Geophyzviz: A visualization approach to multivariate analysis of seismic attributes

Ben Bougher

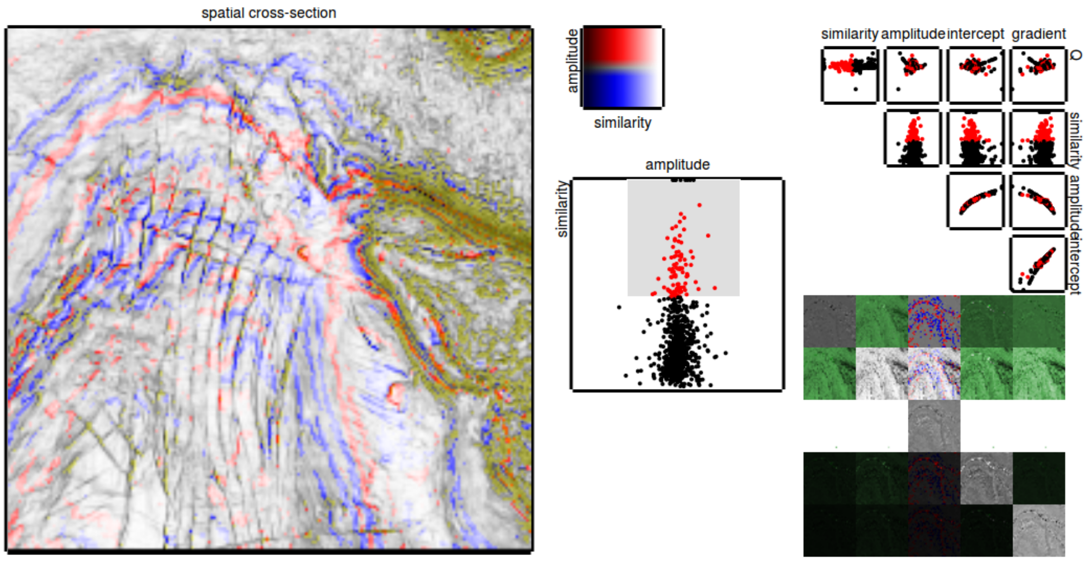

Fig. 1. Geophyzviz: Interactive attribute analysis.

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. We present Geophyzviz, a system that uses linked geospatial and scatter plot views to support visual exploration of high-dimensional multivariate geospatial data. Geophyzviz has two primary detailed views: a scatter plot 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 scatter plot, 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. We present the application of Geophyzviz 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. 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 modulated 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. 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[wiki]. The zoo of specific seismic attributes is large and outside the scope of this paper, so we 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 scatter plots 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 region in the seismic image. This work flow is often pieced together

• *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.xxxx/TVCG.201x.xxxxxxx/

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.

Geophyzviz provides a visualization driven approach to seismic attribute analysis. We integrate scatter plots 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. We 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 scatter plot matrices and juxtaposed small multiples of co-rendered seismic images. We 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 scatter plot, which are then superimposed on the seismic image to allow to for instant geospatial cross-referencing. Geophyzviz is hosted as an open-source project on github and currently serving as a public facing web-app at http://www.geophyzviz.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 CGG'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 [4], 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 [3], but is also documented in the geophysical literature [5]. 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 lets the analyst looks for interesting regions that would not pop out if each attribute was displayed independently.

The literature on seismic attribute visualization is sparse, but there has been much work in generalized multivariate visualization. Burger and Hauser [6] provide an overview of the current state of the art and describe a general set of tools for effective visualization of multivariate data. Geophyzviz makes use all of the tools that are relevant to seismic attributes. Geophyzviz is heavily interactive, uses hybrid-rendering to co-render attributes, and uses layering and fusion to superimpose selected attributes over the seismic image. In the language of the Burger and Hauser, the linked brushes used in Geophyzviz follow a Simviz approach to focus and context handling.

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 [7] and MultiSpec [2] 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, geophyzviz also needs to display multivariate information between attributes.

Table 1. Data abstraction

Year	Submitted	Accepted	Accepted (%)
1994	91	41	45.1
1995	102	41	40.2
1996	101	43	42.6
1997	117	44	37.6
1998	118	50	42.4
1999	129	47	36.4
2000	151	52	34.4
2001	152	51	33.6
2002	172	58	33.7
2003	192	63	32.8
2004	167	46	27.6
2005	268	88	32.8
2006	228	63	27.6

includegraphics[width=1.5in]sample

Fig. 2. Sample illustration.

3 Data abstraction

A seismic volume is formed from a collection of seismic traces, where each trace corresponds to reflected energy vs depth at a given spatial position. The traces are processed to form attributes, which are are the same dimension as the seismic volume. Analysts typically look at cross-sections of the volume called seismic images. Attributes can themselves be plotted as an image, or plotted against each other as a scatter plot.

A seismic image is a snap shot of a continuous wave field, so we use a gridded field idiom for the for image visualization. Each point in the seismic volume can be abstracted as an item which contains attributes. The attributes are the geospatial position of the point as well as the derived seismic attributes.

