Coherence index and curvelet transformation for denoising geophysical data

Hassan Dashtian and Muhammad Sahimi*
Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, California 90089-1211, USA
(Received 26 January 2014; revised manuscript received 17 July 2014; published 17 October 2014)

Geophysical data contain stochastic noise that may mask their useful content. For example, ground roll (GR) is a coherent noise that is present in seismic data. Thus, such data are usually a mixture of useful information and useless coherent noise. The latter masks the relevant geologic information that seismic records contain, and its removal has always been a problem of fundamental importance. We propose a denoising method based on the curvelet transformation (CT), a multiscale transformation with strong directional character that provides an optimal representation of objects that have discontinuities along their edges. An algorithm is presented for processing and denoising of geophysical data. As an example, we apply the method to seismic images that are contaminated with the GR noise. First, the coherence index (CI), which represents a measure of the amount of energy contained in the most coherent modes of Karhunen-Lòeve transform for any given segment of the data, is computed. The contaminated region of the data is then identified as the maximum region of the CI. After demarcating the contaminated segment, the CT is used to eliminate the noise. The method removes the noise with negligible distortion of the data's noncontaminated region. It is also significantly more efficient computationallty than the previous methods. The use of the method is demonstrated by its application to synthetic, as well as actual, seismic data for hydrocarbon reservoirs.

DOI: 10.1103/PhysRevE.90.042810 PACS number(s): 89.90.+n, 05.45.Tp, 93.85.−q, 91.30.Dk

I. INTRODUCTION

An important task in geophysics is to delineate as much information as possible about the structure of the interior of rock. Since direct exploration by penetrating rock is impractical, various types of data, such as seismological, electromagnetic, and gravitational measurements, collectively referred to as geophysical data, are used. Such data provide much insight into the structure and properties of rock and large-scale porous media, such as oil, gas, and geothermal reservoirs and groundwater aquifers. Thus, a major task is making inferences from such data [1]. The method of analysis depends, however, on the volume of the data. If the number of observations or data points is small, one usually uses them to fit a number of parameters in the analytic solutions of the equations of classical physics, such as the wave equation. With the advent of new and precise instruments and rapidly increasing computing power, however, it has become common to have a massive amount of geophysical data that may contain far more information about large-scale porous media than can possibly be modeled by the analytic solutions. For example, a typical reflection seismic marine survey ship can collect about a trillion bits of information over a relatively short time. In addition, the existing analytic solutions have been derived based on the assumption that, at a large enough length scale, any porous medium of interest may be represented by a continuum to which the averaged equations of motions are applicable. In reality, that is usually not the case, as rock and large-scale porous media are heterogeneous over a wide spectrum of distinct length scales [2], and modeling of any phenomenon in them through averaged continuum mechanical equations may be in error.

Similar to many other types of data that contain a stochastic component [3], accurate processing and analysis of geophysical data is a complex task and fraught with difficulties. The difficulties are due to the uncertainties associated with the measurements, as well as various types of what is usually referred to as noise. Consider, for example, seismic data, the focus of the present paper, which play a fundamental role in obtaining information about and insight into the structure of rock [4] and its content. The sources of noise in seismic data are usually divided into two categories. The first one consists of noise due to the experimental, or measurement, errors. Such errors comprise any unexpected perturbation of the environment in which the data are recorded, such as, for example, a geophone that may have malfunctioned, wind, cable vibrations, etc. This type of noise imparts more coherent energy into the data and can be misinterpreted as true data.

The second class of noise is due to modeling of uncertainties. Consider, for example, seismic data, which are often contaminated by random noises caused mainly by the receiver's equipments [5,6]. In seismic wave propagation, the wave fronts [7] are formed by various components, such as direct, reflected, converted, and transmitted waves, which endow the wave front with abundant information about important geologic features. Modeling of uncertainties associated with seismic data is necessary because the physical description and parametrization of rock is incomplete. The incomplete description is caused by the inherent complexity of wave propagation in rock, a highly heterogeneous medium. As seismic data are very complex, they are usually separated into distinct propagation modes that are somehow easier to understand and use. The downside of such a decomposition of the data is that, most of the actual data may consequently be obliterated. For example, the ground roll, a coherent noise that is present in land-based seismic data, is attenuated and unravels the seismic reflections.

*moe@usc.edu

Due to such difficulties, data denoising is an important issue in the processing of geophysical data. There are many denoising methods, most of which serve the same goal, namely, eliminating random noise and useless signal components, while protecting and/or recovering the original information that may be lost by the denoising process, such as the signal's amplitudes and discontinuous contours or edges. In particular, the goal in a high-quality processing of seismic data is having a high signal-to-noise ratio (SNR), high resolution, and high fidelity in which improving the SNR or the peak SNR (PSNR) is the chief task. Collectively, the three desired features are called *3-high guidelines*.

Depending on its variant features, noise in seismic data is generally divided into three main classes. The first, based on the features of the seismic section plane, consists of regular and irregular noise, which are also called coherent noise and random noise, respectively. The second class is based on the spreading mechanism and includes noise in surface, refracted, lateral, and tube waves, as well as multiple waves. Finally, one also has low-frequency, high-frequency, and 50 Hz industrial noise. Among these, elimination of the ground roll (GR) [8], a low-frequency coherent noise, is a fundamental issue.

The GR is the signature of a surface wave, i.e., the Rayleigh dispersive wave with low frequency and low phase and group velocities [9]. As a component of surface wave, the GR does not contain information on deeper subsurface structures associated with, for example, oil reservoirs. An important characteristic of the GR is that its amplitude may be stronger than the waves carrying the relevant information that is due to the reflections in geologic strata.

Over the last several decades, many methods have been proposed for denoising geophysical data, and in particular seismic data. One well-known method is based on the Karhunen-Loève (KL) transformation (Principal Orthogonal Decomposition) [10], which is [11] an optimal orthogonal transformation based on statistical characteristics of the data and has been used in pattern recognition, feature optimization, and denoising of time series [12–14]. Its application to denoising of seismic data and eliminating the GR has been described by Liu [8] and Montagne and Vasconselos [15]. Another denoising method is based on the use of Fourier transform. Canales [16] proposed a scheme in two-dimensional (2D) frequency-spatial domain for suppressing random noise in seismic data. A classical method of denoising is the linear Radon transform [17] (see below) or the so-called τ-p transform [18], which has the advantage of being capable of distinguishing the primary reflection and various types of noise, including random noise.

The aforementioned methods are most accurate for the data collected over a single scale. If geophysical data represent a system at multiple scales, which is almost always the case, then one needs a multiresolution (MR) analysis method. Over the last two decades such MR methods as wavelet transform [19] have been developed and utilized for analyzing and denoising geophysical data. Gendron [20] proposed an application of Bayes' Law of Probability to denoising of seismic data and derived a "best" wavelet packet-basis strategy. Chanerley and Alexander [21], and Corso *et al.* [22] used stationary wavelet transform as an alternative denoising scheme to standard band-pass filtering. Directional or geometric wavelets for data processing have become an important research topic over the last several years. Sahimi and co-workers proposed use of wavelet transforms for phenomena associated with various fluid flow and transport in large-scale porous media [23], characterization of such porous media [24], and transport in disordered solids [25].

Generalizations of wavelet transforms have also been proposed and utilized in various applications. For example, contourlets proposed by Do and Vetterli [26] can capture directional details of smooth anisotropic contours that are intrinsic geometric features of seismic data. Po and Do [27] proposed a denoising scheme for oriented textures, taking into account the dependence on distinct scales, directions, and locations in the contourlet domain. Alternatively, the curvelet transform (CT) [28,29] is a geometric multiscale transform that allows an optimal sparse representation of objects with second-order continuous differentiable singularities. They represent edges and singularities along curves much more precisely with needle-shaped elements, which own very high directional sensitivity and are very efficient for smooth contour capturing. For denoising of geophysical data, and in particular for seismic discontinuity-preserving denoising, curvelets are superior to 2D standard wavelets [5]. Computationally, curvelets are also more efficient as they use much fewer coefficients than wavelets to represent edges or wave fronts for a given accuracy.

