Mapping the distribution of ferric iron minerals on a vertical mine face using derivative analysis of hyperspectral imagery (430–970 nm)

Richard J. Murphy *, Sildomar T. Monteiro

Australian Centre for Field Robotics, The Rose Street Building J04, Department of Aerospace, Mechanical & Mechatronic Engineering, The University of Sydney, NSW 2010, Australia

A R T I C L E   I N F O

Article history:
Received 26 August 2011
Received in revised form 17 July 2012
Accepted 19 September 2012

Keywords:
Mining
Iron ore
Remote sensing
Hyperspectral
Derivative analysis
Banded iron formation

A B S T R A C T

Hyperspectral imagery is used to map the distribution of iron and separate iron ore from shale (a waste product) on a vertical mine face in an open-pit mine in the Pilbara, Western Australia. Vertical mine faces have complex surface geometries which cause large spatial variations in the amount of incident and reflected light. Methods used to analyse imagery must minimise these effects whilst preserving any spectral variations between rock types and minerals. Derivative analysis of spectra to the 1st-, 2nd- and 4th-order is used to do this. To quantify the relative amounts and distribution of iron, the derivative spectrum is integrated across the visible and near infrared spectral range (430–970 nm) and over those wavelength regions containing individual peaks and troughs associated with specific iron absorption features. As a test of this methodology, results from laboratory spectra acquired from representative rock samples were compared with total amounts of iron minerals from X-ray diffraction (XRD) analysis. Relationships between derivatives integrated over the visible near-infrared range and total amounts (% weight) of iron minerals were strongest for the 4th- and 2nd-derivative \( R^2 = 0.77 \) and 0.74, respectively) and weakest for the 1st-derivative \( R^2 = 0.56 \). Integrated values of individual peaks and troughs showed moderate to strong relationships in 2nd- \( R^2 = 0.68–0.78 \) and 4th-derivative \( R^2 = 0.49–0.78 \) spectra. The weakest relationships were found for peaks or troughs towards longer wavelengths.

The same derivative methods were then applied to imagery to quantify relative amounts of iron minerals on a mine face. Before analyses, predictions were made about the relative abundances of iron in the different geological zones on the mine face, as mapped from field surveys. Integration of the whole spectral curve (430–970 nm) from the 2nd- and 4th-derivative gave results which were entirely consistent with predictions. Conversely, integration of the 1st-derivative gave results that did not fit with predictions nor distinguish between zones with very large and small amounts of iron oxide.

Classified maps of ore and shale were created using a simple level-slice of the 1st-derivative reflectance at 702, 765 and 809 nm. Pixels classified as shale showed a similar distribution to kaolinite (an indicator of shales in the region), as mapped by the depth of the diagnostic kaolinite absorption feature at 2196 nm. Standard statistical measures of classification performance (accuracy, precision, recall and the Kappa coefficient of agreement) indicated that nearly all of the pixels were classified correctly using 1st-derivative reflectance at 765 and 809 nm. These results indicate that data from the VNIR (430–970 nm) can be used to quantify, without a priori knowledge, the total amount of iron minerals and to distinguish ore from shale on vertical mine faces.

© 2012 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.

1. Introduction

Over the past three decades, advances in hyperspectral imaging have continued to improve our ability to remotely identify, map and quantify geological materials on the Earth’s surface (Bierwirth et al., 2002; Plaza et al., 2009; Vane and Goetz, 1988). Most commonly, data are acquired from airborne platforms, but recent advances in scanning technology have enabled data to be acquired from sensors mounted on field-based platforms (Kurz et al., 2008, 2011; Monteiro et al., 2009). The use of imagery for open-pit mining applications has enormous potential because it enables the identity or abundance of minerals on vertical mine faces to be determined remotely (Fraser et al., 2006; Murphy et al., 2012; Ramanaidu et al., 2002). This is important for reasons of safety and to develop methods to guide or automate the mining process, particularly in large-scale mining operations such as those used in the extraction of iron ore. In this context, methods need to be
evaluated which enable estimation of relative abundances of iron minerals, without a priori knowledge of the distribution of rock types on the mine face. The primary objective of this paper is to quantify the total amounts of ferric iron minerals (hematite + goethite) on an open pit mine face in the Pilbara region of Western Australia, using hyperspectral imagery acquired from a field-based platform. The region is characterised by deposits of Banded Iron Formation (BIF) containing kaolinite-rich shale units of varying thickness (Thorne et al., 2008).

