algorithms

MDPI

Article

Multi-Component Temporal-Correlation Seismic Data Compression Algorithm Based on the PCA and DWT

Mateus Martinez de Lucena *,†, **Josafat Leal Ribeiro** †, **Matheus Wagner** † and **Antônio Augusto Fröhlich** †

Software/Hardware Integration Lab, Federal University of Santa Catarina, Florianópolis 88040-900, Brazil; josafat@lisha.ufsc.br (J.L.R.); wagner@lisha.ufsc.br (M.W.); guto@lisha.ufsc.br (A.A.F.)

* Correspondence: lucena@lisha.ufsc.br; Tel.: +55-48-3721-9516

† Current address: UFSC/CTC/LISHA, P.O. Box 476, Florianópolis 88040-900, Brazil.

Abstract: Industrial application data acquisition systems can be sources of vast amounts of data. The seismic surveys conducted by oil and gas companies result in enormous datasets, often exceeding terabytes of data. The storage and communication demands these data require can only be achieved through compression. Careful consideration must be given to minimize the reconstruction error of compressed data caused by lossy compression. This paper investigates the combination of principal component analysis (PCA), discrete wavelet transform (DWT), thresholding, quantization, and entropy encoding to compress such datasets. The proposed method is a lossy compression algorithm tuned by evaluating the reconstruction error in frequency ranges of interest, namely 0–20 Hz and 15–65 Hz. The PCA compression and decompression acts as a noise filter while the DWT drives the compression. The proposed method can be tuned through threshold and quantization percentages and the number of principal components to achieve compression rates of up to 31:1 with reconstruction residues energy of less than 4% in the frequency ranges of 0–20 Hz, 15–65 Hz, and 60–105 Hz.

Keywords: compression algorithms; principal component analysis; discrete wavelet transform

check for updates

Academic Editor: Abdulsalam Yassine

Received: 10 December 2024

Revised: 28 December 2024

Accepted: 30 December 2024

Published: 9 January 2025

Citation: Lucena, M.M.d.; Ribeiro, J.L.; Wagner, M.; Fröhlich, A.A. Multi-Component Temporal-Correlation Seismic Data Compression Algorithm Based on the PCA and DWT. *Algorithms* **2025**, *18*, 33. https://doi.org/10.3390/a18010033

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Decision making in industrial applications relies heavily on datasets and data acquisition systems. Statistical analysis for complex systems requires representative datasets and high data volume. Seismic data analysis [1] and data-driven prognostics [2,3], for example, are applications with an increased need for high-frequency data sampling (i.e., vibration, acoustics) that increase storage demands. In specific application scenarios, such as seismic surveys conducted by oil and gas companies, seismic data can amount to tens of terabytes [4]. Consequently, storing and transferring such vast amounts of data pose significant cost and technical challenges [5]. Data in analytics can contain less relevant components depending on the application. In these cases, pruning the dataset can be considered to lower storage, that is, filtering less relevant data. While it is possible to automate such a process, it is challenging to balance meaningful storage reduction and maintain a relevant dataset. Therefore, these datasets tend to be compressed for storage using some form of classical lossless compression, even though lossless compression provides little benefit in reducing storage space.

The seismic signal can be expressed as a linear combination of wavelets with varying bandwidths and time lags, contingent upon subsurface interactions with the source. The use of wavelet decomposition, extensively explored and recognized for its compression applications [6], and principal component analysis (PCA), a widely used method machine learning (ML) in multivariate analysis for dimensionality reduction, have both been studied. The utilization of these techniques in seismic data compression has been documented [7]. Both algorithms aim to generate sparse data representations using as few coefficients as possible for effective compression. This involves transforming and discarding the least representative data and efficiently encoding the remaining information. In this work, we propose the usage of PCA as a filter to reduce the noise associated with seismic data, allowing discrete wavelet transform (DWT) to achieve higher compression rates. The main contributions of this work are as follows:

- The modeling of a compression algorithm that combines PCA and DWT while employing thresholding, quantization, and entropy encoding strategies to increase compression performance;
- The analysis of a compression case study on seismic data from the Jubarte permanent reservoir monitoring.

The remainder of this work is organized as follows: Section 2 presents the related works found in literature regarding the applications of principal component analysis (PCA), DWT, and their combination for data compression applications. Section 3 discusses the fundamentals of data compression using PCA, DWT, quantization, and thresholding, and explains the proposed data compression strategy. Section 4 presents the case study considered to evaluate the performance of the proposed compression strategy when applied to seismic data acquired during the monitoring process of an oil reservoir, as well as the results and discussion. Finally, Section 5 presents the concluding remarks of this work.

2. Related Work

Encoding through transformation represents a highly effective approach to data compression that is widely utilized in various compression tools like JPEG, JPEG-2000, and HEVC [8–10]. The encoding strategies differ based on the type of transform utilized. Commonly employed transforms in data compression include the discrete cosine transform (DCT) [11,12], the discrete wavelet transform (DWT) [6], and the family of lapped transforms [13,14]. Moreover, in seismic data compression, domain-specific transforms such as the seislet transform [15], the dreamlet transform [16], and the urvelet transform [17] still play significant roles; however, they are also very costly operations.

