

Eliana Leite Goldner

**Um algoritmo de menor caminho em
rastreamento de horizontes sísmicos**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Informática do Departamento de Informática da PUC-Rio

Orientador: Prof. Marcelo Gattass

Rio de Janeiro
Junho de 2014



Eliana Leite Goldner

Um algoritmo de menor caminho em rastreamento de horizontes sísmicos

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

Prof. Marcelo Gattass

Orientador

Departamento de Informática — PUC-Rio

Dr. Pedro Mário Cruz e Silva

Instituto Tecgraf — PUC-Rio

Prof. Hélio Côrtes Vieira Lopes

Departamento de Informática — PUC-Rio

Prof. Cristina Nader Vasconcelos

Universidade Federal Fluminense — UFF

Prof. José Eugenio Leal

Coordenador Setorial do Centro

Técnico Científico — PUC-Rio

Rio de Janeiro, 06 de Junho de 2014

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

Eliana Leite Goldner

Graduou-se Bacharel em Engenharia de Computação na Pontifícia Universidade Católica do Rio de Janeiro em dezembro de 2011.

Ficha Catalográfica

Goldner, Eliana Leite

Um algoritmo de menor caminho em rastreamento de horizontes sísmicos / Eliana Leite Goldner; orientador: Prof. Marcelo Gattass. — Rio de Janeiro : PUC-Rio, Departamento de Informática, 2014.

v., 67 f: il. ; 29,7 cm

1. Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Tese. 2. Computação Gráfica;. 3. Processamento de Imagens;. 4. Sísmica;. 5. Menor Caminho em Grafo;. 6. Rastreamento Horizontes. I. Gattass, Marcelo. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Agradecimentos

Ao meu orientador, Marcelo Gattass, pelo apoio, confiança e excelente orientação, que vai muito além desta pesquisa. À Pedro Mário pelas idéias e contribuições e pela oportunidade de trabalhar em um ambiente repleto de problemas interessantes.

Aos meus pais, Maria da Graça Goldner e Lidemar Goldner, pela educação que recebi e pelo amor e apoio irrestritos. À minha irmã, Evelyn Goldner, pelo carinho, incentivo e compreensão nos momentos de ausência e estresse.

Ao meu namorado, Walther Maciel, pelas excelentes idéias, palavras de conforto nos momentos mais difíceis e pela amizade e companheirismo sempre presentes em nossa relação. À sua família pelo carinho, especialmente à sua mãe, Margaret Maciel, por me receber e tratar como sua filha.

A todos os colegas do v3o2, especialmente ao Jéferson Coêlho pelas discussões e ajuda com a reconstrução do traço. A Vanessa Leite e Paula Rodrigues pela amizade e revisões de texto feitas quase imediatamente. Ao intérprete da Petrobras, Fernando Barbosa, pela interpretação do horizonte utilizado nos testes.

Aos meus amigos por compreenderem os longos períodos de ausência e por estarem sempre presentes nos momentos de descontração. Ao Carlos André Campos pelas idéias, apoio e paciência. A Fabíola Maffra pelo incentivo dado mesmo de longe.

À CAPES, ao CNPq, à FAPERJ e à PUC-Rio pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

À todos que acompanharam e participaram desta etapa da minha vida.

Muito obrigada.

Resumo

Goldner, Eliana Leite; Gattass, Marcelo. **Um algoritmo de menor caminho em rastreamento de horizontes sísmicos.** Rio de Janeiro, 2014. 67p. Dissertação de Mestrado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

A interpretação manual de um horizonte sísmico é um processo muito custoso em termos de tempo de trabalho do intérprete, o que incentiva a pesquisa de métodos automáticos, ou semi automáticos, de rastreamento. Dentre as propostas existentes baseadas em correlação, uma limitação conhecida é o uso de abordagens locais para definir as amostras pertencentes ao horizonte rastreado. Esse tipo de abordagem possui bom desempenho em dados onde não há a presença de falhas sísmicas, porém, nas regiões de baixa coerência, característica das regiões ruidosas ou de falhas, ao tomar uma decisão local o rastreador fica suscetível à propagação de erro. O objetivo deste trabalho é avaliar o uso de algoritmos de menor caminho em grafos para a solução do problema de rastreamento de horizontes sísmicos, afim de propor um método de caráter global que seja robusto a diferentes feições sísmicas.

Palavras-chave

Computação Gráfica; Processamento de Imagens; Sísmica; Menor Caminho em Grafo; Rastreamento Horizontes

Abstract

Goldner, Eliana Leite; Gattass, Marcelo. **Evaluation of a short path algorithm for seismic horizon tracking**. Rio de Janeiro, 2014. 67p. MSc. Dissertation — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

The manual interpretation of a seismic horizon is a time consuming process, which drives the research for automatic or semi automatic tracking methods. Among the known propositions that use correlation, there is a common limitation: the usage of local approaches to determine which samples belong to the horizon. This kind of approach performs well in data where there are no seismic faults. However, by using only local information, it is prone to error propagation in low coherency areas, which usually corresponds to fault regions. The goal of this work is to evaluate the performance of shortest path algorithms as a solution for the horizon tracking problem. It intends to propose a global method that is robust to different seismic features.

Keywords

Computer Graphics; Image Processing; Seismic; Shortest Path; Horizon Tracking

Sumário

1	Introdução	13
1.1	Motivação	13
1.2	Objetivo	14
1.3	Contribuições	15
1.4	Estrutura da Dissertação	15
2	Conceitos Básicos	16
2.1	Exploração Sísmica	16
2.2	Rastreamento de Horizontes	21
3	Trabalhos Relacionados	23
3.1	Brown, Morton e Whittle	23
3.2	Zinck, Donias e Guillon	24
3.3	Li, Ma e Du	25
4	Método Proposto	28
4.1	Extração de Picos	30
4.2	Orientação Local	30
4.3	Cálculo de Correlações	31
4.4	Construção do Grafo	33
4.5	Cálculo do Melhor Caminho	36
5	Resultados	38
5.1	Dado Sintético	38
5.2	Dado Real	39
5.3	Avaliação Visual	49
6	Conclusões	53
6.1	Trabalho Futuros	54
	Referências Bibliográficas	57
A	Medidas de Correlação	60
A.1	Correlação de Cossenos	60
A.2	Correlação de Pearson	60
A.3	Correlação de Spearman	61
A.4	Correlação de Bhat <i>et. al.</i>	62

Lista de figuras

1.1	O uso de informação local se torna um problema na presença de falhas. (a) ilustra o rastreamento de um horizonte utilizando uma abordagem gulosa, enquanto (b) ilustra o resultado esperado deste rastreamento.	14
2.1	Etapas de exploração baseadas na sísmica de reflexão.	16
2.2	Aquisição sísmica. Adaptada de Gerhardt (1).	17
2.3	Traço sísmico (esquerda), linha sísmica (centro) e volume sísmico (direita). Adaptado de Silva <i>et al.</i> (2)	19
2.4	Modelo de convolução. Adaptado de Gerhardt (1).	19
2.5	Modelo geológico. Adaptada de Robinson e Treitel (3).	20
2.6	Exemplo da forma como o horizonte se manifesta no dado sísmico. O ponto A pertence a um horizonte formado por picos e o ponto B pertence a um horizonte formado por vales de amplitude.	21
2.7	Horizontes diferentes alinhados por uma falha geológica. Adaptado de Silva (2).	22
4.1	Fluxograma do algoritmo proposto.	29
4.2	Parâmetros do algoritmo de rastreamento proposto.	29
4.3	Esquema ilustrando como o gradiente da fase instantânea pode ser utilizado para determinar a orientação local do horizonte sísmico. Adaptado de Pampanelli <i>et. al.</i> (4).	31
4.4	Dependência entre os tamanho da meia janela e a largura do pulso sendo caracterizado.	33
4.5	Diferentes formas de construção do subtraço utilizado pela medida de correlação.	33
4.6	Grafo criado a partir da seção sísmica. As arestas tracejadas incidem em nós situados em uma região de falha.	35
4.7	Estado final do grafo. Cada nó armazena o custo do melhor caminho encontrado e a amostra vizinha pertencente ao mesmo. As setas verdes mostram o caminho que representa o horizonte.	37
5.1	(a) exhibe o dado sintético utilizado no teste e (b) seu atributo de falha.	38
5.2	Diferentes etapas do algoritmo proposto. Em (c), exibimos o resultado do rastreamento no dado sintético.	39
5.3	Exemplos de entrada utilizada nos testes.	39
5.4	Diferença entre o grafo utilizado no primeiro caso de teste (a) e o grafo proposto (b).	40
5.5	Resultado da comparação entre as abordagens local (laranja) e baseada em menor caminho (verde).	42
5.6	Problema encontrado na primeira versão de teste: a função objetivo não é ótima ao longo do horizonte e o rastreamento troca de horizonte no meio do caminho.	43

5.7	Resultado da comparação entre a utilização de amostras consecutivas no traço discreto (verde) e utilização dos picos do sinal para construção do subtraço (vermelho).	44
5.8	Resultado da comparação entre construção do grafo sem utilização do atributo de falha (vermelho) e do método proposto (azul).	47
5.9	Conjunto de horizontes mapeados com o método proposto através de um domo de sal na <i>crossline</i> 780.	50
5.10	Conjunto de horizontes mapeados na <i>crossline</i> 940. A abordagem proposta se mostrou robusta à presença de falhas.	50
5.11	Conjunto de horizontes mapeados na <i>inline</i> 280. O método proposto se mostrou capaz de atravessar falhas com grandes rejeitos.	51
5.12	Conjunto de horizontes rastreados com o método proposto na <i>inline</i> 385.	51
5.13	Situações onde a abordagem proposta (verde) possui melhor desempenho que a local (vermelho).	52
A.1	Exemplo ilustrando o cálculo do coeficiente c_θ .	61
A.2	Exemplo ilustrando o cálculo do coeficiente r de Pearson.	62
A.3	Exemplo ilustrando o cálculo do coeficiente ρ de Spearman.	63
A.4	Exemplo ilustrando o cálculo do coeficiente κ .	64
A.5	Diagrama de dispersão de dois conjuntos correlacionados linearmente, onde $Y = 2 \times X$. Em (a) as quatro métricas resultam em coeficientes iguais a um e (b) ilustra o efeito de um <i>outlier</i> nos conjuntos em questão. Em (c) e (d) o mesmo conjunto é deslocado para $Y = 2 \times X + 10$.	65
A.6	Diagrama de dispersão de dois conjuntos correlacionados cubicamente. Neste caso, somente as medidas baseadas na ordenação relativa das amostras são capazes de capturar a correlação perfeita entre os conjuntos.	66
A.7	Diagrama de dispersão de dois diferentes pares de traços extraídos de uma seção real. Em (b) e (d) foram adicionados manualmente ruídos aos traços originais ilustrados por (a) e (c).	67

Lista de tabelas

3.1	Resumo e resultados dos trabalhos relacionados	27
5.1	Resultados da utilização de diferentes métricas de correlação aplicadas ao subtraço do sinal discreto.	45
5.2	Resultados da utilização de diferentes métricas de correlação aplicadas ao subtraço construído a partir dos picos do sinal sísmico.	46
5.3	Resumos dos resultados nas seções <i>inline</i> .	48
5.4	Resumos dos resultados nas seções <i>crossline</i> .	48
5.5	Resumos dos resultados nas seções <i>inline</i> + <i>crossline</i> .	48
5.6	Tempo de processamento médio em uma máquina Core i7 1.9Ghz, 8GB RAM.	49

Lista de símbolos

X	Amplitude sísmica
Y	Resultado da transformada de Hilbert; parte imaginária do traço sísmico complexo
ϕ	Fase instantânea
x	Direção <i>inline</i>
t	Tempo
h_2	Tamanho da meia janela vertical
X_a^i	i -ésima amostra do subtraço X_a
w_{ij}	Custo da aresta orientada do nó i para o nó j
α	Peso da orientação local no custo da aresta
c_j	Coeficiente de correlação entre o nó j e a semente
r	Coeficiente de correlação cruzada
ρ	Coeficiente de correlação de postos de Spearman
κ	Coeficiente de correlação de Bhat <i>et. al.</i>
c_θ	Coeficiente de correlação de cossenos

Simplicity is a great virtue but it requires hard work to achieve it and education to appreciate it. And to make matters worse: complexity sells better.

Edsger W. Dijkstra.

1

Introdução

A sísmica de reflexão é um método bastante utilizado no processo de exploração de hidrocarbonetos. Aliada a um processo de interpretação, ela auxilia a construção de um modelo geológico da subsuperfície. Para isso, é necessária a identificação de estruturas geológicas denominadas horizontes e falhas.

A interpretação manual de um horizonte é um processo muito custoso em termos de tempo de trabalho do intérprete, o que incentiva a pesquisa de métodos automáticos, ou semi automáticos, de rastreamento. A automatização deste processo pode reduzir o trabalho do intérprete a pequenos retoques no horizonte mapeado, diminuindo consideravelmente o seu tempo de trabalho.

Em Dorn (5), os métodos de rastreamento automático são classificados em duas categorias: métodos baseados em feições, que procuram por máximos (mínimos ou cruzamentos de zero) de amplitude, e métodos baseados em correlação, que usam uma medida de correlação para guiar o rastreamento de um traço sísmico para o outro. Sob essa perspectiva, propomos um método híbrido que, ao mesmo tempo em que procura por extremos, utiliza medidas de correlação para guiar o rastreamento.

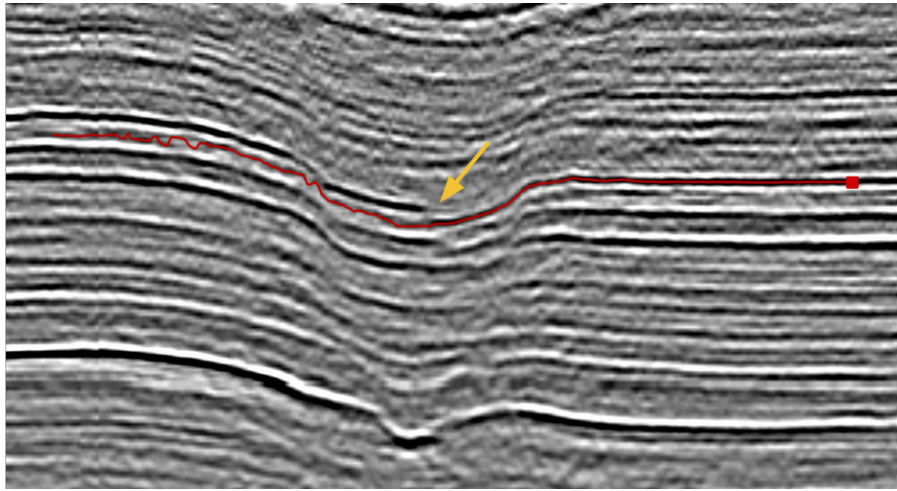
1.1

Motivação

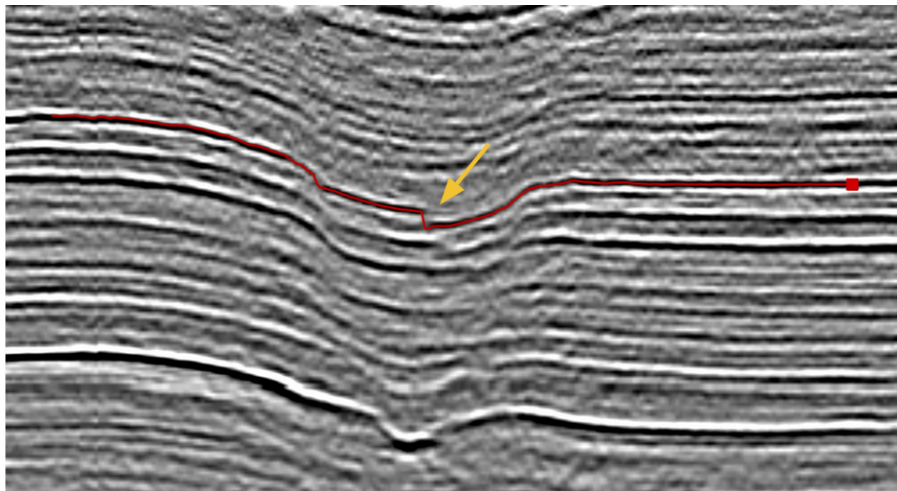
Diversas abordagens são utilizadas para automatizar o rastreamento de horizontes. Este trabalho se inspira naquelas que baseiam-se em sementes, amostras pré-definidas utilizadas para caracterizar o horizonte.

Um problema conhecido dessas abordagens é o uso de informação local para decidir quais amostras pertencem ao horizonte rastreado. Nesses casos, novas amostras são adicionadas a partir das amostras já descobertas nos traços vizinhos. Essa abordagem possui bom desempenho em volumes e seções onde não existe a presença de falhas sísmicas. Porém, nas regiões de baixa coerência, característica das regiões ruidosas ou de falhas, ao tomar uma decisão local, o rastreador pode adicionar uma amostra errada ao horizonte e utilizá-la como parâmetro de decisão para outras amostras, propagando o erro. A Figura 1.1

ilustra como o uso de informação local se torna um problema na presença de falhas. Em 1.1(a) é utilizada uma técnica gulosa que se perde na presença das mesmas e em 1.1(b) tem-se o resultado esperado.



(a)



(b)

Figura 1.1: O uso de informação local se torna um problema na presença de falhas. (a) ilustra o rastreamento de um horizonte utilizando uma abordagem gulosa, enquanto (b) ilustra o resultado esperado deste rastreamento.

1.2

Objetivo

O objetivo deste trabalho é avaliar a utilização de algoritmos de menor caminho em grafos para rastreamento de horizontes em seções 2D, de forma a incorporar um caráter global ao processo. Esperamos propor um método capaz de mapear horizontes mesmo na presença de falhas sísmicas.

1.3

Contribuições

Como contribuições desta dissertação destacam-se:

- A elaboração de um novo método para rastreamento de horizontes que não impõe restrições ao tipo de horizonte sendo mapeado.
- A avaliação do processo de rastreamento utilizando dados públicos e um horizonte gabarito fornecido por um intérprete.
- A proposta de uma nova maneira para cálculo da semelhança entre duas amostras de um volume sísmico, a qual produziu resultados mais estáveis.
- A comparação entre diferentes métricas de correlação.

1.4

Estrutura da Dissertação

Esta dissertação está estruturada em seis capítulos. O Capítulo 2 apresenta uma breve descrição dos conceitos básicos necessários para entendimento deste trabalho. No Capítulo 3 são apresentados trabalhos relacionados, que estudam o mesmo problema tratado nesta dissertação. O Capítulo 4 apresenta o método proposto neste trabalho e o Capítulo 5 os testes e resultados obtidos a partir de sua implementação. Por fim, o Capítulo 6 apresenta as conclusões deste estudo, bem como sugestões de trabalhos futuros.

2

Conceitos Básicos

O conteúdo interdisciplinar deste trabalho torna fundamental a introdução de alguns conceitos básicos da área de exploração sísmica. Neste capítulo, descrevemos, de maneira superficial, os processos de aquisição, processamento e interpretação de um dado sísmico, bem como definimos grandezas e estabelecemos notações.

O texto das próximas seções é fortemente baseado no capítulo correlato de teses anteriores desta linha de pesquisa (2) (1).

