

Compressão Wavelet com truncamento global para dados sísmicos

Victor Martins Gomes¹; Marco Antonio Cetale Santos¹; Denis Lage Ferreira da Silva¹; Rodrigo Bird Burgos²; Djalma Manoel Soares Filho³

¹Departamento de Geologia e Geofísica, Instituto de Geociências, Universidade Federal Fluminense

²Departamento de Engenharia Civil, Universidade Estadual do Rio de Janeiro

³CENPES/Petrobras

Copyright 2014, SBGf - Sociedade Brasileira de Geofísica.

Este texto foi preparado para a apresentação no VI Simpósio Brasileiro de Geofísica, Porto Alegre, 14 a 16 de outubro de 2014. Seu conteúdo foi revisado pelo Comitê Técnico do VI SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem prévia autorização da SBGf.

Abstract

In this work the seismic volume compression is performed using a method with global thresholding in the wavelet transform domain. The wavelet analysis is performed by varying the level of decomposition according to the direction of filtering in order to get best results. After thresholding, the generated sparse matrix is stored in vector format using the CSV format (Compressed Sparse Vector). The results of compression are analyzed from a qualitative point of view, by analyzing the visual quality of the compressed data and quantitative data by the percentage of energy from the compressed data to the original data, the percentage of data compression, the root mean squared error, the peak signal-to-noise ratio and the mean squared error.

Introdução

A exploração, avaliação e monitoramento de reservatórios de hidrocarbonetos utiliza como principal ferramenta os dados sísmicos. No entanto, com o aumento da necessidade de hidrocarbonetos, uma maior quantidade de dados sísmicos foi e é continuamente gerada. O armazenamento destes dados se torna um problema quando dados de uma única aquisição podem atingir dimensões de terabytes. Neste contexto, a compressão de dados sísmicos se torna etapa importante, garantindo que mais dados possam ser armazenados utilizando menos espaço.

Diversos métodos de compressão de dados sísmicos existem atualmente na literatura. Podem-se citar os trabalhos: Luo and Schuster (1992), Villasenor et al. (1996) e Khene and Abdul-Jauwad (2000), onde o fluxo de compressão utilizado pelos métodos envolve o uso da transformada wavelet combinada com alguma técnica para o truncamento dos coeficientes da transformada wavelet do dado.

Os métodos de compressão se classificam em compressão com perda, caso haja uma perda de informação e compressão sem perda, no caso em que a informação é toda conservada. O método de compressão com perda descarta parte da informação dos dados através de truncamento ou codificação. Deseja-se que os erros resultantes da reconstrução do sinal estejam dentro de um padrão aceitável.

O método de compressão utilizado neste trabalho envolveu a transformação do dado sísmico para o domínio

da transformada wavelet, o truncamento (*threshold*) dos coeficientes a partir de um limiar de corte global estabelecido e o armazenamento dos valores em um formato vetorial utilizando o formato CSV (*compressed sparse vector*) de Farzaneh et. al.(2009).

Metodologia

A transformada wavelet envolve a análise multiresolução de um sinal através da expansão deste em relação a uma base de funções (uma função de escala e uma função wavelet), da mesma forma que a transformada de Fourier expande um sinal em relação a uma base ortogonal de senóides (senos e cossenos). No entanto, diferentemente da transformada de Fourier, a transformada wavelet nos permite uma representação do sinal não só no domínio da frequência, mas também permite uma localização temporal (ou espacial dependendo do domínio utilizado) permitindo uma análise local das variações do sinal analisado em vários níveis de resolução (Burgos et al. (2013)).

Através da análise multiresolução introduzida em Mallat (1989) e a condição de aninhamento dos espaços das funções, a função de escala base pode ser representada em termos de uma soma ponderada de translações da sua versão escalada por um fator 2, como apresentado na equação abaixo:

$$v(t) = \sum_n h_{lp} \sqrt{2} v(2t - n), \quad n \in \mathbb{Z} \quad (1)$$

Os coeficientes h_{lp} na equação formam o chamado filtro passa-baixa da análise e o fator $\sqrt{2}$ normaliza a função de escala por um fator dois. A função wavelet base pode ser representada por uma soma ponderada de versões transladadas da função de escala dilatada por um fator 2.

$$v(t) = \sum_n h_{hp} \sqrt{2} v(2t - n), \quad n \in \mathbb{Z} \quad (2)$$

Os coeficientes h_{hp} formam o chamado filtro passa-alta da análise.

