Journal of Applied Geophysics 133 (2016) 70–81

Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

Multicomponent seismic noise attenuation with multivariate order statistic filters

Chao Wang [a,*], Yun Wang [b], Xiaokai Wang [c], Chao Xun [a]

[a] State Key Laboratory of Ore Deposit Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang 550081, China
[b] China University of Geosciences, Beijing 100083, China
[c] Department of Computational Geophysical Science, Xi'an Jiaotong University, Xi'an 710049, China

ARTICLE INFO

Article history:
Received 14 May 2015
Received in revised form 29 June 2016
Accepted 26 July 2016
Available online 2 August 2016

Keywords:
Multicomponent seismic
Noise attenuation
Vector
Multivariate order statistics

ABSTRACT

The vector relationship between multicomponent seismic data is highly important for multicomponent processing and interpretation, but this vector relationship could be damaged when each component is processed individually. To overcome the drawback of standard component-by-component filtering, multivariate order statistic filters are introduced and extended to attenuate the noise of multicomponent seismic data by treating such dataset as a vector wavefield rather than a set of scalar fields. According to the characteristics of seismic signals, we implement this type of multivariate filtering along local events. First, the optimal local events are recognized according to the similarity between the vector signals which are windowed from neighbouring seismic traces with a sliding time window along each trial trajectory. An efficient strategy is used to reduce the computational cost of similarity measurement for vector signals. Next, one vector sample each from the neighbouring traces are extracted along the optimal local event as the input data for a multivariate filter. Different multivariate filters are optimal for different noise. The multichannel modified trimmed mean (MTM) filter, as one of the multivariate order statistic filters, is applied to synthetic and field multicomponent seismic data to test its performance for attenuating white Gaussian noise. The results indicate that the multichannel MTM filter can attenuate noise while preserving the relative amplitude information of multicomponent seismic data more effectively than a single-channel filter.

1. Introduction

In multicomponent seismic data, the conventional vertical component and two horizontal components of the incident wavefield are simultaneously recorded. There are large quantities of information that can be extracted from multicomponent seismic data, but this requires relative amplitudes in all components to remain unchanged through prior processing. To date, there are still many impediments to achieve this objective, such as the faultiness in statics, noise attenuation, and P- and S-waves separation. To obtain an ideal final result, each of these processing steps must be improved with a more effective technology. Thus, noise attenuation as a pre-processing step is highly critical in multicomponent processing to improve the quality of subsequent processes.

Many practical denoising methods have been developed for seismic data. The most-used methods include frequency filtering, time-frequency filtering based on the wavelet transform (Gao et al., 2006) or the S-transform (Pinnegar and Eaton, 2003), median-based filtering (Liu et al., 2006), f-x deconvolution (Marfurt, 2006), matched filtering (Eisner, 2008; Song et al., 2010), and types of f-k filtering. Although these methods can be used for multicomponent seismic data by treating each component as an independent scalar field, the vector relationship of the multicomponent seismic wavefield could be damaged.

Because each component of the multicomponent seismic data reflects a certain property of a common subsurface objective, there are coherent relationships between these components. In certain cases, the linear dependence existing between components can be separated by simply rotating components to a reference frame in which the components are linearly independent (Li and Yuan, 1999). However, it is not always possible to separate wavefields in this way, such as when the medium is anisotropic (Stanton and Sacchi, 2013). Therefore, it would be beneficial to treat all components as a vector wavefield rather than a set of independent scalar fields. At present, only a small number of approaches which treat the multicomponent seismic data as a vector wavefield are proposed. Naghizadeh and Sacchi (2012) proposed a 3-component vector autoregressive (VAR) model and attenuated the random noise of 3-component seismic data in the frequency-space (f-x) domain. The coherencies between the three components of multicomponent seismic data can be effectively determined by VAR modelling. However, certain effective signals are removed from their real data example. Rodriguez et al. (2012) assumed that

* Corresponding author.
E-mail address: srmn28@163.com (C. Wang).

http://dx.doi.org/10.1016/j.jappgeo.2016.07.023
0926-9851/

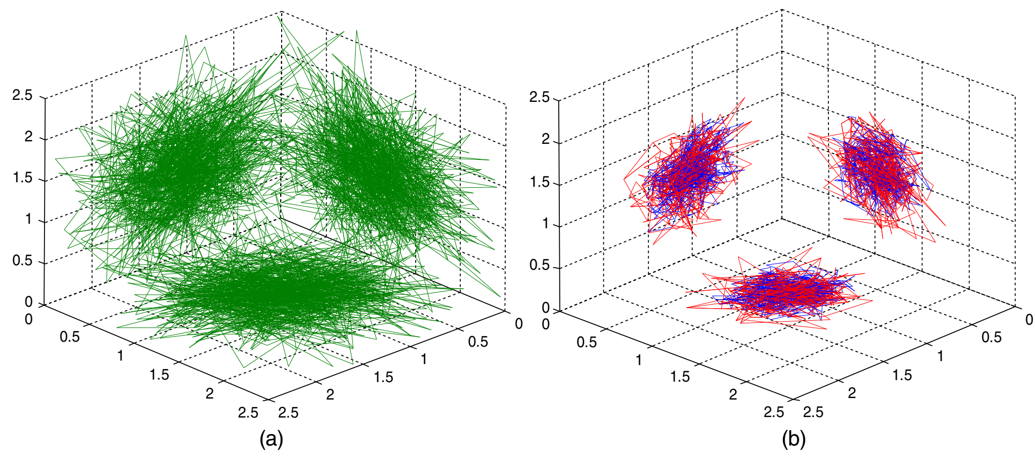

Fig. 1. The projections of the 3D hodogram of a 3-channel signal. (a) The contaminated constant signal. (b) The filtering results of multichannel MTM filter (blue line) and VMF (red line).

