Fuzzy-based Integration of Sentinel-2 MSI Data for Locating Potential Hydrocarbon Microseepage

Ziyong Zhou
College of geosciences, China University of Petroleum Beijing, 102200, China
zzy@cup.edu.cn

Abstract. The vertical migration of hydrocarbons from oil and gas reservoirs to the surface and/or subsurface (i.e., microseepage) can induce mineral alteration that results in concentration of ferrous iron, clay, and carbonate minerals at the Earth’s surface. These altered zones may be detected in remote sensing imagery based on their specific reflectance spectral characteristics; zones where three types of mineral alteration occur simultaneously reflect possible hydrocarbon microseepage locations with the potential for underlying oil and gas reservoirs. In this study, a fuzzy set based approach was used to integrate the results of principal component analysis (PCA) and band ratios (BRs). First, each selected indicative principal component (PC) and BR was treated as a fuzzy set, and a corresponding fuzzy membership function defined. The membership degree of each pixel, which indicates the possibility of the presence of specific altered minerals, was then calculated. Subsequently, the Gamma operator was used to fuse all fuzzy sets to create a new fuzzy set, which was then defuzzied and used to indicate the possible locations of hydrocarbon microseepages. The Sentinel-2 MSI data were used to illustrate the approach and the zones of hydrocarbon microseepage mapped.

1. Introduction
The microseepage of hydrocarbons from oil and/or gas reservoirs can induce mineral alterations at the surface and/or subsurface[1-2]. The dominant minerals induced by hydrocarbon microseepage include ferrous iron, clay minerals, and carbonate minerals [3]. The concentration of these minerals at the surface can be highlighted in remote sensing imagery owing to their specific reflectance spectra in visible and infrared (VIR) and short-wavelength infrared (SWIR) [2,4]. The most used multispectral images include TM/ETM/OLI images of the Landsat series[5] and ASTER images[6]. Sentinel-2 MSI (hereinafter MSI for short) images provide similar bands as to these of TM/ETM/OLI and are available in free. However, the Sentinel-2 platform was launched in June 2015; therefore, few studies to date have used it for hydrocarbon microseepage detection.

Principal component analysis (PCA) and band ratios (BRs) are commonly used methods for the extraction of hydrocarbon microseepage information from multispectral images. The PCA-based method, which was first proposed by Crosta & Moore [7], implements a PC transform of selected original bands and selects the specific principal component (PC), or indicative PC, according to the correlation between reflection spectral characteristics of iron-bearing minerals and hydroxyl minerals and the load matrix (i.e., eigenvector). The selected indicative PCs may be used to highlight pixels that may be related to iron-bearing and hydroxyl-bearing minerals [8]. BRs can also highlight the spatial distribution of a specific mineral by selecting appropriate reflection and absorption bands corresponding to the mineral in question. These ideas can also be adopted to detect altered minerals induced by hydrocarbon microseepages (i.e., ferrous iron, clay, and carbonate minerals). False color compositions
(FCC) of selected PCs and/or BRs can then be used to highlight the occurrence of detected minerals, providing spatial distributions of potential hydrocarbon microseepages. The FCC based approach is a simple and easy way to integrate different images and highlight specific pixels. However, the exact interpretation and determination of subtle colours in FCC images remains an issue. To address this issue, a novel approach was developed in this study to integrate PCs and/or BRs using fuzzy set theory. Here, Sentinel-2 MSI data from the Chu-Sarysu basin, Kazakhstan were used to illustrate the approach.

2. Dataset
Sentinel-2 is a European wide-swath, high-resolution, multi-spectral imaging mission, which carries a Multi-Spectral Instrument (MSI) and provides continuity of SPOT and LANDSAT-type image data. Sentinel-2 MSI data are acquired on 13 spectral bands in the VNIR and SWIR with different spatial resolution. The study region is located in the north of the Chu-Sarysu Basin, Kazakhstan. Figure 1(a) shows a natural colour MSI image.

3. Methodology

3.1 Selection of Principal Components and Band Ratios
According to the reflectance spectra of altered minerals induced by leaking hydrocarbon at the Earth’s surface (i.e., ferrous iron bearing minerals, clay, and carbonate minerals), specific band groups for PCA and BRs (table 1) were selected. As we know, the PCA eigenvectors loadings imply the contribution of the original variables to the principal components. Therefore, examination of their sign (i.e., positive or negative) and magnitude gives an indication of which spectral properties of altered minerals are responsible for the PCs, i.e. indicative PCs for specific minerals. For example, Eigenvector loadings of PCA of input Bands 2, 4, 11 and 12 (group 1) for PC3 of table 2 indicate that it describes the difference between the band 11 and 12. Clay which has the strongest reflectance in band 11 and strongest absorption in band 12 will appear on PC3 as the darkest pixels (negative eigenvector loadings) in this case. It follows that a simple negation of PC3 will indicate the potential altered region of clay by bright pixels. The other potential altered region of ferrous iron bearing minerals and carbonate minerals can also be highlighted by the same way with different band group input for PCA as described in table 1.

