Curvelet transforms and filtering of seismic attributes for reservoir modeling

Matz Haugen and Tapan Mukerji
Department of Energy Resources Engineering
Stanford University

Abstract
Seismic attributes play an important role in reservoir characterization and geostatistical modeling. Geometric seismic attributes capture and enhance spatial features such as faults and channels. Sometimes these attributes may be hidden within specific spectral bands, and may require extraction by spectral decomposition. We explore curvelet transforms and filtering of coherence attributes based on curvelet transforms to extract useful seismic attributes for characterizing reservoir morphology and spatial architecture.

1. Introduction

Because of their exhaustive spatial extent, seismic attributes are important properties that help increase the understanding of a reservoir, and play a key role as secondary data in geostatistical modeling. Seismic attributes, which include reflectivities, velocities, impedances, and others, are derived from seismic data using different processing, analysis, or inversion techniques. Some seismic attributes such as impedance, respond to the reservoir layer properties, and can be related to the local rock properties through rock physics relations. Other attributes, such as reflectivity and AVO gradient respond to contrasts in local rock properties at interfaces between layers. Finally, there are attributes that respond to spatial and morphological characteristics of the reservoir. These geometric seismic attributes include coherence and curvature. Chopra and Marfurt (2007) is a good reference for many of these geometric attributes. Often, the information of interest is hidden within a localized spectral band, and has to be extracted using various techniques of spectral decomposition. For visual interpretation purposes, seismic data can be manipulated or filtered in order to enhance certain features of the reservoir, e.g. the curvature attribute (Al-Dossary and Marfurt, 2006; Chopra and Marfurt, 2007). This paper presents our initial exploration of curvelet transforms and curvelet based filtering of geometric seismic attributes that can be useful for reservoir characterization.

Sharp edges along a seismic amplitude horizon or a time slice could be interpreted as fractures. Currently, there are many techniques to illuminate these features from the data, some of which are incorporated into commercial software. The coherence attribute is one well documented technique that preserves edges in the seismic data (Gersztenkorn and Marfurt, 1999; Bahorich and Farmer, 1995). Both curvature and coherence do a good job. However, these methods still output a broad range of scales. Since fractures are scale
independent to a certain extent, small scaled fractures might overlap large scale fractures and the image might be challenging to interpret. In order to separate the different scales, the image can be decomposed into multiple frequency bands which can be separately inverted back into the spatial domain. The curvelet decomposition is one method that is well suited for this job.

2. Theory of Curvelet Transforms

Curvelets are multi-scale functions that are optimally suited for sparsely reconstructing 2D images with edges. The method was initially developed by E. J. Candès and D. L. Donoho in 1999 (Candès and Donoho, 1999), and was later expanded into the discrete domain (Candès et al., 2006) from which most of the following theory and notation is taken.

In order to explain the Fast Discrete Curvelet Transform (FDCT), some definitions must first be introduced. Because most images are supplied in Cartesian coordinates the corresponding Fourier domain is also in Cartesian coordinates, and consequently, the method is developed in a coordinate system with \([x_1, x_2]\) as spatial coordinates with \(0 < x_1, x_2 < n\) where \(n\) is the number of samples in each direction. Similarly, the spectral domain has coordinates \([\omega_1, \omega_2]\). Now, separate the frequency domain in dyadic corona (concentric squares) by convolving with a window function, \(W_j(\omega_1, \omega_2)\), one for each corona. Because this is an isometric function, a short hand way of writing it could be \(W_j(\omega)\). The function can then be defined in the Fourier domain as

\[
W_j(\omega) = \sqrt{\Phi_{j+1}^2 - \Phi_j^2}, j \geq 0
\]

where \(\Phi\) is a 2D low pass filter at scale \(j\) defined as the product of two one dimensional low pass filters

\[
\Phi_j(\omega_1, \omega_2) = \phi(2^{-j} \omega_1)\phi(2^{-j} \omega_2)
\]

with \(\phi(\omega)\) as a one dimensional low pass filter which is equal to 1 on \([-1/2, 1/2]\) and 0 outside \([-2, 2]\). In addition to these two constraints, the region between \(\pm 1/2\) and \(\pm 2\) is an exponentially decreasing function between 0 and 1 so as to make the filter infinitely differentiable and continuous.

