

# Geophysviz: A visualization approach to multivariate analysis of seismic attributes

Ben Bougher

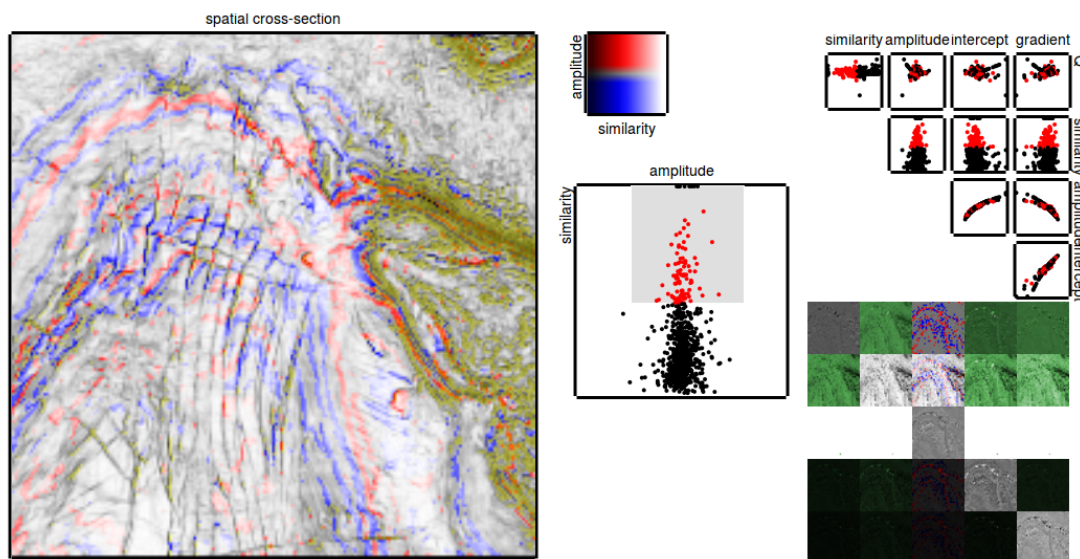


Fig. 1. Geophysviz: Interactive attribute analysis.

**Abstract**—Seismic images are acquired by firing an impulsive source into the earth and recording the backscattered energy. A seismic image can be decomposed into multivariate signal processing attributes which contain unique information about the material properties at the reflection interface. Attribute analysis requires the inherently difficult task of visually interpreting multidimensional geospatial data. I present Geophysviz, a system that uses linked geospatial and scatterplot views to support visual exploration of multivariate geospatial data. Geophysviz has two primary detailed views: a scatterplot which shows the bi-variate distribution between two attributes, and a geospatial image which uses the colour channels to co-render two attributes simultaneously. A select brushing interaction technique supports linked highlighting between the geospatial view and scatterplot, which allows for geospatial referencing of clusters, trends, and outliers. The entire multivariate data space is faceted into a summary view of juxtaposed tiny multiples, allowing for interactive navigation and exploration. Slow transitions are used to switch between detailed scattered plot views to encourage visual tracking of clusters and outliers. I present the application of Geophysviz to a small section of synthetic and measured seismic attribute data from the open F3 seismic data set.

**Index Terms**—seismic, multivariate, visualization

## 1 INTRODUCTION

In exploration geophysics, analysts relate anomalies in seismic data to potential targets of interest such as reservoirs and gas contacts. Seismic data is typically acquired by firing an impulsive source into the earth and recording the reflected energy at receivers at the surface, as shown in Figure 2. The reflection data creates an image of the subsurface of the earth called a seismic image. Reflections are caused by impedance contrasts between sedimentary layers of the earth, where the impedance is a function of the physical properties of the materials at the interface.

Each reflection appears in the data as a scaled, filtered, and modu-

lated copy of the source waveform. The goal of quantitative seismic interpretation is to relate the shape of the reflected waveform to physical properties of the reflection interface, see Figure 3. Unfortunately the transfer function is grossly undetermined which forces analysts to use indirect methods in order to classify reflections.

A seismic attribute can loosely be defined as a quantity derived from seismic data that can be analyzed in order to enhance information that might be more subtle in a traditional seismic image [10]. The zoo of specific seismic attributes is large and outside the scope of this paper, so I will refer to seismic attributes in the general sense of derived data. Each attribute adds a dimension to the seismic image, where every geospatial point in the seismic image now contains a vector of derived attributes. The attributes form a multivariate space, where analysts try to relate trends, anomalies, and clusters to rock classifications and geospatial regions of interest.

The main workhorse for attribute analysis are cross plots, where analysts visually inspect 2D scatterplots of attributes for a chosen section of data. Points that form interesting clusters are then geospatially cross-referenced to see if the cluster corresponds to a geospatial re-

• E-mail: ben.bougher@gmail.com.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.

Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxx/

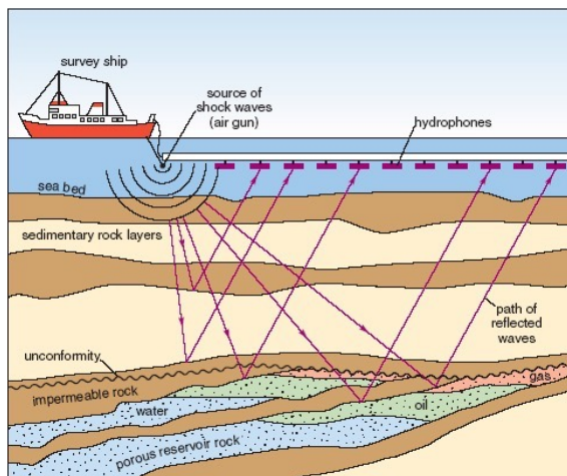


Fig. 2. Marine seismic data acquisition. Energy is reflected off geological interfaces and recorded by receivers at the surface. Diagram source: Open University

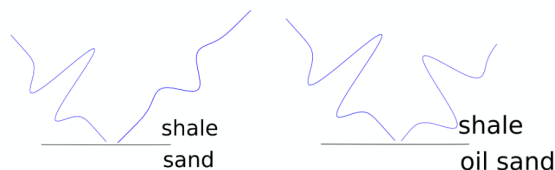


Fig. 3. Example of a reflected waveform changing shape at an interface of geological contrasts.

gion in the seismic image. This work flow is often pieced together using disparate software packages for geospatial visualization and statistical analysis. Data extraction and reformatting is often required to move between applications and rudimentary tools such as side-by-side screen shots are unfortunately used for analysis. Expensive commercial software packages offer some integrated options, but these packages are not general purpose or lack the interactivity required to efficiently visualize the data space.

Geophysviz provides a visualization driven approach to seismic attribute analysis. I integrate scatterplots and geospatial displays as interactive linked views in the same application to allow analysts to quickly explore for data anomalies and seamlessly cross-reference the anomalies to geospatial regions. I increase the information density of the seismic image by using a bi-variate colour map to co-render two attributes, and generalize the tool by providing cross plots and co-rendered images of every combination of attributes. An overview display facets the the entire multivariate data set into scatterplot matrices and juxtaposed small multiples of co-rendered seismic images. I use slow animated transitions which encourages the user to visually track how clusters and points move through the multivariate data space. Finally, brushes are used to interactively select and highlight points on the scatterplot, which are then superimposed on the seismic image to allow to for instant geospatial cross-referencing.

This paper gives an overview of previous work then discusses data and task abstraction and visualization design using a general what, why, how framework. Geophysviz effectiveness is analyzed using a sample dataset and typical use case scenario. Strengths, weaknesses, future work, and lessons learned are then discussed. Geophysviz is hosted as an open-source project on github and is currently serving as a public facing web-app at <http://www.geophysviz.appspot.com>.

## 2 PREVIOUS WORK

There are several existing commercial software packages for seismic attribute analysis, the two most popular being Schlumberger's Petrel and CCG's Hampson-Russel. Both these packages focus on two specific attributes and do not generalize to the entire multivariate attribute space. Viewing geospatial referencing of attributes is poorly linked to the actual seismic image and requires generating and loading new data sets. The software has nested menus and controls for navigation which requires the analyst to switch focus from the data. The work-flow for analyzing two attributes using Petrel is documented in the tutorial video [2], which can serve as a benchmark for user experience.

The idea to co-render attributes in the seismic image was inspired by a blog post by Evan Bianco [4], but is also documented in the geophysical literature [6]. Each seismic attribute is designed to exploit different properties, for example similarity will highlight faults and fractures while the attenuation ( $Q$ ) attribute can indicate the presence of a fluid. Co-rendering two attributes can aid in delineating structure and facies in a seismic image.

The literature on seismic attribute visualization is sparse, but there has been much work in generalized multivariate visualization. Burger and Hauser [7] provide an overview of the current state of the art and describe a general set of tools for effective visualization of multivariate data. In addition to an abstract pipeline for scientific visualization, the authors suggest the following general approaches: Derivations, glyphs, hybrid/multi-method visualization, interaction, layering and fusion, two level rendering, n-D viewing, probing, reduction, machine learning. Not all of these approaches are valid for seismic interpretation, but the design of Geophysviz is heavily based on this framework. Seismic attributes themselves are already derivations, and the dimensionality of the dataset does not require abstraction to level of glyphs. Hybrid/multimethod visualization is the application of several visualization techniques to the same image, which I apply by displaying the same seismic image created from different seismic attributes. Geophysviz is highly interactive, which is described as the most important tool for understanding complex data. I use interactive summary views for data navigation, animated transitions for the movement between detail views, and selection brushes that link multiple displays. I make use of layering and fusion by combining two attributes into a bivariate colour map and superimposing selected targets in a semi-transparent layer. Multiple views encourage comparison and can give contrast to help generate correct understanding. I make heavy use multiple views by faceting the entire multivariate dataset in overview displays consisting of small multiples. Although not currently implemented in Geophysviz, larger seismic datasets will likely require probing, where a smaller subset of the data can be isolated based on user selection. A machine learning approaches to seismic attribute visualization are discussed in the Future work section.