2.1

Exploração Sísmica

A sísmica de reflexão é um método indireto de exploração da subsuperfície da terra. O principal objetivo desse processo é compreender a disposição e comportamento das rochas na região de aquisição de dados, através da observação dos tempos de viagem das ondas sísmicas e suas variações de amplitude, frequência e forma de onda (6).

A utilização desse método traz vantagens como a capacidade de produzir imagens sísmicas que forneçam alta definição das feições geológicas, permitir a cobertura de grandes áreas de aquisição e menor custo quando comparado com um método de aquisição direto, como a perfuração de poços.

Segundo Robinson e Treitel (3), a exploração de hidrocarbonetos (óleo e gás) baseada em sísmica pode ser dividida em três etapas: aquisição, processamento e interpretação.

Apesar desse capítulo não ter o intuito de detalhar o método sísmico, nas subseções seguintes faremos uma breve descrição de cada etapa, a fim de

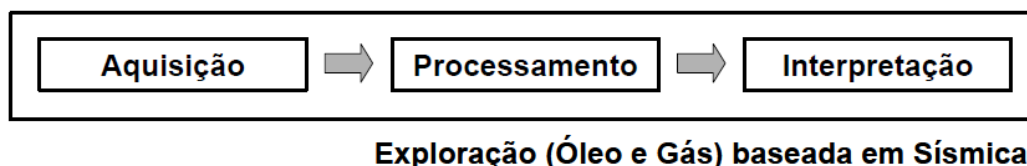


Figura 2.1: Etapas de exploração baseadas na sísmica de reflexão.

contextualizar o leitor para uma melhor compreensão do trabalho. Para uma visão mais aprofundada sugerimos consultar os trabalhos de Gerhardt (1), Machado (7), Robinson e Treitel (3) e Yilmaz (8).

2.1.1

Aquisição sísmica

O processo de aquisição sísmica baseia-se na emissão de onda, a partir de uma fonte colocada na superfície (mar ou terra), para o interior da terra e a captação das reflexões dessas ondas a partir das interfaces (rochas) existentes nos substrato. Em aquisições terrestres é comum usar explosões de dinamite como fonte; em aquisições marinhas são usados normalmente dispositivos pneumáticos como canhões de ar.

Uma vez gerada, a onda sísmica se propaga até alcançar uma interface entre duas camadas de rocha com valores de impedância acústica diferentes. Nesse momento, parte da onda é refratada e continua viajando para baixo e outra parte é refletida e retorna a superfície. A porção da energia refletida é proporcional à diferença de impedância acústica entre os dois meios.

A parte refletida da onda que retorna à superfície é captada pelos receptores (geofones em aquisições terrestres ou hidrofones em marinhas) e gravada nos sismógrafos. O sismógrafo armazena tanto o tempo de chegada da onda quanto a intensidade medida neste momento. Essa intensidade equivale à amplitude da onda refletida e é a principal feição observada nas seções sísmicas.

Após várias detonações variando a posição da fonte e dos receptores, todos os dados armazenados são enviados para serem processados. A Figura 2.2 ilustra os processos de aquisição terrestre e marinha.

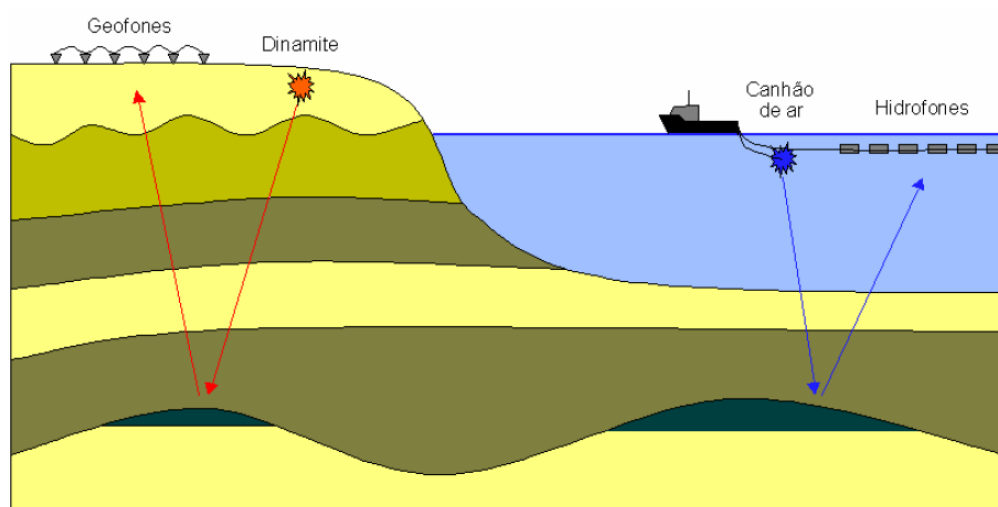


Figura 2.2: Aquisição sísmica. Adaptada de Gerhardt (1).

2.1.2

Processamento

A etapa que segue a aquisição é o processamento dos dados adquiridos, de forma a construir um modelo de dados o mais interpretativo possível. Nessa etapa, alguns erros inerentes ao levantamento sísmico são corrigidos e os dados são reorganizados para formarem uma grade tridimensional com uma amostra de amplitude sísmica em cada vértice da grade (voxel).

Duas das dimensões do conjunto de dados são direções espaciais e estão relacionadas com as posições das fontes e dos receptores. A terceira dimensão é resultado de uma transformação realizada nos dados durante o processamento que simula fonte e receptor em uma mesma posição. Graças a esta transformação, podemos considerar a terceira dimensão do conjunto de dados como sendo temporal e que a propagação da onda é feita apenas na direção vertical. Como podemos considerar que a fonte e o receptor estão na mesma posição na superfície, o tempo de cada amostra corresponde ao tempo que a onda leva para viajar até uma interface mais o tempo da volta à superfície.

Uma coluna de amostras com as mesmas coordenadas espaciais, variando apenas o tempo, é chamada de traço sísmico. Os máximos e mínimos da função de amplitude sísmica do traço são chamados de eventos sísmicos e correspondem aos eventos de reflexão.

A organização das amostras em um dado sísmico é mostrada na Figura 2.3. Do lado esquerdo temos a função de amplitudes do traço sísmico, na qual a única dimensão é a temporal (1D). No centro temos uma seção vertical do conjunto de dados formada por um conjunto de traços sísmicos, que é chamada de linha sísmica (2D), com uma dimensão espacial e a outra temporal. No caso dos dados sísmicos 3D (volume sísmico), formados por várias linhas sísmicas, 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.

Um modelo matemático interessante que descreve bem o efeito do processamento sísmico realizado sobre o dado é o modelo de convolução, ilustrado na Figura 2.4. Neste modelo consideramos a função de amplitude sísmica de cada traço do conjunto de dados como sendo o resultado da convolução de um impulso sísmico com uma função refletividade - a rigor, a função refletividade é uma distribuição de coeficientes de reflexão. Os coeficientes de reflexão são proporcionais à diferença de impedância acústica ¹ entre camadas geológicas adjacentes.

¹A impedância acústica de um material é a sua resistência à passagem do som.

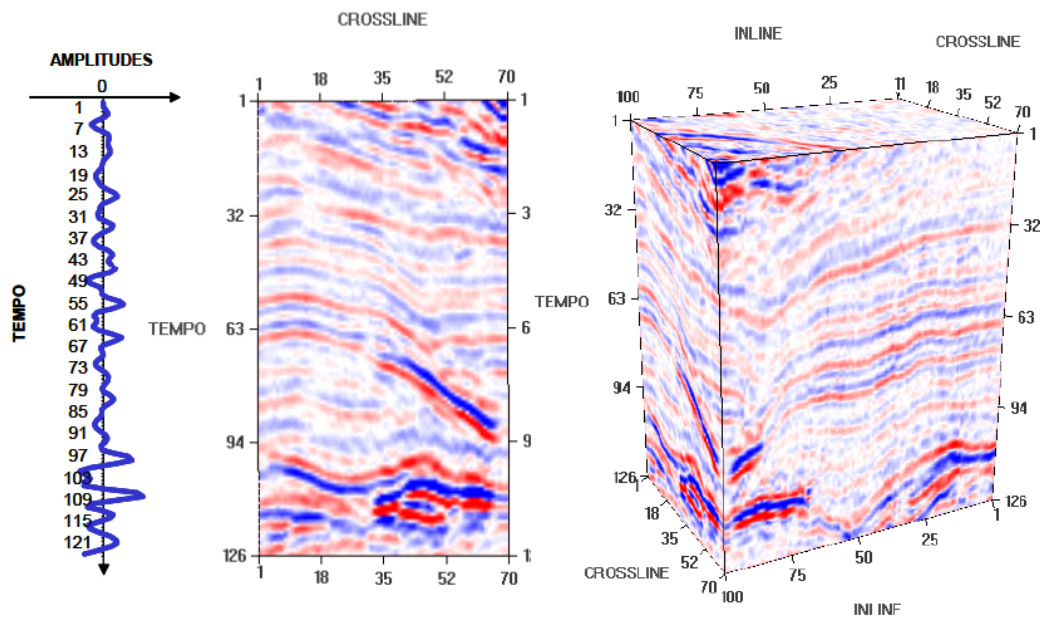


Figura 2.3: Traço sísmico (esquerda), linha sísmica (centro) e volume sísmico (direita). Adaptado de Silva *et al.* (2)

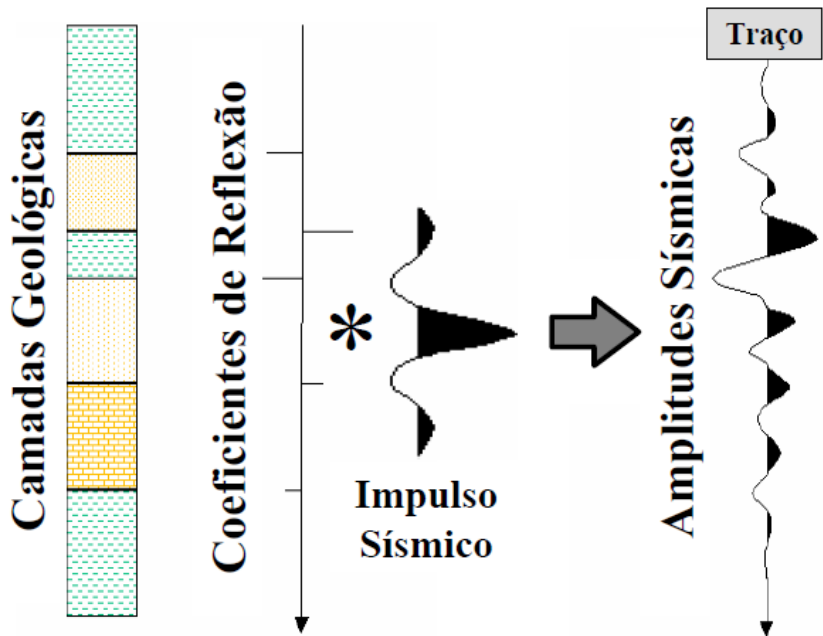


Figura 2.4: Modelo de convolução. Adaptado de Gerhardt (1).

2.1.3

Interpretação

Nesta etapa, o intérprete, em geral um geólogo ou geofísico, analisa os dados sísmicos e tenta criar um modelo que represente a geologia contida na área do levantamento. A Figura 2.5 mostra um modelo geológico que poderia ser resultante da interpretação de uma linha sísmica.

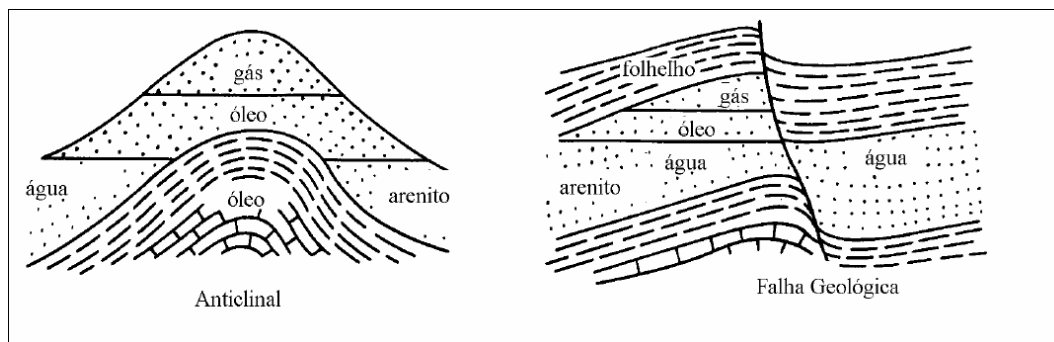


Figura 2.5: Modelo geológico. Adaptada de Robinson e Treitel (3).

A interpretação sísmica pode ser classificada, de acordo com o foco, em dois tipos: estrutural e estratigráfica. A interpretação estrutural basicamente busca identificar as camadas e discontinuidades do solo, enquanto a interpretação estratigráfica busca entender como cada camada foi depositada ao longo do tempo. O processo de interpretação estrutural pode ser dividido em duas principais etapas: a identificação de horizontes sísmicos e o mapeamento de falhas.

Sheriff (9) define um horizonte sísmico como sendo a superfície que separa duas camadas diferentes de rocha, onde, tal superfície (mesmo sem ter sido identificada) está associada com uma reflexão que se estende por uma grande área. Um horizonte se manifesta em um dado sísmico como uma série de eventos (picos ou vales de amplitudes sísmicas) que aparecem de forma consistente traço a traço, como ilustrado na Figura 2.6. Nesta figura, o ponto A representa um pico e o ponto B representa um vale. A partir deste momento, usaremos o termo pico para denominar os dois eventos, com o objetivo de facilitar a leitura do documento. Para distinguir as duas situações, utilizaremos o termo picos positivos para eventos como o ponto A e picos negativos para eventos como o ponto B. Os horizontes sísmicos também são chamados de refletores.

Falhas sísmicas podem ser definidas como uma quebra na continuidade original dos horizontes. Elas são fraturas que causam um deslocamento relativo das rochas, fazendo com que as mesmas percam sua continuidade lateral original. A presença de falhas na região geológica sendo mapeada é uma das principais dificuldades encontradas na identificação de horizontes.

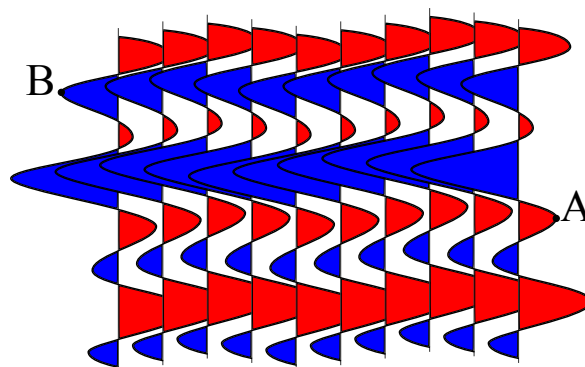


Figura 2.6: Exemplo da forma como o horizonte se manifesta no dado sísmico. O ponto A pertence a um horizonte formado por picos e o ponto B pertence a um horizonte formado por vales de amplitude.

2.2

Rastreamento de Horizontes

O rastreamento de um horizonte consiste em mapear quais os traços do volume sísmico o compõe, identificando em cada traço a amostra que pertence ao horizonte em questão. A identificação de cada amostra é feita a partir da análise de padrões de amplitude que aparecem na vizinhança vertical de cada voxel. Espera-se que esse padrão esteja presente nos traços que compõem o horizonte sendo rastreado.

As dificuldades encontradas no rastreamento de horizontes são, principalmente, a baixa razão sinal-ruído em algumas regiões do dado sísmico, a proximidade entre vários horizontes distintos, a própria complexidade da geologia e a presença de falhas geológicas. Grande parte dos algoritmos tende a confundir horizontes diferentes como se fossem um só durante o rastreamento. O problema de rastreamento de horizontes torna-se ainda mais complexo, pois o horizonte sísmico não é uma superfície de nível dentro do volume de amplitudes sísmicas, por isso métodos tradicionais de extração de superfícies como o *Marching Cubes* (10) não podem ser aplicados diretamente.

A presença de uma falha geológica é um grande complicador do processo, pois a mesma pode alinhar horizontes sísmicos distintos, tornando o rastreamento uma tarefa extremamente difícil. A Figura 2.7 mostra duas seções verticais do dado, (a) *inline* e (b) *crossline*. Em ambas temos a falha maior (linha tracejada preta) e três horizontes, “A”, “B” e “C”, de cima para baixo, interpretados manualmente. Como podemos ver na Figura 2.7(a), os horizontes “A” e “B” foram alinhados pela falha, o mesmo ocorrendo com os horizontes “B” e “C” na Figura 2.7(b).

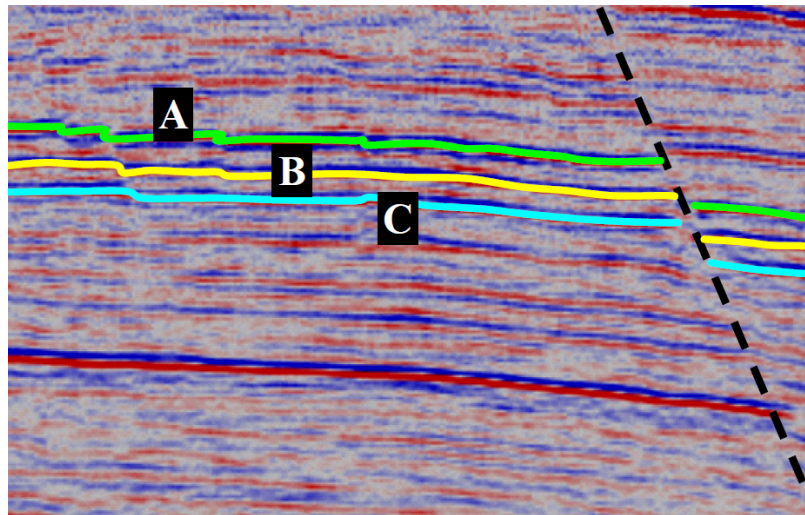
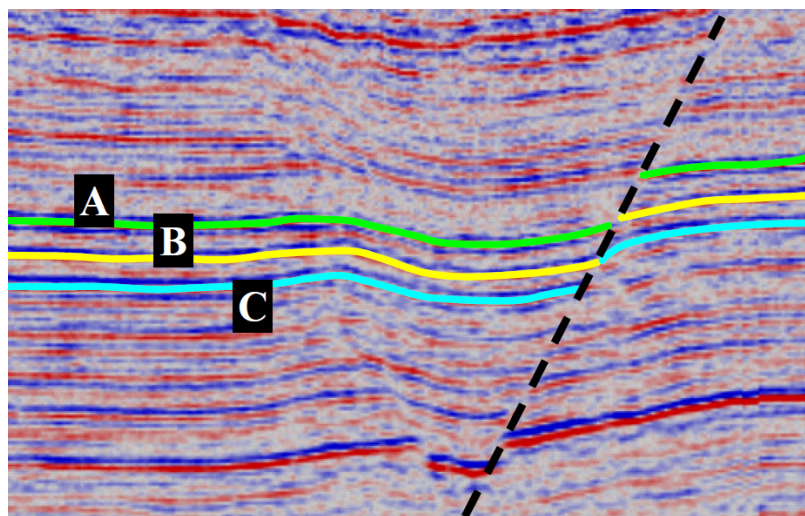
(a) *inline*(b) *crossline*

Figura 2.7: Horizontes diferentes alinhados por uma falha geológica. Adaptado de Silva (2).