Dimensionality reduction-based compression techniques involve converting data from an N-dimensional space to an M-dimensional space, where $M < N$. This transformation exploits redundancy within the original N-dimensional representation. By reducing dimensionality, the method lowers the number of bits needed to encode the initial data while retaining a substantial portion of its information. Consequently, it falls under the category of lossy compression schemes.

Although principal component analysis (PCA) is limited to linear transformations, when applied to compression, it achieves reduced storage demands with a lower computational burden compared to intensive training procedures, such as machine learning-based algorithms. In seismic data compression, PCA has been effectively utilized alongside specific domain-specific enhancements. These include incorporating a pre-processing stage where signals are aligned to optimize cross-correlation [7] and implementing distributed approaches for real-time compression during the acquisition phase [18].

Nuha et al. [7] explore PCA as a dimensionality reduction solution for seismic data. Their results point to a compression ratio of 12:1 while preserving 99% of the signal energy with 0.8% reconstruction error. Most of the compression error prevails from aligning traces with reference data, which allows for reducing the number of principal components. Compression to PCA comprehends finding a pair of eigenvectors and eigenvalues and selecting the highest eigenvalues to promote a projection of the values into the vector. Data decompression can be obtained by projecting the projection coefficients back to the original dimension and adding the average vector. Aligning seismic traces, for instance, shifting traces over time such that patterns are in the same time index (e.g., reflections), promotes a higher uniformity that can improve the compression ratio. In this sense, the authors apply a circular time shift to move patterns inside data (i.e., the trace length is unchanged). Pattern alignment is promoted by maximum amplitude and cross-correlation techniques. Pattern alignment is also an interesting technique that can be applied to other approaches, such as autoencoders, and when searching for compression over time or space. Nevertheless, data alignment based on linear events may lead to the loss of alignment of hyperbolic events, possibly affecting further analysis.

Principal component analysis (PCA) can be enhanced through various combinations with transform encoding techniques. In the literature, PCA has been integrated with the discrete wavelet transform (DWT) as a preliminary filter before transform encoding [19]. Additionally, the DWT has been utilized as a preprocessing stage for PCA, leveraging its ability to decompose the original signal into smaller-dimensional signals to enhance the computational efficiency of PCA. Similarly, PCA has served as a preprocessing step in an image compression workflow, functioning as a pre-filter before employing DCT-based transform encoding, along with quantization and thresholding. Conversely, the DCT has been employed as a preprocessing step for PCA, resulting in loadings that may offer more meaningful insights for operations on compressed data, such as classification purposes [20].

3. Proposed Method

This work proposes the combination of PCA and DWT as a compression algorithm for high-frequency sampled data. This section briefly describes the fundamentals of both algorithms and how their combination leads to a better compression performance.

The combination of the PCA and DWT algorithms, with the support of quantization, entropy encoding, and thresholding, is illustrated in Figure 1 and in pseudo-code in Algorithm 1. The algorithm takes as parameters the principal components count (pc), the energy threshold (et), the quantization level (quant), and the time series to be compressed (data). The first three steps of compression act as a noise filter through PCA and quantization. PCA is a dimensionality reduction operation where the data is represented in principal components. We utilize a truncated form of the PCA, explained in more detail below, that can approximate the original signal even with a reduced number of principal components. In line 1, the number of principal components (pc) determines how many are kept for later reconstruction. The thresholding and quantization remove less representative data from the sparse representation. In line 2, the quantization and thresholding operation verifies each component of the PCA output and changes them to zero according to the specified threshold. In line 3, the PCA decompression returns the signal to the time domain. The DWT, in line 4, performs the wavelet decomposition of the time series and discards wavelets with the least representative wavelet coefficients. The last step, in line 5, is entropy encoding that optimizes the binary representation of the compressed data. The more straightforward decompression steps are shown in Algorithm 2. In line 1, entropy decoding decompresses the wavelet coefficients and, in line 2, the DWT decompression (also named inverse DWT) reconstructs the time series from the wavelet coefficients. The following sections describe the components of the algorithm in more detail.

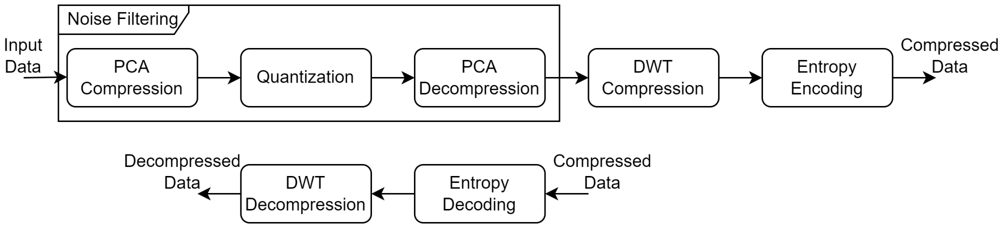

Figure 1. PCA + DWT compression and decompression workflow.

