

# 鉱物組成特定を目的とした可視・短波長赤外域の分光反射特性に基づく 最適マルチスペクトルバンドの設計

岸本将英\*・久保大樹\*・小池克明\*

## Optimal multispectral band estimation for mineral-composition specification using reflectance spectral data in visible and reflective-infrared regions

Masahide Kishimoto\*, Taiki Kubo\* and Katsuaki Koike\*

\*京都大学大学院工学研究科都市社会工学専攻 Department of Urban Management,  
Graduate School of Engineering, Kyoto University, Katsura C1-2, Kyoto 615-8540, Japan.  
E-mail: kishimoto.masahide.75z@st.kyoto-u.ac.jp

キーワード： クラスタ解析, 機械学習, 鉱物組成, リモートセンシング

Key words : Cluster analysis, Machine learning, Mineral composition, Remote sensing

### 1. はじめに

非接触非破壊で迅速かつ容易な鉱物組成の把握は、資源探査やトンネル掘削など、種々の分野で重要となる。ここに、衛星リモートセンシングで用いられる鉱物分布の推定手法が有効に応用できる。その代表的手法として、単一成分を端成分とし、線形混合モデルにより一画素中に含まれる存在量を推定する Linear Spectral Unmixing (LSU) (Settle & Drake, 1993) があげられる。現位置スケールにおけるリモートセンシングでは、主に RGB カラー画像、マルチスペクトル画像、ハイパースペクトル画像等が利用されている。その中でもマルチスペクトル画像はハイパースペクトル画像と比較してコストが小さいとともに、RGB カラー画像よりも多くの情報を得られるため実用性が高い。しかしながら、マルチスペクトル画像は、バンド幅や選択波長域によってその解析精度が異なり、いずれの組み合わせが現位置スケールでの解析に最適であるかは、詳細には検討されていない。そこで本研究では、まず室内試験で得られるデータを用いて機械学習による多クラス分類モデルを作成し、その感度分析を行った。多クラス分類モデルが重視する要素は LSU による端成分含有率推定でも重要であると考えられる。よって、寄与率の大きい要素にマルチスペクトルバンドを設計することを考案し、その妥当性を評価した。

### 2. 解析データと解析手法

#### 2.1 解析データ

多クラス分類モデルの感度分析では、トンネル等のサイトでの一般的な鉱物である石英、モンモリロナイト、ハロイサイト、パオロファイライト、ディッカイト、カオリナイト、方解石、白雲母の 8 種類を選び、それらの 400~2500 nm の反射スペクトルデータを解析の対象とした。解析では USGS と産総研のスペクトルライブラリを使用した。データ数が少ないため、各クラスにおけるデータ間を内挿することによってデータ拡張を行い、さらに生成されたデータ

群にノイズを付加したのもデータに加えた。データ間内挿にはオーバーサンプリング手法の一つである Synthetic Minority Over-sampling Technique (SMOTE) (Fernandez et al., 2018) を適用した。

設計したマルチスペクトルバンドの妥当性の検討では、室内実験によって得られた精製ベントナイトと石灰岩の反射スペクトルを用いた。精製ベントナイトはモンモリロナイト、石英、ハロイサイトが主成分で、石灰岩は方解石が主成分である。含有比を 20% ずつ変化させた精製ベントナイトと石灰岩の混合試料を作成し、それぞれの反射スペクトルを測定した。また、含水率 10%、20%、30% の試料も作成し、同様に測定した。

#### 2.2 多クラス分類モデルの感度分析法

多クラス分類法として、ニューラルネットワークとランダムフォレストを用いた。ニューラルネットワークは入力層、一つの間層、出力層の三層構造とし、中間層のユニット数は 1280 と設定した。ランダムフォレストでは、感度分析しやすいように各決定木の数を試行錯誤的に設定した。学習後に SHapley Additive exPlanations (SHAP) (Lundberg & Lee, 2017) によって分類の寄与度が大きい波長帯を選出した。SHAP はモデル解釈法の一つであり、協力ゲーム理論に基づいている。協力ゲーム理論では複数のプレイヤーが連携できる前提の下でゲームを行う場合、各プレイヤーの行動と配分すべき利得の最適解を数理的に分析する。本研究では SHAP 値の絶対値の平均を重要特徴量として求めた。

#### 2.3 設計したマルチスペクトルバンドの評価法

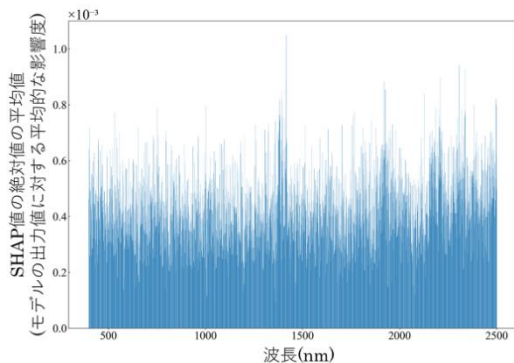
機械学習による分類の寄与度が大きい波長帯を用いて最適なバンドを設定し、設定されたバンド以外の情報を取り除いた。マルチスペクトルバンドの幅は 40 nm に固定した。バンドごとの反射率をそれぞれ平均化し、疑似的なマルチスペクトルデータを作り出し、LSU に基づく最急降下法による端成分含有率推定を行った。最急降下法では含有比が 0 以上 1 以下となるように制約条件を付した。実際の含有比