3

Trabalhos Relacionados

A importância e a complexidade do mapeamento de horizontes incentivaram a pesquisa de diferentes algoritmos para a solução desse problema. Entre as abordagens estudadas encontram-se redes neurais, filtros de associação de dados probabilísticos, teoria dos grafos e crescimento de regiões.

No contexto de propostas para o rastreamento de horizontes 2D, foco desta pesquisa, destacam-se três trabalhos descritos nas seções seguintes.

3.1

Brown, Morton e Whittle

Em Brown, Morton e Whittle (11) os autores descrevem um método modelado para rastreamento do topo do sal, um horizonte caracterizado por grandes variações de amplitude e pequenas descontinuidades. Essa abordagem baseia-se no princípio de Fermat, o qual afirma que a onda, ao se propagar entre dois pontos, segue a trajetória que minimiza o tempo total de percurso.

Para isso, a imagem sísmica é convertida em uma seção de “pseudovelocidade”, por exemplo, adicionando-se um valor à amplitude sísmica para que não existam valores negativos. Então, a equação *eikonal* (12) associada é utilizada para calcular o percurso de menor tempo total entre dois picos selecionados pelo intérprete.

Os testes mostraram que a qualidade dos resultados depende da qualidade da seção de “pseudovelocidade” gerada. Os autores indicam que a seção gerada deve possuir uma boa amostragem e alto contraste de amplitudes, caso contrário, o rastreador preferirá pular de um pico fraco para eventos vizinhos mais fortes.

Apesar dessa dependência, os autores destacam a eficácia do método quando o evento sendo rastreado possui pequenas descontinuidades. Nesse caso, o método conecta com linhas retas as áreas de falhas, pois busca o caminho de menor tempo, gerando horizontes conexos e suaves.

Além disso, o trabalho apresenta uma extensão do método para 3D através de uma triangulação de *Delaunay* das sementes de entrada. Após a

geração dos triângulos, o algoritmo descrito é executado ao longo de cada aresta para formar a superfície 3D.

Embora os resultados apresentados tenham sido satisfatórios, o método proposto pelos autores é modelado para rastreamento do topo do sal. Contudo, a manipulação de horizontes que delimitam o topo e a base do sal não é equivalente à manipulação de horizontes em estrutura de camadas. Dessa forma, não é seguro afirmar que o método produz bons resultados em um ambiente diferente daquele descrito pelos autores.

3.2

Zinck, Donias e Guillon

Zinck, Donias e Guillon (13) apresentam um método para rastreamento de horizontes com descontinuidades causadas por falhas quase verticais. A abordagem proposta requer o conhecimento de dois pontos delimitando o horizonte, bem como a posição da descontinuidade e seu rejeito¹.

Nesse trabalho, o horizonte é modelado como um segmento de curva representado por uma função τ definida entre duas abscissas x_1 e x_2 , com uma descontinuidade conhecida na abscissa $\alpha \in]x_1, x_2[$ e rejeito definido pela equação 3-1:

$$C_c = \lim_{x \rightarrow \alpha^+} \tau(x) - \lim_{x \rightarrow \alpha^-} \tau(x) \quad (3-1)$$

O rastreamento é feito, então, resolvendo uma equação diferencial parcial não linear. Para isso, os autores descrevem uma solução iterativa baseada na equação linear de *Poisson* com condições de contorno incrementais de *Dirichlet*.

Como, em casos reais, a posição e o rejeito da falha dificilmente são conhecidos, os autores apresentam também uma proposta para estimar essas informações. A solução apresentada testa diversos pares de candidatos (η_α, C_d) , amostrados em um intervalo pré determinado, e utiliza um critério de coerência para encontrar a solução ótima.

Dois testes foram realizados para validar o método proposto. O primeiro tinha o objetivo de testar a abordagem de rastreamento proposta e foi executado em uma imagem sintética, onde todos os parâmetros do algoritmo em questão eram conhecidos. O segundo teste foi realizado em um dado real com o objetivo de validar o rastreamento e, também, o método de

¹O rejeito de uma falha geológica caracteriza o movimento relativo dos blocos separados pela mesma. Ele corresponde à distância que separa dois pontos, situados em blocos opostos, que encontravam-se inicialmente juntos.

estimativa da posição e rejeito da falha. Os resultados apresentados exibiram bom desempenho tanto no dado sintético, como no dado real.

Assim como em Brown *et al.* (11), o método proposto pelos autores não foi modelado para rastrear um evento qualquer e sim uma situação particular, na qual o horizonte desejado possui uma descontinuidade quase vertical. A abordagem proposta não prevê em sua modelagem o rastreamento de horizontes contínuos ou com falhas não verticais. Os autores não deixam claro o comportamento do rastreador quando $C_c = 0$, porém, mesmo que ele seja capaz de identificar corretamente o horizonte, muito processamento será utilizado para inferir a posição de uma descontinuidade inexistente, comprometendo a eficiência do método.

3.3

Li, Ma e Du

Li, Ma e Du (14) descrevem um método para reconhecimento de horizontes que utiliza uma combinação de derivada horizontal com operadores morfológicos.

Os autores utilizam um valor limiar para filtrar a derivada horizontal do dado sísmico, de modo a remover picos negativos. Após a filtragem, o dado é utilizado para calcular duas novas imagens através da aplicação das operações morfológicas de erosão e dilatação. Em seguida, uma nova imagem é gerada a partir da razão entre o resultado da operação de erosão e o resultado da operação de dilatação. Nessa nova imagem, as amplitudes de sinais fortes e fracos estarão balanceadas. Uma vez balanceadas, as amplitudes máximas de cada *wavelet* correspondem às posições dos horizontes na imagem original.

Para avaliar a eficácia do método proposto, os autores comparam seus resultados à duas outras abordagens. A primeira é baseada no operador de *Canny* e a segunda é uma rede neural (15). A proposta foi testada em um dado sintético com três horizontes, no qual foi adicionado ruído gaussiano, e em um dado real. Dentre os três métodos avaliados, a proposta exibida pelos autores foi a que melhor conseguiu destacar os horizontes. Essa abordagem se mostrou eficaz para balancear a amplitude de horizontes fortes e fracos e realçar a localização de horizontes.

Através da análise dos resultados exibidos nesse trabalho, é possível concluir que o método proposto trata a imagem como um todo e é eficaz em balancear a amplitude dos horizontes quando a seção sísmica de entrada possui baixa coerência lateral, ou seja, grandes variações de amplitude ao longo de um mesmo evento. Porém, não parece haver distinções entre os máximos de

amplitude encontrados, deixando remanescente a necessidade de outro processo para extrair, separadamente, cada um dos horizontes da imagem.

A Tabela 3.1 apresenta um resumo dos trabalhos descritos neste capítulo e seus respectivos resultados.

Com o intuito de simplificar o processo de rastreamento, os métodos neste capítulo assumem hipóteses que impõem restrições ao ambiente em que devem ser executados. Dessa forma, seus bons desempenhos estão associados às condições assumidas para cada abordagem. As limitações impostas através do tipo de horizonte sendo rastreado ou das condições da seção de amplitude de entrada nos indicam que nenhum dos métodos apresentados possuiria bom desempenho no rastreamento de um caso genérico.

Em busca de produzir um algoritmo capaz de identificar corretamente cada evento, independente das condições geológicas onde ele ocorreu, optamos por utilizar um enfoque embasado em grafos e algoritmos clássicos, de forma a não impor restrições ao ambiente de execução.

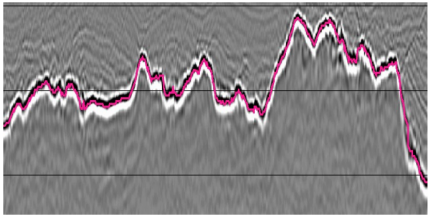
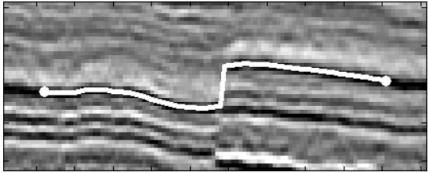
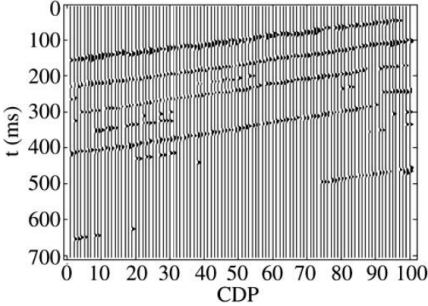
Trabalho	Descrição	Resultado
Brown, Morton e Whittle (11)	Calcula o caminho de menor tempo entre duas amostras em uma seção de “pseudovelocidade”. Foi modelado para rastreamento do topo do sal.	
Zinck, Donias e Guillon (13)	Baseado na equação linear de Poisson com condições de contorno incrementais de Dirichlet, possui o objetivo de rastrear horizontes com descontinuidades devido à presença de falhas.	
Li, Ma e Du (14)	O método é uma combinação de derivada horizontal e operações morfológicas com o objetivo de balancear os sinais de amplitude para determinar a posição dos horizontes.	

Tabela 3.1: Resumo e resultados dos trabalhos relacionados

4

Método Proposto

Neste capítulo apresentamos uma proposta de rastreamento de horizontes 2D de caráter global. Para tal, modelamos o rastreamento de horizontes como um problema de otimização que busca maximizar a similaridade entre as amostras rastreadas e a semente. Esperamos que, ao evitar o uso de medidas de correlação limitadas aos traços vizinhos de cada amostra, o rastreador seja capaz de transpor feições geológicas como falhas, por exemplo.

Nesta proposta, definimos o horizonte sísmico discreto como o caminho de melhor custo entre dois nós de um grafo direcionado acíclico. O peso de cada aresta é dado pela similaridade entre o nó destino e a amostra semente.

Para encontrar o horizonte desejado, utilizamos o algoritmo de menor caminho em DAGs (16) e calculamos o caminho de maior semelhança em um grafo construído a partir da seção sísmica. A Figura 4.1 ilustra o fluxograma do método proposto.

O método recebe como entrada a imagem de amplitudes sísmicas, uma imagem de atributo identificador de falhas e quatro outros parâmetros numéricos: um ponto inicial, um ponto final, o grau de liberdade e a meia janela vertical de correlação.

O ponto inicial é a semente do algoritmo e é utilizado como guia das medidas de semelhança. O ponto final é a amostra que, quando atingida, irá parar o processo de rastreamento. O grau de liberdade, ilustrado pela Figura 4.2(a), é utilizado para definir a conectividade entre amostras de traços vizinhos. Aliado ao atributo identificador de falhas, ele define em quantas amostras, para cima e para baixo, incidirão arestas de saída de cada nó. A meia janela vertical, por sua vez, é utilizada para calcular a correlação entre duas amostras. Ela está ilustrada pela Figura 4.2(b) e seu valor define quantas amostras, para cima e para baixo da amostra de interesse, definirão o subtraço utilizado para o cálculo da correlação.

As seções a seguir descrevem cada uma das etapas ilustradas na Figura 4.1. Nas Seções 4.1 e 4.2 descrevemos o processo de pré-processamento do método. Nessa etapa são calculadas as características do dado de entrada que independem dos parâmetros mencionados anteriormente.

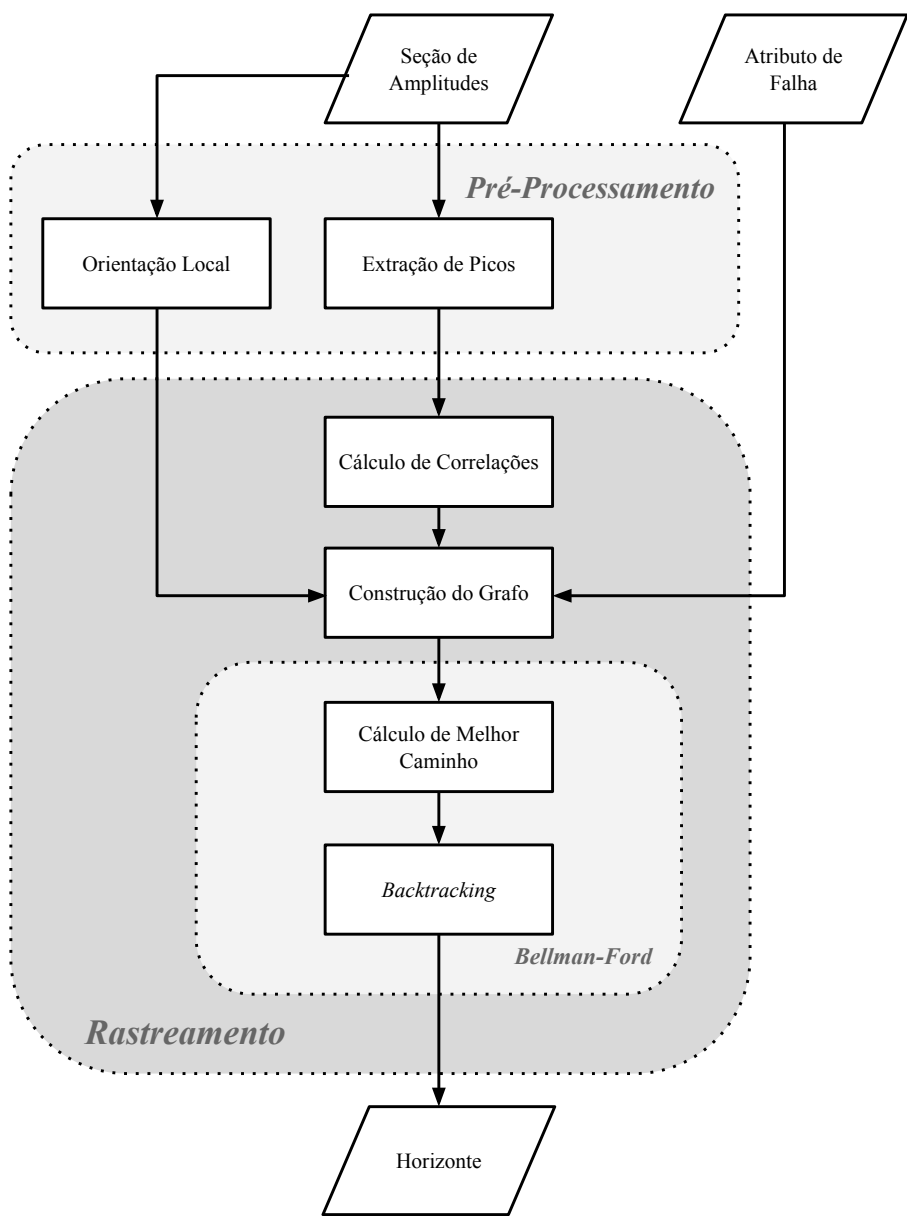
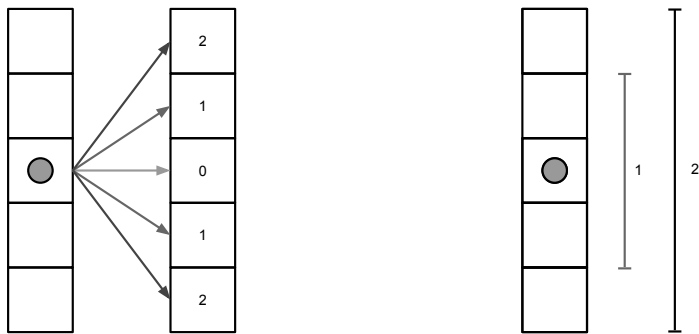


Figura 4.1: Fluxograma do algoritmo proposto.



(a) Grau de liberdade (b) Meia janela vertical

Figura 4.2: Parâmetros do algoritmo de rastreamento proposto.

Em seguida, definimos cada etapa do processo de rastreamento proposto. Na Seção 4.3 descrevemos como é calculada a semelhança entre uma amostra e a semente. Na Seção 4.4 definimos o processo de construção do grafo e na Seção 4.5 descrevemos como, efetivamente, é calculado o melhor caminho.

4.1

Extração de Picos

A primeira característica extraída na fase de pré-processamento do dado é a posição de cada pico do sinal sísmico. Essa informação será utilizada para calcular a semelhança entre duas amostras da seção.

Para estimar a posição e a magnitude de cada extremo do sinal, reconstruímos o sinal discreto com segmentos polinomiais cúbicos. A cada par de amostras a_i e a_{i+1} , utilizamos seus vizinhos a_{i-1} e a_{i+2} para aproximar um polinômio do tipo $p(x) = ax^3 + bx^2 + cx + d$.

Após o cálculo do conjunto de polinômios que descrevem o traço contínuo, os mesmos são consultados para identificar quais pares de amostras representam um pico do sinal sísmico.

Além disso, essa informação é utilizada para ajustar a posição da amostra semente para o pico do sinal mais próximo à posição definida pelo usuário. Dessa forma, garantimos que a amostra inicial pertence a um evento sísmico.

4.2

Orientação Local

Esta etapa do pré-processamento consiste em estimar a orientação local de cada amostra da seção de amplitudes sísmicas. Essa estimativa é utilizada na função de custo com o objetivo de evitar que o rastreador confunda horizontes vizinhos similares.

Os vetores de orientação local de cada amostra são calculados a partir do gradiente da fase instantânea. Dessa maneira, como descrito por Silva *et al.* (17), os horizontes podem ser interpretados como curvas de nível e, por isso, seu gradiente pode ser utilizado para calcular o vetor de orientação. A Figura 4.3 ilustra como o vetor perpendicular ao gradiente da fase instantânea pode ser utilizado para estimar a orientação local.

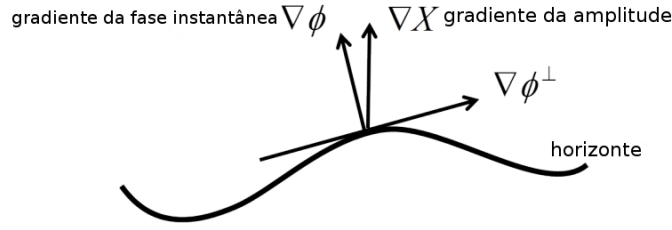


Figura 4.3: Esquema ilustrando como o gradiente da fase instantânea pode ser utilizado para determinar a orientação local do horizonte sísmico. Adaptado de Pampanelli *et. al.* (4).

Apesar da fase instantânea possuir descontinuidades em π e $-\pi$ que causariam artefatos indesejados no campo vetorial, Barnes (18) mostra que é possível calcular seu gradiente através de suas equações analíticas. O cálculo do gradiente nas direções x e t está descrito, respectivamente, nas Equações (4-1a) e (4-1b), onde X é a amplitude sísmica e Y sua transformada de Hilbert 1D (19).

$$\frac{\partial \phi}{\partial x} = \frac{1}{X^2 + Y^2} \left(X \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial x} Y \right) \quad (4-1a)$$

$$\frac{\partial \phi}{\partial t} = \frac{1}{X^2 + Y^2} \left(X \frac{\partial Y}{\partial t} - \frac{\partial X}{\partial t} Y \right) \quad (4-1b)$$

4.3

Cálculo de Correlações

A primeira etapa do processo de rastreamento consiste em calcular uma imagem com a correlação entre cada amostra e a semente. Para isso, os picos previamente extraídos do sinal são utilizados para construir subtraços ao redor de cada amostra de interesse. Com esses subtraços são calculados os coeficientes de correlação, descritos na Seção 4.3.1, entre cada pico e a semente. Essa informação é então salva no *pixel* correspondente ao pico em questão. Uma vez que sabemos que os horizontes correspondem aos extremos do sinal, as demais amostras são inicializadas com o coeficiente de correlação igual a zero.

