GEOPHYSICS, VOL. 85, NO. 2 (MARCH-APRIL 2020); P. O17–O35, 25 FIGS.
10.1190/GEO2019-0223.1

Seismic attribute selection for machine-learning-based facies analysis

Jie Qi[1], Bo Zhang[2], Bin Lyu[1], and Kurt Marfurt[1]

ABSTRACT

Interpreters face two main challenges in seismic facies analysis. The first challenge is to define, or "label," the facies of interest. The second challenge is to select a suite of attributes that can differentiate a target facies from the background reflectivity. Our key objective is to determine which seismic attributes can best differentiate one class of chaotic seismic facies from another using modern machine-learning technology. Although simple 1D histograms provide a list of candidate attributes, they do not provide insight into the optimum number or combination of attributes. To address this limitation, we have conducted an exhaustive search whereby we represent the target and background training facies by high-dimensional Gaussian mixture models (GMMs) for each potential attribute combination. The first step is to choose candidate attributes that may be able to differentiate chaotic mass-transport deposits and salt diapirs from the more conformal, coherent background reflectors. The second step is to draw polygons around the target and background facies to provide the labeled data to be represented by GMMs. Maximizing the distance between all GMM facies pairs provides the optimum number and combination of attributes. We use generative topographic mapping to represent the high-dimensional attribute data by a lower dimensional 2D manifold. Each labeled facies provides a probability density function on the manifold that can be compared to the probability density function of each voxel, providing the likelihood that a given voxel is a member of each of the facies. Our first example maps chaotic seismic facies associated with the development of salt diapirs and minibasins. Our second example successfully delineates karst collapse underlying a shale resource play from north Texas.

INTRODUCTION

Machine learning for seismic facies classification is often described as being a "pattern recognition" problem, in which the objective is to differentiate the pattern exhibited by a target facies from the background seismic reflectivity pattern. Less well recognized is that such facies classification consists of two different pattern recognition problems. For concreteness, let us consider the example of a relatively thick package of marine shale. The first types of patterns are those readily seen by the human interpreter, in which the shale package may be described as being low amplitude, conformal, broadband, and coherent. These patterns are measured by seismic attributes such as the rms amplitude, reflector parallelism, spectral components, and coherence, respectively. The second types of patterns are those measured by machine-learning algorithms whose input consists of multiple seismic attribute volumes. Two attribute volumes can be crossplotted, whereas a third attribute can be assigned to a color in the crossplot. Analysis of more than three attribute volumes requires dimensionality reduction to allow human interaction. The more popular dimensionality reduction techniques include principal component analysis (PCA), independent component analysis (ICA), self-organizing mapping (SOM), and generative topographic mapping (GTM) (Qi and Castagna, 2013; Roy et al., 2014; Roden et al., 2015; Zhao et al., 2018; Lubo-Robles and Marfurt, 2019), in which the projections used in the latter two mapping algorithms are "learned" from the data and form what are called unsupervised machine-learning algorithms.

As noted by Nicolaus Steno (e.g., Cutler, 2003), most sedimentary rocks are deposited layer by layer, such that the bulk of the seismic sedimentary section consists of piecewise conformal, coherent reflections. In contrast, chaotic facies are less common; therefore, they provide a good candidate for seismic facies classification. In theory, the interior of a perfectly homogeneous salt dome should be reflection free. For all but the most carefully acquired and processed 3D

Manuscript received by the Editor 16 April 2019; revised manuscript received 29 October 2019; published ahead of production 24 November 2019; published online 9 January 2020.
[1]The University of Oklahoma, School of Geosciences, Norman, Oklahoma, USA. E-mail: jie.qi@ou.edu; bin.lyu@ou.edu; kmarfurt@ou.edu.
[2]The University of Alabama, Department of Geological Sciences, Tuscaloosa, Alabama, USA. E-mail: bzhang33@ua.edu (corresponding author).

seismic surveys, the interior of a salt dome consists of random noise whose basis is a mixture of inaccurately imaged multiples, converted waves, and head waves. Karst collapse features are also often described as chaotic, with roof collapse over a subterranean void resulting in rotated, broken reflectors separated by random fill. Subsequent diagenetic alteration may further reduce any remaining coherent reflecting horizons. Diagenesis may also remove the internal reflection boundaries of carbonate reefs resulting in a low-amplitude, chaotic seismic response, which Bubb and Hatledid (1977) identify as one of several carbonate buildup indicators. The core of volcanic cones that may consist of multiple pipes and dikes are also chaotic, resulting in an image that is easy to confuse with those of carbonate buildups (Infante-Paez and Marfurt, 2018). Mud volcanoes and shale diapirs often appear as chaotic, with shale diapirs often containing internal blocks of coherent reflectors (Haskell et al., 1999). Mass transport deposits (MTDs) are somewhat more heterogeneous. For well-imaged surveys, the toe of an MTD may look quite chaotic, whereas near the updip end, the MTD may appear as a suite of well-imaged, rotated blocks (Moscardelli and Wood, 2007). On less-well-imaged seismic data, the entire MTD may appear to be chaotic. Meldahl et al. (1999) provide one of the earliest successes in seismic chaotic facies analysis using attributes and machine learning to delineate gas chimneys, where the chaotic pattern is interpreted to be due to either absorption of energy or to decreases in velocity and thus inaccurate imaging, both of which are attributed to the presence of gas in the system.

Interpreters face two main challenges in seismic facies analysis. The first challenge is for a human interpreter to define, or label, the facies of interest. Accurately defining the 3D extent of a given seismic facies takes an understanding of geologic processes and the limits of seismic acquisition, processing, and imaging. Machine learning is based on accurate training data, which in this application is provided by a skilled interpreter defining polygons about facies of interest. The second challenge is to select a suite of attributes that can differentiate a target facies from the background reflectivity. Unfortunately, there are relatively few interpreters who possess a deep understanding of the geology of a given exploration play and a deep understanding of the sensitivity of an ever-expanding collection of seismic attributes to geology.

One partial solution in determining "which attributes work best" is to create a table of images that can be searched, showing the response of each attribute to a given geologic feature. Although useful, the seismic attribute response is a function not only of the geology but also of the seismic data quality, where the interior of a salt dome in the Gulf of Mexico will exhibit one pattern on older, narrow-azimuth time-migrated data and a different pattern on more modern, wide-azimuth depth-migrated data. We therefore propose using a more statistical attribute selection workflow that if successful, could be used most seismic interpreters.

Hampson et al. (2001) use attributes to predict well-log properties such as porosity using a workflow called stepwise multilinear regression. The first step is to compute linear correlations between the desired property and attributes extracted about the well, introducing a certain amount of nonlinearity by creating additional attribute volumes that are squares, square roots, reciprocals, or logarithms of the original candidate attributes. The quality of each attribute is evaluated not only by its correlation but also by its error in predicting values not used in the original correlation. Once ranked by their correlation coefficient, the second step is to perform a multilinear regression of the two best attributes. If the prediction of the log property used for validation improves, we add the third most correlated attribute, continuing the process until the validation shows an increase in error. In Hampson et al.'s (2001) workflow, these n best attributes are then provided to a neural network that generates a nonlinear correlation between the n attributes and the desired log property.

A potential shortcoming in this workflow is that when selecting the two best of n attributes, attributes 2 and 3, may provide a better fit than more highly ranked attributes 1 and 2. This observation suggests a more exhaustive search of all $n!/(n-2)!2!$ two-attribute combinations, validating each combination against data not used in the regressions. The two-attribute test would then be followed by $n!/(n-3)!3!$ three-attribute tests. If one of the three-attribute combinations is better than the two-attribute combinations, we examine four-attribute combinations, and so on, until the validation error increases.

The objective of this paper is to determine which seismic attributes are most useful for the delineation of chaotic seismic facies using modern machine-learning technology. We replace the simple 1D histogram attribute selection criteria used by Qi et al. (2016) to discriminate among salt domes, mass transport complexes, and a background of conformal reflectors with multidimensional histograms represented by Gaussian mixture models (GMMs). We begin with a summary of those attributes that we think are candi-

Candidate attribute for chaotic facies	Salt	Mass transport deposition (MTD)	Karst collapse	Conformal sedimentary background
Reflector dip	Nonparallel	Piecewise parallel, locally subparallel bedding	Parallel at edges, nonparallel at filled sinkholes	Parallel to subparallel
Reflector amplitude	Low reflectivity (except for noise)	Low to high based on impedance contrast within rotated blocks, generally low outside well-imaged blocks	Low to moderate, based on impedance contrast	Low to high based on impedance contrast
Spectral response	Low to moderate spectral response	Variable frequency within rotated blocks; Low in chaotic zones	Broadband response	Broadband response
Texture	High entropy, low homogeneity	Low entropy within rotated blocks; high entropy outside rotated blocks	Moderate entropy, moderate homogeneity	Low entropy, high homogeneity

Figure 1. Table of dip, amplitude, frequency, and texture attribute responses to three chaotic facies as well as to a conformal reflector background. An experienced interpreter would use one or more of these attributes to visually distinguish one facies from another. For this reason, these attributes serve as good candidates for machine-learning-based multiattribute facies classification.

dates able to differentiate chaotic features from the more conformal, coherent background reflectors. Then, we propose a GMM-based attribute selection workflow to evaluate attributes by their histogram models associated with different labeled facies. Next, we construct and then compare GMM clusters of a subset of interpreted (labeled) seismic lines and time slices to determine which of the candidate attributes best differentiate the target chaotic facies from the more conformal, coherent background. Given this suite of attributes, we perform an unsupervised GTM classification, projecting the labeled facies onto it to map our interpretation to the GTM manifold, which in turn can be used to label the entire seismic data volume. Our first example is that of mapping chaotic seismic facies representative of salt. We repeat this workflow for our second example of karst collapse.

SEISMIC ATTRIBUTES

For a chaotic facies, the candidate attributes for computer-assisted seismic facies analysis should be able to measure features such as reflector dip, amplitude or energy, spectral response, and internal textures. Chaotic facies in a given seismic volume usually exhibit similar but not identical attribute responses, allowing a skilled interpreter to differentiate mass transport depositions from turbidities from salt diapirs. We summarize the attribute response of three chaotic facies that are often seen in the Gulf of Mexico and the Fort Worth Basin, Texas (Figure 1).

For machine-learning-based facies analysis, gray-level cooccurrence matrix (GLCM) texture attributes are often quite useful in differentiating channels, salt domes, mass transport complexes, and karst collapse features (Matos et al., 2011; Qi et al., 2016; Di and Gao, 2017; de Lima and Marfurt, 2018; Zhao et al., 2018). GLCM texture attributes statistically measure the lateral change in reflectivity. Compared to geometric attributes such as curvature and aberrancy, the GLCM texture attributes play a more significant role in facies classification, in part because they are more sensitive to heterogeneities in the sedimentary package rather than to a larger scale subsequent tectonic overprint. There are several kinds of GLCM attributes: Some of them are "redundant" or even inferior to coherence, energy, chaos, and other more commonly used attributes (Barnes, 2007, 2016; Marfurt, 2018). For example, GLCM

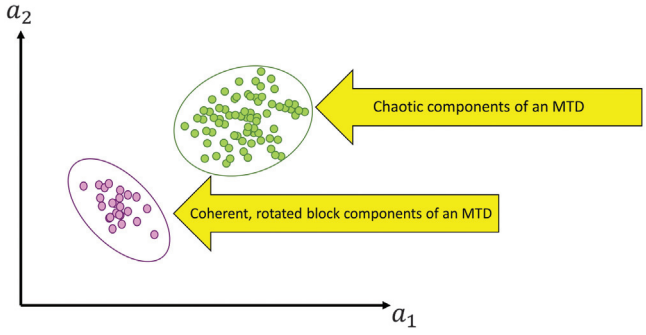

Figure 3. A cartoon showing the necessity of the 2C GMM of an MTD. One Gaussian represents the chaotic components, and the other represents the coherent, rotated block component of an MTD.