Quando trabalhando com sinais sísmicos, imagens e sinais digitais em geral, não é necessário trabalhar diretamente com as funções de escala ou as wavelets. Neste caso, o que deve ser considerado são os filtros de análise, os respectivos de reconstrução e os coeficientes de aproximação e detalhe, os quais podem ser considerados como filtros digitais e sinais digitais, respectivamente (Vaidyanathan, 1993). Em geral a análise (decomposição) e a reconstrução (síntese) do sinal podem ser representadas como dois bancos de filtros de dois canais (Mertins, 1999), como representado na figura 2.

$$a_j(k) = \sum_m h_{lp}(m - 2k)c_{(j+1)}(m) \quad (3)$$

$$d_j(k) = \sum_m h_{hp}(m - 2k)c_{(j+1)}(m) \quad (4)$$

$$c_{(j+1)}(k) = \sum_m c_j(m)g_{lp}(k - 2m) + \quad (5)$$

$$+ \sum_m d_j(m)g_{hp}(k - 2m) \quad (6)$$

Através da observação das equações acima é possível notar que na etapa de análise, os coeficientes de aproximação e detalhe na escala de resolução j podem ser obtidos através da convolução dos coeficientes na escala de resolução $j + 1$ com os filtros $h_{lp}(-n)$ e $h_{hp}(-n)$ seguido de decimação por um fator 2 (preservando-se somente os termos pares). Quanto a etapa de síntese, os coeficientes de aproximação e detalhe em uma escala de resolução $j + 1$ são obtidos através da interpolação do sinal por um fator 2, inserindo-se zeros entre os termos, e em seguida convolvendo o sinal interpolado com os filtros $g_{lp}(n)$ e $g_{hp}(n)$. A implementação do banco de filtros composto por estas estruturas é feita de acordo com o algoritmo (algoritmo piramidal) desenvolvido em Mallat (1989) para a transformada wavelet discreta. A figura dois ilustra a implementação da transformada wavelet utilizando um banco de filtros.

No trabalho de Villasenor et al. (1995), os melhores resultados foram obtidos com os filtros biortogonais 9/7, 13/11 e 10/6. No entanto, apesar da melhor performance geral dos filtros de comprimento ímpar, eles discutem que os filtros com comprimentos par conseguem preservar melhor a localização, forma e intensidade dos impulsos, o que é uma característica desejada quando se analisa sinais sísmicos. Levando em consideração estes resultados os filtros escolhidos foram os biortogonais da classe Cohen-Daubechies-Feauveau (CDF), 9/7 e 10/6, as respostas impulsivas de seus filtros de análise são mostradas n figura 1.

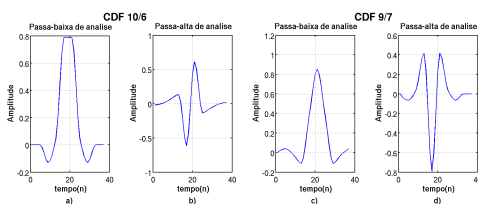


Figura 1: Respostas impulsivas dos filtros passa-alta e passa-baixa do banco wavelet de análise CDF 10/6 (a) e b)) e CDF 9/7 (c) e d)).

A escolha do banco de filtros apropriado para a compressão de dados sísmicos é uma etapa fundamental, uma vez que um filtro com uma melhor relação de perda-ganhos entre a taxa de compressão, o erro de reconstrução do sinal e o total de energia preservado no sinal, pode levar a melhores resultados.

O truncamento (*thresholding*) dos coeficientes da transformada wavelet do dado sísmico e, em geral, de qualquer tipo de dado gera uma matriz esparsa de forma que a opção óbvia é armazenar os valores diferentes de zero e suas posições. Diversas formas de armazenamento de matrizes esparsas existem. Dentre eles vale citar

