. O vetor a_i em A representa o i -ésimo eixo no círculo e o escalar s_i de S corresponde a um fator de escala que multiplica a_i no sistema de coordenadas estelares. Dado um ponto $p=(x_1, \dots, x_n)$ em B , o valor $F_{\{A,S\}}(p)$, chamado de coordenada estelar de p , é calculado da seguinte maneira:

$$F_{\{A,S\}}(p) = \frac{1}{n} \sum_{i=1}^n \frac{x_i - \min_i}{\max_i - \min_i} s_i a_i$$

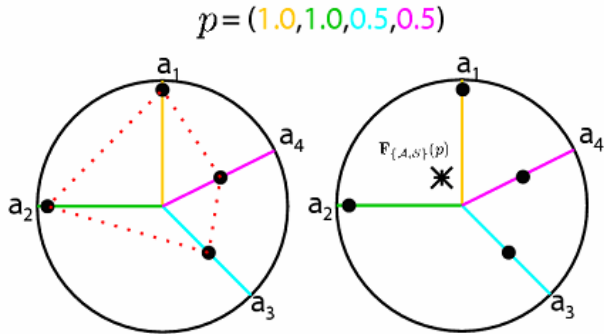


Figura 2 - Valor da coordenada estelar para um ponto $p=(1,1,0.5,0.5)$ para uma determinada escolha de A e S .

Função de transferência n-dimensional

O objetivo de uma função de transferência n-dimensional, em uma determinada imagem, é mapear o pixel ou voxel com n atributos para uma cor e uma opacidade. Formalmente, uma função de transferência n-dimensional é uma função $T:B \rightarrow C$, onde C é o espaço de todas as possíveis propriedades representáveis de cor e opacidade. Geralmente, um ponto no espaço C possui quatro coordenadas: (R,G,B,A) , onde R,G e B correspondem as componentes vermelho, verde e azul da cor e A representa a opacidade.

Explorando com coordenadas estelares

Nesse trabalho propomos uma interface simples e interativa para construir uma função de transferência n-dimensional usando as coordenadas estelares. A idéia é quebrar o desenvolvimento da função em dois estágios: no primeiro estágio o intérprete seleciona interativamente um controle visual para escolher os parâmetros para conjuntos A e S através de operações de rotação e escala:

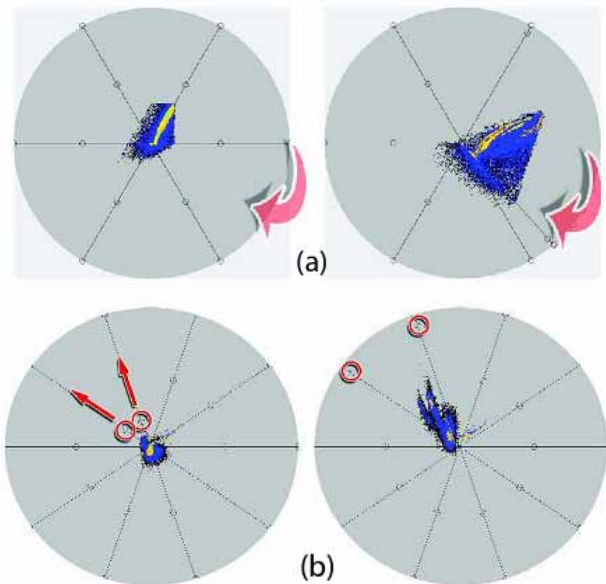


Figura 3 - Construindo um sistema de coordenadas estelares usando (a) rotações dos eixos e (b) escalando os eixos.

No segundo estágio, ele define uma função de transferência $G: R^2 \rightarrow C$ entre o plano e o espaço C das propriedades representáveis. Formalmente, substitui-se a tarefa de se calcular uma função de transferência n-dimensional $T:B \rightarrow C$ pela função composição $GoF_{\{A,S\}}:B \rightarrow C$. Esse procedimento mostra-se muito eficaz pelo fato de simplificar e acelerar a construção de uma função de transferência para uso nos mais diversos cenários. Para se construir a função G é usada uma interface onde o intérprete pode marcar as regiões do plano que ele prefira, como mostra a figura 4.

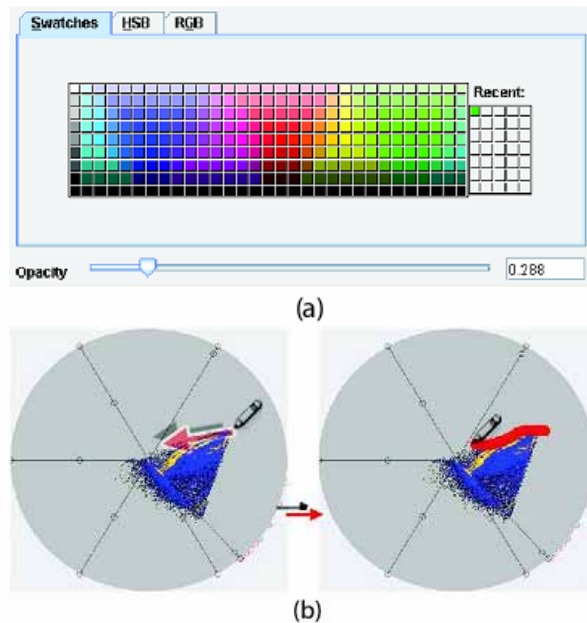


Figura 4: (a) Selecionando o elemento do espaço C . (b) marcando partes do plano que são levados para o elemento selecionado.

