



Aurélio Moraes Figueiredo

**Mapeamento de Eventos Sísmicos baseado em
Algoritmos de Agrupamento de Dados**

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Informática
do Departamento de Informática da PUC-Rio como requisito
parcial para obtenção do título de Doutor em Informática

Orientador: Prof. Marcelo Gattass

Rio de Janeiro
Agosto de 2015



Aurélio Moraes Figueiredo

Mapeamento de Eventos Sísmicos baseado em Algoritmos de Agrupamento de Dados

Tese apresentada ao Programa de Pós-graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC-Rio como requisito parcial para obtenção do título de Doutor em Informática. Aprovada pela comissão examinadora abaixo assinada.

Prof. Marcelo Gattass

Orientador
Departamento de Informática
PUC-Rio

Prof. Cristina Nader Vasconcelos

UFF

Prof. Ruy Luiz Milidui

Departamento de Informática - PUC-Rio

Prof. Helio Côrtes Vieira Lopes

Departamento de Informática - PUC-Rio

Prof. Pedro Mário Cruz e Silva

PUC-Rio

Prof. Fernando Barbosa da Silva

Petrobras

Prof. Prof. José Eugenio Leal

Coordenador do Centro Técnico Científico
PUC-Rio

Rio de Janeiro, 20 de Agosto de 2015

Todos os direitos reservados. Proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Aurélio Moraes Figueiredo

Se formou como Técnico em Metalurgia pela Escola Técnica Pandiá Calógeras, em 1997. É Engenheiro de Computação graduado pela Pontifícia Universidade Católica do Rio de Janeiro, em dezembro de 2004. Em março de 2006 entrou para o Programa de Pós-graduação da PUC-Rio, onde concluiu Mestrado em Informática, em julho de 2007, com ênfase em algoritmos de Aprendizado de Máquina.

Ficha Catalográfica

Figueiredo, Aurélio Moraes

Mapeamento de Eventos Sísmicos baseado em Algoritmos de Agrupamento de Dados / Aurélio Moraes Figueiredo; orientador: Marcelo Gattass. — Rio de Janeiro : PUC–Rio, Departamento de Informática, 2015.

v., 69 f: il. ; 29,7 cm

1. Tese (Doutorado em Informática) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Tese. 2. Horizontes Sísmicos. 3. Falhas Sísmicas. 4. Dados Sísmicos. 5. Análise de Componentes Principais. 6. Análise de Sismofácies. I. Gattass, Marcelo. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 510

Agradecimentos

Este trabalho não é resultado apenas de um esforço individual. Ele nasce de significativas contribuições que recolhi durante minha trajetória profissional, acadêmica e como cidadão, ao lidar com pessoas e instituições que foram fundamentais a essa construção. Por isso aqui vão meus mais sinceros agradecimentos.

À minha família, pelo apoio e pela paciência nos finais de semana perdidos enquanto este trabalho era construído.

Aos amigos e colegas do V3O2, sem os quais essa tese não seria possível.

À PUC-Rio e aos seus professores que ensinam com o mesmo empenho nas matérias mais simples e nos assuntos mais complexos.

Ao meu orientador Marcelo Gattass por todas as idéias e conselhos que foram indispensáveis na elaboração dessa tese.

Ao Tecgraf, por ser uma ótima fonte de problemas difíceis e interessantes.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos de Doutorado.

Resumo

Figueiredo, Aurélio Moraes; Gattass, Marcelo. **Mapeamento de Eventos Sísmicos baseado em Algoritmos de Agrupamento de Dados**. Rio de Janeiro, 2015. 69p. Tese de Doutorado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Neste trabalho apresentamos metodologias baseadas em algoritmos de agrupamento de dados utilizadas para processamento de dados sísmicos 3D. Nesse processamento, os voxels de entrada do volume são substituídos por vetores de características que representam a vizinhança local do voxel dentro do seu traço sísmico. Esses vetores são processados por algoritmos de agrupamento de dados. O conjunto de grupos resultantes é então utilizado para gerar uma nova representação do volume sísmico de entrada. Essa estratégia permite modelar a estrutura global do sinal sísmico ao longo de sua vizinhança lateral, reduzindo significativamente o impacto de ruído e demais anomalias presentes no dado original. Os dados pós-processados são então utilizados com duas finalidades principais: o mapeamento automático de horizontes ao longo do volume, e a produção de volumes de visualização destinados a enfatizar possíveis descontinuidades presentes no dado sísmico de entrada, particularmente falhas geológicas. Com relação ao mapeamento de horizontes, o fato de as amostras de entrada dos processos de agrupamento não conterem informação de sua localização 3D no volume permite uma classificação não enviesada dos voxels nos grupos. Consequentemente a metodologia apresenta desempenho robusto mesmo em casos complicados, e o método se mostrou capaz de mapear grande parte das interfaces presentes nos dados testados. Já os atributos de visualização são construídos através de uma função auto-adaptável que usa a informação da vizinhança dos grupos sendo capaz de enfatizar as regiões do dado de entrada onde existam falhas ou outras descontinuidades. Nós aplicamos essas metodologias a dados reais. Os resultados obtidos evidenciam a capacidade dos métodos de mapear mesmo interfaces severamente interrompidas por falhas sísmicas, domos de sal e outras descontinuidades, além de produzirmos atributos de visualização que se mostraram bastante úteis no processo de identificação de descontinuidades presentes nos dados.

Palavras-chave

Horizontes Sísmicos; Falhas Sísmicas; Dados Sísmicos; Análise de Componentes Principais; Análise de Sismofácies.

Abstract

Figueiredo, Aurélio Moraes; Gattass, Marcelo (advisor). **Mapping Seismic Events using Clustering-based Methodologies**. Rio de Janeiro, 2015. 69p. Doctoral Thesis — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

We present clustering-based methodologies used to process 3D seismic data. It firstly replaces the volume voxels by corresponding feature samples representing the local behavior in the seismic trace. After this step samples are used as entries to clustering procedures, and the resulting cluster maps are used to create a new representation of the original volume data. This strategy finds the global structure of the seismic signal. It strongly reduces the impact of noise and small disagreements found in the voxels of the entry volume. These clustered versions of the input seismic data can then be used in two different applications: to map 3D horizons automatically and to produce visual attribute volumes where seismic faults and any discontinuities present in the data are highlighted. Concerning the horizon mapping, as the method does not use any lateral similarity measure to organize horizon voxels into clusters, the methodology is very robust when mapping difficult cases. It is capable of mapping a great portion of the seismic interfaces present in the data. In the case of the visualization attribute, it is constructed by applying an auto-adaptable function that uses the voxel neighboring information through a specific measurement that globally highlights the fault regions and other discontinuities present in the original volume. We apply the methodologies to real seismic data, mapping even seismic horizons severely interrupted by various discontinuities and presenting visualization attributes where discontinuities are adequately highlighted.

Keywords

Seismic Horizons; Seismic Faults; Seismic Data; Principal Component Analysis; Seismic Facies Analysis.

Sumário

1	Introdução	11
1.1	Objetivo	13
1.2	Motivação	14
1.3	Contribuições	15
1.4	Organização deste Documento	16
2	Trabalhos Relacionados	17
2.1	Métodos baseados em mergulho	18
2.2	Métodos baseados em junção de sub-horizontes	21
2.3	Métodos baseados em processos de otimização global	22
2.4	Limitações gerais	23
3	Métodos Desenvolvidos	25
4	Conclusões	30
A	Anexo A	32
B	Anexo B	38
C	Anexo C	44
D	Anexo D	50
	Referências Bibliográficas	67

Lista de figuras

1.1	Traço sísmico (esquerda), linha sísmica (centro) e volume sísmico (direita).	12
1.2	(a) um horizonte sísmico mapeado a partir da sísmica 3D. (b) equivalente ao horizonte num ambiente real.	13
1.3	(a) falha sísmica identificada na seção sísmica 3D. (b) um modelo ilustrativo do que seria o equivalente na natureza.	13
2.1	Resultado reportado em Blinov e Petrou (4).	18
2.2	Resultado reportado em Lomask e Guitton (18).	19
2.3	Resultado reportado em De Groot et al.	20
2.4	Seções sísmicas mostrando linhas extraídas usando um ponto de controle (curva azul) e 19 pontos de controle (curva verde). Resultado reportado em Wu e Hale (27).	20
2.5	Uma vista 3D das superfícies extraídas usando um ponto de controle (esquerda) e 19 pontos de controle (direita), respectivamente. Ambas as superfícies são coloridas por amplitude. Resultado reportado em Wu e Hale (27).	21
2.6	Processo de mapeamento de horizonte descrito em Borgos et. al. (5). A partir do conjunto de fragmentos de horizontes descobertos pela metodologia, o usuário seleciona um a um aqueles que pertencem ao horizonte a ser mapeado. O método é capaz de agrupar esses fragmentos e remover possíveis discordâncias, gerando o horizonte final.	22
2.7	Resultado reportado em Verney et al.(25).	23
2.8	Aplicação do algoritmo, encontrando posições globais em uma pequena grade sísmica. Resultado reportado em Pauget et al.	23
2.9	(a) Cubo Sísmico, (b) geo-modelo resultante, (c) horizonte mapeado a partir do geo-modelo. Resultado reportado em Pauget et al.(23).	24
3.1	Representando horizontes sísmicos por janelas de vizinhos verticais. Na imagem, a linha azul identifica o horizonte formado por voxels que são picos de amplitude positiva. Tais voxels podem ser representados por uma janela vertical que incluem a si mesmos (no centro) e aos seus vizinhos verticais no mesmo traço sísmico, caso do voxel representado pela janela em vermelho.	26
3.2	Algumas amostras verticais tomadas nos traços representando voxels de entrada (esquerda), que foram classificadas num mesmo grupo devido à sua similaridade. À direita vemos o <i>codevector</i> do grupo resultante, que por sua vez irá representar todas as amostras classificadas nesse grupo.	26

- 3.3 Processo de criação de um atributo de visualização por Seismic Facies Analysis. À esquerda vemos uma seção sísmica de amplitudes retirada do dado sísmico 3D. Depois de criadas as amostras verticais representando cada um dos voxels da seção, as amostras são divididas por semelhança em sub-grupos distintos. Cada sub-grupo recebe uma cor específica. Na porção direita da figura vemos o atributo de visualização resultante, onde o valor de cada voxel do volume de entrada é substituído pelo identificador numérico do sub-grupo onde sua amostra vertical correspondente foi classificada. 27
- 3.4 Ilustração do processo genérico de criação e classificação das amostras de janelas verticais que representam os voxels. Para cada voxel a_{ti} do volume é criada uma amostra contendo seus n vizinhos verticais acima e abaixo de a_{ti} no traço sísmico. Cria-se um *dataset* contendo todas essas amostras verticais. Feito isso, esse *dataset* é processado a partir de um algoritmo de agrupamento de dados, que divide esse conjunto de amostras em m sub-grupos de amostras parecidas entre si. Cada sub-grupo recebe um identificador numérico que o referencia, $\mathbf{C} = \{c_1, \dots, c_m\}$. 27

No que diz respeito ao empenho, ao compromisso, ao esforço, à dedicação, não existe meio termo. Ou você faz uma coisa bem feita ou não faz.

Ayrton Senna da Silva, 1988.

1

Introdução

Um volume sísmico tridimensional é resultante do processo de aquisição indireta dos dados sísmicos de uma determinada região geológica de interesse. Os dados armazenados em um volume desse tipo são função de características particulares da região geológica na qual foram adquiridos, revelando informações estruturais e estratigráficas das subcamadas geológicas existentes.

O processo de aquisição desses dados consiste em produzir artificialmente ondas elásticas de curta duração (impulsos) em pontos específicos da superfície terrestre. À medida em que essas ondas se propagam para o interior do subsolo, a cada vez que encontram uma interface entre camadas de rochas diferentes, parte da energia da onda incidente é refletida de volta à superfície, enquanto outra parte é refratada às camadas mais profundas. A cada reflexão, a amplitude da onda refletida será proporcional à diferença de impedância entre as duas camadas. Quando a onda refletida volta à superfície, sensores específicos registram a variação da amplitude e o tempo gasto nesse processo. Para a aquisição dos dados 3D, esse procedimento é repetido, variando-se as posições dos sensores e da fonte da onda sísmica a cada passo.

A organização das amostras em um dado sísmico é demonstrada na Figura 1.1. A figura apresenta do lado esquerdo um traço sísmico, que idealmente traduz o que acontece na subsuperfície no ponto exatamente abaixo de onde ele foi adquirido. Num traço, portanto, a única dimensão é a temporal (1D). Nesse caso, o traço sísmico é representado através de um sinal ondulante. No centro temos uma seção vertical do conjunto de dados formada por um conjunto de traços sísmicos, que é chamada de linha sísmica (2D). Uma linha sísmica é formada por uma dimensão espacial e outra temporal. Por fim apresentamos os dados sísmicos 3D (volume sísmico), formados por várias linhas sísmicas. Nesse caso temos duas direções espaciais, que são chamadas de inline (direção das linhas sísmicas) e crossline (direção perpendicular às linhas sísmicas), além de uma direção temporal. Vale notar que, ao mostrarmos dados em 2D e 3D a forma de representação do sinal de cada um dos traços sísmicos deixa de ser feita a partir de um sinal ondulante, passando a ser feita através de uma escala

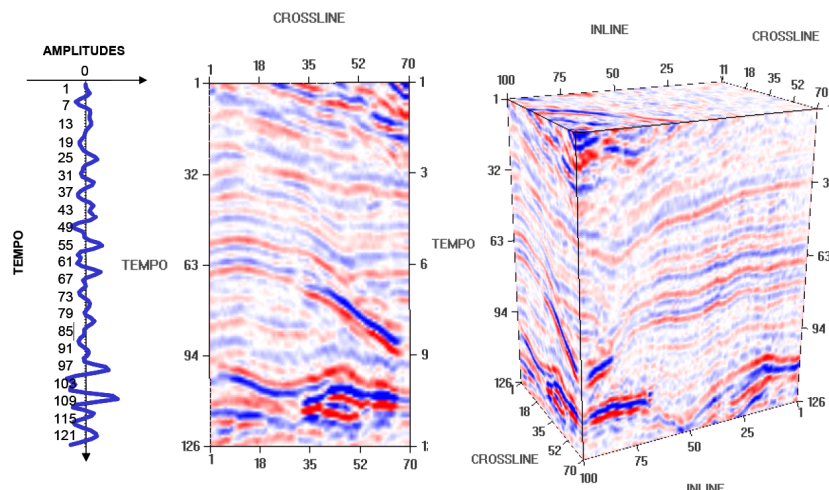


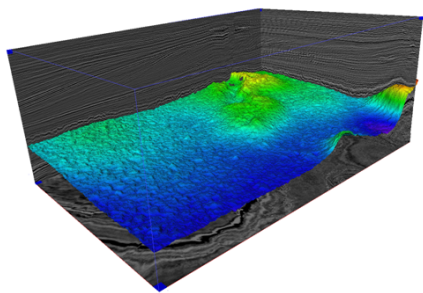
Figura 1.1: Traço sísmico (esquerda), linha sísmica (centro) e volume sísmico (direita).

visual que associa cores aos valores numéricos da amplitude da onda.

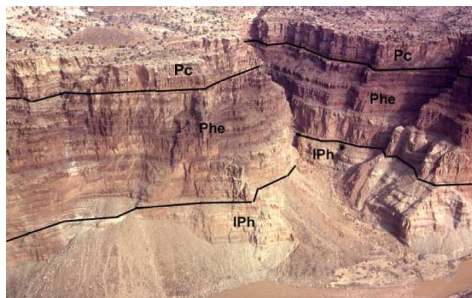
Uma vez adquiridos e processados, os dados sísmicos devem passar por uma fase de interpretação. A interpretação das feições geológicas nesses conjuntos de dados pode indicar situações favoráveis à acumulação de hidrocarbonetos. Durante o processo é necessário fazer a identificação de determinadas estruturas geológicas, tais como interfaces, falhas e domos de sal, entre outras.

Na geologia, define-se uma interface como uma subsuperfície de sedimentos de idade geológica semelhante que se destaca no dado sísmico. O produto da interpretação de uma interface é denominado horizonte. Os horizontes se distinguem por apresentarem características geológicas particulares, tais como a espessura da camada de sedimentos que os define e características físicas do material sedimentar que os compõe, além da sua vizinhança geológica. Nos dados sísmicos, um horizonte se manifesta como uma série de eventos (picos ou vales de amplitudes sísmicas) identificados de forma consistente traço a traço. O rastreamento dessas superfícies consiste em descobrir em quais dos voxels do volume o horizonte está representado, identificando seu conjunto de amostras. A Figura 1.2 ilustra a definição de um horizonte. Em 1.2(a) apresentamos um horizonte sísmico mapeado como uma superfície a partir da identificação dos seus voxels (valores de um gride regular num espaço tridimensional) em um dado sísmico 3D. Em 1.2(b) ilustra-se o que seria o equivalente na natureza.

Falhas podem ser identificadas através de quebras na continuidade dos horizontes. São fraturas que permitem um deslocamento relativo das rochas, fazendo com que elas percam sua continuidade original. O mapeamento de uma falha sísmica consiste em encontrar o conjunto de voxels do volume que evi-



1.2(a): Horizonte sísmico formado pelo seu conjunto de voxels identificados na sísmica.

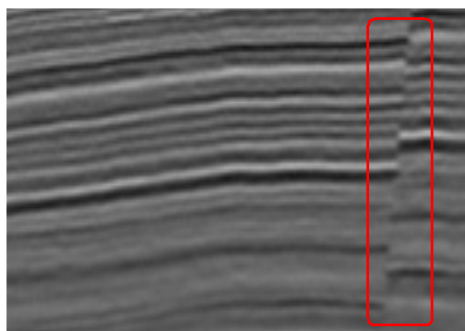


1.2(b): Imagem identificando interfaces in loco.

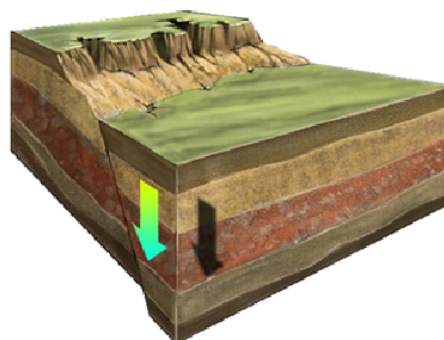
uff.br/geofisica/index.php/interpretacao-sismica

Figura 1.2: (a) um horizonte sísmico mapeado a partir da sísmica 3D. (b) equivalente ao horizonte num ambiente real.

dencia essa quebra de continuidade. A Figura 1.3 apresenta a falha geológica. Em 1.3(a) vemos a falha identificada em uma fatia transversal do volume sísmico 3D, destacada pela janela em vermelho. Em 1.3(b) apresentamos uma ilustração esquemática de uma falha sísmica na natureza.



1.3(a): Falha sísmica identificada na seção sísmica por uma quebra da continuidade original dos horizontes, evidenciada pela janela em vermelho.



1.3(b): Imagem esquemática de uma falha sísmica.

uff.br/geofisica/index.php/interpretacao-sismica

Figura 1.3: (a) falha sísmica identificada na seção sísmica 3D. (b) um modelo ilustrativo do que seria o equivalente na natureza.

1.1 Objetivo

O presente trabalho desenvolve de métodos de mapeamento automático de horizontes sísmicos e atributos visuais indicativos de regiões de descontinuidade lateral existentes em dados sísmicos, tais como falhas sísmicas. No que diz respeito ao mapeamento de horizontes, nos concentramos no desenvolvimento de algoritmos que permitam aos intérpretes o mapeamento de grandes áreas

de dados, capazes de lidar com objetos sísmicos que interrompam a estratigrafia dos horizontes, tais como domos de sal, intrusões ígneas ou regiões de deposição caótica, entre outras.

1.2

Motivação

A fase de interpretação dos dados é uma das principais etapas de todo o processo de descoberta de reservas de óleo e gás, e o mapeamento de horizontes e falhas sísmicas são dois dos processos mais trabalhosos, resultando no grande tempo que as equipes de interpretação levam para entender uma determinada região, para com isso identificar possíveis áreas em que exista maior probabilidade de acumulação de hidrocarbonetos.

Nesse sentido, algoritmos que possam automatizar ou auxiliar a tarefa de rastreamento de horizontes são de grande valia. Tais algoritmos, no entanto, são difíceis de serem desenvolvidos, por diversas razões: relação sinal/ruído inadequada devido a interferências no processo de aquisição de dados sísmicos e complexidade da geologia da área onde o volume de dados foi adquirido, entre outras.

Tradicionalmente, métodos automatizados de rastreamento de horizontes funcionam extraindo horizontes pela correlação das amplitudes locais entre traços vizinhos, mapeando superfícies por estratégias de crescimento de região. O processo de mapeamento começa a partir de um voxel de entrada, denominado voxel semente. Os vizinhos imediatos da semente são identificados e incluídos na superfície. Depois, os vizinhos já descobertos são utilizados como novas sementes para a descoberta de novos vizinhos, seguindo esse processo até que todos os voxels da superfície tenham sido descobertos.

Essa estratégia tende a falhar se o horizonte rastreado apresenta variação brusca de suas amplitudes, inversão da polaridade do seu sinal lateral ou na presença de descontinuidades, tais como falhas sísmicas. O ponto fraco dessas estratégias é que elas utilizam medidas baseadas em informação local. Mapeia-se um horizonte global resolvendo-se milhões de problemas locais. Um único erro no mapeamento se propagará indefinidamente. Devido a essas limitações, a utilização desses métodos é trabalhosa e demorada, limitando-se a regiões com alta qualidade de sinal e geologia relativamente simples.

Recentemente, abordagens globais têm sido propostas destinadas a inferir os modelos geológicos diretamente, sem a necessidade de identificar os horizontes específicos a serem mapeados. Esses novos algoritmos são de natureza global, estimando as superfícies em todo o volume simultaneamente. Tais

métodos são supostamente capazes de explorar a dimensionalidade completa dos dados para rastrear vários horizontes em paralelo, usando todo o conjunto de dados para encontrar uma solução de desajuste mínimo oferecendo, potencialmente, uma solução mais precisa. Vale destacar que essas metodologias se destinam principalmente à produção de modelos em que as superfícies são aproximações daquelas existentes no volume sísmico. Tais modelos, apesar de úteis em alguns contextos, possuem aplicação limitada dentro do fluxo de trabalho da interpretação sísmica voltada à descoberta de hidrocarbonetos.

1.3

Contribuições

Nesse trabalho desenvolvemos metodologias que contornem as limitações das abordagens de mapeamento de horizontes e de produção de atributos de visualização atuais.

