Higher Order Principal Component Analysis for Identifying AVO Anomalies

Stuart Farris and Stewart Levin

ABSTRACT

We use principal component analysis (PCA) to identify amplitude versus offset (AVO) anomalies in a legacy 2D seismic data set. The traditional methods of AVO analysis rely on inversion techniques that require simplified models of the subsurface. Simple models of the earth make assumptions that may lead to incorrect hydrocarbon identification. By making no assumptions about the earth model or how waves propagate in the model, PCA could provide a more robust technique for finding AVO anomalies.

INTRODUCTION

The underlying principles of AVO analysis are presented in Ostrander (1984). In essence, rock formation boundaries with highly contrasting Poisson's ratios can create "bright spots" in seismic offset gathers. These anomalies often indicate trapped hydrocarbon gas which may overlie oil reserves. The promise of having oil reserves light up on seismic gathers has lead the industry to invest significant effort into developing robust detection techniques (Keys and Foster, 1996). But, the propagation of seismic waves through the earth is too complex to model perfectly. Consequently, models make assumptions regarding elasticity, anisotropy, attenuation, and other physical properties (Hill, 2015). These physical assumptions in the modeling process can lead to false-positive or missed reservoir identifications.

Bougher and Herrmann (2016) presented a statistical method for AVO anomaly identification that requires no rock property assumptions. The method transforms migrated common midpoint gathers (CMPs) into their principal components and correctly identifies gas AVO anomalies. Their abstract shows, in synthetic 2 dimensional examples, this PCA method can correctly identify gas AVO anomalies with results comparable to traditional AVO inversion techniques.

We attempt to expand upon their method in multiple ways. Bougher and Herrmann (2016) only examined the first and second principal components for anomalies. We examine higher order principal components with hopes to identify more complex AVO anomalies. Furthermore, we conduct our study on the Mobil Viking Graben data set. This allows us to test PCA on real data which has already been highly scrutinized for AVO anomalies. The legacy AVO results will provide a benchmark for our new method.

METHOD

We use higher order principal components of CMP gathers to identify AVO anomalies.

The goal of PCA is to linearly transform data into a coordinate system where the projection of the data into that system has the highest variance on the first axis, the second highest variance on the second axis, and so on. These axis are known as the principal components. By keeping those components that are most significant, we reduce the number of features in a data set while maximizing the variance of those features. For this application we take single, prestack-migrated CMP gathers, and transform them into their principal components. Therefore, we transform some CMP gather $D \in \mathbb{R}^{i \times j}$, with i time samples and j offsets, into $\tilde{D} \in \mathbb{R}^{i \times j}$, with i time samples and j principal components. Each time sample in a CMP now has values for each principal component. By cross-plotting principal components, we can find time samples that do not match the time vs. offset variance trend of all the time samples. It was shown by Bougher and Herrmann (2016) that outliers in crossplots of the first and second principal components can correspond to AVO anomalies. The crossplot outliers can be visualized by mapping the time samples into image space for each CMP gather.

We extend this idea by cross-plotting higher principal components and analyzing their outliers. We hypothesize that they correspond to more subtle AVO anomalies resulting from anisotropy or other local rock property changes. The more difficult task will be linking newly identified anomalies to the physical phenomenon that caused them. Modeling rock property changes in synthetic data and observing the effects on the principal components will be the focus of future work with this topic.

