Computers & Geosciences 79 (2015) 47–57

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Multiscale hierarchical domaining and compression of drill hole data

E.J. Hill*, J. Robertson, Y. Uvarova

CSIRO, PO Box 1130, Bentley, WA 6102, Australia

ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form
9 March 2015
Accepted 9 March 2015
Available online 11 March 2015

Keywords:
Wavelet transform
Tessellation
Lithochemistry
Filter
Logging

ABSTRACT

New drilling methods, currently under development for minerals exploration, combined with rapid data collection by a range of sensors means that the end-user is confronted with increasingly large data sets. In order to reduce the stream of data into objects which represent meaningful geological features we need to incorporate spatial information into the analysis. Boundary detection incorporating multiscale considerations has previously been carried out using a scale–space plot from a wavelet transform. We present a method which applies a rectangular tessellation to the wavelet transform, this has the advantage of being easier to interpret, as it resembles a geological log. In addition, the tessellation records hierarchical information for different scale objects. The tessellation can be filtered in order to remove unwanted variation (including noise) from the results. When applied to geochemical data, the resulting tessellation provides a basis for classification of lithochemical units that is more reliable than classification by considering individual samples without spatial context.

1. Introduction

Discoveries of mineral resources are declining in Australia and future exploration is being targeted at deep basement rocks that lie beneath barren sedimentary cover. In order to facilitate the collection of geological information on deep basement rocks, low cost, rapid drilling methods are being developed (Hillis et al., 2014). Traditionally, geological information is extracted from the drilling products by visual logging by trained geologists. However, this method is slow and expensive, and more importantly, there is a lack of confidence in the consistency of logging between geologists.

As a result, automated methods of collecting data at the drill site are being developed; designed to be cheaper and faster than sending samples to an analytical laboratory. These include collection of geochemical and mineralogical data from portable X-ray fluorescence spectrometer (XRF) and X-ray diffraction (XRD) devices ("Top-of-hole sensing", Hillis et al., 2014). Converting the streams of geochemical and mineralogical data into geological logs provides a useful visual summary of the information. Geological logs can be used to document important lithological boundaries, classify lithological units and recognise corresponding lithological units in different drill holes in order to build up a 3D model of the local geology.

Traditional methods of classification and discrimination of lithological units based on geochemical data generally do not take spatial information into account (see for example, Mullen, 1983) and this may result in spurious lithological boundaries when the classification is plotted as a down hole geological log (see examples in next section).

A different approach has been taken by geophysicists using downhole geophysical tools such as gamma-ray and resistivity logs, where signal processing techniques have been applied (replacing the time axis with depth) so that the spatial relationships between data points are preserved. In particular, wavelet transforms are popular due to their mathematical efficiency and multiscale-edge detection ability (e.g. Panda et al., 2000; Arabjamaloei et al., 2011; Perez-Munoz et al., 2013). In this paper we propose an alternate method for visualising the wavelet transform of geological data using the wavelet tessellation method of Witkin (1983). The advantage of this method is that it provides a diagram that is similar in appearance to a geological log, and therefore a familiar format to geologists, it preserves the hierarchical multiscale information provided by the wavelet transform and it is a very compact summary of the data.

1.1. Deriving lithochemical units from geochemical data

When a geologist manually logs a drill hole they usually incorporate spatial relationships in their interpretation, in particular, the concept that samples that are in close proximity may be grouped into a single class even if they have somewhat different characteristics. When lithological units are classified using only

---

* Corresponding author.
E-mail addresses: June.Hill@csiro.au (E.J. Hill), Jesse.Robertson@csiro.au (J. Robertson), Yulia.Uvarova@csiro.au (Y. Uvarova).

http://dx.doi.org/10.1016/j.cageo.2015.03.005
0098-3004/

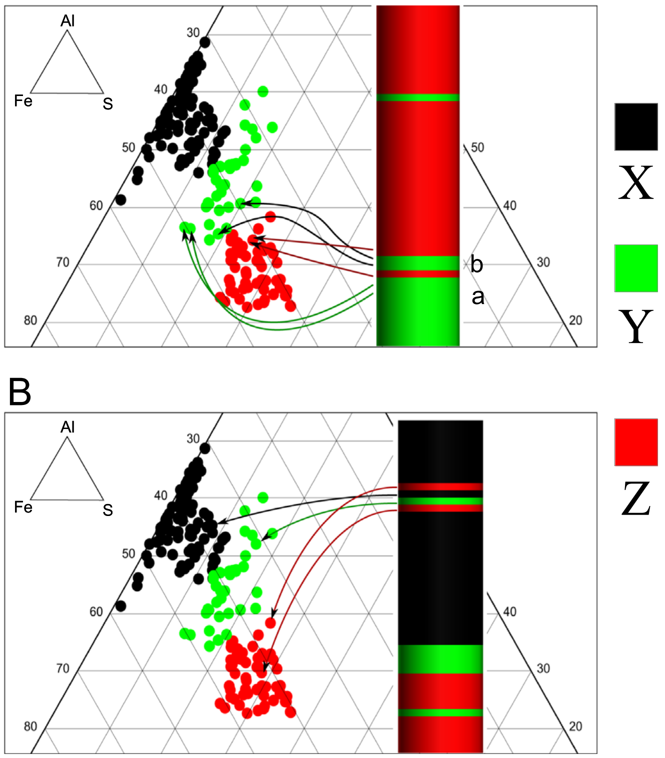

Fig. 1. Data from DETBrukunga02 classified into three groups according to location on a compositional ternary plot (Fe–Al–S). The column plot shows the drill hole intervals coloured according to the same classification scheme. Selected points on the drill hole are linked to their location on the ternary plot. In (A) the section of unit *Y* in the column plot labelled "a" is compositionally distinct from the surrounding unit *Z*, but the section of unit *Y* labelled "b" is very similar to the surrounding unit *Z*. In (B) a narrow section of unit *Y* indicated on the plot may be a mixed sample of the units *X* and *Z*.