Diferentemente dos métodos que revisamos, nosso trabalho deixa de utilizar medidas de similaridade local entre voxels. Para isso, ao invés de mapear eventos sísmicos diretamente no volume de entrada, criamos uma etapa baseada em algoritmos de agrupamento de dados para o pré-processamento desses dados. Essa etapa gera um novo volume, onde as feições sísmicas, de forma geral, tornam-se mais evidentes.

Utilizando essa nova representação de dados, apresentamos métodos em que o processo de busca por voxels semelhantes se baseia em informação global. No volume que produzimos, os voxels fazem referência aos grupos gerados durante o processo de agrupamento de dados, que por sua vez armazenam informações a respeito dos voxels vizinhos mais similares.

Essa característica nos permite apresentar algoritmos em que as sementes de mapeamento de horizontes podem se manter fixas, minimizando a probabilidade de propagação dos erros de mapeamento.

Nossos métodos de mapeamento de horizontes se adequam tanto ao mapeamento de um horizonte particular ou ao mapeamento dos dados sísmicos como um todo, e produz superfícies que contornam adequadamente as discontinuidades presentes nos dados, conforme demonstrado nos resultados que apresentaremos.

Utilizando o mesmo pré-processamento de transformação do volume sísmico de entrada num volume de clusters, apresentamos um método destinado a produzir volumes de atributos onde regiões de descontinuidade lateral tendem a ser enfatizadas. Tal atributo, ao ser visualizado, facilita a identificação dessas descontinuidades por parte dos intérpretes.

Durante o desenvolvimento da tese, não encontramos nenhum tipo de benchmark que pudesse ser utilizado na avaliação de quaisquer algoritmos destinados ao mapeamento automático de horizontes. Para que nossos resultados possam ser avaliados de maneira mais abrangente, criamos um benchmark público de horizontes e o tornamos disponível na internet em www.tecgraf.puc-rio.br/seismic/repo.

1.4

Organização deste Documento

O presente trabalho foi redigido na forma de capítulos, com esta introdução geral contendo nosso objetivo, nossa motivação e as nossas contribuições. Após isso, apresentamos um capítulo destinado aos trabalhos relacionados, onde fazemos uma revisão bibliográfica destinada a caracterizar o estado da arte em nossa área. As metodologias que desenvolvemos, os resultados e as conclusões específicas estão apresentados na forma de artigos científicos apresentados em conferências internacionais, ou submetidos a periódicos internacionais. Por fim, no último capítulo apresentamos nossas conclusões.

2

Trabalhos Relacionados

Algoritmos de aprendizado de máquina são bastante utilizados no fluxo de trabalho da exploração de hidrocarbonetos, como visto em (6) e (14).

Boa parte dos métodos apresentados na literatura para o mapeamento de horizontes sísmicos usam combinações de mapeamento 2D ou 3D com a utilização de filtros locais específicos. Dentre esses destacamos os principais.

O'Malley and Kakadiaris (21) apresentam um filtro específico denominado orientation-isotropy adaptive (OIA) Gaussian. Seus resultados sugerem que a metodologia funciona adequadamente para o caso de mapeamento de horizontes contínuos. No entanto, concluem o trabalho enfatizando que os algoritmos de mapeamento de horizonte baseados em filtros de busca local são insuficientes para interpretar muitos tipos de horizontes, especialmente aqueles que são descontínuos ou situados em áreas de discordâncias. Outros trabalhos que se destacam pela utilização de filtros específicos são vistos em Li et al. (17) e em Yu et al. (28).

Admasu e Toennies (1) apresentam um algoritmo projetado para correlacionar horizontes através de falhas que pode ser usado para mapeamento de horizonte. No entanto, como relatado pelos autores, por se basear em filtragem local o método produz resultados inconsistentes em vários casos.

Aurnhammer and Tonnies (2) apresentam um método que utiliza algoritmos genéticos para mapear horizontes e é capaz de correlacionar porções distintas de horizontes através de falhas sísmicas em 2D.

Recentemente, abordagens globais têm sido propostas visando a interpretação completa dos dados sem a necessidade da escolha de horizontes específicos. Hoyes e Cheret (15) apresentam uma revisão dos métodos de interpretação globais para mapeamento de horizontes. Tais metodologias podem ser divididas da seguinte forma: métodos baseados em mergulho, métodos baseados em junção de sub horizontes, e métodos baseados em processos de otimização global. A seguir detalhamos cada um deles.

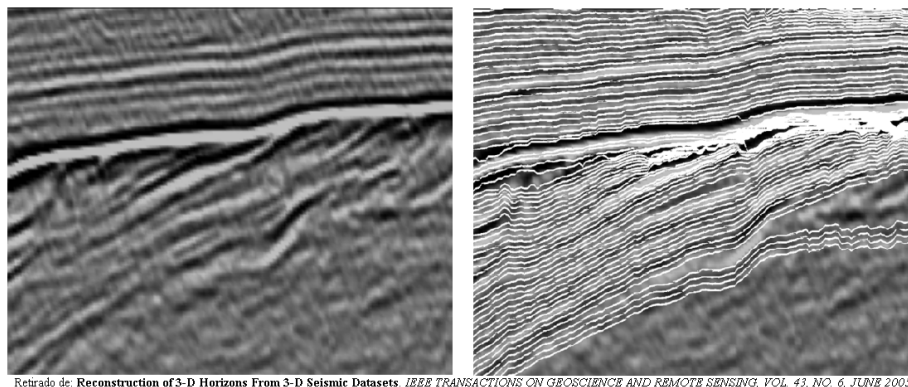


Figura 2.1: Resultado reportado em Blinov e Petrou (4).

2.1

Métodos baseados em mergulho

Esses algoritmos trabalham buscando estimativas dos vetores normais dos horizontes ao longo dos voxels do volume. A partir das estimativas mapeiam em paralelo os horizontes identificados e que possam ser mapeados ao longo de todo o dado. Dentre suas limitações destaca-se o fato de que os objetos mapeados são, na verdade, aproximações, ou superfícies de tempo geológico relativo, não sendo um horizonte propriamente dito. Dentre essas metodologias, destacam-se as citadas abaixo.

Blinov e Petrou (4) propuseram um método que se concentrava na extração de superfícies "quase-horizontais" em dados sísmicos 3D. Depois de estimar a normal dos horizontes, usa-se um filtro local destinado a suavizar o conjunto de normais obtidas, reduzindo ruído e tornando o sinal mais homogêneo. Num segundo passo o algoritmo utiliza os dados das normais aproximadas para mapear simultaneamente todos os horizontes que tenham sido detectados no cubo sísmico. Os autores reportam resultados somente em fatias 2D (Figura 2.1). O artigo não apresenta uma análise do comportamento do método face a descontinuidades normalmente presentes em dados reais, tais como falhas sísmicas ou domos de sal. Além disso, o artigo omite informações importantes relativas aos dados utilizados para reportar os resultados, tais como a largura da área mapeada.

Lomask e Guitton (18) apresentam um método destinado à horizontalização automática de dados sísmicos. Assim como em Fomel et al. (10), o método envolve o cálculo local das normais dos voxels em todo o dado, utilizando uma técnica específica de estimativa de mergulho. Os dados de mergulho são utilizados para a descoberta inicial do modelo, onde possíveis discordâncias (ou mudanças de profundidade) são minimizadas por um sistema de mínimos

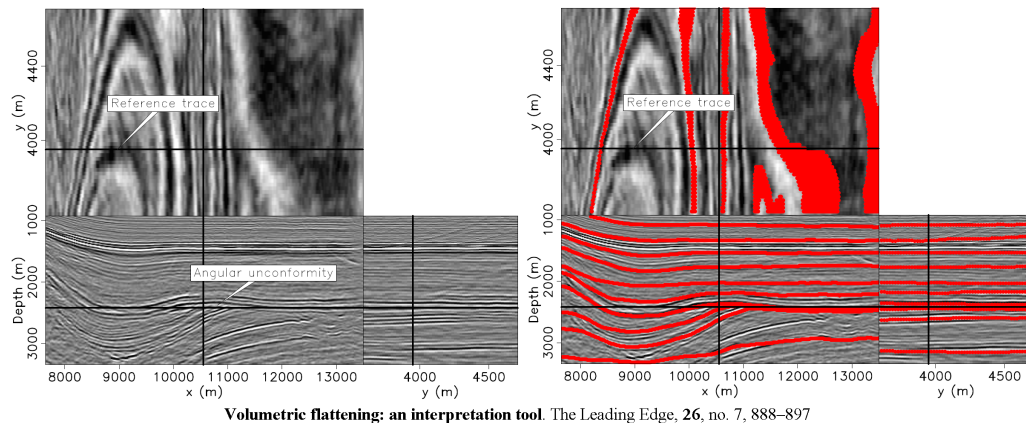


Figura 2.2: Resultado reportado em Lomask e Guitton (18).

quadrados não-linear. De posse das estimativas do comportamento dos horizontes, os dados são subsequentemente horizontalizados, de forma a produzir um volume próximo ao volume de tempo geológico relativo. Os autores apresentam resultados indicando a eficiência do método em tratar zonas de discordância (Figura 2.2). O texto faz menção ao fato de que a metodologia não é capaz de lidar com interrupções do sinal sísmico adequadamente. Para que o método processe dados que possuam falhas sísmicas, tais falhas tem que ser fornecidas como entrada ao modelo.

De Groot et al. (13) e (24) descrevem um método de interpretação voltado ao mapeamento de sequências de eventos crono-estratigráficos (superfícies de tempo geológico relativo aproximado), que funciona a partir das estimativas de mergulho dos voxels. Para funcionar o método recebe como entradas horizontes de topo e base delimitando a zona a ser estudada. Tais horizontes, bem como possíveis discontinuidades como falhas e domos de sal, precisam ser mapeados convencionalmente e fornecidos como dados de entrada. Na Figura 2.3 apresentamos um resultado típico dessa metodologia. Na parte esquerda da figura vemos uma seção 2D dos horizontes de topo e base fornecidos como restrições de entrada. Na parte esquerda de 2.3 apresentamos os resultados fornecidos pelo método. O artigo não apresenta superfícies mapeadas em três dimensões.

Wu e Hale (26) e (27) descrevem o algoritmo mais competitivo dentre todos os que revisamos durante o desenvolvimento da tese. Os autores apresentam dois métodos destinados à construção de horizontes sísmicos: gerando as superfícies uma de cada vez, ou gerando um volume completo delas. Tais horizontes são encontrados por resolução de equações diferenciais parciais que por sua vez são obtidas a partir de um volume de normais (mergulho) contendo estimativas do vetor normal dos voxels do volume.

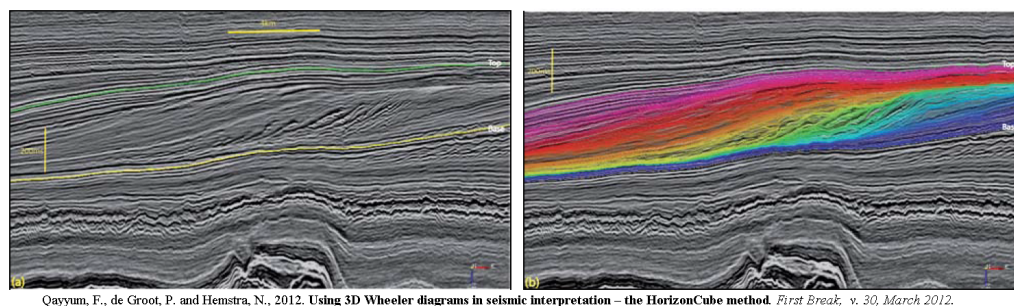


Figura 2.3: Resultado reportado em De Groot et al.

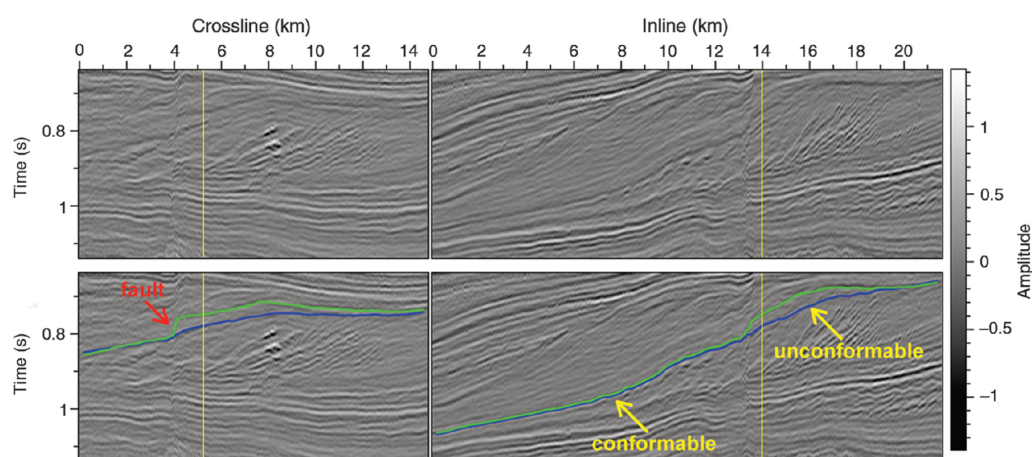


Figura 2.4: Seções sísmicas mostrando linhas extraídas usando um ponto de controle (curva azul) e 19 pontos de controle (curva verde). Resultado reportado em Wu e Hale (27).

No texto, os autores admitem que estimativas dos vetores normais não são precisas o suficiente para obter automaticamente o mapeamento adequado dos horizontes ao redor de zonas de discordância, falhas sísmicas, ou em áreas onde uma imagem é ruidosa.

No entanto, o aspecto mais significativo do seu método é que ele permite que o usuário ajude na construção do horizonte, especificando de forma interativa, durante a interpretação, conjuntos de pontos de controle que definem as restrições do modelo, como pode ser visto nas Figuras 2.4 e 2.5. O trabalho também apresenta um método que constrói o volume de tempo geológico relativo, ou volume horizontalizado, a partir dos horizontes descobertos.

Os autores concluem afirmando um defeito dos métodos, o fato de eles não produzirem automaticamente os espaços que representam o rejeito existente entre as duas porções do horizonte interrompidas por uma dada falhas.

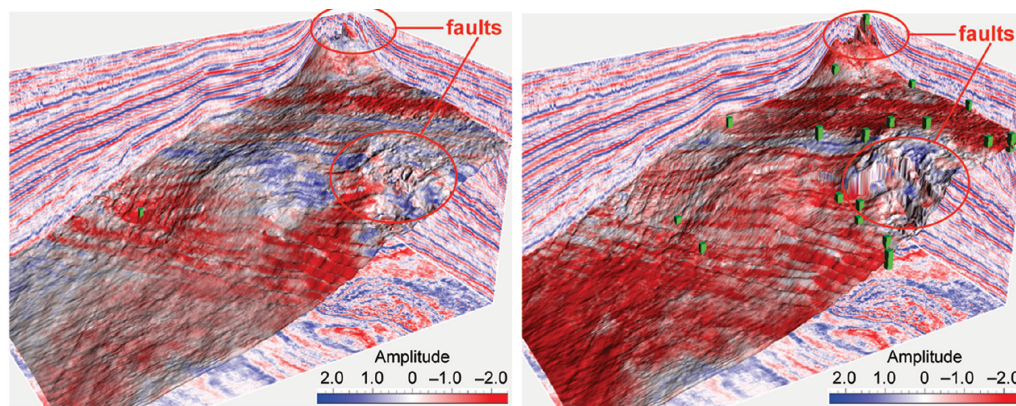


Figura 2.5: Uma vista 3D das superfícies extraídas usando um ponto de controle (esquerda) e 19 pontos de controle (direita), respectivamente. Ambas as superfícies são coloridas por amplitude. Resultado reportado em Wu e Hale (27).

2.2

Métodos baseados em junção de sub-horizontes

Essa sub-classe de métodos de mapeamento busca encontrar relacionamentos topológicos que possam produzir sub-superfícies sísmicas por semelhança. Num segundo passo, esses pedaços de horizontes são mesclados para formar horizontes maiores ordenados cronologicamente. Dentre esses métodos destacamos os descritos resumidamente abaixo.

Borgos et al. (5) apresentam um método que constrói sequências de horizontes a partir de pré-horizontes descobertos. Primeiramente os traços sísmicos do volume são aproximados por polinômios, de forma a melhorar a qualidade do sinal e remover ruído. Depois, todos os picos de máximo e mínimo de amplitudes são identificados. Por fim o algoritmo utiliza técnicas específicas de classificação dos voxels destinadas a agrupar eventos sísmicos que sejam vizinhos e que tenham formato de onda similar. O trabalho manual dos intérpretes consiste em selecionar e combinar os vários fragmentos que pertençam aos horizontes de interesse, formando a superfície final. Na figura 2.6 apresentamos um resultado reportado no artigo que descreve o método. Na parte esquerda da figura podemos visualizar os fragmentos de horizontes. Na porção direita da figura vemos um horizonte resultante depois de o usuário ter manualmente selecionado os fragmentos que o formam.

Bakke et al. (3) e Gramstad et al. (12) propuseram um método que utiliza algoritmo de alinhamento de sequências para encontrar horizontes. Verney et al. (25) desenvolveram um método de interpretação sísmica que constrói e agrupa fragmentos de horizontes utilizando uma abordagem baseada em visão cognitiva, uma sub-área da visão computacional destinada a construir métodos

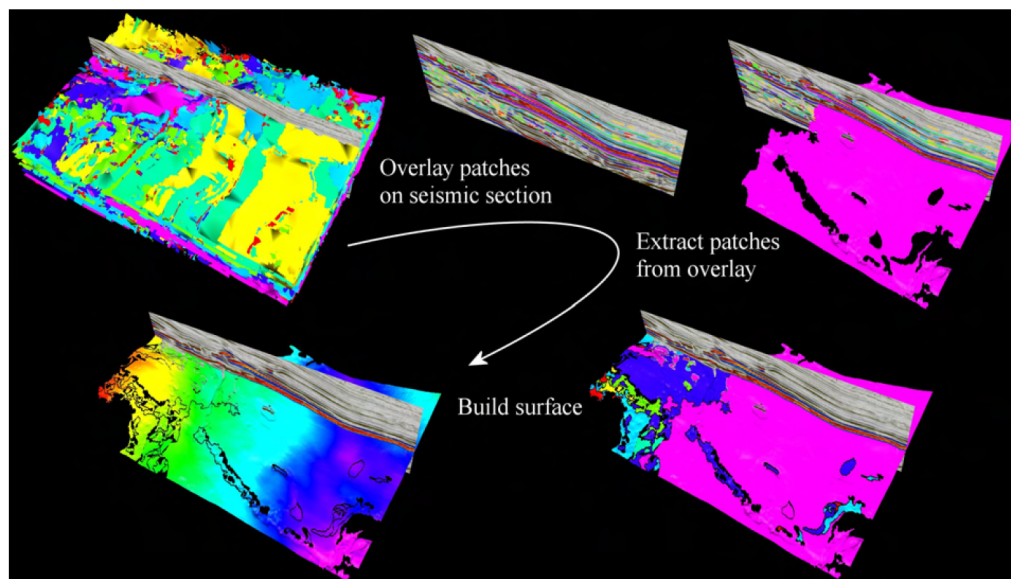


Figura 2.6: Processo de mapeamento de horizonte descrito em Borgos et. al. (5). A partir do conjunto de fragmentos de horizontes descobertos pela metodologia, o usuário seleciona um a um aqueles que pertencem ao horizonte a ser mapeado. O método é capaz de agrupar esses fragmentos e remover possíveis discordâncias, gerando o horizonte final.

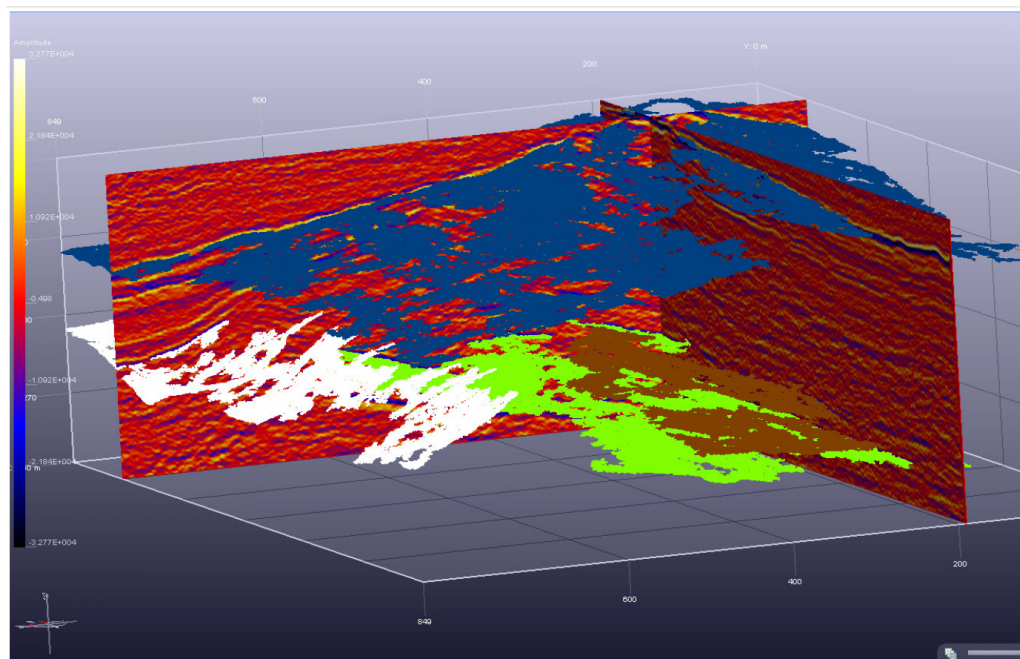
computacionais baseados na cognição e que visam alcançar funcionalidades de detecção, localização, e reconhecimento de objetos. Segundo os autores, tais habilidades são adequadas ao problema interpretação sísmica. Na Figura 2.7 apresentamos um dos resultados reportados.

2.3

Métodos baseados em processos de otimização global

Essa classe de métodos modela os dados de entrada buscando minimizar uma função de custo global que geralmente usa como entrada uma série de arestas ligando voxels que supostamente pertencem ao mesmo horizonte no volume 3D. Esses conjuntos de arestas são descobertos utilizando alguma métrica de correlação entre voxels que estejam situados em traços sísmicos diferentes. A cada aresta associa-se um grau de correlação entre os voxels que ela está ligando. O melhor modelo é então estimado movendo as arestas localmente de forma que a função de custo atinja um mínimo global.

