Churning seismic attributes with principal component analysis

Satinder Chopra[+] and Kurt J. Marfurt[†]*

[+]Arcis Seismic Solutions, Calgary; [†]The University of Oklahoma, Norman

Downloaded 12/02/14 to 205.196.179.237. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Summary

Seismic attributes are an invaluable aid in the interpretation of seismic data. Different attributes are derived for different purposes. For example there are discontinuity attributes for fault interpretation, impedance and AVO-derived attributes for lithology interpretation, spectral decomposition frequency volumes to quantify tuning effects and help identify hydrocarbons, and many others. Extracting all the potential information hidden in the seismic data using a single attribute almost never occurs. Therefore a combination of attributes or multiattribute analysis is carried out to gauge more information overall than what is possible with any one attribute

.

When attempting multiattribute analysis usually attributes of a similar kind are used, i.e. coherence and curvature attributes for fault interpretation, or impedance, lambda-rho, mu-rho, or other similar kind of attributes for lithology or fluid interpretation. However, in doing so we may be limiting ourselves to a subset of the information, say structural vs. stratigraphic. There are several ways of combining multiple attributes, with visualization in RGB and HLS color space coupled with transparency being one of the more powerful means. Unfortunately, such color display is limited to three and with transparency four attributes. One of the methods commonly used for this purpose is principal component analysis, which essentially 'churns' the different attributes and yields one or two volumes that represent the maximum variation in the input attributes. Such analysis reduces redundancy but projects the interpreter into mathematical vs. physical space, such that the resulting images can be difficult to understand. We present the results of our investigation into the combination of attributes that should be used for such an analysis.

Introduction

Principal component analysis (PCA) is a useful statistical technique that has found many applications including image compression and pattern recognition in data of high dimensionality. We are familiar with the usual statistical measures like mean, standard deviation and variance, which are essentially one-dimensional. Such measures are calculated one attribute at a time with the assumption that each attribute is independent of the others. In reality, many of our attributes are coupled through the underlying geology, such that a fault may give rise to lateral changes in waveform, dip, peak frequency, and amplitude. Less desirably, many of our attributes are coupled mathematically, such as alternative measures of coherence (Barnes, 2007) or of a suite of closely spaced spectral components. The amount of attribute redundancy is measured by the covariance matrix. The first step in multiattribute analysis is to subtract the mean of each attribute from the corresponding attribute volume. If the attributes have radically different units of measure, such as frequency measured in Hz, envelope measured in mV, and coherence without dimension, a Z-score normalization is required. The element C_{mn} of an N by N covariance matrix is then simply the cross-correlation between the m^{th} and n^{th} scaled attribute over the volume of interest. Mathematically, the number of linearly independent attributes is defined by the value of eigenvalues and eigenvectors of the covariance matrix. The first eigenvector is a linear combination that represents the most variability in the scaled attributes. The corresponding first eigenvector represents the amount of variability represented. Commonly, each eigenvalue is normalized by the sum of all the eigenvalues, giving us a percentage of the variability represented.

Covariance matrices are routinely used by marketers to identify trends. For example, if you just order 24 baby bottles online, the good people at google will soon present you with a pop-up panel to buy baby diapers and adult sleep aids, fleshing out the first eigenvector of a new parent. While the computation of multiattribute eigenvectors and eigenvalues does not provide the same physical insight into the data as multiattribute display, it does reduce the number of attributes used for subsequent analysis. For this reason principal component analysis is the first step in "fancier" clustering techniques such as self-organizing maps, generative -topographic maps, and support-vector machine analysis.

By convention, the first step is to order the eigenvalues from the highest to the lowest. The eigenvector with the highest eigenvalue is the principal component of the data set (PC1); it represents the vector with maximum variance in the data and also represents the bulk of the information that would be common in the attributes used. The eigenvector with the second-highest eigenvalue, called the second principal component, exhibits lower variance and is orthogonal to PC1. PC1 and PC2 will lie in the plane that represents the plane of the data points. Similarly, the third principal component (PC3) will lie in a plane orthogonal to the plane of the first two principal components. In the case of N truly random attributes, each eigenvalue would be identical, and equal to $1/N$. Since seismic attributes are correlated through the underlying geology and the band limitations of the source wavelet, the first two or three principal components will almost always represent the vast majority of the data variability.