Furthermore, each concentric square can be separated into trapezoids (see Figure 1) using a band pass filter

\[
V_j(\omega) = V(2^{[j/2]} \omega_2 / \omega_1)
\]
which isolates a rectangle and where the exponent is rounded down to the nearest integer.

Now, we can define a scale-dependent window in the Fourier domain

\[ \tilde{U}_j(\omega) = W_j(\omega)V_j(\omega) \]

which isolates tiles in the spectrum where \( \{\omega_1, \omega_2 : 2^j \leq \omega_2 \leq 2^{j+1}, -2^{-j/2} < \omega_1 < 2^{-j/2}\} \). In order to salvage a trapezoid in the Cartesian coordinates, the window \( V_j(\omega) \) must be modified to be dependent on where in the corona the trapezoid is located, i.e. dependent on an azimuthal angle \( \theta \in [-\pi/2, \pi/2] \), which again defines a set of slopes in the Fourier space: \( \tan(\theta) = l \cdot 2^{-[j/2]} \) with \( l = -2^{[j/2]} \ldots, 2^{[j/2]} - 1 \). Now we can define

\[ \tilde{U}_{j,l}(\omega) = W_j(\omega)V_j(S_\theta \omega) \]

where

\[ S_\theta = \begin{pmatrix} 1 & 0 \\ -\tan \theta & 1 \end{pmatrix} \]

is the shearing matrix applied on the rectangular filter \( V_j(\omega) \).

With a complete 2D window function, it can be shown that

\[ \sum_{\text{all angles}} \sum_{\text{all scales}} \left| \tilde{U}_{j,l}(\omega) \right|^2 = 1 \]

Figure 1: A decomposition of the 2D frequencies into a set of digital tiles using the filter \( \tilde{U}_{j,l}(\omega) \)
The continuous version of this window function is given in (Candès and Donoho, 2005a and 2005b), including a proof of the normalization of $\tilde{U}_{j,l}(\omega)$.

Keeping in mind that in the end, a basis is desired from which a set of functions can be linearly combined with a known set of coefficient to reconstruct the image, the discrete curvelet can be defined in the 2D Fourier domain as

$$
\hat{\phi}^D_{j,l,k}(\omega) = \tilde{U}_{j,l}(\omega) \cdot e^{2\pi i \bar{v} \cdot \bar{k}} / \sqrt{L_{1,j,l} \cdot L_{2,j,l}}
$$

where $\bar{v}$ is the vector $(v_1, v_2) = (\omega_1 / L_{1,j,l}, \omega_2 / L_{2,j,l})$ and $\bar{k}$ is the position vector $(k_1, k_2)$ with $0 \leq k_1 < L_{1,j,l}$ and $0 \leq k_2 < L_{2,j,l}$. $L_{1,j,l}$ and $L_{2,j,l}$ are the integer lengths and widths over which the window function $\tilde{U}_{j,l}(\omega)$ is "approximately" non-zero. The reason for the approximation is that the window filter is an infinitely smooth function that is very close to zero outside its specified area. Notice that the construction of $\hat{\phi}^D_{j,l,k}(\omega)$ ensures that it is periodic in $L_{1,j,l}$ and $L_{2,j,l}$. Another way of saying this is that the position vector is wrapped around the origin, which is why the technique is called FDCT by wrapping.

In order to obtain $\hat{\phi}^D_{j,l,k}(\omega)$, the inverse 2DFFT can be computed. However, each of the functions needs coefficients so that they constitute a linear basis. These coefficients are obtained by taking the inner product between the function $f(\omega_1, \omega_2)$ and each of the curvelet function as follows

$$
e^D(j,l,k) = \sum_{\omega_1, \omega_2} \hat{f}(\omega_1, \omega_2) \hat{\phi}^D_{j,l,k}(\omega_1, \omega_2)
$$

where the range of $(\omega_1, \omega_2)$ is $-L_{1,j,l} / 2 \leq \omega_1 < L_{1,j,l} / 2$ and $-L_{2,j,l} / 2 \leq \omega_2 < L_{2,j,l} / 2$. $D$ stands for discrete.