The first generation of the CT was based on a block ridgelet transform [28]. Ridgelets, wedgelets, and other generalizations of the WTS were studied extensively by Donoho and co-workers [28–30]. Ridgelets are not, however, true ridge functions that take on the form $f(\mathbf{F}\cdot\mathbf{x})$, with $\mathbf{F}$ being a fixed vector and $\mathbf{x}$ a variable one, both in d-dimensional space. As a result, the geometry of the ridgelets was itself unclear, leading to important limitations for the first-generation curvelets. More advanced CT was then proposed [29], defined directly via frequency partitioning without using the block ridgelet transform, which makes it more efficient for data processing. Others [5,6] used the CT to map multidimensional data onto almost orthogonal localized multidimensional prototype waveforms that varied in directional and spatial-temporal content. They also proposed [5,6] a nonparametric transform-based recovery method to exploit the compression of seismic data using CT. Fast discrete CT has also been developed [31].

Herrmann and co-workers [32–35] used the CT for image recovery and processing, as well as denoising of seismic data. Yarham *et al.* [34] used the CT to remove the GR from seismic data by a two-step procedure. First, the major component of GR was identified using a multiscale separation, directional or frequency-altering method. Next, the GR was removed through a block-coordinate relaxation method [36]. In another study, Yarham and Herrmann [35] proposed a method for removal of the GR from the data based on CT, showing that by using a Bayesian separation method and the CT more coherent noise can be removed. Neelmani *et al.* [37] discussed applications of the CT, as well as a hybrid CT-WT for removing random and coherent noise from 3D seismic data of carbonate reservoirs. Three main steps are taken when applying the CT to eliminate noise. One first decomposes the data or signal in the curvelet space. Then the coefficients in angular sectors in the curvelet space that correspond to the noise are identified and eliminated (set to zero). Finally, one reconstructs the denoised data. Olhede and Walden [38] used the maximal overlap discrete

WT and Hilbert energy spectrum to analyze nonstationary signals in time-frequency domain. They showed that their decomposition method may be used to denoise signals with Gaussian noise.

All the current variations of the CT have, however, the serious drawback that the region to be filtered out must be picked by hand, a procedure that not only is labor intensive, but also relies on the judgment of the person performing the denoising. All the aforementioned studies, as well as other reports on the removal of the GR using multiscale transforms, also apply their algorithms to the *entire seismic image* without taking into account the actual location of the noise. Not only is this inefficient computationally, in particular one must analyze actual seismic data that are typically very voluminous, it may also damage the valuable information contained in the image.

Such methods also share two main features. First, they decompose the traces of the signal using a one-dimensional spectral technique, such as Fourier, wavelet, or curvelet transform. Second, they utilize an attenuation factor in order to remove a part of the undesirable frequencies (or scales and tones) in the transformed space. But the attenuation is always done blindly and cannot guarantee that the truly useful data are preserved. The attenuation is based on the assumption that if the noise is concentrated in a given range of the frequencies (scales), one should decrease the amplitude of the coefficients that correspond to such frequencies (scales), which is more often than not unjustified. Indeed, the attenuation is carried out without complete knowledge of the data. Although some may consider this as an advantage, we believe that a basic understanding of general features of any set of seismic data is essential to their correct and efficient processing.

Thus, in this paper we propose a significant improvement of the CT method for removal of the coherent noise in which the data segment to be filtered out is selected automatically as an optimization process. To do so, we utilize the coherence index (CI), which represents a measure of the amount of energy contained in the most coherent modes for any given segment of the data. The optimal segment is then one for which the CI is maximum, thus ensuring that the CT removes the coherent noise in a most efficient and accurate way. Furthermore, introducing a quantitative criterion for selecting the optimal segment to be filtered out has the considerable advantage of yielding an automatic, unsupervised scheme for demarcating and efficiently removing the coherent noise. We focus on the demarcating the noisy region by using the K-L transform, and then applying the CT filtering algorithm to the noisy region. Thus, the new algorithm employs the advantages of the CT that have been reported by others, while imposing an intelligent selection criterion. We emphasize that although our main motivation here concerns the suppression of coherent noise in seismic data, we wish to emphasize that the proposed method is applicable to other problems and types of data in which one seeks to identify and eventually remove coherent structures embedded in a complex pattern. In particular, the proposed optimized CT filter yields an image processing technique that is highly suitable for images in which the undesired coherent features degrade the quality of the information that can be extracted from the image.

The rest of this paper is organized as follows. In Sec. II we briefly describe the CT and discuss its main properties. Section III presents the proposed optimized CT procedure for selecting the noise-contaminated region. The application to 2D synthetic and actual seismic results are presented and discussed in Sec. IV, and the computational efficiency of the method is discussed in Sec. V. The paper is summarized in Sec. VI.

II. CURVELET TRANSFORMS

A classical problem in image processing is the reconstruction of images from indirect noisy measurements, that is, recovering an object $f \in L_2(\mathcal{R}^2)$ from a data base y of the form

$$y(u) = K[f(u)] + z(u), \tag{1}$$

where K is a linear operator defined by $K[f(u)] = \int k(u,x) f(x)\, dx$, with $f(x)$ being the object of interest. K can be, for example, a Radon transform [17,39], a collection of the line integrals,

$$K[f(u)] = \mathcal{R}[f(u)] = \int_{L_{\theta,t}} f(x)\, dx, \tag{2}$$

where the line is defined by $L_{\theta,t} = x_1 \sin\theta + x_2 \cos\theta = t$. Here $z(u)$ is the noisy part that may be stochastic or deterministic.

In a manner similar to the fact that it was recognized that Fourier methods are not suitable for all purposes and, in particular, for image processing and multiscale systems and, consequently, the WTs were introduced for analyzing such systems, development of alternatives to the wavelet analysis has also attracted considerable attention. This has been motivated by many problems in which one must study interesting phenomena that occur along *curves* or *sheets*, such as edges in a 2D image. Although wavelets are suitable for dealing with images in which interesting phenomena, such as singularities, are associated with *exceptional* points, they are ill-suited for detecting, organizing, or providing a compact representation of intermediate dimensional structures. Thus, there have been vigorous research efforts for developing better adapted alternatives to the wavelets by combining ideas from geometry and traditional multiscale analysis [5,6]. A special member of this emerging family of multiscale geometric transforms is the CT, which was developed in an attempt to overcome inherent limitations of the traditional multiscale representations, such as those with the wavelets.

At a conceptual level, the CT is a multiscale pyramid with many directions and positions at each length scale, and needle-shaped elements over fine scales. The pyramid is nonstandard, however, in that the curvelets have geometric features that distinguish them from the wavelets. For example, curvelets follow a parabolic scaling relation according to which at scale 2^j each element has an envelope aligned along a "ridge" of length $2^{-j/2}$ and width 2^{-j}. Roughly speaking, the data acquisition geometry separates the curvelet expansion of an object into two pieces,

$$f = \sum_{n \in \text{good data}} \langle f, \varphi_n \rangle \varphi_n + \sum_{n \notin \text{HTD}} \langle f, \varphi_n \rangle \varphi_n, \tag{3}$$

where φ_n is called the *mother* curvelet (see below). The first term on the right side represents that part of the image or data that is not affected by the noise and can be recovered

easily, while the second part, the HTD, represents the part for which it is hard to distinguish between the true signal or data and the noise and, thus, cannot be recovered accurately. The important point is that one can probably reconstruct the recoverable part with an accuracy similar to what one would achieve if one had the complete data. It can be shown [5,6] that for some statistical models that allow for discontinuities in the object to be recovered, there are simple algorithms based on the shrinkage of the curvelet biorthogonal decompositions, which achieve optimal statistical rates of convergence. That is, there are no other estimating procedure that, in an asymptotic sense, yield a fundamentally better reconstruction. But, to realize this potential, and deploy this "technology" to a wide range of problems, one would need a fast and accurate discrete CT operating on digital data, which we now describe.