と推定含有比との二乗平均平方根誤差 (RMSE) を求め、その大小によって有効性を評価した。ここで、反射スペクトルデータに対する前処理として、リモートセンシングや信号処理で広く用いられる正規化処理、およびハルコーシェント処理 (以下、HQ 処理) (Tanaka et al., 2019) の 2 手法を選び、推定精度を比較した。地球観測衛星 Terra, Landsat 8 搭載の高分解能光学センサであるそれぞれ ASTER, OLI の可視・反射赤外域におけるバンドについても同様の処理を行い、比較の対象とした。

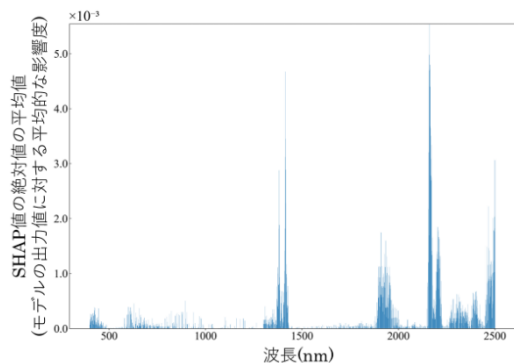
### 3. 結果と考察

#### 3.1 多クラス分類モデルの感度分析結果

ニューラルネットワーク、ランダムフォレストともに学習精度は高く、得られた学習器より SHAP を用いて特徴重要度を求めた。ニューラルネットワーク、ランダムフォレストにおける重要特徴度分布をそれぞれ第 1 図、第 2 図に示す。いずれの手法でも 1400 nm 付近、1900 nm 付近、2100~2500 nm 付近の特徴重要度が高いことが示されているが、ランダムフォレストの感度分析の方がより明瞭に重要特徴量が現れた。よって、ランダムフォレストの波長域に対応するマルチスペクトルバンドを設定した。設定したバンドは 400-440 nm, 1300-1340 nm, 1390-1430 nm, 1430-1470 nm, 1890-1930 nm, 2120-2160 nm, 2160-2200 nm, 2200-2240 nm, 2380-2420 nm, 2420-2460 nm, 2460-2500 nm の 11 バンドである。



第 1 図 ニューラルネットワークにおける SHAP に基づく重要特徴分布

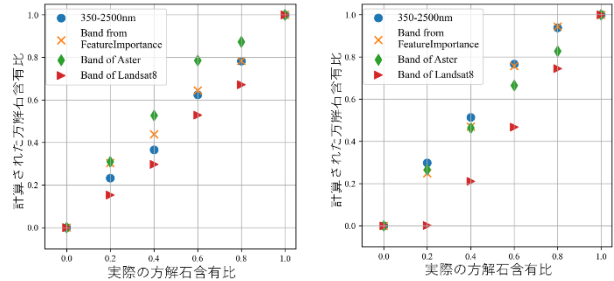


第 2 図 ランダムフォレストにおける SHAP に基づく重要特徴分布

#### 3.2 設計したマルチスペクトルバンドの評価結果

正規化処理, HQ 処理による RMSE はそれぞれ 0.060, 0.115 であり、可視反射赤外域の全波長を用いた解析と同様の精度が得られたことがわかった。方解石の実際の割合と LSU に基づく端成分含有比の関係を第 3 図と第 1 表に示す。一方、含水状態に対しては RMSE が 0.4 以上となり、正確

な推定ができなかった。その原因として水酸基の伸縮および撓曲による反射光の吸収が考えられる。このような鉱物組成特定の際の諸条件の影響については今後の検討課題であり、影響が軽減できれば推定精度がさらに向上すると期待できる。



第 3 図 実際の方解石含有比, および各マルチスペクトルバンドに基づき最急降下法によって計算された方解石含有比の比較. 前処理として正規化処理 (左), HQ 処理 (右) を用いた結果.

第 1 表 方解石の実際と推定による含有比との RMSE による比較

Band type	RMSE	
	正規化処理	HQ 処理
350-2500 nm	0.028	0.132
重要特徴量	0.060	0.115
Terra ASTER	0.130	0.058
Landsat 8 OLI	0.092	0.155

### 4. まとめと今後の課題

ランダムフォレストの SHAP 値は特徴重要度が明瞭に表れることから、最適マルチスペクトルバンドの設計には決定木系のアルゴリズムの一つであるランダムフォレストを用いるのが良いと考えられる。また機械学習によって得られた重要特徴分布に基づくマルチスペクトルバンドは、元のハイパースペクトルデータよりも情報量が大きく損失しているが、ハイパースペクトルでの解析と同程度の精度で鉱物含有比推定を行うことが可能になった。しかし、含水状態では大幅に推定精度が低下した。前処理や解析手法による推定精度の違いについても検討し、より正確な鉱物含有比推定を行うことが今後の課題である。

### 文 献

- Fernandez, A, Garcia, S, Herrera, F & Chawla, NV (2018) SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary, *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905.
- Lundberg, SM & Lee, S-I (2017) A unified approach to interpreting model predictions, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, Curran Associates Inc., Red Hook, NY, USA, pp. 4768–4777.
- Settle, JJ & Drake, NA (1993) Linear mixing and the estimation of ground cover proportions, *International Journal of Remote Sensing*, vol. 14, no. 6, pp. 1159–1177.
- Tanaka, S, Tsuru, H, Someno, K & Yamaguchi, Y (2019) Identification of alteration minerals from unstable reflectance spectra using a deep learning method, *Geosciences*, vol. 9, no. 5, 195.