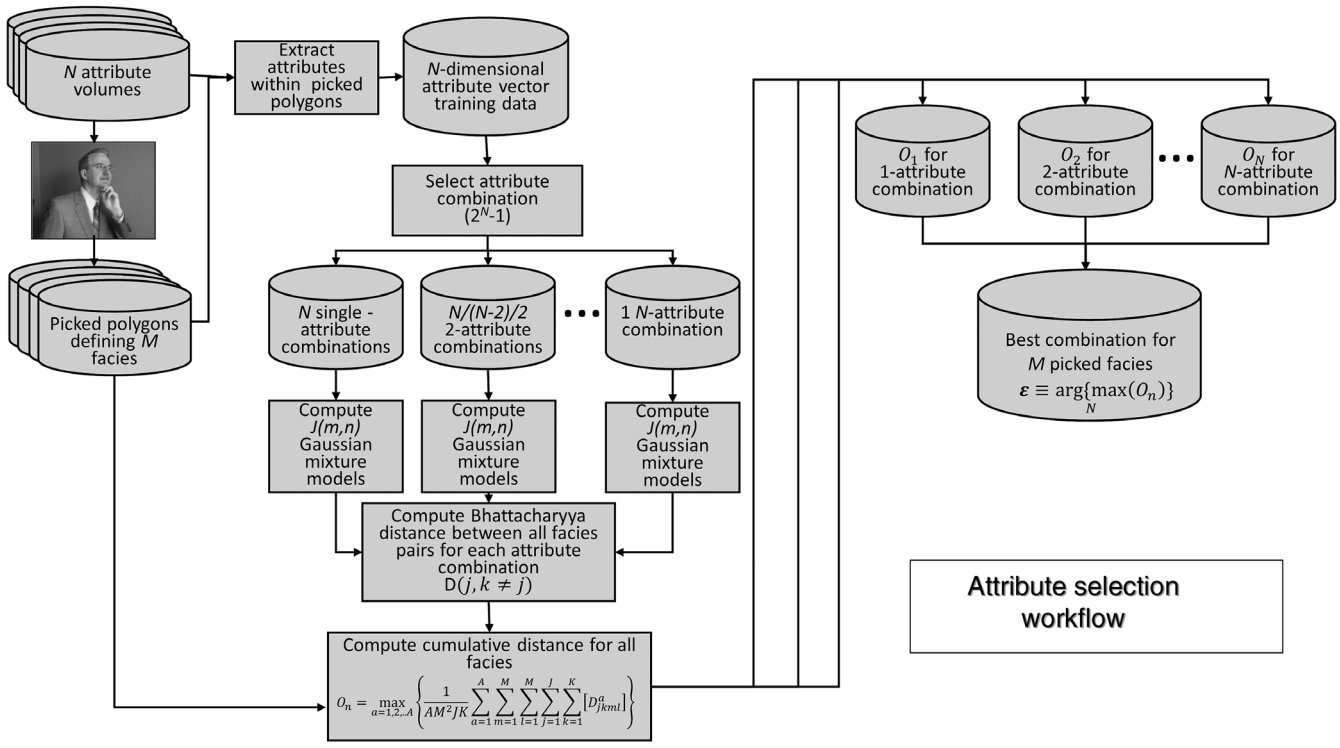

Figure 2. Workflow illustrating the steps used in our GMM-based attribute selection. Begin by defining training data by picking facies to be differentiated. Each voxel within the picked polygons provides an N-dimensional attribute vector. Next, select one of $2^N - 1$ possible attribute combinations and compute the GMM for each facies, and compute the Bhattacharyya distances between the GMM representations of each facies. Then, sum up the distances between all facies to compute a cumulative distance for this attribute combination. Repeat the process for all other attribute combinations. The winning attribute combination is that with the highest average cumulative distance.

entropy has a strong inverse correlation to GLCM homogeneity (Infante-Paez and Marfurt, 2019).

The most useful statistical measures of the seismic spectrum include the peak amplitude, peak frequency, spectral mean, spectral skewness, spectral slope, and spectral bandwidth. Among these attributes, the peak amplitude and peak frequency are broadly used to estimate thin-bed tuning in seismic interpretation. The other spectral measurements are less directly tied to a specific geologic model but provide useful input to machine-learning algorithms. Just as the peak frequency is a better estimate of thin-bed tuning than the instantaneous frequency, the spectral bandwidth computed from spectral components is a more accurate measure than the instantaneous bandwidth. Using synthetic models, Zhang et al. (2008) show that the spectral slope (the trend of the spectrally balanced data at a voxel) can be correlated to upward fining and coarsening, whereas spectral roughness is a measure of the subseismic resolution heterogeneity.

Not all attributes are useful in discriminating our target facies. Others are useful to human interpreters but are not amenable to machine-learning analysis. Human interpreters commonly use the phase and cosine of phase to map discontinuities and unconformities. However, upon further consideration, what they really do is use these attributes to identify vertical and lateral changes in the phase. Such changes are more directly measured by frequency, parallelism, and discontinuity attributes. The values of phase (ranging between −180° and +180°) themselves do not provide a good discriminator. Similarly, the absolute values of dip magnitude and dip azimuth cannot discriminate between rotated fault blocks in the MTD and the steeply dipping flanks of a minibasin. Lubo-Robles et al. (2019) using an exhaustive search of attributes based on probabilistic neural networks find that the events internal to a salt dome gave rise to anomalous curvature values. We feel that a more direct measure of the chaotic orientation within a salt dome is nonparallelism.

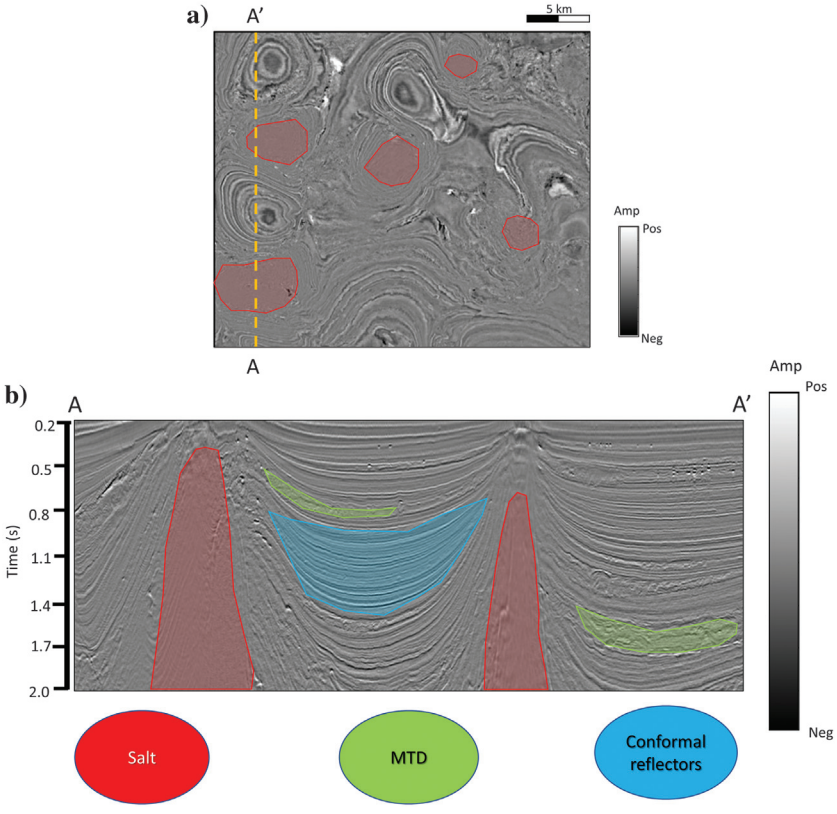

Figure 4. (a) Time slice at $t = 1.22$ s and (b) vertical slice along line AA′ through the seismic amplitude volume. The red polygons define salt diapir voxels that will be used to construct the GMM of the salt facies in multiattribute space. The green polygons define MTD facies, and the blue polygons define background conformal reflectors. Salt facies exhibit the weak envelope, low frequency, and nonparallel reflector dip. When defining training data, interpreters only pick those voxels in which they have greater confidence.

ATTRIBUTE SCALING

GMMs and GTM assume that the N-dimensional attribute vectors can be represented by a suite of N-dimensional Gaussian distributions. Clearly, a bimodal distribution can be represented by a 2C GMM. Nevertheless, many of our more useful attributes (e.g., energy, deviation of energy gradient, spectral roughness, and GLCM variance) exhibit a log-normal distribution that is rendered more Gaussian by using the logarithm of the attribute in the computation. Coherence c and GLCM entropy E are skewed toward a fixed upper limit. Scaling these two attributes to be log(1-c) and log(5-E) results in distributions that better approximate the Gaussian assumption.

ATTRIBUTE SELECTION

K-means (Forgy, 1965) is one of the earliest clustering algorithms. The interpreter hypothesizes the number of clusters K, after which K multidimensional Gaussians distributions are constructed to represent the data.

GMMs also use Gaussian distributions to represent seismic facies, but now each facies may be represented by more than one Gaussian, allowing one to represent multimodal distributions. By construction, GMMs provide a posterior probability that any particular observations (or voxel in 3D data) belong to a given mixture model. In statistics, a multivariate distribution of data vector $\mathbf{x}_i$ on the parameters α_k modeled by GMMs is

$$p(\mathbf{x}_i|\psi) = \sum_{k=1}^{K} \alpha_k g(\mathbf{x}_i|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (1)$$

where $\boldsymbol{\mu}_k$ is the mean, $\boldsymbol{\Sigma}_k$ is the covariance matrix for Gaussian distribution g, and α_k is the kth weight, where

$$\sum_{k=1}^{K} \alpha_k = 1. \quad (2)$$

The kth multidimensional Gaussian probability function is

$$g(\mathbf{x}_i|\mu_k, \boldsymbol{\Sigma}_k) = \frac{1}{|\boldsymbol{\Sigma}_k|^{\frac{1}{2}}(2\pi)^{\frac{n}{2}}} \times \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)^T\right), \quad (3)$$

where the symbol T indicates the transpose of a matrix and n is the number of dimensions (or attributes) analyzed, where $i = 1 \ldots n$. Hardisty

(2017) and Wallet and Hardisty (2019) show how to compute the optimum number of GMMs to represent multiattribute data in a seismic survey, where the objective is to determine if different seismic facies naturally clump into different areas of n-dimensional space, allowing them to be color coded and displayed.

Building on Hardisty's (2017) and Wallet and Hardisty's (2019) work, we will use K-means clustering to generate initial clustering models. We apply a maximum of 300 iterations of a refinement technique to cluster attribute vectors into an attribute space defined by the n-by-n covariance matrix. The K-means provides the initial multivariate means $\boldsymbol{\mu}_k^i$, covariance matrices $\boldsymbol{\Sigma}_k^i$, and weights α_k^i used in equation 1. If the number of components K is known, expectation maximization (EM) is able to determine the best model parameters. When the data vectors are assumed to be well-represented by a single GMM that matches the underlying assumption of GMM, classification EM (CEM) better classifies data vectors than traditional EM (Celeux and Govaert, 1995; Hardisty, 2017). In general, the EM algorithm for mixture Gaussian distributions involves K-means and GMMs. In this paper, we use CEM and stochastic EM (SEM) (Celeux and Govaert, 1995) to learn the mixture parameters $\boldsymbol{\mu}_k^i$, covariance $\boldsymbol{\Sigma}_k^i$, and α_k^i. Like the conventional EM algorithm, CEM and SEM require partitioning the input data, in which the prediction of the data not used in defining the Gaussians is used to compute the posterior probability according to the covariance matrix. Each element w_{ik} (the posterior probability) of the $n \times n$ responsibility matrix (Hathaway, 1986) $\mathbf{W}$ is given by

$$w_{ik} = \frac{\alpha_k g(\mathbf{x}_i|\mu_k, \boldsymbol{\Sigma}_k)}{p(\mathbf{x}_i|\psi)}. \tag{4}$$

CEM accumulates the responsibility matrices $\mathbf{W}$ to create K partitions by assigning each data component to the cluster that provides the highest posterior probability. For each cluster, the mixture parameters of the most likely cluster are using a maximum log-likelihood estimate

$$L(\psi) = \sum_{k=1}^{K}\sum_{i=1}^{N} z_{ik} \log\{\alpha_k g(\mathbf{x}_i|\mu_k, \boldsymbol{\Sigma}_k)\}, \tag{5}$$

where z_{ik} is an indicator that is equal to one only if the data vector $\mathbf{x}_i$ belongs to cluster k (Hardisty, 2017). The SEM algorithm randomly assigns partitions to a cluster associated with the posterior probabilities in the responsibility matrix, which helps to avoid suboptimal solutions provided by the CEM algorithm. Thus, we use SEM to initialize the CEM algorithm, which in turn provides a final partition and GMM. The parameterizations of each covariance matrix $\boldsymbol{\Sigma}_k^i$ associated with each GMM can be controlled resulting in a reduction in the number of parameters. We consider nine modules of covariance matrices introduced by Celeux and Govaert (1995) and Hardisty (2017), and then we use Schwarz's (1978) Bayesian information criterion (BIC) to compare models of differing complexity:

$$\text{BIC} = \log(L(\psi)) - \frac{1}{2} T \log(S), \tag{6}$$

where T is the number of estimated parameters and S is the number of training voxels. Higher values of BIC indicate more confidence in a given covariance parameterization.

Some insight into the attribute expression of a given facies or "what works" reduces the number of combinations to be evaluated. For each collection of attributes ($n = 2$, $n = 3$, $n = 4$, ...) we generate n-dimensional GMMs of each of the M user-defined seismic facies. We then compare models by multiplying the GMM for each facies against the others, and we sum the results. The attribute selection that provides the largest summed distance (or least overlap) is the best combination for that value of n. Then, we validate this combination by predicting facies not used in constructing the original GMMs.