A. Continuous-time curvelet transforms

In this paper we study 2D systems with a spatial variable x, with ω being the frequency domain variable, and r and θ the polar coordinates in that frequency domain. We begin with a pair of windows $W(r)$ and $V(t)$, which we call the *radial window* and the *angular window*, respectively. Both are smooth, nonnegative, and real-valued functions, with W taking on positive real arguments and supported in $r \in (1/2,2)$ and V taking real arguments and supported in $t \in [-1,1]$. The two windows satisfy the admissibility conditions [5,6],

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, \quad \text{with} \quad r \in \left(\frac{3}{4},\frac{3}{2}\right) \tag{4}$$

and

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1, \quad \text{with} \quad t \in \left(-\frac{1}{2},\frac{1}{2}\right). \tag{5}$$

Then, for each j we introduce a frequency window U_j, defined in the Fourier domain by

$$U_j(r,\theta) = 2^{-3j/4} W(2^{-j} r) V\left(\frac{2^{\lfloor j/2 \rfloor}\theta}{2\pi}\right), \tag{6}$$

where $\lfloor \cdot \rfloor$ represents the integer part of the number. Thus, the support of U_j is a polar wedge defined by the support of W and V, the radial and angular windows, applied with scale-dependent window widths in each direction. To obtain real-valued curvelets, we work with the symmetric version of Eq. (4), namely, $U_j(r,\theta) + U_j(r,\theta+\pi)$.

We then define the waveform $\varphi_i(\mathbf{x})$ by means of its Fourier transform, $\widehat{\varphi}_i(\omega) = U_j(\boldsymbol{\omega})$, with the understanding that $U_j(\omega_1,\omega_2) = U(\boldsymbol{\omega})$ represents the window defined in the polar coordinate system. In a manner similar to the wavelets, one may think of $\varphi_i(\mathbf{x})$ as the mother curvelet in the sense that all the curvelets at scale 2^{-j} are obtained by rotation and translation of $\varphi_i(\mathbf{x})$. We then introduce an equispaced sequence of rotation angles, $\theta_l = 2\pi \times 2^{-\lfloor j/2 \rfloor}$, with $l = 0, 1, \ldots,$ such that $\theta_l = 0 < 2$. Note that the spacing between consecutive angles is scale dependent. We also introduce a sequence of translation parameters, $k = (k_1,k_2) \in Z^2$. With the given notations, the curvelets are defined as a function of $\mathbf{x} = (x_1,x_2)$ at scale 2^{-j}, orientation θ_l, and position $x_k^{(j,l)} = R_{\theta_l}^{-1}(k_1 \times 2^{-j}, k_2 \times 2^{-j/2})$ by [5,6,29]

$$\varphi_{j,l,k}(\mathbf{x}) = \varphi_j\left[R_{\theta_l}\left(\mathbf{x} - \mathbf{x}_k^{(j,l)}\right)\right], \tag{7}$$

where R_{θ_l} represents a rotation by θ radians, and $R_{\theta_l}^{-1}$ its inverse, as well as its transpose,

$$R_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad R_\theta^{-1} = R_\theta^{\mathrm{T}} = R_{-\theta}. \tag{8}$$

A curvelet coefficient is then simply the inner product of an element $f \in L^2(R^2)$ and a curvelet $\varphi_{j,l,k}$,

$$c(j,l,k) = \langle f, \varphi_{j,l,k} \rangle = \int f(\mathbf{x}) \overline{\varphi_{j,l,k}}(\mathbf{x})\, d\mathbf{x}, \tag{9}$$

where the overline represents complex conjugate. Since digital CTs operate in the frequency domain, it is useful to apply Plancherel's theorem (according to which the integral of a function's squared modulus is equal to the integral of the squared modulus of its frequency spectrum) and express the inner product as an integral over the frequency plane,

$$\begin{aligned} c(j,l,k) &:= \frac{1}{(2\pi)^2} \int \hat{f}(\boldsymbol{\omega}) \overline{\widehat{\varphi}_{j,l,k}}(\boldsymbol{\omega})\, d\boldsymbol{\omega} \\ &= \frac{1}{(2\pi)^2} \int \hat{f}(\boldsymbol{\omega}) U_j(R_{\theta l}\boldsymbol{\omega}) \exp\left(i\left\langle \mathbf{x}_k^{j,l}, \boldsymbol{\omega} \right\rangle\right) d\boldsymbol{\omega}, \end{aligned} \tag{10}$$

where the hat sign indicates the Fourier transform of the function. As in the theory of wavelets, we also have coarse-scale elements. We introduce the low-pass window W_0 obeying

$$|W_0(r)|^2 + \sum_{j=0} |W(2^{-j} r)|^2 = 1, \tag{11}$$

and for $(k_1,k_2) \in Z$ define the coarse-scale curvelets as

$$\varphi_{j_0,k}(\mathbf{x}) = \varphi_{j_0}(\mathbf{x} - 2^{-j_0}\mathbf{k}), \quad \widehat{\varphi}_{j_0}(\boldsymbol{\omega}) = 2^{-j_0} W_0(2^{-j_0}|\boldsymbol{\omega}|). \tag{12}$$

Hence, coarse-scale curvelets are nondirectional. The complete CT consists of the fine-scale directional curvelets $(\varphi_{j,l,k})_{j=j_0,l,k}$ and of the coarse-scale isotropic "father" wavelets $(\Phi_{j_0,k})_k$. It is the behavior of the fine-scale directional elements that is of interest here.

B. Curvelet filter for noise removal

In general, the contaminated region of the data or image may have any shape, such as, for example, a square, rectangle, or any other polygon. But, as already mentioned, due to its dispersive nature the GR noise in a seismic image typically appears as a fanlike coherent structure (see Figs. 1 and 9 below). For the purpose of using CT we need at least a quadrangle region. So we need at least four points in the original image in order to demarcate a quadrilateral contaminated region to be denoised. In this study, we assume for simplicity that the noisy section has a square shape, but, in general, if the contaminated region is not a square, we can shift and/or stretch the sides of the polygon to convert it to a square sector. The data points inside the polygon are then mapped onto the corresponding locations in the new rectangular domain, with the mapping carried out via a cubic convolution interpolation technique [40].

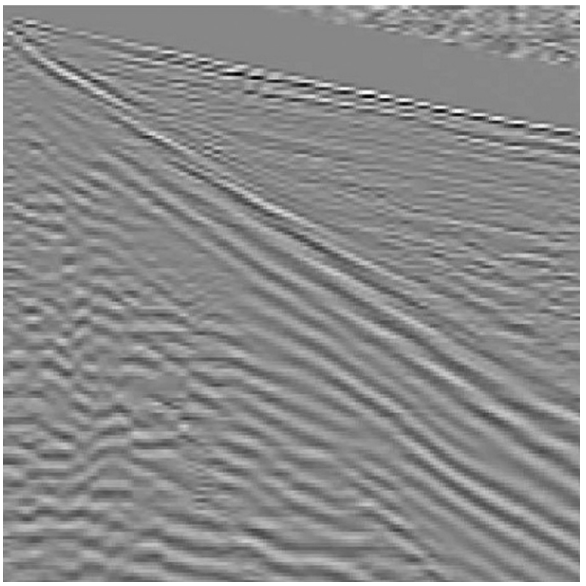

FIG. 1. A typical synthetic seismic image containing ground-roll noise [15].