Esta é a estratégia adotada em Pauget et al. (23). Num primeiro passo, os traços sísmicos são avaliados com o objetivo de construir um grid regular no plano XY, o plano da fatia sendo analisada. A construção desse grid leva em conta cada voxel de cada um dos traços, e dependendo do tamanho da fatia pode requerer uma sub-amostragem de forma a reduzir o custo computacional.



Verney et al. (2008). An approach of seismic interpretation based on cognitive vision. 70th EAGE Conference and Exhibition.

Figura 2.7: Resultado reportado em Verney et al.(25).

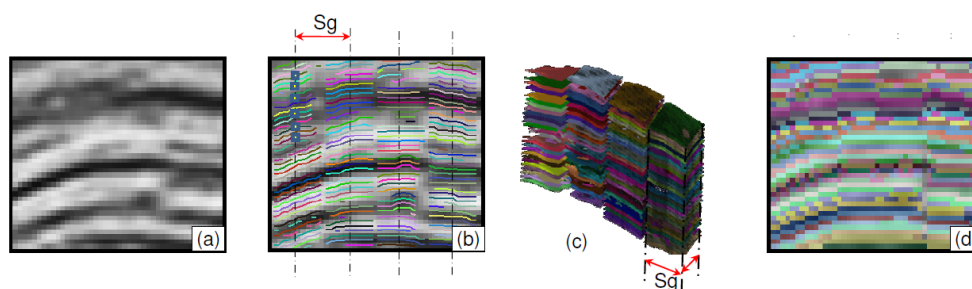


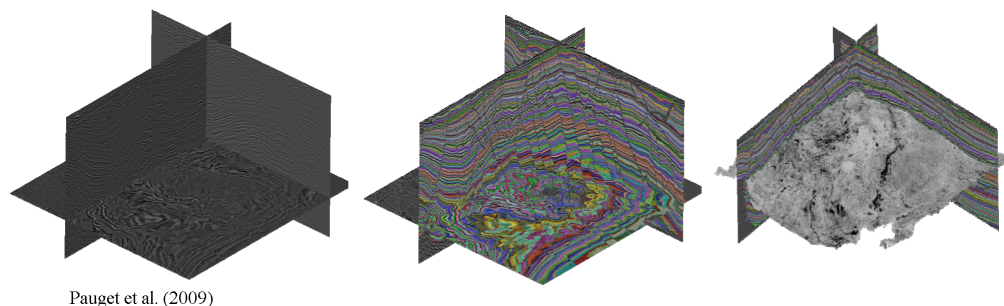
Figura 2.8: Aplicação do algoritmo, encontrando posições globais em uma pequena grade sísmica. Resultado reportado em Pauget et al.

Uma vez que o grid foi criado, seus pontos são correlacionados traço a traço para cada par de traços vizinhos. Feito isso o método é capaz de criar um modelo aproximado da sísmica em questão, fatia a fatia. O processo é ilustrado na figura 2.8. Um resultado reportado pelos autores pode ser visualizado na figura 2.9.

2.4

Limitações gerais

De forma geral os métodos de interpretação atuais alcançam resultados satisfatórios somente em dados que reflitam a natureza estratificada da sísmica. Tais métodos não são capazes de lidar com objetos sísmicos que interrompam essa estratigrafia, como domos de sal, intrusões ígneas ou regiões de deposição



Pauget et al. (2009)

Figura 2.9: (a) Cubo Sísmico, (b) geo-modelo resultante, (c) horizonte mapeado a partir do geo-modelo. Resultado reportado em Pauget et al.(23).

caótica. Essas limitações são contornadas caso os modelos recebam como entrada esse objetos já interpretados por algum método externo. Segundo Hoyes e Cheret (15) a maioria dos métodos atuais requerem dados de entrada de alta qualidade a fim de produzir um modelo válido da subsuperfície geológica sendo estudada.

Por fim, esses métodos se destinam à produção de modelos em que as superfícies são aproximações daquelas existentes no volume sísmico. Apesar de úteis em vários contextos, os horizontes produzidos possuem aplicação limitada dentro do fluxo de trabalho de interpretação sísmica para descoberta de hidrocarbonetos (16).

Os resultados que apresentaremos nos artigos que constam nessa tese sugerem que nossos métodos resolvem boa parte dessas limitações. Ao longo do texto, apresentamos horizontes formados somente por máximos ou mínimos locais, sendo portanto totalmente feis ao dado sísmico. Além disso, esses horizontes foram mapeados adequadamente, respeitando as discontinuidades presentes nos dados sísmicos, tenham sido elas falhas, domos de sal ou regiões de deposição caótica, entre outras.

3

Métodos Desenvolvidos

Esta pesquisa se baseia na representação dos voxels por janelas verticais de vizinhos no mesmo traço sísmico, criando uma assinatura que caracteriza horizontes ao longo de sua vizinhança lateral, conforme ilustrado na Figura 3.1. Esta representação é comumente utilizada no campo da Seismic Facies Analysis (Marroquín, Brault e Hart (19)), onde amostras verticais são agrupadas por semelhança, como ilustrado na Figura 3.2.

Em nosso primeiro algoritmo, dado um volume sísmico de entrada, criamos vetores de amostras verticais de amplitudes para representar cada um dos seus voxels. Utilizamos esses vetores como entradas de um algoritmo de *clustering* baseado no Growing Neural Gas (GNG) (11), criando sub grupos que representam grupos de assinaturas semelhantes, ou seja com pequenas variações de amplitudes verticais. Esse processo é ilustrado na Figura 3.3. De posse do conjunto de *clusters* criamos um novo volume de mesmas dimensões espaciais do volume sísmico de entrada, onde o valor de amplitude de cada voxel do volume de entrada foi substituído pelo número identificador do grupo onde sua amostra vertical correspondente foi classificada. Por fim definimos formalmente uma função de similaridade que se baseava na distribuição dos grupos no espaço de amostras, o que consequentemente tornava a função de similaridade auto-adaptável. Utilizamos o volume de *clusters* para mapear horizontes num método global. A íntegra do texto desse trabalho, apresentado na conferência da European Association of Geoscientists and Engineers (EAGE) 2013 (9), está no Anexo A.

O processo de classificação das amostras em *clusters*, passando a ser representadas pelo seu *cluster* correspondente, filtra discordâncias existentes no dado original. Verificamos que no volume de *labels* as regiões de falha ou outras discontinuidades tendem estar melhor definidas, conforme apresentamos na Figura 3.4.

A partir dessa observação desenvolvemos um método que usava a mesma metodologia de *clustering* descrita no primeiro artigo, porém com o objetivo de produzir um atributo de visualização destinado a enfatizar regiões de des-

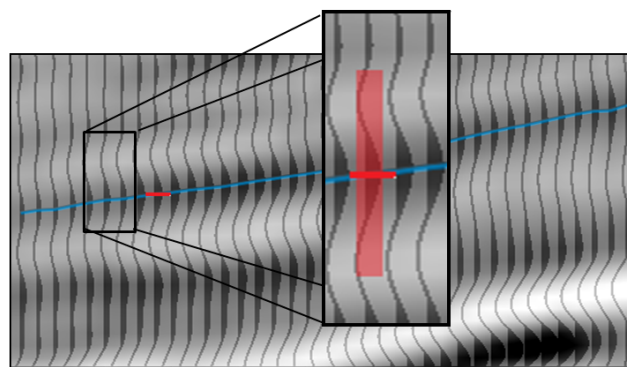


Figura 3.1: Representando horizontes sísmicos por janelas de vizinhos verticais. Na imagem, a linha azul identifica o horizonte formado por voxels que são picos de amplitude positiva. Tais voxels podem ser representados por uma janela vertical que incluem a si mesmos (no centro) e aos seus vizinhos verticais no mesmo traço sísmico, caso do voxel representado pela janela em vermelho.

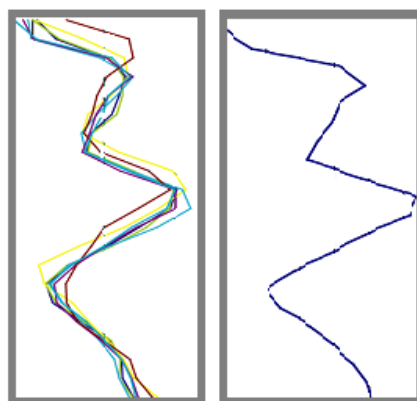


Figura 3.2: Algumas amostras verticais tomadas nos traços representando voxels de entrada (esquerda), que foram classificadas num mesmo grupo devido à sua similaridade. À direita vemos o *codevector* do grupo resultante, que por sua vez irá representar todas as amostras classificadas nesse grupo.

continuidade presentes no dado sísmico de entrada. Testamos essa metodologia em diversos dados sísmicos, e comparamos o resultado de visualização obtido com métodos disponíveis na literatura destinados a produzir atributos de visualização semelhantes, tais como o atributo de curvatura (20). Os resultados que obtivemos foram competitivos, enfatizando algumas áreas da sísmica que não eram ressaltadas pelos outros métodos. Essa foi a base do segundo trabalho que produzimos, apresentado na conferência da Society of Exploration Geophysicists (SEG) em 2014 (7). A íntegra do trabalho está apresentada no Anexo B.

Os resultados alcançados pelo método de mapeamento de horizontes se mostraram competitivos nos dados de que dispúnhamos. Tais dados, apesar de reais e de dimensões consideráveis (650 inlines por 950 crosslines, numa

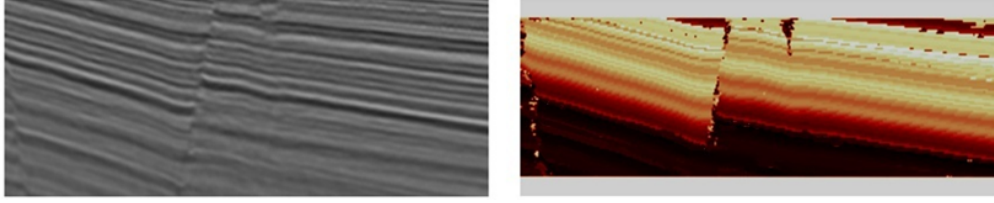


Figura 3.3: Processo de criação de um atributo de visualização por Seismic Facies Analysis. À esquerda vemos uma seção sísmica de amplitudes retirada do dado sísmico 3D. Depois de criadas as amostras verticais representando cada um dos voxels da seção, as amostras são divididas por semelhança em sub-grupos distintos. Cada sub-grupo recebe uma cor específica. Na porção direita da figura vemos o atributo de visualização resultante, onde o valor de cada voxel do volume de entrada é substituído pelo identificador numérico do sub-grupo onde sua amostra vertical correspondente foi classificada.

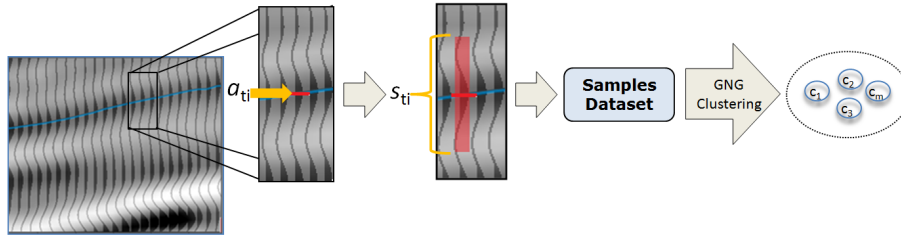


Figura 3.4: Ilustração do processo genérico de criação e classificação das amostras de janelas verticais que representam os voxels. Para cada voxel a_{ti} do volume é criada uma amostra contendo seus n vizinhos verticais acima e abaixo de a_{ti} no traço sísmico. Cria-se um *dataset* contendo todas essas amostras verticais. Feito isso, esse *dataset* é processado a partir de um algoritmo de agrupamento de dados, que divide esse conjunto de amostras em m sub-grupos de amostras parecidas entre si. Cada sub-grupo recebe um identificador numérico que o referencia, $\mathbf{C} = \{c_1, \dots, c_m\}$.

área total de 16 por 24 quilômetros, ou 384 km^2), eram pequenos quando comparados a alguns dados disponíveis num ambiente de produção atual. Ao avaliarmos o desempenho do método nesses dados de produção, os resultados ficaram um pouco aquém do que esperávamos.

O fato de o número de *clusters* que criamos ser uma aproximação trazia inconvenientes. Amostras contíguas representando voxels num mesmo traço poderiam ser classificadas em um mesmo *cluster*. Verificamos essa ocorrência principalmente em torno dos picos de amplitude (imediatamente acima e abaixo dos picos).

Horizontes são caracterizados por 4 tipos de padrões: picos de amplitudes positivas (máximos locais), picos negativos (mínimos locais), pontos de valor nulo com cruzamento negativo para positivo, ou valor nulo com cruzamentos positivo para negativo.

Essas observações nos levaram ao desenvolvimento de um método de rastreamento de horizontes destinado a mapear horizontes com maior precisão em dados sísmicos maiores. Essa nova abordagem também utiliza amostras verticais para representar os voxels dos horizontes. Contudo, o processamento é executado somente para o sub conjunto dos voxels que façam parte dos horizontes de interesse: horizontes formados por picos, ou formados por cruzamentos de zero. Os resultados que alcançamos se mostraram bastante competitivos. No que diz respeito às dimensões da área mapeada, ao tamanho em memória do volume de dados processado, e ao número de horizontes que apresentamos, não encontramos na literatura resultados semelhantes. O nosso principal dado de testes representava uma área de 42 por 42 km, totalizando 1.764 km^2 , com aproximadamente 7 Gb de dados. O texto descrevendo o algoritmo, apresentado na conferência da Society of Exploration Geophysicists (SEG) 2014 (8), está disponível em detalhes no Anexo C.

Durante o processo de desenvolvimento da tese, fizemos algumas constatações importantes:

1. Nenhum dos trabalhos da área disponibilizava um *benchmark* que permitisse a comparação direta do desempenho alcançado entre as metodologias que apresentamos e quaisquer outras disponíveis na literatura;
2. Deveríamos desenvolver um critério objetivo para selecionar os tamanhos das amostras verticais utilizadas para processar cada dado. Em nenhum momento apresentamos algum critério objetivo que permitisse encontrar os parâmetros de entrada que os algoritmos, tais como quantos voxels utilizar nas janelas verticais, quantos *clusters* formar nos processos de agrupamento, ou mesmo quantos processos de agrupamento formar;
3. Houveram sugestões para que executássemos testes de desempenho substituindo o volume de amplitudes de entrada por algum volume de atributos equivalente, dentre os vários que estão disponíveis na literatura (fase instantânea, envelope, cosseno da fase instantânea, entre outros);
4. Por último foi observado que ainda não havíamos realizado processos de extração de *features* atualmente utilizadas na literatura de *machine learning* (PCA, Deep Learning, entre outras). Supostamente, a utilização de tais algoritmos deveria gerar um novo vetor de características a partir de cada amostra de voxels verticais, de forma a enfatizar as características dos vetores de entrada, melhorando o desempenho.

Procuramos responder a todas essas perguntas e relatamos seus resultados no último dos artigos escritos durante o processo de desenvolvimento da tese, submetido para Computers and Geosciences. Nesse artigo aplicamos várias técnicas de texturização dos horizontes, utilizando alguns atributos, como o atributo de curvatura proposto em Martins et al. (20). Também extraímos características do dado sísmico de entrada utilizando Análise de Componentes Principais como descrito em Partridge (22). O artigo está disponível no Anexo D.

4

Conclusões

Nessa tese foram apresentados o processo de desenvolvimento e os resultados alcançados por uma estratégia baseada em algoritmos de agrupamento de dados destinada ao rastreamento automático de horizontes sísmicos e à produção de atributos de visualização destinados a enfatizar discontinuidades focado principalmente na identificação de falhas.

As estratégias apresentadas são desenvolvidas a partir do pré-processamento dos dados de entrada, e baseadas em dois conceitos: utilização de dados de entrada que contenham implicitamente a caracterização do problema e a utilização de métodos de extração de *features* buscando melhorar essa caracterização. Assim, busca-se uma melhor auto-organização dos dados no que diz respeito à sua semelhança. Os métodos posteriores usam os dados resultantes dessa etapa de agrupamento de dados.

No que diz respeito aos processos de mapeamento de horizontes sísmicos, os processos de agrupamento e posterior auto-organização passam a viabilizar que a identificação dos voxels de um determinado horizonte possa ser realizada através de critérios de medição que não se baseiam apenas em informação local. Tais informações tendem a ser implicitamente disponibilizadas pela distribuição das amostras de entrada nos grupos, ao longo do espaço amostral. Isso também permite que os voxels dos horizontes sejam encontrados utilizando-se de sementes que podem se manter fixas durante boa parte do processo de mapeamento, minimizando a probabilidade de que hajam erros de mapeamento. Os resultados alcançados apresentados ao longo do texto demonstraram a competitividade do método, mesmo quando aplicado a volumes sísmicos caracterizados pelo alto nível de ruído presente nos dados.

Em relação ao processo de criação do atributo de discontinuidades, os resultados alcançados levam a crer que a estratégia merece um estudo mais aprofundado.

De maneira geral, melhorias podem ser pensadas, tais como a substituição de PCA por processos de extração de *features* não lineares, tais como de mecanismos de *Deep Learning*. Também devemos investigar outras aplicações

para os dados resultantes do agrupamento, tais como a produção de volumes de mergulho, por exemplo. Enfim, existe uma vasta quantidade de testes e melhorias a serem realizados.

A

Anexo A

Surface Mapping Using Auto-Adaptable Similarity Measures.

Resumo expandido apresentado na conferência da European Association of Geoscientists and Engineers (EAGE), 2013.

We-06-04

Surface Mapping Using Auto-adaptable Similarity Measures

A.M. Figueiredo* (Tecgraf/PUC-Rio), P.M. Silva (Tecgraf/PUC-Rio) & M. Gattass (Tecgraf/PUC-Rio)

SUMMARY

In this expanded abstract, we introduce a new method that finds the global structure of seismic events and allows the automatic mapping of sub-surface data. Seismic reflection data is transformed from the amplitude space into a multi-dimensional amplitude space. In order to minimize the interference of noise and uncertainties naturally present in the data, we apply a clustering procedure. After this step each voxel is represented by its corresponding cluster label. Based on these clusters, we compute a similarity function optimized for each particular dataset. This similarity function is non-local and auto-adaptable. The sub-surface mapping is provided by this function.

The experimental results indicate the efficiency of the proposed method and illustrate its advantages.

Introduction

Automated extraction of subsurface features is a valuable tool in the interpretation of seismic data. The features can be horizons, faults or sub-volumes. A horizon is characterized by seismic reflection properties in a depositional environment and can be represented as a three-dimensional surface between rock layers (Faraklioti and Petrou, 2004).

There are many methods to extract horizons. Automated horizon-mapping methods generally use seed-based auto-tracking and extract horizons by correlating the local amplitude between neighboring traces. Li, Ma and Du (2012) present a method that identifies horizons using a combination of horizontal derivative and mathematical morphology. Bakke, Gramstad and Sønneland (2012) describe a method that examines the seismic data with the same approach used in string search, where misalignments are tolerated. Yu, Kelley and Mardanova (2011) propose an algorithm that combines pick and trace selection to obtain horizon surfaces. Recently, global approaches have been proposed to compute geological models directly from the seismic data without picking. Among them, the Horizon Cube (De Groot, Huck, De Bruin, Hemstra, and Bedford, 2010) uses local dip-azimuth information calculated from the input seismic cube, filtering the data to improve dip information. Hoyes and Cheret (2011) present a review summarizing global interpretation methods for 3D horizon mapping.

As far as we know, the previous works use predefined similarity functions to find neighbors of a single horizon. This work presents a new method where the similarity is adaptable and takes the inherent variability of the data into account. This similarity is obtained by means of a clustering procedure based on the Growing Neural Gas algorithm (GNG) (Fritzke, 1995). The outline of the proposed method is as follows: (a) we first define a vector of vertical neighboring amplitudes for each voxel in the volume; (b) we use these n -dimensional vectors and the GNG algorithm to create groups (clusters) that represent similar vertical amplitude variations, and, with these clusters, we create a new volume where each voxel amplitude is replaced by its corresponding cluster label; (c) we then define a function that measures similarity between two arbitrary clusters; and finally (d) we show how the output volume can be used by a procedure that is capable of mapping all possible horizons.

Dataset of Samples and the Clustering Procedure

The lateral continuity of sedimentary layers is represented in seismic data by the relatively small amplitude variations among voxels of a same horizon over the lateral neighborhood. Vertical windows contextualize the voxels against their vertical neighbors, creating a signature that characterizes horizon portions over these lateral areas. This representation is commonly used in the field of Seismic Facies Analysis (Marroquín, Brault and Hart, 2009), and it is the basis of our methodology. We firstly create a dataset \mathbf{S} of samples, where each amplitude value, a , in the seismic volume is seen as a vector of vertical neighbor values, i.e., the vector is composed of a in the central position and its immediate neighbors above and below in the seismic trace. Formally, for each i -th voxel of trace t , we compute an element of \mathbf{S} by:

$$\mathbf{s}_{ti} = \{a_{t(i-k)}, \dots, a_{t(i-1)}, a_{ti}, a_{t(i+1)}, \dots, a_{t(i+k)}\}, \quad \mathbf{s}_{ti} \in \mathbb{R}^{2k+1} \quad (1)$$

After the creation of this dataset, we use a specific clustering algorithm, known as Growing Neural Gas (Fritzke, 1995), to divide \mathbf{S} into p distinct groups of samples. GNG creates a self-organizing map of clusters whose structure is constructed spatially reflecting the sample distribution of the input dataset. Ideally p should be close to the number of layers contained in the data, and we use the number of amplitude peaks of the central trace of the volume to estimate its value. After the clustering procedure, each cluster's position represents the subset of samples from the dataset that was classified into this cluster.

Creating the Output Volume

For each of the p clusters of samples we attribute a numeric label varying from 0 to $(p-1)$ and use the labels to create an output volume where the value of each voxel is replaced by the label representing the cluster where its corresponding sample vector is classified. Despite the lack of lateral information stored with the samples, neighboring horizon voxels represented by their corresponding vertical neighbors' samples will share similarities with respect to their neighboring amplitudes, and probabilistically tend to be located in a same cluster, or at least in closely positioned clusters. Consequently, in the output volume, fault regions as well as horizons will be better defined, as illustrated in Figure 1.