The ferric iron minerals hematite (\(\alpha-FeO\)) and goethite (\(\alpha-FeOOH\)) in the BIF have distinct spectral curves in the visible near-infrared (VNIR) caused by absorptions induced by crystal field transitions at about 465 nm (\(6A_{1g} \rightarrow 4T_{1g}, 4E_g\)), 650 nm (\(6A_{1g} \rightarrow 2T_{2g}\)), and 850–950 nm (\(6A_{1g} \rightarrow 4T_{1g}\); Townsend, 1987). The shape or magnitude of the spectral curve of iron oxides can change due to numerous factors, including replacement of iron by aluminium (Buckingham and Sommer, 1983), particle size variations causing transparence spectral features in different parts of the spectrum and particle shape (reviewed in Cudahy and Ramanaidou, 1997; Ramanaidou et al., 2008). Distribution of iron minerals can be mapped by quantifying the intensity of these VNIR features. Ideally, to quantify intensity of absorption features they should be completely resolved in the spectrum (i.e. the centre and the shortand long-wave boundaries of the feature should be present). To do this, one would require measurement of reflectance between 400 and 1314 nm (the latter being the long wavelength limit of the Fe\(^{3+}\) feature centred at about 850–950 nm). Some researchers do not use the long-wavelength limit of the absorption feature but define the ‘boundary’ of the feature to be at shorter wavelengths i.e. within the absorption feature itself (e.g. 1050 nm; Haest et al., 2012). However, in data acquired by a VNIR imaging sensor, spectra at this wavelength are extremely noisy due to a fall-off in sensor sensitivity and atmospheric absorption by water. VNIR sensors are generally sensitive up to about 970 nm and therefore truncate the long-wave part of this feature. To measure longer wavelengths than 970 nm would require a separate shortwave infra-red (SWIR) sensor, thus increasing the equipment costs and the time required to acquire and process the data into a unified data-cube. We therefore test the hypothesis that the distribution of iron minerals can be quantified using only VNIR data (430–970 nm).

Quantifying the intensity of VNIR absorption features as an index of the abundance of iron minerals can be done using various methods which either use discrete bands or the whole spectral curve. The objective of all these methods is to quantify the strength of the iron absorption features whilst minimising the effects of variations in illumination, viewing geometry and changes in the background slope of spectra which are known to occur, for example, as a result of variations in grain size of the absorbing minerals (Cudahy and Ramanaidou, 1997) or organic content (Ben-Dor et al., 1999). In the simplest case, ratios of reflectance are used to quantify changes in the slope of spectra between the centre of an absorption feature and nearby wavelengths (bands) where iron or other minerals do not absorb (Crowley et al., 1989; Rowan and Mars, 2003; Rowan et al., 1974). Ratios, are very effective at removing variations in illumination across a scene. Appropriate bands, however, need to be selected a priori and because only a few bands (normally 2–3) are used, ratios do not exploit all of the available data. This can lead potentially to reduced accuracy and specificity compared with methods which use the whole spectral curve, as is the case for amounts of green vegetation (Elvidge and Chen, 1995). Furthermore, ratios are only effective where there is a constant spectral slope or gradient among samples. Variability in slope can be caused by variations in grain-size of minerals, organic material and water content (Clark, 1999). These factors are unlikely to be consistent among samples. Removal of the spectral continuum (Clark and Roush, 1984) has been used to normalise variations in the brightness of the rock. This is done by defining a spectral continuum for each spectrum and then dividing the spectrum by the continuum. Calculation of the depth (i.e. strength) of the mineral absorption features is then straightforward (e.g. Murphy, 1995). The defined continuum, however, would also contain information on, and therefore be impacted by, variations in background slope of the spectra, potentially causing changes in the apparent depth of absorption features.

Derivative analysis presents an elegant way of removing background slope effects of spectra and minimising effects of variations in illumination caused by sun-angle cloud cover or topography (Tsai and Philpot, 1998). Vertical mine faces are topographically complex and are subject to changes in incident illumination during the course of the day which can mask spectral variation due differences in the type or abundance of minerals. Derivative analysis of image spectra was used in this study to minimise these extraneous effects whilst enhancing changes in reflectance across wavelengths due to mineralogy.

Spectral derivatives have been used successfully to quantify iron oxides (Deaton and Balsam, 1991), but mainly in sediments or soils (e.g. Barranco et al., 1989; Kosmas et al., 1984; Scheinost et al., 1998). Several different measures of the amounts of iron minerals, calculated from the derivatives of spectra acquired in the laboratory are compared with total amounts of iron minerals determined from XRD analysis. These measures include the integration of the whole derivative curve (430–970 nm) and integration of specific peaks and troughs in the derivative spectrum associated with iron (Fe\(^{3+}\)) absorption. These same methods are then applied to hyperspectral imagery of a mine face. The secondary objective of this paper is to demonstrate a method to distinguish ore from shales, an important waste product in the mining process. Shales in the Pilbara are characterised by moderate to abundant amounts of the clay mineral kaolinite. Kaolinite can be mapped using its diagnostic SWIR absorption feature at 2196 nm (e.g. Kruse et al., 2003). We demonstrate an alternative approach which exploits residual slope differences between ore and shale that are present in the 1st derivative spectrum between 430 and 970 nm, thus removing the need for a SWIR sensor. Both VNIR and SWIR image data are collected, but the SWIR data are used only to validate the map of shale produced using only VNIR bands.