Figure 2 shows our attribute selection workflow. In addition to defining N candidate attributes, the interpreter defines the M facies of interest by drawing a suite of polygons on the time and vertical slices of the seismic amplitude or attribute images, in which each

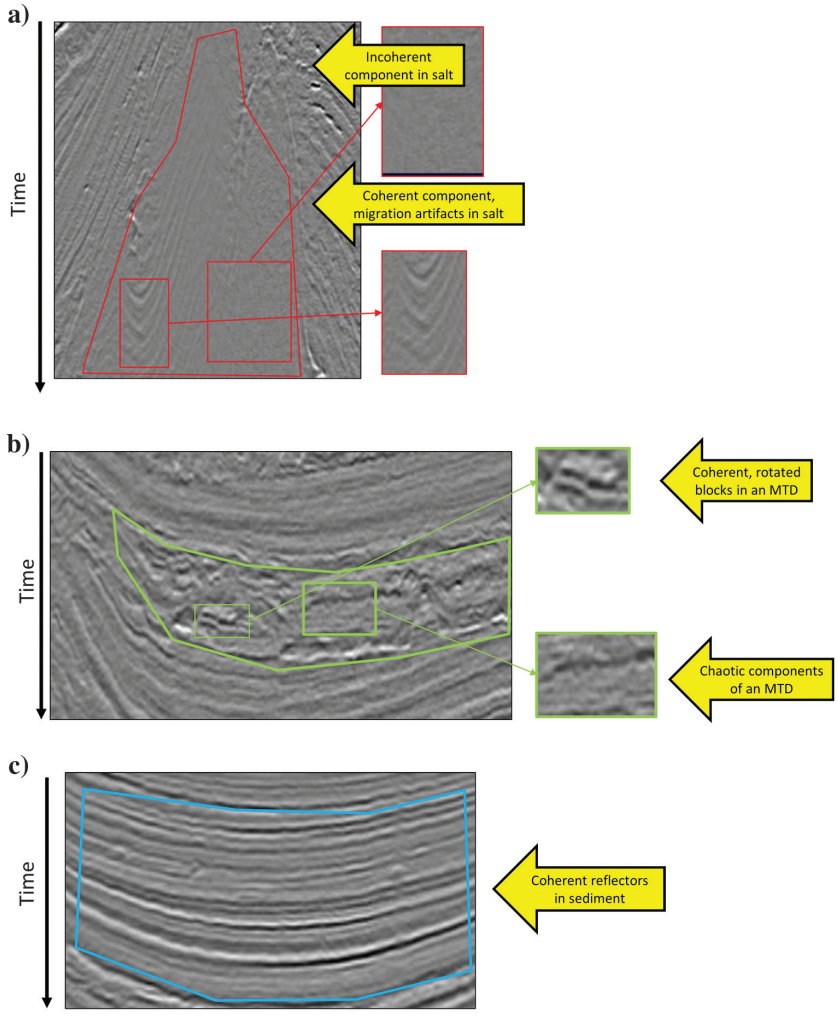

Figure 5. Magnified vertical slices from Figure 5b of picked (a) salt, (b) MTD, and (c) conformal background facies. Because salt and MTDs may contain coherent and incoherent voxels, a mixture of Gaussians (or GMM), rather than a single Gaussian is required to represent these facies.

facies can be represented by multiple polygons. The result is $N*M$ individual volumes that represent the N candidate attributes and the M facies. Note that the size of the interpreter-defined data is a small fraction of the size of the original data volume, allowing us to conduct an exhaustive search for the best attribute combination. The workflow iteratively selects a different number of n attributes from the N candidate attributes. When only considering one selected attribute, there are only N possible attribute combinations. For two selected attributes, there are $N!/(N–2)!/2!$ attribute combinations. The total number of different selected attribute combinations from the N candidate attributes will be $2^N - 1$. We compute GMMs $[\alpha_k, \mu_k, \mathbf{\Sigma}_k]$ that represent the picked voxels for each attribute combination and each facies. Skilled interpreters can accurately draw one or multiple polygons about a target facies. Facies such as MTDs and salt containing coherent noise events will require representation by more than one Gaussian (Figure 3). Numerical experimentation shows that we can limit the maximum cluster number to each GMM in our two applications.

After computing GMMs for each facies and each attribute combination, we use the Bhattacharyya distance (Mak and Barnard, 1996) to measure the similarity between each cluster pair. For two GMMs j and k residing in n-dimensional attribute space, the distance D^n_{jk} between them is

$$D^n_{jk} = \frac{1}{8}(\boldsymbol{\mu}_\mathbf{k} - \boldsymbol{\mu}_\mathbf{j})^T \left[\frac{\mathbf{\Sigma_k} + \mathbf{\Sigma_j}}{2}\right]^{-1} (\boldsymbol{\mu}_\mathbf{k} - \boldsymbol{\mu}_\mathbf{j}) + \frac{1}{2}\ln\frac{\left|\frac{\mathbf{\Sigma_k}+\mathbf{\Sigma_j}}{2}\right|}{\sqrt{|\mathbf{\Sigma_k}||\mathbf{\Sigma_j}|}}, \tag{7}$$

where $\boldsymbol{\mu}_\mathbf{k}$, $\boldsymbol{\mu}_\mathbf{j}$ and $\mathbf{\Sigma_k}$, $\mathbf{\Sigma_j}$ are the mean and covariances of the GMM clusters k and j from two different facies. Equation 7 measures the differences of size, shape, orientation, and position of the clusters in the GMM space, where the larger the value, the less overlap and the greater dissimilarity between two clusters. The similarity of all cluster pairs is measured by the Bhattacharyya distance. The Bhattacharyya distance measures similarities between clusters that reside in the same n-dimensional attribute space, and it will not be compared to GMM clusters in n - 2, $n-1$, $n+1$, $n+2$, ... dimensional attribute space. When comparing two facies GMMs with two clusters each, the Bhattacharyya distance between each cluster of one GMM and each cluster of the second GMM needs to be computed, resulting in an average Bhattacharyya distance that measures the similarity between these two GMMs. Therefore, to evaluate which dimensionality n is best, we define the average cumulative distance of each attribute combination,

$$O^n_{a_n} = \max_{A_n}\left\{\frac{1}{A_n M^2 JK}\sum_{a=1}^{A_n}\sum_{m=1}^{M}\sum_{l=1}^{M}\sum_{j=1}^{K}\sum_{k=1}^{K} \times [D^n_{mljk}] | n = 1, 2, \ldots, N\right\}, \tag{8}$$

where M is the number of the picked facies; A_n is the number of possible n-attribute combinations; K is the maximum number of clusters per facies; and the index n ranges from 1 to N, where N is the number of input candidate attributes. We determine the optimum attribute combinations in each n-selected attribute combination by comparing the value of the average cumulative distance $O^1, O^2, \ldots, O^N$. Finally, the number of clusters and attribute combination $\mathbf{r}$ for the picked M facies by

$$\mathbf{r} \equiv \arg\{\max_N(O^n_{a_n})\}, \tag{9}$$

Input attributes
1) Coherence
2) Spectral bandwidth
3) Covariance of dip and energy gradient
4) GLCM entropy
5) GLCM variance
6) Energy gradient deviation
7) Spectral roughness
8) Reflector convergence
9) Dip deviation

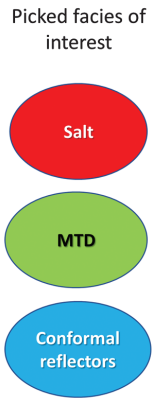

Figure 6. List of the candidate attributes and the picked facies of interest. Coherence measures the similarity between traces, which is also sensitive to strong random noise. Spectral bandwidth and spectral roughness are statistic measures of the spectra. GLCM entropy and GLCM variance are texture attributes and measure lateral variations of seismic amplitude along reflector dip. Dip deviation, energy deviation, and covariance of dip and energy gradient measure vertical and lateral changes of reflector dip, reflector energy gradient and the two together. Reflector convergence measures vertical changes of the reflector dip.

DATA CONDITIONING PRIOR TO MACHINE-LEARNING-ASSISTED SEISMIC FACIES ANALYSIS

In general, any data conditioning process that facilitates interactive interpretation also facilitates machine-learning algorithms.

Subset n of N selected attributes	Number of combinations A_n	Number of clusters J_n	Number of the comparisons
1	9	54	1434
2	36	216	23220
3	84	504	126756
4	126	756	285390
5	126	756	285390
6	84	504	126756
7	36	216	23220
8	9	54	1434
9	1	6	15

Figure 7. List of each subset n ($n = 1, 2, \ldots, N$) selected attributes and the corresponding number of combinations, number of clusters, and number of comparisons in each subgroup.

Size of attribute subset, n	1	2	3	4	5	6	7	8	9
Best attribute combination	3	1,4	1,4,6	1,2,4,6	1,2,4,5,6	1,3,4,5,8,9	1,2,4,5,6,7,9	1,2,3,4,6,7,8,9	1,2,3,4,5,6,7,8,9
Cumulative distance, O_n	9.65	9.71	14.66	15.01	15.07	15.62	16.14	15.70	10.62

Figure 8. Table showing the best n-attribute combination providing the highest average cumulative distance using equation 8. Note that the 7-attribute combination {1, 2, 4, 5, 6, 7, 9} providing the highest distance of 16.14 attribute indices are labeled in Figure 8.

Selected attributes
1) Coherence
2) Spectral bandwidth
4) GLCM entropy
5) GLCM variance
6) Energy deviation
7) Spectral roughness
9) Dip deviation

Figure 9. List of the seven attributes providing the highest cumulative distance measurement.

Common postmigration data conditioning algorithms include spectral balancing and structure-oriented filtering prior to attribute computation. Filters can also be applied to the seismic attributes themselves, where fault enhancement algorithms form more easily interpreted fault surfaces from discontinuous coherence anomalies. For 3D seismic facies analysis, conventional image processing filters include suppressing random noise, sharpening texture edges, removing unwanted features, and segmenting specific images. Low-pass and band-pass filters are some of the more common image processing techniques used to suppress noise. However, these filters also blur the image edges that otherwise are needed to separate two geologic formations. In contrast, Kuwahara et al. (1976) filtering not only removes random noise but it also preserves image edges. Luo et al. (2002) use Kuwahara filtering to sharpen edges and improve the continuity of seismic amplitude volumes. Qi et al. (2015, 2016) apply Kuwahara filtering to precondition seismic attribute volumes for subsequent machine-learning facies analysis. The smoothness, edge preservation, and noise suppression are controlled by the size of the analysis window. Recently, two generalizations including anisotropic Kuwahara filtering (Kyprianidis et al., 2009) and adaptive Kuwahara filtering (Bartyzel, 2015), have been used in image processing, both of which use the data statistics to adjust the size of the analysis window. We provide details of this new algorithm in Appendix A.

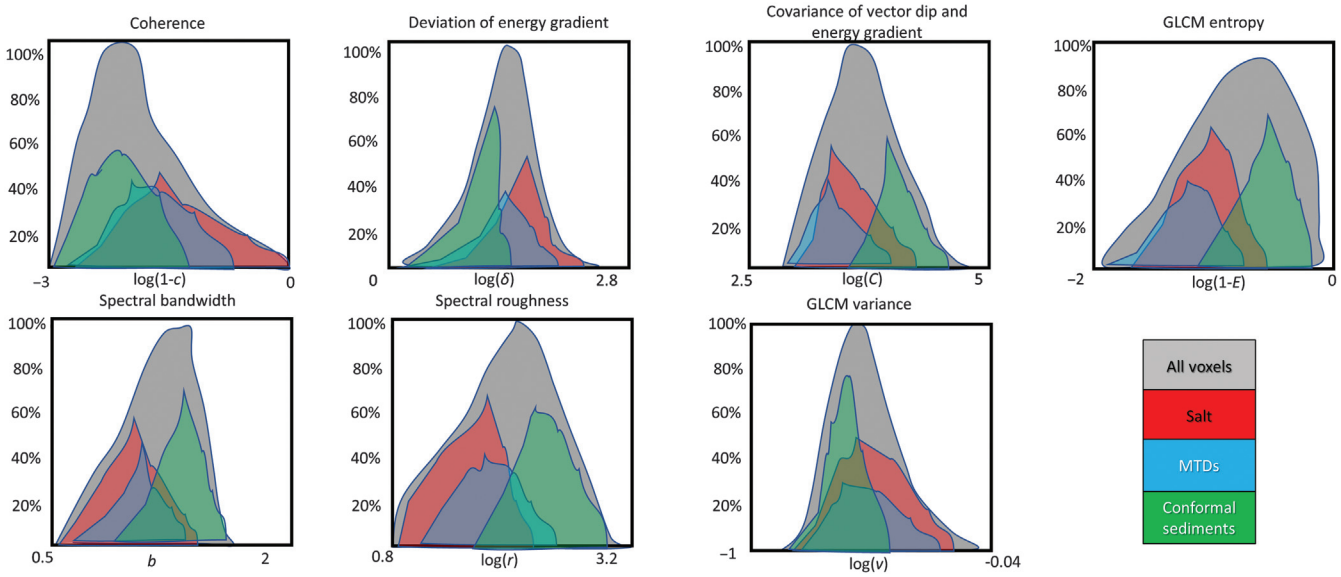

Figure 10. Histograms of the seven selected attributes for the entire data volume. The red masks indicate most of the salt voxels located on the histograms of different attributes. Although most of the salt can be defined by a simple threshold, some of the MTDs as well as other chaotic facies such as turbidites may fall within these zones and potentially be misclassified as salt.