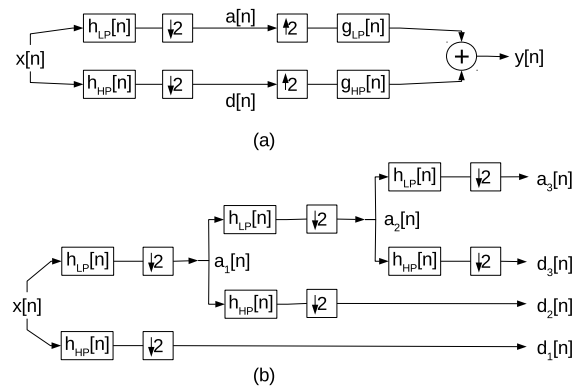


Figura 2: Implementação da transformada wavelet utilizando banco de filtros. (a) mostra a análise e síntese wavelet utilizando banco de filtros em um nível de decomposição wavelet. (b) mostra a análise wavelet utilizando banco de filtros em três níveis de decomposição.

os formatos mais comuns: CRS (Compressed Row Storage), CCS (Compressed Column Storage), MSR (Modified Sparse Row) e BCRS(Block Compressed Row Storage).

Apesar dos métodos citados serem os mais difundidos, o que mais satisfaz os objetivos deste trabalho foi o método CSV (Compressed Sparse Vector), desenvolvido no trabalho de Farzaneh et al. (2009) como tentativa de encontrar um método eficiente e de menos custo em relação aos tradicionais CRS e MSR (Modified Sparse Row).

O método representa a matriz esparsa utilizando dois vetores, um com todos os elementos diferentes de zero e um com a localização destes elementos de acordo com a contagem coluna a coluna de todos os elementos da matriz. Após armazenados todos os valores, os valores do número de linhas e colunas da matriz original são associados aos vetores de elementos e de posições, respectivamente.

Resultados

Para testar o método de compressão aqui proposto, utilizou-se um pequeno volume sísmico já submetido a etapa de processamento sísmico, contendo 407 *cross-lines*, 48 *inlines* e 1801 amostras temporais. O volume sísmico é ilustrado na figura 3.

A etapa de truncamento avaliou três porcentagens de corte para os coeficientes da transformada wavelet do dado, utilizando os dois filtros wavelet escolhidos: 90%, 92,5% e 95%. A Tabela 1 mostra o resultado da análise quantitativa da compressão.

As medidas quantitativas de erro utilizadas foram: o erro quadrático médio do dado comprimido (EQM), a razão pico de sinal/ruído (PSNR), a porcentagem de energia do dado comprimido em relação ao dado original (PE), a porcentagem de compressão do dado (PC) e a raiz quadrada do erro quadrático médio (RMS).

Utilizando um limiar de corte global de 95% foi possível atingir uma porcentagem de compressão de 6,24%

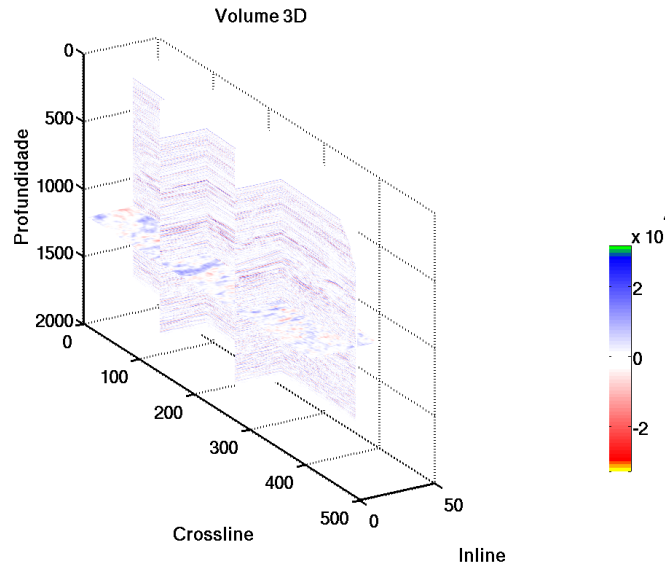


Figura 3: Volume Sísmico mostrando inline 32, crosslines 135 e 270 e slice temporal 1200

Tabela 1: Análise quantitativa da compressão

| FILTROS | CDF 10/6 | CDF 9/7 | CDF 10/6 | CDF 9/7 | CDF 10/6 | CDF 9/7 |
|-----------|--------------|--------------|--------------|--------------|--------------|--------------|
| CORTE | 90% | | 92,5% | | 95% | |
| EQM | 3,84e +05 | 3,56e +05 | 5,55e +05 | 5,14e +05 | 8,46e +05 | 7,87e +05 |
| PSNR (dB) | 67,75 | 68,08 | 66,15 | 66,49 | 64,32 | 64,63 |
| PE | 95,77% | 95,77% | 94,11% | 94,11% | 91,40% | 91,40% |
| PC | 12,49% | 12,49% | 9,36% | 9,36% | 6,24% | 6,24% |
| RMS | 619,81 | 596,61 | 744,80 | 716,70 | 919,82 | 887,27 |