4.3.1

Coefficiente de Pearson

O coeficiente de correlação de Pearson (20), r , também chamado de correlação cruzada, é a medida de similaridade comumente utilizada pelos métodos de rastreamento automático (5). Ela é uma medida de correlação

linear entre os vetores de entrada X_a e X_b . Seu valor encontra-se no intervalo $[-1, 1]$. O valor de r positivo indica que X_b cresce quando o valor de X_a cresce. Se r é igual a um, os conjuntos possuem correlação positiva perfeita, o que implica que a relação entre as amostras de cada conjunto pode ser descrita por uma equação linear. Por outro lado, se r é negativo, o valor de X_b decresce quando X_a cresce. O valor de r igual a zero indica que os conjuntos não demonstram correlação linear entre si.

A correlação cruzada pode ser definida pela Equação (4-2), onde X_a e X_b representam as sequências de entrada, \bar{X}_a e \bar{X}_b , suas médias aritméticas e X_a^i e X_b^i , suas i -ésimas amostras, respectivamente.

$$r = \frac{\sum_{i=1}^n (X_a^i - \bar{X}_a)(X_b^i - \bar{X}_b)}{\sqrt{\sum_{i=1}^n (X_a^i - \bar{X}_a)^2} \sqrt{\sum_{i=1}^n (X_b^i - \bar{X}_b)^2}} \quad (4-2)$$

4.3.2

Construção do Subtraço

As sequências utilizadas pela medida de correlação são construídas a partir da meia janela vertical, h_2 , recebida como parâmetro. Com este valor, são construídos subtraços de tamanho $2h_2 + 1$ formados pela amostra de interesse e $2h_2$ outras amostras do mesmo traço.

Em geral, os métodos que baseiam-se em correlação utilizam uma janela de amostras consecutivas no traço para extrair o subtraço utilizado no cálculo do coeficiente de Pearson (20). Isso pode não ser um problema ao comparar subtraços vizinhos, porém, como propomos incorporar uma abordagem global, a semelhança é sempre medida em comparação com a semente. Neste caso, o uso de amostras consecutivas pode ser prejudicial caso a largura do pulso de um mesmo horizonte varie.

A Figura 4.4 ilustra o problema em questão. Uma meia janela $h_2 = 2$ é capaz de representar a largura do pulso destacado no traço A, porém, no traço B, este valor extrapola os limites do pulso, prejudicando a comparação entre estas duas amostras similares. Essa abordagem cria uma dependência entre o tamanho da meia janela e a largura do pulso sendo analisado.

Para tentar eliminar essa dependência, optamos por utilizar somente os picos do sinal para construir cada subtraço. Dessa forma, evitamos o efeito da discretização ilustrado na Figura 4.4 e construímos os subtraços de maneira mais robusta.

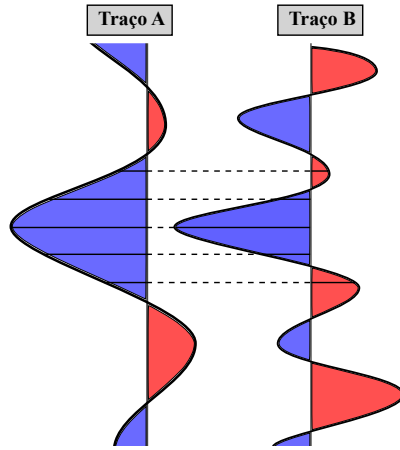


Figura 4.4: Dependência entre o tamanho da meia janela e a largura do pulso sendo caracterizado.

A Figura 4.5 ilustra as diferentes formas de construção do subtraço. As amostras vermelhas representam picos positivos, amostras azuis representam picos negativos e amostras cinzas não correspondem a nenhum extremo do sinal. Na Figura 4.5(a) o subtraço é construído a partir de amostras consecutivas, enquanto na Figura 4.5(b) somente os picos são utilizados.

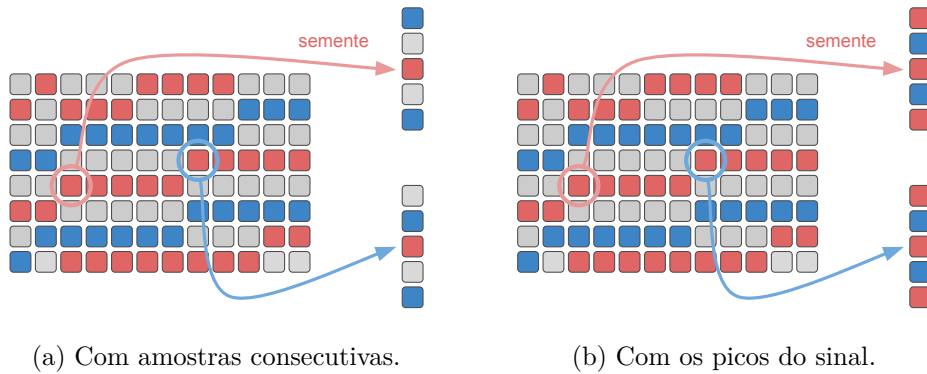


Figura 4.5: Diferentes formas de construção do subtraço utilizado pela medida de correlação.

4.4

Construção do Grafo

O grafo utilizado é construído implicitamente durante o cálculo do melhor caminho, porém a fim de clareza, essas duas etapas são descritas separadamente. Na Seção 4.4.1 definimos como é feita a inserção de arestas em um grafo onde cada *pixel* da imagem sísmica é interpretado como um nó. Na Seção 4.4.2, por sua vez, definimos como é calculado o peso de cada aresta do grafo.

4.4.1

Definição da Conectividade

Nesta etapa, definimos quais arestas conectam os nós do grafo. Consideramos, para tal, todos os *pixels* da imagem sísmica, inclusive aqueles que não correspondem a picos do sinal, de forma a permitir a construção de uma curva final mais suave. Esperamos que a correlação nula associada a estes nós seja suficiente para desencorajar sua inclusão desnecessária no melhor caminho.

A conectividade entre os nós do grafo é definida em função do grau de liberdade, L , recebido como parâmetro e do atributo identificador de falhas.

Cada nó, a partir da semente, é conectado a no máximo $n = 2L + 1$ nós com arestas orientadas no sentido do ponto final. Esse conjunto de arestas de saída incide em um conjunto de nós consecutivos centrados verticalmente com o nó de origem. Em regiões de falha, são inseridas n arestas de saída. Porém, em pontos onde o atributo de falha está abaixo de um limiar, as arestas são adicionadas de forma a evitar o cruzamento de horizontes vizinhos. Para isso, a partir do vizinho imediatamente ao lado, os nós são consultados para cima e para baixo adicionando-se arestas até que haja uma inversão na polaridade do sinal.

Essa restrição pode interromper o crescimento do grafo antes de alcançar o traço final. Isso aconteceria caso todos os vizinhos de cada nó possuísem polaridade inversa à semente. Nesse caso, buscamos nos traços seguintes um vizinho válido, até que se chegue ao traço final.

Esse conjunto de regras garante que a solução só possui uma amostra por traço, um requisito do problema em questão. A Figura 4.6 ilustra um grafo construído com $L = 1$. Nela, a cor de cada nó representa o sinal da amplitude da respectiva amostra na imagem sísmica e a semente está destacada com uma borda preta. As arestas tracejadas incidem sobre uma região de falha e não possuem restrições em relação a amplitude do nó destino, enquanto as demais só conectam nós de amplitudes compatíveis com a semente.

4.4.2

Função de Custo

O peso de cada aresta do grafo é definido por uma função de custo que associa um número real a um par de *pixels* da imagem sísmica. Este valor busca refletir a probabilidade dos dois *pixels* em questão estarem em um mesmo evento sísmico. Esperamos que, quanto maior o valor da função objetivo para dois pontos, maiores as chances destes pontos pertencerem ao mesmo horizonte.

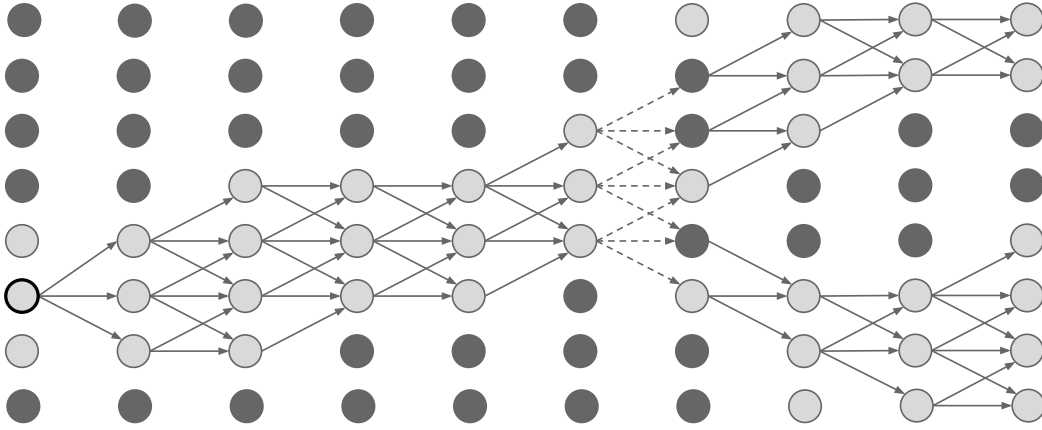


Figura 4.6: Grafo criado a partir da seção sísmica. As arestas tracejadas incidem em nós situados em uma região de falha.

Duas importantes características de um horizonte são consideradas na função de custo proposta: seu mergulho¹ e sua identidade vertical. A primeira é medida pela orientação local, e é utilizada pois, deseja-se que o caminho rastreado seja coerente com a orientação aparente do dado. A segunda é sua assinatura vertical, o padrão de amplitudes que representam as diferentes camadas depositadas ao longo do tempo. Essa característica é representada pelo coeficiente de correlação.

Dessa forma, o custo associado à cada aresta do grafo é dado pela Equação (4-3), onde w_{ij} é o custo de aresta orientada do nó i para o nó j , θ é o ângulo entre a aresta e a orientação local do nó i , c_j é o coeficiente de correlação entre o nó j e a semente e, α é o peso da orientação local no custo da aresta, com valor entre 0 e 1. O valor de w_{ij} está no intervalo $[-1, 1]$.

$$w_{ij} = \alpha \cos \theta + (1 - \alpha)c_j \quad (4-3)$$

A função de custo proposta busca misturar os benefícios das abordagens globais e locais. Sua primeira parcela, de caráter local, incorpora informação a respeito do *pixel* corrente e seu vizinho imediato e está associada à aresta que conecta estes dois nós. A segunda parcela, por sua vez, possui caráter global e está associada somente ao nó corrente, uma vez que compara este *pixel* com a semente definida para caracterizar o horizonte.

¹Ângulo com o qual uma camada de rocha se desvia da horizontal.

4.5

Cálculo do Melhor Caminho

Uma vez definido o grafo, o próximo passo consiste em calcular o custo do melhor caminho da semente ao ponto final. Este processo pode ser dividido em duas principais etapas: o percorrimento e o *backtracking*.

4.5.1

Percorrimento

A primeira etapa consiste em encontrar, para cada nó do grafo, o custo do melhor caminho até a semente. Para isso, o grafo é percorrido utilizando-se o princípio de relaxamento (16), onde uma aproximação do melhor custo entre dois nós é gradualmente substituída por valores mais precisos até alcançar o ótimo.

Como o grafo construído é direcionado e acíclico, podemos facilmente percorrê-lo em ordem topológica (21) a partir da semente. Para cada nó visitado i , o custo do melhor caminho encontrado até cada um de seus vizinhos, j , é atualizado segundo a Equação (4-4), onde opt_i é o custo do melhor caminho da semente ao nó i e w_{ij} é o peso da aresta direcionada (i, j) .

$$opt_j = \begin{cases} opt_j, & \text{se } opt_j \geq opt_i + w_{ij} \\ opt_i + w_{ij}, & \text{caso contrário} \end{cases} \quad (4-4)$$

Além disso, todo vizinho cujo melhor caminho encontrado inclui a aresta (i, j) , armazena uma referência para o nó i .

4.5.2

Backtraking

O último passo do rastreamento consiste em definir os nó pertencentes ao melhor caminho encontrado, que vai da semente ao ponto final. Para isso, percorremos o grafo a partir do ponto final seguindo a informação salva no passo anterior, que aponta para o vizinho que pertence ao melhor caminho encontrado. Todos os nós consultados são salvos e ao alcançar o ponto inicial o caminho de melhor custo é conhecido.

Neste passo, opcionalmente, pode ser feito um ajuste *subpixel* dos pontos rastreados. Como a amostragem da imagem sísmica não segue os extremos do sinal, uma aproximação do ponto rastreado para o pico mais próximo pode fornecer uma curva mais suave e precisa.

A Figura 4.7 ilustra o estado final do grafo. Cada nó armazena o custo do melhor caminho partindo da semente até a si. Sendo assim, só seria necessário o

conhecimento do traço final do rastreamento e o *backtracking* poderia começar a partir da amostra com maior semelhança acumulada. Porém, a fim de evitar que o rastreamento confunda horizontes muito próximos, o *backtracking* é realizado a partir de um ponto final fornecido como entrada, dando mais robustez ao método.

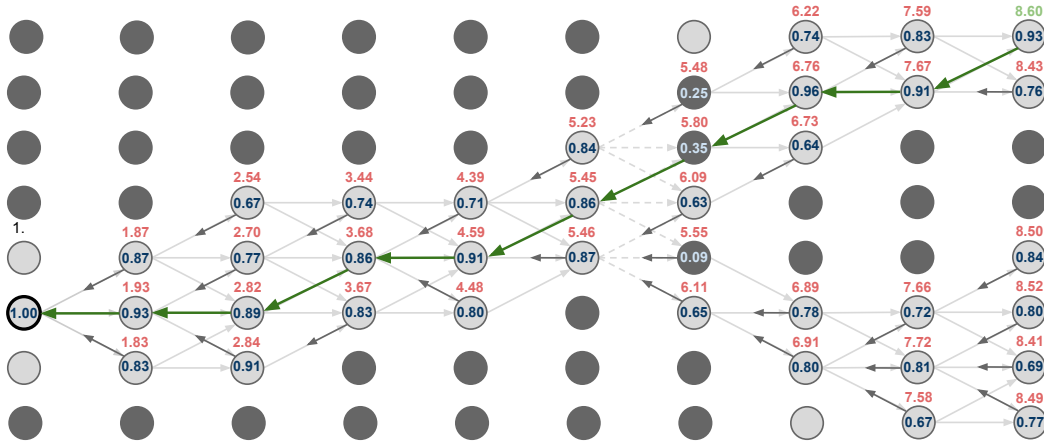


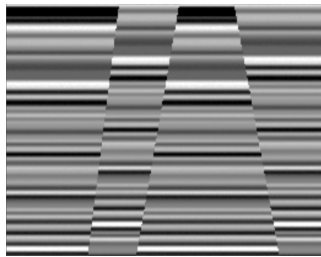
Figura 4.7: Estado final do grafo. Cada nó armazena o custo do melhor caminho encontrado e a amostra vizinha pertencente ao mesmo. As setas verdes mostram o caminho que representa o horizonte.

5 Resultados

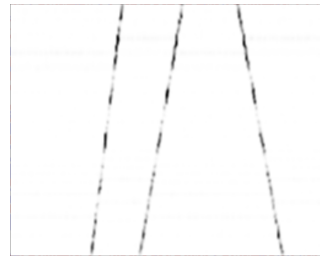
Neste capítulo apresentamos cada um dos testes realizados para avaliar a utilização de um algoritmo de menor caminho para o rastreamento de horizontes.

5.1 Dado Sintético

O primeiro teste realizado tem por objetivo a prova de conceito do método proposto. Para tal, criamos um dado sintético onde o resultado esperado é conhecido. Para a criação do dado sintético, extraímos um traço do volume *F3 Block* disponibilizado no site do Opendtect (22). Então, repetimos esse traço lado a lado de forma a gerar uma imagem. Em seguida, deslocamos verticalmente alguns desses traços de forma a simular uma falha. Finalmente, adicionamos ruído branco à imagem. A Figura 5.1 exibe a seção de amplitudes criada e seu respectivo atributo identificador de falhas.



(a) Seção de amplitudes



(b) Atributo de falha

Figura 5.1: (a) exibe o dado sintético utilizado no teste e (b) seu atributo de falha.

A Figura 5.2 exibe os resultados do rastreamento de um horizonte no dado sintético. Em cada imagem, a escala de cores varia de branco, o menor valor, a preto. Em 5.2(a) exibimos a imagem de correlações calculada. Note que somente os extremos do sinal sísmico possuem um coeficiente de correlação calculado. A Figura 5.2(b) ilustra o grafo construído para este rastreamento. A intensidade de cada *pixel* é inversamente proporcional ao custo do melhor caminho encontrado até o mesmo. É possível perceber que o grafo só é

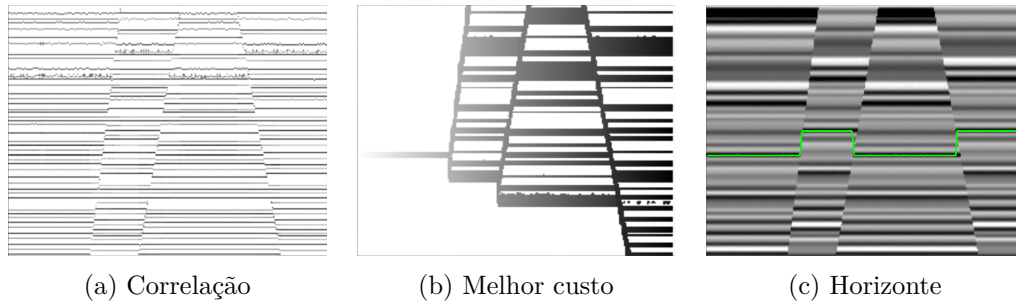


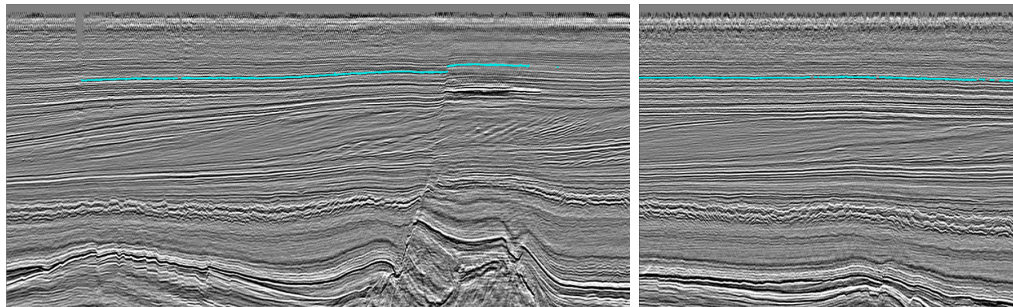
Figura 5.2: Diferentes etapas do algoritmo proposto. Em (c), exibimos o resultado do rastreamento no dado sintético.

expandido ao atingir uma região de falha, onde são abertas ramificações para encontrar o melhor caminho. Finalmente, a Figura 5.2(c) exibe o horizonte rastreado. Como esperado, o método é capaz de pular as discontinuidades encontradas ao longo do horizonte.