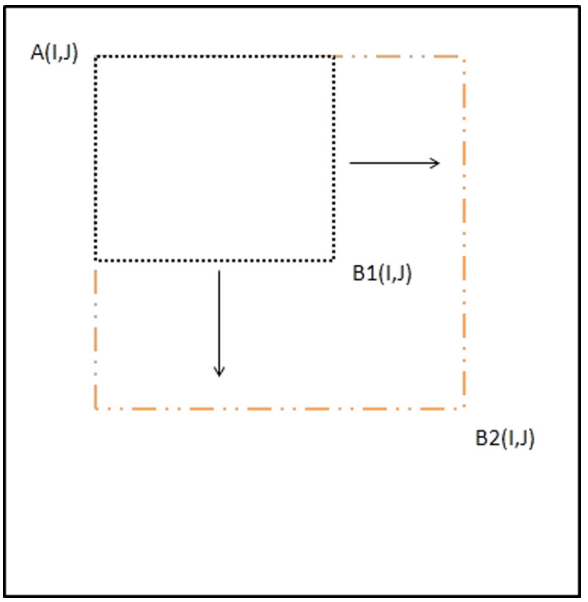

FIG. 2. (Color online) Demarcating the noisy region. Points A and B1 are the starting points, and B2 is the final point, which in this study is set to be the lower-right corner of the seismic image. We use l,k instead of I, J throughout the paper.

The space-time localization of the GR allows the possibility of applying the CT to suppress the noise, leaving intact the uncontaminated region. In order to demarcate the noisy region, we begin from a point in the left-upper section of the image and enlarge the alignment region, both vertically and horizontally. This is done by adding more pixels to the rows and columns of the matrix that we use in the KL transform (see below). Next, we use the horizontal lines to demarcate the noisy region. The process is repeated by starting from the four corners in the original image. Once the region contaminated by the GR has been demarcated, we map each sector onto a horizontal rectangular region by shifting and stretching along the time axis. We then use the CT to denoise the demarcated region, i.e., set the corresponding curvelet coefficients to zero, and then reconstruct (transform) it back to the original image. To implement the procedure on a computer, a rectangular area of seismic image is used as the initial noisy region, and then expand the its dimensions. Figure 1 presents a set of data, and Fig. 2 shows the rectangular area.

Suppose we have chosen l sectors to demarcate the different wave trains in the contaminated region of the original data, and let $\{\theta_1, \ldots, \theta_N\}$ be the set of the parameters characterizing the respective alignment functions that define the sectors. Let us denote by $\hat{A}_k$, $k = 1, \ldots, l$ the matrix representing the kth transformed sector obtained from linear mapping of the respective original sector, as discussed above. For each transformed sector $\hat{A}_k$ we compute its KL transform. For the data matrix $\mathbf{A}$ the KL transform is defined as the $m \times n$ matrix $\mathbf{\Psi}$ (with m and n being the dimensions of the seismic image), given by

$$\mathbf{\Psi} = \mathbf{U}^{\mathrm{T}}\mathbf{A}, \tag{13}$$

where the columns of $\mathbf{U}$ are the eigenvectors of $\mathbf{\Gamma} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$,

$$\mathbf{U} = (\mathbf{u}_1\ \mathbf{u}_2 \cdots \mathbf{u}_m), \tag{14}$$

The original data are recovered by the inverse KL transform,

$$\mathbf{A} = \mathbf{U}\mathbf{\Psi}. \tag{15}$$

We then calculate a *coherence index* C_k for that sector, defined as the relative energy contained in its first KL mode:

$$C_k = \frac{\lambda_1^k}{\sum_{i=1}^{r_k} \lambda_i^k}, \tag{16}$$

where λ_i^k are the eigenvalues of the correlation matrix $\widehat{\Gamma}_k = \hat{A}_k \hat{A}_k^{\mathrm{T}}$, with r_k being the rank of $\hat{A}_k$. As defined, C_k represents the relative weight of the most coherent mode in the KL expansion of the transformed sector $\hat{A}_k$. Note that a quantity analogous to the CI was presented in Ref. [15] for the KL filter.

Next, we introduce an overall CI, $C(\theta_1, \ldots, \theta_N)$, for the entire demarcated region, defined as the arithmetic average of CI of all the sectors:

$$C(\theta_1, \ldots, \theta_N) = \frac{1}{l} \sum_{k=1}^{l} C_k. \tag{17}$$

As the name suggests, the CI is a measure of the amount of coherent energy contained in the selected demarcated region given by the parameters $\{\theta_i\}_{i=1}^N$. Thus, the higher C, the larger is the energy contained in the most coherent modes in a given region. For the purpose of removing the coherent noise, it is, therefore, mostly favorable to pick the region with the largest possible C. Thus, we propose the following criterion to select the optimal region to be filtered: vary the parameters $\{\theta_i\}$ over some appropriate range and then choose the values θ_i^* that maximize C; that is,

$$C(\theta_1^*, \ldots, \theta_N^*) = \max_{\{\theta_i\}} C(\theta_1, \ldots, \theta_N). \tag{18}$$

Once we have selected the optimal region, given by the parameters $\{\theta_i\}_{i=1}^{N}$, we simply apply the CT filter to the region, the procedure for which is described next.

By definition, the coherence CI is defined as the amount of coherent energy contained in a given region. The energy E of a database is defined as the sum of the eigenvalues of symmetric matrix Γ:

$$E = \sum_i \lambda_i. \tag{19}$$

Thus, λ_i and $E_i = \lambda_i/E$ as the energy and relative energy captured by $\mathbf{u}_i$, the ith eigenvector. Next, consider the covariance matrix of $\mathbf{A}$, defined by

$$\mathrm{cov}(\mathbf{X}) = \langle [\mathbf{X} - \langle \mathbf{X} \rangle][\mathbf{X} - \langle \mathbf{X} \rangle]^{\mathrm{T}} \rangle. \tag{20}$$

If we compare the covariance matrix defined by Eq. (20) with the matrix $\mathbf{\Gamma} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$, we see that they are very similar, implying that $\mathbf{\Gamma}$ is a covariance-like matrix as well. This allows us to interpret λ_i as the variance of $\mathbf{u}_i$. Therefore, larger λ_i and, thus, larger E_i indicate higher coherence in the KL mode of $\mathbf{u}_i$. As a result, the higher the coherence energy, the larger CI will be in the selected region.

C. Denoising

In this section we describe in some detail the denoising methodology. Elimination of the coherent noise, such the GR noise, consists of three main steps: (1) decomposition of the data in the curvelet space, (2) identification and removal of the GR coefficients in the angular sectors in the curvelet space, and (3) reconstruction of the data after the coherent noise is removed.

(1) Curvelet analysis and decomposition of the data. The seismic data are decomposed in the curvelet space. This is accomplished by applying the fast discrete CT, as already mentioned, which consists, roughly speaking, of computing the following inner products in the Fourier domain. Based on Plancherel's theorem we have

$$C_{j,l,k} = \langle f, \varphi_{j,l,k} \rangle = \langle \hat{f}, \widehat{\varphi}_{j,l,k} \rangle. \tag{21}$$

(2) Identification and removal of the GR noise. Next, we identify and erase (set to zero) the curvelet coefficients of the angular sections that correspond to the GR noise. Following Eq. (1), we split the curvelet decomposition of the signal or data f into two parts,

$$f = \sum_j \sum_{l \in \mathrm{GR}} \sum_k \langle f, \varphi_{j,l,k} \rangle \varphi_{j,l,k} + \sum_j \sum_{l \notin \mathrm{GR}} \sum_k \langle f, \varphi_{j,l,k} \rangle \varphi_{j,l,k}. \tag{22}$$

We emphasize that the selection of the GR coefficients is done only by the l index, the angular coefficients. The GR in the seismic image is formed by almost vertical lines. Therefore, the GR in the curvelet space is represented by the directional components near the horizontal direction.