2. Materials and methods

2.1. Study area

The study area is located at the West Angelas mine in the Hamersley Province of the Pilbara Region of Western Australia. The province is characterised by extensive (80,000 km\(^2\)) Banded Iron Formation (BIF), composed of multiple layers of iron oxides including magnetite (Fe\(_{3}\)O\(_4\)) and hematite, alternating with bands of carbonate and chert (Thorne et al., 2008). At the West Angelas mine, operated by Rio Tinto, supergene enrichment has resulted in the formation of a hematite-goethite and goethite iron ore deposit. Extensive beds of West Angelas shale and manganiferous shale are also present and are considered to be waste (gangue). Mining is conducted via conventional drill and blast open pit operations. The principal minerals/rock types considered in this study are hematite, goethite, shale and manganiferous shale (Fig. 1).

2.2. Laboratory reflectance spectrometry and XRD analysis

Reflectance spectra (350–2500 nm) were acquired using a halogen light-source from the cleaned surface of cores of rock (10 cm wide), obtained from the mine site. Each spectrum was acquired from an area of 3.6 cm\(^2\), and was the average of 30 replicate
2.4.1. Derivative analysis of VNIR data

A common approach to calculate spectrum derivatives is to first apply a smoothing filter to reduce the noise and then perform the derivatives using a finite difference approximation (Orfanidis, 1996). The Savitzky-Golay filter, also known as least-squares smoothing filter, is a more elaborate approach that fits polynomials and differentiates them analytically (Savitzky and Golay, 1964).

This analytical approach is used in this study, using the corrected coefficients of Steiner (1972). Savitzky-Golay filters are attractive for use with hyperspectral data because they preserve the relative widths and heights of spectral signatures in noisy spectrometric data (Tsai and Philpot, 1998). In the Savitzky-Golay method, a frame of data points surrounding the current point (wavelength or spectral band) is fit to a polynomial using local least-squares regression. The function value of the current band is retained, while the function values of the others in the frame are discarded. For derivative analysis, an appropriate smoothing window must be set which resolves absorption features, yet is effective at suppressing spectral noise. The amplitude of derivatives changes with the width of the smoothing window (Fig. 2). To standardise the results across laboratory and imaging sensors, a smoothing filter of 60 bands was used. This number of bands was selected so that noise in image data was minimised, but the (Fe$^{3+}$) absorption features were preserved. Previous work by Monteiro et al. (2009) showed that by varying the width of the smoothing window of the derivative it is possible to control the degree of smoothness of spectra prior to classification. The authors found that with a relatively low (40 dB) signal-to-noise ratio it was possible to classify ore-bearing rocks of the same types as the ones described here.

Derivatives were calculated from field and image spectra between 430 and 970 nm. 1st- and 2nd-derivatives were calculated using a 2nd-order polynomial and 4th-derivatives were calculated using a 4th-order polynomial. To aid visualisation, 4th derivative spectra were inverted, so that peaks and troughs were consistent with the second-derivatives. The effect of increasing order of differentiation on estimates of absorption by iron minerals were tested using 2nd- and 4th-derivatives.

Fig. 2. Effect of increasing width of smoothing window on the amplitude of spectral derivatives. The first derivative of goethite is shown as an example: 20 bands (—); 40 bands (—-); 60 bands (—-). The original reflectance spectrum of goethite is shown for comparison (—). Note the amplitude of the derivative peaks and troughs decreases with increasing width of smoothing filter.
2.4.2. Measures of iron absorption from spectral derivatives

Different measures of derivative reflectance were tested for their ability to quantify the total abundance (% weight) of ferric iron minerals (hematite + goethite) as determined from quantitative XRD. The same measures were also derived from hyperspectral imagery of the mine face. Derivatives were integrated by summation over the range \( i \) to \( N \) bands:

\[
\sum_{i=1}^{N} |\delta(\lambda)|
\]

where \( \delta \) is the 1st-, 2nd- or 4th-derivative

First, the entire derivative spectrum between 430 and 970 nm was integrated either side of the zero-baseline (i.e. the area bounded by the peaks and troughs; \( \text{Int}_{430–970} \)). Due to the effects of spectral noise towards longer wavelengths, derivative spectra were also integrated over a restricted wavelength range of 430–757 nm (\( \text{Int}_{430–757} \)). Second, individual peaks and troughs in the spectrum were integrated (Fig. 3). This was not done for the first derivative because peaks were not fully differentiated about the zero-baseline. The relationship between total abundance of ferric iron minerals obtained from quantitative XRD analysis and these different derivative measures was explored using linear regression. The derivative measures calculated from image data were compared with predictions of the distribution of iron on the mine face based on the field mapping of the ore zones.