Seismic attribute analysis is closely linked to hyper-spectral imaging, as they both involve analysis of high dimensional geospatial data. The commercial software package ENVI [1] by exelisvis provides hyper-spectral visualization and automated image segmentation, while packages Gerbil [8] and MultiSpec [3] provide open-source options. Ideas on general visualization and interactivity can borrowed from the hyper-spectral community, however there are fundamental differences between hyper-spectral images and seismic attributes. Hyper-spectral data is considered to be multidimensional, where the spectral bands are analyzed independently from each other. Classifications are made directly on the spectral curves corresponding to each pixel. Seismic attributes on the other hand are multivariate, where interpretation requires analysis of relationships between attributes. In addition to display a multidimensional image, geophysviz also needs to display multivariate information between attributes.

## 3 DATA ABSTRACTION

Seismic data is generated by using an impulsive source to create a wavefield in the Earth, which is sampled by a grid of receivers at the surface. Each receiver records 1D timeseries data called a trace, which is measurement of reflected energy vs depth at a given spatial position. The collection of traces form a seismic volume, shown in Figure

4, where cross-sections through the volume are called seismic images. A seismic image can reveal structure about where reflection interfaces are, but more information is required to determine physical properties about the Earth. The seismic volume is processed to create seismic attributes, which are the same dimension as the seismic volume. Attributes can themselves be plotted as an image, or plotted against each other as crossplots. Patterns and trends in crossplots can be heuristically linked to physical properties of the Earth.

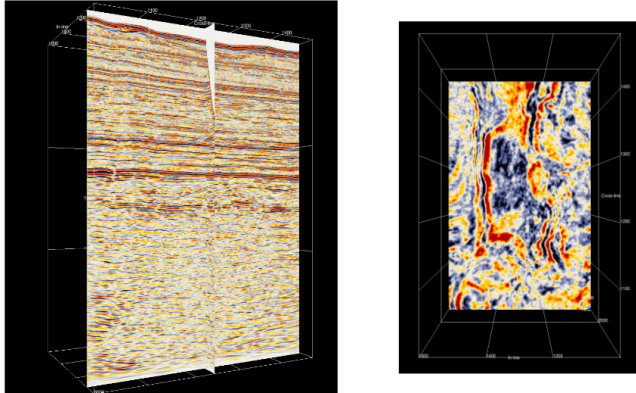


Fig. 4. Example of a seismic data set and seismic images formed by taking cross-sections.

Abstractly, seismic data is a 3D static spatial scalarfield where each cell corresponds to the reflected energy at a latitude, longitude, and depth. Seismic attributes are derived from the seismic data, and the full seismic attribute dataset form a multivariate 3D tensor field where each cell contains the derived seismic attribute values. Analysis is performed on a 2D cross-section of the attribute dataset, which is a 2D multivariate tensor field. When seismic attributes are analyzed independently of their spatial position, the dataset can be flattened into a list of items, where each item contains the attribute values from a given cell.

#### 4 TASK ABSTRACTION

The general task of seismic analysis is to find anomalous regions in seismic images that may be an indicator resource reservoirs or something of geological interest. Seismic images alone do not contain enough information to reveal these regions, so seismic attributes are used to aid analysis. Multivariate analysis of attributes is heuristically related physical properties of the Earth, so analysts look for and compare visual trends and anomalous clusters in attribute scatterplots, see Figure 5. These trends and clusters are geospatially cross-referenced to the seismic image to see if they correspond to a region of structural interest. These tasks are abstracted in actions and targets as described in [9].

##### 4.1 Actions

At the highest level, the analyst is consuming the data with the goal of discovering new information and generating hypothesis' about the data. When searching for targets, the analyst does not know precisely where or what they are looking for, so they need to explore the dataset. Once they have found a potential target/targets on a scatterplot they need to identify points related to the target, compare the selected target against other scatterplot views of the multivariate space, and summarize the target on a geospatial view.