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. Scatter plots of seismic attributes can reveal visual trends and anomalous clusters which can be detected and further analyzed. Points of interest are compared in other attribute plots and hypothesis are formed using the analyzed a priori information of trend classifications. The selected points are geospatially cross-referenced to the seismic image to see if they correspond to a region of structural interest.

At the highest level, the analyst is consuming the data in order discover new information and generate hypothesis' from the seismic data. Although the analyst has some a priori knowledge of patterns to search for, primarily both the type of target and location are unknown so the analyst needs to explore the data. Once the targets have been identified on a single scatter plot, the targets are then compared in other scatter plots. The targets are then summarized by viewing their geospatial locations. There are many types of targets the analyst is looking for as regions of interest manifest as correlations, outliers, trends, and features depending on the geological situation and the attributes measured.

5 Solution

The solution focuses on integrating and interacting with two main views, the geospatial image display and the scatter plot display.

Seismic images are typically displayed in variable density images, or heat maps where the value of the grid points are mapped to the colour channel via a sequential colour map. We take advantage of the magnitude channels saturation and lightness to optionally co-render two attributes in the seismic image. This information dense technique allows an analyst to directly locate correlations and similarities between two attributes in one view. All possible co-rendering combinations are faceted into juxtaposed small multiples in an overview display, which encourages comparisons and data navigation. We used diverging colour maps for +/- attributes and sequential colour maps for magnitude attributes.

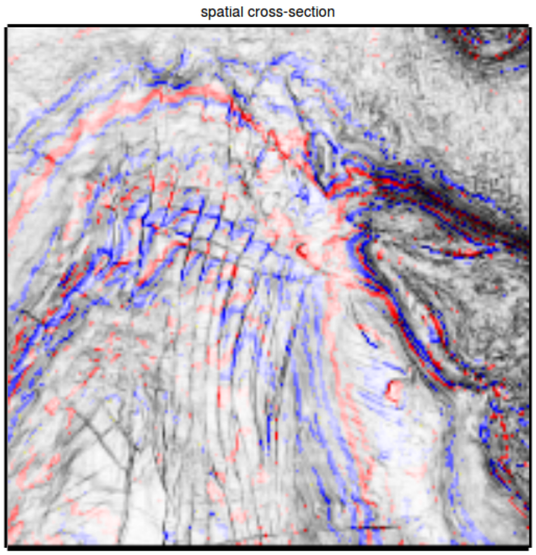

Fig. 3. The spatial cross-section view with amplitude and similarity attributes co-rendered using a bivariate colour map.

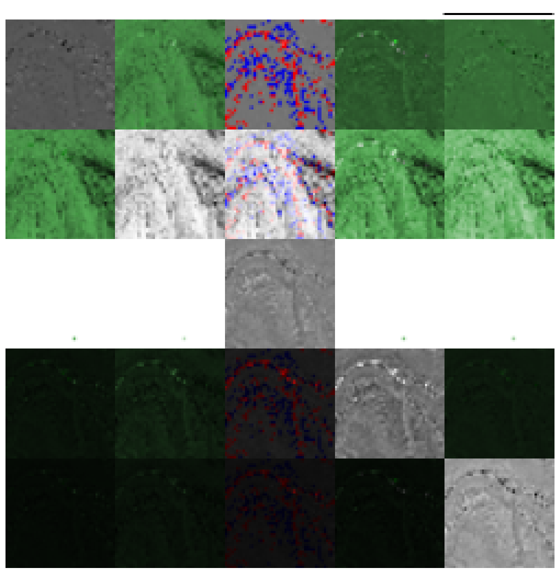

Fig. 4. The overview panel for attribute correndering views.

The scatter plot display spatially encodes the attribute values in order to visualize trends, correlations, outliers, and clusters. There is typically ¡ 10 attributes in a data set, so the choice to display all possible scatter plots in a scatter plot matrix is a logical decision. The scatter plot matrix also serves as an overview, where clicking a single scatter plot changes the focus of the detailed view. The transitions between detailed scatter plot views is chosen to be a noticeably slow animation to encourage the analyst to track targets as they move between plots.

All the views are linked via an interactive selection brush. Analysts use a brush to select targets on the detailed scatter plot view. This highlights the targets in all of the scatter plots which allows for side by comparisons. The brush also links to the image display, where the

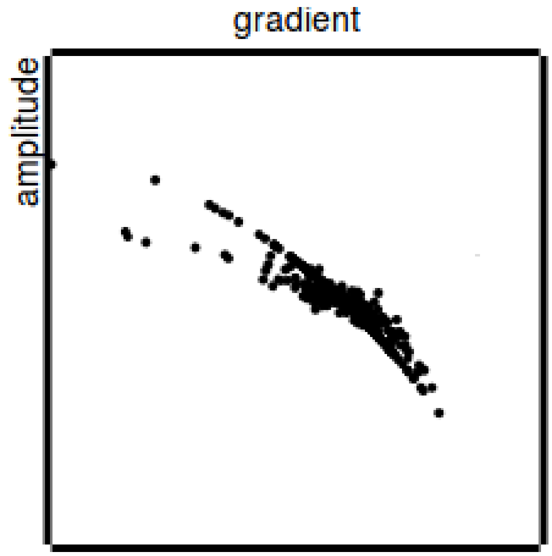

Fig. 5. Detail view of a scatter plot of intercept and amplitude attributes.

selected targets are superimposed as a semi-transparent mask allowing for instant geospatial referencing and target summary.