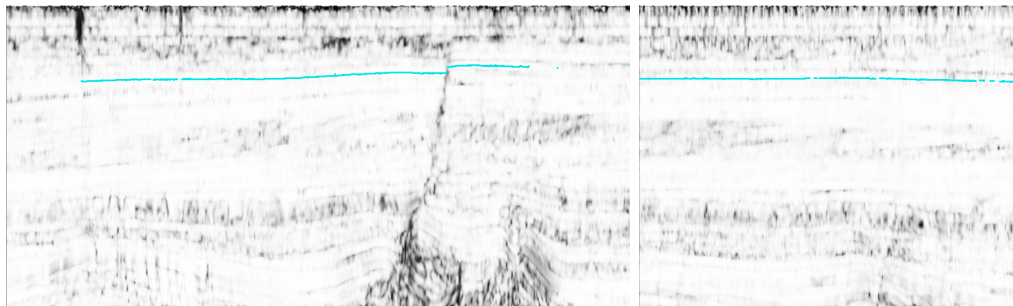
5.2

Dado Real

O conjunto de testes exibido nesta seção utilizou como gabarito um horizonte interpretado no volume *F3 Block* (22), fornecido pelo intérprete da Petrobras, Fernando Barbosa. A Figura 5.3 ilustra uma seção de amplitudes do volume em questão, assim como o atributo de falhas utilizado e o horizonte interpretado.



(a) Amplitudes na *inline* 240 e na *crossline* 700



(b) Atributos de falha na *inline* 240 e na *crossline* 700

Figura 5.3: Exemplos de entrada utilizada nos testes.

Para a realização dos testes, foram selecionadas dezenove seções na direção *inline* e vinte e duas seções na direção *crossline*. Para cada uma delas, calculamos sua interseção com o horizonte interpretado de forma a construir um conjunto de dados de teste. Em seguida, executamos variações do método de forma a analisar diferentes aspectos da abordagem proposta.

Para cada bateria de testes, rastreamos os horizontes em cada seção com os seguintes parâmetros: $L = 1$, $\alpha = 0.5$, meia janela variando no intervalo de $[1, 20]$, ponto inicial igual a amostra mais à esquerda do gabarito e ponto final igual a amostra mais à direita.

Em seguida, para cada horizonte rastreado, calculamos a distância vertical traço a traço entre a amostra rastreada e o gabarito. Destes valores, calculamos média, desvio padrão e taxa de acerto. A taxa de acerto é dada pela quantidade de amostras que estavam a uma tolerância de dois *pixels* do gabarito dividida pela quantidade total de amostras avaliadas.

Além disso, como base de comparação para o método baseado em menor caminho, utilizamos um algoritmo simples guloso, comumente encontrado. O algoritmo guloso inicia a busca pelo melhor caminho na amostra semente, consultando seus vizinhos para encontrar a amostra com maior coeficiente de correlação. Essa amostra é, então, considerada a nova semente e o processo se repete até alcançar o traço final. Esse algoritmo está disponibilizado no sistema v3o2, desenvolvido através de uma parceria entre o Instituto Tecgraf e a Petrobras.

No primeiro teste realizado, foram relaxadas duas características do método proposto. O grafo construído nessa versão de teste não utiliza o atributo de falha de forma a impor restrições à inserção de arestas entre traços. Dessa forma, ao distanciar-se da semente, mais amostras são candidatas a pertencer ao melhor caminho. Para facilitar a leitura deste documento, chamaremos este grafo de grafo cheio. A Figura 5.4 ilustra a diferença entre o grafo cheio e o grafo proposto no Capítulo 4.

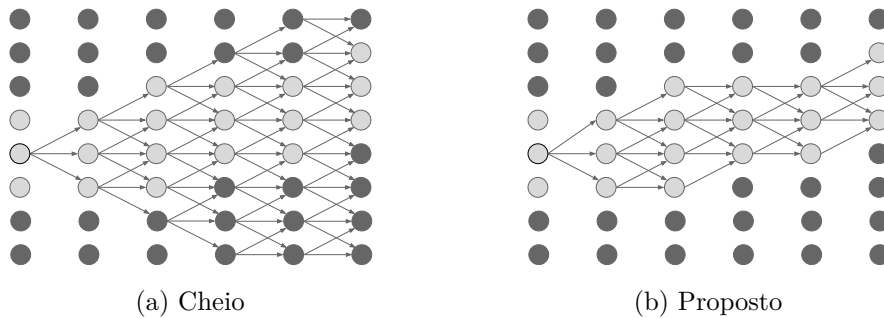


Figura 5.4: Diferença entre o grafo utilizado no primeiro caso de teste (a) e o grafo proposto (b).

Além disso, as janelas utilizadas pela medida de correlação foram extraídas a partir de amostras consecutivas no traço discreto, como ilustrado na Figura 4.5(a). Essa mesma modelagem é utilizada pelo método guloso, o que produz não só uma comparação sob condições mais similares, como também uma avaliação da abordagem puramente baseada no algoritmo de menor caminho. Para comparação com os demais métodos apresentados ao longo deste capítulo, chamaremos este método de Imediato, pois ele é o algoritmo de menor caminho imediato derivado do algoritmo guloso descrito.

Os resultados do teste estão exibidos na Figura 5.5, onde cada imagem representa o histograma dos resultados calculados em três cenários: (a) seções *inlines*, (b) *crosslines* e (c) nas duas direções. Essa separação ocorre pela grande diferença na largura das seções em cada direção. A dimensão de fatias *inline* é 950×460 , enquanto a de fatias *crossline* é 580×460 .

Nos histogramas apresentados, o eixo horizontal representa a distância de cada amostra rastreada à seu respectivo gabarito. É possível perceber que a abordagem baseada em menor caminho, ao contrário do que era esperado, não possui desempenho superior ao método guloso.

Também é interessante notar a característica multimodal dos histogramas, a qual é atribuída aos rastreamentos que se deslocaram para horizontes vizinhos. Essa característica é visível, principalmente, nas seções *inline*, onde, aparentemente, o rastreador global se concentrou em diferentes horizontes vizinhos.

Para compreender o motivo da qualidade inferior dos resultados, analisamos os coeficientes de correlação calculados onde o rastreamento falhou. Percebemos, então, que a medida de similaridade utilizada não é sempre mais alta ao longo do horizonte de interesse. Essa característica, aliada à liberdade do grafo cheio, dá ao rastreador de menor caminho a tendência de confundir horizontes vizinhos similares, como ilustrado na Figura 5.6.

Dessa forma, identificamos dois principais pontos de melhorias: a função objetivo e a construção do grafo. As seções 5.2.1 e 5.2.2 descrevem os testes realizados para identificar como cada uma das características do método proposto traz benefícios aos resultados.

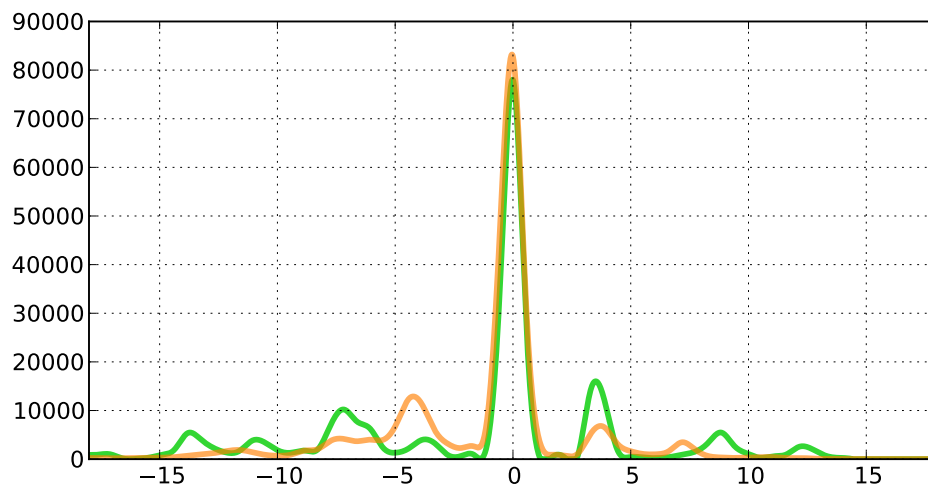
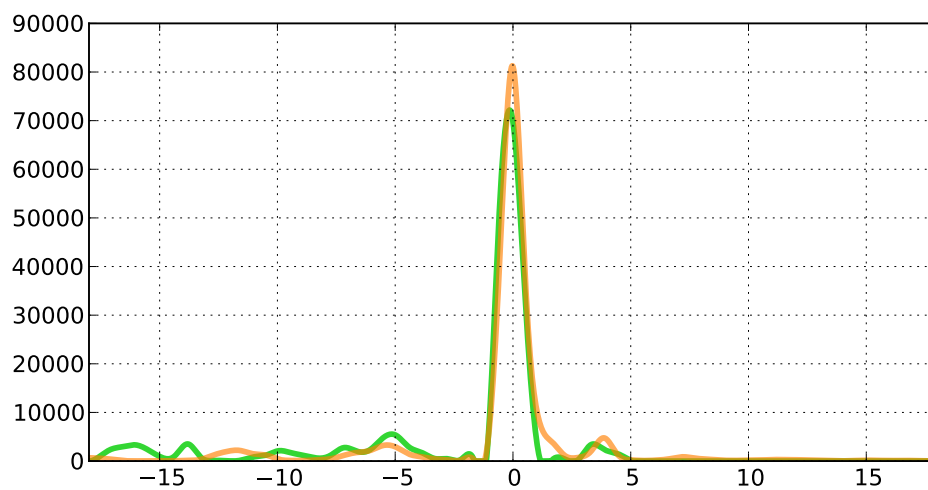
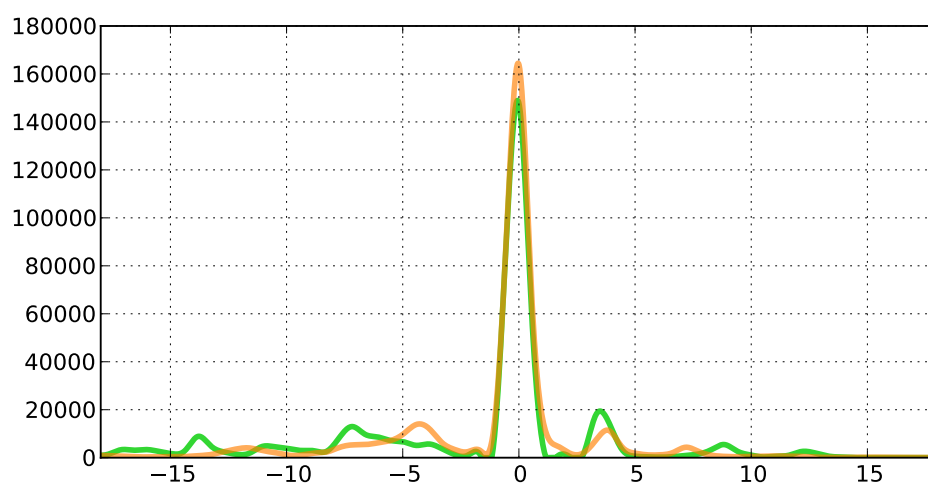
(a) *inline*(b) *crossline*(c) *inline + crossline*

Figura 5.5: Resultado da comparação entre as abordagens local (laranja) e baseada em menor caminho (verde).

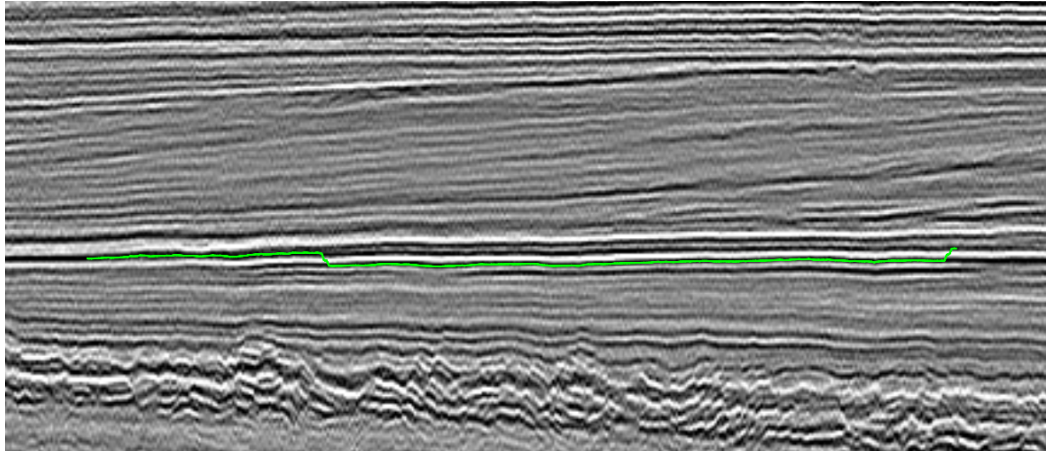


Figura 5.6: Problema encontrado na primeira versão de teste: a função objetivo não é ótima ao longo do horizonte e o rastreamento troca de horizonte no meio do caminho.

5.2.1

Avaliação da Função Objetivo

Estudamos dois aspectos da função objetivo. O primeiro diz respeito à entrada e avalia a diferença entre construir as janelas de correlação a partir amostras consecutivas no traço e a partir dos extremos do sinal. O segundo, refere-se à medida de similaridade utilizada para comparar duas amostras da seção sísmica.

Extração dos Picos do Sinal

A Figura 5.7 ilustra o resultado do mesmo conjunto de testes descrito anteriormente, porém, utilizando a extração de picos para a construção das janelas de correlação. Para comparação com os demais testes realizados, chamaremos este método de Picos.

É possível perceber que esta abordagem traz uma grande melhoria aos resultados. A quantidade de amostras próximas à semente é significativamente maior, o que indica a maior estabilidade desta abordagem.

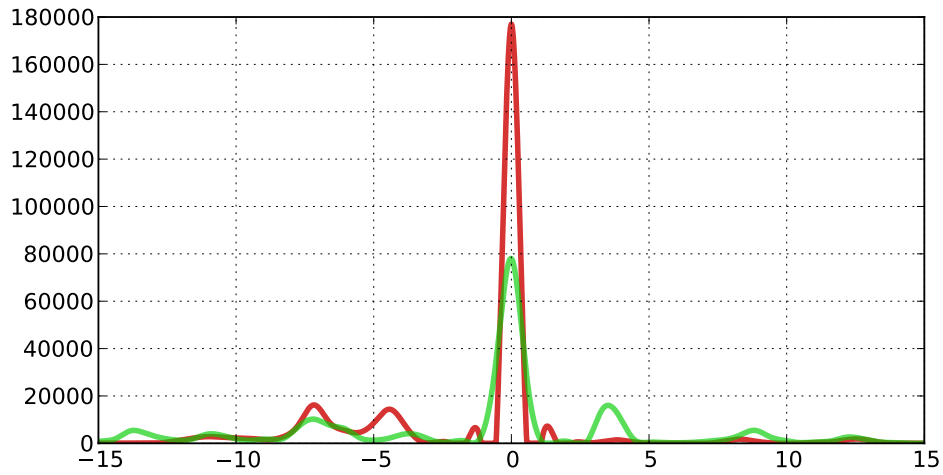
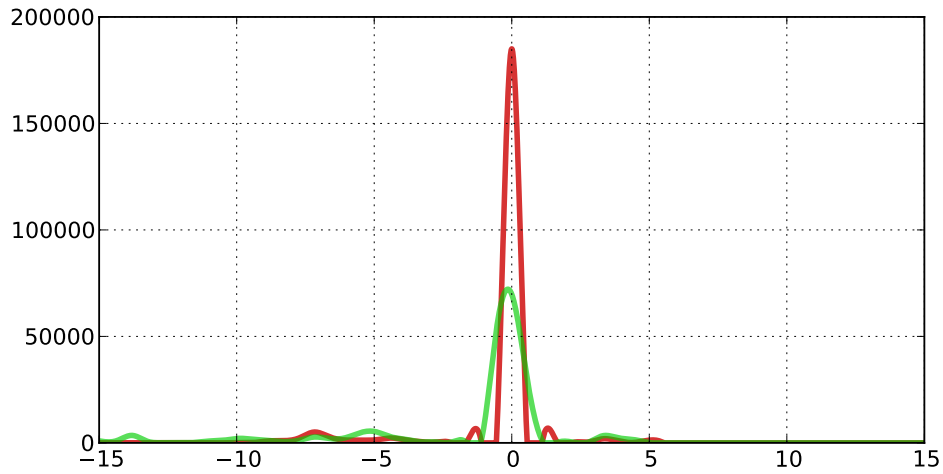
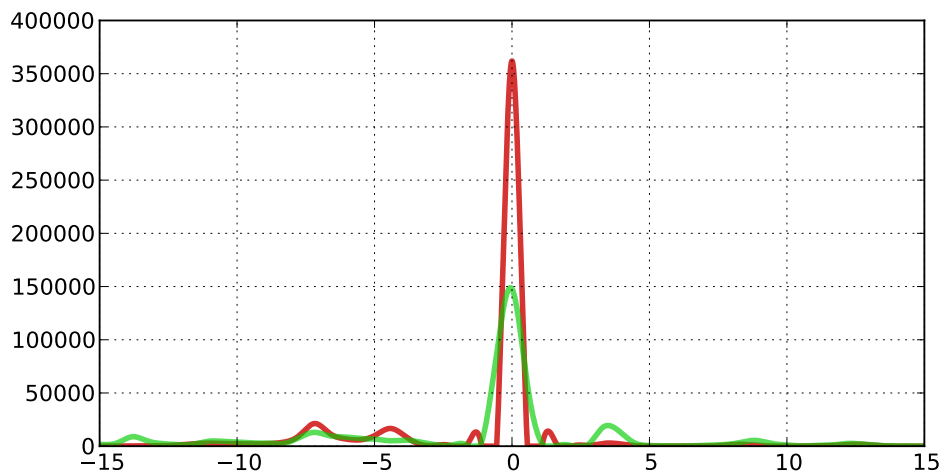
(a) *inline*(b) *crossline*(c) *inline + crossline*

Figura 5.7: Resultado da comparação entre a utilização de amostras consecutivas no traço discreto (verde) e utilização dos picos do sinal para construção do subtraço (vermelho).

Métricas de Correlação

Adicionalmente, estudamos diferentes métricas de correlação, com o objetivo de encontrar uma que se adaptasse às condições encontradas na sísmica.

Além da correlação cruzada (20), outras três medidas foram analisadas. A primeira, a correlação de cossenos c_θ (23), representa o ângulo entre os subtraços de entrada. A segunda, o coeficiente de correlação de Spearman ρ (24), pode ser definida como a correlação de Pearson aplicada aos postos dos dados de entrada. Já a terceira, κ , é uma medida baseada na ordenação relativa da intensidade das amostras dos conjuntos em comparação, proposta por Bhat *et. al.* (25). Em Aurnhammer *et al.* (26) os autores sugerem o uso desse último coeficiente de correlação e argumentam que o uso de postos para medir semelhança no tipo de dado proveniente do volume sísmico traz vantagens como robustez a *outliers* e variações de intensidade não lineares. No Apêndice A encontra-se a explicação detalhada de cada uma destas medidas.

Realizamos os mesmos testes descritos anteriormente variando a métrica de correlação utilizada pelo rastreador. A Tabela 5.1 exibe os resultados dos testes executados com o subtraço do sinal discreto. Neste cenário, as medidas parecem possuir comportamentos bastante diferentes. Há uma grande disparidade não só entre as médias de cada métrica, mas também entre seus desvios padrão, que variam de $\sigma = 6.278$ a $\sigma = 12.515$.