Observou-se que os melhores resultados são atingidos quando os níveis de decomposição são maiores na segunda dimensão (dimensão temporal). Especificamente, utilizou-se 3 níveis de decomposição para a direção *inline*, 7 níveis de decomposição para a dimensão temporal e 4 para a direção *crossline*. Aumentando-se ainda mais o nível de decomposição para a segunda dimensão observa-se um aumento gradual mas muito lento no percentual de energia retido e razão de pico sinal/ruído. Como consequência este aumento gerará aumento na porcentagem de compressão, desta forma, observou-se que até 9 níveis de decomposição obtém-se uma boa relação entre a energia retida, a PSNR e o tamanho do sinal comprimido.

A análise qualitativa do dado comprimido permite observar que em geral, a qualidade visual, apesar da presença de artefatos gerados pela compressão, é bem preservada. As feições mais significativas das seções sísmicas são bem preservadas, não se notando diferenças entre estas no dado comprimido e no não comprimido. A análise qualitativa não produz resultados significativos a respeito da diferença entre o uso dos dois filtros, uma

vez que a qualidade visual do dado comprimido com o CDF 10/6 é, visualmente, indistinguível ao comprimido pelo CDF 9/7. As figuras 4 e 5 mostram o *inline* 32 do volume sísmico original e do volume comprimido e, também, uma comparação entre os traços 135 destes *inlines*, onde em azul é o traço original e em vermelho o comprimido.

Através de uma análise quantitativa dos resultados observa-se que os dois filtros conseguem preservar a mesma porcentagem de energia no sinal comprimidos enquanto atingindo a mesma porcentagem de compressão para um mesmo limiar de corte. No entanto, em geral, o filtro CDF 9/7 acaba por se mostrar mais apropriado para a compressão de dados sísmicos em relação ao filtro CDF 10/6, obtendo maiores razões de pico sinal/ruído e possuindo menores erros RMS e quadrático médio.

Os resultados foram similares aos encontrados em Villaseñor et al. (1995), onde o filtro CDF 9/7 obteve melhores resultados para compressão que o CDF 10/6, mas no caso deste trabalho a análise de compressão foi realizada para dados sísmicos não imagens.

A maior porcentagem de compressão do dado (6.24%) apresentou os resultados menos satisfatórios, de um ponto de vista quantitativo enquanto foi capaz de preservar as feições mais importantes do dado sísmico e, apesar da presença de artefatos gerados pela compressão, o dado comprimido é, de um ponto de vista visual, praticamente indistinguível do original.

Conclusões

A transformada wavelet de um sinal contém as informações mais relevantes a respeito do sinal armazenadas em um número pequeno de coeficientes em relação ao total de coeficientes da transformada. Esta propriedade esta associada ao uso da transformada wavelet

em diversos métodos de compressão existentes na literatura (Villasenor et al., 1996; Khene and Abdul-Jauwad, 2000), uma vez que o truncamento dos coeficientes menos significantes não influenciará em grandes proporções as características do sinal reconstruído em relação ao original.

O dado sísmico é, em geral, um dado de banda limitada e dominada pelas menores frequências. Desta forma um método de compressão baseado na transformada wavelet se torna útil uma vez que as informações mais significantes do dado se encontram nos coeficientes da transformada associados a filtragem passa-baixa. Estes coeficientes também são os que apresentam os maiores valores entre os coeficientes, assim o truncamento dos menores coeficientes da transformada wavelet será responsável por descartar a informação redundante associada aos coeficientes de detalhe relacionados a filtragem passa-alta.

O método de compressão utilizado neste trabalho envolveu o truncamento dos menores coeficientes da transformada wavelet do sinal baseado em um limiar de corte global (90, 92,5 e 95 %).