Para facilitar uma exploração visual dos dados introduz-se o uso de cores. A idéia é a de subdividir o espaço do sistema coordenado em pequenas regiões e atribuir uma cor a cada um, de acordo com a densidade dos pontos, que sejam projetados pelas coordenadas estelares. Para subdividir o domínio, propõe-se duas estratégias: uma subdivisão quadrada regular ou uma subdivisão radial. Em ambos os casos, atribui-se uma cor a cada região de acordo com o número dos pontos projetados na região. Supondo que são dados um vetor de q padrão $t=(t_1..t_q)$ e um vetor de $q + 1$ cores $c=(c_1..c_{q+1})$. A cor c_i é atribuída a uma região que tenha menos pontos de t_{i+1} e mais pontos que t_i . A figura 5 ilustra tal idéia.

As figuras 3 e 4, anteriores, foram geradas dessa forma: as regiões amarelas representam regiões densas na projeção.

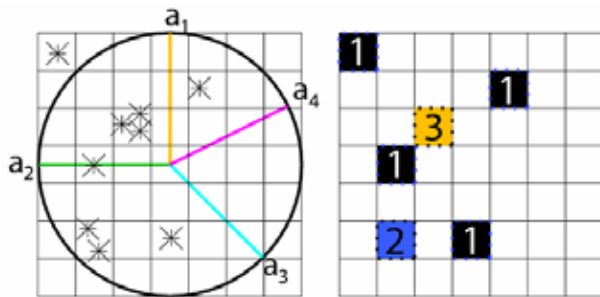


Figura 5: Colorindo segundo a freqüência de projeções, regiões do plano são coloridas segundo a densidade de pontos projetados.

Dessa maneira o intérprete possui um ótimo guia para determinar a função G como mostra a figura 4.

Resultados

Aplicou-se com sucesso a técnica aqui apresentada em dados sísmicos usando multiatributos. As figuras seguintes ilustram alguns dos resultados obtidos, tanto para visualização quanto para segmentação dos dados.

Conclusões

Coordenadas estelares associadas a funções de transferências são de ferramentas poderosas que permitem ao geofísico de petróleo estudar aglomerações de dados, escolhe-los apropriadamente de acordo com o problema sendo tratado e visualizá-las de modo interativo. Assim é possível partindo de dados sísmicos escolher regiões de interesse, com fácies por exemplo, reconhecer a sua importância de modo a auxiliá-lo na caracterização de reservatório.

Agradecimentos

Agradecemos à Petrobras pelo contínuo suporte ao Projeto Sismolitofacies, através de convênio tecnológico com a PUC-Rio, dentro do qual são desenvolvidas as teorias aqui ilustradas.

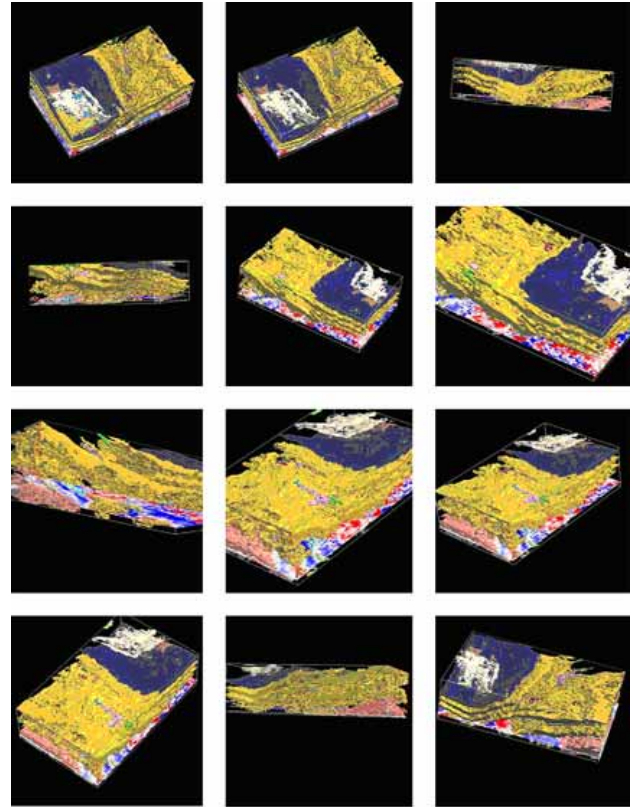


Figura 6 – Usando a função de transferência gerada a partir de 8 atributos é possível gerar a superfície envolvente com resolução subvoxel para visualização da segmentação efetuada.

Referências

A. O. Artero and M. C. F. de Oliveira. Viz3d: effective exploratory visualization of large multidimensional data sets. In Proceedings of the 17th Brazilian Symposium on Computer Graphics and Image Processing, 2004, pp. 340–347.

A. Bordignon, R. de Castro, H. Lopes, T. Lewiner, G. Tavares. Exploratory Visualization Based on Multidimensional Transfer Functions and Star Coordinates. In Proceedings of the 19th Brazilian Symposium on Computer Graphics and Image Processing, 2006, pp. 273-280.

E. Kandogan. Visualizing multi-dimensional clusters, trends, and outliers using star coordinates. In Proceedings of the 7th ACM SIGKDD, 2001, pp. 107–116.

J. Kniss, S. Premoze, M. Ikits, A. Lefohn, C. Hansen, and E. Praun. Gaussian transfer functions for multi-field volume visualization. In IEEE Visualization, volume 11, 2003, pp. 497 – 504.

M. C. F. de Oliveira and L. Levkowitz. From visualization to visual data mining: A survey. IEEE Transactions on Visualization and Computer Graphics, 9, 2003, pp.378–394.