Algorithm 1 PCA + DWT compression algorithm.

Input: pc, et, quant, data
Output: compressed_data

```
temp = PCA_compress(pc, data)
temp = Quantization(quant, temp)
temp = PCA_decompress(temp)
temp = DWT_compress(et, temp)
compressed_data = Entropy_encoding(temp)
return compressed_data
```

Algorithm 2 PCA + DWT decompression algorithm.

Input: compressed_data
Output: decompressed_data

```
temp = Entropy_decoding(compressed_data)
temp = DWT_decompress(temp)
return temp
```

3.1. Principal Component Analysis

The application of PCA is a machine learning (ML) algorithm that aims to identify a basis for an N-dimensional vector space where the elements are uncorrelated and arranged based on their ability to explain the variance within the data [21]. By discarding components that contribute less to explaining the dataset's variance, PCA facilitates data compression through dimensionality reduction while preserving the most pertinent information. The original vector T is decomposed into a matrix of principal components X and a weights matrix W, as shown in Equation (1). This maps the dataset vector x from the original space p to a new space $\hat{p}$, where W is a p-by-p matrix of weights whose columns are the eigenvectors of Equation (2). The compression happens as only the first L eigenvectors are kept, producing the truncated transformation shown in Equation (3); the matrix T_L has n rows but only L columns [22].

$$T = XW \tag{1}$$

$$X^T X \tag{2}$$

$$T_L = XW_L \tag{3}$$

The specific decomposition used in this paper is the singular value decomposition (SVD) in Equation (4), where Σ is an n-by-p rectangular diagonal matrix of positive number $\sigma_{(k)}$ called the singular values of X; U is an n-by-n matrix, the columns of which are orthogonal unit vectors of length n called the left singular vectors of X; and W is a p-by-p matrix whose columns are orthogonal unit vectors of length p and called the right singular vectors of X. We can then obtain a truncated $n \times L$ matrix T_L through Equation (5) [23].

$$X = U\Sigma W^T \tag{4}$$

$$T_L = U_L \Sigma_L = XW_L \tag{5}$$

3.2. Discrete Wavelet Transform

The DWT is a process that quantifies how much energy is contained in specific frequency bands at particular times within a signal. The wavelet is a complex bounded signal, i.e., it integrates to zero and declines to zero amplitude at some distance. The decomposition of the signal creates a collection of daughter wavelets. The daughter wavelets are compressed or expanded versions of the mother wavelet, and each daughter wavelet extends across a different section of the original signal. Daughter wavelets are organized in hierarchies where wavelets in level n_i have half as many wavelets, with half the frequency and twice the length of the wavelets in level n_{i-1}. The number of levels increases until there are as many wavelets as samples in the original signal. Each daughter wavelet in this collection is associated with a coefficient that relates it to the decomposed signal.

DWT compression happens as the discretized signal is represented only with the coefficients of the daughter wavelets. The less representative coefficients, i.e., coefficients of wavelets that show a lower relation to the original signal, are filtered, thus resulting in a signal with a sparse representation. Filtering the coefficients can be achieved through various methods; we utilize thresholding and quantization in this paper.

3.3. Thresholding

In data compression, thresholding is a process where samples of a digital discrete time signal with an absolute value lower than a specified threshold are set to zero. This threshold might align with the lowest absolute quantization level, particularly when quantization is also applied. In the compression method introduced in this work, the threshold is determined as a percentage of the maximum absolute value found among the signal samples. This threshold is then applied to both the loadings and transform coefficients obtained from the PCA procedure. Different thresholds are applied to each due to the anticipated substantial differences in the numerical value ranges of the elements within each matrix.

3.4. Quantization

In digital signal processing, quantization is mapping the co-domain of a signal into another set, the quantized co-domain, with fewer elements. While there are many methods of calculating quantization, this work employs uniform quantization. The distance between two quantization levels, the inter-level distance, is parametrized as a percentage of the sample standard deviation of the signal's constituent samples. As part of the proposed compression algorithm comprises the loadings and transform coefficient matrices resulting from PCA, quantization is applied to both matrices with individual inter-level distances.

3.5. Entropy Encoding

Entropy encoding is the final step in the proposed compression algorithm. Entropy encoding is a lossless compression scheme that minimizes the amount of bits necessary to represent data. The chosen implementation is the Deflate algorithm, used in many well-known file formats, such as PNG and zip, and libraries, such as gzip and zlib [24]. The Deflate algorithm uses a combination of LZ77 and Huffman coding [25]. The LZ77 algorithm builds a dictionary and replaces repeated bit strings by references to the dictionary. Huffman coding is used to represent the most frequent symbols with the least amount of bits.

4. Case Study

A case study was performed to assess the compression efficacy of the proposed approach using a dataset of active seismic data acquired during an oil and gas reservoir monitoring survey employing ocean bottom nodes (OBNs). The following sections will offer an overview of the dataset employed for evaluation, the experiments conducted, the performance evaluation metrics employed, and the findings from the experiments, accompanied by a discussion of the results.

