Flavia Medeiros dos Anjos

Reorganização e Compressão de Dados Sísmicos

DISSERTAÇÃO DE MESTRADO

DEPARTAMENTO DE INFORMÁTICA

Programa de Pós-Graduação em Informática

Rio de Janeiro, agosto de 2007

Pontifícia Universidade Católica
do Rio de Janeiro

Flavia Medeiros dos Anjos

Reorganização e Compressão de Dados Sísmicos

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio.

Orientadores: Eduardo Sany Laber
Pedro Mário Cruz e Silva

Rio de Janeiro

Agosto de 2007

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Flavia Medeiros dos Anjos

Reorganização e Compressão de Dados Sísmicos

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Informática da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Eduardo Sany Laber
Orientador
Departamento de Informática - PUC-Rio

Pedro Mário Cruz e Silva
Co-orientador
TeCGraf - PUC-Rio

Marcelo Gattass
Departamento de Informática - PUC-Rio

Carlos Alves da Cunha Filho
Petrobras

Prof. José Eugenio Leal
Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 9 de agosto de 2007

Flavia Medeiros dos Anjos

Graduou-se em Engenharia de Computação pela Pontifícia Universidade Católica do Rio de Janeiro, onde continuou seus estudos no programa de Mestrado em Informática. Sempre com interesse nas áreas de algoritmos e computação gráfica participou de projetos no laboratório LEARN e desde 2003 atua com desenvolvimento e pesquisa no laboratório TeCGraf.

Ficha Catalográfica

Anjos, Flavia Medeiros dos

Reorganização e compressão de dados sísmicos / Flavia Medeiros dos Anjos ; orientadores: Eduardo Sany Laber, Pedro Mário Cruz e Silva. – 2007.

78 f. : il. (col.) ; 30 cm

Dissertação (Mestrado em Informática)–Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007.

Inclui bibliografia

1. Informática – Teses. 2. Dados sísmicos 3. Sub-volumes. 4. Transferência disco-memória. 5. Agrupamento. 6. Compressão. 7. Reorganização. 8. K-medianas. I. Laber, Eduardo Sany. II. Silva, Pedro Mário Cruz e III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

às minhas avós Helena e Guilhermina e ao meu avô Hermano, que tenha encontrado sua paz.

Agradecimentos

À minha família que me apoiou desde o início.

Ao meu orientador Eduardo Laber que acreditou e depositou sua confiança em mim.

Aos colegas e amigos do TeCGraf e da Petrobras, em especial à equipe do v3o2, que me ajudaram e forneceram recursos para o desenvolvimento do trabalho.

À CNPq pelo apoio financeiro.

Aos meus amigos que sempre torceram pelo meu sucesso.

A todos aqueles que me acompanharam ao longo desta trajetória.

Resumo

dos Anjos, Flavia Medeiros; Laber, Eduardo Sany. **Reorganização e Compressão de Dados Sísmicos.** Rio de Janeiro, 2007. 78p. Dissertação de Mestrado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Dados sísmicos, utilizados principalmente na indústria de petróleo, costumam apresentar dimensões de dezenas de gigabytes e em alguns casos, centenas. Este trabalho apresenta propostas de manipulação destes dados que ajudem a contornar problemas enfrentados por aplicativos de processamento e interpretação sísmica ao trabalhar com arquivos deste porte. As propostas se baseiam em reorganização e compressão. O conhecimento do formato de utilização dos dados permite reestruturar seu armazenamento diminuindo o tempo gasto com a transferência entre o disco e a memória em até 90%. A compressão é utilizada para diminuir o espaço necessário para armazenamento. Para dados desta natureza os melhores resultados, em taxa de redução, são das técnicas de compressão com perda, entre elas as compressões por agrupamento. Neste trabalho apresentamos um algoritmo que minimiza o erro médio do agrupamento uma vez que o número de grupos tenha sido determinado. Em qualquer método desta categoria o grau de erro e a taxa de compressão obtidos dependem do número de grupos. Os dados sísmicos possuem uma coerência espacial que pode ser aproveitada para melhorar a compressão dos mesmos. Combinando-se agrupamento e o aproveitamento da coerência espacial conseguimos comprimir os dados com taxas variando de 7% a 25% dependendo do erro associado. Um novo formato é proposto utilizando a reorganização e a compressão em conjunto.

Palavras-chave

dados sísmicos; sub-volumes; transferência disco-memória; agrupamento; compressão; reorganização; k-medianas

Abstract

dos Anjos, Flavia Medeiros; Laber, Eduardo Sany. **Reorganization e Compression of Seismic Data.** Rio de Janeiro, 2007. 78p. MSc Dissertation - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Seismic data, used mainly in the petroleum industry, commonly present sizes of tens of gigabyte, and, in some cases, hundreds. This work presents propositions for manipulating these data in order to help overcoming the problems that application for seismic processing and interpretation face while dealing with file of such magnitude. The propositions are based on re-organization and compression. The knowledge of the format in which the data will be used allows us to restructure storage reducing disc-memory transference time up to 90%. Compression is used to save storage space. For data of such nature, best results in terms of compression rates come from techniques associated to information loss, being clustering one of them. In this work we present an algorithm for minimizing the cost of clustering a set of data for a pre-determined number of clusters. Seismic data have space coherence that can be used to improve their compression. Combining clustering with the use of space coherence we were able to compress sets of data with rates from 7% to 25% depending on the error associated. A new file format is proposed using re-organization and compression together.

Keywords

seismic data, sub-volume, disc-memory transference, clustering, compression, reorganization, k-medians

Sumário

Lista de figuras

Lista de tabelas