It has been shown in Candès and Donoho (2002) that the curvelets form a tight frame with a reconstruction formula

$$f = \sum_{j,l,k} \langle \varphi_{j,l,k}, f \rangle \varphi_{j,l,k}
$$

and with a Parseval relation over $L_2$

$$\sum_{j,l,k} |\langle \varphi_{j,l,k}, f \rangle|^2 = \|f\|_{L^2}^2 \forall f \in L^2(R^2)$$
For convenience, we can define a wrapping operator which takes each trapezoid and wraps it around the origin. Call this operator \( w \). This allows the inverse FFT to be applied on a rectangle centered around the origin as opposed to the location of \( \tilde{U}_{j,l}(\omega) \). A short summary of the FDCT can now be presented.

1. Apply the 2D FFT on the input image, \( f(x_1, x_2) \), yielding \( f(\omega_1, \omega_2) \) with
   \[-n/2 \leq (\omega_1, \omega_2) < n/2.\] \( n^2 \) is the number of pixels.
2. At each scale \( j \) and \( l \), obtain the product \( \tilde{U}_{j,l}(\omega) \hat{f}(\omega_1, \omega_2) \)
3. Wrap this around the origin by \( w(\tilde{U}_{j,l}(\omega) \hat{f}(\omega_1, \omega_2)) \) where the vector \( (\omega_1, \omega_2) \) has changed its range to \(-L_{n,j,l}/2 \leq \omega_i < L_{n,j,l}/2\) for \( i = \{1,2\} \)
4. Apply the inverse 2D FFT on this rectangular region to obtain the coefficients \( c_D^{j,l}(j,l,k) \).
5. Use desired amount of coefficients to recreate the modified image:
   \[ f_{\text{med}} = \sum_{j,l,k} c_D^{j,l}(j,l,k) \varphi_{j,l,k}^{p}(\omega_1, \omega_2) \]

This theory has been applied in many aspects of geophysical signal processing, including seismic data recovery (Herrmann and Hennenfent, 2008), primary multiple-separation (Herrmann et al., 2007) and wavefield extrapolation (Lin and Herrmann, 2007). The 3D extension of the FDCT is in principle similar to the 2D theory and is explained in Ying and Demanet, (2005).

One of the reasons that curvelets are so successful in the field of geophysics is because the edges produced by seismic reflections are highly anisotropic, and thus require anisotropically scaled basis functions. Curvelets follow a parabolic scaling relationship where \( \text{length} \propto 2^{-j/2} \) and \( \text{width} \propto 2^{-j} \) yielding \( \text{width} \propto \text{length}^2 \). This is due to the rectangular window functions \( V_j(\omega) \). Seismic reflectors have a vertical resolution that scales as the wavelength, \( \lambda \), while the lateral resolution is related to the Fresnel zone which scales as \( \sqrt{Z} \) where \( Z \) is the depth. This gives rise to a parabolic anisotropic scaling that is similar to the scaling relation for curvelets. In addition to this, the curvelet is also geometrically steerable so that it can adapt to the direction of the edge discontinuity. One curvelet is shown in Figure 2, and as one can see, the curvelet is localized both in space and frequency which makes it suitable to reconstruct features of similar characteristics. This aids the sparse reconstruction of the image. However, it is recognized that not all constructions behave in this fashion, something that can limit the curvelet's potential.
On the whole, curvelets offer a different way of separating the frequency spectrum. Wavelet transform decomposition is an alternative method which can be used, even though these wavelets are not suitable for preserving curved edges, they do perform well on point discontinuities, for example along edges of an image.

Pure Fourier decomposition is another method which can be used. However, curvelet decomposition is the most sparse of the three mentioned methods when considering application on seismic data. To illustrate this more clearly, consider a binary function $f$ which has discontinuities along the edges and across a curve in $C^2$. It can then be shown that the squared difference between the approximate reconstruction of $f$ using $n$ number of Fourier basis functions and $f$ itself obeys the following relationship.

$$\| f - f_n^F \|^2_{L_2} \propto n^{-1/2}, n \to \infty$$

where $f_n^F$ are the functions with the $n$ largest coefficients.

With a similar starting function E.J Candès and D. L. Donoho have shown that curvelets have a convergence rate

$$\| f - f_n^T \|^2_{L_2} \propto n^{-2}, n \to \infty$$

which is an optimally sparse deconstruction meaning that there does not exist any decomposition that has a faster convergence rate. Wavelets land in between the Fourier and Curvelet transform with a convergence rate $\propto n^{-1}$ (Candès and Donoho, 2002).
It is important to note that the function $f$ is very simple and in real seismic examples, both multi-scale edge and point discontinuities are present.