Métrica	Média	Desvio Padrão	Acerto (%)
Cossenos	-2.619	8.595	63.45
Pearson	-2.151	6.278	60.02
Spearman	-4.134	10.675	55.05
Bhat <i>et. al.</i>	-6.585	12.515	52.68

Tabela 5.1: Resultados da utilização de diferentes métricas de correlação aplicadas ao subtraço do sinal discreto.

Em seguida, testamos cada uma das métricas utilizando o subtraço construído a partir dos extremos do sinal sísmico. A Tabela 5.2 mostra os novos resultados obtidos. Nesse caso, o desempenho de todas as métricas foi bastante similar, exceto pelo coeficiente κ , cujos resultados, apesar da melhora, são inferiores às outras medidas. Este resultado indica que, diante da estabilidade proporcionada pela extração dos picos, a utilização das diferentes métricas estudadas não traz benefício significativo ao método.

Métrica	Média	Desvio Padrão	Acerto (%)
Cossenos	-1.634	3.624	74.21
Pearson	-1.448	3.752	73.37
Spearman	-1.623	3.531	70.19
Bhat <i>et. al.</i>	-3.113	6.847	72.62

Tabela 5.2: Resultados da utilização de diferentes métricas de correlação aplicadas ao subtraço construído a partir dos picos do sinal sísmico.

5.2.2

Avaliação da Construção do Grafo

Os resultados exibidos nesta seção ilustram o desempenho do método proposto no Capítulo 4. Como visto, além da extração de picos para a medida de correlação, o crescimento do grafo é limitado pelo atributo de falha. Para quantificar as possíveis melhorias obtidas com a limitação do grafo, o método proposto é comparado com a versão apresentada na seção 5.2.1, onde não há restrição durante a construção do grafo e também é utilizado o subtraço formado pelos extremos do sinal.

Os resultados, exibidos na Figura 5.8, indicam que a abordagem proposta traz benefícios, principalmente, na direção *inline*. Isso acontece porque o tamanho do caminho nesta direção é maior, o que dá margem para o rastreador pular de horizonte no grafo cheio.

As Tabelas 5.3, 5.4 e 5.5 exibem as estatísticas de cada um dos testes apresentados. Nos três cenários, o método proposto obteve melhores resultados que a abordagem local avaliada. A distância média entre pontos rastreados e seus respectivos gabaritos se manteve menor do que 1; o maior desvio padrão encontrado é, aproximadamente, $\sigma = 3 \text{ pixels}$; e a taxa de acerto calculada com uma tolerância de dois *pixels* está em torno de 90%.

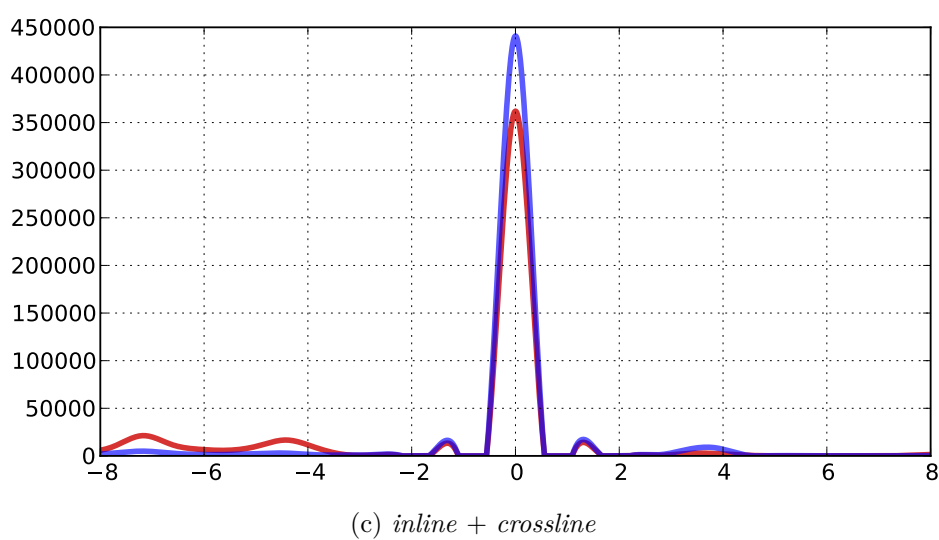
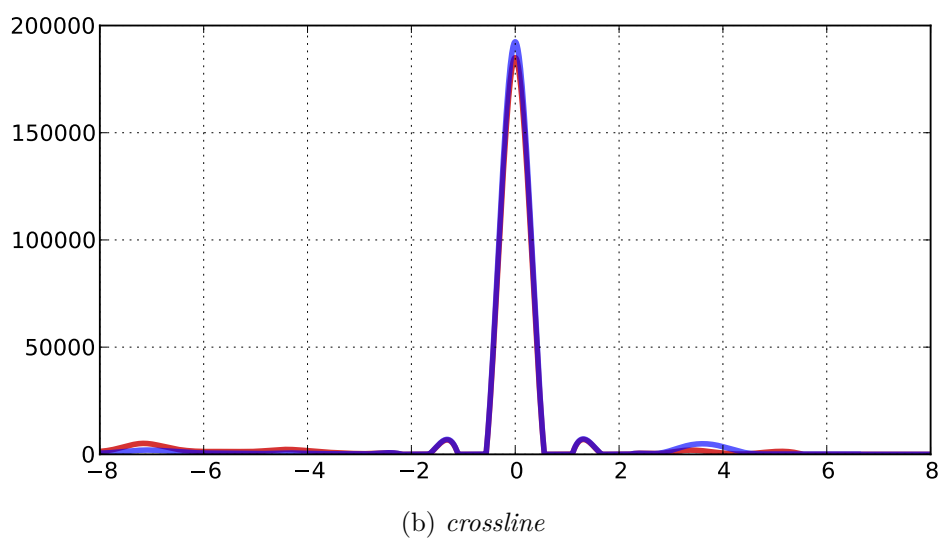
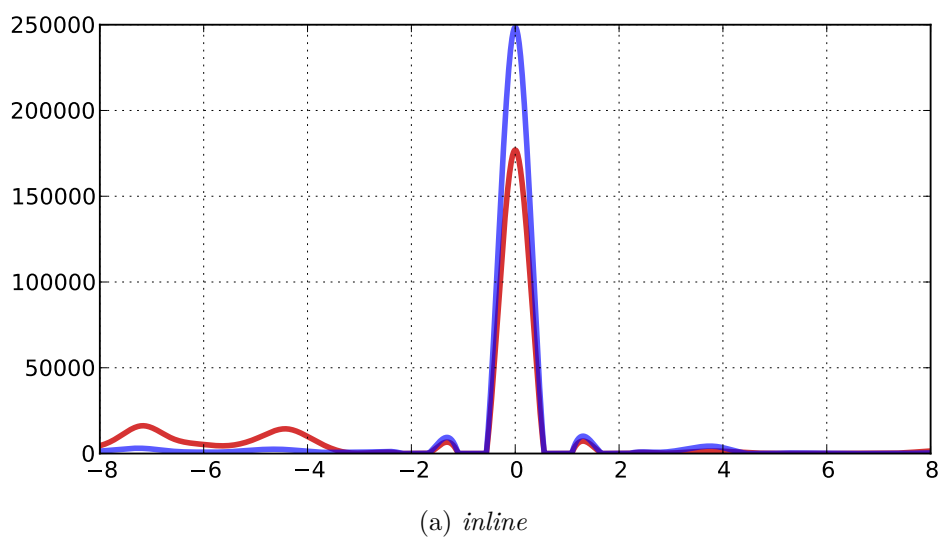


Figura 5.8: Resultado da comparação entre construção do grafo sem utilização do atributo de falha (vermelho) e do método proposto (azul).

Método	Média	Desvio Padrão	Acerto (%)
Imediato	-1.968	6.964	50.71
Picos	-1.989	4.374	62.62
Proposto	0.172	2.997	87.97
Guloso	-1.295	4.406	61.64

Tabela 5.3: Resumos dos resultados nas seções *inline*.

Método	Média	Desvio Padrão	Acerto (%)
Imediato	-2.396	5.212	72.46
Picos	-0.726	2.564	87.73
Proposto	-0.094	1.566	91.39
Guloso	-0.684	4.207	80.10

Tabela 5.4: Resumos dos resultados nas seções *crossline*.

Método	Média	Desvio Padrão	Acerto (%)
Imediato	-2.151	6.278	60.02
Picos	-1.448	3.752	73.37
Proposto	0.058	2.490	89.43
Guloso	-1.033	4.332	69.54

Tabela 5.5: Resumos dos resultados nas seções *inline* + *crossline*.

5.2.3

Tempo de Processamento

Para avaliar a eficiência do método proposto, medimos o tempo de processamento para cálculo do horizonte rastreado em cada uma das seções disponíveis para teste. A Tabela 5.6 exibe o tempo médio em segundos calculado nas direções *inline* e *crossline* para cada valor de meia janela utilizado nos testes.

Os resultados indicam que a variação no tamanho da meia janela de correlação não produz impactos significativos na eficiência do método. No entanto, a dimensão da seção utilizada possui impacto direto no tempo de processamento. Como esperado, o tempo de processamento na direção *inline* é aproximadamente o dobro do tempo utilizado nas seções *crossline*, assim como a relação entre as dimensões das imagens nas duas direções. Ainda assim, o método se mostrou eficiente o bastante para sua utilização de maneira interativa.

Meia Janela	Tempo Médio (s)	
	Inline	Crossline
1	0.0384	0.0225
2	0.0385	0.0219
3	0.0383	0.0220
4	0.0384	0.0223
5	0.0384	0.0222
6	0.0379	0.0220
7	0.0380	0.0219
8	0.0381	0.0225
9	0.0376	0.0223
10	0.0378	0.0223
11	0.0390	0.0221
12	0.0385	0.0222
13	0.0394	0.0222
14	0.0391	0.0226
15	0.0387	0.0219
16	0.0395	0.0222
17	0.0382	0.0221
18	0.0383	0.0231
19	0.0385	0.0222
20	0.0378	0.0218

Tabela 5.6: Tempo de processamento médio em uma máquina Core i7 1.9Ghz, 8GB RAM.

5.3

Avaliação Visual

O objetivo desta seção é apresentar o comportamento do algoritmo proposto em diferentes regiões do dado sísmico. Os horizontes exibidos foram rastreados no mesmo volume utilizado anteriormente, *F3 Block* (22), porém o dado foi pré-condicionado com uma filtragem orientada (27).

As Figuras 5.9, 5.10, 5.11 e 5.12 mostram diferentes conjuntos de horizontes rastreados nas direções *inline* e *crossline*. Percebe-se que o método proposto se manteve robusto frente a diferentes feições sísmicas.

Além disso, na Figura 5.13, encontram-se exemplos de situações onde a abordagem proposta possui melhor desempenho que a local. Nessas regiões, tipicamente, existe alguma descontinuidade no horizonte que provoca propagação de erro no método guloso.

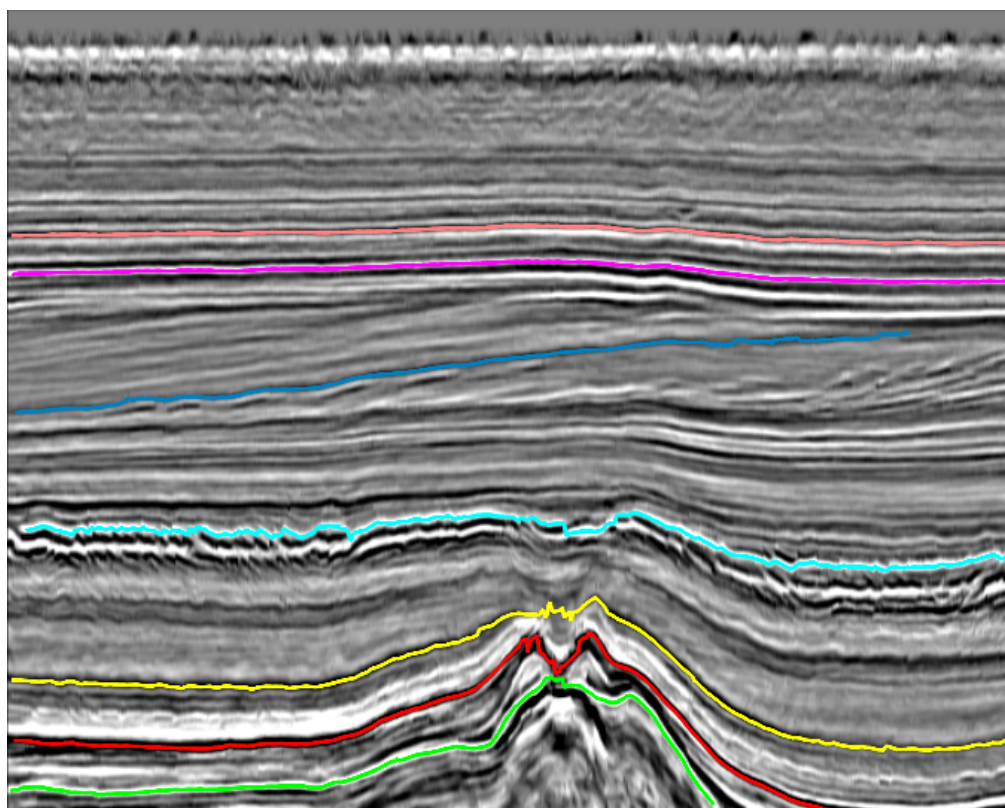


Figura 5.9: Conjunto de horizontes mapeados com o método proposto através de um domo de sal na *crossline* 780.

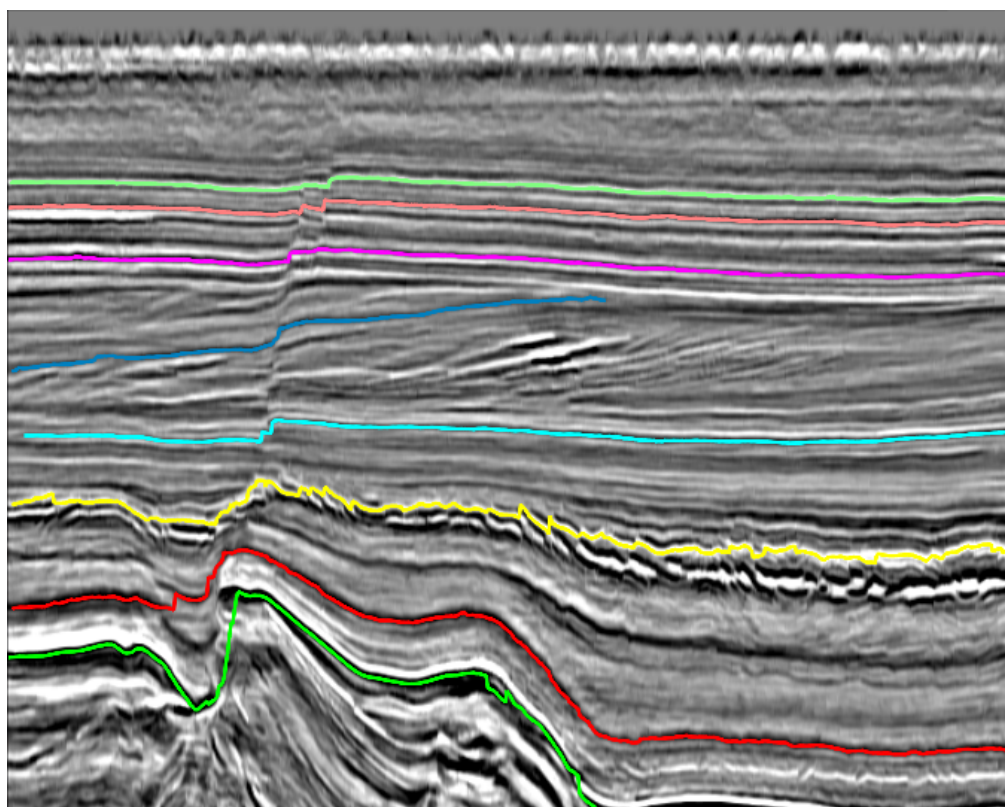


Figura 5.10: Conjunto de horizontes mapeados na *crossline* 940. A abordagem proposta se mostrou robusta à presença de falhas.

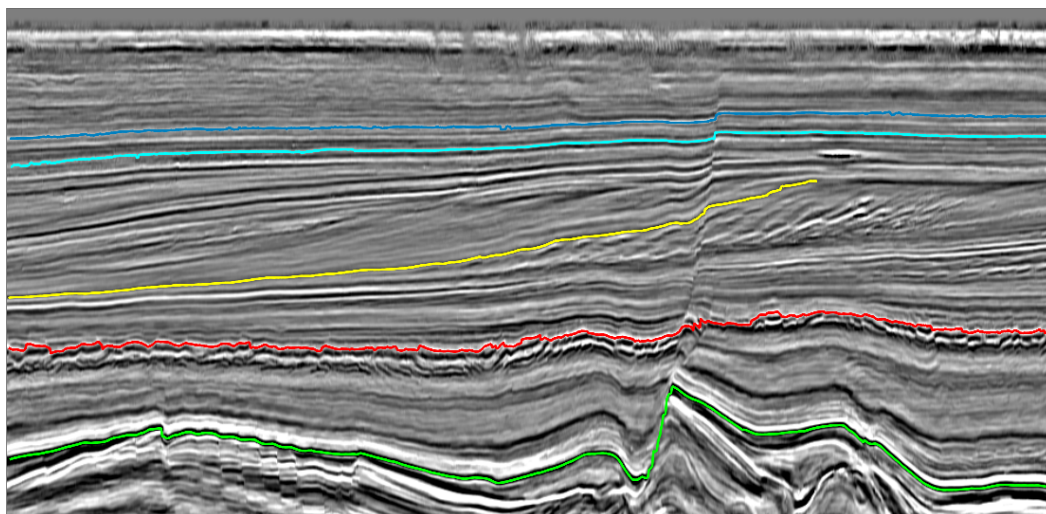


Figura 5.11: Conjunto de horizontes mapeados na *inline* 280. O método proposto se mostrou capaz de atravessar falhas com grandes rejeitos.