6 Implementation

Geophyzviz 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 clients browser. The back-end data server is written in python and the visualization is built on top of the d3 charting library in java script. Other than generic external libraries, all code was written by the primary author.

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. The Google App Engine service was used as the server framework and cloud host.

The client side visualization code uses the html5 canvas for image plots and d3 data bindings to SVG elements for the scatter plots. The co-rendered image is formed from a simple HSL blending algorithm where an HSL array is generated from one attribute, and the lightness channel is modulated by the amplitude of the second attribute. Interactivity drives AJAX calls to back-end to dynamically update the client-side data.

7 Results

We 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.

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. Once they notice something of interest, they click on the small plot to bring the data into the detail view. At this point the analysts will spend time free playing, switching between detail views and watching how points rearrange themselves through the multivariate attribute space. Once they have explored the space searching for targets of trends, features, outliers, and clusters, they can query a target using the interactive selection brush. Selecting a target on the scatter plot immediately highlights the affected points on the plots the in the scatter matrix. The analyst can then compare the juxtaposed plots to see if the selected target also forms trends in other scatter plot views. Once a target has been analyzed in the multivariate scatter plots, the analyst will focus their

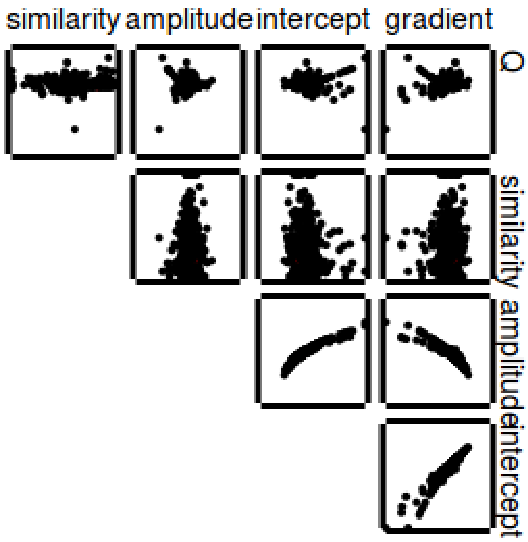

Fig. 6. Scatterplot matrix of overviewing all attribute combinations.

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. The analysts can then explore attribute co-rendering options by clicking on the small multiple images. If a target pops out on particular attribute co-rendering views, the analyst can make an interpretation about the geological significance of the target.

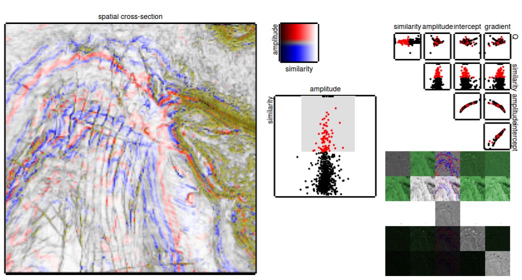

Fig. 7. Analysis of an anamolous region.

8 Discussion and future work

We applied principals of information visualization to address a broken work-flow is exploration geophysics. We developed an integrated and interactive approach to seismic attribute analysis which used linked views and faceted small multiples to let analysts explore the multi-dimensional, mutlivariate 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 the 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 scatter plot 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. We made simple design choices of colour maps, where we 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, but others require more attention.

The biggest lesson I learned during this project is to use the data as the interface. 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.

Although we designed Geophyzviz for data consumption tasks, we 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 we can foresee many situations where an analyst would save selections and generate a collection of multiple targets. The produced output of the tool would would be a segmented seismic image where the segments are derived from attribute analysis. It would relevant to perform automated clustering and let the analysts use the tool for quality assurance of the clustering algorithm.

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

We presented Geophyzviz, 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 scatter plots of many points.

References

[1] Itt visual information solutions: Envi. http://www.ittvis.com/ProductServices/ENVI.aspx, Jun 2010.

[2] Multispec: A freeware multispectral image data analysis system. https://engineering.purdue.edu/ biehl/MultiSpec/index.html, feb 2015.

[3] E. Bianco. Corendering attributes and 2d colourmaps. http://www.agilegeoscience.com/blog/2015/5/27/corendering-attributes-and-2d-colourmaps, May 2015.

[4] Cegal. Easy seismic and avo analysis. https://www.youtube.com/watch?v=VrMYiRNF9GQ, Sept 2013.

[5] S. Chopra. Interpreting fractures through 3d seismic discontinuity attributes and their visualization. *CSEG Recorder*, 34(8):99–108, Oct. 2009.

[6] R. Fuchs and H. Hauser. Visualization of multi-variate scientific data. *Computer Graphics Forum*, 28(6):1670–1690, 2009.

[7] E. A. J. Jordan.