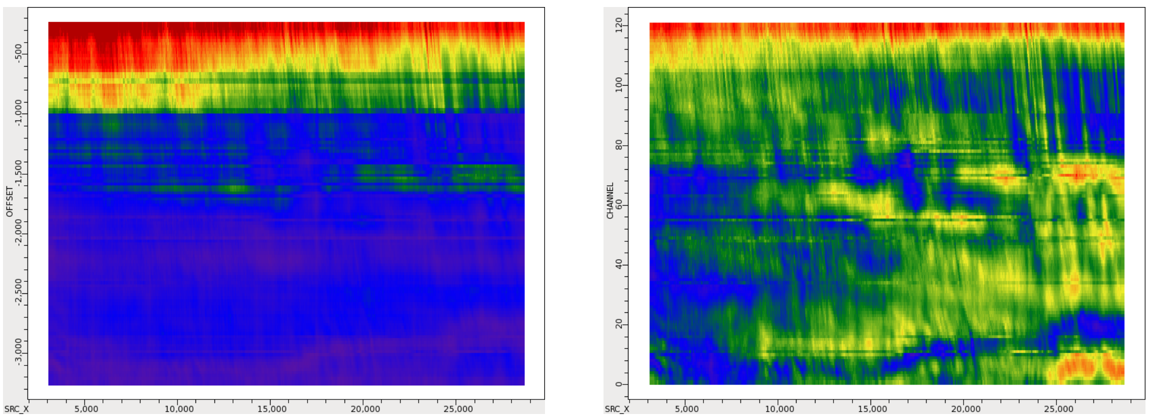
DATA

Here we evaluate our method using the public Viking Graben data set collected by Mobil and provided by the Society of Exploration Geophysicists (SEG). This data set was released for industry and academic institutions to test "new ideas about the use of seismic inversion or AVO methods for detecting hydrocarbons" (Keys and Foster, 1996). We continue the investigation of AVO methods on this data set with our new PCA technique. The Stanford Exploration Project (SEP) is very familiar with this data set, namely in reports 80 through 84. With the legacy results from SEG and SEP reports, we will be able to assess the feasibility of our PCA anomaly detection.

DATA PROCESSING

To prepare the Viking Graben data for PCA, preprocessing steps must be considered. Notice that these steps do not preserve amplitudes for quantitative AVO analysis.

Amplitude variations caused by inconsistent source strengths and hydrophone sensitivities can distort AVO results Berlioux and Lumley (1994). To mitigate this issue we balance the amplitudes of traces for both shot and receiver variations using OpenCPS (Open Geophysical, Inc., 2016). To balance shot strengths, we compute the root-mean-square (RMS) for all shot gathers. Then we subtract the amplitudes of a particular shot from the average RMS of all shots. Likewise, we compute the RMS of all receiver gathers then subtract amplitudes of each receiver from the average RMS of all receiver gathers. This ensures no individual trace can bias the average shot or receiver RMS value. Figure 1 illustrates the RMS amplitudes for each trace before and after amplitude balancing. Most of the streaks in the vertical direction, caused by source strength variation, and horizontal direction, caused by hydrophone sensitivity variation, have been removed.



(a) (b)

Figure 1: RMS amplitude values plotted in shot vs offset domain using a) original data b) data after shot and receiver strength correction [**CR**]

Our PCA method assumes flattened, migrated CMP gathers. To obtain these gathers, we use the prestack Kirchhoff migration in OpenCPS with the velocity model found in (Ecker and Lumley, 2001) illustrated in Figure 2. This migration method is advantageous since it preserves amplitudes reasonably well and keeps our data in the CMP domain.

We also attempt surface-related-multiple-elimination (SRME) using the workflow outlined in OpenCPS. It is unclear if multiples affect our PCA method. Intuitively, it seems multiples contaminate migrated gathers and distort the amplitude variations we wish to identify. Regardless, we will attempt our method with and without SRME.

Finally, we are ready for PCA. We use the PCA code in the University of Tokyo's C Clustering Library (Michiel de Hoon et al., 2013).

PRELIMINARY RESULTS

Figures 3, 4, and 5 show crossplots of various principal components for a single CMP gather. The CMP gather was selected from the center of the line to test the viability of this method before applying the technique to the full collection of gathers.

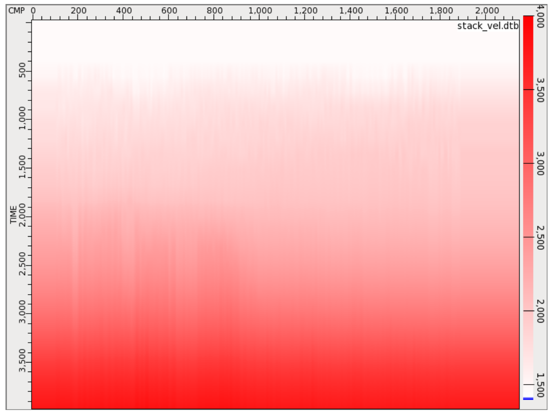
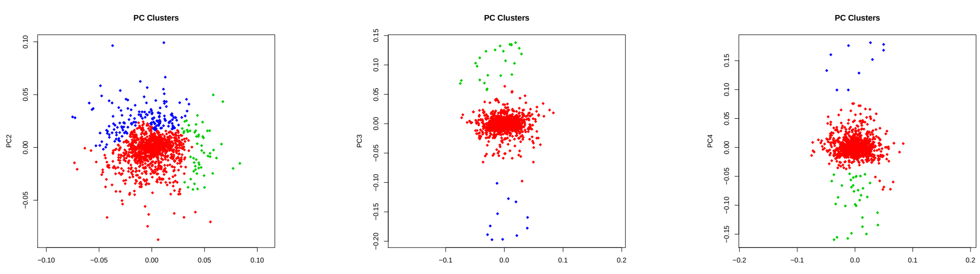


Figure 2: Stacking velocities used for SMRE and prestack Kirchhoff migration [**CR**]

In order to automatically identify outliers in the crossplots, we used a hierarchical clustering technique to group samples based on euclidean distance in the principal component space. Here we have chosen to cluster the samples into three groups which are identified by their color in the principal component crossplots. The optimal number of clusters needs to be determined if we want to quickly identify anomalies in many CMP gathers.