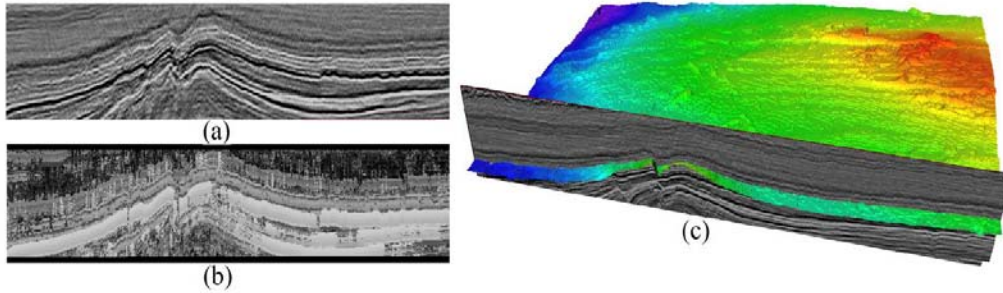


Figure 1: a slice of seismic data (a) and its counterpart in the output volume (b). In (c) an example of the same slice in contrast with a mapped horizon.

Similarity Measurement in the Output Volume

The similarity measurement of two given clusters is in the core of the proposed horizon-mapping algorithm. The use of Euclidean distances between two clusters does not take the local sample density into account. Usually this density varies significantly among the clusters. Clusters located in high-density regions are relatively closer to their neighbors than those located in regions where sample density is lower. Thus, in a given dataset, the meaning of the Euclidean distance is relative. To measure similarity, we use the Mean Square Error (MSE) to scale the distances among two arbitrary pairs. We define the function D_{xy} between two arbitrary cluster positions, \mathbf{c}_x and \mathbf{c}_y , simply as:

$$D_{xy} = \frac{1}{\text{MSE}_x} \|\mathbf{c}_y - \mathbf{c}_x\| \quad (2)$$

where MSE_x is the Mean Square Error of the cluster \mathbf{C}_x .

Samples in the Output Volume and Sequence Alignment

As we do not know exactly the number p of clusters to be created, in the output volume there are cases when contiguous samples are classified into a same cluster. There are obvious cases, for instance around the amplitude peaks (i.e., immediately above and below). When looking for lateral neighbors, using a single cluster label causes ambiguity. To overcome this drawback, we transform the output volume again to represent for each voxel not only its label id_{ii} , but the vertical neighborhood of labels \mathbf{f}_{ii} . This neighborhood has the same structure as \mathbf{s}_{ii} in the original volume. That is:

$$\mathbf{f}_{ii} = \{id_{t(i-q)}, \dots, id_{t(i-1)}, id_{ii}, id_{t(i+1)}, \dots, id_{t(i+q)}\}, \mathbf{f}_{ii} \in \mathcal{R}^{2q+1} \quad (3)$$

Given two samples \mathbf{f}_{i1} and \mathbf{f}_{i2} , to compute the similarity level between them we propose an approach similar to the one used in the Seismic DNA method (Bakke, Gramstad and Sønneland, 2012), which is generally referred to as the sequence alignment problem (Kleinberg and Tardos, 2005). It consists of, upon receiving two strings, finding what is called a “matching”, that is, aligning the strings where there are no crossing pairs. In our case the p cluster labels are the symbols used to construct the strings, \mathbf{f}_{i1} and \mathbf{f}_{i2} . When the same identifier id occurs in a given position in the strings, this position has no alignment cost. If the ids differ, the algorithm adds the cost of aligning the two identifiers. This cost is computed by Equation 2. Note that two contiguous symbols of string \mathbf{f}_{i2} may be paired up with a same symbol contained in \mathbf{f}_{i1} , i.e., the mapping between symbols may be one to many and

many to one. This happens due to several geological reasons: one sample could have been taken in a dip, which might have stretched its values; the horizons do not have the same thickness along the seismic traces, etc.

The Horizon-Mapping Procedure

Automated methods generally extract horizons solving a series of local problems, using a small portion of the data at a time. This strategy leads to poor mapping quality, as any errors can propagate during the region-growing procedure.

The procedure that maps arbitrary horizons avoids this drawback by using fixed seeds, and can be described as follows: (a) the user picks, from an arbitrary seismic trace t , a voxel that belongs to the horizon to be mapped. This voxel, seed voxel s , has a corresponding sample in the output volume, \mathbf{f}_{ts} . (b) The user defines a similarity threshold T . (c) The sequence-alignment procedure presented above uses \mathbf{f}_{ts} to find the most similar sample in each of its neighboring traces. If the best candidate is within the threshold value T , this sample is added to the set of samples that compose the horizon. (d) The procedure continues, testing neighbors of the already discovered voxels using the same seed \mathbf{f}_{ts} . (e) After all the voxels similar to \mathbf{f}_{ts} have been discovered, a new sample \mathbf{f}'_{ts} is chosen randomly from already discovered samples, and the process continues up to the point when \mathbf{f}'_{ts} does not discover any new samples.

Results

The proposed methodology was implemented in language C++ and tested with many different seismic datasets. Despite the presence of different seismic events, which usually makes sub-surface mapping a harder task, the proposed method is able to properly detect horizons, as shown in Figures 2 and 3. Figure 2 shows eleven horizons mapped throughout the whole data. Figure 3 presents nine horizons mapped in a specific region of the data. As we can observe, the method yields satisfactory results regarding the existence of faults and discontinuities. In fact, our method is able to interrupt the mapping procedure in the presence of discontinuities. Due to space limitations we present the results obtained with the 3D dataset known as F3 Block, downloaded from the Opendtect website.

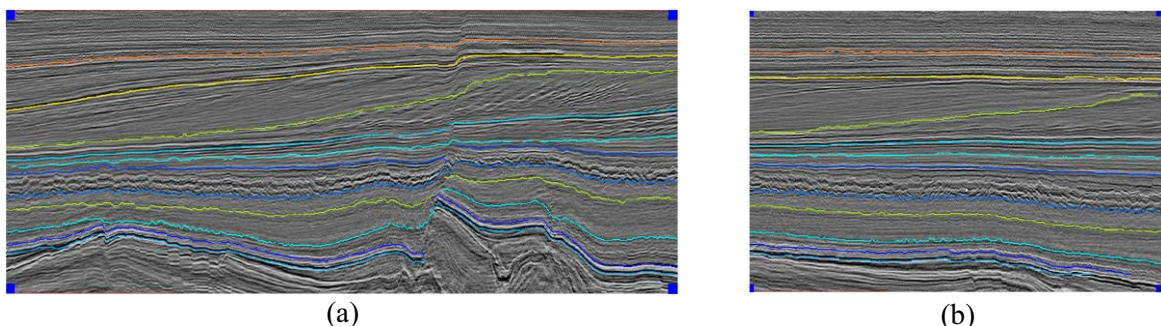


Figure 2: Results of proposed method. The 11 mapped horizons are in contrast with two slices, (a) inline number 265, and (b) crossline number 680. The first 300 milliseconds of the original data were not included in the picture. Each color indicates a different horizon. Note that the horizons were correctly interrupted by the faults present in the seismic data.

In these results, we used 19 amplitude values for the training window ($k = 9$, Eq. 1) and 11 labels to construct the strings ($q = 5$, Eq. 3). For this input data, we defined the number of clusters as 240. The processing time for the clustering procedure was around 5 minutes on an Intel Core i7-3960X.

Conclusions

In this paper, we proposed a new solution to the problem of automatically mapping horizons in 3D seismic data. Our method is based on a clustering procedure that uses the Growing Neural Gas

algorithm and causes the lateral similarity measures to be implicitly distributed over the clusters of a GNG graph. The method uses non-local searching to look for neighboring samples, avoiding the drawbacks of local procedures, and is capable of mapping even horizons that are severely interrupted by seismic faults or other discontinuities. We presented experiments indicating its efficiency and illustrated its output. The experimental results reported here suggest that clustering-based procedures are a good alternative solution for the task of automatically mapping seismic horizons.

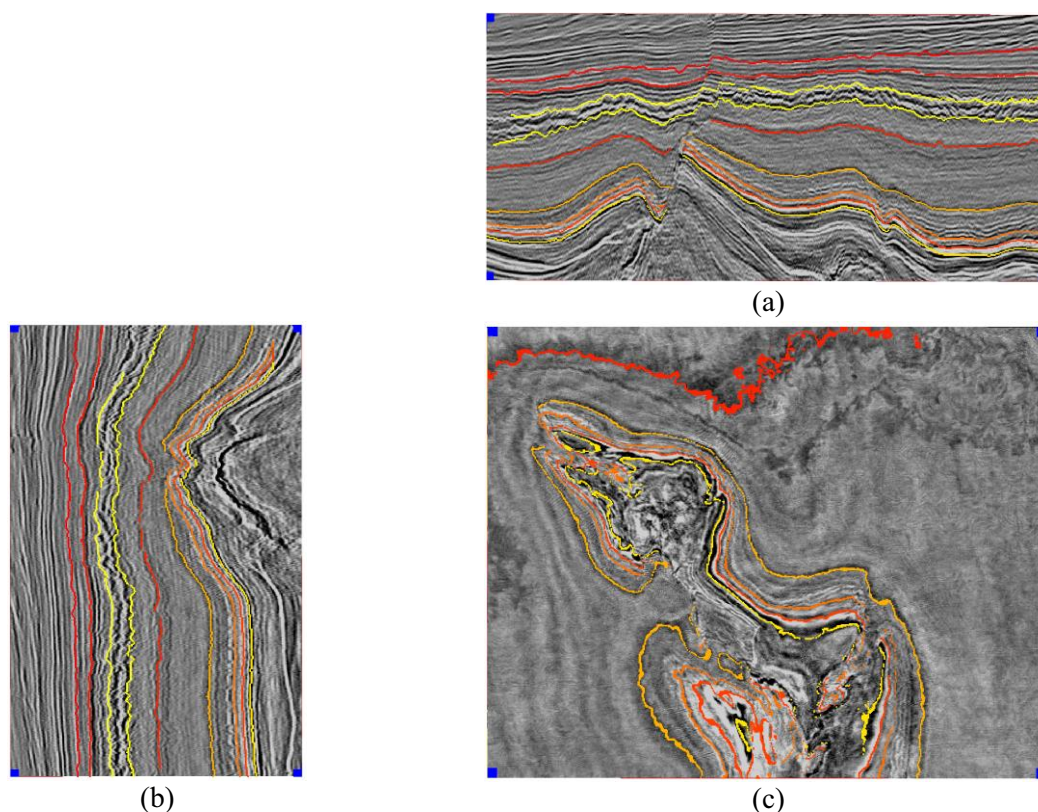


Figure 3: Results achieved mapping 9 different horizons, showed in contrast with three slices, (a) inline number 299, (b) crossline number 801, and (c) a time slice placed at 1458 milliseconds.

References

- Bakke, J., Gramstad, O. and Sønneland, L., 2012. Seismic DNA - A Novel Seismic Feature Extraction Method Using Non-local and Multi-attribute Sets, 74th EAGE Conference & Exhibition incorporating SPE EUROPEC.
- De Groot, P., Huck, A., De Bruin, G., Hemstra, N., and Bedford J., 2010. The horizon cube: A step change in seismic interpretation: The Leading Edge, v. 29/9, p. 1048-1055.
- Faraklioti, M. and Petrou, M., 2004. Horizon picking in 3D seismic data volumes. Machine Vision and Applications, 15(4):216–219.
- Fritzke, B., 1995. A growing neural gas network learns topologies. Advances in Neural Information Processing Systems 7, MIT Press, Cambridge MA, pp. 625-632.
- Hoyes, J. and Cheret, T., 2011. A review of global interpretation methods for automated 3D horizon picking, Leading Edge, vol. 30, no. 1, pp. 38–47.
- Kleinberg, J., Tardos, E., 2005. Algorithm Design, Wesley Publishing Co.
- Li, L., Ma, G., and Du, X., 2012. New Method of Horizon Recognition in Seismic Data, IEEE Geoscience And Remote Sensing Letters, Vol. 9, No. 6.
- Marroquin, I. D., Brault J. J., and Hart B. S., 2009. A visual data-mining methodology for seismic facies analysis: Part 1 - Testing and comparison with other unsupervised clustering methods. Geophysics, 74, no. 1, P1–P11.
- Yu, Y., Kelley, C. and Mardanova, I., 2011. Automatic horizon picking in 3D seismic data using optical filters and minimum spanning tree, SEG Expanded Abstracts.

B

Anexo B

Minimal Similarity Accumulation Attribute for Fault Enhancement.

Resumo expandido apresentado na conferência da Society of Exploration Geophysicists (SEG), 2014.

Minimal Similarity Accumulation Attribute for Fault Enhancement

A.M. Figueiredo*, G.M. Faustino, J.P. Peçanha, P.M. Silva & M. Gattass, Tecgraf/PUC-Rio

SUMMARY

In this work we present a new method for seismic fault enhancement on volumetric grids called Minimal Similarity Accumulation (MSA). The seismic reflection data is transformed into a multi-dimensional amplitude space and a clustering procedure is applied in order to minimize the noise interference. Thereafter, we use an algorithm to compute the MSA for each voxel in the volume. This is done by applying an auto-adaptable function that uses the voxel neighboring information. The MSA measurement globally highlights the fault regions presented in the original volume. To validate the proposed method we use the volume of the Netherlands offshore F3 block downloaded from the Open Seismic Repository. In order to assess the proposed method and illustrate its advantages, a set of vertical 2D slices of the seismic data are presented providing a comparison between our results and images manually interpreted by a geologist. Finally, we conclude that the proposed method is sufficiently accurate.

INTRODUCTION

The extraction of sub-surface features, such as horizons, faults and channels is a key subject in seismic data analysis. The properly mapping of these properties has a great importance in oil and gas exploration. In particular, fault detection plays an important role in hydrocarbon reservoir modelling. The discontinuities caused in the rock layers by faults should be identified since it introduces relevant changes in the geological system.

The seismic data analysis has been traditionally done by manual interpretation processes. This time-consuming task requires well-trained interpreters so that the extracted information can be used in further steps. Automatic fault enhancement is, then, a valuable tool in the seismic data interpretation process. This very useful approach highlights regions hard to identify visually in the sub-volume and reduces the time needed for fault mapping.

There are several works in the literature focused on fault attribute enhancement. Bajorich and Farmer (1995) presented an attribute which computes coherence coefficients from seismic amplitudes of adjacent traces. This work is known as the first generation of coherence attributes, and is followed by Marfurt *et al.* (1998), which achieved results with a better vertical resolution. In a different approach, the method presented by Gersztenkorn and Marfurt (1999) highlights coherence attributes by computing the eigenvalues of the covariance matrix. This type of attribute is very useful for faults and fractures detection and reaches the results by measuring the lateral changes in waveform (Chopra and Marfurt, 2007).

Many other attributes have been proposed with the aim of en-

hancing faults in seismic data. We may quote attributes based on dip-magnitude and dip-azimuth (Bendar, 1998), volumetric curvature computations (Martins *et al.*, 2012), variance cube (Pepper and Van Bommel, 2000), among others. Several other studies are supported by image processing theory to identify faults and channels. In this approach, edge detection procedures and digital filters are commonly used (Aqrawi and Boe, 2011, Jing *et al.*, 2007). In a recent work, Pampanelli *et al.* (2013) took advantage of the strong mathematical background inherent to these approaches and used the first-order directional derivative to enhance discontinuities along horizons.

In this extended abstract, we present a new attribute for global fault-enhancement that we call Minimal Similarity Accumulation (MSA). To find this attribute, we first perform a clustering procedure considering vertical windows of samples. This is done aiming to reduce noise interference in the input-data. After that, each voxel is represented by its corresponding cluster. An auto-adaptable function is used to measure the similarity in the clustering space. The final attribute is given by the sum of the minimal similarity between a voxel and its neighbours. To validate the proposed method we use the volume of the Netherlands offshore F3 block downloaded from the Open Seismic Repository. The results shown that this new attribute is very efficient for highlight fault regions on seismic data.

SEISMIC FAULT ENHANCEMENT

In this section, we describe the proposed method to compute a seismic attribute that enhances faults in a volumetric grid. We use the seismic amplitude volume as the input data. Our procedure is composed of three steps: (1) sample dataset creation; (2) volume clustering; and (3) minimal similarity accumulation. An overview of the proposed method is depicted in Figure 1.

Sample Dataset Creation

It is well known that the lateral continuity of sedimentary layers is represented in seismic data by the relatively small amplitude variations among voxels of a same horizon over the lateral neighborhood. Vertical windows contextualize the voxels against their vertical neighbors, creating a signature that characterizes horizon portions over these lateral areas (Figueiredo *et al.* 2013). For that reason, we create a representation of the input-volume as follows. Let $V = \{v_1, \dots, v_m\}$, where $v_i \in \mathbb{R}$ contains the seismic amplitude values, and $|V| = m$ is the total number of voxels in the input-volume. In this step, we create the dataset $A = \{a_1, \dots, a_m\}$, where each element $a_i \in A$ is a vector of vertical neighbor values taken from V . Thus, $a_i = \{v_{i-k}, \dots, v_i, \dots, v_{i+k}\} \in \mathbb{R}^{2k+1}$ represents the i -th voxel as illustrated in Figure 2. Here, $k \in \mathbb{N}$ is a constant value that indicates the size of the vertical half-window. This representation is commonly used in the field of Seismic Facies Analysis (Marroquín, Brault and Hart, 2009).

Minimal Similarity Accumulation Attribute for Fault Enhancement

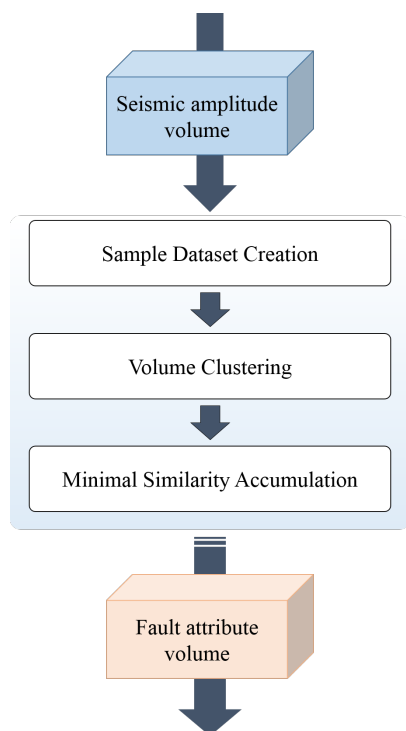


Figure 1: Overview of the proposed method: each step is denoted by a box. The cubes in blue and orange represent the input and output data, respectively.

Volume Clustering

The main goal of this step is to partition A into p clusters containing samples that share similar features. Here we use a specific clustering algorithm, known as Growing Neural Gas (GNG) (Fritzke, 1995). Ideally, p should be close to the number of layers in the seismic data. As Figueiredo *et al.* (2013), we use the number of amplitude peaks of the central trace of the input-volume to estimate p . Let $C = \{c_1, \dots, c_p\}$, where $c_j \subset A$, be the clusters set generated by the GNG algorithm. Its important to mention that the cluster position $\vec{c}_j \in \mathbb{R}^{2k+1}$, provided by GNG, is spatially located as to properly represent its corresponding set of samples. As a consequence of the clustering procedure, all samples in a same cluster are represented

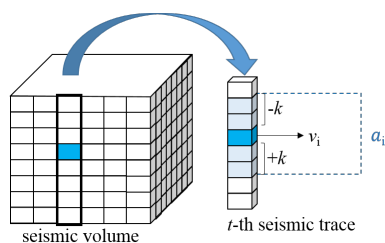


Figure 2: Input volume representation: each voxel v_i is represented by a vertical vector of size $2k+1$.

by its cluster position.

Despite the lack of lateral information stored with the samples, neighboring horizon voxels are represented by their corresponding vertical neighbors samples. These samples share similarities with respect to their neighboring amplitudes, and probabilistically tend to be located in a same cluster, or at least in closely positioned clusters. Therefore, fault regions as well as horizons are better defined in the clusters set.

Minimal Similarity Accumulation

In this step, we compute the proposed fault attribute. The similarity measurement is the key point of the proposed algorithm. Clusters located in regions of high samples' density are closer than those located in lower density regions. The Euclidean distances does not take this density into account. To overcome this issue, we use the Root Mean Square Error (RMSE) to scale the Euclidean distance between arbitrary pairs of clusters. As Figueiredo *et al.* (2013), the similarity measure described above is auto-adaptable, and takes into account the inherent variability of the data. The similarity between two clusters is described in the following equation:

$$S(c_x, c_y) = \frac{1}{RMSE_{c_x}} \|\vec{c}_x - \vec{c}_y\|; 1 \leq x, y \leq p, \quad (1)$$

where, $\|\vec{c}_x - \vec{c}_y\|$ is the Euclidean distance between vectors \vec{c}_x and \vec{c}_y , and $RMSE_{c_x}$ denotes the Root Mean Square Error of the cluster c_x .

The use of a single cluster may not be sufficiently accurate to compute the similarity between voxels of neighbouring traces. This occurs when contiguous samples in a trace are classified into the same cluster. We are able to solve this drawback using more than one cluster. Moreover, it emphasizes the differences and similarities between the elements. For this reason, we apply the same representation proposed in Step 1. That is, each voxel $v_i \in V$ is represented by a vertical neighbourhood of size ε as $\vec{\delta}_i = \{c_{i-\varepsilon}, \dots, c_i, \dots, c_{i+\varepsilon}\}$.

We can estimate the similarity between two given vector $\vec{\delta}_i$ and $\vec{\delta}_j$ based on this representation. This is done through a particular implementation of the sequence alignment problem (Kleinberg and Tardos, 2005). The algorithm is straightforward: it receives two vectors $(\vec{\delta}_i, \vec{\delta}_j)$ and finds a matching between this sets, i.e., an alignment where there are no crossing pairs. If the same cluster occurs in two given positions, it has no alignment cost, otherwise, the algorithm adds the cost of aligning this clusters applying the Equation 1.

To compute the final attribute value $MSA(v_i)$ for $v_i \in V$, we consider N_i as a window with lateral half-size r , centred on v_i . Suppose that v_i is located in the plane z , as shown in Figure 3. For each $v_j \in N_i$ we apply the sequence alignment procedure to find the most similar sample between $\vec{\delta}_i$ and $\{\vec{\delta}_{j(z-1)}, \vec{\delta}_j, \vec{\delta}_{j(z+1)}\}$, i.e., the one with minimal similarity value considering a small vertical neighbourhood.