(3) Reconstruction of the data. After erasing the GR coefficients from the second term of Eq. (18), one obtains the denoised signal. We note that this process is performed without any artificial attenuation (no leakage and excellent localization), while the angular coefficients corresponding to the GR are completely removed from the seismic image.

III. RESULTS AND DISCUSSION

We used two synthetic and two actual sets of seismic data with distinct noise content in order to test the accuracy of the method. The first sample data were taken from Ref. [15], which had been provided by the Brazilian petroleum company PETROBARS. The second synthetic image was provided by Geokinetics, a geophysical service company. The real seismic data were provided by the National Iranian Oil Company and Geokinetics. A seismic section consists of several traces, where a trace is the recorded signal from a single geophone (receiver). Standard imaging techniques are used to generate seismic images from a set of traces. Usually, a Ricker wavelet, sometimes referred to as the "Mexican hat" wavelet,

$$\psi(x) = \frac{2}{\pi^{1/4}\sqrt{3\sigma}} \left(1 - \frac{x^2}{\sigma^2}\right) \exp\left(-\frac{x^2}{2\sigma^2}\right),$$

which is essentially the second Hermite function, is used to generate the synthetic traces and seismic sections. Then, coherent noise, mimicking the GR noise, with a lower frequency range is added to the traces in order to create noisy images. In a seismic image, the horizontal axis represents the distance between the seismic source (the device for generating seismic energy) and the receiver. The vertical axis corresponds to time. A seismic image can be displayed by gray, colored, or wiggle modes, and here we use a gray mode to show seismic images. The gray levels in the images change linearly from black to white, as the amplitude of the seismic signal varies from minimum to maximum. Owing to its dispersive nature, the GR in a seismic image appears as a fanlike structure, which is visible in Figs. 1, 5, 8, and 10.

We first applied the method to the synthetic data and image shown in Fig. 1. We began from a point in the x-y plane at the image's top left side, and for each choice of k and l in Eqs. (14) and (15) we applied the procedure described in the previous section, in order to compute CI(k,l) for the corresponding region. In Fig. 3 we show the energy surface CI(k,l), indicating that it possesses a sharp peak, hence demonstrating that the criterion based on CI is indeed quite discriminating with respect to the positioning of the lines demarcating the region contaminated by the coherent GR noise. The global maximum

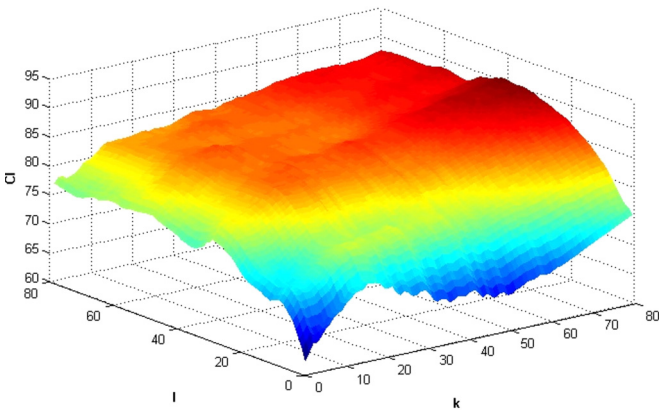

FIG. 3. (Color online) Normalized coherence index (CI), calculated by the KL transform. Higher CI corresponds to the coherent noise region.

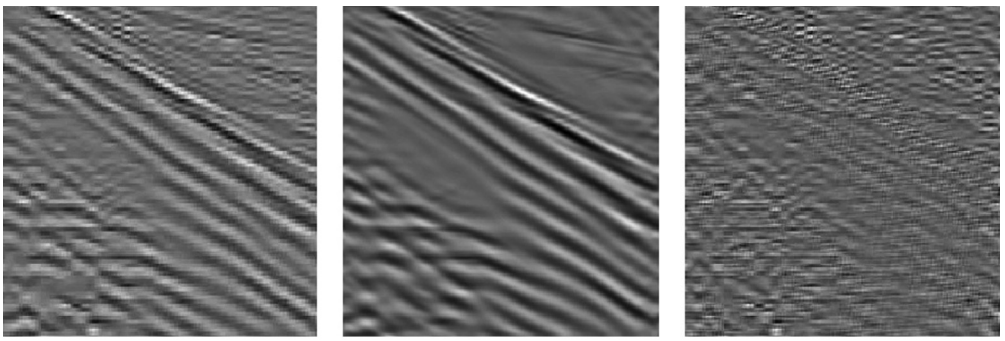

FIG. 4. (a) The demarcated region of the data shown in Fig. 1, corresponding to the maximum of CI in Fig. 3. (b) The denoised image of part of the data shown in (a), using the curvelet transform. (c) The noise removed from the image in (a).

of CI in Fig. 3 is located at $k = 80$ and $l = 40$. To confirm the position of the global maximum of CI, we carried out the same procedure for the other three corners of the image and computed the corresponding CIs. To do so, one needs the coordinates of the four corners of the rectangle that represents the noisy region. If we repeat the KL transform starting from the end of the image matrix, we obtain a new pair (k,l) that corresponds to another corner of the noisy region. Since we need only the coordinates of two corners in a rectangle to demarcate it, the KL transform is repeated only for the lower right side of the image. Indeed, one may assume that, macroscopically, the GR region in the seismic image has a triangular shape. However, if the image is inspected at finer scales, the GR region appears as one made of almost vertical lines. The directional character of the curvelet transform (CT) allows sharp erasing of such sectors. Here, for simplicity, we assumed a rectangular shape for the noisy zone, but in general linear mapping can be used to transform the identified region into a rectangular one.

Figure 4(a) presents the demarcated region that we identified in the synthetic data of Fig. 1 using Fig. 3. One clearly sees that the GR wave trains appear mostly as horizontal events. After locating the most coherent region, we used the CT to remove the GR noise from the demarcated region. In Fig. 4(b) we present the result. Figure 4(b) indicates clearly that by removing only the GR noise from the demarcated region, the main horizontal events corresponding to the noise have been greatly suppressed. Since we have a new image with much less GR noise, we can again apply the algorithm to identify and treat a larger noisy area of the image for large seismic images with a various level of coherent noise. Figure 4(c) depicts the amount of data that has been taken out as the noise from the original image. In fact, Fig. 4(c) is obtained by subtracting Fig. 4(b) from Fig. 4(a).

To demonstrate the difference between our proposed algorithm and other methods that also utilize the CT, we present another example in Fig. 5. Figure 5(a) indicates that the seismic image is polluted with a high amount of coherent GR noise. We therefore used the CT to denoise the entire image, as

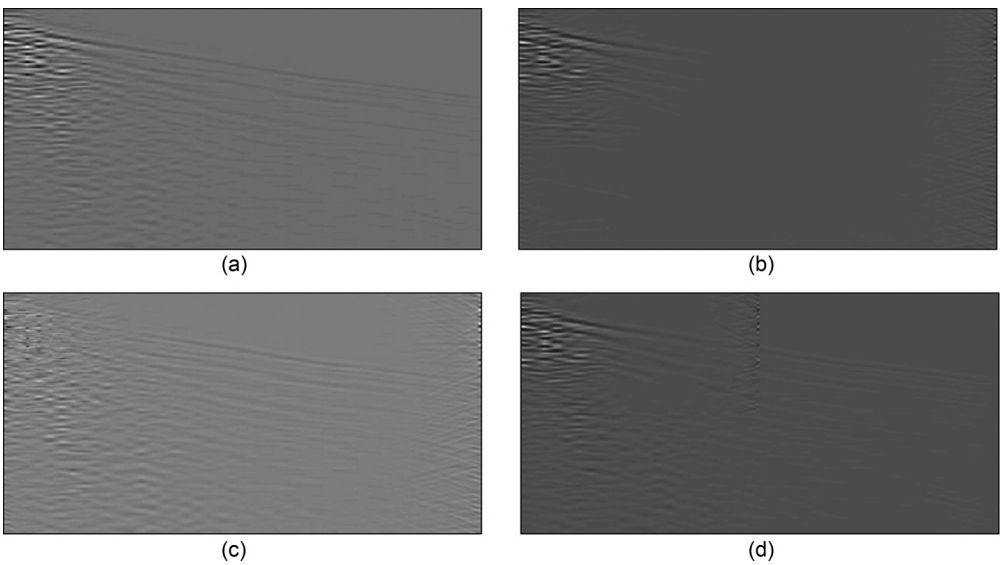

FIG. 5. (a) Synthetic seismic data with GR noise (courtesy of Geokinetics). (b) The denoised image of the data in (a). In this case we used the CT to denoise the *entire image*, but, as shown in the figure, some of the useful data have also been removed from the image. (c) The data removed from the image in (a). (d) The denoised image using the proposed method. In this image only the coherent sections, corresponding to the high values of the CI, have been denoised.