| Minerals                      | Band groups for PCA | Band ratios | Highlighting pixel in BR |
|-------------------------------|---------------------|-------------|--------------------------|
| Bleaching of red beds         | \                  | 2/4         | bright                   |
| Ferrous iron                  | 2+,4+,8−,11+        | 4/8         | bright                   |
| Clay and carbonate            | 2,4,11+,12−         | 11/12       | bright                   |
| Clay and carbonate            | 3,4,11+,12−         | \           | \                       |
| Ferrous iron, clay and carbonate | 4,8−, 11+,12−     | \           | \                       |

Note: ‘+’ indicates the reflection bands, and ‘−’ indicates absorption bands.

| Minerals                      | Band groups for PCA | Band ratios | Highlighting pixel in BR |
|-------------------------------|---------------------|-------------|--------------------------|
| Bleaching of red beds         | \                  | 2/4         | bright                   |
| Ferrous iron                  | 2+,4+,8−,11+        | 4/8         | bright                   |
| Clay and carbonate            | 2,4,11+,12−         | 11/12       | bright                   |
| Clay and carbonate            | 3,4,11+,12−         | \           | \                       |
| Ferrous iron, clay and carbonate | 4,8−, 11+,12−     | \           | \                       |

Note: ‘+’ indicates the reflection bands, and ‘−’ indicates absorption bands.
3.2 Membership function of PCs and BRs

Fuzzy set theory has been rapidly developed since it was first proposed and has found wide application for image processing and analysis [9]. Fuzzy set based integration of remote sensing images need solve two problems: (1) definition of the membership function and calculation of the membership degree; that is, the degree to which each pixel belongs to a certain fuzzy set; and (2) the fuzzy fusion algorithm.

According to shape, the membership function can be roughly classified into one of three categories: S-type, Z-type, or bell-type; and each type has various varieties in actual application.

If the bright pixels of indicative PCs or BRs highlights a specific mineral (i.e., the greater the value of the pixel, the greater is the possibility that this pixel relates to the corresponding mineral), then a non-linear increasing (S-type) membership function can be defined as follows [10]:

\[
\mu(x) = \begin{cases} 
0 & \text{if } x \leq 0 \\
\frac{1}{1 + \left(\frac{x - f_2}{f_1}\right)^{f_1}} & \text{other}
\end{cases}
\]

if a dark pixel of indicative PCs or BRs highlights a specific mineral (i.e., the smaller the value of that pixel, the greater is the possibility of that pixel relating to the corresponding mineral, then a simple negation is implemented and the membership function can also be defined as equation (1).

\[ x \text{ in equation (1) is the value of a pixel, } \mu(x) \text{ is the corresponding membership, } f_1 \text{ is the spread which determines the curve gradient, and } f_2 \text{ is the midpoint which corresponds to the value for which the membership is 0.5. In our case, } f_1 = 3, f_2 = \text{mean} + 1.5\sigma \text{ (mean – average of an image, and } \sigma-\text{standard deviation), implies that if the value of a pixel is greater than mean+1.5} \sigma, \text{ then this pixel more likely highlights the potential location of altered mineral. We can thus calculate the membership degrees of each pixels of indicative PCs “image” and BRs “image” using corresponding membership function. Figure 1 (b)-(h) shows the “image” of membership degree of different fuzzy sets (i.e. indicative PCs and BRs). The bright pixels in figure 1(b)-(h) highlight the possible locations of corresponding minerals.} \]
3.3 Fuzzy integration of PCs and BRs

For the fuzzy fusion algorithm for two or more fuzzy sets, the AND, OR, fuzzy algebraic product, fuzzy algebraic sum, and fuzzy Gamma operator can be used. Zhou et al. described each operator in detail [11]; here, only the Gamma operator was used. Given $n$ fuzzy sets, namely, $A_1, A_2, \ldots, A_n$, and their corresponding memberships, $\mu_{A_1}, \mu_{A_2}, \ldots, \mu_{A_n}$, the Gamma operator of fuzzy sets $A_1, A_2, \ldots, A_n$ is defined as:

$$
\mu_{A_1 \odot A_2 \odot \ldots \odot A_n} = \left(1 - (1 - \mu_{A_1}) \cdot (1 - \mu_{A_2}) \cdot \ldots \cdot (1 - \mu_{A_n})\right)^\gamma \cdot \left(\mu_{A_1} \cdot \mu_{A_2} \cdot \ldots \cdot \mu_{A_n}\right)^{1-\gamma}
$$

(2)

where the exponent $\gamma$ is a value between [0, 1]. Given the two special cases here:

- If $\gamma = 1$, then $\mu_{A_1 \odot A_2 \odot \ldots \odot A_n} = 1 - (1 - \mu_{A_1}) \cdot (1 - \mu_{A_2}) \cdot \ldots \cdot (1 - \mu_{A_n})$, and $\mu_{A_1 \odot A_2 \odot \ldots \odot A_n}$ is greater than $\max(\mu_{A_1}, \mu_{A_2}, \ldots, \mu_{A_n})$, implying the maximum distribution of possible altered minerals.

