Submission code: ZP5XZQWYVB

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Seismic Data Compression with OBNZip: Implications for 4D Seismic Analysis

Josafat Leal Filho, Pablo Machado Barros (Cenpes-Petrobras), Mateus Lucena (Software/Hardware Integration Lab - UFSC), André Bulcão (Petrobras), Matheus Wagner (Software/Hardware Integration Lab - UFSC), Alexandre Augusto Cardoso da Silva (CENPES - Petrobras), Antônio Augusto Fröhlich (Software/Hardware Integration Lab - UFSC), Ubiratan Jose Furtado (Petrobras)

Seismic Data Compression with OBNZip: Implications for 4D Seismic Analysis

Please, do not insert author names in your submission PDF file.

Abstract Summary

Efficient data storage and transmission in geophysical workflows is highly dependent on seismic data compression due to the massive multi-dimensional dataset involved. This work shows the application of OBNZip, a scalable modular compression framework that leverages the sparsity, non-stationarity, and spatial redundancy found in seismic data. We focus on a configuration of the compressor utilizing Discrete Wavelet Transform (DWT), coefficient thresholding, fixed-point quantization, and entropy encoding, chosen after extensive evaluation with passive, microseismic, and active-source datasets from the accelerometer data from Ocean Botton Sensor. Performance analysis shows that compression efficiency scales with spatial dimension. OBNZip achieves high compression ratios—up to 5:1- while maintaining signal fidelity with 4D reconstruction relative errors as low as 2×10^{-3}. And offers a robust, adaptable solution for seismic data compression, with future extensions aimed at integrating machine learning-based encoders and supporting real-time, low-power geophysical deployments.

Introduction

In the context of oil and gas exploration, 4D seismic is an indispensable tool for monitoring temporal reservoir changes to optimize recovery strategies and enhance operational efficiency. The sophistication of modern acquisition technologies allows for the detection of subtle variations in acoustic impedance, on the order of 2% or less Cruz et al. (2021). This high resolution, however, generates massive datasets that present significant storage and transmission challenges, making effective data reduction critical for lowering costs and accelerating analysis to enable more agile decision-making.

Current data compression strategies range from traditional signal processing techniques, such as the Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and dreamlet transform exploit the energy compaction properties of seismic signals in transformed domains Averbuch et al. (2001); Fajardo et al. (2015); Geng et al. (2009); Radosavljević et al. (2020) and sampling-based approaches Balakrishnan and Sheeba (2020); Rubin et al. (2016), to modern machine learning models like autoencoders and recurrent neural networks Helal et al. (2021); Horstmann et al. (2022); Nuha et al. (2019). While traditional methods often require extensive parameter tuning and domain expertise, machine learning solutions can incur high computational costs and may lack robustness across diverse conditions. This highlights a persistent need for lightweight, adaptive architectures suitable for real-time deployment in low-power, resource-constrained environments Hoffmann and Fröhlich (2022).

To address these challenges, we propose OBNZip, a modular and scalable seismic data compression framework tailored to the structural properties of seismic signals. OBNZip combines classical and modern strategies through a multi-stage pipeline, for example, involving Discrete Wavelet

Transform, coefficient thresholding, fixed-point quantization, and entropy encoding. Its component-based design allows flexible deployment across CPUs, GPUs, and embedded platforms, making it adaptable to varied operational contexts. We evaluate OBNZip using passive, microseismic, and active-source data sets from the accelerometer data from Ocean Botton Sensor, demonstrating that it achieves compression ratios up to 5:1 with minimal information loss, as indicated by 4D reconstruction relative errors as low as 2×10^{-3}. With the correct configuration, the framework preserves essential geological features, making it suitable for time-lapse imaging. Additionally, performance analysis shows that its efficiency scales with spatial dimensions, and that implementation trade-offs between CPU and GPU configurations can be optimized based on workload. OBNZip thus provides a robust foundation for seismic data reduction and opens pathways for future integration with learning-based encoders and deployment in resource-constrained geophysical systems.

System Architecture

The proposed architecture presents a modular, component-based framework for Big Data Systems (BDS), designed for adaptability across varied data formats and deployment environments. At its core is the Data_Segment abstraction, which unifies individual data points with compression metadata to support uniform processing in heterogeneous contexts. The system accommodates both push- and pull-based data transmission and scales via horizontal and vertical parallelism. Input and output adapters handle data ingestion and output conversion, enabling format-agnostic operation while preserving metadata required for decompression. Compression is performed through configurable, multi-stage pipelines asynchronously in separate threads and connected by thread-safe queues. Each stage implements both compress() and decompress() methods to ensure symmetry and enable automated from a given compression configuration. The architecture's emphasis on modularity, configurability makes it highly suitable for edge, cloud, and hybrid environments.