The final value of our attribute is given by the sum of these

Minimal Similarity Accumulation Attribute for Fault Enhancement

minimum similarities, as presented in Equation 2.

$$MSA(v_i) = \sum_{\forall v_j \in N_i} \min \left\{ \begin{array}{l} \bar{S}(\vec{\delta}_i, \vec{\delta}_{j(z-1)}), \\ \bar{S}(\vec{\delta}_i, \vec{\delta}_{j(z)}), \\ \bar{S}(\vec{\delta}_i, \vec{\delta}_{j(z+1)}) \end{array} \right\} \quad (2)$$

Where \bar{S} is the similarity estimation given by the sequence alignment procedure aforementioned.

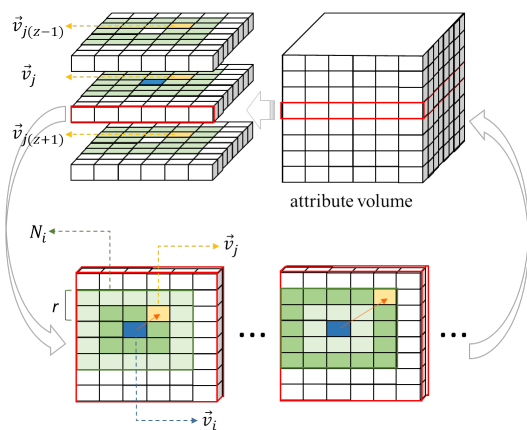


Figure 3: Voxel neighborhood: the voxel v_i is located at plane z . The $v_j \in N_i$ can be positioned on planes $z-1$, z , or $z+1$.

RESULTS

In this section we present some results and discussions. In order to evaluate the proposed method, use the volume of the Netherlands offshore F3 block downloaded from the Open Seismic Repository. In addition, a manual labeled set of faults surfaces have been identified by a geologist. An example set of results showing a vertical 2D slice of the seismic data are presented providing a comparison between the automatically generated results produced using the proposed method, with manually labeled images. These results are shown in Figure 4. As can be seen from the comparative results there is a broad agreement between the placement of many of the faults from the automatic detection and manually labelled methods.

Figure 7a shows an interesting area of the 3D seismic volume located around the main seismic fault identified in the data. A sub-volume is defined from 300 to 1840 milliseconds. The obtained result for this region is presented in Figure 7b. As can be noted, areas of high discontinuities are correctly enhanced. Figure 5 shows a crossline taken from the seismic data in contrast with a time slice, located at 1312 milliseconds. As can be noted, the fault region is acceptably highlighted on the fault attribute. A 3D view of a sub-volume of Netherlands offshore F3 block with the obtained attribute in time slice at 1736 milliseconds is presented in Figure 6.

The results presented in this section suggest that the proposed method can be efficiently used to improve the visualization of fault systems.

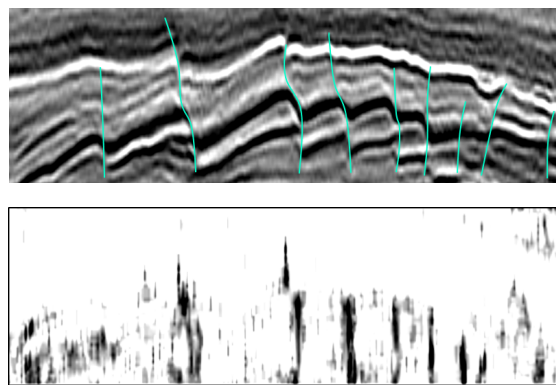


Figure 4: Obtained result for the inline 418: (a) estimated faults region; and (b) the equivalent manually labelled faults.

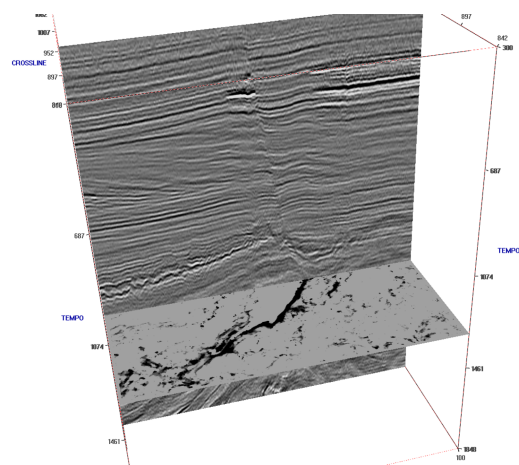


Figure 5: The proposed attribute in the time slice at 1312 milliseconds. The vertical slice shown here corresponds to the crossline 964. As we can observe, our attribute correctly highlights the fault region.

CONCLUSIONS AND FUTURE WORKS

In this extended abstract, we present a new attribute for global fault-enhancement that we call Minimal Similarity Accumulation (MSA). Our method creates a multi-dimensional representation for each voxel. These representations are used in a clustering procedure in order to globally organize the voxels of the seismic data. As a result the noise interference in the input-data is reduced. The clusters are then used to compute the minimal similarity between a sample and its neighbourhood.

We presented results illustrating the effectiveness of the proposed methodology. The experimental results suggest that the methodology (MSA) is good alternative solution for the task of seismic fault enhancement on 3D seismic data. Future works involve extending the described methodology to fully automatic fault surface extraction.

Minimal Similarity Accumulation Attribute for Fault Enhancement

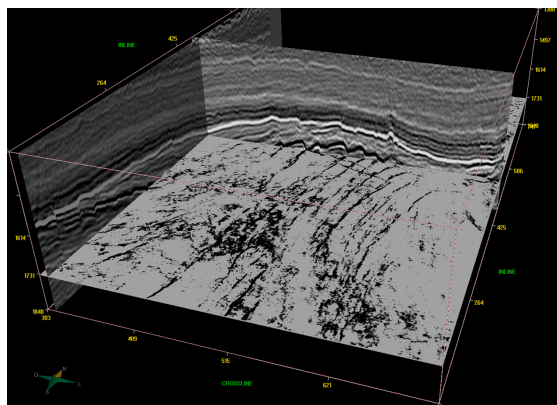
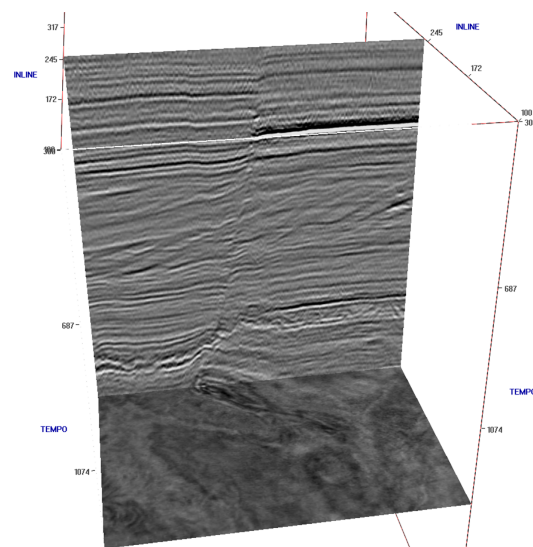


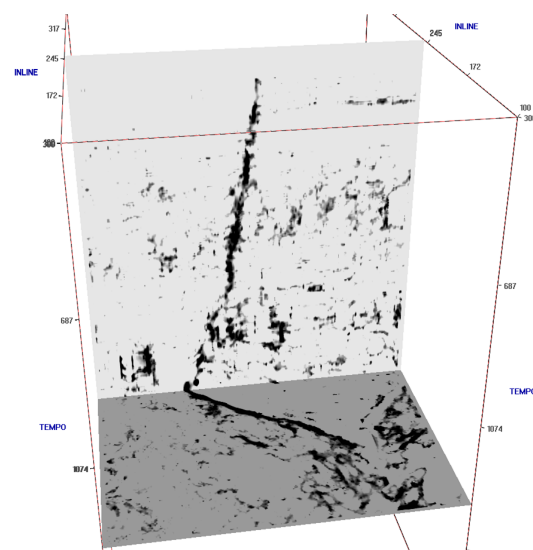
Figure 6: Fault region highlighted by our attribute in time slice at 1736 milliseconds.

ACKNOWLEDGMENTS

The authors would like to thank the computational geophysics team of Tecgraf Institute for discussions about researcher hypothesis and Petrobras for technical contributions. Aurlio M. Figueiredo would like to thank CNPq (National Research Council of Brazil) for scholarship support.



(a) Input seismic data



(b) Proposed attribute for fault enhancement

Figure 7: Results of the proposed method as for Netherlands offshore F3 block. This figure show the inline 252 and time slice at 1312 milliseconds.

<http://dx.doi.org/10.1190/segam2014-1615.1>

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

- Agrawi, A. A., and T. H. Boe, 2011, Improved fault segmentation using a dip guided and modified 3D Sobel filter: 81st Annual International Meeting, SEG, Expanded Abstracts, 999–1003.
- Bahorich, M., and S. Farmer, 1995, 3D seismic discontinuity for faults and stratigraphic features: The coherence cube: *The Leading Edge*, **14**, 1053–1058, <http://dx.doi.org/10.1190/1.1437077>.
- Bednar, J. B., 1998, Least squares dip and coherency attributes: *The Leading Edge*, **17**, 775–778, <http://dx.doi.org/10.1190/1.1438051>.
- Chopra, S., and K. Marfurt, 2007, Seismic attributes for fault/fracture characterization: 77th Annual International Meeting, SEG, Expanded Abstracts, 1520–1524.
- Figueiredo, A. M., P. M. Silva, and M. Gattass, 2013, Surface mapping using auto-adaptable similarity measures: 75th Conference & Exhibition, EAGE, Extended Abstracts, We 06 04.
- Fritzsche, B., 1995, A growing neural gas network learns topologies, in G. Tesauero, D. S. Touretzky, and T. K. Leen, eds., *Advances in neural information processing systems*: MIT Press, 625–632.
- Gersztenkorn, A., and K. Marfurt, 1999, Eigenstructure-based coherence computations as an aid to 3D structural and stratigraphic mapping: *Geophysics*, **64**, 1468–1479, <http://dx.doi.org/10.1190/1.1444651>.
- Jing, Z., Z. Yanqing, C. Zhigang, and L. Jianhua, 2007, Detecting boundary of salt dome in seismic data with edge-detection technique: 77th Annual International Meeting, SEG, Expanded Abstracts, 1392–1396.
- Kleinberg, J., and E. Tardos, 2005, *Algorithm design*: Wesley Publishing Co.
- Marfurt, K., R. Kirlin, S. Farmer, and M. Bahorich, 1998, 3-D seismic attributes using a semblance-based coherency algorithm: *Geophysics*, **63**, 1150–1165, <http://dx.doi.org/10.1190/1.1444415>.
- Marroquín, I. D., J. J. Brault, and B. S. Hart, 2009, A visual data-mining methodology for seismic facies analysis: Part 1 — Testing and comparison with other unsupervised clustering methods: *Geophysics*, **74**, no. 1, P1–P11, <http://dx.doi.org/10.1190/1.3046455>.
- Martins, L., P. M. Silva, and M. Gattass, 2012, A method to estimate volumetric curvature attributes in 3D seismic data: 74th Conference & Exhibition, EAGE, Extended Abstracts, E023.
- Pampanelli, P., P. M. Silva, and M. Gattass, 2013, A new volumetric fault attribute based on first order directional derivatives: 13th International Congress of the Brazilian Geophysical Society & EXPOGEF, Expanded Abstracts, 1621–1625.
- Van Bemmelen, P., and R. E. F. Pepper, 2000, Seismic signal processing method and apparatus for generating a cube of variance values: U. S. patent 6,151,555.

C

Anexo C

A Seismic Facies Analysis Approach to Map 3D Seismic Horizons.

Resumo expandido apresentado na conferência da Society of Exploration Geophysicists (SEG), 2014.

A Seismic Facies Analysis Approach to Map 3D Seismic Horizons

A.M. Figueiredo*, Tecgraf/PUC-Rio; F. B. Silva, Petrobras; P.M. Silva, Tecgraf/PUC-Rio; Ruy L. Milidiú, PUC-Rio; M. Gattass, Tecgraf/PUC-Rio

Summary

We present a clustering based methodology used to process 3D seismic data and map automatically a dense amount of horizons. It firstly transforms seismic reflection data from the amplitude space into multi-dimensional amplitude spaces. Then it uses different instances of a specific clustering procedure to create various distinct datasets of clusters. Each voxel is represented by a set of cluster labels, minimizing the interference of noise and uncertainties naturally present in the seismic data. The method finds the global structure of lateral amplitude reflections. It is capable of mapping a great portion of the seismic horizons present in the data, even when severely interrupted by seismic faults, salt domes and other discontinuities.

Introduction

Seismic Horizons are geologically significant surfaces that can be extracted from 3D seismic data. Horizons refer to those seismic reflectors representing stratal surfaces of constant geologic time (Wu and Hale, 2013). Generally, local automated horizon-mapping methods use seed-based, auto-tracking, and extract horizons by correlating the local amplitude between neighboring traces. Yu and Kelley (2011) combine pick and trace selection to obtain horizon surfaces. Li et al. (2012) identify horizons using a combination of horizontal derivative and mathematical morphology. Recently, global approaches have been proposed to compute global geological models from the seismic data. Hoyes and Cheret (2011) present a review summarizing global interpretation methods for 3D horizon mapping. Wu and Hale (2013) present a horizon-extraction method that uses seismic normal vectors to extract globally optimized horizons. Figueiredo et al. (2013) unveil a global methodology that transforms seismic data from the amplitude space into a multi-dimensional space, using clusters from vertical windows of samples to map horizons. Their method organizes the seismic data according to a global and auto-adaptable criterion, and maps horizons using fixed seeds. This avoids a major weakness of many methods, i.e., solve huge series of local problems, leading to poor mapping quality.

This work presents an auto-adaptable method to map horizons globally distributed along the seismic data, using a methodology that resembles the field of Seismic Facies Analysis (Marroquín et al., 2009), as it shares some similarities with the methodology presented by Figueiredo et al. (2013). Our routine aims to partition the data into groups of similar seismic trace shapes providing a natural

clustering structure. Patterns in a given cluster resemble each other more than in other clusters. In order to find seismic horizons, we vary the size of the analysis window and create different representations for each amplitude sample. We then classify each representation through a corresponding clustering procedure, based on the Growing Neural Gas algorithm (GNG) (Fritzke, 1995). The different voxels' representations and the information provided by the clustering procedures are used to develop a similarity measure between voxels that dismiss any kind of parameters or thresholds given by the user. Using the clusters' information, the methodology finds a list of good seeds from what horizons disposed along all the seismic data can be adequately mapped in a totally automatic manner.

The outline of the proposed method is as follows: (a) We first define the type of amplitude samples to process. Then we create datasets composed by vectors of vertical neighboring amplitudes to represent each voxel of this type of interest. (b) We use these datasets and the GNG algorithm to create groups (clusters) that represent similar vertical amplitude variations. (c), We create, using these clusters, a new volume where each amplitude sample is replaced by its corresponding vector of cluster labels, and post process this Labels' Volume removing bad cluster classifications. (d) We present a similarity measure between two given voxels, based on its corresponding sets of labels. (e) We use the similarity criterion to build a list of high-quality seeds used to map the horizons along all the seismic data and describe our horizon mapping algorithm that uses this set of high-quality seeds to track the surfaces dismissing any kind of external information.

Selecting Type and Datasets of Samples

Our methodology transforms seismic reflection data from the amplitude space into multi-dimensional amplitude spaces. In order to minimize the amount of data to be processed, the method initially receives one amongst four types of voxels that will compose the horizons to be mapped: positive amplitude peaks, negative amplitude peaks, negative to positive zero crossing, or positive to negative zero crossing points. We examine all the volume data and construct a set, referred as the set **M**, containing only the voxels of the desired type. Then we create feature vectors to represent each voxel stored in **M**. In order to contextualize these voxels against its vertical neighbors, we use the same representation commonly used in Seismic Facies Analysis (Marroquín et al., 2009), i.e., by windows of trace shapes that are vectors of vertical neighbors. In order to capture different features, we vary the window

A Seismic Facies Analysis Approach to Map 3D Seismic Horizons

size, creating r representations for each voxel. The window sizes to be used are defined in the set $\mathbf{N}=\{n_0, n_1, \dots, n_{r-1}\}$. Consider a_{ti} as the $(i$ -th) voxel of a trace (t) , ($a_{ti} \in \mathbf{M}$). We compute each of its r corresponding vertical windows by:

$$\mathbf{s}_{tij} = \{a_{t(i-n_j)}, \dots, a_{t(i-1)}, a_{ti}, a_{t(i+1)}, \dots, a_{t(i+n_j)}\}, \quad \mathbf{s}_{tij} \in \mathcal{R}^{2n_j+1} \quad (1)$$

We group these vectors according to its size creating r datasets of vertical windows, that is, the samples created using $n_j \in \mathbf{N}$ are stored on the corresponding dataset \mathbf{D}_j . This process is illustrated in Figure 1.

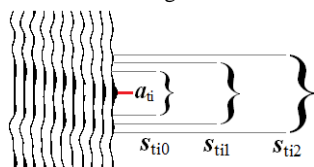


Figure 1: Each sample a_{ti} from the subset to be processed (e.g., negative peaks) is represented by its r vertical windows. Here, $\mathbf{N}=\{n_0, n_1, n_2\}$, with $r=3$, giving rise to the samples \mathbf{s}_{ti0} , \mathbf{s}_{ti1} , and \mathbf{s}_{ti2} , stored into its corresponding datasets, \mathbf{D}_0 , \mathbf{D}_1 and \mathbf{D}_2 .

Clustering the Datasets of Samples

Each dataset \mathbf{D}_j is divided, giving rise to a corresponding set of clusters. The objective is to separate the samples into distinct groups that share similar features without using a priori information to guide the samples classification. The clustering is obtained through the use of a specific clustering algorithm, known as Growing Neural Gas (Fritzke, 1995). We avoid classifying various samples from the same seismic trace into the same group, defining the number of clusters to be slightly greater than the number of layers contained in the data. Despite the lack of lateral information stored with the samples, neighboring layer voxels represented by their corresponding vertical neighbors' samples will share similarities with respect to their neighboring amplitudes, and probabilistically tend to be located in the same cluster, or at least in closely positioned clusters. After the clustering step, each cluster receives a unique numeric label id .

Representing Voxels by Sets of Labels

We use the labels to represent each voxel a_{ti} , ($a_{ti} \in \mathbf{M}$), by its equivalent set of labels, $\mathbf{E}_{ti}=\{id_0, id_1, \dots, id_{r-1}\}$ containing the ids of the clusters where its r corresponding vertical vectors were classified. This is used to create a new volume called as the Labels' Volume. On this new volume the voxels that do not compose the set \mathbf{M} are replaced by a dummy value id_{inv} , indicating that they are invalid to map horizons. The voxels from \mathbf{M} are replaced by its corresponding set of labels \mathbf{E}_{ti} . Then a post processing is used to eliminate bad classifications. Ideally, voxels of

different layers should not be classified into the same cluster. We eliminate these occurrences. If any label occurs in two or more sets of labels from a same seismic trace, the repeated label is removed on both sets.

Measuring Similarity on the Labels' Volume

We then define a function capable of measuring the similarity between two arbitrary voxels using the Labels' Volume. Given two voxels a_1 and a_2 , we define the similarity function $S(a_1, a_2)$, using its corresponding set of labels \mathbf{E}_1 and \mathbf{E}_2 . Considering \mathbf{E}_1 and \mathbf{E}_2 as two sets composed of integer elements, we define $S(a_1, a_2)$ as the cardinality of the intersection between the two sets, where bigger values of S indicate greater similarity:

$$S(a_1, a_2) = |\mathbf{E}_{a1} \cap \mathbf{E}_{a2}|, \quad 0 \leq S(a_1, a_2) \leq r \quad (2)$$

The Horizon Mapping Procedure

Our method uses the clusters' location to map the horizons. The lack of lateral information stored with the vertical samples and the clusters' auto organization along the samples space provide important advantages: (1) The clusters implicitly store information about its corresponding samples' set, and the clusters disposition along the samples space reflects the local density of the vertical samples. Neighboring layer voxels represented by their corresponding vertical neighbors' samples share relative similarities with respect to their neighboring amplitudes, and probabilistically tend to be located in same clusters, or at least in closely positioned clusters. (2) At same time, areas of relative breaks in the lateral continuity are represented into the vertical samples' space by a relative lack of similarity with respect to its neighboring vertical samples and are naturally classified into different clusters.

Before describe the horizon mapping procedure we need to define our concept of immediate neighbor voxels. Considering a_{ti} as the i -th voxel of a trace t , we define the immediate neighbors of a_{ti} as the voxels $a'_{t(i-1)}$, a'_{ti} and $a'_{t(i+1)}$, the $(i-1)$ -th, i -th, and $(i+1)$ -th voxels on the immediate neighbor trace t' .