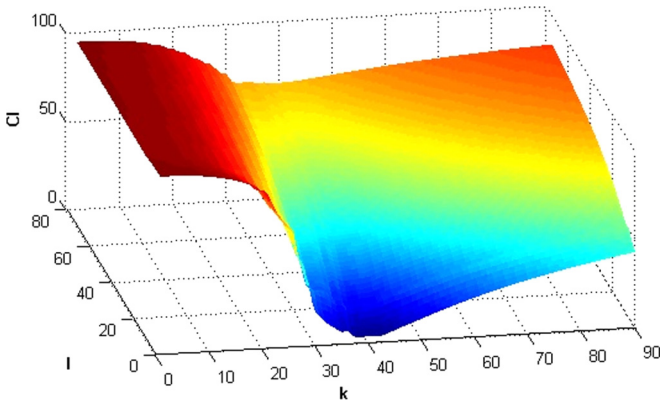

FIG. 6. (Color online) The computed coherence index CI for the seismic data of Fig. 5(a).

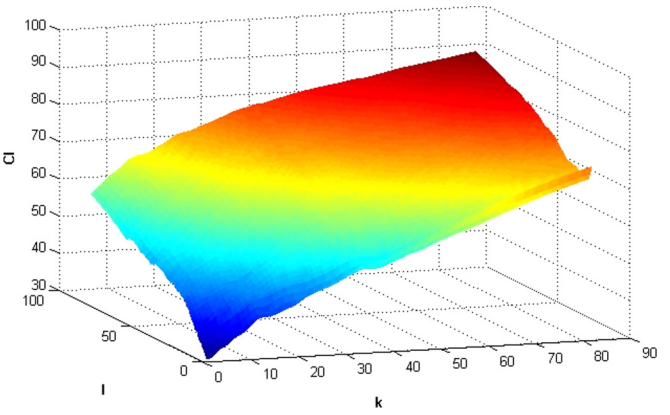

FIG. 8. (Color online) The computed coherence index CI for the seismic data of Fig. 7.

other researchers have done in the past. The denoised image is shown in Fig. 5(b), while the removed noisy data are shown in Fig. 5(c). Although the goal is removing only the GR from the entire image, the CT has also removed some of the useful information from the seismic image. This becomes clear if we compare Figs. 5(a), 5(b), and 5(c). The CT, as has been applied in the past, has removed some useful details and information from the right side of the image, whereas the coherent GR noise exists mostly on the image's left side.

Next, we applied our proposed algorithm to denoise Fig. 5(a). First, the corresponding CI for the image was computed. The result is displayed in Fig. 6. As Fig. 6 indicates, the maximum CI is linked with the upper left side of the image, occurring at $l = 80$ and $k = 14$. We then denoised the demarcated region; the denoised image is shown in Fig. 5(d). In fact, only the upper left side of the image has been denoised, while the rest of the image has remained intact. Comparing Figs. 5(b) and 5(d), it is obvious that there is valuable information preserved in Fig. 5(d).

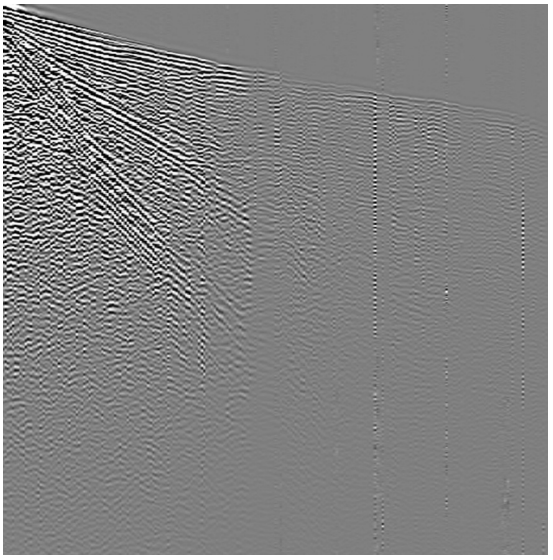

FIG. 7. Actual seismic data for an oil reservoir in southern Iran with GR noise.

We then applied the proposed method to actual seismic data, shown in Fig. 7. The corresponding computed CI is shown in Fig. 8. Carrying out the procedure for denoising using the computed CI, the image yields the demarcated, denoised, and removed data images shown in Figs. 9(a), 5(b), and 9(c). Figure 9(b) shows the filtered image after removing the GR from the demarcated region. The computed CI indicates that the coherent GR noise exists in the entire image, and, thus, we denoised the entire image. If one wishes to filter out the GR noise further, one may consider successively lower values of the CI. Iterating the scheme too many times is not, however, recommended, as it would begin to degrade the relevant data as well.

For all the previous images we used one-half of the seismic shot to demonstrate the applicability of the CI for demarcating the coherent noise region. The assumption behind the procedure is that the coherent noise begins from one side of the rectangle and propagates to another side. Figure 10 presents a single of a second actual seismic shot that contains the GR noise. As Fig. 10 indicates, the assumption is not applicable in this case, since the noise begins from the middle of the upper side of the image and propagates towards the two vertical sides, as well as the lower side of the image. In this case we split the image in two halves and computed the CI for both sides separately. We concatenate the two resulting CIs. The overall CI for the data shown in Fig. 10 is presented in Fig. 11. We then applied the CT to remove the noise from the demarcated region of the image, which is at the center of upper part. Figure 12(a) shows the denoised region inside the image. Next, we repeat the procedure for the purpose of denoising the other parts of the image. The demarcated region for the new image [Fig. 12(a)] is located at the center of the lower part of the image. The denoised result using the CT is displayed in Fig. 12(b).

IV. COMPUTATIONAL EFFICIENCY

To demonstrate the computational efficiency of the method, we compare the denoised data obtained with the traditional CT and our proposed CT denoising algorithm. The data used are the real seismic data for a hydrocarbon reservoir in southern Iran. For a given noisy image, the results demonstrate that

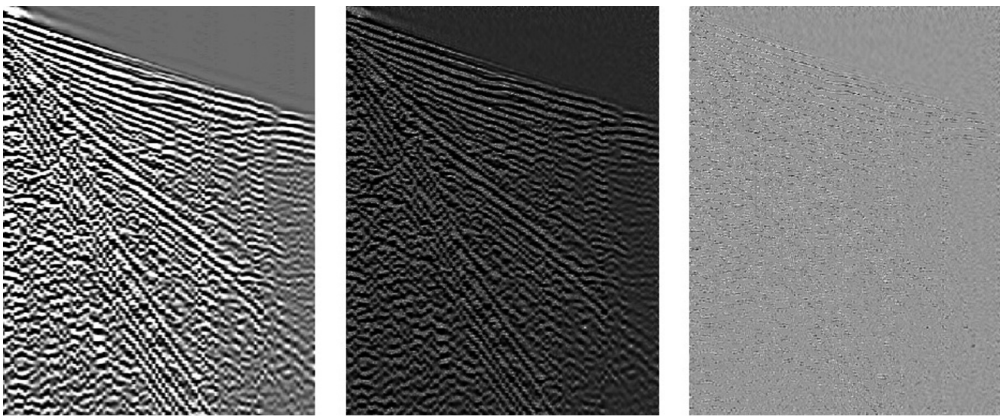

FIG. 9. (a) The demarcated region of the data shown in Fig. 7, corresponding to the maximum of CI in Fig. 8. (b) The denoised image of part of data shown in (a), using the curvelet transform method. (c) The data removed from (a).

our proposed CT method has significant advantages over the traditional scheme based on the CT, not only in terms of the visual effects, but also in terms of the computation time.