##### 4.2 Targets

Generally the targets are geospatially localized multivariate trends, anomalies, and clusters. A potential target would be set items on a single attribute scatterplot. These items would arrange themselves differently on other scatterplots in the multivariate attribute space, and would also correspond to cells in the geospatial grid. A valid target would stand out as an anomalous cluster or trend in a scatterplot, and

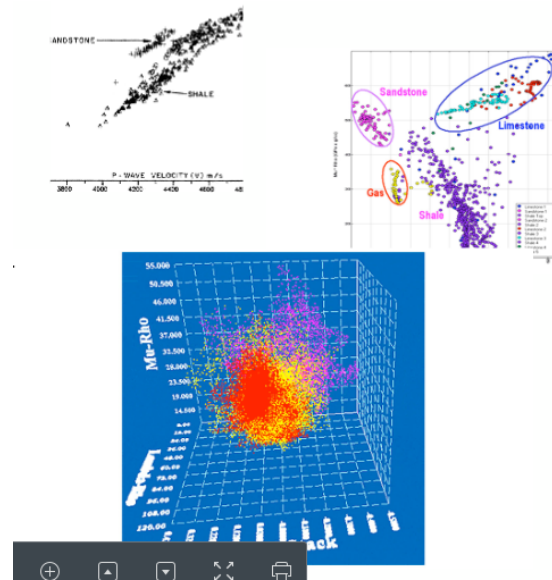


Fig. 5. Examples demonstrating how rocks are classified based on scatterplot trends. Figures originally published in [5] and [6].

also correspond to localized clusters of cells that form a geologically relevant shape in the geospatial plot.

## 5 SOLUTION

The solution focuses on integrating and interacting with five core components: the geospatial image display, the geospatial overview, the scatterplot display, the scatterplot overview, and the linked target selection brush.

### 5.1 Geospatial display

Seismic images are spatial fields that are best displayed as variable density images/heat maps. Where the value of the grid points are mapped to the colour channel via a sequential colour map. I take advantage of the two magnitude colour channels saturation and lightness to optionally co-render two attributes in the same seismic image. This information dense technique allows an analyst to directly locate correlations and similarities between two attributes in one view. I form the co-rendered image from a simple HSL blending algorithm which generates an HSL array from a primary attribute, and modulates the lightness channel by the amplitude of a secondary attribute. For attributes that are +/-, I use a diverging colourmap that moves from blue to white to red for the hues, and use a sequential colourmap with green hues for magnitude data. I do not render images with +/- data as secondary attributes as the lightness channel is magnitude only. The geospatial display panel is shown in Figure 6.

### 5.2 Geospatial overview

The content of the geospatial display, shown in Figure 7, is controlled by the geospatial overview panel, which facets every co-rendering possibility into a matrix of small multiples. The geospatial overview serves the purpose of data navigation as clicking on a small multiple will update the focus of the geospatial display. Geophysviz is built without control panels and menus and uses the data itself as the interface whenever possible. In addition to navigation, the juxtaposed small multiples in the geospatial overview serve as analysis view; the analyst is encouraged to look for contrasts between images. The diagonal column of the matrix is uni-variate data with one attribute plotted in grey scale.

### 5.3 Scatterplot display

The scatterplot display, shown in Figure 8, spatially encodes the attribute values in order to visualize targets consisting of trends, corre-

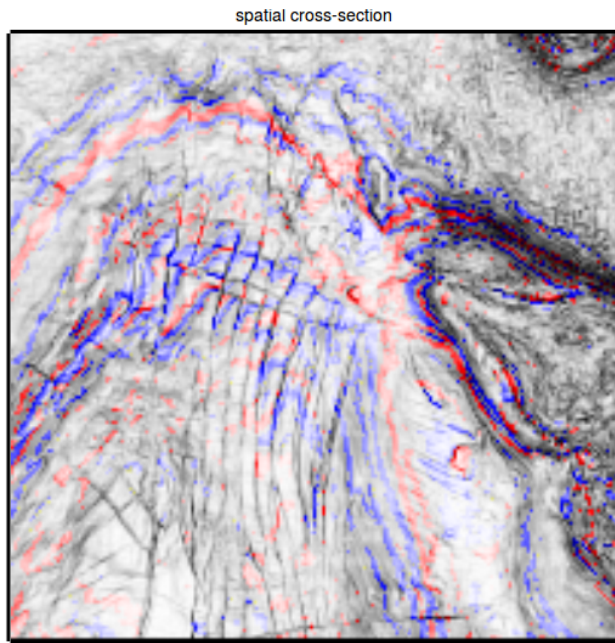


Fig. 6. The spatial cross-section view with amplitude and similarity attributes co-rendered using a bivariate colour map.

lations, outliers, and clusters. The scatterplot is typically a focal point of the attribute analysis workflow and is therefore positioned in the centre of the visualization. As the analyst navigates to another set of attributes, slow transitions are deliberately used to encourage the analyst to track a target as it moves through the multivariate space. The choice of using 2D scatterplots over 3D is to avoid occlusion and ambiguity when using selection brushes.