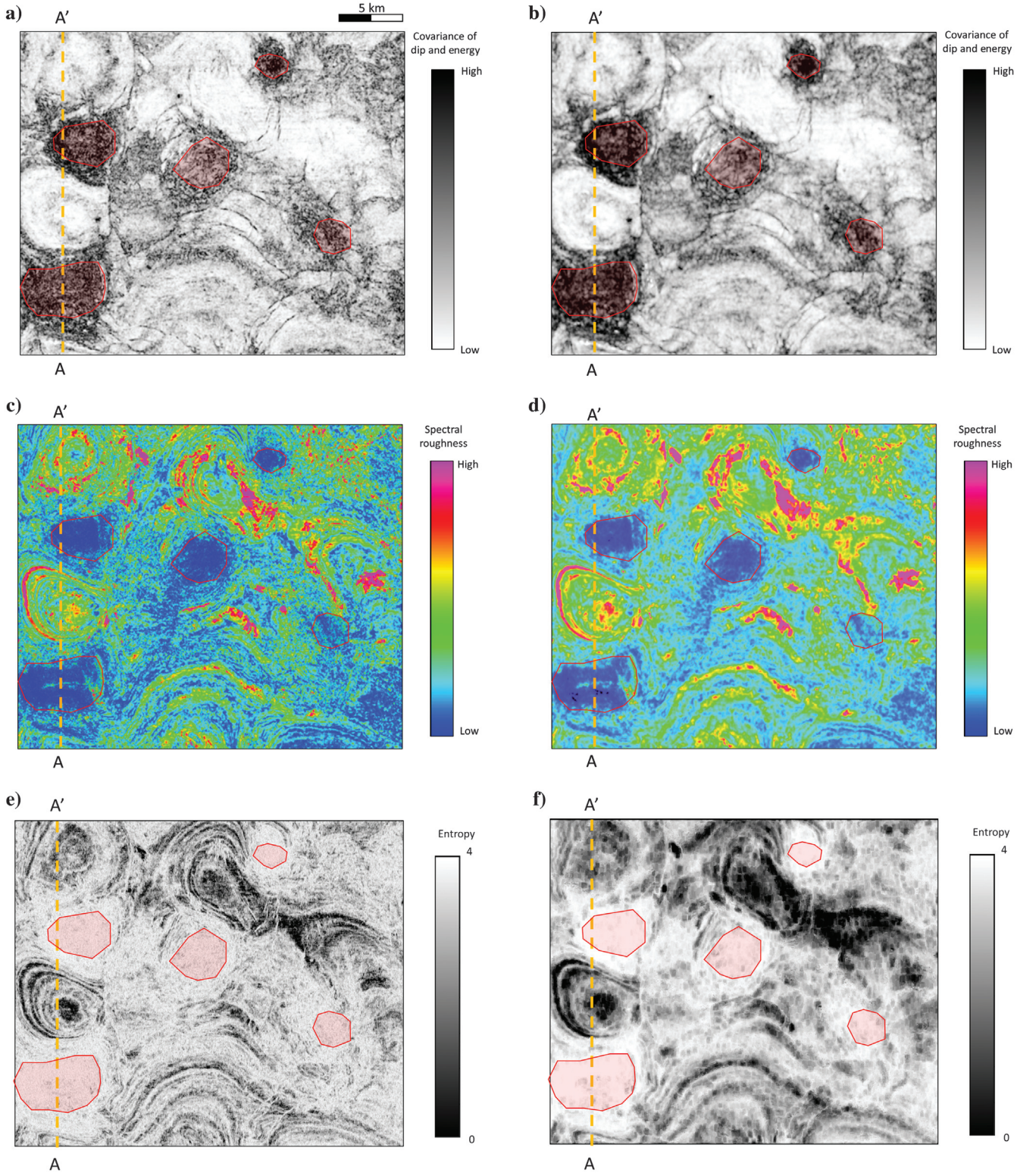

Figure 11. Time slices at $t = 1.22$ s through (a) the covariance of vector dip and energy gradient, (b) the covariance of the dip and energy gradient after 3D adaptive Kuwahara filtering, (c) the spectral roughness, (d) the Kuwahara-filtered spectral roughness, (e) the GLCM entropy, and (f) the Kuwahara-filtered GLCM entropy. Note that Kuwahara filtering increases the discrimination between salt facies and the generally conformal background facies.

GTM

Like SOM, GTM is considered to be an unsupervised classification algorithm. More accurately, like PCA and ICA, SOM and GTM are projection algorithms. The first two eigenvectors constructed from the N attributes represent a 2D plane that best fits the data distribution in N-dimensional space. Projecting each data vector onto this plane and summing the results gives the first two principal components. SOM and GTM deform the original plane into a 2D manifold that better fits the data in the N-dimensional space. In general, not all of the data vectors are used to construct the manifold. In our examples, we decimate the data by a factor of five in each of the three dimensions, reducing the amount of data needed to construct the manifold by a factor of 125. We find that such a decimation still represents all but the smallest geologic anomalies. MTDs, salt domes, and conformal sediments are the three most common facies found in our data volume, and they are well-represented by the decimated data. If we wished to use GTM to identify less common patterns (e.g., progradation, turbidites, and bright spots), we would wish to augment the decimated data with these data vectors to assure that they are well-represented by the projection. The GTM classification technique represents the PDFs of attribute data vectors (2D histogram) by a 2D color bar. Seismic facies can be recognized by the different color in the unsupervised learning. We will show that selecting attributes using our GMM-based workflow exhibits better facies maps.

Unlike supervised learning algorithms such as multilinear feed forward neural networks (Meldahl et al., 1999), probabilistic neural networks (Lubo-Robles et al., 2019), and random forest decision trees (Kim et al., 2019), GTM provides no indication of what the projections mean geologically. Nevertheless, some level of supervision can be provided in two ways. First, the attribute selection process can result in facies that overlap or are separated on the manifold. Second, after projection, we can provide a posteriori supervision by projecting vectors corresponding to each of the labeled facies onto the manifold. Because GTM is probabilistic, we can compute the likelihood that any given data vector at a voxel is like one of the target facies. Obviously, there will be some data vectors that do belong to one of the target facies. These voxels can be mapped in an "other" facies. Greater details on the working of GTM can be found in Roy et al. (2014) and Zhao et al. (2015).

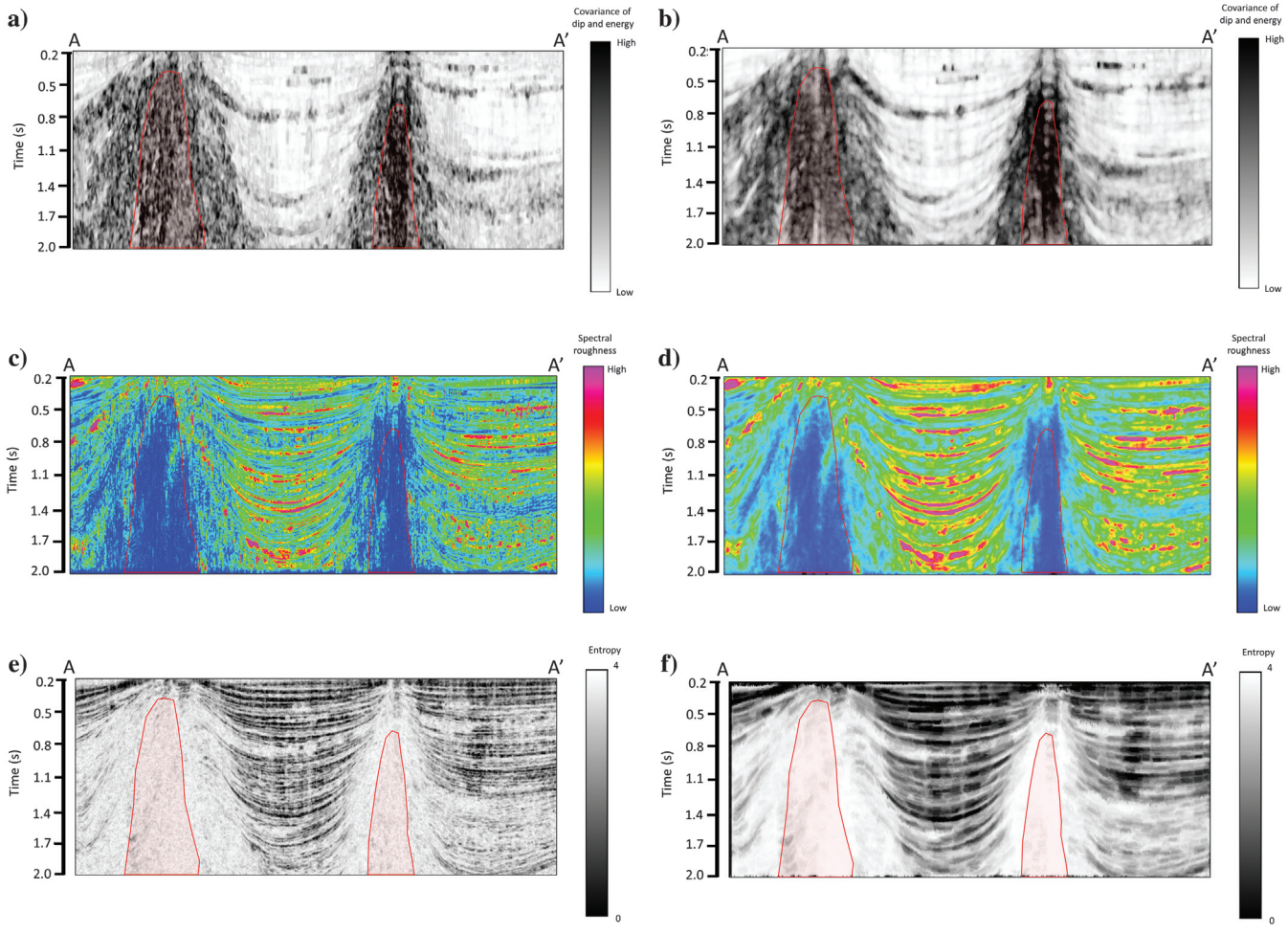

Figure 12. Vertical slices along AA′ through (a) the covariance of the vector dip and energy gradient, (b) the covariance of the dip and energy gradient after 3D adaptive Kuwahara filtering, (c) the spectral roughness, (d) the Kuwahara-filtered spectral roughness, (e) the GLCM entropy, and (f) the Kuwahara-filtered GLCM entropy.

APPLICATION 1: DIFFERENTIATING SALT AND MTDS FROM CONFORMAL SEDIMENTS

Our first application is to determine which attributes best differentiate salt from MTDs from conformal sediments for a marine survey over the Louisiana shelf. The survey covers approximately 8000 km^2 with 37.5 × 25 m bins. With heterogeneous sediment loading associated with the Mississippi River and its predecessors, minibasins grow and push down on the salt, which then rises to form salt diapirs. The change in slope gives rise to MTDs. Additional seismic facies observed in the data set include undeformed shale, interbedded sand and siltstone, and shallow and deepwater channels.

Seismic amplitude is almost always the initial "attribute" used in seismic interpretation in which a skilled interpreter can readily identify most of the larger geologic features by the spatial variation of seismic amplitude and phase. Useful seismic attributes provide quantitative geometric, dynamic, or statistical measures of the same features seen by the interpreter. Figure 4a shows a time slice at $t = 1.22$ s, and Figure 4b shows a vertical slice along line AA′ through the seismic amplitude volume. The polygons indicate the three picked facies of interest ($M = 3$), which are salt diapirs, conformal reflectors, and MTDs. The seismic expression of salt in the GOM data is of low amplitude, low frequency, and vertically and laterally chaotic. Because the data are prestack time rather than depth migrated, the steeply dipping salt boundaries may not align with the termination of relatively horizontal reflectors. The seismic expression of parts of the MTDs is also chaotic and of low frequency. However, the better imaged parts of the MTDs are more organized, showing rotated fault blocks containing reflectors exhibiting mixed energy and frequency that are a function of the internal layering. The picked salt facies in Figure 5a contains random, incoherent seismic noise that appear chaotic; however, mismigrated coherent noise events appear internal to the salt dome. The MTD facies in Figure 5b shows coherent, rotated reflector blocks as well as less coherent gravity flows sliding into the minibasin. The magnified image of the background seismic pattern shown in Figure 5c consists of coherent, conformal layered sediments. We define (or label) the three facies of interest on 5 time slices, 5 crossline slices, and 5 inline slices for a total of 15 slices or less than 1% of the data to be classified. In general, more labeled data provide a more accurate classification, but with increased computation cost. The total number of labeled voxels is small (totally 0.014% compared of the data available), allowing us to evaluate the 2^N–1 attribute combinations at a cost that is significantly smaller than the GTM projection algorithm.

Familiarity with the seismic expression of salt domes, MTDs, and conformal reflectors suggests that attributes sensitive to conformity, coherence, frequency, and texture may be good candidates for machine-learning-assisted seismic facies classification. Therefore, we propose candidate attributes that measure conformity (parallelism and/or reflector convergence), continuity (coherence), seismic textures (GLCM textures), rms amplitude or energy, and properties of the seismic spectra (peak frequency, bandwidth, and slope), for a total of the nine candidate attributes ($N = 9$) shown in Figure 6. Interactive interpretation using 3D visualization indicated that the deviation of vector dip, the deviation of energy gradient, and the covariance of vector dip and energy gradient highlighted chaotic

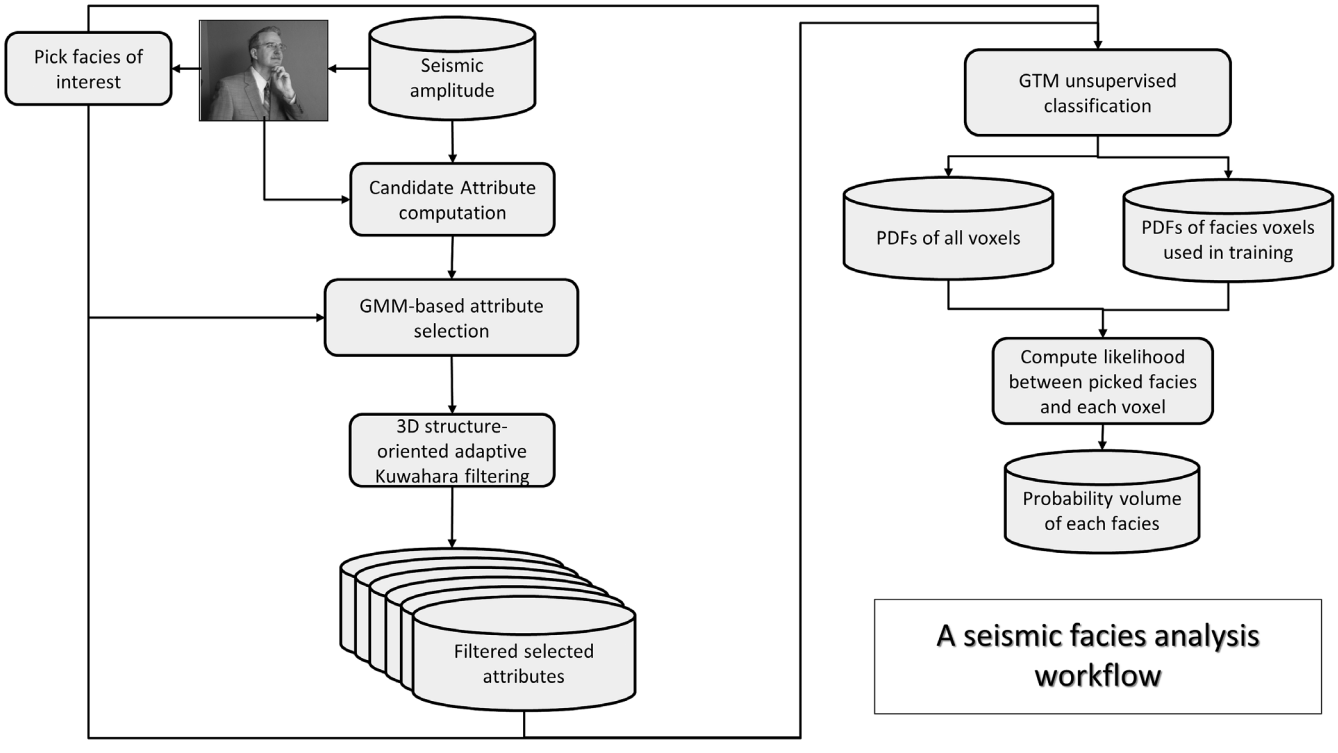

Figure 13. Workflow illustrating the steps used in a GMM attribute selection computer-assisted seismic facies classification. After selection, the chosen attributes are subjected to adaptive Kuwahara filtering. In this workflow, we use GTM but one can also use the PDFs from probabilistic neural networks or construct probabilities from SOM algorithms. For GTM, the comparison of the probability volumes is computed using the Bhattacharyya coefficient.

salt areas, high-energy salt boundaries, and nonparallel random seismic events, respectively. The spectral bandwidth and spectral roughness attributes are statistical measures of the spectrum that can further help to differentiate the seismic response of the salt from other facies.