3. Method

In this paper curvelets transforms have been incorporated to filter the coherence seismic attributes. There are a number of ways in which one can calculate coherence. Eigenstructure coherence was the method implemented on this data set, for which the reader is referred to the excellent paper by A. Gersztenkorn and K. J. Marfurt (1999). Other methods include semblance-based coherence and cross-correlation based coherence estimation (Marfurt and Kirlin, 1998). While eigenstructure coherence maximizes the lateral resolution it only accounts for the changes in the reflector waveform. This is also the case for cross correlation coherence. On the other hand, the semblance based method also accounts for the reflector amplitudes as well as the waveform (Chopra and Marfurt, 2007).

Coherence is a measure of structural order. Consider a small search cube, $C$, with the dimensions of one's choosing within a large cube of seismic traces with a total of $M$ inlines and $N$ crosslines with $O$ samples. If the search cube is reshaped from this $3^{rd}$ rank tensor to a matrix such that each column contains one trace from either inline or crossline, then the ratio of the largest eigenvalue of this matrix over the total sum of its eigenvalues will be a measure of coherence with values ranging from zero to one. For example, if the matrix has rank 1, the coherence will be at maximum one because all of the traces are just multiples of one trace. This is mediated by the largest eigenvalue being the only non-zero eigenvalue. As soon as the traces become independent of each other, the coherence will be less than one. So, depending on the search cube, the coherence value will give information about the amount of change within the cube. A large search cube will account for changes on a large scale and vice versa.

The coherence can be calculated for the post-stack seismic data, partial angle gathers, for seismic data filtered by the curvelet transform. The advantage of this is that multiple scales of the data set can be isolated individually. This can be advantageous when wanting to filter out the small fractures that sometimes obfuscate the larger picture.

In the example shown below, three frequency bands will be created, the high, middle and low frequencies. Each of the spatial frequency bands contain a set of frequencies which can be obtained by using the knowledge of the architecture of the window function, $W_j(\omega_1, \omega_2)$. 
At each scale $j$, the dual grid is separated by $W_j$, with each scale having an outward length from the origin larger than the previous scale by a factor of 2 in each direction, $x$, $y$ and $z$ (see Figure 1). Furthermore, the Euclidian distance on the dual grid from the origin to the coefficient represents the frequency of the corresponding curvelet. From Figure 1, the distance to the midpoint of each window function can be obtained. Call this distance $d$. A good approximation of the minimum and maximum frequency for the frequency band contained within the window would be $d/2$ and $2d$ respectively. Furthermore, the dual grid has the reciprocal units of the spatial grid. Given that each trace is separated at $\Delta x = 12.5m$ intervals, the dual grid unit, $\Delta f$, in the frequency domain will be the inverse of this divided by the total number of sample points, i.e. $\Delta f = 1/(N\Delta x)$. In this case the discretization in the x and y directions are the same.

With this information, each frequency band can be bounded. In the E-W direction, the mid band wavelength (inverse of spatial frequency) ranges from 40m to about 600m. The low band includes all wavelengths longer than 600m and the high band represents the shorter wavelengths, less than 40m. The numbers associated with the N-S direction are about the same size as the E-W direction.

As there is only one post-stack cube used as the input (full stack), a 3D curvelet transform will suffice. After this is done, the coherence analysis will be performed on each of the three resulting cubes. This will enhance the contrast along the various curved edges. A similar workflow has been implemented by Zhang et al. (2008).

4. Results and Discussion

The examples shown below are exploratory exercises initiated to familiarize ourselves with the curvelet transform software and to illustrate the effects of curvelet filtering based on scales and angles. The seismic dataset used is from offshore Equatorial Guinea, West Africa with 142 inlines and 135 crosslines (each spaced at 12.5m intervals). Each trace has 1500 time samples at 2 ms intervals. The time slice considered was at 1158 ms which has seismic amplitudes shown in Figure 3. The horizontal dimensions are 2.5 by 2.5 km (see (Dutta et al., 2007) for more background)
Figure 3: A horizontal slice showing amplitudes from a seismic cube taken at 1158 ms.

The slice is along a horizontal section covering a channel deposit system. When separating the cube into different scales, the scales used in the curvelet transform are grouped into three bands.