2.4.3. Separating ore from shale

Variations in 1st-derivative reflectance between rock types enabled wavelengths to be selected to separate ore-bearing rocks from shale, which is considered to be waste. Differences in background slope among spectra of different rock types were still evident in 1st-derivative spectra (spectra were not evenly distributed about the zero base-line). Further differentiation (in 2nd- or 4th-derivatives) removed differences in background slope, leaving only information related to specific absorption features. 1st-derivative spectra were therefore most useful in distinguishing rock types based upon differences in the background slopes of spectra. Preliminary analysis of laboratory and image spectra identified three wavelengths (702, 765 and 809 nm) in the 1st-derivative spectrum which could potentially be of use in separating ore from waste (Fig. 3).

2.5. Validation

2.5.1. Distribution of iron

Numerical validation of the relationship between total amounts of iron (hematite + goethite) done using field spectra of rock

---

**Fig. 3.** Laboratory spectra of dominant rock types at the West Angelas: goethite (- -), hematite (- -); West Angelas shale (- -); manganiferous shale (- -). (a) Reflectance. Centres of absorption by Ferric iron (Fe\(^{3+}\)) are indicated, (b) 1st-derivative reflectance, (c) 2nd-derivative reflectance, (d) 4th-derivative reflectance peaks (P) and troughs (T) used to quantify amounts of ferric iron are numbered consecutively in (c–d). Location of the bands used to separate ore from shale are indicated by vertical lines in (b).
samples (Section 3.2). Ideally, estimates of the abundance of these minerals derived from the imagery should be compared with amounts measured (using XRD) from rocks sampled directly from the mine face. Because of the relatively coarse spatial resolution of the sensor, it would be necessary to acquire numerous rock samples from within several quadrats on the mine face. Average amounts of iron from all rock samples within each quadrat could then be compared with the estimates derived from the imagery, averaged over each quadrat. Mine safety regulations severely restrict direct interaction with the mine face so it was not possible to sample in this way. An alternative approach to validation of the imagery was therefore designed. A large body of knowledge about the geology of the local area and the mine pit has been accumulated by Rio Tinto, including an extensive database of the typical mineral abundances for each of the rock types that were mapped in each of the geological zones identified in the field (see Fig. 1). From this database, hypotheses could be formulated about how iron should be distributed among the different zones on the mine face. These distributions can then be compared with those measured from the imagery. This approach to validation of the image distribution of iron depends upon an accurate delineation and mapping of the geological zones on the mine face. To achieve this, an experienced mine geologist with a working knowledge of the particular mine pit in question, was used to verify the geological mapping (Fig. 1). Based on our a priori knowledge of the abundance of hematite and goethite in the rocks present in these zones the following hypotheses can be made:

**Hypothesis 1.** Zone 2 (manganiferous shale) will have the smallest amount of iron of all zones.

**Hypothesis 2.** Zones 1 and 2 will each have less iron than Zones 3–5.

**Hypothesis 3.** Zone 5 (high grade ore) will have the greatest amount of iron of all zones, including the other ore zones (Zones 3, 4 and 6).

If Hypotheses 1–3 are supported then the image-derived distribution of iron will be consistent with the geological mapping in the field. To test these hypotheses, statistics were generated from two Regions of Interest (ROI; 60 × 60 pixels) defined for each Zone.

### 2.5.2. Separation of ore from shales

Maps of the distribution of shale were validated by: (i) comparing them to the distribution of kaolinite as mapped by the SWIR sensor and (ii) comparing them to the distribution of shale from the geological mapping made in the field.

Images of 1st-derivative reflectance reflectance at 702 nm, 765 nm and 805 nm were compared with maps of kaolinite distribution as mapped by the depth of the kaolinite absorption feature at about 2196 nm. The depth of this feature is proportional to the amount of kaolinite and is therefore an independent and reliable indicator of manganiferous shale (with abundant kaolinite) and West Angelas shale (with moderate kaolinite). If the 1st-derivative reflectance was effective at separating shales, then the distribution of pixel values would show similar patterns to those mapped by the depth of the kaolinite feature. The kaolinite feature at about 2196 nm (ABD$_{2196}$) was isolated using automated feature extraction (Kruse, 1988), and its depth determined in accordance with Clark and Roush (1984). Mean values of the 1st-derivative reflectance at 702, 765 and 809 nm and ABD$_{2196}$ were extracted from common (spatially collocated) ROI, randomly located within the images (24 ROI from the ore zones and 24 from the shale zones). Plots of ABD$_{2196}$ against derivative reflectance were made from the mean ROI values; these were used to define appropriate thresholds for classification.