A comparação entre os resultados da compressão utilizando os filtros CDF 9/7 e 10/6 mostrou que o CDF 9/7, em geral, atinge melhores resultados. No entanto a análise visual dos dados comprimidos, utilizando os dois filtros, praticamente não apresentou distinções entre os resultados para os limiares de corte utilizados neste trabalho, de forma que os dados comprimidos ficaram muito semelhantes ao original para todos os limiares.

A compressão foi responsável por retirar em grande parte a informação redundante do dado sísmico, de forma que as feições mais significantes nas seções sísmicas se tornam mais evidentes. Desta forma, um resultado aceitável com porcentagem de compressão de 6,24% foi atingido.

Os resultados apresentados são os preliminares da pesquisa em compressão de dados sísmicos utilizando a transformada wavelet. Atualmente, métodos de codificação estão sendo analisados a fim de substituir o truncamento global, além de obter melhores porcentagens de compressão.

Agradecimentos

O primeiro autor deste trabalho gostaria de agradecer ao PIBIC/UFF pelo auxílio financeiro recebido através de bolsa de iniciação científica. Os autores gostariam de agradecer a PETROBRAS ao suporte dado ao Laboratório de Imageamento Sísmico e Inversão Sísmica (ISIS) do Departamento de Geologia e Geofísica da Universidade Federal Fluminense.

Referências

- Burgos, R. B., M. A. C. Santos, and R. R. e Silva, 2013, Deslauriers–dubuc interpolating wavelet beam finite element: Finite Elements in Analysis and Design, **75**, 71 – 77.
- Farzaneh, A., H. Kheiri, Mehdi, and A. Shahmersi, 2009, An efficient storage format for large sparse matri-

ces: Commun.Faculty of Science. University of Ankara.Series A1: Mathematics and Statistics.

- Khene, M., and S. H. Abdul-Jauwad, 2000, Adaptive seismic compression by wavelet shrinkage: Statistical Signal and Array Processing, 2000. Proceedings of the Tenth IEEE Workshop on, 544–548.
- Luo, Y., and G. Schuster, 1992, Wave packet transform and data compression: SEG Technical Program Expanded Abstracts, 1187–1190.
- Mallat, S., 1989, A theory for multiresolution signal decomposition: the wavelet representation: Pattern Analysis and Machine Intelligence, IEEE Transactions on, **11**, 674–693.
- Mertins, A., 1999, Signal analysis: Wavelets, filter banks, time-frequency transforms and applications: John Wiley & Sons, Inc. (330 páginas).
- Vaidyanathan, P., 1993, Multirate systems and filter banks: Prentice Hall. Prentice Hall Signal Processing Series. (944 páginas).
- Villasenor, J., B. Belzer, and J. Liao, 1995, Wavelet filter evaluation for image compression: Image Processing, IEEE Transactions on, **4**, 1053–1060.
- Villasenor, J., R. A. Ergas, and P. L. Donoho, 1996, Seismic data compression using high-dimensional wavelet transforms: Data Compression Conference, 1996. DCC '96. Proceedings, 396–405.