4.1. Dataset Description

The dataset utilized comprises raw seismic traces structured within a SEG-Y file. These traces encompass pressure measurements acquired via hydrophones mounted on permanent reservoir monitoring (PRM), strategically positioned across an irregular grid on the seafloor during an oil and gas reservoir monitoring survey. The Jubarte PRM (permanent reservoir monitoring) records 441 shot lines, spaced 25 m apart, with a shot interval of 25 m [26]. The files used for analysis come from 171 OBNs, each containing 5120 samples acquired at a 500 Hz sampling rate. The proposed strategy was applied to a single batch of 2000 traces while varying parameters of the compression algorithms.

4.2. Experiments

To evaluate the performance of the proposed method, we performed multiple compressions using the individual algorithms and the proposed combination. The three configurations were PCA configured through principal components and quantization, DWT configured through energy threshold and quantization, and PCA with DWT configured with all three. All configurations were iterated with principal components ranging from 5 to 50 in increments of 5 and quantization and energy threshold ranging from 0 to 10%. We also compared the algorithms with the ZFP compressor [27,28], which has been used in large dataset compression applications, and the SZ3 compressor [29–31] for scientific datasets.

4.3. Compression Performance Metrics

The performance of the compression algorithms was evaluated by the compression rate (CR), normalized residue energy (NRE), and normalized filtered residue energy (NFRE) metrics. The CR is defined as the ratio between the number of bytes that compose the input signal and the number of bytes that compose the output signal, defined in Equation (6), in which P is the number of bytes of the input signal and R is the number of bytes in the output signal.

$$CR = \frac{P}{R} \tag{6}$$

The NRE is the ratio between the energy of the residue, the difference between the original signal and the decompressed signal, and the energy of the original signal, as defined in Equation (7), in which x_i is the ith sample of the original signal and $\hat{x}_i$ is the ith sample of the reconstructed signal.

$$NRE = \frac{\sum_{i=0}^{N}(x_i - \hat{x}_i)^2}{\sum_{i=0}^{N} x_i^2} \tag{7}$$

By computing the NRE after filtering the original and decompressed signals to extract only their contents in a specific frequency band, the NFRE in said frequency band is obtained.

4.4. Results and Discussion

The evaluation of the compression algorithms shows that not all configurations lead to better performance. Figure 2 shows the configurations where the combination of PCA filtering with the DWT compression improved on both individual compressions. The lines drawn indicate the combination of PCA and DWT configurations directly translating into a PCA + DWT configuration. The increase in compression rate shown also exhibits an increase in the NRE, which leads to a trade-off between lower data storage and maintaining data accuracy. PCA filtering with 5 and 10 components appear in the configurations with higher compression rates. This means that the filtering is more aggressive than PCA with a higher count of principal components. Higher principal component configurations also achieved high compression when combined with DWT but with a decrease in compression rate. This leads us to conclude that the filtering in these cases hindered the DWT decomposition instead of contributing to it.

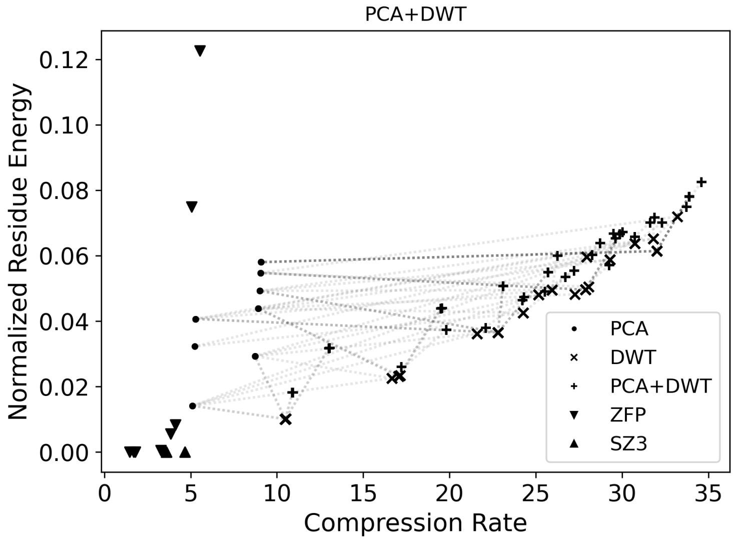

Figure 2. NRExCR relation of PCA + DWT combinations that achieved higher CR.

Figures 3 and 4 show the NFRE for the frequency ranges of 0–20 Hz and 15–65 Hz, respectively. Seismic survey data analysis puts greater emphasis on low frequency signal where the seismic response to the surveying process is more dominant. Therefore, we consider the NRE at the ranges of 0–20 Hz and 15–65 Hz. Figures 2–4 also show the comparison of the compression algorithm with the ZFP compression algorithm. ZFP has configurations for fixed-precision, fixed-rate, and fixed-tolerance, which were all individually varied. The figures show that, independent of the frequency range, ZFP can maintain a lower NRE, but at the cost of much lower compression rates than the proposed algorithm. The same can be observed for the SZ3 compressor, although it obtained negligible NRE values in the 10^{-11} range.