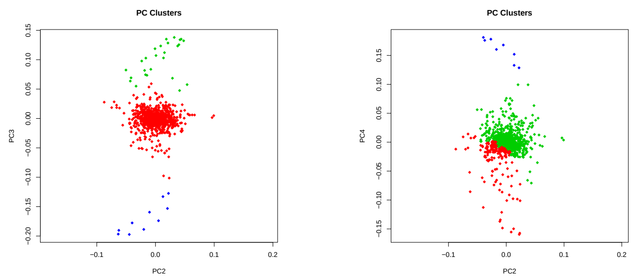
(a) (b) (c)

Figure 3: Crossplots of principal components a) 1 vs. 2, b) 1 vs. 3, and c) 1 vs. 4 [**CR**]

DISCUSSION

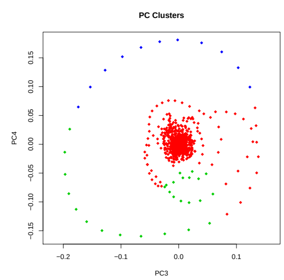
In Figure 3a we do not see outlier clusters that are as distinct as we expect. If there was a significant AVO anomaly, we would expect it to be a clear outlier in the crossplot of the principal components 1 vs 2, as shown in Bougher and Herrmann (2016). It is possible that the CMP gather selected for testing did not pass through a significant anomaly. It is also possible a step in processing prior to PCA dampened any existing anomalies.

Higher order principal component crossplots show more promise. The clusters



(a) (b)

Figure 4: Crossplots of principal components a) 2 vs. 3 and b) 2 vs. 4 [**CR**]



(a)

Figure 5: Crossplot of principal components 3 vs. 4 [**CR**]

clearly separate in many of the crossplots shown in Figures 3b-4b. But, it is difficult to rely on these results until we can reproduce the results we expect in principal components 1 vs. 2.

WHAT'S NEXT

The next steps for this project are clear. First, we need to refine our preprocessing until we are confident that amplitudes are not compromised beyond the point of AVO analysis. Next, we need to replicate the PCA cross-plotting with a larger number of gathers that we know interact with AVO anomalies. We know these crossplots should show distinct outliers. Assuming successful testing on a subset of CMP gathers, we will apply PCA to all gathers. The outliers identified from all the gathers will then be mapped and visualized in the migrated image space. The results in the image space can then be compared to the legacy AVO analysis results found in the Viking Graben data.

The more difficult task will be matching anomalies found with PCA to the physical phenomenon that caused them. A synthetic model will have to be used to induce various physical anomalies, such as density, velocity, or anisotropy. With this model we can forward model a 2D seismic experiment and evaluate principal component anomalies.

ACKNOWLEDGMENTS

We must first acknowledge Mobil and SEG for providing the legacy Viking Graben data. Further, borrowing the PCA code from the University of Tokyo allowed us to avoid reinventing the wheel with regards to a statistical method that is well understood. Finally, we must thank all of the sponsors of the SEP for their continued funding and support.

REFERENCES

Berlioux, A. and D. Lumley, 1994, Amplitude balancing for AVO analysis: SEP-80: Stanford Exploration Project, 349–360.

Bougher, B. and F. Herrmann, 2016, AVA classification as an unsupervised machine-learning problem, *in* SEG Technical Program Expanded Abstracts 2016, 553–556, Society of Exploration Geophysicists.

Ecker, C. and D. E. Lumley, 2001, Velocity analysis of the Mobil AVO dataset: SEP-82: Stanford Exploration Project.

Hill, S., 2015, Topics in Seismic Analysis.

Keys, R. G. and D. J. Foster, 1996, Comparison of Seismic Inversion Methods on a Single Real Data Set. Open file publications, No. no. 3: Society of Exploration Geophysicists.

Michiel de Hoon, Seiya Imoto, and Satoru Miyano, 2013, The C Clustering Library: Technical report, The University of Tokyo, Institute of Medical Science, Human Genome Center.

Open Geophysical, Inc., 2016, OpenCPS 3.3 Manual.

Ostrander, W., 1984, Plane-wave reflection coefficients for gas sands at nonnormal angles of incidence: Geophysics, **49**, 1637–1648.