Extracting the training voxels of the three picked facies from the nine candidate attributes is the first step of the attribute selection workflow. The total number of possible combinations using nine attributes is $2^N - 1 = 511$. Different attribute combinations may require different numbers of component clusters. Although a "homogeneous" seismic facies like salt may often be well-represented by a single Gaussian, more heterogeneous seismic facies like an MTD may require two or more Gaussian distributions. To decrease computation cost and the influence of seismic random noise, we limited the maximum number clusters for each attribute combination for each facies to two. One GMM is computed for each facies for each of A_n attribute combination when using a subset n of the N attributes. The number of clusters in a subset is J_n. Values of A_n and J_n as well as the number of comparisons are shown in Figure 7. We compute the Bhattacharyya distance given in equation 7 between each GMM cluster in each n-selected attribute combination to compare the similarity of each two-cluster GMM pair. The total number of such combinations is $J!/(J-2)!2!$. Then, we compute the commutative distance to evaluate each candidate attribute combination, and then we average the distance by the number of subgroup possible combinations, the number of picked facies, and the number of computed GMM clusters. Figure 8 shows the average cumulative distance of each possible number of selected attributes from the nine total attributes. The maximum distance is within the seven selected attributes of coherence, spectral bandwidth, GLCM entropy, GLCM variance, energy deviation, spectral roughness, and dip deviation (Figure 9). We examine histograms of the seven selected attributes (Figure 10). The red masks indicate most of the salt voxels located on the histograms of the selected attributes, the blue masks indicate most of the MTD voxels, and the green masks indicate most of the conformal background voxels.

To smooth the attribute response and suppress the overprint of seismic noise, we next apply the 3D adaptive Kuwahara filter to the selected attributes. Figure 11 shows the time slices through covariance of the structure dip and energy, spectral roughness, and GLCM entropy attributes and associated 3D adaptive Kuwahara filtered attributes. Figure 12 shows the vertical slices through the same attributes as shown in Figure 11. By comparing the attributes before and after filtering, we note that the 3D adaptive Kuwahara filter sharpens facies edges and smooths the interior texture of facies,

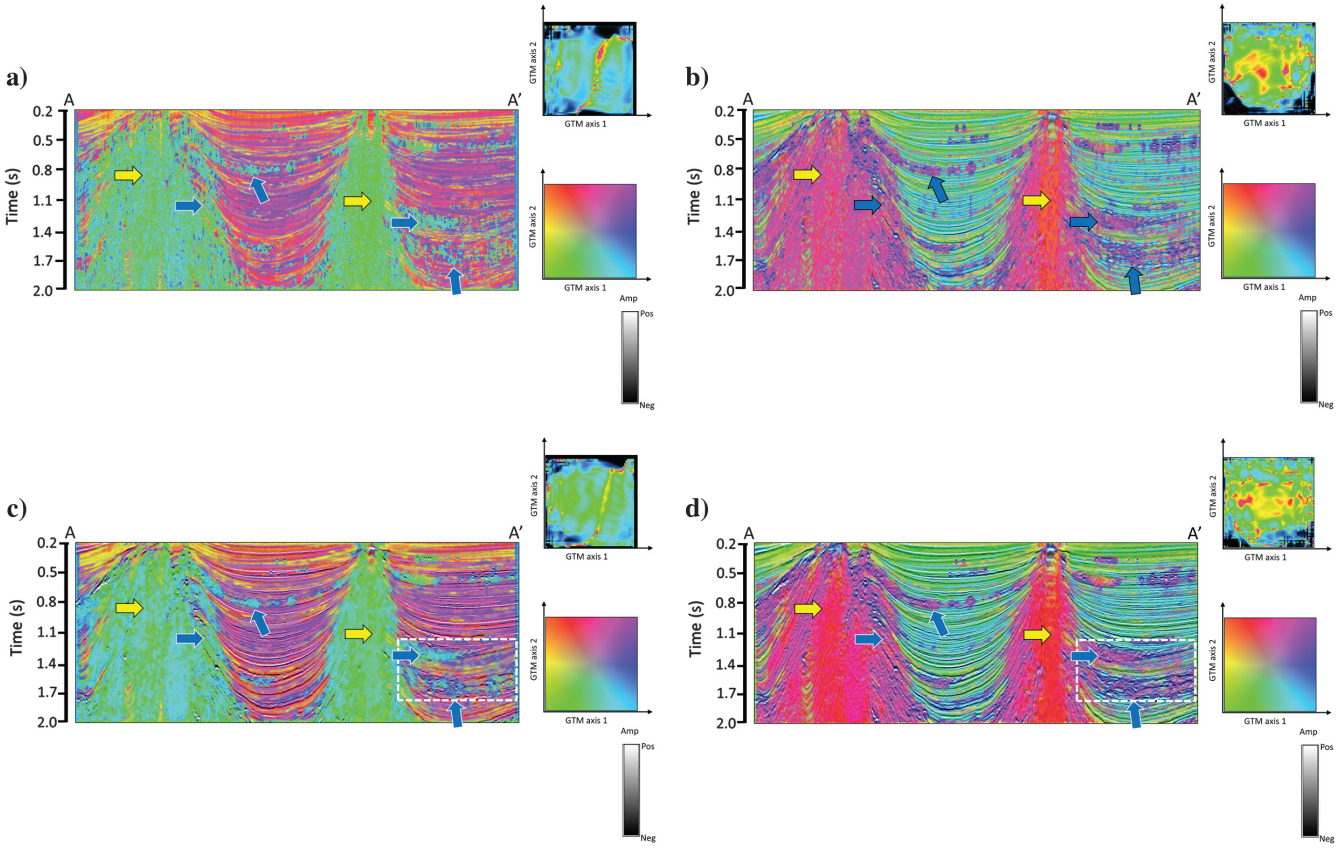

Figure 14. Vertical slices along AA′ through a seismic amplitude corendered with the GTM classification (a) using the interpreter-selected attributes as input, (b) using the seven best attributes using the GMM-based workflow, (c) using the same input attributes as in (a) after 3D adaptive Kuwahara filtering, and (d) using the same input attributes as in (b) after a 3D adaptive Kuwahara filtering. The interpreter-selected attributes are energy deviation, GLCM variance, spectral bandwidth, spectral roughness, and reflector convergence. The 2D histogram indicates the distribution of training voxels corresponded with the 2D color bar. Note that the GTM with the workflow selected attributes after 3D adaptive Kuwahara filtering exhibits the best facies map. The yellow arrows indicate the salt dome, and the blue arrows indicate MTD.

which results in more coherent, blocked seismic facies, more amenable to machine-learning classification.

We then apply a seismic facies analysis workflow to generate the probability volume of salt facies from a machine-learning-based multiattribute classification method. The workflow (Figure 13) is an extension of Qi et al.'s (2016) workflow in which we now use GMMs to select the attributes and an adaptive rather than a fixed-size 3D Kuwahara filter. We next compare the GTM mappings from the seven optimum attributes, and we compare it to the GTM mapping constructed from five attributes based on our experience and geologic insight: energy deviation, GLCM variance, spectral bandwidth, spectral roughness, and reflector convergence. We then project the labeled data voxels onto the two different GTM latent spaces, and the first GTM axis and the second GTM axis are represented by the posterior probability mean projections on the 2D latent space. Figure 14a shows the vertical slices of the crossplot of the first GTM axis and the second GTM axis computed from the interpreter-selected attributes, whereas Figure 14b shows the GTM result computed from the workflow-selected attributes. Figure 14c and 14d shows the results computed from the same suite of attributes as shown in Figure 14a and 14b but after 3D adaptive Kuwahara filtering. We note that the GTM result using the workflow selected attributes can better discriminate salt facies (the yellow arrows) from the chaotic parts of MTD facies (the blue arrows). MTD facies are more easily recognized on the GTM result computed from the workflow-selected attributes rather than the GTM result within the interpreter-selected attributes as the inputs. Application of the Kuwahara filtered attributes to the multiattribute classification results in smoother seismic facies and sharper facies edges. The cropped slices (the dashed white rectangle defined in Figure 14c and 14d) in Figure 15 highlight the improvements of mapping MTD facies of the GTM computed by the workflow selected attributes after the Kuwahara filtering, which exhibit less seismic noise. Figure 16a shows the time slice through the GTM within the interpreter-selected attributes as the inputs, and Figure 16b shows the time slice through the GTM within the workflow-selected attributes as the inputs before 3D adaptive Kuwahara filtering. Figure 16c and 16d shows the same time slices after 3D adaptive Kuwahara filtering. Note that the GTM within the workflow-selected attributes after Kuwahara filtering exhibits the best seismic facies. The last step of the seismic facies analysis workflow is to compute the probability volume associated with the facies. The Bhattacharyya coefficient provides a probability measure of how similar any given data vector at a voxel is to each of the labeled facies projected onto the 2D latent space. By examining the probability of each of the labeled facies (including the other facies), we obtain the probability that the data vector belongs to each of the facies. Simple thresholding above a sufficiently high Bhattacharyya coefficient provides a segmented image of each facies in 3D. Figure 17a shows the salt probability volume corendered with the seismic amplitude computed from the Bhattacharyya coefficient result within the interpreter-selected and Kuwahara-filtered attributes as the inputs, and Figure 17b shows the salt probability volume computed from the GTM result within the workflow-selected and Kuwahara-filtered attributes as the inputs. We note that the salt probability volume with the workflow-selected attributes exhibits improved salt delineation and fewer misclassified voxels than workflows using the five attributes the authors thought to be best and the fixed-size Kuwahara filter. Figure 17c shows the responsibility of total training voxels, picked salt voxels, MTD voxels, conformal sediment voxels, and other facies on the latent space.

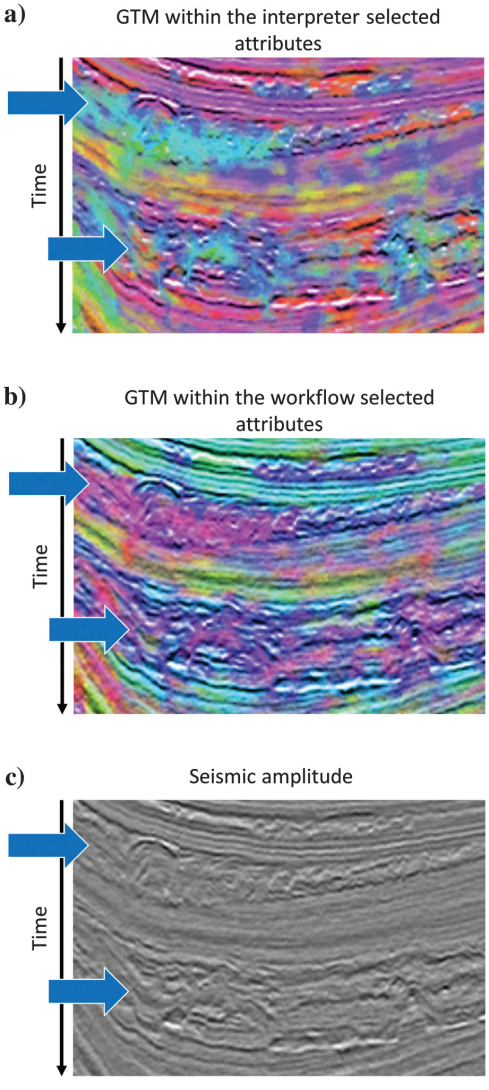

Figure 15. The magnified areas shown by the dashed white rectangles in Figure 14c and 14d, indicated by the dashed white rectangular in Figure 16c and 16d, exhibits the improvements of MTD mapping. The blue arrows indicate MTD.