non-spatial features spurious lithological boundaries may result. Two examples of which are illustrated in Fig. 1, which shows a ternary plot of geochemical data from an experimental drill hole from the Deep Exploration Technologies Cooperative Research Centre (DET CRC) project in Brukunga, South Australia. In the figure hard boundaries have been assigned to the ternary plot in order to separate different lithological units (without reference to any spatial information). These lithochemical classifications can be plotted in a column to represent the lithological units in the drill hole. Spurious lithological boundaries can result from a unit whose chemical variation crosses a classification boundary. Fig. 1A shows an example of this problem: if the spatial location had been taken into account the unit labelled "b" would probably have been assigned to the same rock type as the surrounding units due to their similarity in composition. Spurious lithological boundaries can also occur when a sample is a mixture of two rock types. For example, in Fig. 1B it is possible that the unit classified as unit *Y* may actually be a mixture between adjacent *X* and *Z* units.

To correct this spatial data problem we propose to first determine the boundaries of the lithological units and then classify the units by using all the data points between a pair of boundaries as a single lithological unit rather than classifying each data point separately. We do not address the problem of classification in this paper, only the problem of boundary identification.

When logging a drill hole geologists do more than simply detect boundaries and classify units, they also make decisions regarding splitting and grouping of adjacent lithological units (see for example, the *combining* versus *separating* and *single* versus *many* dichotomies documented by Shipley et al., 2013). The degree to which units are grouped depends on the scale of the study. It is desirable to be able to summarise the data in a manner that not only preserves information at different scales (so that it can be extracted as required) but also summarises the hierarchical structure of the higher level groupings. The tessellation method of Witkin (1983) was chosen as it provides both multiscale and hierarchical information in a succinct format. Furthermore, we propose a user-controlled filtering procedure that allows undesirable low level variation to be removed from the tessellated log as required.

1.2. *Multiscale edge detection*

Edge detection algorithms are used to identify points of sharp variation in a measured signal. These edges (or boundaries) separate regions of low variation, which can be interpreted to represent objects in the signal. Edge detection algorithms have been widely used in image analysis to help identify objects in images; however, traditional edge detecting algorithms work at a predefined scale (e.g. Marr and Hildreth, 1980; Canny, 1986). This means that the scale of edge detection has to be determined before the calculation proceeds. If too small a scale is selected, then we may have a very large number of edges, many of which may represent noise or information of no interest at the scale of the problem under consideration. If, on the other hand, we select a scale that is too large, we may lose useful information about interesting small scale features. Using a continuous wavelet transform to detect edges has the advantage that it works simultaneously across a range of scales. Multiscale edge detection methods for images using the continuous wavelet transform were developed by Mallat (1991), Mallat and Hwang (1992) and Mallat and Zhong (1992). Their aim was to be able to compress image files by removing unwanted noise and preserving important edges. This is directly analogous to deriving geology logs from analytical data, where the geologist only wishes to detect important changes in the measured property and ignore weak variations.

The wavelet transform has been used for detecting significant lithological boundaries in geophysical log data (typically from petroleum or groundwater wells). These include applications to gamma ray, sonic and resistivity logs (e.g. Arabjamaloei et al., 2011; Cooper and Cowan, 2009; Davis and Christensen, 2013). In the petroleum industry, the technique is referred to as blocking or zoning, the aim of which is to smooth the log data in regions of relatively homogeneous rock properties (i.e. regions of weak variation) while leaving the sharp boundaries (i.e. points of strong variation). A comprehensive overview of recent applications in the use of the wavelet transform for analysing well log data is given in Perez-Munoz et al. (2013).

A wavelet is a small oscillation that is finite in extent and has a specified frequency. The wavelet transform is a convolution of the wavelet with the signal. The wavelet can be translated in space (along the signal) and stretched (to operate at different scales). There are a large number of different wavelet shapes which are suitable for different purposes. For example, Chandrasekhar and Eswara Rao (2012) performed a comparison study using a number of wavelet transforms, which demonstrated that the Gaussian wavelet transforms were the most useful in delineating lithological boundaries from geophysical logs.

We use the Gaussian wavelet transform (known as the DoG or Hermitian wavelet) for edge detection as it has certain properties which make it favourable for this application, these are discussed further in Section 2. The second order (i.e. second derivative) DoG is also known as the Mexican hat wavelet. Using the DoG wavelet as an edge detector is equivalent to applying a Gaussian smoothing function to the signal at various scales and then detecting inflection points in the smoothed signal using the derivatives of the signal. Inflection points are interpreted to represent edges of