Saleh and de Bruin (2000) demonstrated the extraction of AVO attributes from distorted offset-dependent amplitudes, and went on to show that these attributes were more robust displaying improved ability to identify fluid effects.

Tingdahl and Hemstra (2003) discussed the estimation of fault orientation using principal component analysis on seismic attributes such as dip, azimuth, coherence, or meta-attributes derived from the others. All the considered attributes have a common property in that they have high values at the position of the faults and exhibit low vales elsewhere.

© 2014 SEG
DOI http://dx.doi.org/10.1190/segam2014-0235.1

Downloaded 12/02/14 to 205.196.179.237. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Singh (2007) discussed the application of principal component analysis on AVO-derived attributes for lithofacies discrimination and fluid detection. In particular this study found that PC2 was a robust discriminator of lithofacies in comparison with other attributes such as acoustic impedance, shear impedance and other attributes derived from a combination of these two attributes.

Guo et al. (2009) compute 86 spectral components ranging from 5 to 90 Hz using a matching pursuit technique described by Liu and Marfurt (2007). Next PCA was performed on the 86 spectral components by forming an 86x86 covariance matrix. Thereafter, the covariance matrix is decomposed into 86 eigenvalue-eigenvector pairs. They found that the first 3 components account for most of the spectral variance seen along the horizon of interest, with the remaining components accounting for about 17 percent of the data variance. This way the dimensionality reduction was brought down from 86 to 3.

The choice of the attributes that are selected for the principal component analysis would of course be dependent on the goal of the exercise that has to be performed. For the applications mentioned above, the work carried out by Tingdahl and Hemstra (2003) focused on the fault orientations, and so the choice of the attributes input into the exercise were all that had some kind of definition of the faults. The facies determination exercise by Singh (2007) used the AVO attributes, but since the fluid discrimination was the objective, the PC2 component was the component showing variability above the background represented by PC1. Similarly, the spectral components analysis done by Guo et al. (2009) uses the frequency volumes generated by spectral decomposition of the seismic data to generate eigenspectra (spectra that best represent the response within the zone of interest). Interestingly, sometimes we see such analysis carried out on disparate input datasets, comprising the discontinuity attributes and the lithology attributes. There might be merit in arguing that the two separate classes of attributes would bring different characteristics into the 'pot ingredients' that could be churned beneficially. This is the motivation for the present work.

We have tried to compare the first (PC1) and the second (PC2) principal component displays using a varied mix of the input attributes, comprising the following attributes amongst others.

Discontinuity attributes: Coherence and curvature attributes are commonly used for interpreting faults, fractures, reef edges, channels, etc. (Chopra and Marfurt, 2007). Coherence, most-positive curvature (long-wavelength), most-negative curvature (long-wavelength), most-positive curvature (short-wavelength), most-negative curvature (short-wavelength), are the commonly used discontinuity attributes.

Sweetness attribute: Is usually considered a sand/shale indicator. Itself a "meta-attribute" sweetness is defined as the amplitude envelope divided by the instantaneous frequency, the sweetness attribute exhibits high values of sweetness for hydrocarbon-saturated sands. This would be because the presence of hydrocarbons enhances the amplitude envelope and lowers the frequency somewhat in Tertiary basins (Hart, 2008).

GLCM texture attributes: (energy, entropy and homogeneity): GLCM (grey-level co-occurrence matrix) texture attributes are useful for the determination of seismic facies analysis. GLCM energy is a measure of textural uniformity in an image, GLCM entropy is a measure of disorder or complexity of the image, and GLCM homogeneity is a measure of the overall smoothness of the image. More information on these attributes can be found in Chopra and Alexeev (2005) and Yenugu et al. (2010).

Spectral decomposition: frequency attributes (spectral magnitude components, peak frequency, peak magnitude). Spectral decomposition refers to the transformation of seismic data into individual frequency components within the seismic bandwidth. The derived frequency data has found application for the interpretation of bed thickness, discontinuities and distinguishing fluids in the reservoirs (Partyka et al.,1999).