All the analysis were performed using Matlab (R2012a). In particular, all the results presented in this paper were obtained on a 1.70 GHz Asus notebook computer with 2.00 GB of memory. When denoising the real seismic data, the computation time for the conventional CT was 114 CPU sec, whereas the our proposed algorithm took only 19 CPU sec to denoise the same image. Moreover, we note that the proposed algorithm preserves the fragile seismic textures with higher precision than any other method. The conventional CT took 59 CPU sec to denoise the synthetic seismic data, whereas the corresponding time for the proposed algorithm was 14 CPU sec. At this point we do not have enough synthetic and real data to formulate more precisely the effect of the image size on the performance of our proposed algorithm. But, it is clear that, in general, we can state unconditionally that the computation time for the proposed algorithm is much less than that of the conventional CT. Moreover, we believe that the improvement in the computational efficiency with 3D seismic data and, more importantly, with 4D seismic data (3D seismic data collected repeatedly over a period of time) will be even better than what we report here, which we hope to demonstrate in a future paper.

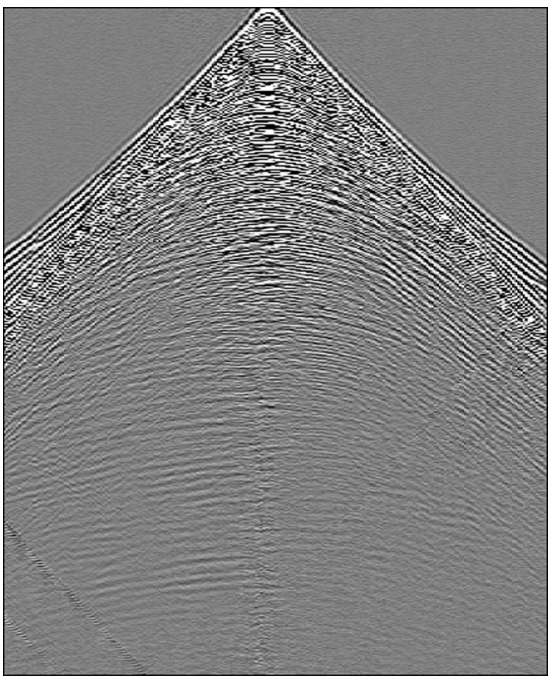

FIG. 10. Actual seismic data, one shot with the GR noise, for a shallow oil reservoir (courtesy of Geokinetics).

V. SUMMARY

A new algorithm based on the curvelet transform was introduced for processing geophysical data that are contaminated with coherent noise. A great advantage of the algorithm is its local nature, meaning that only the contaminated segment of the data is processed for denoising, hence allowing the removal of the coherent noise without distorting the rest of the data. The method uses an optimization process whereby the region of the coherent noise is selected, so as to maximize an appropriately defined coherence index. As an application, we applied the method to removing the ground roll, a coherent noise, from 2D seismic images. Both synthetic and real seismic data were analyzed.

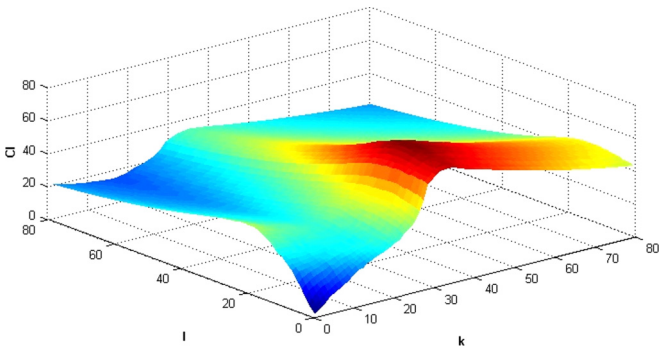

FIG. 11. (Color online) The computed coherence index CI for the entire image shown in Fig. 10.

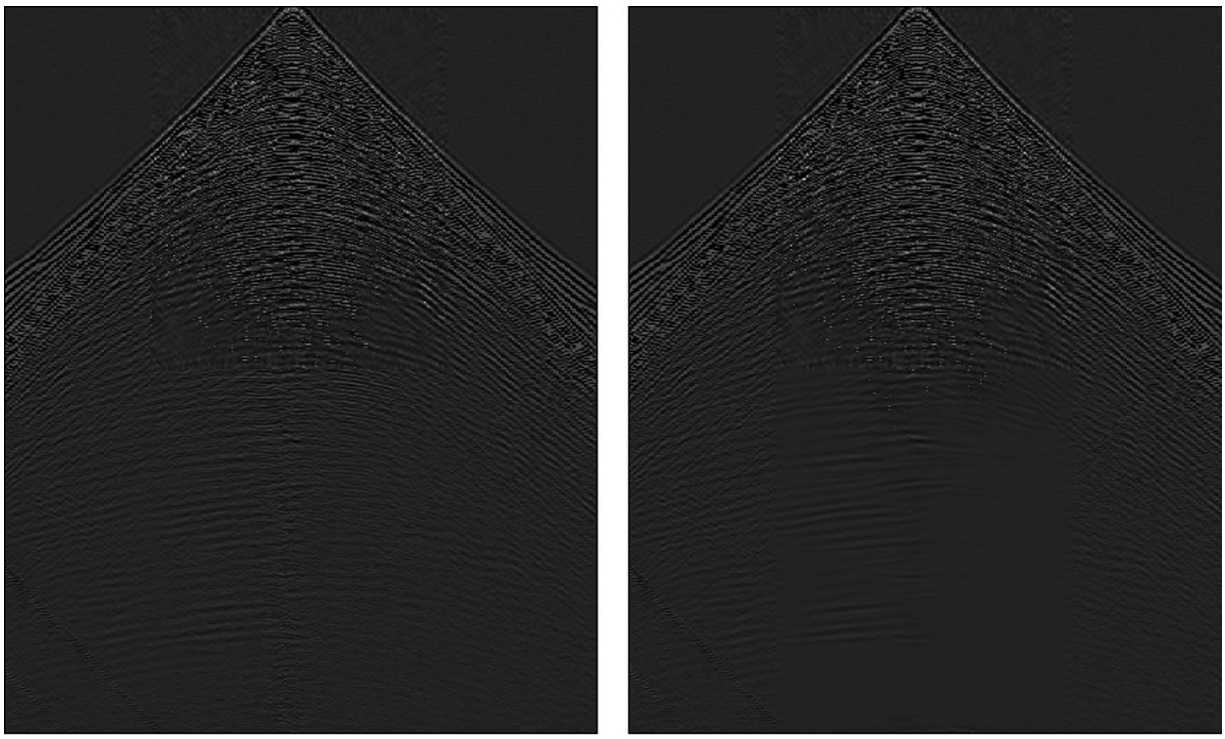

FIG. 12. (a) Demarcating and denoising, (a) upper part and (b) lower part of the image for the seismic data shown in Fig. 10.

The only input that the method requires is the alignment functions to be used in the computations, as well as the number of cycles that one wishes to repeat the procedure in order to make sure that all or most of the coherent noise has been removed from the selected region. These may vary, depending on the specific application. Once the choices are made, however, the denoising task proceeds in an automated fashion. The image-processing technique described in this study is also relevant for other applications, where coherent structures embedded in a complex spatiotemporal pattern need to be identified in an refined manner.