Each scale is isolated using the window function $W_j(\omega)$. For this particular example, 25 scales were used of which the lowest third of the scales were grouped into the “low” frequency band (wavelengths longer than 600m). Similarly, the middle (~ 40m to 600m) and high (shorter than 40m) frequency bands have the corresponding third of the scales. It is important to note that the bands do not have to have equal number of scales. This can be changed at the discretion of the analyst. The angles of the curvelets can also be restricted but in this case all angles are used.

The frequency-banded images in addition to the coherence calculations are shown in Figure 4.
Figure 4: A curvelet deconstruction of the seismic data taken at 1158 ms (shown in bottom 3 images) with the corresponding eigenstructure coherence (top 3 images).

The high coherence values (white) show that there might be a potential channel located within the ellipse. Because this is a timeslice, the channel might be discontinuous at certain places. The large scale coherence map shows a broad perimeter of the channel. Each of the frequency bands were analyzed using different search windows. In the high frequency case, the search cube was 3-by-3 square of traces in the horizontal plane with 5 samples, i.e. a 10 ms time period. The middle frequency band was calculated using the same horizontal dimensions as in the high frequency case with a 22 ms period, while the low frequencies had a 5-by-5-by-21 search matrix corresponding to a spatiotemporal block of 12.5 m-by-12.5 m-by-42 ms. In each case, the size of the search cube reflects the amount of change occurring within the designated area.
Another interesting case is the change of the coherence with respect to offset. The result of this is shown in Figure 5, where only the mid-band of the curvelets is included. As seen, the far offset is quite dissimilar to the near- and mid offsets, indicated by the lack of coherence in the potential river area. The near and mid offsets have some subtle differences but there is no significant advantage of one compared to the other.

Figure 5: Coherence (top) and seismic amplitude (bottom) with changing offsets.
Finally as another example, an angular filter is considered. In order to illustrate the concept, a simple example will be analyzed. Consider a channel system that has two sets of channels running more or less perpendicular to each other as depicted in Figure 6.

![Figure 6: A fluvial system where the rivers run approximately perpendicular, one set running E-W and one running N-S.](image)

At first glance, the structure of the system could now be decomposed with respect to the orientation of the channels. A simple case would be to separate the E-W component from the N-S component. This can be done using only the curvelet coefficients corresponding to a limited set of angles, which in this case was chosen to be the [-45,45] degrees for the N-S filter and [45,135] degrees for the E-W. The angles refer to the trend along the width of the individual curvelets (see Figure 2) and the results of these two filters are shown in Figure 7.

![Figure 7: An angular scale dependent filter based on curvelet transform.](image)
This particular result was obtained from a filter that not only removed curvelet coefficients depending on orientation but also on scale. At the fine scale level, the curvelets tend to delineate sharp features regardless of orientation of the feature. This produced slight contours of the unwanted channels. To remove this, the finest third of the scales were removed from the reconstruction images.

One important note is that the channel system in merely a training image with logical values, making the channels easy to spot as they have sharp discontinuities along their edges. The above analysis is just presented to illustrate the potential of the technique. Further investigation should be conducted in order to assess the application of this technique on a real reservoir.

5. Conclusions

By separating the frequency spectrum the observer is allowed to separate a multi-scale image into many isolated single scaled images. This offers a new ability to differentiate larger features like faults from smaller fractures and noise. Although the curvelets do a decent job in separating the features in the image, the number of scales used, \( j \), was arbitrary. If the number of scales increases, the original image could still be divided in three, but in this case each band will be a linear combination of a number of scales.

As the offset does affect the result, there might be a potential for performing the curvelet transform on the cube before stacking the offsets. This means that multiple cubes must be handled simultaneously, lending themselves to the idea of a 4D curvelet transform, which has yet to be developed.

Furthermore, since curvelets are geometric, the original image can be filtered such that only certain directions of the curvelets are used to reconstruct the image. This might be desirable if the data includes a dissonance of multidirectional features.

Another interesting idea is the curvelet's potential with geostatistical simulation algorithms like FILTERSIM. At the moment the filters used are just directional derivatives of the image. With the FDCT there might be some valuable information in the frequency spectrum.

Acknowledgements

We would like to thank the team behind Curvelab (www.curvelab.org) for producing excellent and easy-to-use software. We thank Tanima Dutta for help with the data. We gratefully acknowledge support from the sponsors of the Stanford Center for reservoir Forecasting.
References