The dataset chosen for this exercise is from central Alberta, Canada. We focus on the Mannville channels at a depth of z=1150 to 1230 m that are filled with interbedded units of shale and sandstone. On the 3D seismic volume, these channels show up at a mean time of t=1000 ms plus or minus 50 ms. In Figure 1 we show the coherence attribute display depicting the channel edges. We would like to populate these channels with better definition, showing internal variation within their bounds. For this we begin with the principal analysis (PCA) on one kind of attributes. In Figure 2, we show the PCA performed on discontinuity attributes and find that the 'Sobel on coherence' attribute dominates and that the resulting displays depict better definition of the channels. In Figure 3 we show PCA performed on lithology attributes and notice that GLCM energy attribute has a higher contribution and exhibits higher values within the channels. In Figure 4 we show the spectral decomposition attributes and notice that the dominant attribute is the 60 Hz frequency volume.

Having done this we bring together all the attributes which have shown higher contribution in the previous three steps. We show the results in Figure 5. Notice again the 'Sobel on coherence' dominates, but if this attribute is taken out of the attribute mix and replaced with coherence attribute, we notice a more balanced contribution from the first two attributes (Figure 6).

Conclusions

Principal component analysis carried out on a similar kind of attributes is a good dimensionality reduction tool. When performed on discontinuity attributes, one kind of geologic feature dominates PCA1, with PCA2 and PCA3 enhancing artifacts associated with numerical differences in the computation rather than geology. We therefore hypothesize that one may wish to use a more balanced mix of attributes in the pot, so that PC2 and PC3 yield additional geologic insight. Similar is the case with lithology and spectral decomposition attributes.

While the data reduction in principal component analysis is powerful, the results are always computed in mathematical space. Principal components of spectral components will be some kind of spectra with measurement of magnitude as a function of frequency. Linear components of attributes having different units of measure are much harder to visualize, though combinations of low peak frequency, strong negative amplitude at near offset, a negative-amplitude gradient, and a high coherence are a well-known hydrocarbon indicator.

© 2014 SEG

DOI http://dx.doi.org/10.1190/segam2014-0235.1

Principal components are routinely used in facial recognition and monetary transfers to identify patterns. In this application, the most important eigenvectors are computed to represent the bulk of all faces or monetary transactions. Then the attribute vector of interest (a suspected criminal or in our case, a good well) are then projected against this smaller number of eigenvectors. This small number of principal component coefficients are then compared to the smaller number of principal component projection volumes.

Acknowledgements

We wish to thank Arcis Seismic Solutions, TGS, for encouraging this work and for permission to present these results.

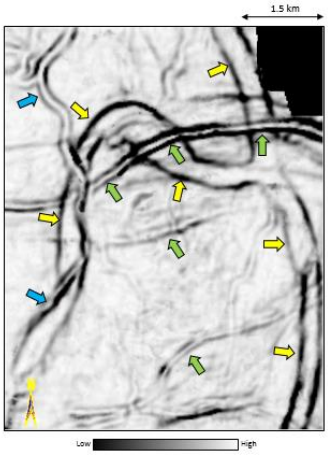


Figure 1: Strat-slice at a level close to a horizon picked at 1020ms from a coherence volume showing different channels. Some can be seen clearly (yellow arrows), other are not so clear (blue and green arrows).

Table 1

Attributes used: coherence[1], Sobel on coherence[2], most-pos curv (long-wave)[3], most-neg curv (long-wave), most-pos curv (short-wave), most-neg curv (short-wav)							
	Eigenvalue	Eigenvector					
PCA	Percentage	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5	Attribute 6
PCA 1	86.03307	-0.07927	-0.99673	-0.01401	-0.02564	0.02609	0.02209
PCA 2	8.81348	0.04594	0.00605	0.41212	0.23798	0.79240	0.37879
PCA 3	2.47449	-0.75469	0.04918	-0.44904	-0.29689	0.29158	0.16877

Figure 2: Results of principal component analysis carried out on the seismic discontinuity attributes shown in Table 1. The relative contribution of the first three principal components (PC1, PC2 and PC3) are indicated therein, with the attributes marked with superscripts, represents 97% of the variability in the six attributes.

A horizon slice at a level close to 1020ms from (a) PC1, and (b) PC2 volumes are shown. Notice, that the contribution from the attribute 'Sobel on coherence' is high (86%) and dominates the display as well.