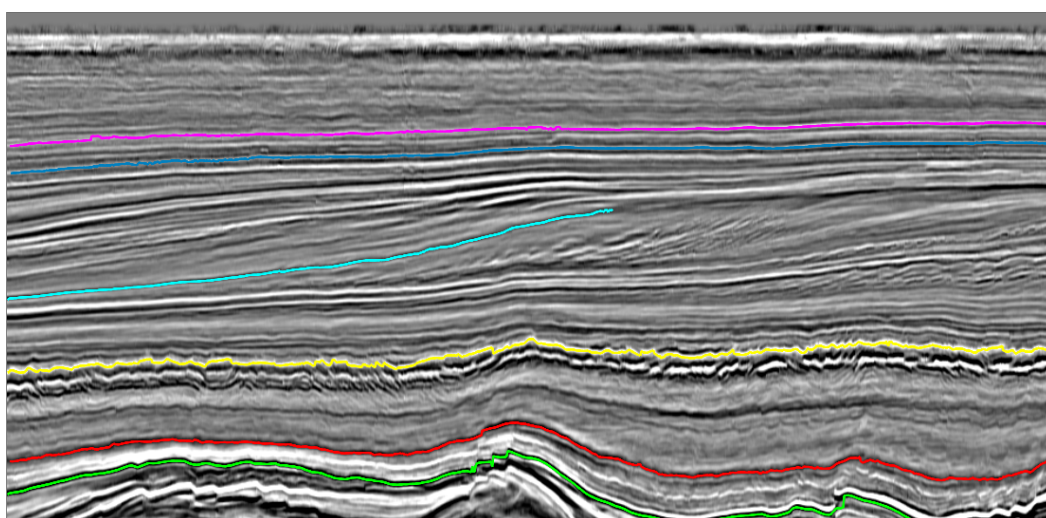
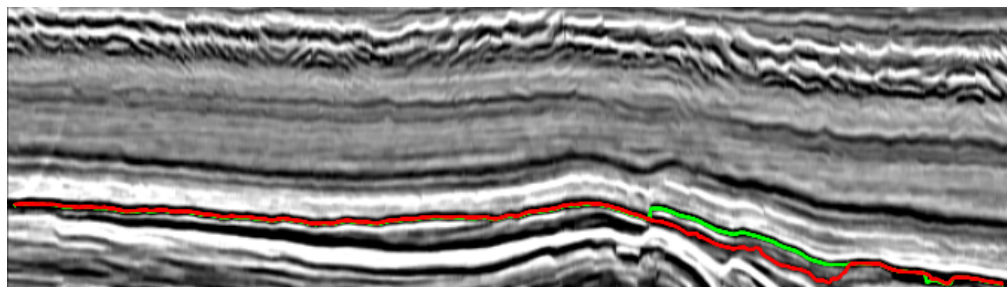
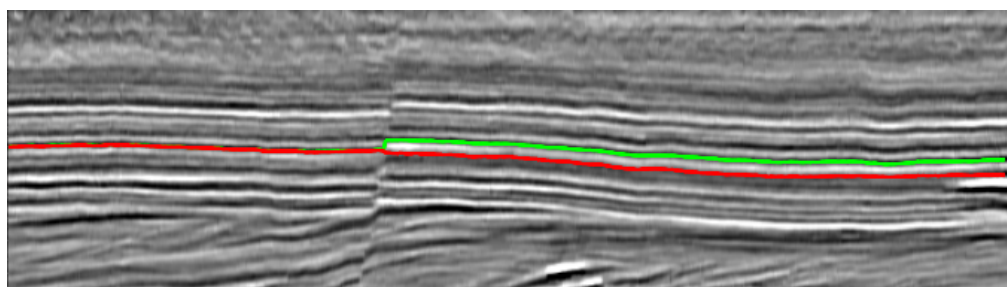


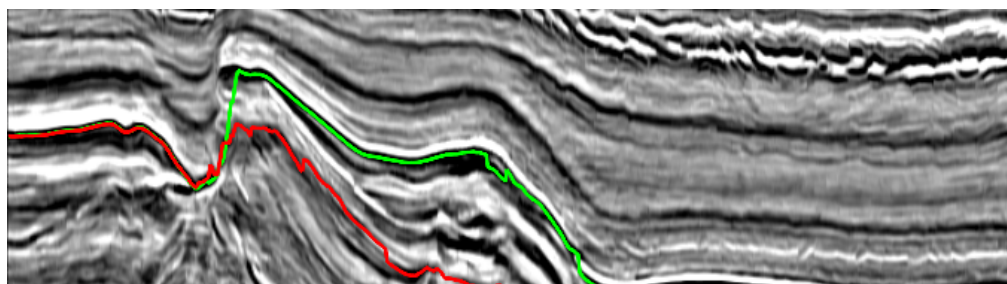
Figura 5.12: Conjunto de horizontes rastreados com o método proposto na *inline* 385.



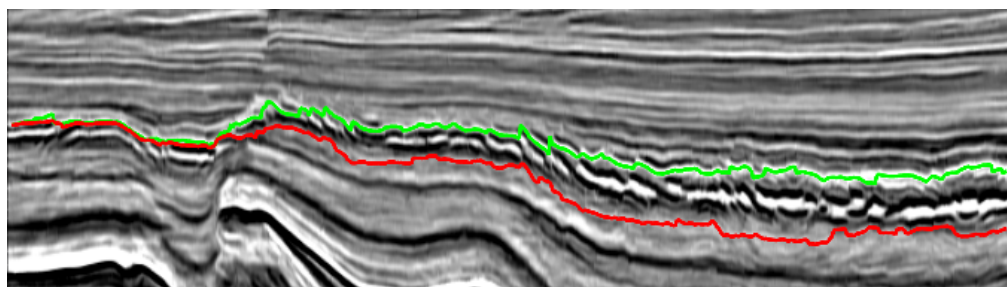
(a)



(b)



(c)



(d)

Figura 5.13: Situações onde a abordagem proposta (verde) possui melhor desempenho que a local (vermelho).

6

Conclusões

Neste trabalho, apresentamos uma proposta de rastreamento 2D de horizontes sísmicos baseada no algoritmo de menor caminho em grafos de *Bellman-Ford* (16). Para isso, modelamos o horizonte discreto como o caminho de melhor custo em um grafo cujo peso de cada aresta representa a similaridade entre o nó destino e a amostra semente.

Além disso, apresentamos uma avaliação da utilização de uma abordagem baseada em algoritmos de menor caminho no processo de rastreamento. Para isso, comparamos o impacto de diferentes características do método proposto, de forma a apontar as vantagens e desvantagens do processo global.

A partir dos resultados exibidos no Capítulo 5, podemos concluir que a função objetivo proposta não é sempre máxima ao longo do horizonte. Por essa razão, o rastreamento no grafo cheio mostrou uma fraqueza da abordagem global. Em regiões coerentes, o melhor caminho atravessa horizontes vizinhos similares ao horizonte em questão. Nestas regiões, a abordagem local se mostrou mais estável. Isso acontece porque, além de limitar o número de opções por onde o caminho pode percorrer, ela atualiza a informação da semente nestas regiões seguras. Por outro lado, a abordagem global tem sua maior vantagem nas regiões de falha, uma vez que ela é capaz de incorporar uma visão abrangente da seção de forma a permitir que a descontinuidade seja atravessada sem que o caminho seja comprometido.

Observamos, também, que os efeitos da discretização do sinal sísmico possuem grande impacto sobre a medida de similaridade utilizada comumente, a correlação cruzada. Nos testes onde a correlação foi calculada a partir de um subtraço construído a partir de amostras consecutivas no traço discreto, os resultados apresentaram altos valores de desvio padrão, inclusive na abordagem gulosa. A utilização dos picos do sinal para construção do subtraço trouxe melhorias, principalmente em termos de estabilidade, traduzidas a partir dos menores desvios padrão apresentados.

Ainda assim, a medida de correlação utilizada não se mostrou sempre máxima ao longo do horizonte sendo rastreado. Dessa forma, para utilização de um método puramente global, ainda há necessidade de um estudo aprofundado,

de forma a encontrar uma função objetivo bem adequada ao rastreamento de horizontes.

Adicionalmente, foram estudadas diferentes métricas de similaridade encontradas na literatura. Apesar de cada métrica ter apresentado comportamento diferente com a utilização de um subtraço com amostras consecutivas do sinal discreto, com a utilização dos picos suas estatísticas se mostraram bastante similares. Isso indica que, dentre as métricas avaliadas, nenhuma traz benefícios quando comparada à correlação cruzada.

Observamos que uma construção mais inteligente do grafo faz com que o processo baseado em menor caminho possua melhores resultados. A limitação de seu crescimento a partir do atributo de falha foi capaz de eliminar cenários onde o rastreador trocava de horizonte, pois esta construção permite que saltos no caminho ocorram apenas em regiões de falha.

Essa limitação do crescimento do grafo cria uma dependência entre a qualidade do método e do atributo identificador de falhas, porém a dependência não é grave, pois para o método, é importante apenas a localização aproximada de pontos onde não existem falhas. Dessa forma, não é de vital importância a utilização de um atributo extremamente preciso.

De forma geral, concluímos que, idealmente, têm-se um método híbrido das duas abordagens. A versão proposta tenta mesclar os benefícios de cada uma, utilizando o atributo de falha para simular uma abordagem local nas regiões coerentes do dado, através da limitação do tamanho do grafo. Ao atingir uma região de falha, no entanto, o método proposto busca as vantagens do método global, permitindo que se busque o melhor caminho num grafo cheio.

6.1

Trabalho Futuros

A partir dos resultados e conclusões observados ao longo deste trabalho, destacamos os seguintes trabalhos futuros.

6.1.1

Estudo da Função Objetivo

O maior problema encontrado para a utilização do algoritmo de menor caminho foi a função objetivo proposta, que não se mostrou eficaz em caracterizar um horizonte sísmico. Dessa forma, o primeiro trabalho futuro sugerido é o estudo aprofundado das características do sinal sísmico, para que se possa definir uma função de custo bem adequada ao problema em questão.

Além disso, os altos valores de desvios padrão apresentados pela abordagem gulosa podem indicar que a discretização do sinal sísmico possui

outros impactos na medida de correlação, além da dependência entre tamanho da meia janela e largura do pulso. Dessa forma, é interessante, também, avaliar outros possíveis problemas decorrentes da discretização do sinal.

6.1.2

Definição de múltiplas sementes

Outra possível melhoria no método proposto consiste em incorporar a definição de múltiplas sementes. Uma fraqueza da função de custo proposta é a medida de similaridade calculada sempre em relação ao ponto inicial do horizonte. Isso tende a funcionar bem no entorno da semente, porém, ao distanciar-se da mesma, não é possível garantir que as condições da sísmica se mantêm. Dessa forma, a utilização de múltiplas sementes pode ajudar a caracterizar o horizonte em diferentes regiões.

Implementamos uma versão simples de teste que utiliza a informação de diversas sementes para realizar pequenos rastreamentos. Essa implementação mostrou bons resultados, corrigindo de maneira eficiente o rastreamento em pontos de erro, porém essa informação pode ser utilizada de maneira mais inteligente na função de custo. Para isso, deve ser definida uma forma de interação entre as diferentes sementes, de modo que o algoritmo não perca o caráter global.

6.1.3

Identificação de Interrupções no Horizonte

O algoritmo atual espera a existência de uma continuidade lateral entre as amostras do horizonte. Apesar do algoritmo permitir a busca por vizinhos onde o traço adjacente não possui amostras válidas, essa situação é bastante rara e este tratamento não leva em consideração as características da sísmica na região. Dessa forma, o rastreador tende a adicionar, desde a semente até o ponto final, uma amostra por traço.

Essa modelagem pode não ser sempre verdade. Processos de erosão, por exemplo, podem fazer com que a camada seja interrompida em alguns traços e depois continue. Seria desejável, então, incorporar ao algoritmo a capacidade de decidir se o horizonte sendo mapeado está presente ou não em cada traço. Um tratamento simples para esse problema consiste na definição de um valor mínimo para o coeficiente de correlação, de modo que amostras com coeficiente abaixo desse limiar não seriam inseridas ao resultado final.

6.1.4

Extensão para 3D

Finalmente, o último trabalho sugerido é a extensão do algoritmo para 3D. Apesar das possíveis melhorias descritas acima, a metodologia proposta se mostrou precisa e estável. Dessa forma, seria interessante avaliar a utilização da mesma abordagem no ambiente 3D. Para isso, seria necessária a busca por uma superfície de melhor custo.

Referências Bibliográficas

- [1] GERHARDT, S. **Aspectos da visualização volumétrica de dados sísmicos**. Rio de Janeiro, Brasil, 1998. Dissertação de Mestrado - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.
- [2] SILVA, P. **Visualização Volumétrica de Horizontes em Dados Sísmicos 3D**. Rio de Janeiro, Brasil, julho 2004. Tese de Doutorado - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.
- [3] ROBINSON, E. A.; TREITEL, S. **Geophysical signal analysis**, volume 31. Prentice-Hall New Jersey, 1980.
- [4] PAMPANELLI, P. C. P.; SILVA, P. M. ; GATTASS, M. **A new volumetric fault attribute based on first order directional derivatives**. In: Thirteenth International Congress of The Brazilian Geophysical Society. SBGF, 2013.
- [5] DORN, G. **Modern 3-d seismic interpretation**. The Leading Edge, v.17, n.9, p. 1262–1262, 1998.
- [6] SHERIFF, R. E.; GELDART, L. P. **Exploration seismology**, volume 1995. Cambridge university press Cambridge, 1982.
- [7] MACHADO, M. **Segmentação de dados sísmicos via hyperstack para visualização**. Rio de Janeiro, Brasil, 2000. Dissertação de Mestrado - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.
- [8] YILMAZ, Ö.; DOHERTY, S. M. **Seismic data processing**, volume 2. Society of Exploration Geophysicists Tulsa, 1987.
- [9] SHERIFF, R. E. **Encyclopaedic dictionary of exploration geophysics**. Society of Exploration Geophysicists, Tulsa, Oklahoma, 1991.
- [10] LORENSEN, W. E.; CLINE, H. E. **Marching cubes: A high resolution 3d surface construction algorithm**. In: Proceedings of the

- 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '87, p. 163–169, New York, NY, USA, 1987. ACM.
- [11] BROWN, M. P.; MORTON, S. A.; WHITTLE, G. ; OTHERS. **Seismic event tracking by global path optimization**. 76th Mtg. Soc. of Expl. Geophys, p. 1063–1067, 2006.
- [12] BORN, M.; WOLF, E. **Principles of Optics, student expanded edition**, p. 117–120. Cambridge University Press, 7th. ed., 1999.
- [13] ZINCK, G.; DONIAS, M.; GUILLON, S. ; LAVIALLE, O. **Discontinuous seismic horizon tracking based on a poisson equation with incremental dirichlet boundary conditions**. In: Image Processing (ICIP), 2011 18th IEEE International Conference on, p. 3385–3388, 2011.
- [14] LI, L.; MA, G. ; DU, X. **New method of horizon recognition in seismic data**. Geoscience and Remote Sensing Letters, IEEE, v.9, n.6, p. 1066–1068, 2012.
- [15] HUANG, K.-Y.; CHANG, C.-H.; HSIEH, W.-S.; HSIEH, S.-C.; WANG, L. K. ; TSAI, F.-J. **Cellular neural network for seismic horizon picking**. In: Cellular Neural Networks and Their Applications, 2005 9th International Workshop on, p. 219–222. IEEE, 2005.
- [16] CORMEN, T. H.; STEIN, C.; RIVEST, R. L. ; LEISERSON, C. E. **Introduction to Algorithms**, p. 585–594. McGraw-Hill Higher Education, 2nd. ed., 2001.
- [17] SILVA, P. M.; MARTINS, L. ; GATTASS, M. **Horizon indicator attributes and applications**. In: SEG Technical Program Expanded Abstracts 2012, p. 1–6, 2012.
- [18] BARNES, A. E. **Theory of 2-d complex seismic trace analysis**. Geophysics, v.61, n.1, p. 264–272, 1996.
- [19] BRACEWELL, R. N. **Fourier transform and its applications**, p. 359–367. McGraw-Hill Education, 3rd. ed., 1980.
- [20] HOGG, R. V.; MCKEAN, J. ; CRAIG, A. T. **Introduction to mathematical statistics**, p. 73–79. Pearson Education, 4th. ed., 2005.
- [21] CORMEN, T. H.; STEIN, C.; RIVEST, R. L. ; LEISERSON, C. E. **Introduction to Algorithms**, p. 549–550. McGraw-Hill Higher Education, 2nd. ed., 2001.

- [22] **Opendtect website.** <http://opendtect.org>.
- [23] SALTON, G.; MCGILL, M. **Introduction to modern information retrieval.** McGraw-Hill computer science series. McGraw-Hill, 1983.
- [24] KENDALL, M.; GIBBONS, J. **Rank correlation methods.** A Charles Griffin Book. E. Arnold, 1990.
- [25] BHAT, D. N.; NAYAR, S. K. **Ordinal measures for image correspondence.** Pattern Analysis and Machine Intelligence, IEEE Transactions on, v.20, n.4, p. 415–423, 1998.
- [26] AURNHAMMER, M.; MAYORAL, R. **Improving seismic horizon matching by ordinal measures.** In: Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, volume 3, p. 642–645. IEEE, 2004.
- [27] DABER, R.; AQRAWI, A. ; LIMITED, S. **Petrel 2010: Interpreter's Guide to Seismic Attributes.** Schlumberger, 2010.

A

Medidas de Correlação

O objetivo deste apêndice é detalhar cada uma das métricas de correlação avaliadas neste trabalho. Além de suas definições, fazemos uma comparação entre seus comportamentos quando calculadas com conjuntos correlacionados de diferentes maneiras.

A.1

Correlação de Cossenos

A correlação de cossenos (23) é a medida do cosseno do ângulo formado pelos vetores de entrada X_a e X_b . Por essa razão, ela é uma medida de orientação, e não magnitude, o que é adequado para caracterizar as variações de amplitude ao longo de um horizonte. Essa métrica é uma medida de correlação linear entre as variáveis de entrada e assume valor igual a 1 quando os conjuntos possuem correlação linear positiva perfeita e valor igual a -1 quando sua correlação é negativa e perfeita. Porém, a correlação de cossenos não é invariante a deslocamentos. Em um par de conjuntos relacionados pela equação $X_b = X_a$, se X_b for deslocado para $X_b = X_a + 1$, por exemplo, o coeficiente de similaridade de cossenos será alterado.

O cálculo da correlação de cossenos está descrito na Equação A-1:

$$c_\theta = \cos(\theta) = \frac{X_a \cdot X_b}{\|X_a\| \|X_b\|} = \frac{\sum_{i=1}^n X_a^i X_b^i}{\sqrt{\sum_{i=1}^n (X_a^i)^2} \sqrt{\sum_{i=1}^n (X_b^i)^2}} \quad (\text{A-1})$$

A Figura A.1 ilustra o cálculo do coeficiente de similaridade de cossenos para dois subtraços extraídos de uma seção de amplitude real.

A.2

Correlação de Pearson

O coeficiente de correlação de Pearson (20), r , foi a métrica utilizada no método proposto e está descrita no Capítulo 4.

A Figura A.2 ilustra o cálculo do coeficiente de correlação de Pearson nos mesmos subtraços da Figura A.1. A correlação de Pearson é equivalente à

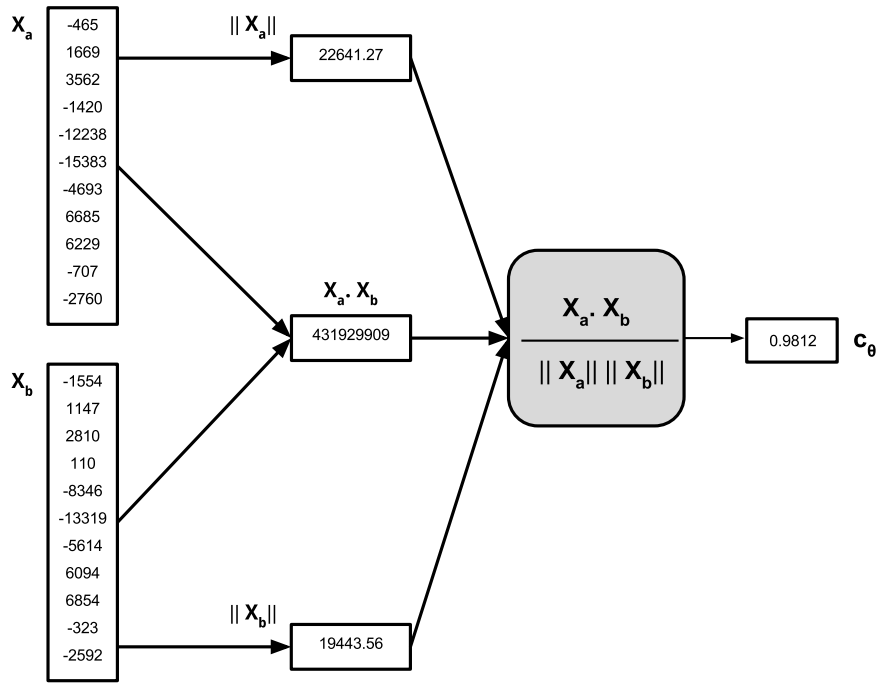


Figura A.1: Exemplo ilustrando o cálculo do coeficiente c_θ .

correlação de cossenos quando calculada sobre dados centralizados (quando a média do conjunto é subtraída de cada amostra).