The main theme of this paper was to introduce an algorithm based on the CT for processing of geophysical data, but the same procedure can be used along with wavelet and contourlet transformations, and even a scheme that combines the wavelet and curvelet transforms, in denoising of geophysical data, each of which has its own advantages for application to real data.

ACKNOWLEDGMENTS

H.D. is grateful to the Viterbi School of Engineering at the University of Southern California for a doctoral fellowship. We are also grateful to Lee Bell of Geokinetics for providing us with the data of Figs. 5 and 10. This work was also supported in part by the Research Partnership to Secure Energy for America (RPSEA) Consortium.

[1] W. Menke, *Geophysical Data Processing: Discrete Inverse Theory*, 3rd ed. (Academic Press, San Diego, 2012); J. F. Claerbout, *Fundamentals of Geophysical Data Processing* (Blackwell Scientific, Palo Alto, CA, 1985).

[2] M. Sahimi, *Flow and Transport in Porous Media and Fractured Rock*, 2nd ed. (Wiley-VCH, Weinheim, 2011).

[3] R. Friedlich, J. Peinke, M. Sahimi, and M. R. Rahimi Tabar, Phys. Rep. **506**, 87 (2011).

[4] T. A. Tafti, M. Sahimi, F. Aminzadeh, and C. G. Sammis, Phys. Rev. E **87**, 032152 (2013), and references therein.

[5] G. Hennenfent and F. J. Herrmann, Comput. Sci. Eng. **8**, 16 (2006).

[6] F. J. Herrmann and G. Hennenfent, Geophys. J. Inter. **173**, 233 (2008).

[7] F. Shahbazi, A. Bahraminasab, S. M. Vaez Allaei, M. Sahimi, and M. R. Rahimi Tabar, Phys. Rev. Lett. **94**, 165505 (2005); S. M. Vaez Allaei and M. Sahimi, *ibid.* **96**, 075507 (2006); A. Bahraminasab, S. M. Vaez Allaei, F. Shahbazi, M. Sahimi, M. D. Niry, and M. R. Rahimi Tabar, Phys. Rev. B **75**, 064301 (2007); R. Sepehrinia, A. Bahraminasab, M. Sahimi, and M. R. Rahimi Tabar, *ibid.* **77**, 014203 (2008); R. Sepehrinia, M. R. Rahimi Tabar, and M. Sahimi, *ibid.* **78**, 024207 (2008).

[8] X. Liu, Geophysics **64**, 564 (1999).

[9] O. Yilmaz, *Seismic Data Processing* (Society of Exploration Geophysicists, Tulsa, OK, 2003).
[10] K. Karhunen, Ann. Acad. Sci. Fennicae Ser. A **37**, 1 (1947); M. Loève, *Probability Theory*, Vol. 2, 4th ed. (Springer, New York, 1978).
[11] B. R. Hunt and O. Kubler, IEEE Trans. Acoustics Speech Signal Process. **32**, 592 (1984).
[12] K. V. M. Fernando and H. Nicholson, IEEE Proc. **127**, 155 (1980).
[13] I. F. Jones, Ph.D. thesis, University of British Columbia, Vancouver, Canada, 1985, http://circle.ubc.ca/handle/2429/25831.
[14] S. Ribaric and I. Fratric, IEEE Trans. Pattern Anal. Machine Intelligence **27**, 1698 (2005).
[15] R. Montagne and G. L. Vasconcelos, Phys. Rev. E **74**, 016213 (2006).
[16] L. L. Canales, http://www.onepetro.org/mslib/servlet/onepetropreview?id=SEG-1984-0525.
[17] S. R. Deans, *The Radon Transform and Some of Its Applications* (Wiley, New York, 1983).
[18] E. K. Mary, J. H. Alistair, and A. O. John, Geophysics **55**, 1202 (1990); Z. Binzhou and S. A. Greenhalgh, *ibid.* **59**, 1133 (1994).
[19] P. Kumar and E. Foufoula-Georgiou, Rev. Geophys. **35**, 385 (1997); M. Sahimi, Granular Matter **3**, 3 (2001); Comput. Sci. Eng. **5**, 75 (2003).
[20] P. Gendron, in *Proc. International Workshop on Wavelets and Statistics, ISDS Statistics Week at Duke University* (Duke University, Durham, North Carolina, 1997), p. 12.
[21] A. A. Chanerley and N. A. Alexander, in *Proc. Sixth Conference on Computational Structures Technology*, edited by B. H. V. Topping and Z. Bittnar (Civil-Comp Press, Stirling, UK, 2002), p. 107.
[22] G. Corso, P. Kuhn, L. Lucena, and Z. Thomé, Physica A **318**, 551 (2003).
[23] A. R. Mehrabi and M. Sahimi, Phys. Rev. Lett. **79**, 4385 (1997); F. Ebrahimi and M. Sahimi, Transp. Porous Media **57**, 75 (2004); M. R. Rasaei and M. Sahimi, Comput. Geosci. **13**, 187 (2009), and references therein.
[24] M. Sahimi and M. Hashemi, Geophys. Res. Lett. **28**, 611 (2001); H. Dashtian, G. R. Jafari, Z. Koohi Lai, M. Masihi, and M. Sahimi, Transp. Porous Media **90**, 445 (2011), and references therein.
[25] E. Pazhoohesh, H. Hamzehpour, and M. Sahimi, Phys. Rev. B **73**, 174206 (2006), and references therein.
[26] M. N. Do and M. Vetterli, in *Beyond Wavelets*, Vol. 4, edited by G. V. Welland (Academic Press, Amsterdam, 2003), p. 83; IEEE Trans. Image Process. **14**, 2091 (2005).
[27] D. D. Y. Po and M. N. Do, IEEE Trans. Image Process. **15**, 1610 (2006).
[28] E. J. Candès and D. L. Donoho, in *Curves and Surface Fitting: Sait-Malo* 1999, edited by A. Cohen, C. Rabut, and L. Schumaker (Vanderbilt University Press, Nashville, TN, 2000), p. 105.
[29] D. L. Donoho, Ann. Statist. **27**, 859 (1999); SIAM J. Math. Anal. **31**, 1062 (2000).
[30] E. J. Candès and D. L. Donoho, Commun. Pure Appl. Math. **57**, 219 (2004); see also E. J. Candès and F. Guo, Signal Process. **82**, 1519 (2002).
[31] E. J. Candès, L. Demanet, D. L. Donoho, and L. Ying, Multiscale Modeling Simulation **5**, 861 (2006); J. Ma and G. Plonka, IEEE Signal Process. Mag. **27**, 118 (2010).
[32] F. J. Herrmann, D. Wang, G. Hennenfent, and P. P. Moghaddam, Geophysics **73**, A1 (2008).
[33] F. J. Herrmann, P. P. Moghaddam, and C. Stolk, Appl. Comput. Harmonic Analysis **24**, 150 (2008).
[34] C. Yarham, U. Boeniger, and F. J. Herrmann, Society of Exploration Geophysicists (SEG) Technical Program, Expanded Abstracts **25**, 2777 (2006).
[35] C. Yarham and F. J. Herrmann, Society of Exploration Geophysicists (SEG) Technical Program, Expanded Abstracts **27**, 2576 (2008).
[36] S. Sardy, A. G. Bruce, and P. Tseng, J. Comput. Graph. Statis. **9**, 361 (2000).
[37] S. Olhede and A. T. Walden, Proc. R. Soc. Lond. A **460**, 955 (2004).
[38] E. J. Candés and D. L. Donoho, *Wavelet Applications in Signal and Image Processing VIII, 108*, edited by A. Aldroubi, A. F. Laine, and M. A. Unser, Proc. SPIE 4119.
[39] S. Park and R. Schowengerdt, Comput. Vis. Graph. Image Proc. **23**, 258 (1983).