To determine how well the 1st-derivative reflectance at 702, 765 and 809 nm corresponded to the distribution of shales, they were each compared with the geological maps made in the field. Standard statistics, describing the performance of classification were generated from 1st-derivatives at 702, 765 and 809 nm for areas of ore and shale identified in the field mapping. Using the plots of ABD$_{2196}$ against 1st-derivative reflectance, appropriate thresholds for classification were determined as the mid-point between the median ROI values separating ore from shale classes (1.382 for 702 nm, 0.549 for 765 nm and –0.207 for 809 nm). Pixel values below and above their respective thresholds were classified, respectively, as ore and shale. The classified pixels were extracted from Zones 1–6 in the image and compared with the distribution of ore and shale mapped in the field using a one-against-all approach. Zone 3 was not included in this analysis because it contained a highly heterogeneous mix of ore and shale which varied at small spatial scales, making it impossible to assign class labels. Definitive class labels against which the classified maps were to be compared were assigned according to the geological mapping made in the field (i.e. ore vs. shale). Statistics, including accuracy, precision, recall and the Kappa coefficient of agreement (Kappa; Congalton et al., 1983; Hudson and Ramm, 1987) were generated. Precision is a measure of the quality of the results predicted and is the number of positive results predicted divided by the total number of results returned (Olson and Delen, 2008). Recall measures the quantity of positive results predicted by the classifier. It is the number of positive results predicted divided by the total number of results that should have been returned (van Rijbergen, 1979). For Kappa, a value of zero means no agreement and a value of one perfect agreement. A value of 0.75 is considered to be a very good to excellent classification (Monserud and Leemans, 1992).

### 3. Results

#### 3.1. Laboratory and image reflectance spectra and their derivatives

The four minerals/rock types examined as part of this study (hematite, goethite, West Angelas shale and manganiferous shale) had distinctive spectral curve shapes (Fig. 3a). Absorption features associated with Fe$^{3+}$ are evident in all spectra. The spectrum of manganiferous shale, however, has only a weak shoulder due to Fe$^{2+}$ at wavelengths greater than 820 nm. Peaks and troughs in the 1st-derivative spectrum represent, respectively, the inflection points of the upward and downward slopes of absorption features (Fig. 3b). Because the reflectance spectrum of manganiferous shale monotonically increases with wavelength, the 1st-derivative of this spectrum never dips below the zero baseline. With increasing order of differentiation, background slopes of spectra are progressively removed and the peaks and troughs become more symmetrically centred around the zero-baseline (e.g. the 2nd- and 4th-derivatives; Fig. 3c and d).

Image spectra from areas of the mine face of known composition closely resemble laboratory spectra from cores of rock (cf. Figs. 3 and 4). Although there are brightness differences between laboratory and image spectra (e.g. West Angelas shale), their curve shapes and the wavelength positions of the peaks and troughs in derivative spectra are consistent.

#### 3.2. Relationship between XRD and derivative estimates of iron

Regression analyses of XRD estimates of total abundance of iron minerals on the different spectral derivative measures of iron absorption show moderate to strong linear relationships (Fig. 5;...
Table 1). There was no significant difference in the strength of the relationship between XRD estimates and the Int430–970 and Int430–757 measures, respectively (Table 1). The weakest relationship was for the 1st-derivative measures Int430–970 and Int430–757, which explained only 56% and 54% of the variance in the data, respectively. Int430–970 and Int430–757 generated from 2nd- and 4th-derivatives showed strong, linear and highly significant relationships with total amounts of iron minerals. Integration of individual peaks and troughs in the derivative spectra, in some cases, marginally improved the strength of the relationship over integration over broader spectral regions (e.g. for the 2nd-derivative, Trough 1 gave a marginally stronger relationship than did the Int 430–970 and Int430–757 measures). The weakest relationships were found in the peaks and troughs towards longer wavelengths (e.g. Trough 2 in the 2nd-derivative and Peak 3 and Tough 3 in the 4th-derivative).