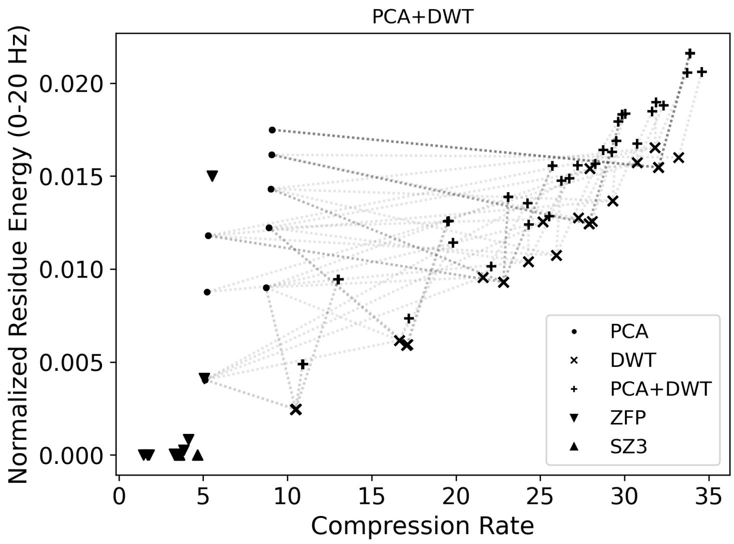

Figure 3. NFRExCR relation of PCA + DWT combinations that achieved higher CR in the 0–20 Hz range.

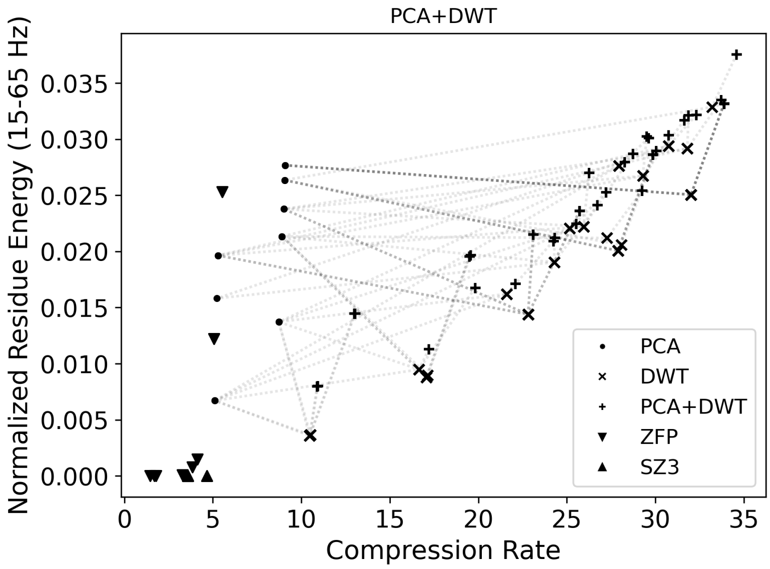

Figure 4. NFRExCR relation of PCA + DWT combinations that achieved higher CR in the 15–65 Hz range.

Table 1 shows the configurations with the best performance in each metric, and Table 2 shows the metrics achieved by each configuration. The tables are separated for readability. The last line of the table shows the configuration with the smallest NRE in the 15–65 Hz range. While the combination did not achieve an NRE lower than 2%, which can be a good discernment metric for seismic survey data, we can further analyze the compression quality by visualizing the data.

Table 1. Best performing configurations of the PCA + DWT compressor.

Name	Configuration
Configuration 1	5 Principal Components, Energy Threshold = 10%, Quantization = 8%
Configuration 2	5 Principal Components, Energy Threshold = 0%, Quantization = 10%
Configuration 3	10 Principal Components, Energy Threshold = 10%, Quantization = 6%
Configuration 4	10 Principal Components, Energy Threshold = 0%, Quantization = 6%

Table 2. Best performing configurations of the PCA + DWT compressor results.

Configuration	Compression Rate	NRE	NRE (0–20 Hz)	NRE (15–65 Hz)	NRE (60–105 Hz)	NRE (100–250 Hz)
Configuration 1	**34.5865**	**8.25e-2**	2.06e-2	**3.75e-2**	**2.19e-2**	1.02e-2
Configuration 2	33.8814	7.80e-2	**2.16e-2**	3.31e-2	1.84e-2	**1.22e-2**
Configuration 3	31.8692	7.16e-2	1.89e-2	3.21e-2	1.86e-2	8.4e-3
Configuration 4	<u>23.1024</u>	<u>5.06e-2</u>	<u>1.38e-2</u>	<u>2.15e-2</u>	<u>1.18e-2</u>	<u>8.2e-3</u>

