Signal Processing 141 (2017) 309–321

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Simultaneous denoising and interpolation of 2D seismic data using data-driven non-negative dictionary learning

Mohammad Amir Nazari Siahsar[a,*], Saman Gholtashi[b], Vahid Abolghasemi[c], Yangkang Chen[d,*]

[a] Young Researchers and Elite Club, Mashhad Branch, Islamic Azad University, Mashhad, Iran
[b] School of Mining, Petroleum and Geophysics Engineering, Shahrood University of Technology, Shahrood, Iran
[c] School of Electrical and Robotic Engineering, Shahrood University of Technology, Shahrood, Iran
[d] Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States of America

ARTICLE INFO

Article history:
Received 26 January 2017
Revised 26 May 2017
Accepted 13 June 2017
Available online 21 June 2017

Keywords:
2D seismic data
Denoising
Interpolation
Sparse coding
Dictionary learning
Versatile non-negative matrix factorization

ABSTRACT

As a major concern, the existence of unwanted energy and missing traces in seismic data acquisition can degrade interpretation of such data after processing. Instead of analytical dictionaries, data-driven dictionary learning (DDL) methods as a flexible framework for sparse representation, are dedicated to the problem of denoising and interpolation. Due to their meaningful geometric repetitive structures, seismic data are intrinsically low-rank in the time-space domain. On the other hand, noise and missing traces increase the rank of the noisy data. Therefore, the clean data, unlike noise and missing traces, can be modeled as a linear combination of a few elements from a learned dictionary. In this paper, a parts-based 2D DDL scheme is introduced and evaluated for simultaneous denoising and interpolation of seismic data. A special case of versatile non-negative matrix factorization (VNMF) is used to learn a dictionary. In VNMF, smoothness constraint can improve interpolation, and sparse coding helps improving denoising. The proposed method is tested on synthetic and real-field seismic data for simultaneous denoising and interpolation. Through experimental results, the proposed method is determined to be an effective and robust tool that preserves significant components of the signal. Comparison with four state-of-the-art methods further verifies its superior performance.

1. Introduction

Remote data gathering of the subsurface earth's physical properties in numerical measurements is one of the geoinformation that needs to be processed well. Some goals of processing such data are to monitor earthquakes and tsunamis and manage natural resources such as energy, fresh water and etc. In this context, geophysical methods are useful to keep away inessential digging of earth layers. These methods discover or deduce the existence and location of economically useful geological reservoirs, such as fossil fuels and hydrocarbons accumulations or determination of the earth's crust and core structure. Among geophysical techniques, seismic reflection data acquisition is a well-known technique to map the subsurface distribution of stratigraphy and its structure that are images of the earth layers. Due to the existence of unwanted coherent and incoherent energy (noise) [1], missing traces [2], and the economic and physical limitations in seismic data acquisition [3], researchers are motivated to introduce and investigate reconstruction and interpolation methods to achieve reliable high-quality seismic data. These methods are crucial before subsequent processing steps such as seismic migration and inversion [4,5].

Hitherto, several techniques and algorithms have been proposed over decades for noise attenuation and trace interpolation, either simultaneously or separately [1,6–8]. Most noise suppression methods are applied in a transform domain (with orthogonal fixed dictionary) where clean data and noise are more distinguishable; the sparser representation results in the better performance [9,10]. We use sparse representation to refer to the domain in which the data can be represented by only a few coefficients. The orthogonal dictionary is composed of basis function of the sparsifying transform. These techniques were applied in time-scales transforms such as wavelet, and curvelet [11–16], or time-frequency transforms [17–21].

There are many cases where the seismic data are incomplete and some traces are missing. The process of recovering the missing

* Corresponding authors.
E-mail addresses: a.nazari.s@gmail.com, a.nazari.s@mshdiau.ac.ir (M.A. Nazari Siahsar), saman.gh.88@gmail.com (S. Gholtashi), vahidabolghasemi@gmail.com (V. Abolghasemi), chenyk2016@gmail.com (Y. Chen).

http://dx.doi.org/10.1016/j.sigpro.2017.06.017
0165-1684/

samples are called interpolation. Sparse transforms are also used to interpolate missing traces in seismic records [3,22,23]. In [23], L1-norm minimization problem with sparsity constraint is successfully solved to restore seismic data. The idea of these methods is that the missing traces result in non-sparse artifacts in transform domain where the data are represented sparsely.

In addition to the mentioned methods for denoising and interpolation, there are algorithms utilizing the inherent low-rank property of the seismic data such as [3,24,25]. Based on this premise, several techniques are proposed for reconstruction of the seismic data in different domains such as multichannel singular spectrum analysis (MSSA) [1,26], damped MSSA [27–29] and also randomized MSSA [30]. Classical f-x SSA operates in the frequency-space domain by embedding spatial data at a frequency slice into a Hankel matrix. Noise and missing traces generally increase the rank of this meaningful matrix extracted from the data [31]. Therefore, the Hankel matrix of the corrupted observations, which should be low-rank in the case of the clean data, is restored using a truncated singular value decomposition (SVD) technique. They used an iterative thresholding algorithm similar to the method projected onto convex set (POCS) [32,33] for simultaneous denoising and interpolation. Another method in this category denoises the data using low-rank and sparse decomposition in the transform domain where the data are approximately low-rank and the noise are sparse [34]. Nazari Siahsar et al. [35] introduced a low-rank and sparse decomposition scheme for seismic data denoising in the transform domain. This method starts by transforming the seismic trace into the sparse domain using synchrosqueezing transform [36]. Then, the method is followed by low-rank and sparse array decomposition based on the mixed norm optimization [37–39]. Matrix completion is another technique that is introduced for 3D seismic data interpolation by a rank-reduction formulation [3]. This patch-mapping method solves a nuclear-norm minimization problem to restore the data. Furthermore, several fixed-basis sparsity-promoting transforms have already been proposed in the literature for processing seismic data including the Fourier transform [40,41], the Radon transform [42,43], the high-order sparse radon transform [44], the wavelet and curvelet transform [45,46], the shearlet transform [47,48], the seislet transform [49–52] and the EMD-seislet transform [53].