3.3. Distribution of iron from hyperspectral imagery

Images of the distribution of iron minerals are shown in Fig. 6. Box plots of values extracted from within ROI for each geological zone are shown in Fig. 7. Integrated derivative values across the whole spectral curve (Int430–970) gave almost identical results to values integrated up to 757 nm (Int430–757). This was true for all orders of derivative generated from the image data. Henceforth, therefore, only Int430–970 is considered. Visual examination of the images showed that Int430–970 constructed from the 1st-derivative, had different patterns of distribution of iron than did the same measure constructed from 2nd- or 4th-order derivatives (cf. Fig. 6a with b and c). Fig. 6a shows that the 1st-derivative does not show the large differences in the amounts of iron between the zones of manganiferous shale (Zone 2) and ore (Zones 3–6). The zone with the smallest median amount of iron was Zone 3 (Fig. 7a), therefore Hypothesis 1 was rejected. The median amounts of iron in the shale zones (Zones 1 and 2) was greater than in the mixed shale and ore zone (Zone 3; Fig. 7a), leading to the rejection of Hypothesis 2. The amount of iron in Zone 5 was less than Zone 4, leading to the rejection of Hypothesis 3. Thus, the patterns of distribution of iron mapped from the 1st-derivative were not consistent with patterns expected from the geological mapping made in the field. Results from the 2nd- and 4th-derivatives (Figs. 6b and c and 7b and c) were similar to each other, but different to results from the 1st-derivative. The smallest amount of iron was found in Zone 2 (Hypothesis 1 accepted); the amounts of iron in Zones 1 and 2 were less than in Zones 3–6 (Hypothesis 2 accepted) and the greatest amount of iron was found in Zone 5 (Hypothesis 3 accepted). Acceptance of all hypotheses indicates that the pattern of distribution among zones as mapped by the 2nd- and 4th-derivative was entirely consistent with patterns expected from the geological mapping made in the field.
Symbols represent different types of rock: goethite ($\alpha$); hematite ($\beta$); West Angelas shale (○); manganiferous shale (●). (a) 1st-derivative reflectance, (b) 2nd-derivative reflectance, (c) 4th-derivative reflectance.

4. Discussion

Increasingly, hyperspectral data acquired from field based platforms are being exploited in dangerous environments to acquire information about nearby – but inaccessible – objects. This relatively new modality of use introduces new opportunities and challenges. Open pit mines are extremely challenging places to acquire information about nearby – but inaccessible – objects. This relatively new modality of use introduces new opportunities and challenges. Open pit mines are extremely challenging places to acquire hyperspectral data from field-based platforms. Dust, high-temperatures and safety considerations necessitate that data used for operational mining should be acquired in the shortest possible interval of time. To this end we have shown that the distribution of iron can be quantitatively estimated using only VNIR data.

Analysis of individual peaks and troughs in 2nd- and 4th-derivative spectra gave mixed results (not shown). Images from individual peaks and troughs were all noisier than Int$_{430–970}$ calculated from the same derivative. Integration of P1, P2, T1 and T2 from 2nd-derivatives showed a similar distribution to Int$_{430–970}$, however, P3 showed only gross differences in distribution of iron among geological zones. For the 4th-derivative, P1 gave almost identical results to Int$_{430–970}$. P2 and T1 were very noisy and showed only gross differences in distribution of iron among geological zones. P3 and T2 contained only noise.

3.4. Separation of ore from shale

To provide an independent validation of the distribution of shale as mapped by the first derivative, values were compared with amount of kaolinite on the mine face as mapped by the depth of the diagnostic kaolinite Al–OH feature at 2196 nm (ABD$_{2196}$). Values of the 1st-derivative at 702, 765 and 809 nm were separately plotted against ABD$_{2196}$ (i.e. average values within each of 24 ROI from the ore zones and 24 ROI from the shale Zones; Fig. 8). 1st-derivative reflectance at 702 nm did not show any relationship with ABD$_{2196}$ (Fig. 8a). Good separation of shales from ore was provided by 1st-derivative reflectance at 765 nm (Fig. 8b), however, the best separation was provided by 1st-derivative reflectance at 809 nm (Fig. 8c).

Visual comparison of 1st-derivative reflectance at 702 nm with the geology maps made in the field showed that it did not separate ore from shale (Fig. 9a). At 765 nm, the 1st-derivative reflectance of the West Angelas (Zone 1) and manganiferous shale (Zone 2) had a consistently larger 1st-derivative reflectance than did the ore zones (Zones 3–6), providing good separation between the ore and shale (Fig. 9b). Reflectance at 809 nm showed a similar pattern to that at 765 nm (Fig. 9c). Statistical analyses of the classified maps of ore and shale showed that there was excellent agreement between the classification made from a simple level-slice of the derivative reflectance at 765 and 809 nm and the geological mapping made in the field (Table 2). Both of these maps had an accuracy and precision greater than 98%. Recall, a measure of the sensitivity, was greater than 97%. Kappa was greater than 95% indicating a near-perfect classification performance. 1st-derivative reflectance at 702 nm provided a poor classification (Table 2).