APPLICATION TO KARST COLLAPSE FEATURES

Our second test application is to a wide-azimuth survey acquired in the Fort Worth Basin. The data were prestack time migrated with a 55×55 ft bin size to image the Barnett Shale and the underlying Ellenburger carbonate formation. The Ellenburger is characterized by karst and solution collapse features (Qi et al., 2014) that continue into the overlying Barnett Shale target. In general, horizontal wells that intersect these deformed areas produce high amounts of water and are uneconomic. Figure 18 shows the time slice through the seismic amplitude volume, in which the Ellenburger carbonate is the dominant formation. Using seismic attributes and seismic amplitude, we draw the yellow polygons to label the larger karst features. The input nine candidate attributes, and two picked facies — karst collapse features and unaltered carbonate, are shown in

Figure 19. Figure 20 shows the average cumulative distance of each possible selected attribute numbers through the GMM-based attribute selection workflow. We note that when selecting coherence, spectral bandwidth, GLCM entropy, GLCM variance, spectral roughness, and dip deviation attributes, the average cumulative distance is highest, which means that this combination is the optimum combination of the candidate attributes (Figure 21).

Figure 22a shows the GTM classification within the optimum attributes selected from the attribute selection workflow. The contract between karst facies and conformal background is also increased. Figure 22b shows the GTM classification within the 3D adaptive Kuwahara filtered optimum attributes as the inputs. Note that Figure 22b better exhibits karst facies than Figure 22a, in which karst collapse facies is piecewise smoothed. Then, we apply the same technique that has been used in the first example, and the final Bhattacharyya coefficient associated with karst collapse features is shown in Figure 23.

DISCUSSION

PCA is commonly used to reduce the dimensionality of the attribute volumes being analyzed, where in our example, the first three linear combinations of the nine attributes represent 80% of the data. The limitation of this approach is that features of geologic interest may represent less than 20% of the data and be poorly represented by the statistically most important linear combinations. This limitation was recognized by Roden and Chen (2017) who use PCA followed by a 5 × 5-neuron (prototype vector) SOM algorithm to model a Gulf of Mexico survey. They investigate which data vectors (voxels) were not well-represented by these neurons and find that many corre-

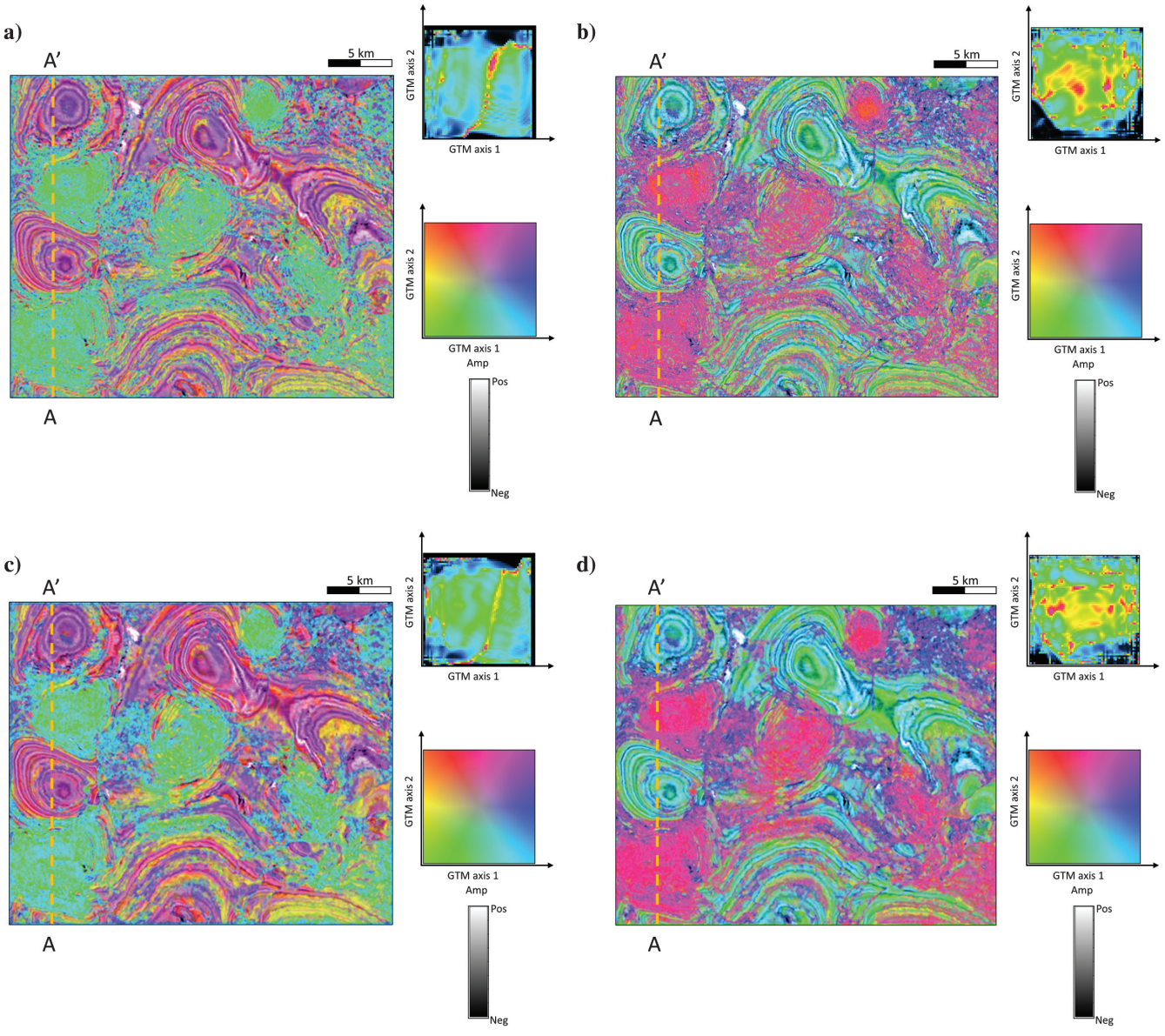

Figure 16. Time slices at $t = 1.22$ s through the GTM classification as in Figure 16. Note that the GTM with the workflow-selected attributes after 3D adaptive Kuwahara filtering exhibits the best facies map.

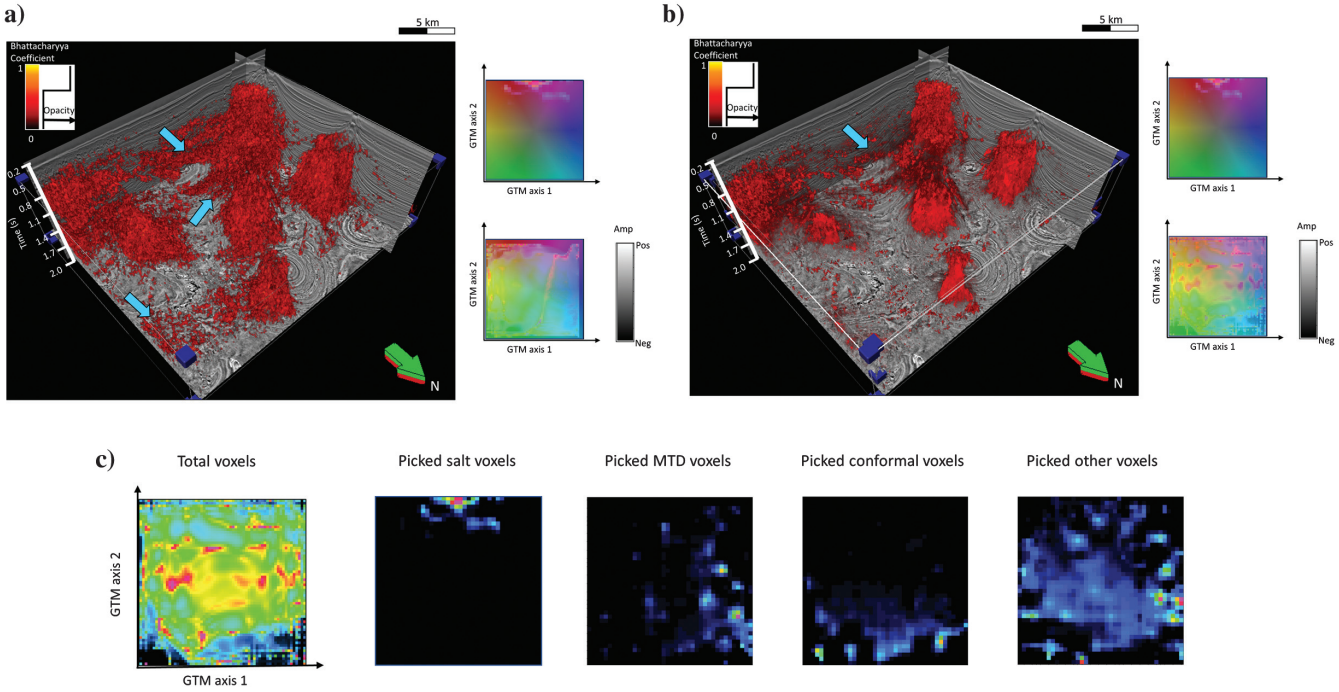

Figure 17. The 3D salt probability volume corendered with the seismic amplitude computed through the workflow (Figure 15) (a) using the interpreter-selected attributes as the inputs, (b) using the GMM-based attribute selection workflow-selected attributes as the inputs, and (c) the responsibility of total training, picked salt, MTD, conformal facies, and other facies voxels on the latent space. The suite of attributes is filtered by the Kuwahara filter. Note that the results in (b) exhibit less noise. The 2D histogram at the upper right shows the distribution of the picked salt voxels, whereas the histogram at the middle right shows the distribution of voxels representing all of the facies. The blue arrows indicate that the wrong classified voxels may be due to seismic noise or the facies within the similar histogram as salt.

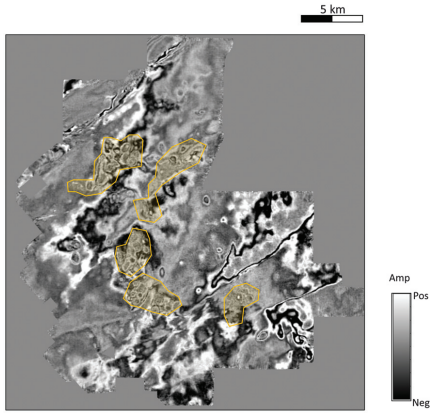

Figure 18. Time slices at $t = 0.72$ s through the seismic amplitude volume mapping karst collapse features on the Fort Worth Basin. Note that the yellow polygons indicate karst collapse features that are also painted facies of interest for this data set.

Input attributes
1) Coherence
2) Spectral bandwidth
3) Covariance of dip and amplitude
4) GLCM entropy
5) GLCM variance
6) Energy deviation
7) Spectral roughness
8) Reflector convergence
9) Dip deviation

Picked facies of interest

Karst collapse features

Conformal sediments

Figure 19. List of the candidate attributes and the picked facies.

Size of attribute subset, n	1	2	3	4	5	6	7	8	9
Best attribute combination	9	1,4	2,8,9	1,4,8,9	1,4,5,7,9	1,2,4,5,7,9	1,2,3,4,5,7,9	1,2,3,4,5,6,7,9	1,2,3,4,5,6,7,8,9
Cumulative distance, O_n	2.75	4.77	6.02	7.01	7.57	8.39	8.02	8.28	4.97

Figure 20. List of the average cumulative distance of each n selected attributes.

sponded to (anomalous) hydrocarbon bright spots of interest. In contrast, the SOM (16-sample waveform classification) software described by Coléou et al. (2003) begins with an initial classification using (by default) 12 prototype vectors defined on a 1D shoelace manifold in 16D space. The user is then able to introduce additional fixed waveforms (neurons) by averaging the seismic data about anomalous producing wells, or by generating synthetics of interest. In our workflow, the interpreter biases the GTM mapping by adding data vectors corresponding to the target facies that augment the otherwise uniformly sampled "training" data used to construct the manifold.

Our workflow consists of several steps: (1) labeling the target facies, (2) using an understanding of the seismic expression of the target features to select a (perhaps large) suite of candidate attributes, (3) using an adaptive Kuwahara filter to smooth and block the attribute response, (4) selecting an attribute subset that best differentiates the target facies using the metric described in this paper, (5) constructing a GTM from the decimated input data, augmented if necessary with the labeled data, (6) projecting all of the data onto the color-coded GTM manifold and latent space, and finally (7) projecting the labeled data onto the same manifold, thereby defining the likelihood that any given data vector (voxel) belongs to a specific labeled facies.

At present, we can confidently state that this workflow is a good candidate for multiattribute facies classification. However, we cannot claim that it is the best workflow. Working on the same data volume, Kim et al. (2019) find good results using supervised learning using nonlinear statistics and a random forest decision tree algorithm. Lubo-Robles et al. (2019) show promising results using Gaussian statistics and a supervised probabilistic neural network algorithm. Convolutional neural networks (CNNs) (e.g., Waldeland et al., 2018; Zhao, 2018) also show considerable promise though significantly less insight in discriminating seismic facies. The similarities between CNN and our workflow are striking in that CNN has multiple convolutional steps (i.e., it generates its own "attributes") and several pooling steps (similar to our successful application of the Kuwahara filter). A potential advantage of our GTM workflow over these three competing supervised learning workflows is that GTM will generate a projection that best represents all of the data used in its training, somewhat minimizing the chance of statistically important unlabeled facies being misclassified. Significant work will need to be conducted to determine which of these four candidates work best for a given problem. For the Gulf of Mexico data analysis shown here, the cost of the GMM attribute selection using the 60 labeled slices (81,204 labeled data vectors), nine candidate attributes, and 511 potential attribute combinations was 1 h. The training of the GTM using 22,207 decimated data vectors was 0.5 h. The projection of all of the data onto the manifold takes 0.75 h. In principle, it would make more sense to define the attribute combination that best separates the target facies when projected onto the 2D GTM manifold rather than in the N-dimensional attribute space. At present, such an analysis would take 511*(0.5 + 0.75) or 638 h to compute, thereby dominating the computation.