Methodology

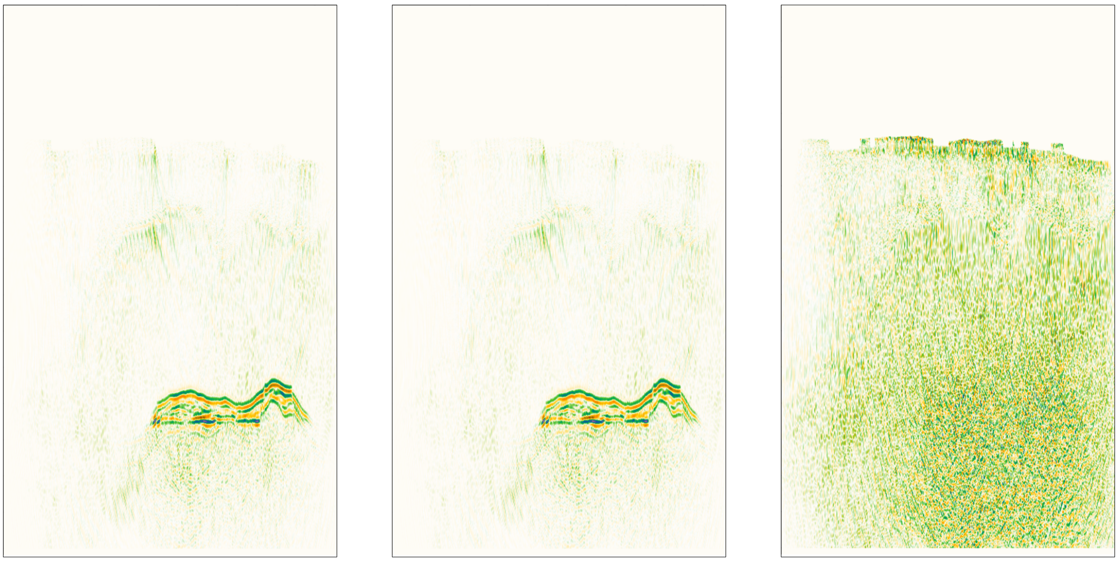
We selected datasets encompassing different acquisition conditions, such as, passive monitoring, microseismic activity, and active-source experiments from accelerometer data from Ocean Botton Sensor, to evaluate the compression framework. The dataset was curated to reflect varying signal complexities and entropy levels to assess the adaptability and robustness of the framework. Passive data contained examples of entropy extremes, where the measured entropy represents environment complexity. Microseism data contained a peak in natural seismic noise to evaluate performance under ambient, high-entropy conditions. Active data comprised sustained and continuous seismic acquisition to assess performance in structured, high-volume acquisition scenarios. All data segments were processed using a standardized compression pipeline, beginning with a Discrete Wavelet Transform (DWT) followed by a thresholding step. During compression, the values less than 0.25% of DWT coefficients with the highest absolute values were removed to enforce sparsity while preserving critical high-energy components.

To check the implications on 4D seismic, we obtained 4Ds images through the difference of seismic 3D images obtained by reverse time migration (RTM) from the base and monitor cases after compression/decompression of synthetic seismograms.

Results

The analysis of compression performance across data types, computational devices, and spatial dimensions reveals consistent trends in efficiency and scalability. The compression ratios increase with spatial dimension for types of data, highlighting the ability of the algorithm to exploit redundancy and spatial correlations in larger datasets. While decompression bit rates also decline with larger spatial dimensions, the reduction is more modest. Device-specific performance shows CPUs generally provide higher compression ratios and faster decompression, especially at smaller scales, whereas GPUs excel at medium to large scales due to their parallelism. Compression bit rates on GPUs eventually surpass those of CPUs.

Compression outcomes also vary significantly by data type. Passive data achieves the highest compression ratios, up to 78.1 on the CPU, indicating substantial spatial redundancy and resulting in low compression bit rates. Microseism data shows strong improvements with scale, with compression ratios increasing from 2.65 to 45.8 and bit rates dropping from over 1125 Mb/s to 26.4 Mb/s. Active data demonstrates more gradual gains. These findings underscore the need to align compression strategies with both data characteristics and hardware capabilities.