Table 1

<table>
<thead>
<tr>
<th>Index</th>
<th>1st-derivative</th>
<th>2nd-derivative</th>
<th>4th-derivative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Int$_{430–970}$</td>
<td>0.56 $&lt;0.01$</td>
<td>0.74 $&lt;0.0001$</td>
<td>0.77 $&lt;0.0001$</td>
</tr>
<tr>
<td>Int$_{430–757}$</td>
<td>0.54 $&lt;0.01$</td>
<td>0.76 $&lt;0.0001$</td>
<td>0.76 $&lt;0.0001$</td>
</tr>
<tr>
<td>Peak 1</td>
<td>–</td>
<td>–</td>
<td>0.73 $&lt;0.0001$</td>
</tr>
<tr>
<td>Peak 2</td>
<td>–</td>
<td>–</td>
<td>0.70 $&lt;0.0001$</td>
</tr>
<tr>
<td>Peak 3</td>
<td>–</td>
<td>–</td>
<td>0.74 $&lt;0.0001$</td>
</tr>
<tr>
<td>Trough 1</td>
<td>–</td>
<td>–</td>
<td>0.78 $&lt;0.0001$</td>
</tr>
<tr>
<td>Trough 2</td>
<td>–</td>
<td>–</td>
<td>0.68 $&lt;0.001$</td>
</tr>
<tr>
<td>Trough 3</td>
<td>–</td>
<td>–</td>
<td>0.49 $&lt;0.01$</td>
</tr>
</tbody>
</table>

Fig. 5. Regression analyses of amount (% weight) of Ferric iron ($\text{Fe}^{3+}$) on the absolute derivative reflectance integrated over the wavelength range 430–970 nm. Symbols represent different types of rock: goethite (△); hematite (●); West Angelas shale (○); manganiferous shale (●). (a) 1st-derivative reflectance, (b) 2nd-derivative reflectance, (c) 4th-derivative reflectance.
Derivative in quantifying distribution of ferric iron minerals. This is likely due to background slope effects being present in the 1st-derivative spectrum but which had been removed by further differentiation in the higher-order derivatives (Tsai and Philpot, 1998). A positive slope means that the 1st-derivative spectrum has a positive bias, even though no appreciable absorptions were present in the spectrum. Integration of the spectral curve mean that such biases would be translated into large abundances of iron even though only small amounts of iron, were present. In 2nd- and 4th-order derivatives, the deviation around the zero base-line was related only to absorptions present in the spectrum, with low-frequency spectral slope effects being largely eliminated. The distribution of amounts of iron on the mine face calculated from the 1st-derivative was therefore not consistent with field observations. This draws into question the use of 1st-derivatives to quantitatively estimate abundances of minerals in any rocks which may have variable spectral slopes. We had anticipated that the 1st-derivative would yield less accurate results than higher-order derivatives (as predicted by the smaller coefficient of determination; Table 1) but the very large differences in distribution of iron on the mine face as mapped by former compared with the latter was surprising. In this study, the abundance of total iron (i.e. hematite + goethite) was estimated. Work is underway to derive independent estimates for goethite and hematite which take into account the inherent variations in brightness between these minerals.

Reflectance spectra have absorption features which contain information on the identity of an absorbing material as well as its concentration. A major advantage of derivative analysis is that background (i.e. low frequency) slope effects are removed with increasing order of differentiation, leaving only information related to absorption by specific materials. Information on background slope can still remain in 1st derivative spectra of many materials (e.g. sediment: Murphy et al., 2005). We exploit differences in the slope of 1st-derivative spectra of ore and shale to separate these classes of rock. It was initially expected that quantifying amounts of iron absorption on the mine face by integrating the spectral curve would, by default, delineated ore from shale (ore should have been separated from shale because it has more iron).

However, West Angelas shale contains (or is coated with) some iron (see Fig. 6b and c) it was not delineated from the ore using this method. Ore zones were instead separated from shale using information on spectral slope retained within the 1st-derivative spectrum. The 1st-derivative spectra at 765 nm and 809 nm are brighter in the shales because, their zero-order image spectra are still, overall, increasing in reflectance, unlike goethite or hematite. The distribution of shales on the mine face as mapped by the first-derivative reflectance at 765 nm and 809 nm strongly resembles that mapped by the depth of the kaolinite feature at 2196 nm. A classification based on the 1st-derivative was extremely effective at mapping distribution of shale and ore, indicating the potential of the 1st-derivative for separating rock types on the basis of differences in spectral curve shape in the VNIR.

Derivatives have the disadvantage of increasing spectral noise with increasing order of derivative. We attempted to minimise the effects of noise by generating the derivatives using a combined differentiation and smoothing approach rather than simple change in slope with respect to wavelength. A second limitation with derivatives is that an appropriate wavelength interval needs to be specified. For broad spectral features like crystal field absorptions a relatively wide wavelength interval needs to be set to detect the change in slope. The optimal wavelength interval would change depending upon the spectral resolution of the sensor and the absorption characteristics of the minerals being mapped.