the representation coefficients of the 3 components, which are detected at the same time and position, have the same distribution pattern in the wavelet transform domain, and therefore they proposed a denoising method based on a complex wavelet transform constrained by group sparsity to simultaneously attenuate the random noise of 3-component microseismic data. Stanton and Sacchi (2013) construct a quaternion with the real and imaginary parts of the x- and y-components in the frequency domain; next, they propose a method based on quaternion Fourier transform to reconstruct the x- and y-components simultaneously.

Several multivariate order statistic filters have been demonstrated to be effective for colour image processing (Nikolaidis and Pitas, 1996), in which the colour images are described as vector images. This class of nonlinear filters has many different forms in terms of their approaches to ordering multivariate data, such as the multichannel α-trimmed mean filter, the multichannel MTM filter (Pitas and Tsakalides, 1991), the vector median filter (VMF) (Lucat and Siohan, 1997; Astola et al., 1990; Xu et al., 2014), and the vector directional filter (VDF) (Trahanias et al., 1996). VMF is also applied to seismic data: Liu (2013) used VMF to attenuate noise in an azimuth vector field derived from a migrated seismic image, Huo et al. (2012) used VMF to separate blended seismic data recorded with simultaneous source acquisition technology However, only single-component seismic data was considered in both of their studies. Theory and experiments indicate that VMF has good performance for random noise with long-tailed distributions, whereas it performs poorly for random noise with short-tailed distributions. Some experiments also demonstrated that a multichannel MTM filter is more effective than a VMF filter for Gaussian noise removal (Pitas and Venetsanopoulos, 1992).

A seismic wavefield is essentially a vector field and is usually recorded with a multi-component geophone array. It is natural to develop a multivariate order statistic filter for multicomponent seismic data. However, there are still many special problems that must be studied to ensure that good results can be obtained. In this paper, we will address these issues according to the characters of multicomponent seismic data.

2. Methods

2.1. Multivariate order statistic filters

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be N random vector samples in a p-dimensional space. In order to facilitate the implementation of filtering, N is usually an odd number. Each sample, $\mathbf{x}_i$ $(i = 1, \ldots, N)$ is denoted as $[x_{1i}, x_{2i}, \ldots, x_{pi}]^T$, where the superscript T means transpose. There is not a natural basis for ordering multivariate data. Barnett (1976) discussed the problem of multivariate data ordering. He proposed several approaches to order multivariate data. In this study, we only introduce marginal ordering and reduced ordering, which will be used later in this paper.

In marginal ordering, each component of the p-dimensional vector sample is ordered independently:

$$x_{j(1)} \le x_{j(2)} \le \ldots \le x_{j(N)}, \quad j = 1, 2, \ldots, p, \tag{1}$$

where, $x_{1(1)}, x_{2(1)}, \ldots, x_{p(1)}$ are the minimal elements in each channel, and $x_{1(N)}, x_{2(N)}, \ldots, x_{p(N)}$ are the maximal elements in each channel. The ith marginal order statistic is the vector $\boldsymbol{x}_{(i)} = [x_{1(i)}, x_{2(i)}, \ldots, x_{p(i)}]^T$. The second subscript in parentheses is to distinguish it from the subscript before ordering. The cumulative density function and the probability density function of the marginal order statistics are discussed in Galambos (1975); Pitas (1990), and Pitas and Tsakalides (1991).

For reduced ordering, the definition of generalized distance needs to be introduced first:

$$d_i = (\boldsymbol{x}_i - \boldsymbol{a})^T \Gamma^{-1} (\boldsymbol{x}_i - \boldsymbol{a}), \tag{2}$$

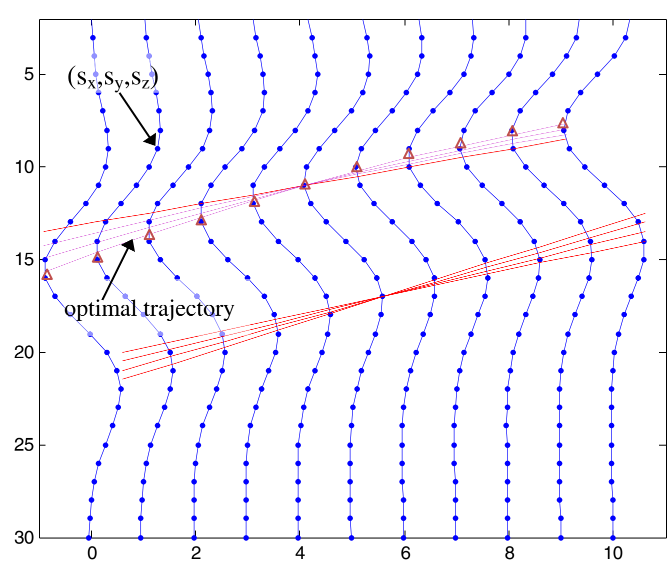

Fig. 2. Illustration of the straight trajectories (red) and curved trajectories (purple). Each point represents a vector. The vector points chosen as the input data of a filter are marked with triangle symbols.