Figures 5–7 help us visualize the effects of decompression. The rows show the original dataset, the decompressed dataset, and the difference between the original and decompressed; the first column shows the traces in the time domain and the second column shows the traces in the frequency domain. Figure 5 shows the results for the configuration with the highest compression rate for the compression algorithm, Figure 6 shows the result for the DWT configuration with the highest compression rate, and Figure 7 shows the results for a configuration of the algorithm that has a more balanced compression rate and NRE. In Figures 5 and 6, we can see that the compression was very aggressive, in the sense that NRE was relatively high (6 to 8 % NRE). This causes visual artifacts to disappear in the decompressed data that can clearly be seen in the original data. As a result, the difference shows a visual pattern that is similar to the original data, which could mean that some

relevant data have been lost. However, in the frequency domain, it is harder to identify a pattern, and the difference has an appearance closer to noise.

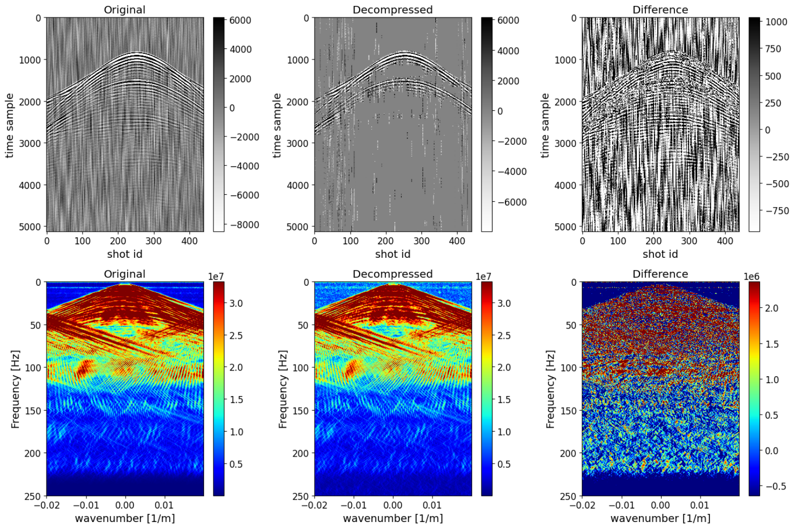

Figure 5. Reconstruction difference of PCA + DWT + LZW + quantization, five principal components energy threshold = 10% quantization = 8%.

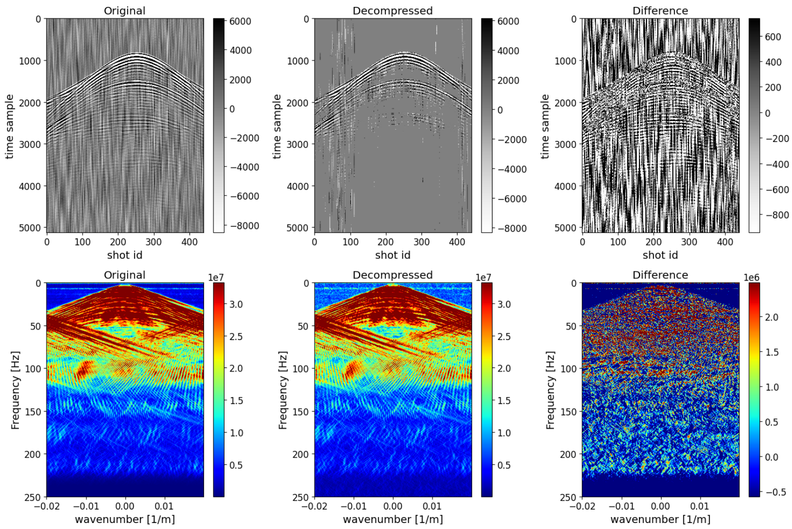

Figure 6. DWT + LZW + quantization, energy threshold = 10% quantization = 8%.

We can compare further with another configuration of the PCA and DWT compression algorithm shown in Figure 7. The 40 principal components energy threshold = 0% quantization = 2% configuration achieved 10.07:1 compression with an NRE of 0.016. This compression is much less aggressive and appears much closer to the original data in the decompressed plot. Also important to note, the difference plot shows no visual artifacts from the original dataset. We can see in the relative difference that the noise presents a random spectral behavior that does not reproduce the patterns in the original signals. The randomness means that the noise is uncorrelated from the input signal. Therefore, the compressed signals are conserving the signal of interest.