A.3 Correlação de Spearman

O coeficiente de correlação de postos de Spearman (24), ρ , é definido como a correlação de Pearson aplicada sobre os postos dos dados de entrada. Ao contrário da correlação cruzada, essa medida não assume dependência linear entre os conjuntos e tenta descrever a relação entre as variáveis através de uma função monótona arbitrária.

A Equação A-2 define o cálculo do coeficiente ρ , onde \bar{x}_a e \bar{x}_b são as médias aritméticas dos conjuntos de postos x_a e x_b , x_a^i é o posto correspondente à amostra X_a^i dado pela posição de X_a^i no vetor X_a ordenado e x_b^i é o posto correspondente à amostra X_b^i dado pela posição de X_b^i no vetor X_b ordenado.

$$\rho = \frac{\sum_{i=1}^n (x_a^i - \bar{x}_a)(x_b^i - \bar{x}_b)}{\sqrt{\sum_{i=1}^n (x_a^i - \bar{x}_a)^2} \sqrt{\sum_{i=1}^n (x_b^i - \bar{x}_b)^2}} \quad (\text{A-2})$$

A Figura A.3 ilustra o cálculo do coeficiente ρ . Apesar de parecido, o coeficiente de correlação de Spearman traz vantagens em relação ao coeficiente de Pearson. Por exemplo, alterações nos dados de entrada que não modifiquem os valores dos postos não alteram o coeficiente de Spearman, mas alteram o

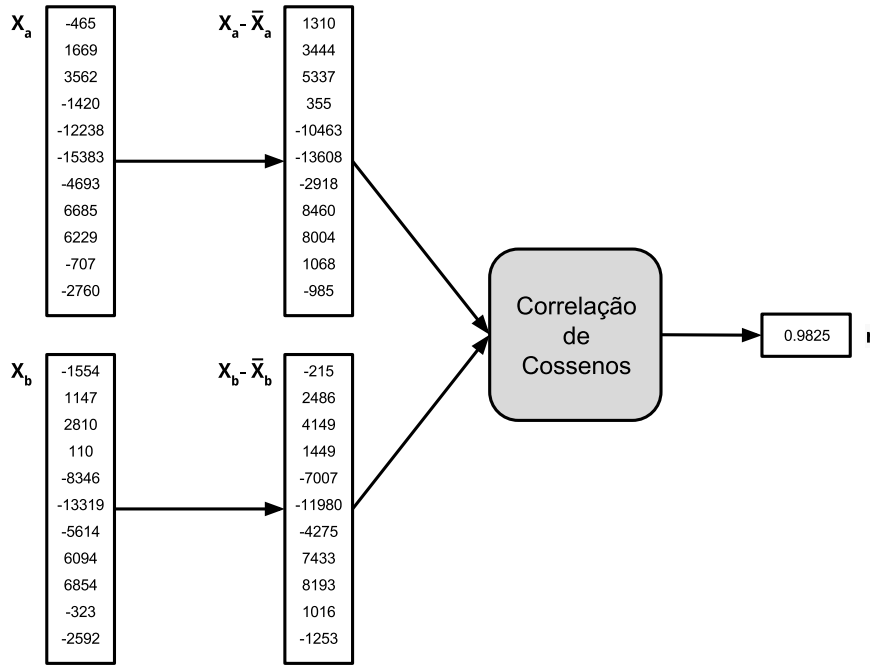


Figura A.2: Exemplo ilustrando o cálculo do coeficiente r de Pearson.

coeficiente de Pearson. Essa característica é vantajosa no âmbito da sismica porque permite captar variações de amplitude não lineares ao longo de um horizonte. Além disso, por utilizar postos, ele é menos sensível a *outliers*, causados pelo ruído inerente ao dado sísmico.

A.4

Correlação de Bhat *et. al.*

Essa medida de correlação foi originalmente proposta por Bhat *et. al.* (25) para algoritmos de correspondência de imagens. Ela é baseada na ordenação relativa da intensidade das amostras dos conjuntos em comparação.

Em Aurnhammer *et. al.* (26) os autores sugerem o uso do coeficiente de correlação, κ , proposto por Bhat *et. al.* e argumentam que o uso de postos para medir correlação para o tipo de dado proveniente do volume sísmico traz vantagens como robustez a *outliers* e alterações de intensidade não lineares.

O primeiro passo para computar o coeficiente κ é calcular a permutação composta, definida em A-3, onde x_a^{-1} representa a permutação inversa de x_a , definida por: se $x_a^i = j$ então $(x_a^{-1})^j = i$.

$$s^i = x_b^k, k = (x_a^{-1})^i \quad (\text{A-3})$$

“O valor de s representa o *ranking* de X_b em relação ao de X_a . Se os conjuntos possuem correlação positiva perfeita, s deve ser igual à *permutação*

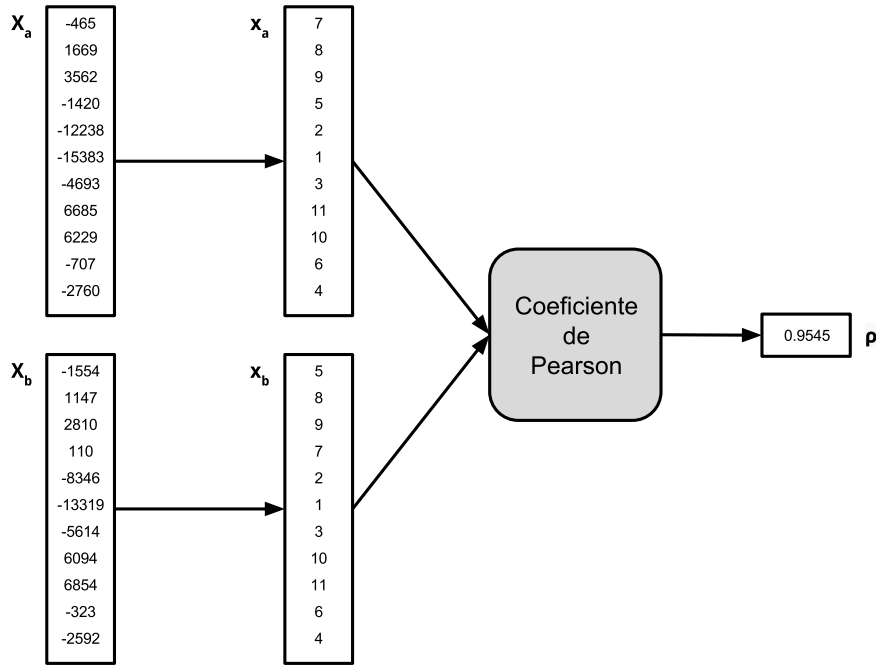


Figura A.3: Exemplo ilustrando o cálculo do coeficiente ρ de Spearman.

identidade, dada por $u = (1, 2, \dots, n)$ (25). Em seguida, é calculada uma medida de distância entre s e u , definida pela Equação A-4, onde $J(B) = 1$ se B é verdadeiro e $J(B) = 0$ caso contrário.

$$d_m^i = i - \sum_{j=1}^i J(s^j \leq i) \quad (\text{A-4})$$

O valor de cada componente do vetor d_m estima quantos elementos precedentes em s estão fora de posição. Finalmente, o coeficiente de correlação κ é dado pela Equação A-5.

$$\kappa = 1 - \frac{2 \max_{i=1}^n d_m^i}{\lfloor \frac{n}{2} \rfloor} \quad (\text{A-5})$$

A Figura A.4 ilustra o cálculo do coeficiente κ . Bhat *et. al.* destacam que essa medida de correlação foi projetada para possuir baixa sensibilidade à ruídos aleatórios e distorções na classificação das amostras, o que Aurnhammer *et. al.* destacam como uma vantagem em relação ao coeficiente de Spearman. Porém, os autores também observam que essa invariância pode se tornar uma deficiência ao comparar conjuntos não correspondentes, pois ela reduz o poder discriminatório de κ .

As Figuras A.5, A.6 e A.7 ilustram como cada uma das métricas descritas acima se comportam em conjuntos correlacionados de diferentes maneiras.

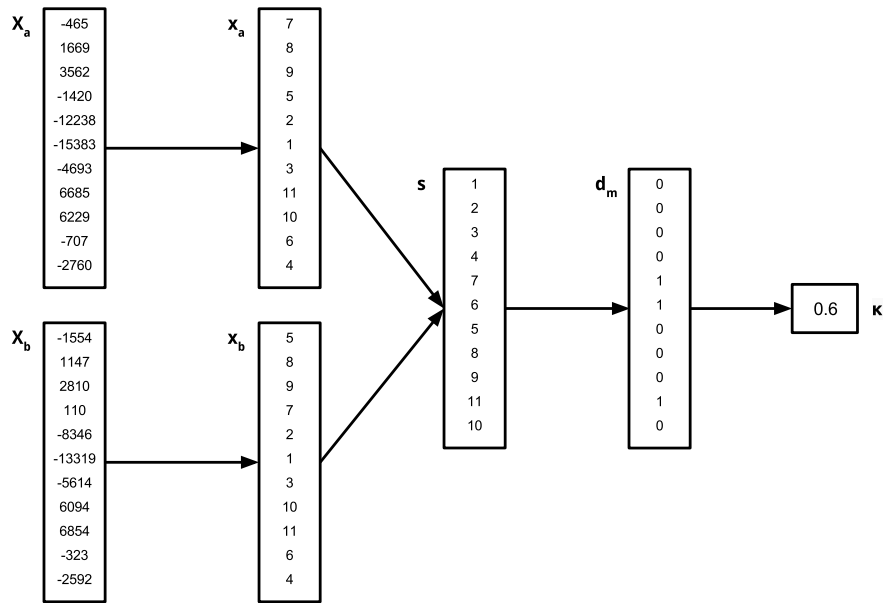


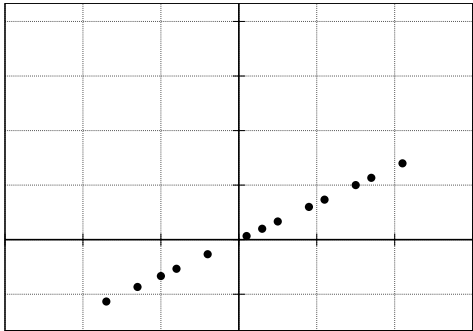
Figura A.4: Exemplo ilustrando o cálculo do coeficiente κ .

Em A.5(a) os conjuntos possuem correlação linear perfeita e as quatro métricas resultam em coeficientes iguais a um. A Figura A.5(b) ilustra o efeito da presença de um *outlier* em cada uma das medidas. Em A.5(c), o mesmo conjunto é deslocado e a correlação de cossenos é alterada mesmo quando não existem *outliers*.

A Figura A.6, por sua vez, ilustra dois conjuntos correlacionados cubicamente. Neste caso, somente as medidas baseadas na ordenação relativa das amostras são capazes de captar a correlação perfeita entre os conjuntos.

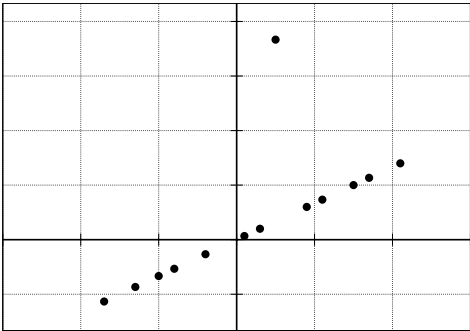
Já as Figuras A.7(a) e A.7(c) representam o diagrama de dispersão de dois pares de traços extraídos de uma seção sísmica real. Nos dois casos, os diagramas apresentam uma tendência linear, de modo que as métricas de Pearson e Cossenos se mostram eficazes em capturar a correlação entre os conjuntos. Em A.7(b) e A.7(d) uma amostra de cada diagrama foi manualmente deslocada para simular a presença de ruído. Neste caso, o coeficiente de Spearman é mais eficaz, sofrendo menos alteração em seu valor do que as medidas de correlação lineares.

É interessante notar na Figura A.7 que, quando comparado às outras métricas, o coeficiente de correlação de Bhat *et. al.* possui valor muito inferior, o que pode nos levar a acreditar que essa medida não seria apropriada para o método em questão. Porém, apesar de baixo, seu valor permanece invariante à presença dos ruídos inseridos manualmente, o que pode ser interessante no âmbito da sísmica, onde a presença de ruídos é uma característica do dado de amplitudes.



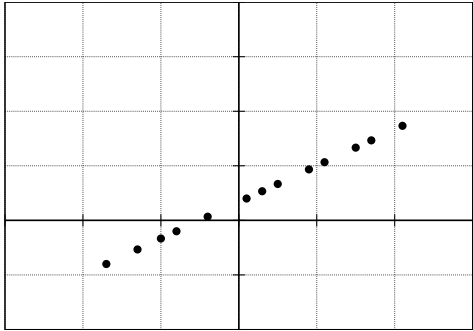
Pearson	1.0000
Spearman	1.0000
Cossenos	1.0000
Bhat <i>et. al.</i>	1.0000

(a)



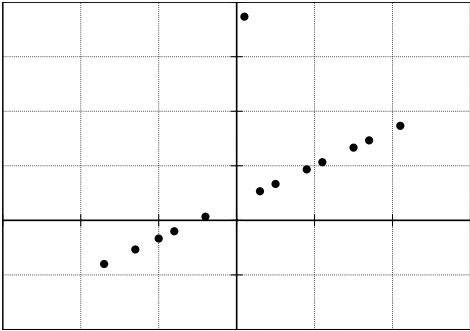
Pearson	0.5151
Spearman	0.8559
Cossenos	0.6996
Bhat <i>et. al.</i>	0.6667

(b)



Pearson	1.0000
Spearman	1.0000
Cossenos	0.9341
Bhat <i>et. al.</i>	1.0000

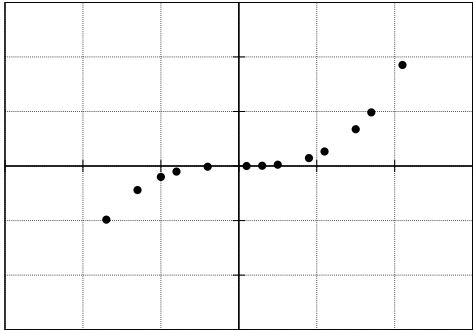
(c)



Pearson	0.4269
Spearman	0.7310
Cossenos	0.6367
Bhat <i>et. al.</i>	0.6667

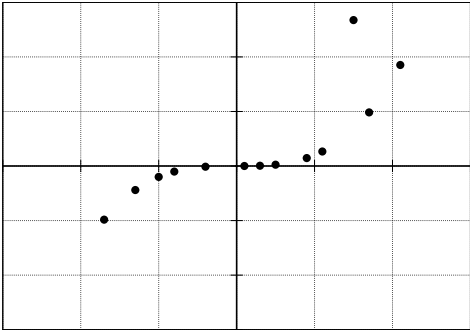
(d)

Figura A.5: Diagrama de dispersão de dois conjuntos correlacionados linearmente, onde $Y = 2 \times X$. Em (a) as quatro métricas resultam em coeficientes iguais a um e (b) ilustra o efeito de um *outlier* nos conjuntos em questão. Em (c) e (d) o mesmo conjunto é deslocado para $Y = 2 \times X + 10$.



Pearson	0.8692
Spearman	1.0000
Cossenos	0.8989
Bhat <i>et. al.</i>	1.0000

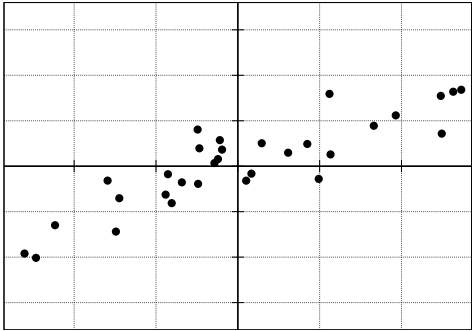
(a)



Pearson	0.7645
Spearman	0.9712
Cossenos	0.8176
Bhat <i>et. al.</i>	0.6667

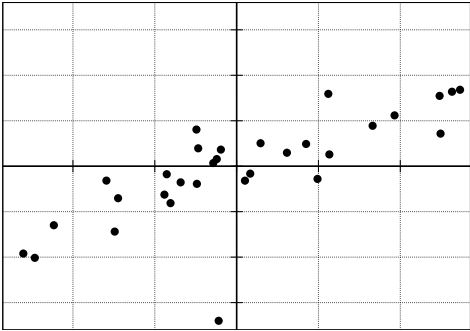
(b)

Figura A.6: Diagrama de dispersão de dois conjuntos correlacionados cubicamente. Neste caso, somente as medidas baseadas na ordenação relativa das amostras são capazes de capturar a correlação perfeita entre os conjuntos.



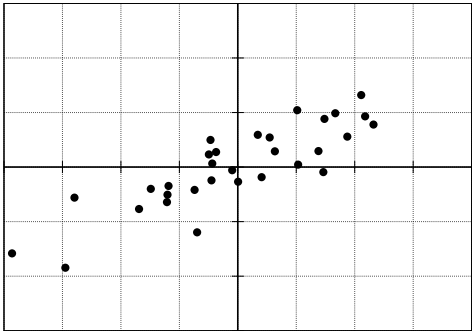
Pearson	0.8766
Spearman	0.8484
Cossenos	0.8763
Bhat <i>et. al.</i>	0.4667

(a)



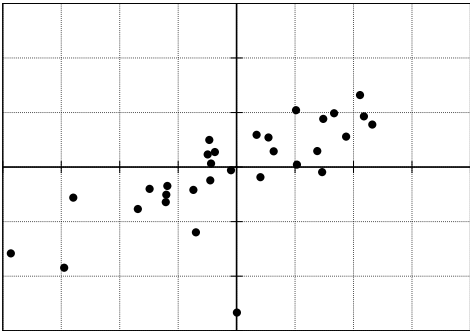
Pearson	0.7629
Spearman	0.8169
Cossenos	0.7596
Bhat <i>et. al.</i>	0.4667

(b)



Pearson	0.8549
Spearman	0.8617
Cossenos	0.8540
Bhat <i>et. al.</i>	0.4667

(c)



Pearson	0.7184
Spearman	0.8153
Cossenos	0.7185
Bhat <i>et. al.</i>	0.4667

(d)

Figura A.7: Diagrama de dispersão de dois diferentes pares de traços extraídos de uma seção real. Em (b) e (d) foram adicionados manualmente ruídos aos traços originais ilustrados por (a) e (c).