#### 5.4 Scatterplot overview

The number of attributes in a seismic dataset is typically less than 12, so the entire multivariate scatterplot space can be visualized in a scatterplot matrix, see Figure 9. Similar to the geospatial overview, I use the scatterplot overview data as the navigation interface for changing the focus of the scatterplot display. The scatterplot matrix also serves as an analysis tool by allowing contrast comparisons.

#### 5.5 Target selection brush

All of the views are linked via an interactive selection brush. Analysts use a brush to select a target on the scatterplot view. This highlights the target points in each cell of the scatterplot overview, which encourages side-by-side comparisons. The brush also links to the geospatial display, where the selected targets are superimposed as a semi-transparent mask allowing for instant geospatial referencing and target summary. The target brush in action is shown in Figure 10.

### 6 IMPLEMENTATION

Geophysviz is developed as a web application where data processing and formatting is handled on a cloud-hosted back-end server and visualizations are generated on the client's browser. I chose to use the web app paradigm as analysts can use and demo the tool without downloading and building any software. The back-end data server is written in Python and the visualization is built on top of the D3 charting library in JavaScript.

The back-end is relatively simple, and primarily handles data requests from the client and basic pre-processing such as normalization, image down-sampling for the small multiples, and array masking for the superimposed image. I used the webapp2 framework for URL routing and Google App Engine as a cloud host.

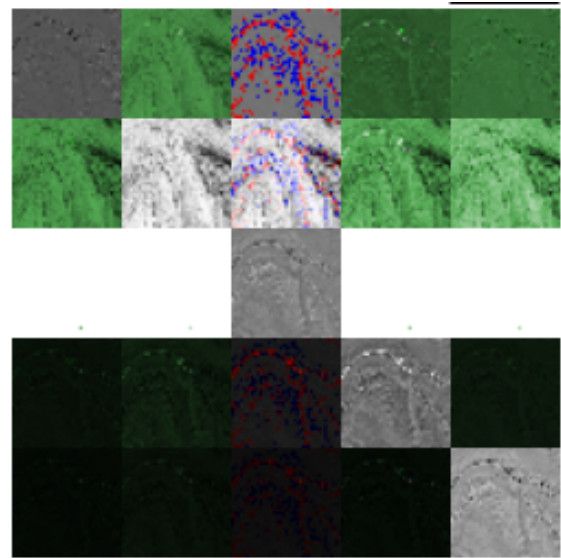


Fig. 7. The overview panel for attribute co-rendering views. Click on an individual plot focuses the data on the geospatial plot.

The client side visualization code uses the html5 canvas for image plots and d3 data bindings to SVG elements for the scatterplots. Interactivity drives AJAX calls to back-end to dynamically update the client-side data. The angularJS framework was used as the client-side architecture.

The source code is open-source and hosted on GitHub.

### 7 RESULTS

I demonstrate the prototype application with the open F3 seismic data set collected offshore New Zealand. The data was processed for amplitude and coherency attributes, while intercept, gradient, and attenuation (Q) attributes were synthesized for the demonstration. A single depth slice image from the dataset was used.

The typical scenario is an analyst trying to find geologically interesting regions and form a hypothesis based on analysis of seismic attributes. Starting from the initial view, the analyst will scan the overview plots and compare the images looking for features to pop out. In the F3 example, the Q vs amplitude scatterplot, Figure 11, shows a particularly interesting trend. Having spotted a potential target, they click on the small plot to bring the data into the detail view. Using knowledge of the local geology, the analyst would expect a parabolic relationship between these two attributes, but there is a cluster of points displaying anomalous behaviour. At this point the analysts will spend time free playing, switching between detail views and watching how the target points rearrange themselves through the multivariate attribute space. Noticing that the target points are all on one side of the similarity vs Q scatterplot, they query the target using the interactive selection brush. See Figure 12. Selecting a target on the scatterplot immediately highlights the affected points on each plot in the scatter matrix, shown in Figure 13. The analyst can then compare the juxtaposed plots to see if the selected target also forms trend of interest in other scatterplot views. This particular target shows anomalous behaviour on the intercept vs gradient scatterplot, which can often be an indication of the presence of hydrocarbons. See Figure 14. Once a target has been analyzed in the multivariate scatterplots, the analyst will focus their attention on the image plot. The selection brush is linked to the image plot, where points from the selected target are superimposed as a semi-transparent mask. Ideally a valid target will form a geologically relevant shape on the image plot, as shown in 15. The analysts can then explore attribute co-rendering options by clicking on the small multiple images. If the overlaid target pops out on particular attribute co-rendering views, the analyst can make an in-