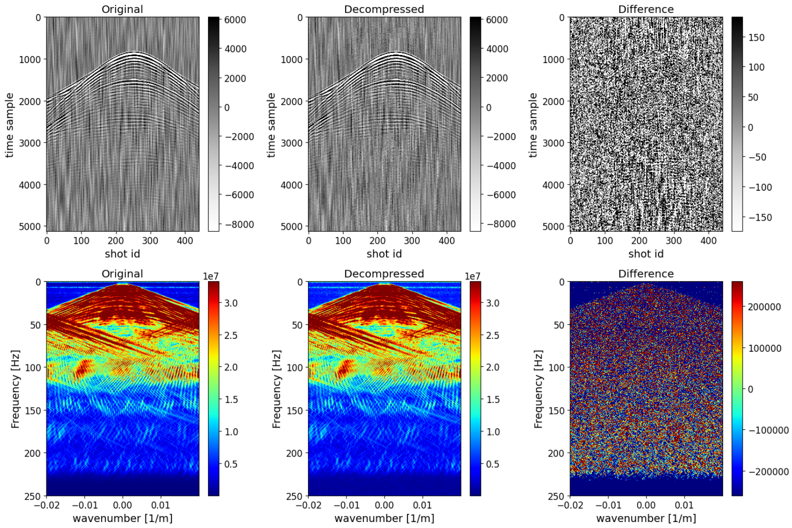

Figure 7. Reconstruction difference of PCA + DWT + LZW + quantization, 40 principal components energy threshold = 0% quantization = 2%.

5. Conclusions

This work presented a compression algorithm combining principal component analysis (PCA) and discrete wavelet transform (DWT) for improved data compression. The main algorithms are supported by thresholding, quantization, and entropy encoding strategies that prune less representative data from the dataset. We discussed the fundamentals of data compression based on the involved algorithms and the proposed algorithm combining them. The proposed method was then evaluated on a dataset of seismic data collected during an oil and gas survey leading to compression ratios of up to 31:1 while preserving relevant characteristics of data in frequencies of interest in the ranges of 0–20 Hz, 15–65 Hz, and 60–105 Hz, even though the compression leads to high reconstruction errors in total.

Unfortunately, not all configurations of the algorithm produce better compression ratios than their individual counterparts, meaning that the algorithm must be tuned for the specific dataset being compressed. However, this also means that the user can adjust how much reconstruction error the compression can tolerate. In essence, the compression strategy outlined in this study streamlines the management of the extensive data volumes generated by seismic reservoir monitoring surveys carried out by oil and gas companies. However, further work is required to demonstrate that the information loss from aggressive compression does not significantly interfere with downstream data processing to extract pertinent insights from the seismic data acquired during such surveys.

Author Contributions: M.M.d.L., conceptualization, writing—original draft preparation, review, and editing; J.L.R., writing—review and editing; M.W., writing—review and editing; A.A.F., supervision, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially funded by Petróleo Brasileiro S.A. under grant agreement no. 2022/0082 and the Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC).

Data Availability Statement: The data used in this paper are part of the Jubarte Permanent Reservoir Monitoring available at [26].