The seismic horizon mapping procedure constructs a list \mathbf{L} of seed voxels ordered by priority. The priority of each voxel a , ($a \in \mathbf{M}$), is given using the similarity function S , summing up the similarity values between a and the more similar immediate neighbor on the neighboring traces. The horizon-mapping procedure uses fixed seeds, mapping a dense set of seismic horizons. It can be described as follows: (a) The method takes the voxel of greatest priority from \mathbf{L} . We call this seed voxel as s_{ti} , the $(i$ -th) voxel of a trace (t) . It has a corresponding set of labels in the Labels' Volume, \mathbf{E}_{ti} . (b) We use the similarity function S to find the

A Seismic Facies Analysis Approach to Map 3D Seismic Horizons

most similar immediate neighbor voxel in each immediate neighboring trace of s_{ti} . If the similarity value between this best candidate and s_{ti} is greater than one, the candidate is added to the set of samples that compose the horizon. The best candidate is marked as discovered and can not be used as part of any other horizon. Besides that, this best candidate is removed from L . (c) The procedure continues, using S to test the immediate neighbors of the already discovered voxels against the fixed seed s_{ti} . (d) After all the voxels similar to s_{ti} have been discovered, a new sample s'_{ti} is chosen randomly from already discovered voxels, and the process continues up to the point when the new seeds s'_{ti} does not discover any new samples. Once the horizon is completely mapped, the algorithm takes again the voxel of greatest priority from L , and begins the mapping of a new horizon, repeating the steps (a) through (d).

many applications to the methodology. The contour of the horizons around the salt domes can be post processed as to estimate its surface. The information provided by horizons around the seismic faults could be used to estimate its dip and plane direction. In Figure 3, we present fundamental details about the features of the mapped horizons. A time slice is positioned at 3200 milliseconds 3(a), where it is possible to verify the existence of salt domes and faults on the seismic data. As showed in 3(b), the horizon surface effectively follows the signal continuity. Fault zones as well as the contour of radial faults around the salt domes are characterized in the horizon surface. 3(c) shows the seismic horizon in contrast with the time slice. The horizons are mapped with voxel precision. In these examples, all the mapped voxels are negative amplitude peaks. We then apply the volumetric seismic curvature

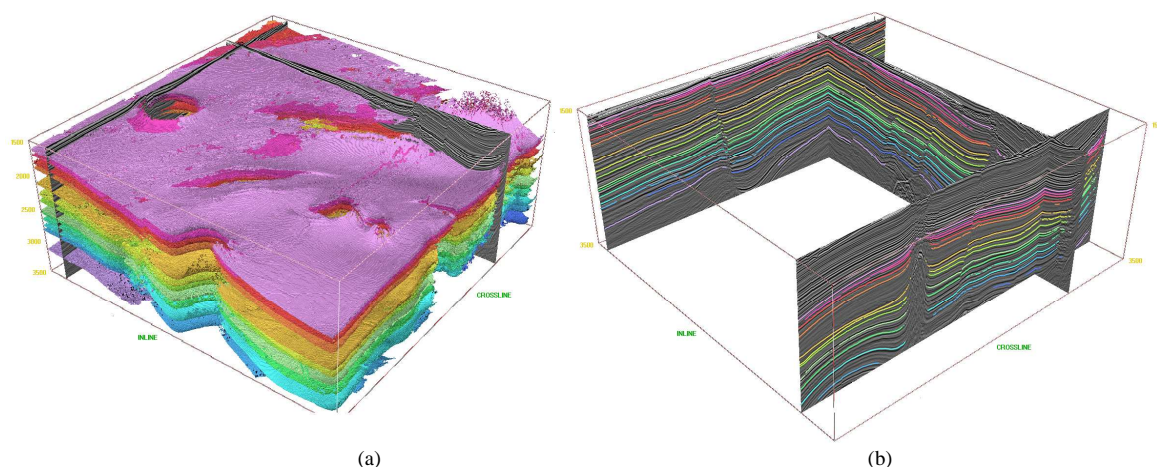


Figure 2: Results of proposed method. The main horizons, obtained mapping negative peaks are showed in (a). Due to space limitations, we only present the horizons that extend along all the seismic data. The same horizons are showed in contrast with three orthogonal slices in (b). Note that the horizons are correctly interrupted by any discontinuities present in the seismic data.

Results

We tested the proposed methodology using different datasets. Due to space limitations, we present here the results achieved using a dataset from an offshore area in Brazil. This post-stack data comprises 42x42 km, a total area of 1764 km². The seismic cube extended from 1500 to 3500 milliseconds. Figure 2(a) illustrates a typical result achieved mapping only negative amplitude peaks. The method correctly mapped a dense amount of seismic horizons, following the seismic signal with voxel precision. As noted, the mapping accurately stopped where the lateral continuity of the signal degrades such as around seismic faults, salt domes and other discontinuities. Figure 2(b) presents the same surfaces in contrast with 3 seismic slices. We did not use any kind of seismic data conditioning. The horizons visualized were obtained without using any post processing step. The results in 2(b) suggest that there are

attribute along the horizon surface. We used a specific implementation of this attribute, described in Martins et al. (2012). It is presented in 3(d). The attribute shows the highest values on the interface between the horizon surface and the salt domes. Curvature values along the horizon confirm a satisfactory contour of the faults along the horizon surface.

The set of parameters we used in the results reported here are defined below. The set of amplitude samples was composed by the negative peaks. The set of vertical sample sizes, already defined in the $2n+1$ dimensional space, were {25, 25, 25, 27, 27, 27, 29, 29, 29}. The corresponding number of clusters created using each dataset was {93, 98, 104, 102, 107, 96, 99, 109, 111}. The processing time was around 2890 seconds. We used a hp Z820 workstation.

A Seismic Facies Analysis Approach to Map 3D Seismic Horizons

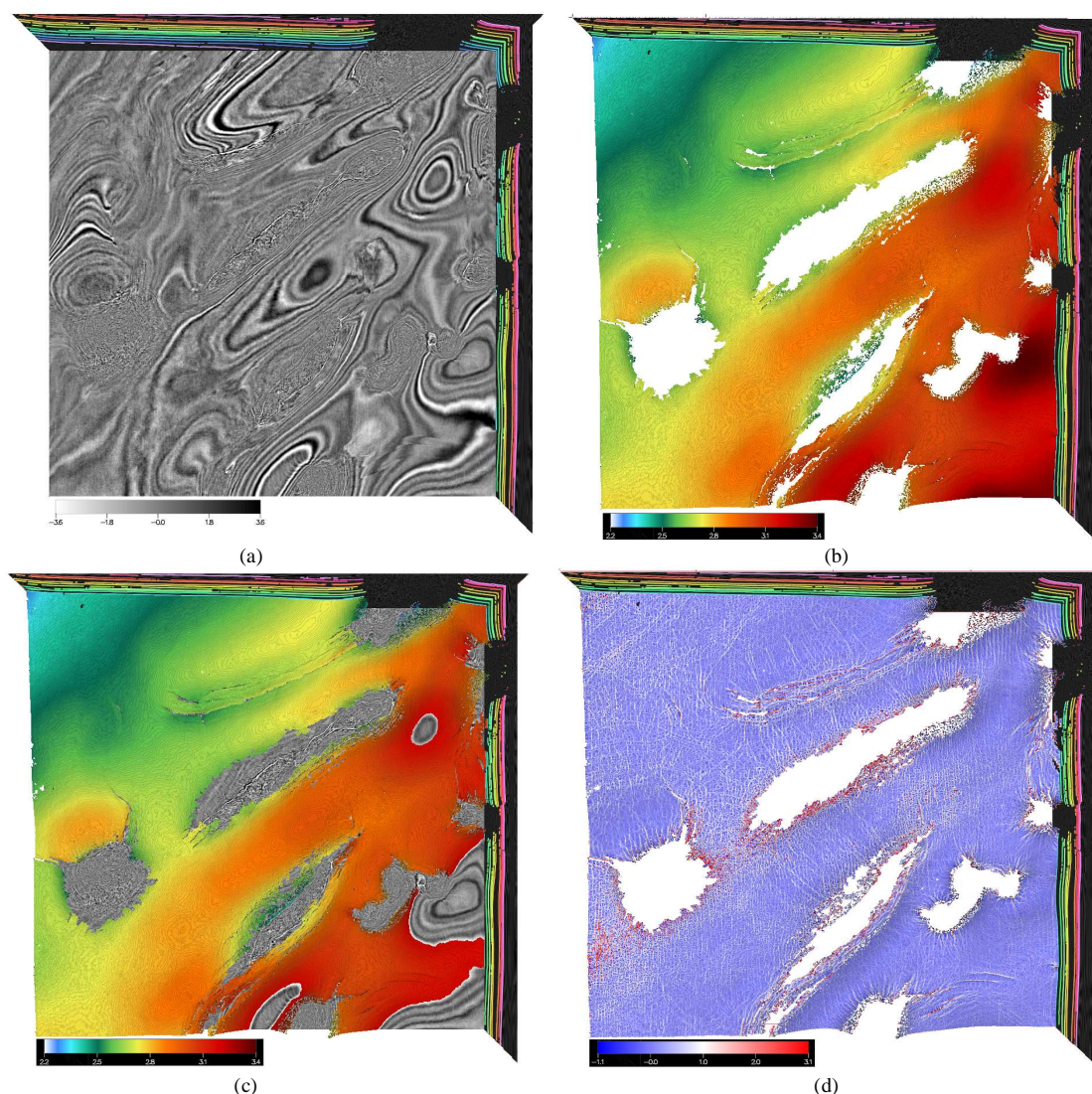


Figure 3: Detailed view of a mapped seismic horizon. A time slice positioned at 3200 milliseconds where salt domes and seismic faults can be identified (a). The horizon procedure maps horizons with voxel precision and stops adequately when the signal degrades. Even radial faults around the salt domes are correctly delineated (b). In (c), we show the seismic horizon in contrast with the time slice. The seismic curvature applied into the horizon surface attests the quality of the mapped procedure (d).

Conclusions

We proposed a solution based on clustering to the problem of automatically mapping horizons in 3D seismic data. Our method creates many representations for each voxel, as to generate various clustering procedures based on the Growing Neural Gas algorithm. The technique organizes the set of voxels according to a global criterion. It is capable of effectively map horizons severely interrupted by discontinuities. We presented experiments indicating its efficiency and illustrated its output. The experimental

results suggest that clustering-based procedures are a good solution for the task of automatically mapping horizons.

Acknowledgments

The authors would like to thank the computational geophysics team of Tecgraf Institute and interpreters from Petrobras for technical contributions. We thank specially Alexandre Kolisnyk by the indispensable assistance. The seismic images shown in this paper were provided by Petrobras.

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

- Bakke, J., O. Gramstad, and L. Sønneland, 2012, Seismic DNA — A novel seismic feature extraction method using non-local and multi-attribute sets: 74th Conference & Exhibition, EAGE, Extended Abstracts, E024.
- de Groot, P., A. Huck, G. de Bruin, N. Hemstra, and J. Bedford, 2010, The horizon cube: A step change in seismic interpretation: *The Leading Edge*, **29**, 1048–1055, <http://dx.doi.org/10.1190/1.3485765>.
- Faraklioti, M., and M. Petrou, 2004, Horizon picking in 3D seismic data volumes: *Machine Vision and Applications*, **15**, no. 4, 216–219, <http://dx.doi.org/10.1007/s00138-004-0151-8>.
- Figueiredo, A. M., P. M. Silva, and M. Gattass, 2013, Surface mapping using auto-adaptable similarity measures: 75th Conference & Exhibition, EAGE, Extended Abstracts, We 06 04.
- Fritzke, B., 1995, A growing neural gas network learns topologies, *in* M. I. Jordan, Y. LeCun, and S. A. Solla, eds., *Advances in neural information processing systems*: MIT Press, 625–632.
- Gramstad, O., J. Bakke, and L. Sønneland, 2012, Seismic surface extraction using iterative seismic DNA detection: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0600.1.
- Hoyes, J., and T. Cheret, 2011, A review of global interpretation methods for automated 3D horizon picking: *The Leading Edge*, **30**, 38–47, <http://dx.doi.org/10.1190/1.3535431>.
- Li, L., G. Ma, and X. Du, 2012, New method of horizon recognition in seismic data: *IEEE Geoscience and Remote Sensing Letters*, **9**, no. 6, 1066–1068, <http://dx.doi.org/10.1109/LGRS.2012.2190039>.
- Marroquín, I. D., J. J. Brault, and B. S. Hart, 2009, A visual data-mining methodology for seismic facies analysis: Part 1 — Testing and comparison with other unsupervised clustering methods: *Geophysics*, **74**, no. 1, P1–P11, <http://dx.doi.org/10.1190/1.3046455>.
- Martins, L. O., P. M. Silva, and M. Gattass, 2012, A method to estimate volumetric curvature attributes in 3D seismic data: 74th Conference & Exhibition, EAGE, Extended Abstracts, E023.
- Wu, X., and D. Hale, 2013, Extracting horizons and sequence boundaries from 3D seismic images: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1440–1445.
- Yu, Y., C. Kelley, and I. Mardanov, 2011, Automatic horizon picking in 3D seismic data using optical filters and minimum spanning tree (patent pending): 81st Annual International Meeting, SEG, Expanded Abstracts, 965–969.

D

Anexo D

An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction.

Artigo submetido ao periódico Computers & Geosciences, 2015.

An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction

A.M. Figueiredo, *Tecgraf/PUC-Rio*; F. B. Silva, *Petrobras*; P.M. Silva, *Tecgraf/PUC-Rio*; Ruy L. Milidiú, *PUC-Rio*; M. Gattass, *Tecgraf/PUC-Rio*

Abstract

We present a new clustering based method to process 3D seismic data and map 3D horizons automatically. The method uses global similarity measurements to organize horizon voxels into clusters. The approach is very robust when mapping seismic horizons severely interrupted by seismic faults, salt domes and other discontinuities. The proposed algorithm replaces the volume voxels creating feature samples to represent the seismic trace signatures. We use PCA and cosine of instantaneous phase to enhance the quality of the signatures. These samples are processed through different clustering procedures. The resulting cluster maps are used to find the global structure of lateral amplitude reflections. The cluster mappings strongly reduce the impact of noise and small disagreements found in the voxels of all horizons. We propose a public benchmark of seismic horizons to validate the results.

1. Introduction

In the 3D seismic data acquisition process, short-duration elastic waves (impulses) are artificially produced at specific points on the Earth's surface. As these waves propagate towards the interior of the Earth, every time they reach an interface between two different layers, part of the incident wave reflects back and another part is refracted to deeper layers. The amplitude of the reflected wave is proportional to the impedance difference between the two layers. When the wave reaches the Earth's surface, a sensor records its amplitude variation with time. This procedure is repeated, varying both impulse and sensor positions. After numerical seismic processing, the recorded waves are registered creating a 3D image of the Earth's sub-surface. An important task of the seismic interpretation process consists of identifying interfaces between two contiguous layers, composing sub-surfaces named seismic horizons. They are identified in the data following peaks, troughs or zero crossings along neighbor seismic traces.

In this work we present an auto-adaptable method to map horizons distributed along the seismic data, based on a local variations similar to the ones presented in Seismic Facies Analysis (Marroquín et al., 2009). Our approach first contextualizes each voxel considering its vertical neighborhood amplitudes and then uses a set of clustering based procedures to group the samples representing each voxel, creating groups of voxels with similar seismic signature. The samples are classified according to a global proximity criterion that takes the entire dataset into account, dismissing any kind of local filter. This global criterion naturally filters noise distortions normally present into seismic data. As it uses only vertical neighborhood information to create the samples, the procedure that maps the horizons does not use any lateral similarity measure, except by its immediate neighborhood between the voxels. Actually, the lateral information is implicitly stored over the sample distribution into the clusters.

Generally, automated horizon-mapping methods use seed-based, auto-tracking, and extract horizons by correlating the local amplitude between neighboring traces. In their work, O'Malley and Kakadiaris (2004) propose a filter-based strategy using a specific filter named orientation-isotropy adaptive (OIA) Gaussian. They presented results suggesting this method works well with continuous horizons, but concluded by acknowledging that filter-based horizon-picking algorithms are insufficient for interpreting some horizon types. An algorithm designed to correlate horizons across faults was presented by Admasu and Toennies (2004), who introduce an automatic method that aligns locations of

the left fault plane onto the right fault plane. This algorithm can be used with filter-based picking methods composing a partial solution to the problem. However, the correlation fails in some cases, as reported by the authors.

Blinov and Petrou (2005) propose a method for quasi-horizontal surface extraction in 3D seismic data. It estimates the local direction of the horizons at each point of the 3D data and uses a filter to smooth the signal reducing the noise level. In a second step, the algorithm simultaneously maps the horizons detected into the seismic cube.

Yu and Kelley (2011) combine pick and trace selection to obtain horizon surfaces. Li et al. (2012) present a methodology that identifies horizons using a combination of horizontal derivative and mathematical morphology. Bakke, Gramstad and Sønneland (2012) describe a method that examines the seismic data with the same approach used in string search, where misalignments are tolerated.

Recently, global approaches have been proposed to compute geological models from the seismic data. Among them, the Horizon Cube (De Groot et al. (2010)) uses local dip-azimuth information calculated from the input seismic cube, filtering the data to improve dip information. Hoyes and Cheret (2011) present a review that summarizes global interpretation methods for 3D horizon mapping.

Wu and Hale (2015) describe two methods for construction of seismic horizons: generating surfaces one at a time, or generating a volume of horizons. The surfaces are found solving partial differential equations, which in turn are obtained from sets of approximate normal vectors for each of the volume voxels. In the text, the authors observe that approximations of normal vectors are not accurate enough to properly map horizons automatically, around areas of disagreement, seismic faults, or in areas where an image is noisy. The most significant aspect of this method is that it allows the user to assist the horizon procedure, specifying interactively sets of control points that define the model constraints. Figueiredo et al. (2013, 2014) present global methodologies capable of transform seismic data from the amplitude space into a multi-dimensional space, using clusters from vertical windows of samples to map horizons. The methods organize the seismic data according to a global and auto-adaptable criterion, and maps horizons using fixed seeds. This avoids a major weakness of many methods, i.e., solve huge series of local problems that leads to poor mapping quality.

In this work we improve the method presented in (Figueiredo et. al. 2014). Our approach uses the cosine of instantaneous phase attribute and applies Principal Component Analysis (PCA) (Partridge and Calvo, 1997) into the original datasets of trace shapes in order to improve the quality of the original samples. Then we propose a measurement to estimate the quality of the clusters used to map the seismic horizons. Based on this measurement, we show that using cosine of instantaneous phase attribute as well as PCA greatly improve the mapping of seismic horizons. Finally, we present our horizon mapping procedure, and report results using public seismic data, allowing a better evaluation of our work.

The outline of the paper is the following: in Section 2 we define the subset of voxels to process using a cosine of instantaneous phase version of the entry seismic data. In Section 3 we describe how to represent these voxels by vectors of trace shapes (vertical neighboring amplitudes) and use PCA to extract its principal components. In Section 4 we use these datasets and a specific clustering algorithm called Growing Neural Gas (GNG) (Fritzke, 1995) to create groups (clusters) that represent similar signal variations. Based on the clustering classification, we define a function to measure similarity between two given voxels. Section 5 presents a method to infer the quality of the clustering procedures. We also discuss how this measurement can be used to define training parameters according to the input data in Section 6. In Section 7 we describe our horizon mapping algorithm. Finally, in Section 8 we present results, and conclusions.

2. Selecting the type of Samples to Process

Our method characterizes the voxels of interest through windows of vertical neighbors, a representation commonly used in the field of Seismic Facies Analysis (Marroquín et. al., 2009). It contextualizes each voxel considering its vertical neighboring amplitudes, and creates a signature that adequately characterizes the voxels along the data. However, in this work in order to minimize the amount of voxels to be processed, the method separates the volume into four main groups of voxels: positive amplitude peaks, negative amplitude peaks, negative to positive zero crossing, and positive to negative zero crossing points. The horizons to be mapped are formed by one of these types of voxels.

3. Voxel Representation and Principal Component Analysis

Once the group of horizons to be mapped is defined, all the voxels of this group are assemble in a dataset \mathbf{M} . For each voxel in this dataset it is critical to define the size of the vertical analysis window that closely follows the local stratigraphy to avoid contaminating the feature vectors with amplitudes unrelated to the voxel of interest. As the stratigraphy features vary along the volume data, we change the window sizes, creating r feature vectors for each voxel. The window sizes are defined in the set $\mathbf{N}=\{n_0, n_1, ..., n_{r-1}\}$. Consider a_{ti} as the (i -th) voxel of a trace (t), ($a_{ti} \in \mathbf{M}$). We compute each of its r corresponding vertical windows by:

$$s_{tij} = \{a_{t(i-n_j)}, ..., a_{ti}, a_{t(i+1)}, ..., a_{t(i+n_j)}\}, s_{tij} \in R^{2n_j+1} \quad (1)$$

We store the feature vectors according to its size, creating r datasets of vertical windows. The samples created using $n_j \in \mathbf{N}$ are stored on the corresponding dataset \mathbf{D}_j . We then use the actual datasets of trace shapes and apply a statistical technique called Principal Component Analysis (PCA) (Partridge and Calvo, 1997) to reduce the dimension of the trace shapes and capture relevant features.

Basically, PCA uses an orthogonal transformation that converts a dataset of possibly correlated variables into another dataset formed by feature vectors of linearly uncorrelated variables, called principal components. This transformation is defined so that the first principal component captures the largest variance. Each succeeding components have the highest possible variance under the constraint that they are uncorrelated with the previous components. According to Chopra and Marfurt (Chopra and Marfurt, 2014), PCA is the first step in many clustering techniques such as self-organizing maps, generative-topographic maps, and support-vector machine analysis.

We create a corresponding PCA procedure to each of our r datasets of vertical windows. At the end of this step, for each dataset \mathbf{D}_j we create a corresponding dataset \mathbf{D}'_j , as depicted in Figure 1.

4. Clustering the Datasets of Samples

For each dataset \mathbf{D}'_j we apply a clustering procedure to separate the samples into distinct groups that share similar trace shapes. By our construction, a vertical sample does not contain information regarding its side neighborhood. However, due to existing natural similarity of traces belonging to the same horizon, samples tend to be classified in the same cluster. In this work we use a clustering algorithm known as Growing Neural Gas (GNG) (Fritzke, 1995). Nevertheless, similar results can be achieved by other clustering algorithms such as k-means (Hartigan (1975)).

After the GNG execution, each cluster receives a unique numeric label id . We associate to each voxel a_{ti} , ($a_{ti} \in \mathbf{M}$), an equivalent set of labels, $\mathbf{E}_{ti} = \{id_0, id_1, \dots, id_{r-1}\}$ containing the ids of the clusters where its r corresponding feature vectors were classified (Figure 1).

Using the sets of labels, we state a function to measure the similarity between two arbitrary elements. Given two voxels a_1 and a_2 , we define $S(a_1, a_2)$ as the cardinality of the intersection between its corresponding sets \mathbf{E}_{a1} and \mathbf{E}_{a2} :

$$S(a_1, a_2) = |E_{a1} \cap E_{a2}|, 0 \leq S(a_1, a_2) \leq r, \quad (2)$$

where higher values of S indicate greater similarity.

5. Inferring the Quality of Clustering Procedures

In our method, the set of parameters to be defined are: the group of voxels to be processed, the r window sizes, the number of clustering procedures to create and the number of clusters generated by each clustering procedure.

As the quality of the tracked horizons is directly related to the appropriate configuration of these parameters, we define a measure that estimates the quality of an arbitrary configuration.

Our algorithm separates the datasets of samples into meaningful groups. Samples classified into the same cluster representing voxels that are lateral neighbors on the seismic data have a high probability to be part of the same horizon. Similarly, two or more vertical neighbor voxels on a seismic trace should not be part of the same horizon and cannot be classified into the same cluster. Using our similarity function S , an arbitrary voxel should have at most one horizon neighbor in each neighbor trace. This is the base of the measurement we use to verify the quality of a given set of parameters.