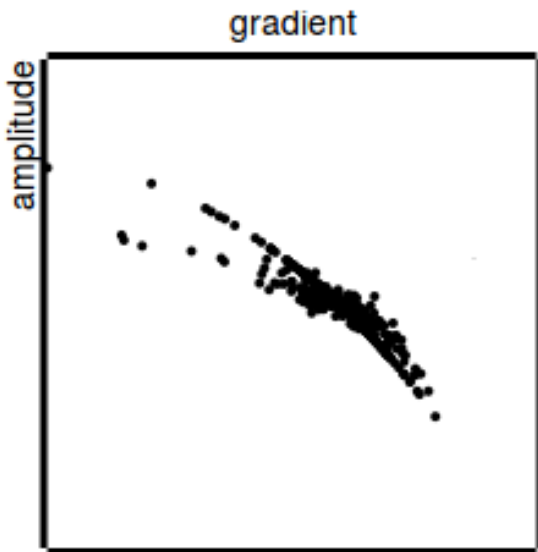


Fig. 8. Detail view of a scatterplot of intercept and amplitude attributes.

terpretation about the geological significance. In this example, the amplitude-similarity co-rendering, shown in 16, is the most interesting. Analysts may wish to further assess the highlighted region as a potential drilling target.

I recently demoed the application to a small group of geophysicists who were excited by the design of the tool and the open-source licencing. The main limitation with the tool is scalability to industry sized problems and handling multiple input file formats. There were also requests for the ability to select multiple targets as well have more free form selection brushes.

## 8 DISCUSSION AND FUTURE WORK

I applied principals of information visualization to address a broken workflow is exploration geophysics. I developed an integrated and interactive approach to seismic attribute analysis which used linked views and faceted small multiples to let analysts explore the multidimensional, multivariate data space.

The greatest strength is the information density and interactivity of the display. The overview of juxtaposed small multiples allows the user to see the entire multivariate data space in one view. The interactivity allows the analyst to quickly focus on targets to form and test hypothesis. The interactivity is data-driven, so the analyst can change views and make selections by clicking and selecting the data rather using controls outside the visualization. The data as interface design choice lets the user focus on the data analysis task without having switch to a control panel to change the view.

The data chosen for the demonstration was a manageable size, but interactive scatterplot displays will be a challenge with larger data sets. The d3 charting library provides easy methods for interacting with SVG elements, but this approach quickly becomes inefficient as the number of points becomes large. For larger seismic images a more efficient plotting approach using low-level graphics card libraries like OpenGL or WebGL will be a necessity. Additional approaches to embedding data and down-sampling may be required to display larger seismic images.

Choosing adequate colour maps for each possible co-rendering is another difficult challenge that has not been fully addressed. I made simple design choices of colour maps, where I normalized data and used diverging and sequential maps where reasonable. Attributes have a large spread of values and dynamic ranges, so each combination of co-rendering requires attention to colour map detail. The current implementation has some very effective colour maps such as the amplitude-similarity plot, but others require more attention.

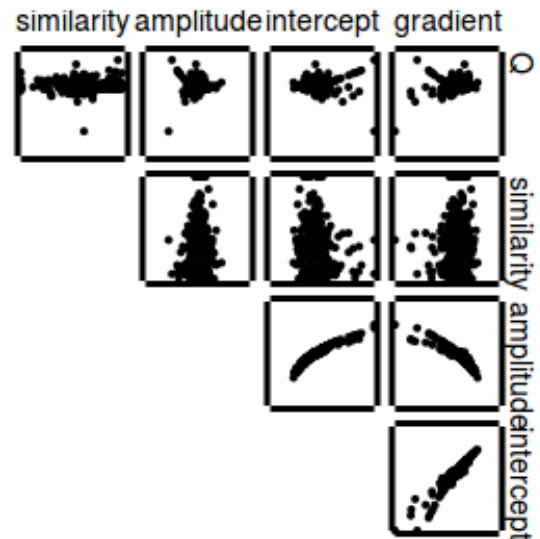


Fig. 9. Scatterplot matrix of over viewing all attribute combinations. Clicking on an individual plot changes the focus of the scatterplot display.

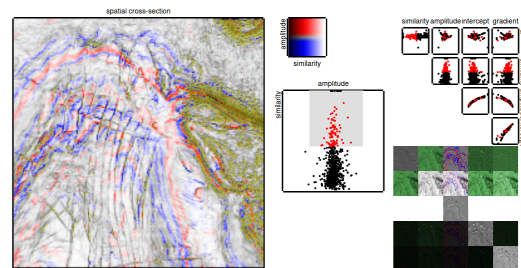


Fig. 10. Using the target selection brush to highlight a target on multiple plots.

Although I designed Geophysviz for data consumption tasks, I can extend the functionality to include data production tasks. After the analyst has explored the data and generated hypothesis, they may wish to create annotations and classifications of the targets. The current tool allows selection of just one target at a time, but I can foresee many situations where an analyst would save selections and generate a collection of multiple targets. The produced output of the tool could be a segmented seismic image where the segments are derived from attribute analysis.