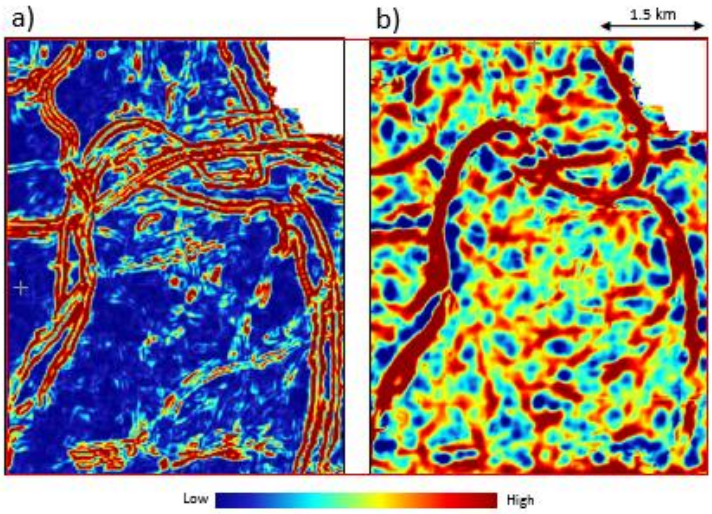
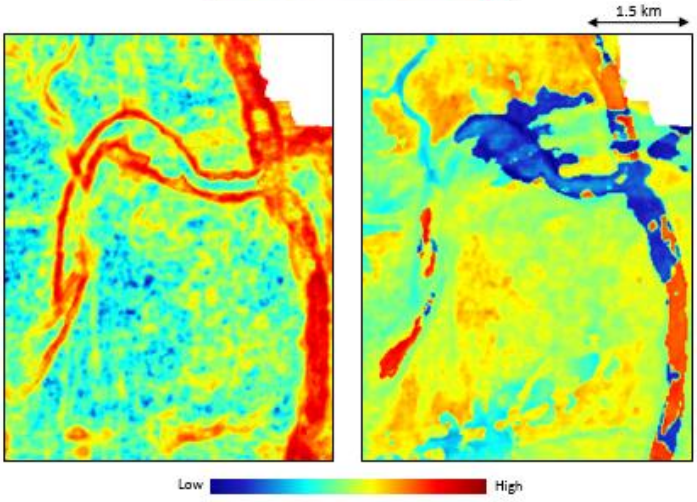


Table 2

Attributes used: Sweetness, GLCM_energy, GLCM_entropy[1], GLCM_homog, peak-freq[2], peak-mag[3]							
	Eigenvalue	Eigenvector					
PCA	Percentage	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5	Attribute 6
PCA 1	53.14759	-0.31010	-0.38292	0.62950	-0.56768	-0.19226	-0.04179
PCA 2	27.89072	0.03073	0.08388	-0.13044	0.10973	-0.97193	-0.12815
PCA 3	10.99169	-0.29869	0.08871	-0.09185	0.02378	0.13171	-0.93628

Figure 3: Results of principal component analysis carried out on the seismic lithology attributes shown in Table 2. The relative contribution of the first three principal components (PC1, PC2 and PC3) are indicated therein, with the attributes marked with superscripts, represents 92% of the variability in the six attributes.

A horizon slice at a level close to 1020ms from (a) PC1, and (b) PC2 volumes are shown. Notice, that the contribution from the attribute 'GLCM_entropy' is somewhat higher (53%) that the other two.



Downloaded 12/02/14 to 205.196.179.237. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

© 2014 SEG

DOI http://dx.doi.org/10.1190/segam2014-0235.1

Downloaded 12/02/14 to 205.196.179.237. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Table 3

Attributes used: Frequency volumes at 50[1] Hz, 55Hz, 60[1] Hz, 65 Hz, 70 Hz, 75[1] Hz.							
	Eigenvalue	Eigenvector					
PCA	Percentage	Attribute 1	Attribute 2	Attribute 3	Attribute 4	Attribute 5	Attribute 6
PCA 1	85.90909	-0.37823	-0.42890	-0.44522	-0.44167	-0.40621	-0.31864
PCA 2	12.54183	0.65245	0.36318	-0.64843	-0.22208	-0.49722	-0.47427
PCA 3	1.38579	-0.54339	0.22008	0.49734	0.30095	-0.54383	-0.54362

Figure 4: Results of principal component analysis carried out on the spectral decomposition attributes shown in Table 3. The relative contribution of the first three principal components (PC1, PC2 and PC3) are indicated therein, with the attributes marked with superscripts, represents 99% of the variability in the six attributes.