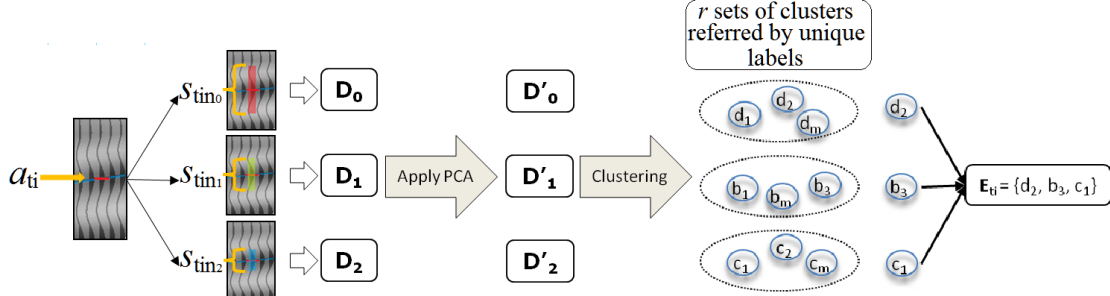


Figure 1: Each sample a_{ti} from the subset to be processed (e.g., negative peaks) is represented by its r vertical windows, where the half size of each window is defined on the set \mathbf{N} of half sizes. Here, $r=3$, with $\mathbf{N}=\{n_0, n_1, n_2\}$, giving rise to the samples s_{tin0} , s_{tin1} , and s_{tin2} , stored into its corresponding datasets, \mathbf{D}_0 , \mathbf{D}_1 and \mathbf{D}_2 . After applying the PCA on these initial datasets, we create its corresponding version formed by the principal components, \mathbf{D}'_0 , \mathbf{D}'_1 and \mathbf{D}'_2 . Each of these datasets is processed by an instance of the GNG clustering algorithm creating a corresponding set of clusters. Each cluster receives a unique label id . After all the clustering and classification procedures have been taken place, each sample a_{ti} is represented by its corresponding vector of labels, \mathbf{E}_{ti} containing the ids of the clusters where its r corresponding samples were classified.

Consider a_{ti} as the $(i$ -th) voxel of a trace (t) , ($a_{ti} \in \mathbf{M}$). We define as candidates all the voxels $a' \in \mathbf{M}$ included into a cube with half size h and centered on a_{ti} , except by the voxels above and below a_{ti} on the same trace (t) . To identify *conflicted voxels* we use the similarity function S to find the number of

horizon neighbors of a_{ii} on each column of the cube: if there is an unique candidate a' so that $S(a_{ii}, a') > 0$, it is considered to be a horizon neighbor. If there are more candidates considered as horizon neighbors, this occurrence is counted as a conflict.

We examine the entire volume, summing up for each voxel $a_{ii} \in \mathbf{M}$ the number of horizon neighbors and the number of conflicts. After that, we divide these numbers by the total number of voxels in the set \mathbf{M} . We call this metric as $CvC(h)$, that is, continuity versus conflicts using a window of half size h .

This measurement is used to find the ideal training parameters of our method. The best configuration maximizes the rates of lateral continuity while minimizing the number of conflicts.

Once a configuration is considered adequate, all the encountered conflicts are removed from the vectors of labels before the execution of the horizon mapping procedure.

6. Discussion

In this section we use the metric previously presented to assess possible configurations that generates best clustering setups. Using the $CvC(h)$ metric we show that the cosine of instantaneous phase attribute improves the general performance of the method. We also show that the use of PCA to create feature vectors from the vertical samples minimizes the average occurrence of conflicts and maintains the lateral continuity almost unchanged.

Several clustering algorithms such as GNG and k-means, use the Euclidean distance as its similarity function. The Euclidean metric is not adequate to our subject since it creates distortions on the distribution of clusters formed by vertical amplitude samples. Seismic regions characterized by intense seismic signal generally have amplitude peaks with large magnitude. The measurement of the Euclidean distances between these samples results in high values, causing an over distribution of clusters into the regions of the samples' space where they are located. The reverse is evidenced in areas with low average amplitudes of the seismic signal. Using the cosine of instantaneous phase to create the vertical samples partially corrects the distortions. This attribute preserves the initial shape of the seismic wave normalizing the amplitude peaks. We report comparative results using the 3D dataset known as F3 Block, in the North Sea, Netherlands, which is publicly available via dGB's Seismic Repository (<https://opendtect.org/osr>).

We created six instances of the GNG clustering algorithm, processing negative peaks. The vertical windows were created using a half-size of 10 voxels. We execute the GNG incremental training procedure, measuring continuities and conflicts $CvC(3)$ at each 10 clusters.

In Table 1 we show results using vertical samples from the amplitude volume data. In Table 2 we show the results applying the PCA into the samples of the amplitude volume, extracting four principal components to each vertical sample. Finally, in Table 3 we use the cosine of instantaneous phase and PCA to extract four principal components.

The use of PCA on the vertical samples considerably reduces the number of conflicts, as verified comparing Tables 1 and 2. According to the tables, we should use 80 clusters to produce a continuity factor around 45% when taking samples that uses amplitudes (Table 1). In this case the rate of conflicts is 0.80%. Using PCA and cosine of instantaneous phase, we can obtain the same level of continuity using only 40 clusters, with conflicts rate of 0.14%. This behavior is maintained as we vary the number of clusters.

We performed comparative tests in different seismic data getting similar results. It suggests that the use of PCA and cosine of instantaneous phase generate a significant improvement over the other arrangements.

An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction

Clusters	Amplitude	
	Mean Continuity (%)	Mean Conflicts (%)
40	52.01	1.380
50	49.60	1.118
60	47.30	1.024
70	47.17	0.958
80	45.09	0.806
90	43.56	0.695
100	42.61	0.597

Table 1: Average continuities and average conflicts per sample when varying the number of clusters using dataset of vertical amplitude samples directly.

Clusters	Amplitude with PCA	
	Mean Continuity (%)	Mean Conflicts (%)
40	37.35	0.439
50	34.04	0.322
60	32.53	0.247
70	30.34	0.193
80	29.19	0.168
90	27.97	0.141
100	26.86	0.127

Table 2: Average continuities and average conflicts per sample when varying the number of clusters using dataset formed by the first 4 principal components from vertical amplitude samples.

Clusters	Inst. Cos. Phase with PCA	
	Mean Continuity (%)	Mean Conflicts (%)
40	44.15	0.148
50	33.22	0.105
60	30.08	0.082
70	28.83	0.069
80	26.86	0.053
90	25.84	0.048
100	24.09	0.040

Table 3: Average continuities and average conflicts per sample when varying the number of clusters using dataset formed by the first 4 principal components from vertical samples taken from the Cosine Instantaneous Phase.

7. Mapping Seismic Horizons

The clustering processes group similar samples into same groups, without take into consideration any kind of lateral similarity measures. This strategy provides some important advantages:

1. The clusters implicitly store information about its corresponding samples' set, and its disposition along the samples space reflects the local density of its samples. The distribution of the datasets of samples through the clusters are data driven and auto adaptable;
2. Neighboring layer voxels represented by their corresponding feature vectors share relative similarities, and probabilistically tend to be located in same clusters, or at least in closely positioned clusters. At same time, areas of relative breaks in the lateral continuity are represented into the samples' space by a relative lack of similarity with respect to its neighboring samples and tends to be naturally classified into different clusters;
3. As a same cluster stores considerable number of samples, it is possible to use the similarity function S with a fixed seed voxel and look for horizon neighbors located over a considerable lateral area.

Before describe the horizon mapping procedure we need to define our concept of immediate neighbor voxels. Given an arbitrary seismic trace t , we define generally as t'_0, \dots, t'_7 as its 8 immediate trace neighbors on the 3D data. Considering a_{ti} as the i -th voxel of a trace (t), we define the immediate neighbors of a_{ti} as the voxels $a'_{t'(i-1)}$, $a'_{t'_i}$, and $a'_{t'(i+1)}$, the $(i-1)$ -th, i -th, and $(i+1)$ -th voxels on every of the immediate neighbor traces t'_n .

We describe our horizon-mapping using two procedures. The first method receives an initial voxel a_{init} , ($a_{init} \in \mathbf{M}$) and returns a set of voxels \mathbf{H} containing the discovered voxels, including a_{init} . It constructs a priority queue \mathbf{L} of pairs $\langle priority, a_x \rangle$, where a_x , ($a_x \in \mathbf{M}$), refers generically to the already discovered voxels. The *priority* is found through the similarity function S , measuring $S(a_{init}, a_x)$. It can be described as follows:

$\mathbf{H} \leftarrow findSubHorizon(a_{init})$

step 1: the method receives the initial voxel a_{init} , ($a_{init} \in \mathbf{M}$). This voxel is inserted in \mathbf{L} . Its priority is set to one;

step 2: take and remove the voxel a_{ti} of greatest priority from \mathbf{L} . Insert a_{ti} into the set \mathbf{H} ;

step 3: for each immediate neighbor trace from a_{ti} , t'_0, \dots, t'_7 , use the function S and a_{init} to find the most similar neighbor, $a_{neighbor}$ between the candidates $a'_{t'(i-1)}$, $a'_{t'_i}$, or $a'_{t'(i+1)}$. If $S(a_{init}, a_{neighbor}) > 0$ and if ($a_{neighbor} \notin \mathbf{H}$), then $a_{neighbor}$ is inserted in \mathbf{L} , and its priority is the similarity value;

step 4: repeat the steps (2) through (4) until \mathbf{L} is empty;

step 5: return the set \mathbf{H} of discovered voxels.

The second algorithm receives a seed voxel a_{seed} , ($a_{seed} \in \mathbf{M}$) to begin with. It iterates using *findSubHorizon*, and composes the complete surface using the returned sub-horizons. Beginning with a_{seed} , at each step a new seed is chosen randomly from already discovered voxels. As it receives new horizon portion, it looks for possible errors comparing the sub-horizon with the already discovered voxels. The procedure returns a set \mathbf{H}_{final} containing the 3D coordinates of the discovered horizons. It can be described as follows:

$H_{final} \leftarrow \text{mapHorizon}(a_{seed}, n_{max})$

step 1: create an integer counter n and set its value to 0;

step 2: $H \leftarrow \emptyset$ and make $H \leftarrow \text{findSubHorizon}(a_{seed})$;

step 3: compare the voxels from the sets H and H_{final} . If there are voxels from a same seismic trace that have different depth, discard the set H and increment n by one. Otherwise, set n to zero, and copy the voxels from H into H_{final} ;

step 4: take randomly a new a_{seed} from the voxels in H_{final} ;

step 5: repeat the steps (2) through (4) until n equals n_{max} .

We can complete small horizon holes that were not mapped using the mean depth of its discovered neighbor voxels. When there are horizons' portions that are not continuous in the data, it is necessary to add new seed voxels manually to complete the mapping.

8. Results

The proposed method was implemented in language C++ and tested with many different seismic datasets. Despite the presence of different seismic events, which usually makes subsurface mapping a harder task, the proposed method is able to properly detect horizons. In fact, our method is able to interrupt the mapping procedure in the presence of discontinuities. Besides that, the methodology presented another advantage: even considering that in few cases the method was not able to map more than a few voxels from the seed it received, errors were not possible to propagate in all the tests we conducted. It is an important result once error propagation can invalidate the posterior utilization of the surfaces that were mapped.

We present the results obtained using two different datasets. The first is a 3D dataset known as F3 Block, downloaded from the Opendtect website. Moreover, we present results achieved using a dataset from an offshore area in Brazil. This post-stack seismic data comprises 42 per 42 km, a total area of 1.764 km². The seismic cube extended from 1.500 to 3.500 milliseconds.

Concerning the results reported using the F3 Block, we define below the parameters we used. The set \mathbf{M} was composed by the negative peaks. We produced 24 GNG clustering procedures. Half were created using datasets in which the vertical samples contained 19 samples. The other half used datasets with vertical samples containing 21 samples. To create the datasets, we used PCA, extracting 4 principal components. All the clustering procedures had around 85 clusters. We did not use any kind of pre-conditioning algorithm in order to improve the quality of the seismic data.

After training and classification steps, we measured the continuity versus conflicts using a window of half size 3, obtaining a continuity value of 35.82% and conflicts rate of 0.085%. We used a hp Z820 workstation.

To measure the accuracy of the mapped surfaces, we used a public benchmark constructed for the F3 Block. This benchmark contains 5 horizons interpreted and validated line by line by an experienced interpreter geophysicist. Besides the fact that the benchmark is not exact, once it was manually mapped in its greatest part, as far as we know it is the most precise public set of horizons available for this seismic data. It is available to download at www.tecgraf.puc-rio.br/repo/seismic.

An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction

We used our methodology to map the same 5 seismic horizons. For each horizon we compared our results with the version provided by the benchmark. Table 4 resumes our results, where we compare the 5 horizons one by one, measuring the rates of false positives, false negatives and RMS error. In the table, N_{BEN} refers to the number of voxels available into the benchmark version. N_{Map} refers to the number of voxels mapped using our algorithm.

Idx	N_{BEN}	N_{Map}	False Pos(%)	False Neg(%)	RMS
0	525181	546865	4.89	0.76	0.680
1	538549	540850	1.57	1.15	0.916
2	467893	496896	9.81	3.61	1.185
3	512494	513105	5.38	5.26	1.349
4	494339	488615	3.19	4.35	1.495

Table 4: A summary of the results achieved when comparing the 5 horizons available on the F3 benchmark and its corresponding counterparts mapped using our method. In the table, N_{BEN} refers to the number of voxels available into the benchmark version. N_{Map} refers to the number of voxels mapped using our algorithm. We measured the rates of false positives, false negatives and **RMS** error between benchmark coordinates and that mapped by our algorithm.

Concerning the number of seeds used to map the presented horizons; this number was not greater than 6 in any of the reported results. In fact the first two horizons (Idx 0 and Idx 1) were mapped using only took 2 seeds each. Horizons number 2 and number 3 were the most difficult to be mapped, as they are located into a specific region of the data characterized by having high occurrences of lateral breaks, presenting a very low quality of the seismic signal. Horizons number 2 and 3 took 5 and 6 different seeds to be mapped, respectively. The seeds were not chosen in order to minimize the amount necessary to map the horizons. Actually, all horizons were mapped using the same steps: the user chooses a unique seed to begin with. Once the method returns the resulting surface the user examine the surface and adds a new seed into the necessary areas.

The Figures 2 and 3 show visual comparisons made between the horizons. In these figures, the left column presents the benchmark version of the horizons, and the column on the right presents its corresponding version mapped using our methodology. In Figure 3, we present an intersection of the horizons in contrast with two seismic slices, allowing a better visual comparison. As we can note, the mapped versions are highly competitive.

We also present results achieved using a dataset from an offshore area in Brazil, comprising 42 per 42 km, or 1.764 km². The seismic cube extended from 1500 to 3500 milliseconds.

In this case, instead of insert each seed manually, we used an algorithm that creates the sets of seeds from which the horizons were tracked. It constructs a list \mathbf{L} of seed voxels ordered by priority. The priority of each voxel a , ($a \in \mathbf{M}$), is given using the similarity function S , summing up the similarity values between a and the more similar immediate neighbor on the neighboring traces.

Figure 4 illustrates a typical result achieved mapping only negative amplitude peaks, where we present the intersection of the surfaces in contrast with 3 seismic slices. The method correctly mapped a dense amount of seismic horizons, following the seismic signal with high precision. As noted, the mapping accurately stopped where the lateral continuity of the signal degrades, such as around seismic faults, salt domes and other discontinuities. Again, we did not use any kind of seismic data pre-conditioning.

An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction

We apply a volumetric seismic curvature attribute along the horizon's surface, using a specific implementation described in Martins et al. (2012). The result is presented in Figure 5. In Figure 5(a) we show the resultant mapped surface and in Figure 5(b) we show the same surface after applying the maximum curvature attribute on the horizons' surface. As we can verify, the application of this attribute revealed radial faults around salt domes, emphasized into the surface.

In Figure 6, we again applied maximum curvature on the horizon surface revealing, in this case, footprints along the acquisition direction, as well as radial faults.

Figure 7 presents another result that can only be achieved by mapping surfaces with high precision, fitting the seismic signal adequately. It is possible to identify various pockmarks present on the horizon's surface, as well as a very well delineated fault.

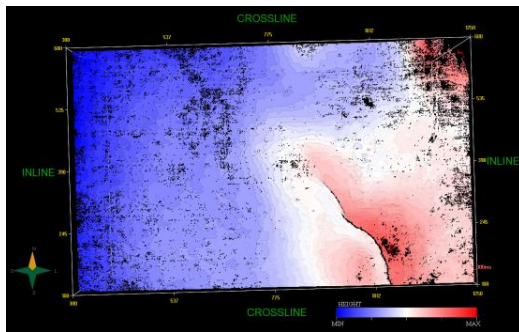
9. Conclusions

We proposed a solution based on clustering to the problem of automatically mapping horizons in 3D seismic data. Our method creates many representations for each voxel, as to generate various clustering procedures based on the Growing Neural Gas algorithm. The technique organizes the set of voxels according to a global criterion. It is capable of effectively map horizons severely interrupted by discontinuities. We presented experiments indicating its efficiency and illustrated its output. The experimental results suggest that clustering-based procedures are a good solution for the task of automatically mapping horizons.

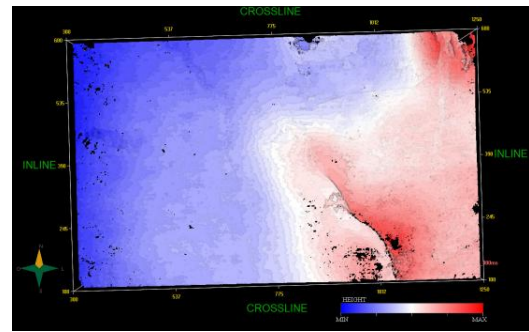
Acknowledgments

The authors would like to thank the computational geophysics team of Tecgraf Institute and interpreters from Petrobras for technical contributions. We thank specially João Paulo Peçanha, Geisa Faustino and Leonardo Martins by the indispensable assistance. The seismic images shown in this paper were partially provided by Petrobras.

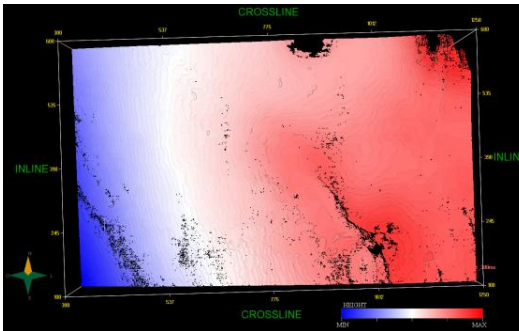
An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction



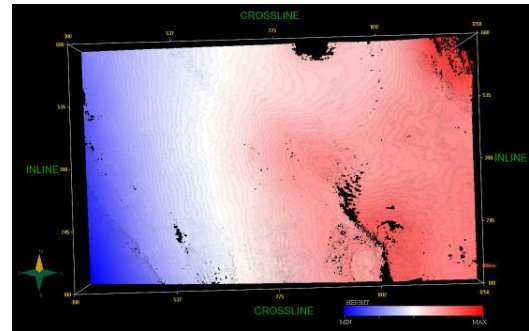
(a) Benchmark Idx 0.



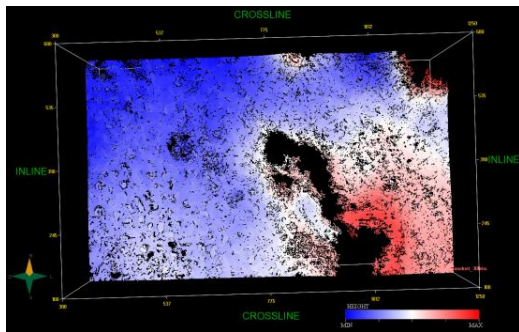
(b) Mapped version of Idx 0.



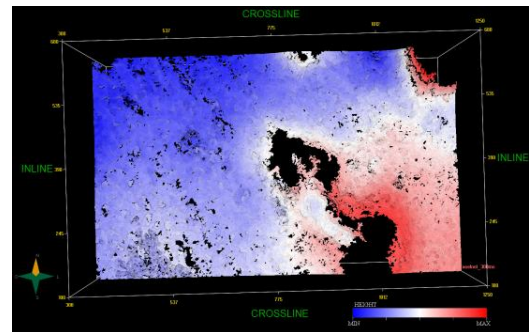
(c) Benchmark Idx 1.



(d) Mapped version of Idx 1.



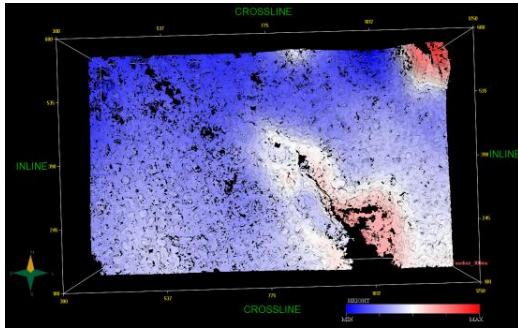
(e) Benchmark Idx 2.



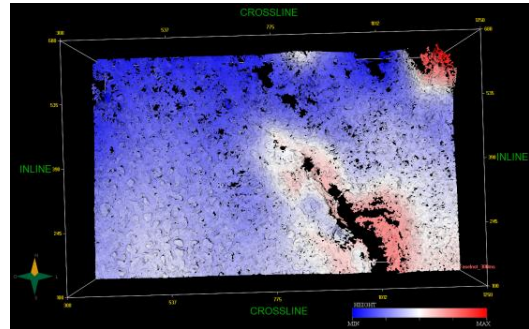
(f) Mapped version of Idx 2.

Figure 2: On the left column we see the 3 first horizons available into the benchmark. The column on the right exhibits its corresponding versions mapped using our methodology.

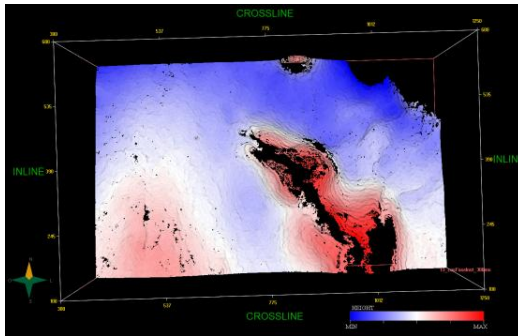
An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction



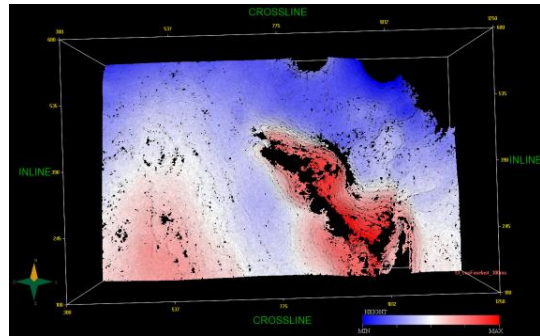
(a) Benchmark Idx 3.