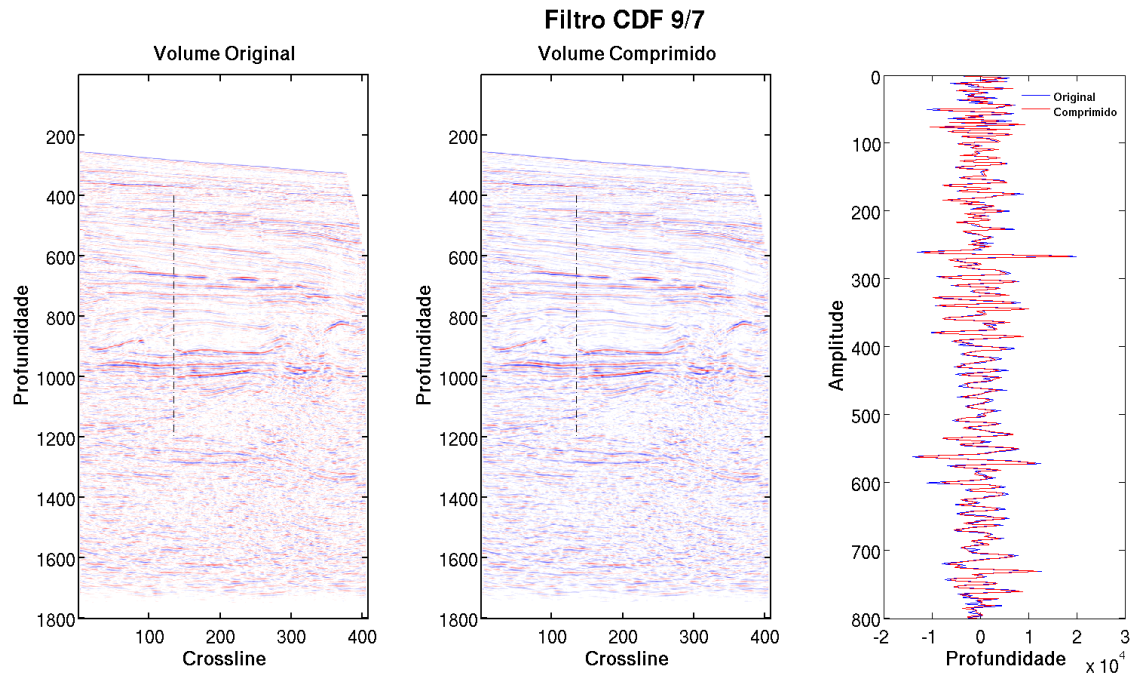


Figura 4: Comparação entre os inlines 32 do volume original (imagem mais a direita) e comprimido (imagem central) utilizando CDF 9/7. A imagem mais a esquerda mostra uma comparação entre os traços 135 destes inlines, de azul o traço original e de vermelho o comprimido.

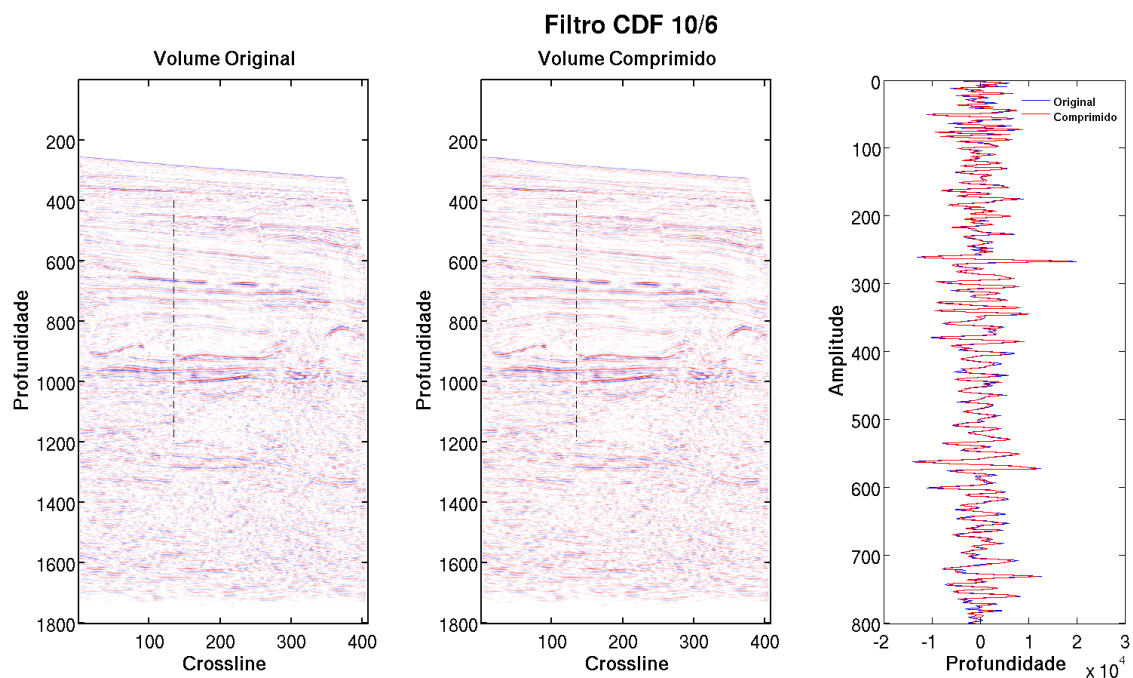


Figura 5: Comparação entre os inlines 32 do volume original (imagem mais a direita) e comprimido (imagem central) utilizando CDF 10/6. A imagem mais a esquerda mostra uma comparação entre os traços 135 destes inlines, de azul o traço original e de vermelho o comprimido.