Selected attributes
1) Coherence
2) Spectral bandwidth
4) GLCM entropy
5) GLCM variance
7) Spectral roughness
9) Dip deviation

Figure 21. List of the selected attributes through the attribute selection workflow.

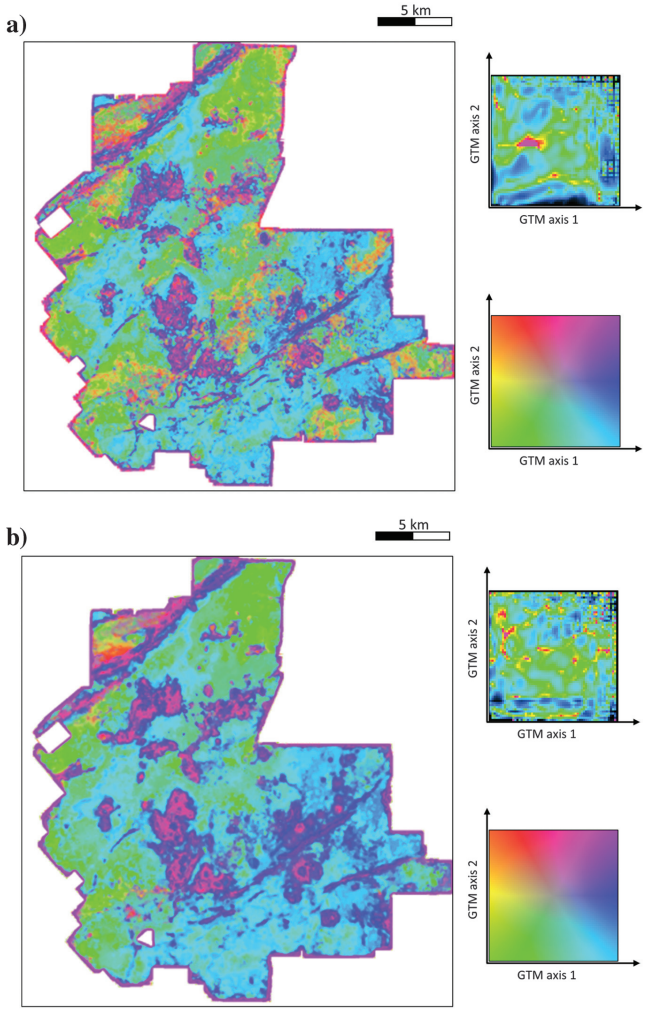

Figure 22. Time slices at $t = 0.72$ s through (a) the GTM classification with the workflow attributes as input before Kuwahara filtering and (b) the GTM classification with the adaptive Kuwahara-filtered attributes as input. Note the GTM classification with the adaptive Kuwahara-filtered attributes as input exhibits clearer and smoother karst collapse facies.

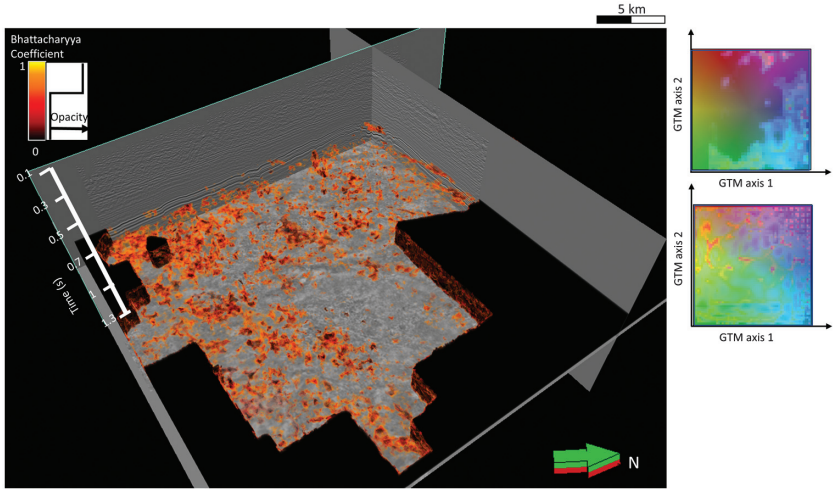

Figure 23. The 3D view of seismic amplitude corendered with the probability volume of karst collapse features computed through the workflow in Figure 15.

CONCLUSION

Although human interpreters routinely use seismic attributes as an aid to seismic facies classification in an interactive environment, it is unclear which attributes are best to differentiate seismic facies in a machine-learning environment. We have introduced a GMM-based attribute selection workflow for computer-assisted seismic facies classifications and applied it to two different data volumes containing a mix of chaotic, semichaotic, and conformal seismic facies. Unlike stepwise multilinear regression, we use an exhaustive search over candidate attributes to determine the optimum number and combination of attributes that should be used. In general, heterogeneous seismic facies such as salt and MTDs need to be represented by more than one Gaussian distribution. We find that the GMM-based attribute selection using a combined Classification and SEM algorithm, coupled with an averaged and summed Bhattacharyya distance measure allows us to evaluate all possible variations of attribute combinations to determine which combinate best separates the labeled seismic facies of interest in N-dimensional space.

When trying to differentiate chaotic features such as salt and MTDs, we found that attributes that are more statistical, including spectral bandwidth, dip deviation (non-parallelism), coherence, and GLCM textures worked well, whereas geometric attributes, such as the dip magnitude and curvature, worked poorly. We also found that a 3D adaptive Kuwahara filter helped smooth the internal response and sharpened the edges of seismic facies and sharpened the boundaries between adjacent facies.

Human interpreters are excellent at identifying seismic features and differentiating one feature from another and from noise. However, most human interpreters may find it challenging to describe the mechanics of how they do this identification and differentiation to a novice interpreter. An advantage of attribute-driven facies classification (using GMM with GTM, multilayer feed-forward neural network (MLFN), probabilistic neural network (PNN), or random forest decision trees) over deep learning CNNs is that the resulting successfully classified images provide greater insight into which features (attributes) allow us to differentiate one seismic facies from another, or, in the case of misclassification, what additional features (attributes) may need to be generated.

ACKNOWLEDGMENTS

We would like to thank PGS and Marathon Oil for providing a license to the data used in this research. We would also like to thank Schlumberger for the license of Petrel provided to the University of Oklahoma for use in research and education.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

APPENDIX A

DATA ADAPTIVE KUWAHARA IMAGE PROCESSING

When used for seismic facies analysis, the fixed-window Kuwahara window size needed for one facies may be inappropriate for a second facies. Likewise, as the seismic resolution decreases with depth, a fixed window size may oversmooth shallow features and undersmooth deeper features. To address this shortcoming, we modify the analysis window of the 3D structure-oriented Kuwahara filter to be adaptive along the lateral and vertical axes. Lin et al. (2014) use an adaptive window for coherence computation that define the lateral and vertical size of the analysis window based on the smoothed peak frequency. The traditional $5 \times 5 \times 5$-voxel 3D Kuwahara filter searches 27 overlapping subwindows and applies the median m in the overlapping window that has the smallest mean-normalized standard deviation σ/μ.

We follow Lin et al. (2014) and define the size of the adaptive Kuwahara searching window by the smoothed average power spectrum at each voxel. The average power spectrum is

$$P_{\text{avg}}(t,f) = \frac{1}{J(2I+1)} \sum_{j=1}^{J} \sum_{i=-I}^{I} u_j^2(t + i\Delta t, f), \qquad \text{(A-1)}$$

where t is the time sample, I is the half-length of the vertical analysis window, and u_j is the spectral magnitude of the jth trace. The average power spectrum is averaged over all traces and $2I + 1$ vertical samples. The peak frequency at time t is the frequency at which it is associated with the maximum power spectrum. The resolution of spectrally balanced data is determined not only by the peak frequency but also by the highest usable frequency. Suppose that the average power spectrum of the analysis points with its H neighboring voxels is f; first, we need to define the reference frequency f_{ref} of the adaptive analysis window at a given time t:

$$f_{\text{ref}}(t) = f_p(t) - \frac{\left(\sum_{f=1}^{F_p} P_{\text{avg}}(t,f) - p\sum_{f=1}^{F} P_{\text{avg}}(t,f)\right)}{P_{\text{avg}}(t,F_p)}\Delta f, \qquad \text{(A-2)}$$

where p is a percentile of the average power spectrum (in our work $P = 70\%$), Δf is the frequency increment between analysis window sizes in the computation of the average time-variant power spectrum, and f_p is the corresponding frequency. For a depth-migrated data set, time t in equation A-2 can be replaced by depth d and equation A-2 will become the depth-variant reference frequency. With the defined time or depth variant $f_{\text{ref}}(t)$, we can then define the corresponding time period or wavelength. For time-migrated data, the window height ΔT of the adaptive analysis window at a given time t can be defined as

$$\Delta T(t) = \frac{bP}{2f_{\text{ref}}(t)}. \tag{A-3}$$

Parameter $b = 0.01$ is a fixed prewhitening factor. Using a reference velocity $v(t)$, the two-way traveltime window height ΔT corresponds to a spatial vertical height ΔZ,

$$\Delta Z(t) = v(t)\frac{\Delta T(t)}{2}. \tag{A-4}$$

Equations A-3 and A-4 define an analysis window varying vertically for time-migrated data. Then, we use the local structure dip to define lateral variations of an analysis window. Therefore, when the window height ΔT increasing, the window width and length would also be increasing. If we assume the shape of an analysis window to be cylinder rather than rectangular, the width and the length of an analysis window at the crossline and inline directions are

$$\Delta X(t) = \frac{\Delta Z(t)}{w_x}, \tag{A-5}$$

and

$$\Delta Y(t) = \frac{\Delta Z(t)}{w_y}, \tag{A-6}$$

where w_x and w_y are the bin size along the crossline and inline directions. For depth-migrated data, wavelengths are used to define the vertical and lateral window size. Because the corresponding frequency is continuous, we need a taper to scale the analysis window. Lin (2016) shows an example that uses tapered analysis windows in the x-, y-, and z-directions, resulting in scaled traces.

Figure A-1 shows a cartoon of Kuwahara overlapping windows containing the red analysis point. In each window, we compute the median m, the mean μ, and the standard deviation σ. The result of the Kuwahara searching window w_l at an analysis point t is

$$d = \text{Arg}\left(\min_{d\in\{w_l|l\in 1,2,3\ldots L\}}\left(\frac{\sigma_d}{\mu_d}\right)\right), \tag{A-7}$$

where d is the index of the sub-Kuwahara searching analysis window. Thus, the filtered attribute will be the value of m_d associated with the window having the minimum value of normalized standard deviation σ_d/μ_d.

The 3D structure-oriented adaptive Kuwahara filter searches all windows containing a given voxel. The filtered attribute will have a smoothed facies that is shown as a blurred internal texture and sharpened edges between each facies. The smoothness is defined by a local average spectrum. The result of 3D adaptive Kuwahara filtering may somewhere look like the result of the fixed-window Kuwahara filtering computed with a large analysis window. However, one of the advantages is that the computation cost of the adaptive window is less than a large fixed window because not any subwindows should be equally large. Figure A-2a shows an example of application of the 3D adaptive Kuwahara filter to the deviation of vector dip in salt delineation. We first apply the original 3D Kuwahara filtering results computed with a fixed $3\times3\times3$ analysis window, which is shown in Figure A-2b. Note that the red polygon indicates salt, and the white color texture indicates another feature. Kuwahara filtering smooths the internal details of each feature and also sharpens the boundary. Figure A-2c shows the original 3D Kuwahara filtering result computed with a fixed $5\times5\times5$ analysis window, and Figure A-2d shows the result of our 3D adaptive Kuwahara filter. Note that the original Kuwahara filtering with a large window results in a better internal material, but a blurred edge. However, because we are using adaptive windows defined by a local average spectrum, the new adaptive Kuwahara filtering not only smooths low-frequency internal textures of salt but it also sharpens the high-frequency and energy salt boundaries.

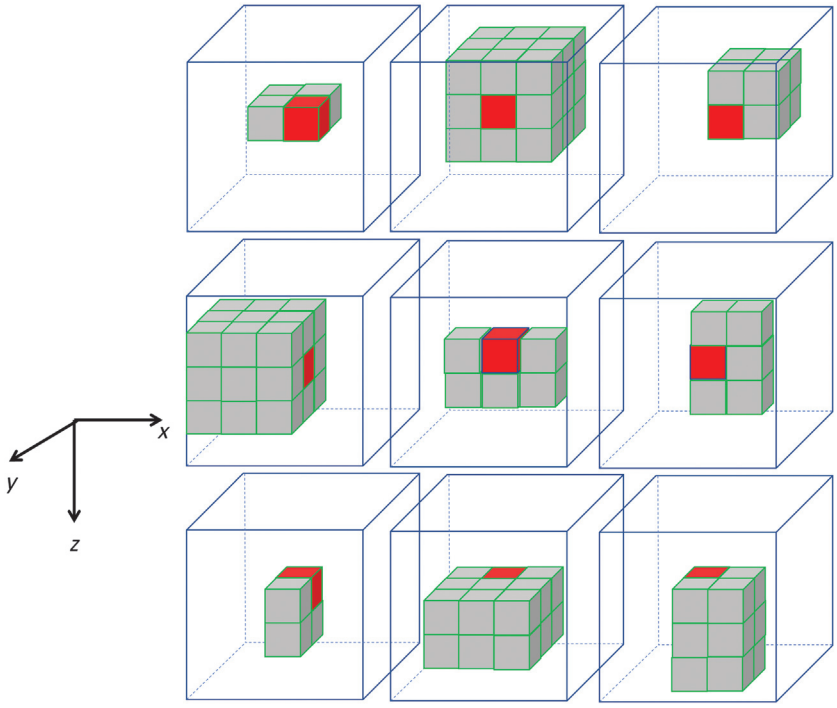

Figure A-1. Cartoon of 3D adaptive Kuwahara filtering. The $5\times5\times5$ or 125-sample analysis window is centered about the red voxel. The cartoon shows only nine of the $3\times3\times3$ or 27 subwindows that contain the red cube. The size of subwindow is adaptive, and increases in size with decreasing seismic resolution, defined by the local frequency of the seismic amplitude. The output is the median of the subwindow that has the smallest mean-normalized standard deviation.