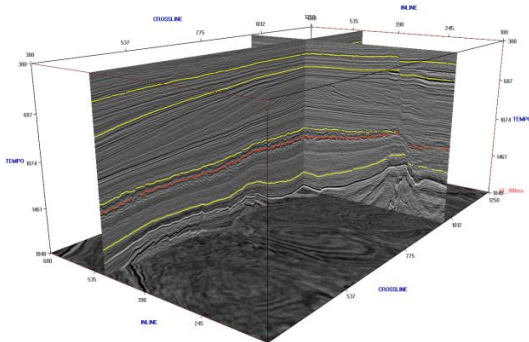
(b) Mapped version of Idx 3.



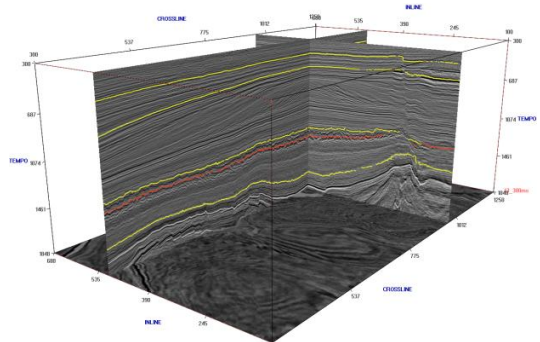
(c) Benchmark Idx 4.



(d) Mapped version of Idx 4.



(e) Benchmark version.



(f) Mapped version.

Figure 3: On the left column we see the 3 first horizons available into the benchmark. The column on the right exhibits its corresponding versions mapped using our methodology.

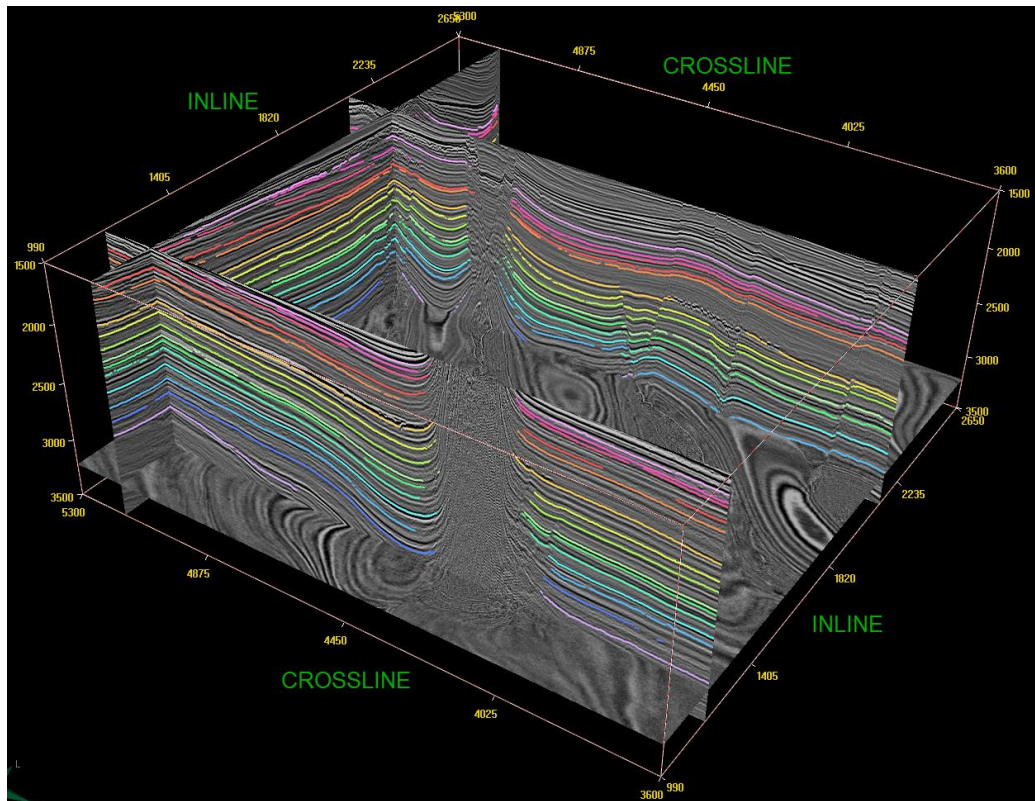


Figure 4: Results of proposed method. Mapped horizons in contrast with 3 seismic slices. The seismic cube extended from 1500 to 3500 milliseconds. Each color indicates a different horizon. Note that the horizons were correctly interrupted by the discontinuities present in the seismic data.

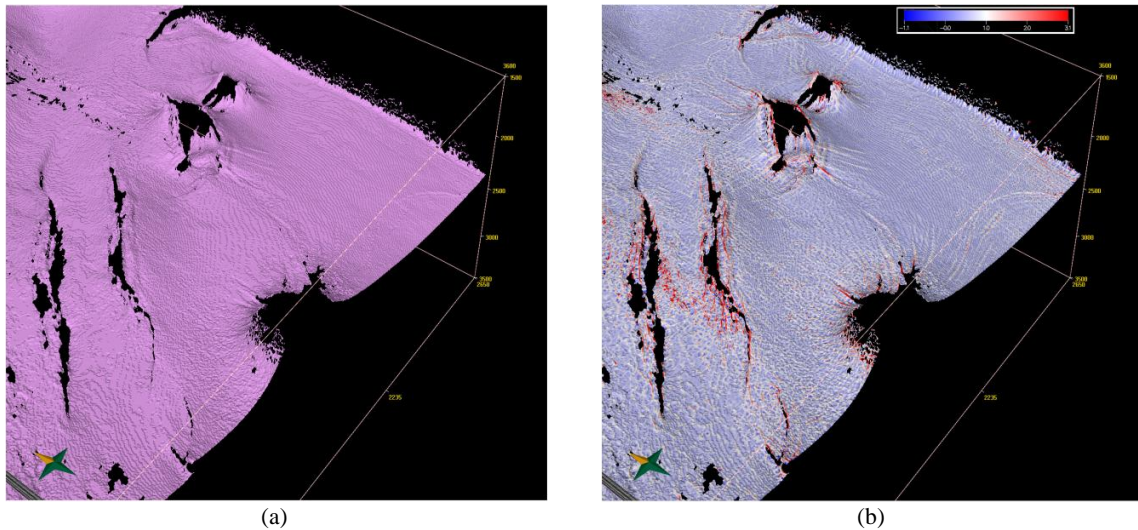


Figure 5: Mapping horizons with high precision provides advantages. In (a) we see a mapped surface that surrounds seismic events interrupting the seismic signal. In (b) we apply maximum curvature attribute on the horizons' surface, what provides important information: in this case it emphasized radial faults represented into the surface.

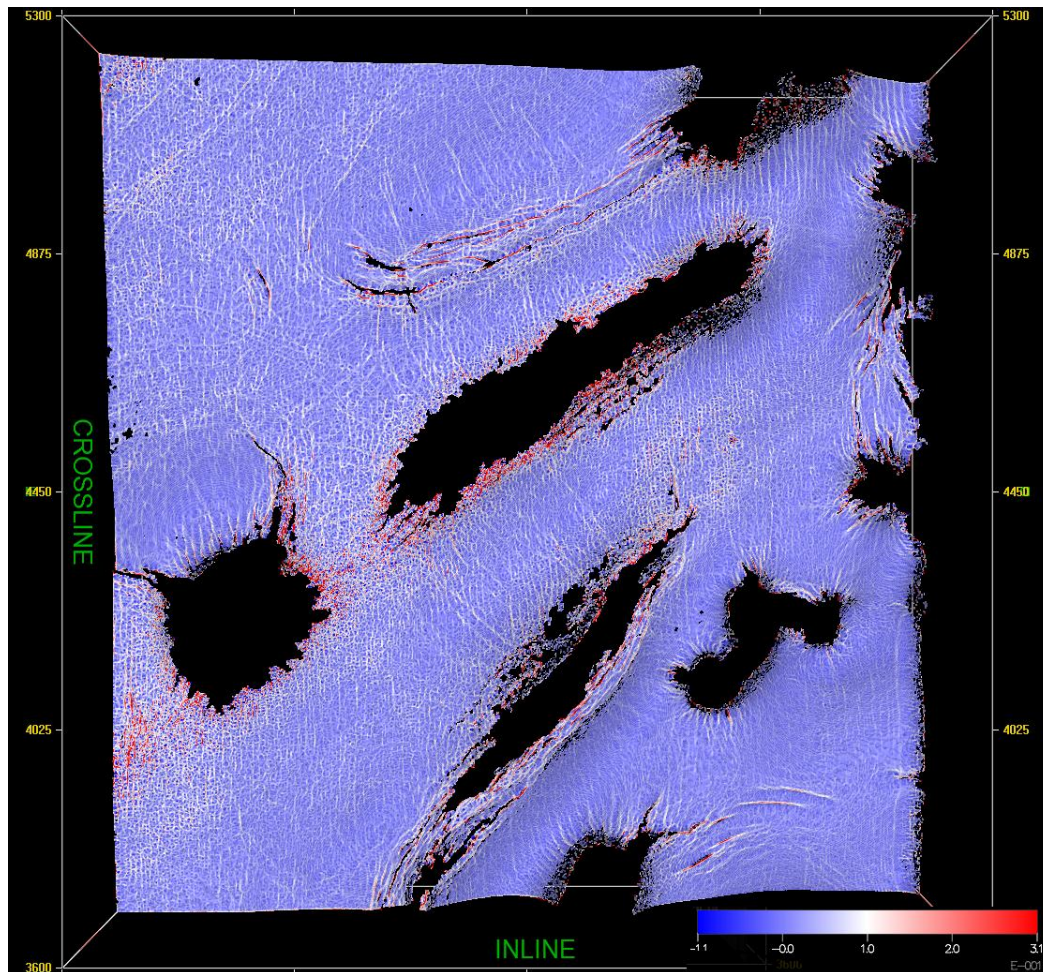


Figure 6: We present a case where the horizon extended along all the seismic data. The surface correctly respected the seismic events interrupting its signal, in this case seismic faults and salt domes. We again applied curvature attribute into the surface emphasizing radial faults and foot prints along the acquisition direction.

References

- Admasu, F., and Toennies, K., 2004. **Automatic method for correlating horizons across faults in 3D seismic data.** IEEE Conference on Computer Vision and Pattern Recognition, Washington DC, June 2004.
- Aurnhammer, M and Tonnes, K (2005). **A Genetic Algorithm for Automated Horizon Correlation Across Faults in Seismic Images.** *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, VOL. 9, NO. 2, APRIL 2005.
- Bakke, J., Gramstad, O. and Sønneland, L., 2012. **Seismic DNA - A Novel Seismic Feature Extraction Method Using Non-local and Multi-attribute Sets,** *74th EAGE SPE EUROPEC.*

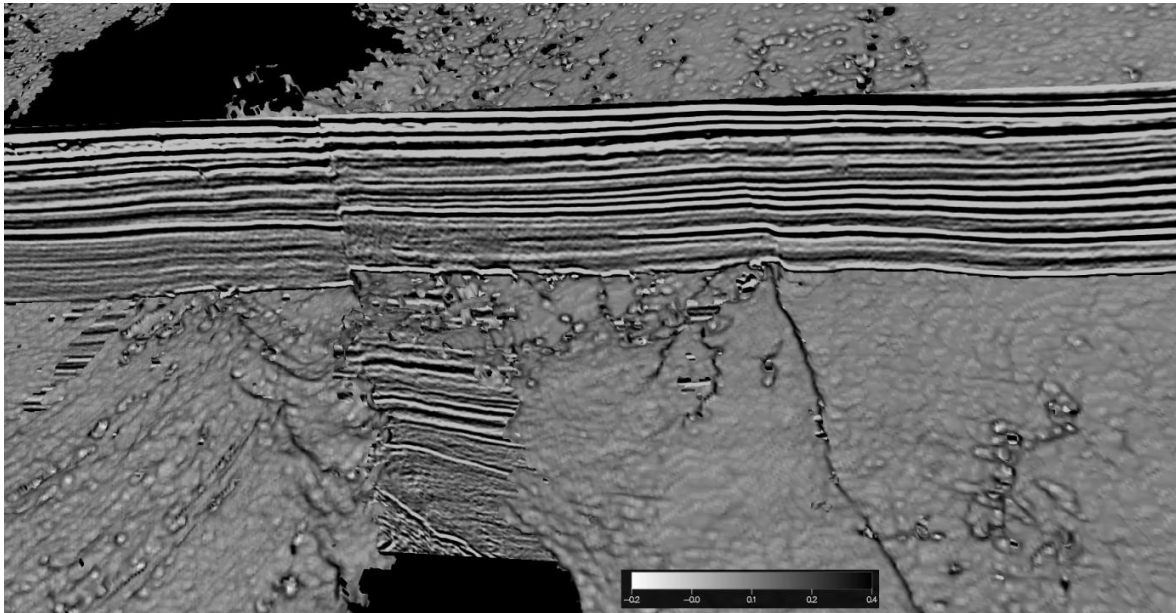


Figure 7: Mapping horizons with great precision revealed pockmarks. We also see a delineated fault.

- Borgos, H. G., T. Skov, T. Ramden, and L. Sonneland, 2003. **Automated geometry extraction from 3D seismic data**. 73rd Annual International Meeting, SEG, Expanded Abstracts, 1541–44.
- Chopra, S., Marfurt, K. J., 2014. **Churning seismic attributes with principal component analysis**.
- Figueiredo, A. M., Silva P. M., and Gattass M., 2013, **Surface mapping using auto-adaptable similarity measures**. 75th EAGE Conference.
- Figueiredo A.M., Silva F. B., Silva P.M., Milidiú Ruy L., Gattass M., 2014, **A seismic facies analysis approach to map 3D seismic horizons**. 84th Annual International Meeting, SEG, Expanded Abstracts.
- Fomel, S., 2010. **Predictive painting of 3D seismic volumes**. GEOPHYSICS, VOL. 75, NO. 4, P. A25–A30.
- Gramstad, O., Bakke, J. and Sonneland, L., 2012. **Seismic surface extraction using iterative Seismic DNA detection**. 82nd Annual International Meeting, SEG, *Expanded Abstracts*.
- Hoyes, J. and Cheret, T., 2011. **A review of "global" interpretation methods for automated 3D horizon picking**. *Leading Edge*, vol. 30, no. 1, pp. 38–47.
- Karimi, P., Fomel, S., 2015. **Stratigraphic coordinates - a coordinate system tailored to seismic interpretation**. *Geophysical Prospecting*, vol 63, 3, march 2015
- Li, L., Ma, G., and Du, X., 2012. **New Method of Horizon Recognition in Seismic Data**. *IEEE Geoscience And Remote Sensing Letters*, Vol. 9, No. 6.
- Lomask, J., and A. Guitton, 2007, **Volumetric flattening: an interpretation tool**. *The Leading Edge*, 26, no. 7, 888–897.
- Marroquín, I. D., Brault J. J., and Hart B. S., 2009. **A visual data-mining methodology for seismic facies analysis: Part 1 - Testing and comparison with other unsupervised clustering methods**. *Geophysics*, 74, no. 1, P1–P11.
- Martins, L. O., Silva, P. M. and Gattass, M., 2012. **A Method to Estimate Volumetric Curvature Attributes in 3D Seismic Data**. 74th EAGE Conference & SPE EUROPEC 2012

An Auto-adaptable Clustering-based Approach for 3D Seismic Horizon Extraction

- O'Malley, S. and Kakadiaris, I., 2004. **Towards robust structure-based enhancement and horizon picking in 3-d seismic data.** IEEE Int. Conf. on Computer Vision and Pattern Recognition, pages 482–489, July 2004.
- Partridge, M., Calvo, R., 1997. **Fast Dimensionality Reduction and Simple PCA.** Intelligent Data Analysis, 2, 292-298.
- Pauget, F., S. Lacaze, and T. Valding, 2009. **A global interpretation based on cost function minimization.** 79th Annual International Meeting, SEG, Expanded Abstracts, 2592-96.
- Verney, P., M. Perrin, M. Thonnat, and J. F. Rainaud, 2008, **An approach of seismic interpretation based on cognitive vision.** 70th EAGE Conference and Exhibition.
- Yu, Y., Kelley, C. and Mardanova, I., 2011. **Automatic horizon picking in 3D seismic data using optical filters and minimum spanning tree.** SEG Expanded Abstracts.
- Wu, X., Hale, D., 2013. **Extracting horizons and sequence boundaries from 3D seismic images.** 83rd Annual International Meeting, SEG, Expanded Abstracts.
- Wu, X., Hale, D., 2015. **Horizon volumes with interpreted constraints.** GEOPHYSICS, VOL. 80, NO. 2 P. IM21–IM33.
- | Hartigan, J. A., 1975. **Clustering Algorithms.** New York: Wiley.

Referências Bibliográficas

- [1] ADMASU, F.; TOENNIES, K. **Automatic method for correlating horizons across faults in 3d seismic data.** In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, Washington DC, June 2004.
- [2] AURNHAMMER, M.; TONNIES, K. **A genetic algorithm for automated horizon correlation across faults in seismic images.** IEEE Transactions on Evolutionary Computation, 09(2), Apr. 2005.
- [3] BAKKE, J.; GRAMSTAD, O. ; SØNNELAND, L. **Seismic dna - a novel seismic feature extraction method using non-local and multi-attribute sets.** In: PROCEEDINGS OF THE 74TH EAGE CONFERENCE, Copenhagen, 2012.
- [4] BLINOV, A.; PETROU, M. **Reconstruction of 3-d horizons from 3-d seismic datasets.** IEEE Transactions on Geoscience and Remote Sensing, 43(6), June 2005.
- [5] BORGOS, H. G.; SKOV, T.; RAMDEN, T. ; SONNELAND, L. **Automated geometry extraction from 3d seismic data.** In: PROCEEDINGS OF THE 73RD SEG ANNUAL MEETING, Dallas, Oct. 2003.
- [6] CHOPRA, S.; MARFURT, K. J. **Churning seismic attributes with principal component analysis.** In: PROCEEDINGS OF THE 84ND SEG ANNUAL MEETING, Denver, Oct. 2014.
- [7] FIGUEIREDO, A.; FAUSTINO, G.; PEÇANHA, J.; SILVA, P. ; GATTASS, M. **Minimal similarity accumulation attribute for fault enhancement.** In: PROCEEDINGS OF THE 84ND SEG ANNUAL MEETING, Denver, Oct. 2014.
- [8] FIGUEIREDO, A.; SILVA, F. B.; SILVA, P.; MILIDIÚ, R. L. ; GATTASS, M. **Seismic facies analysis approach to map 3d seismic horizons.** In: PROCEEDINGS OF THE 84ND SEG ANNUAL MEETING, Denver, Oct. 2014.

- [9] FIGUEIREDO, A.; SILVA, P. ; GATTASS, M. **Surface mapping using auto-adaptable similarity measures.** In: PROCEEDINGS OF THE 75TH EAGE CONFERENCE, London, June 2013.
- [10] FOMEL, S. **Predictive painting of 3d seismic volumes.** Geophysics, 75(4):A25—A30, 2010.
- [11] FRITZKE, B. **A growing neural gas network learns topologies.** In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 7. MIT Press, Cambridge MA, 1995.
- [12] GRAMSTAD, O.; BAKKE, J. ; SØNNELAND, L. **Seismic surface extraction using iterative seismic dna detection.** In: PROCEEDINGS OF THE 82ND SEG ANNUAL MEETING, Houston, Oct. 2012.
- [13] GROOT, P. D.; HUCK, A.; BRUIN, G. D. ; HEMSTRA, N. **The horizon cube: A step change in seismic interpretation.** The Leading Edge, 29(9), Apr. 2010.
- [14] HARTIGAN, J. A. **Clustering Algorithms.** Wiley, New York, 1975.
- [15] HOYES, J.; CHERET, T. **A review of "global"interpretation methods for automated 3d horizon picking.** The Leading Edge, 30(1), 2011.
- [16] KARIMI, P.; FOMEL, S. **Stratigraphic coordinates - a coordinate system tailored to seismic interpretation.** Geophysical Prospecting, 63(3), Mar. 2015.
- [17] LI, L.; MA, G. ; DU, X. **New method of horizon recognition in seismic data.** IEEE Geoscience And Remote Sensing Letters, 9(6), 2012.
- [18] LOMASK, J.; GUITTON, A. **Volumetric flattening: an interpretation tool.** The Leading Edge, 27(7), 2007.
- [19] MARROQUÍN, I. D.; BRAULT, J. J. ; HART, B. S. **A visual data-mining methodology for seismic facies analysis: Part 1 - testing and comparison with other unsupervised clustering methods.** Geophysics, 74(1):P1–P11, 2009.
- [20] MARTINS, L. O.; SILVA, P. M. ; GATTASS, M. **A method to estimate volumetric curvature attributes in 3d seismic data.** In: PROCEEDINGS OF THE 74TH EAGE CONFERENCE, Copenhagen, 2012.

- [21] O'MALLEY, S.; KAKADIARIS, I. **Towards robust structure-based enhancement and horizon picking in 3-d seismic data.** In: IEEE INT. CONF. ON COMPUTER VISION AND PATTERN RECOGNITION, p. 482—489, July 2004.
- [22] PARTRIDGE, M.; CALVO, R. **Fast dimensionality reduction and simple pca.** Intelligent Data Analysis, 2:292–298, 1997.
- [23] PAUGET, F.; LACAZE, S. ; VALDING, T. **A global interpretation based on cost function minimization.** In: PROCEEDINGS OF THE 79TH SEG ANNUAL MEETING, Houston, Oct. 2009.
- [24] QAYYUM, F.; GROOT, P. D. ; HEMSTRA, N. **Using 3d wheeler diagrams in seismic interpretation – the horizon cube method.** First Break, 30(1), Mar. 2012.
- [25] VERNEY, P.; PERRIN, M.; THONNAT, M. ; RAINAUD, J. F. **An approach of seismic interpretation based on cognitive vision.** In: PROCEEDINGS OF THE 70TH EAGE CONFERENCE, Rome, June 2008.
- [26] WU, X.; HALE, D. **Extracting horizons and sequence boundaries from 3d seismic images.** In: PROCEEDINGS OF THE 83TH SEG ANNUAL MEETING, Houston, Oct. 2013.
- [27] WU, X.; HALE, D. **Horizon volumes with interpreted constraints.** Geophysics, 80(2):IM21–IM33, 2015.
- [28] YU, Y.; KELLEY, C. ; MARDANOVA, I. **Automatic horizon picking in 3d seismic data using optical filters and minimum spanning tree.** In: PROCEEDINGS OF THE 81TH SEG ANNUAL MEETING, San Antonio, Sept. 2011.