- If $\gamma = 0$, then $\mu_{A_1 \odot A_2 \odot \ldots \odot A_n} = \mu_{A_1} \cdot \mu_{A_2} \cdot \ldots \cdot \mu_{A_n}$, and $\mu_{A_1 \odot A_2 \odot \ldots \odot A_n}$ is smaller than $\min(\mu_{A_1}, \mu_{A_2}, \ldots, \mu_{A_n})$, implying the minimum distribution of possible altered minerals.

4. RESULTS AND DISCUSSION

As described above, the bright pixels in each of images in Figure 1 highlight the possible locations of clay, carbonate or ferrous-iron-bearing minerals that relate to hydrocarbon microseepage. Therefore, the bright pixels in the fused images of seven fuzzy sets in figure 1 highlight zones of simultaneous occurrence of ferrous-iron-bearing minerals, clay minerals, and carbonate minerals, which gives a more reliable suggestion of location of hydrocarbon microseepage.
As described in (2), the fused results of the Gamma operator depend on the exponent $\gamma$; and figure 2 shows fused results of four indicative PCs and three BRs given $\gamma = 0.7, 0.8,$ and $0.9.$ The greater the value of $\gamma$, the wider the distribution of the potential zone of hydrocarbon microseepage, while the lower the reliability. Due to the trade-off between the distribution and reliability, we finally used $\gamma = 0.7$ to identify the potential microseepage sites. Finally, the fused fuzzy set was defuzzified and reclassified into two levels according to membership (i.e., $\leq 0.5, > 0.5$), allowing for the assessment of potential hydrocarbon microseepage, as shown in figure 3.

Figure 2. Gamma fused result of 4 PCs and 3 BRs given (a) $\gamma = 0.7$, (b) $\gamma = 0.8$, and (c) $\gamma = 0.9$.

Figure 3. Potential location of hydrocarbon microseepage (The blue pixels).

Figure 3 shows an overlay of the mapped hydrocarbon microseepages and natural color image, where the blue pixels indicate the potential location of hydrocarbon microseepage.

5. CONCLUSIONS
This study attempted to integrate PCA and BR results using fuzzy set theory in order to detect altered minerals, and through them hydrocarbon microseepages, from MSI data. The proposed approach was able to quantitatively synthesize the PCs and/or BRs and to visualize the distribution of possible altered zones. Nevertheless, some issues require further study; for example, during the construction of
membership functions for fuzzy fusion, there was no explicit criterion for the selection of $f_1$ and $f_2$ for the non-linear increasing (S-type) membership function. Based on the results of this study, values of $f_1 = 3$ and $f_2 = \text{mean} \pm 1.5 \sigma$ appear to be reasonable. The other main issue involves the selection of a $\gamma$ value for the Gamma operator; generally speaking, $\gamma = 0.7$–0.9 appears to be applicable for most instances.

References
[1] Van der Meer, F., Van Dijk, P., Van der Werff, H., Yang, H. (2002) Remote sensing and petroleum seepage: a review and case study, Terra Nova, 14:1–17.
[2] Schumacher, D., Abrams, M. A. (Eds.). (1996) Hydrocarbon migration and its nearsurface expression, AAPG Mem., no. 66.
[3] Donald, F., Saunders, K., Ray, B., Thompson, C. K. (1999) Model for hydrocarbon microseepage and related near-surface alterations, AAPG Bull., 83:170–185.
[4] Ben, L. (2017) Remote sensing techniques for onshore oil and gas exploration, Leading Edge, 2017:24-32.
[5] Zhang, G. F., Zheng, Z., Shen, X. H., Zou, L. J., Huang, K. Y. (2011) Remote sensing interpretation of areas with hydrocarbon microseepage in northeast china using landsat-7/ETM+ data processing techniques, Int. J. Remote Sens., 32: 6695-6711.
[6] Lammoglia, T., De Souza Filho, C. R. (2013) Unraveling hydrocarbon microseepages in onshore basins using spectral–spatial processing of advanced spaceborne thermal emission and reflection radiometer (ASTER) data, Surv. Geoph, 2013:349–373.
[7] Crosta, A. P., Moore, J. McM. (1989) Enhancement of Landsat thematic mapper imagery for residual soil mapping in SW Minas Gerais State, Brazil: A prospecting case history in greenstone belt terrain, In: Proceedings of the 7th Thematic Conference on Remote Sensing for Exploration Geology. Calgary, pp. 1173-1187.
[8] Loughlin, W. P. (1991) Principal component analysis for alteration mapping, Photogramm Eng. Rem. S., 57(9), 1163-1169.
[9] Isabelle, B. (2015) Fuzzy sets for image processing and understanding, Fuzzy Set Syst., 281: 280–291.
[10] Tsoukalas, L.H., Uhrig, R.E. (1997) Fuzzy and neural approaches in engineering, John Wiley and Sons, New York.
[11] Zhou, Z.Y., Yu, H. Y., Gu, X. D. (2018) Fuzzy fusion of geological and geophysical data for mapping hydrocarbon potential based on GIS, Petrol. Geosci., 24:131–141.