Seismic attribute analysis consists of finding clusters and patterns in multivariate data and looking for textured facies in images, both tasks can be aided by machine learning algorithms. Unsupervised clustering algorithms can be applied to the attribute dataset and the analyst would use geophysviz to vet the results. As analysts select target and verify targets, a database of labelled can be formed which can train a classifier in a supervised learning context. Machine learning would likely not replace the analyst, but it can serve as aid in assisting them analyze large datasets.

The biggest lesson I learned during this project is to display as much data in the visualization as possible. The first iteration of the concept consisted of only detailed views and a control panel for navigation and interaction. Replacing the control panel with overview displays which themselves drive interactivity and navigation greatly increased the information density of the display. Using the data as the interface allows the analyst to interact intuitively while maintaining context.

The focus of the design was on seismic attributes, but the application generalizes to any multidimensional geospatial data sets. Immedi-

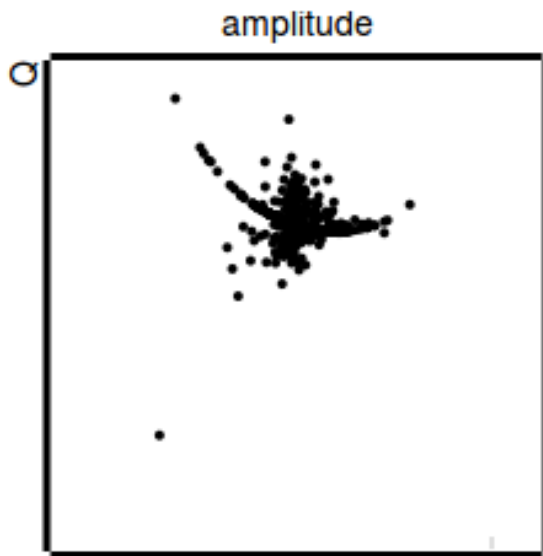


Fig. 11. Anomalous behaviour in the amplitude vs Q scatterplot. The cluster of points in the middle does not obey the expected trend.

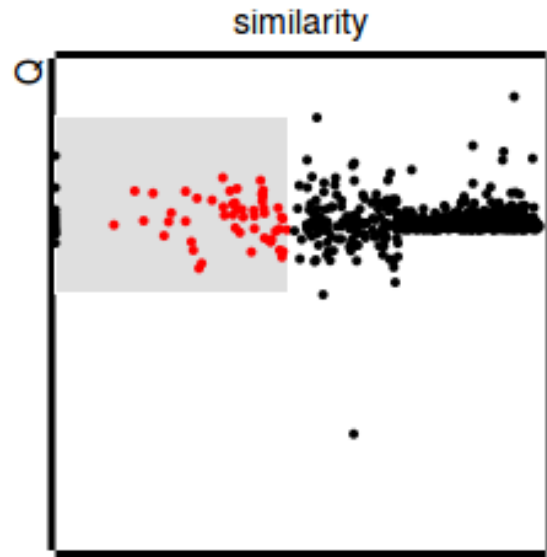


Fig. 12. A cluster have points transitions entirely to left side of a similarity vs Q scatterplot.

ate applications are hyper-spectral image analysis and remote sensing. More abstractly the tool can be used to visualize features in image classification and segmentation problems.

## 9 CONCLUSION

I presented Geophysviz, a data visualization tool for exploring multidimensional, multivariate geospatial data. The tool was successfully demonstrated on the F3 seismic data set and received general positive feedback from geophysical interpreters. Although useful for small data sets, future work is required to scale to scatterplots of many points.

## REFERENCES

- [1] Itt visual information solutions: Envi. <http://www.itvis.com/ProductServices/ENVI.aspx>, Jun 2010.
- [2] Easy seismic and avo analysis. <https://www.youtube.com/watch?v=VrMYiRNF9G> Sept 2013.
- [3] Multispec: A freeware multispectral image data analysis system. <https://engineering.purdue.edu/biehl/MultiSpec/index.html>, feb 2015.
- [4] E. Bianco. Corendering attributes and 2d colourmaps. <http://www.agilegeoscience.com/blog/2015/5/27/corendering-attributes-and-2d-colourmaps>, May 2015.
- [5] J. P. Castagna, M. L. Batzle, and R. L. Eastwood. Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. *Geophysics*, 50(4):571–581, 1985.
- [6] S. Chopra. Interpreting fractures through 3d seismic discontinuity attributes and their visualization. *CSEG Recorder*, 34(8):99–108, Oct. 2009.
- [7] R. Fuchs and H. Hauser. Visualization of multi-variate scientific data. *Computer Graphics Forum*, 28(6):1670–1690, 2009.
- [8] J. Jordan and E. Angelopoulou. Gerbil - a novel software framework for visualization and analysis in the multispectral domain. In *VMV 2010: Vision, Modeling and Visualization*, pages 259–266, Nov. 2010.
- [9] T. Munzer. *Visualization Analysis and Design*. CRC Press, 2014.
- [10] Wikipedia. Seismic attribute — Wikipedia, the free encyclopedia, 2015. [Online; accessed 12-Dec-2015].