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Ikelle, L.T.; Amundsen, L. *Introduction to Petroleum Seismology*; Society of Exploration Geophysicists: Tulsa, OK, USA, 2018.
2. Lu, B.; Li, Y.; Wu, X.; Yang, Z. A review of recent advances in wind turbine condition monitoring and fault diagnosis. In Proceedings of the 2009 IEEE Power Electronics and Machines in Wind Applications, Lincoln, NE, USA, 24–26 June 2009; pp. 1–7.
3. Javed, K.; Gouriveau, R.; Zerhouni, N.; Nectoux, P. Enabling Health Monitoring Approach Based on Vibration Data for Accurate Prognostics. *IEEE Trans. Ind. Electron.* **2015**, *62*, 647–656. [CrossRef]
4. Payani, A.; Abdi, A.; Tian, X.; Fekri, F.; Mohandes, M. Advances in seismic data compression via learning from data: Compression for seismic data acquisition. *IEEE Signal Process. Mag.* **2018**, *35*, 51–61. [CrossRef]
5. Mohammadpoor, M.; Torabi, F. Big Data analytics in oil and gas industry: An emerging trend. *Petroleum* **2020**, *6*, 321–328. [CrossRef]
6. Fajardo, C.; Reyes Torres, O.; Ramirez, A. Seismic Data Compression Using 2D Lifting-Wavelet Algorithms. *Ingeniería Cienc.* **2015**, *11*, 221–238. [CrossRef]
7. Nuha, H.H.; Liu, B.; Mohandes, M.; Deriche, M. Seismic Data Compression Using Signal Alignment and PCA. In Proceedings of the 2017 9th IEEE-GCC Conference and Exhibition (GCCCE), Manama, Bahrain, 8–11 May 2017; pp. 1–6.
8. Wallace, G.K. The JPEG still picture compression standard. *IEEE Trans. Consum. Electron.* **1992**, *38*, xviii–xxxiv. [CrossRef]
9. Marcellin, M.W.; Gormish, M.J.; Bilgin, A.; Boliek, M.P. An overview of JPEG-2000. In Proceedings of the DCC 2000 IEEE, Data Compression Conference, Snowbird, UT, USA, 28–30 March 2000; pp. 523–541.
10. Sze, V.; Budagavi, M.; Sullivan, G.J. High efficiency video coding (HEVC). In *Integrated Circuit and Systems, Algorithms and Architectures*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 39, p. 40.
11. Nuha, H.; Liu, B.; Mohandes, M.; Balghonaim, A.; Fekri, F. Seismic data modeling and compression using particle swarm optimization. *Arab. J. Geosci.* **2021**, *14*, 2542. [CrossRef]
12. Radosavljević, M.; Xiong, Z.; Lu, L.; Hohl, D.; Vukobratović, D. High Bit-Depth Seismic Data Compression: A Novel Codec Under the Framework of HEVC. *IEEE Access* **2020**, *8*, 114443–114459. [CrossRef]
13. Bui-Tran, V.; Nguyen, T.; Tran, T.; Duval, L. GenLOT optimization techniques for seismic data compression. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Los Alamitos, CA, USA, 5–9 June 2000; Volume 4, pp. 2111–2114.
14. Duval, L.; Nagai, T. Seismic data compression using GULLOTS. In Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221), Salt Lake City, UT, USA, 7–11 May 2001; Volume 3, pp. 1765–1768.
15. Gan, S.; Wang, S.; Chen, Y.; Chen, X.; Huang, W.; Chen, H. Compressive sensing for seismic data reconstruction via fast projection onto convex sets based on seislet transform. *J. Appl. Geophys.* **2016**, *130*, 194–208. [CrossRef]
16. Geng, Y.; Wu, R.; Gao, J. Dreamlet transform applied to seismic data compression and its effects on migration. In *SEG Technical Program Expanded Abstracts 2009*; SEG: Tulsa, OK, USA, 2009; pp. 3640–3644.
17. Liu, W.; Cao, S.; Chen, Y.; Zu, S. An effective approach to attenuate random noise based on compressive sensing and curvelet transform. *J. Geophys. Eng.* **2016**, *13*, 135–145. [CrossRef]
18. Liu, B.; Mohandes, M.; Nuha, H.; Deriche, M.; Fekri, F. A Distributed Principal Component Analysis Compression for Smart Seismic Acquisition Networks. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3020–3029. [CrossRef]
19. Ranjan, R.; Kumar, P. An Improved Image Compression Algorithm Using 2D DWT and PCA with Canonical Huffman Encoding. *Entropy* **2023**, *25*, 1382. [CrossRef] [PubMed]
20. Hassan, H.J.; Wali, M.K.; Shujaa, M.I. Hybrid of DCT And PCA For Image Size Compression. In Proceedings of the 2023 IEEE International Conference on Computer and Applications (ICCA), Yangon, Myanmar, 27–28 February 2023; pp. 1–8.
21. Jolliffe, I. *Principal Component Analysis*; Springer Series in Statistics; Springer: Berlin/Heidelberg, Germany, 2002.
22. Bengio, Y.; Courville, A.; Vincent, P. Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [CrossRef] [PubMed]
23. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
24. Roelofs, G. PNG Lossless Image Compression. In *Lossless Compression Handbook*; Academic Press: Cambridge, MA, USA, 2002; pp. 371–390.
25. Deutsch, P. *DEFLATE Compressed Data Format Specification Version 1.3*; Technical Report; RFC: Marina del Rey, CA, USA, 1996.
26. Thedy, E.A.; Filho, R.; Johann, W.L.; Seth, S.N.; Souza, S.C.; Murray, P.E. Jubarte Permanent Reservoir Monitoring—Installation and First Results. In Proceedings of the 13th International Congress of the Brazilian Geophysical Society, EXPOGEF, Rio de Janeiro, Brazil, 26–29 August 2013; Society of Exploration Geophysicists and Brazilian Geophysical Society: Houston, TX, USA, 2013; Volume 8.
27. Lindstrom, P. Fixed-Rate Compressed Floating-Point Arrays. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 2674–2683. [CrossRef] [PubMed]

28. Diffenderfer, J.; Fox, A.L.; Hittinger, J.A.; Sanders, G.; Lindstrom, P.G. Error Analysis of ZFP Compression for Floating-Point Data. *SIAM J. Sci. Comput.* **2019**, *41*, A1867–A1898. [CrossRef]
29. Liang, X.; Zhao, K.; Di, S.; Li, S.; Underwood, R.; Gok, A.M.; Tian, J.; Deng, J.; Calhoun, J.C.; Tao, D.; et al. SZ3: A Modular Framework for Composing Prediction-Based Error-Bounded Lossy Compressors. *IEEE Trans. Big Data* **2023**, *9*, 485–498. [CrossRef]
30. Zhao, K.; Di, S.; Dmitriev, M.; Tonellot, T.L.D.; Chen, Z.; Cappello, F. Optimizing Error-Bounded Lossy Compression for Scientific Data by Dynamic Spline Interpolation. In Proceedings of the 2021 IEEE 37th International Conference on Data Engineering (ICDE), Chania, Greece, 19–22 April 2021; pp. 1643–1654. [CrossRef]
31. Liang, X.; Di, S.; Tao, D.; Li, S.; Li, S.; Guo, H.; Chen, Z.; Cappello, F. Error-Controlled Lossy Compression Optimized for High Compression Ratios of Scientific Datasets. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 438–447. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.