Recently, machine learning techniques are used to find (learn) a suitable transform domain which the data can be represented sparsely [54,55]. One such technique for this purpose is K-SVD [56]. Methods, in this case, assume that a clean data can be represented as a sparse linear combination of the atoms in an overcomplete dictionary instead of a predefined analytical one (e.g., wavelets). Therefore, the noise of the corrupted data can be reduced by approximating the clean data using the mentioned atoms [57–59]. Simultaneous denoising and interpolation using learning based approach was first proposed for image restoration by [60,61]. Unfortunately, it is not easy to find suitable fixed dictionaries that enable us to model the complex local structures of the data, hence, data-driven dictionary learning methods are introduced that allow us to capture the morphology of the noisy data itself [62,63]. Generally, there are two approaches to find dictionaries: (i) using a clean data as training samples to reconstruct another test image [64]; (ii) using noisy data itself as training samples and adapted the method to reconstruct the data [65]. In the former case, it is difficult to find a proper clean data, which help to reconstruction, so, using noisy data itself (i.e. second approach), can be a better solution in dictionary learning.

Due to their meaningful geometric repetitive structures, such as events and dips, seismic data are intrinsically low-rank in the time-space domain [3,35]. On the other hand, noise and missing traces increase the rank of the noisy data [1,3,31,35]. Therefore, assuming a learned dictionary, the clean data, unlike noise and missing traces, can be modeled as a linear combination of a few elements from the dictionary. The learning-based approaches also used to update and infer the dictionaries for seismic data processing [66–68]. Cai et al. [69] proposed a data-driven tight frame (DDTF) based dictionary learning algorithm for image restoration that is faster than the traditional K-SVD based approach, which was then applied to seismic data restoration by Liang et al. [66]. Zhu et al. [70] and Chen et al. [4] combined the dictionary learning based sparse transform with the fixed-basis transform, which is called double-sparsity dictionary, to better adapt to seismic data. In double sparsity [71], the base dictionary is not sparse, such as DCT or wavelets, etc., only the adaptive layer is sparse. Recently, Zhu et al. [72] introduced a joint seismic data denoising and interpolation using a masking strategy in the sparse representation of the dictionary. Chen [67] applied a new dictionary learning algorithm based on the sequential generalized K-means (SGK) to seismic noise attenuation that is faster than its alternative: the K-SVD algorithm.

In [73], Bekouche and Ma introduced a patch-based data-driven dictionary learning algorithm with sparse coding to attenuate random noise in 2D seismic data. In order to decrease the time cost of their algorithm, they picked a limited number of patches randomly from the whole data to learn a dictionary. Therefore, as a drawback, it can be considered that the method did not use whole data information to learn a dictionary and therefore the learned dictionary does not contain all information about the data. In this work, we will increase the quality of the seismic data to tackle the mentioned drawback by using whole data information to learn a dictionary.

In this paper, we introduce a simultaneous denoising and interpolation algorithm for 2D seismic data which relies on a parts-based data-driven non-negative dictionary learning (DNDL) algorithm with sparsity and smoothness constraints. Our proposed method takes the signal characteristics into account and does not require prior knowledge. The atoms of the dictionary in this algorithm are generated based on a special case of versatile non-negative matrix factorization (VNMF) [74]. The VNMF model introduces a sparsity-promoting dictionary learning minimization problem with smoothness constraint, which combines dictionary learning and sparse coding [75]. The sparsity-inducing norm minimization can help us to improve the noise reduction ability of the algorithm. Furthermore, there is a relation between non-negativity and sparsity constraints in which non-negativity can induce sparsity [76,77]. Also, by applying the non-negativity constraints to dictionary and coefficients matrices, VNMF can learn the data part by part [77]. In fact, VNMF can approximate the data using only an additive combination of multiple basis (with no subtractions). Since, in dictionary learning algorithms, one seeks to find atoms of the data, the parts-based feature can be helpful to obtain the parts (atoms) of the data. In addition to scaling each atom in the dictionary, the smoothness constraint can be useful for interpolation by smoothing the data. It has been proven that using both L1 and L2 norms constraints on atoms can remove correlated variables, simultaneously [74].

In general, seven merits of using this method are described as follows (our contributions in this paper are marked by $*$):

(1) Working with data in time-space domain.
(2*) VNMF combines dictionary learning and sparse coding, that finds the basis of low-dimensional space from noisy data itself (data-driven feature).
(3*) Utilizing L1 and L2 norms constraints on dictionary matrix to (i) achieve sparser representation, and (ii) obtain more efficient interpolation.
(4*) Using non-negativity constraint to induce sparsity on transform domain in seismic data and reduce the solution space.