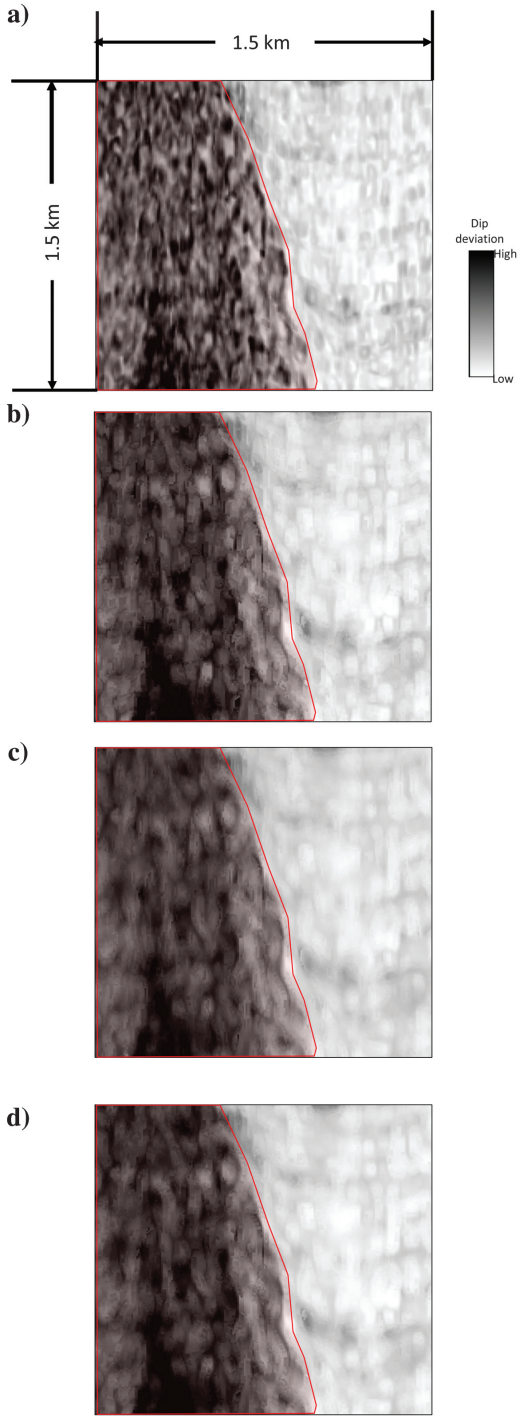

Figure A-2. Cropped vertical slices through the deviation of the vector dip attribute (a) before Kuwahara filtering, (b) after 3D Kuwahara filtering with a fixed $3 \times 3 \times 3$ analysis window, (c) after 3D Kuwahara filtering with a fixed $5 \times 5 \times 5$ analysis window, and (d) after adaptive Kuwahara filtering in which the window size varies between $3 \times 3 \times 3$ and $5 \times 5 \times 5$. In this example, the black anomalies indicate chaotic salt diapirs, whereas the white anomalies indicate relatively conformal facies. Kuwahara filtering smooths the interior response of salt while sharpening the edges. A small fixed analysis window results in less smoothing but greater sharpening, whereas a large fixed analysis window results in greater smoothing but less sharpening. The adaptive Kuwahara filtering better smooths the interior response of the salt and sharpens the edges.

REFERENCES

Barnes, A. E., 2007, Redundant and useless attributes: Geophysics, **72**, no. 3, P33–P38, doi: 10.1190/1.2716717.

Barnes, A. E., 2016, Handbook of poststack seismic attributes: SEG, Geophysical References Series 21.

Bartyzel, K., 2015, Adaptive Kuwahara filter: Signal, Image and Video Processing, **10**, 663–670, doi: 10.1007/s11760-015-0791-3.

Bubb, J. N., and W. G. Hatledid, 1977, Seismic recognition of carbonate buildups in seismic stratigraphy, *in* C. E. Payton, ed., Applications to hydrocarbon exploration: AAPG Memoir 26, 185–204.

Celeux, G., and G. Govaert, 1995, Gaussian parsimonious clustering models: Pattern Recognition, **28**, 781–793, doi: 10.1016/0031-3203(94)00125-6.

Coléou, T., M. Poupon, and K. Azbel, 2003, Unsupervised seismic facies classification: A review and comparison of techniques and implementation: The Leading Edge, **22**, 942–953, doi: 10.1190/1.1623635.

Cutler, A., 2003, The seashell on the mountaintop: A story of science, sainthood, and the humble genius who discovered a new history of the earth: Penguin Publishing Group.

de Lima, R. P., and K. J. Marfurt, 2018, Principal component analysis and K-means analysis of airborne gamma-ray spectrometry surveys: 88th Annual International Meeting, SEG, Expanded Abstracts, 2277–2281, doi: 10.1190/segam201.

Di, H., and D. Gao, 2017, Nonlinear gray-level co-occurrence matrix texture analysis for improved seismic facies interpretation: Interpretation, **5**, no. 3, SJ31–SJ40, doi: 10.1190/INT-2016-0214.1.

Forgy, E. W., 1965, Cluster analysis of multivariate data: Efficiency vs interpretability of classifications: Biometrics, **21**, 768–769.

Hampson, D. P., J. S. Schuelke, and J. A. Quirein, 2001, Use of multiattribute transforms to predict log properties from seismic data: Geophysics, **66**, 220–236, doi: 10.1190/1.1444899.

Hardisty, R., 2017, Unsupervised seismic facies using Gaussian mixture models: M.S. thesis, The University of Oklahoma.

Haskell, N., S. E. Nissen, M. J. S. Hughes, J. Grindhaug, S. Dhanani, R. P. Heath, J. D. Kantorowicz, L. Antrim, M. Cubanski, R. Nataraj, M. Schilly, and S. Wigger, 1999, Delineation of geological drilling hazards using 3D seismic attributes: The Leading Edge, **18**, 373–382, doi: 10.1190/1.1438301.

Hathaway, R. J., 1986, Another interpretation of the EM algorithm for mixture distributions: Statistics & Probability Letters, **4**, 53–56, doi: 10.1016/0167-7152(86)90016-7.

Infante-Paez, L., and K. J. Marfurt, 2018, In-context interpretation: Avoiding pitfalls in misidentification of igneous bodies in seismic data: Interpretation, **6**, no. 4, SL29–SL42, doi: 10.1190/INT-2018-0076.1.

Infante-Paez, L., and K. J. Marfurt, 2019, Using machine learning as an aid to seismic geomorphology, which attributes are the best input: Interpretation, **7**, no. 3, SE1–SE18, doi: 10.1190/INT-2018-0096.1.

Kim, Y., R. Hardisty, and K. J. Marfurt, 2019, Attribute selection in seismic facies classification: Application to a Gulf of Mexico 3D seismic survey and the Barnett Shale: Interpretation, **7**, no. 3, SE281–SE297, doi: 10.1190/INT-2018-0246.1.

Kuwahara, M., K. Hachimura, S. Eiho, and M. Kinoshita, 1976, Digital processing of biomedical images: Plenum Press, 187–203.

Kyprianidis, J. E., H. Kang, and J. Döllner, 2009, Image and video abstraction by anisotropic Kuwahara filtering: Computer Graphics Forum, **28**, 1955–1963, doi: 10.1111/j.1467-8659.2009.01574.x.

Lin, T., 2016, Seismic attribute analysis of depth-migrated data: Ph.D dissertation, University of Oklahoma.

Lin, T., B. Zhang, S. Zhan, Z. Wan, F. Li, H. Zhou, and K. J. Marfurt, 2014, Seismic attributes of time- vs. depth-migrated data using self-adaptive windows: 84th Annual International Meeting, SEG, Expanded Abstracts, 1659–1662, doi: 10.1190/segam2014-1579.1.

Lubo-Robles, D., T. N. Ha, S. Lakshmivarahan, and K. J. Marfurt, 2019, Supervised seismic facies classification using probabilistic neural networks: Which attributes should the interpreter use?: 89th Annual International Meeting, SEG, Expanded Abstracts, 2273–2276, doi: 10.1190/segam2019-3216841.1.

Lubo-Robles, D. R., and K. J. Marfurt, 2019, Independent component analysis for reservoir geomorphology and unsupervised seismic facies classification in the Taranaki Basin, New Zealand: Interpretation, **7**, no. 3, SE19–SE42, doi: 10.1190/INT-2018-0109.1.

Luo, Y., M. Marhoon, S. Al-Dossary, and M. Alfaraj, 2002, Edge-preserving smoothing and applications: The Leading Edge, **21**, 136–158, doi: 10.1190/1.1452603.

Mak, B., and E. Barnard, 1996, Phone clustering using the Bhattacharyya distance: Proceeding of Fourth International Conference on Spoken Language Processing 4, 2005–2008, doi: 10.1109/ICSLP.1996.607191.

Marfurt, K. J., 2018, Seismic attributes as the framework for data integration throughout the lifespan of oil field: SEG, Distinguished Instructor Series 508.

Matos, M. C., M. Yenugu, S. M. Angelo, and K. J. Marfurt, 2011, Integrated seismic texture segmentation and cluster analysis applied to channel delineation and chert reservoir characterization: Geophysics, **76**, no. 5, P11–P21, doi: 10.1190/geo2010-0150.1.

Meldahl, P., R. Heggland, B. Bril, and P. de Groot, 1999, The chimney cube, an example of semi-automated detection of seismic objects by directive attributes and neural networks — Part 1: Methodology: 69th Annual International Meeting, SEG, Expanded Abstracts, 931–934, doi: 10.1190/1.1821262.

Moscardelli, L., and L. Wood, 2007, New classification system for mass transport complexes in offshore Trinidad: Basin Research, **20**, 73–98, doi: 10.1111/j.1365-2117.2007.00340.x.

Qi, J., M. Cahoj, A. AlAli, L. Li, and K. Marfurt, 2015, Segmentation of salt domes, mass transport complexes on 3D seismic data volumes using Kuwahara windows and multiattribute cluster analysis: 85th Annual International Meeting, SEG, Expanded Abstracts, 1821–1825, doi: 10.1190/segam2015-5876831.1.

Qi, J., and J. P. Castagna, 2013, Application of a PCA fault-attribute and spectral decomposition in Barnett shale fault detection: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1421–1425, doi: 10.1190/segam2013-0674.1.

Qi, J., T. Lin, T. Zhao, F. Li, and K. J. Marfurt, 2016, Semisupervised multiattribute seismic facies analysis: Interpretation, **4**, no. 1, SB91–SB106, doi: 10.1190/INT-2015-0098.1.

Qi, J., B. Zhang, H. Zhou, and K. J. Marfurt, 2014, Attribute expression of fault-controlled karst — Fort Worth Basin, TX: Interpretation, **2**, no. 3, SF91–SF110, doi: 10.1190/INT-2013-0188.1.

Roden, R., and C. Chen, 2017, Interpretation of DHI characteristics with machine learning: First Break, **35**, 55–63.

Roden, R., T. Smith, and D. Sacrey, 2015, Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps: Interpretation, **3**, no. 4, SAE59–SAE83, doi: 10.1190/INT-2015-0037.1.

Roy, A., A. S. Romero-Peláez, T. J. Kwiatkowski, and K. J. Marfurt, 2014, Generative topographic mapping for seismic facies estimation of a carbonate wash, Veracruz Basin, southern Mexico: Interpretation, **2**, no. 1, SA31–SA47, doi: 10.1190/INT-2013-0077.1.

Schwarz, G., 1978, Estimating the dimension of a model: The Annals of Statistics, **6**, 461–464, doi: 10.1214/aos/1176344136.

Waldeland, A. U., A. C. Jensen, L.-J. Gelius, and A. H. S. Solberg, 2018, Convolutional neural networks for automated seismic interpretation: The Leading Edge, **37**, 529–537, doi: 10.1190/tle37070529.1.

Wallet, B. C., and R. Hardisty, 2019, Unsupervised seismic facies using Gaussian mixture models: Interpretation, **7**, no. 3, B62–B72, doi: 10.1190/INT-2018-0119.1.

Zhang, K., K. J. Marfurt, and Y. Guo, 2008, Volumetric application of skewed spectra: 78th Annual International Meeting, SEG, Expanded Abstracts, 919–923, doi: 10.1190/1.3063789.

Zhao, T., 2018, Seismic facies classification using different deep convolutional neural networks: 88th Annual International Meeting, SEG, Expanded Abstracts, 2046–2049, doi: 10.1190/segam2018-2997085.1.

Zhao, T., V. Jayaram, A. Roy, and K. J. Marfurt, 2015, A comparison of classification techniques for seismic facies recognition: Interpretation, **3**, no. 4, SAE29–SAE58, doi: 10.1190/INT-2015-0044.1.

Zhao, T., F. Li, and K. J. Marfurt, 2018, Seismic attribute selection for unsupervised seismic facies analysis using user-guided data-adaptive weights: Geophysics, **83**, no. 2, O31–O44, doi: 10.1190/geo2017-0192.1.