(a) 4D seismic anomaly of original dataset

(b) 4D seismic anomaly of decompressed dataset

(c) Difference of 4D seismic anomaly

Figure 1: Comparison of migrated 4D seismic anomaly sections. (a) Anomaly derived from the original dataset; (b) anomaly obtained from the decompressed dataset; and (c) the absolute difference between the original and decompressed anomalies. This visual analysis highlights the fidelity of the decompression process in preserving key seismic features.

Additionally, as mentioned in the introduction, maintaining subtle amplitude variations is critical in 4D seismic analysis. For a particular case, the compression method achieves a relative error of approximately 2×10^{-3} between the 4D obtained with the original synthetic seismograms and the 4D obtained with the same seismograms after compression/decompression. The compression ratio was 5:1, with residuals resembling random noise and preserving key structural features. This confirms the method's ability to retain signal integrity, crucial for accurately detecting subsurface changes. Figure 1 presents the original dataset, the decompressed version, and their corresponding difference. It is evident that the 4D anomaly is perceptible in both the original and decompressed datasets, with the primary structural features preserved. These features are manifested as parallel or quasi-parallel reflectors. In contrast, the difference image exhibits a noisy appearance apparently devoid of any geological coherence, further validating the fidelity of the data compression and reconstruction process.

Other tests were performed with inclusion of swell-type noise extracted from real data in the base and monitor seismograms. The 4D signals are satisfactorily preserved for signal-to-noise ratios similar to those of real seismic data, indicating a good reconstruction of the signals of interest. However, it is important to emphasize that the 4Ds for the cases with noise were obtained without pre-processing to attenuate this noise. In some cases, noise can affect the process of obtaining the seismic images and the corresponding 4D.

Conclusion

OBNZip proves to be effective for seismic data compression in embedded, workstation, and server environments. It offers adaptable pipelines, good compression ratios, and low overhead. For 4D seismic with synthetic seismograms, it maintains integrity when configured for moderate to low-loss compression. Investigations with actual seismic acquisition data will then be necessary to validate the method.

References

Averbuch, A., F. Meyer, J.-O. Stromberg, R. Coifman, and A. Vassiliou, 2001, Low bit-rate efficient compression for seismic data: IEEE Transactions on Image Processing, **10**, 1801–1814.

Balakrishnan, P., and V. Sheeba, 2020, Compressed encoding of vibration signals using extremum sampling: SN Applied Sciences, **2**.

Cruz, N. M., J. M. Cruz, L. M. Teixeira, M. M. da Costa, L. B. de Oliveira, E. N. Urasaki, T. P. Bispo, M. de Sá Jardim, M. H. Grochau, , and A. Maul, 2021, Tupi nodes pilot: A successful 4d seismic case for brazilian presalt reservoirs: The Leading Edge, **40**, 886–896.

Fajardo, C., O. Reyes Torres, and A. Ramirez, 2015, Seismic data compression using 2d lifting-wavelet algorithms: Ingeniería y Ciencia, **11**, 221–238.

Geng, Y., R. Wu, and J. Gao, 2009, *in* Dreamlet transform applied to seismic data compression and its effects on migration: SEG, 3640–3644.

Helal, E. B., O. M. Saad, A. G. Hafez, Y. Chen, and G. M. Dousoky, 2021, Seismic data compression using deep learning: IEEE Access, **9**, 58161–58169.

Hoffmann, J. L. C., and A. A. Fröhlich, 2022, Embedding anomaly detection autoencoders for wind turbines: Presented at the 2022 IEEE 27th Conference on Emerging Technologies & Factory Automation (ETFA), IEEE.

Horstmann, L. P., M. Wagner, R. M. Scheffel, and A. A. Fröhlich, 2022, Dealing with incomplete datasets with a confidence attribution algorithm: Measurement, **199**, 111509.

Nuha, H. H., A. Balghonaim, B. Liu, M. Mohandes, and F. Fekri, 2019, Seismic data compression using deep neural network predictors: Presented at the SEG Technical Program Expanded Abstracts 2019, Society of Exploration Geophysicists.

Radosavljević, M., Z. Xiong, L. Lu, D. Hohl, and D. Vukobratović, 2020, High bit-depth seismic data compression: A novel codec under the framework of hevc: IEEE Access, **8**, 114443–114459.

Rubin, M. J., M. B. Wakin, and T. Camp, 2016, Lossy compression for wireless seismic data acquisition: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, **9**, 236–252.