The purpose of this study was to demonstrate the potential of hyperspectral VNIR imagery to yield quantitative information on the distribution and abundances of iron minerals and to separate major components (ore and shale) in the mining process. Other studies have used derivative analysis to obtain information about iron minerals, particularly in soils (e.g. Kosmas et al., 1984), however these studies have been limited to using only field spectra. In this study we applied derivative analysis to hyperspectral imagery of a mine face. To the best of our knowledge, this is the first study to apply derivative analysis to hyperspectral imagery obtained from a field-based platform for geological applications. Hyperspectral imagery is a more difficult test of this approach than laboratory spectra because of the increased amounts of noise in the data and the variable illumination and viewing geometries.
involved. Although integration over some individual Peaks or Troughs (e.g. T1 in 2nd-derivative and P2 in the 4th-derivative) showed the strongest relationship with total amounts of iron minerals obtained from XRD analysis, these measures, when calculated from hyperspectral imagery, produced much noisier images than did integration over broader spectral regions (e.g. Int430–970).

Methods used in this draw upon knowledge about the physical processes which governed the absorption of light by iron minerals. Spectral features (in this case integrated values of 1st-, 2nd- and 4th-derivatives) were selected to measure of the intensity of absorption. Such manual selection of features may not, however, be the most efficient, effective or parsimonious way to select the optimal constellation of features required to predict amounts of iron or to separate different types of rock using hyperspectral data. Methods designed to detect spectral anomalies in hyperspectral data could also be used to identify iron minerals from hyperspectral data (Du and Zhang, 2011; Matteoli et al., 2010). This paper has underscored the fact that different orders of derivative (including the original reflectance) bring out different aspects of information contained within spectra. A more general approach to feature selection would be to automate the process within a machine-learning framework so that the most optimal features are ‘learned’ from the data. Some work has already been done in this regard (e.g. Monteiro et al., 2009; Murphy et al., 2012) and further work is underway.

Fig. 7. Box plots of amounts of hematite and goethite (iron) estimated by Int430–970. Data are from ROI randomly located within each geological Zones (1–6) mapped in the field. Image data have been converted to amounts of hematite + goethite using the regression coefficients derived from the data presented in Fig. 4. The line bisecting each box is the median. The lower and upper boundaries of the box represent the 25th and 75th percentile respectively. Whiskers indicate the 10th and 90th percentiles. The black circle represents the 5th and 95th percentile of the outliers. (a) 1st-derivative, (b) 2nd-derivative, (c) 4th-derivative.

Fig. 8. Plot of depth of the absorption feature at 2196 nm (ABD2196) against derivative reflectance. Values are averages from regions of interest, randomly-positioned within the ore zones (Zones 3–6; n = 24), West Angelas shale (Zone 1; n = 12) and manganiferous shale (Zone 2; n = 12). Vertical lines show optimal derivative values for separating shale from ore calculated as the difference between median values for each class. (a) 702 nm. (b) 765 nm. (c) 809 nm.
5. Conclusions

(i) Total abundance (% weight) of iron minerals from quantitative XRD analysis was tightly correlated with measures from 2nd- and 4th-derivatives of laboratory spectra. These measures were obtained by integrating the derivative curve over the entire VNIR wavelength range and over specific parts of the spectrum corresponding to specific Fe$^{3+}$ absorption features.

(ii) Ore and shale (waste) could be separated using the 1st-derivative reflectance at 765 and 809 nm. The distribution of shale, thus mapped, was entirely consistent with a map of the distribution of kaolinite (a major component and specific indicator of shales in the study area), derived from SWIR data. Standard statistics of validation (accuracy, precision, recall and Kappa) showed excellent agreement with the field mapping.

(iii) Total amounts of iron can be quantified and that ore can be separated from shale using only the VNIR sensor module of a hyperspectral imaging system. The significance of this is that data can be collected using only a single imaging sensor, thus reducing costs associated with equipment, data collection and analysis.

(iv) Hyperspectral imagery provides information on the identity, abundance and distribution of minerals on vertical mine faces. This information is invaluable in informing, guiding and automating the mining process.

(v) Derivative analysis was found to be an effective tool for suppressing topographically-induced variations in illumination across the mine face, enabling the distribution of iron to be mapped.

(vi) 2nd- and 4th-derivatives were most effective in quantifying amounts of iron minerals in accordance with predictions based on a field mapping and a priori knowledge. The 1st-derivative was a poor estimator of amount of iron minerals.

Acknowledgements

This work has been supported by the Rio Tinto Centre for Mine Automation and the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government. We thank the Reviewers which have improved the quality of this paper. The help of Sven Schneider and Katie Silversides in collecting spectra for XRD analysis is gratefully acknowledged. Particular thanks are given to James Batchelor for his help in the field.

References