A horizon slice at a level close to 1020ms from (a) PC1, and (b) PC2 volumes are shown. Notice, that the contribution from the attribute 'Frequency volume 60 Hz' is way higher (86%) that the other two.

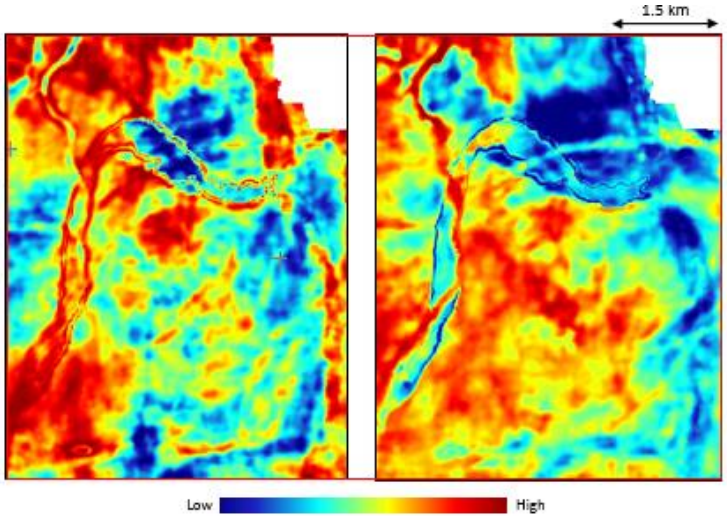


Table 4

Attributes used: Sobel on coherence[1], GLCM_entropy[2], Frequency vol. (60Hz)[3]				
	Eigenvalue	Eigenvector		
PCA	Percentage	Attribute 1	Attribute 2	Attribute 3
PCA 1	76.32059	-0.24025	0.96987	-0.04030
PCA 2	12.09813	0.95684	0.22962	-0.17854
PCA 3	11.58129	0.16352	0.08136	0.98308

Figure 5: Results of principal component analysis carried out on the spectral decomposition attributes shown in Table 4. The relative contribution of the first three principal components (PC1, PC2 and PC3) are indicated therein, with the attributes marked with superscripts, represents 100% of the variability in the three attributes, as these were the three input attributes.

A horizon slice at a level close to 1020ms from (a) PC1, and (b) PC2 volumes are shown. Notice, that the contribution from the attribute 'Sobel on coherence' is way higher (76%) that the other two, and the display in (a) is consequently dominated.

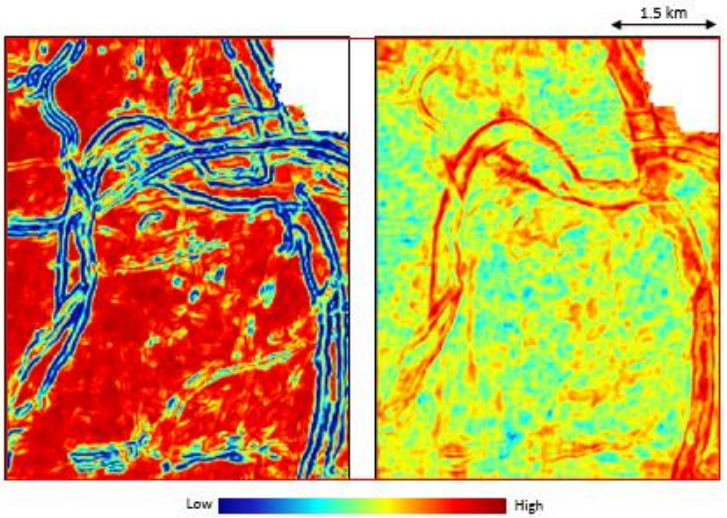
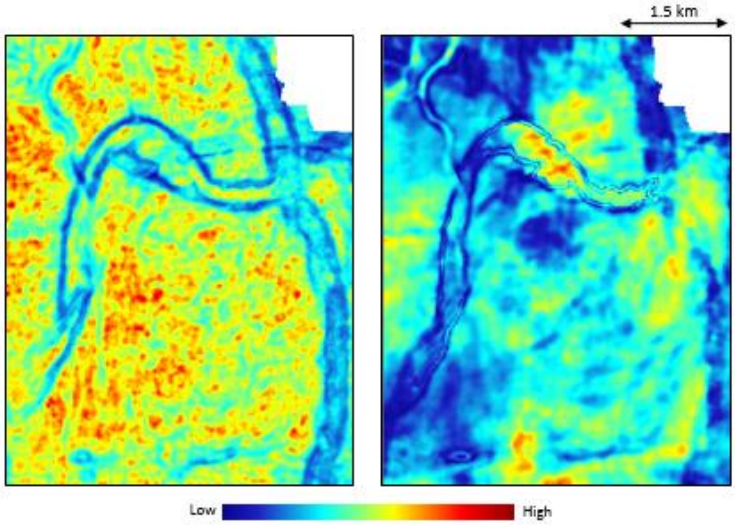