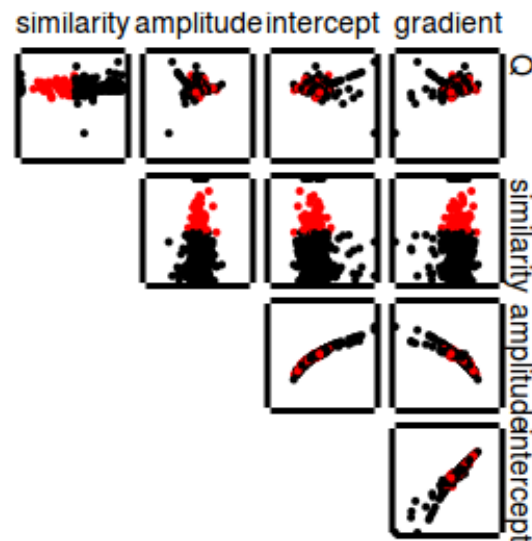


Fig. 13. Selecting the cluster of points using the target selection brush.

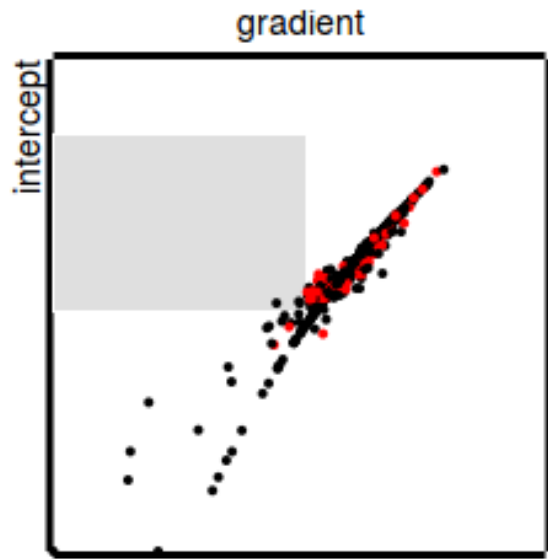


Fig. 14. Bimodal behaviour in the intercept gradient plot. The highlighted trend could be a hydrocarbon indicator.

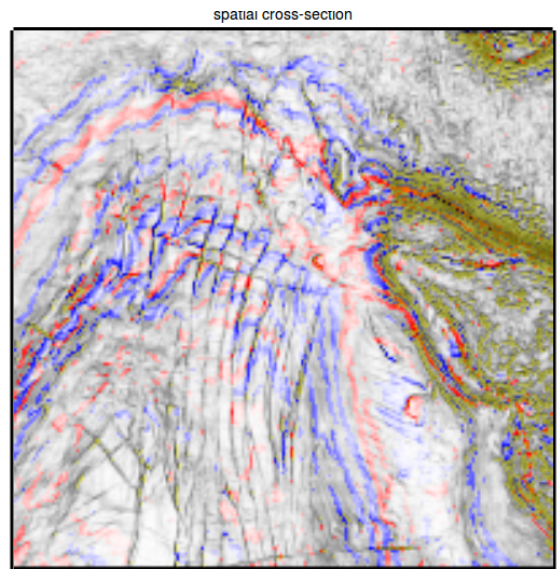


Fig. 16. The selected target superimposed on an amplitude-similarity co-rendered geospatial plot. The selected region corresponds to features in the image.

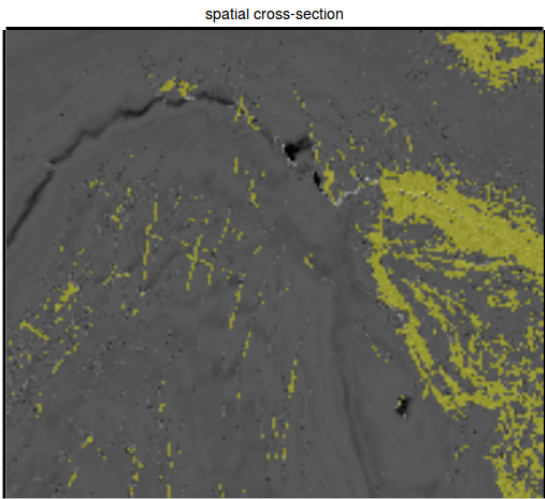


Fig. 15. The selected target superimposed on the geospatial plot of the Q attribute. The selected target is geospatially localized in a geologically relevant shape.