Table 5

Attributes used: Coherence[1], GLCM_entropy[2], Frequency vol. (60Hz)[3]				
	Eigenvalue	Eigenvector		
PCA	Percentage	Attribute 1	Attribute 2	Attribute 3
PCA 1	54.56645	-0.98098	-0.12595	-0.14787
PCA 2	39.24743	-0.12685	0.99192	0.00333
PCA 3	6.18613	-0.14690	-0.01546	-0.98903

Figure 6: Results of principal component analysis carried out on the spectral decomposition attributes shown in Table 5. The relative contribution of the first three principal components (PC1, PC2 and PC3) are indicated therein, with the attributes marked with superscripts, represents 100% of the variability in the three attributes, as these were the three input attributes.

A horizon slice at a level close to 1020ms from (a) PC1, and (b) PC2 volumes are shown. Notice, that the contribution from the attribute 'GLCM_entropy' is higher (54.5%) that the other two.



© 2014 SEG

DOI http://dx.doi.org/10.1190/segam2014-0235.1

http://dx.doi.org/10.1190/segam2014-0235.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Barnes, A. E., 2007, Redundant and useless seismic attributes: Geophysics, **72**, no. 3, P33–P38.

Chopra, S., and V. Alexeev, 2005, Application of texture attribute analysis to 3D seismic data: CSEG Recorder, **30,** no. 7, 28–32, http://209.91.124.56/publications/recorder/2005/09sep/sep05-texture-attribute-analysis.pdf.

Chopra, S., and K. J. Marfurt, 2007, Seismic attributes for prospect identification and reservoir characterization: SEG.

Guo, H. K., K. J. Marfurt, and J. Liu, 2009, Principal component spectral analysis: Geophysics, **74**, no. 4, P35–P43, http://dx.doi.org/10.1190/1.3119264.

Hart, B. S., 2008, Channel detection in 3D seismic data using sweetness: AAPG Bulletin, **92**, no. 6, 733–742, http://dx.doi.org/10.1306/02050807127.

Liu, J. L., and K. J. Marfurt, 2007, Multi-color display of spectral attributes: The Leading Edge, **26**, 268–271, http://dx.doi.org/10.1190/1.2715047.

Partyka, G. A., J. M. Gridley, and J. Lopez, 1999, Interpretational applications of spectral decomposition in reservoir characterization: The Leading Edge, **18**, 353–360, http://dx.doi.org/10.1190/1.1438295.

Saleh, S. J., and J. A. Bruin, 2000, AVO attribute extraction via principal component analysis: 70th Annual International Meeting, SEG, Expanded Abstracts, 126–129.

Singh, Y., 2007, Lithofacies detection through simultaneous inversion and principal component attributes: The Leading Edge, **26**, 1568–1575, http://dx.doi.org/10.1190/1.2821944.

Tingdahl, K., and N. Hemstra, 2003, Estimating fault-attribute orientation with gradient analysis, principal component analysis and the localized Hough-transform: 73rd Annual International Meeting, SEG, Expanded Abstracts, 358–361.

Yenugu, M. K., K. J. Marfurt, and S. Matson, 2010, Seismic texture analysis for reservoir prediction and characterization: The Leading Edge, **29**, 1116–1121, http://dx.doi.org/10.1190/1.3485772.

Downloaded 12/02/14 to 205.196.179.237. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/