

**科学研究費助成事業 研究成果報告書**

平成 28 年 6 月 10 日現在

機関番号：17301

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330199

研究課題名(和文) 教師データが限られた衛星画像のための半教師つき高精度画素分類アルゴリズム

研究課題名(英文) Algorithm of semi-supervised classification for remotely sensed images with restricted training data

研究代表者

喜安 千弥 (KIYASU, Senya)

長崎大学・工学研究科・教授

研究者番号：20234388

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：衛星画像を対象として限られた教師データから高精度のカテゴリ分類を実現するために、クラスタリングに基づく教師データの拡張、および、マルチスペクトル画像における分光特性の空間的な変動とカテゴリの空間的な分布状態に着目した、半教師つき手法による分類方法を開発した。開発したアルゴリズムを衛星画像に適用して実験を行った結果、教師データが限定された場合においても分類精度の向上を実現することができた。

研究成果の概要(英文)：We developed a method of semi-supervised classification for remotely sensed multispectral images to improve the accuracy of classification. The method includes expansion of training data based on clustering, classification considering the variation of spectral characteristics according to the location in the image, and classification considering the characteristics of spatial distribution of each category of objects in the image. We confirmed the validity of the method using remotely sensed data observed from satellites.

研究分野：パターン情報処理工学

キーワード：半教師付き分類 リモートセンシング マルチスペクトル画像

### 1. 研究開始当初の背景

人工衛星から地表を観測したマルチスペクトル画像を用いて、画素のカテゴリー分類を行い、土地被覆図を作成することは、リモートセンシングにおける重要な処理である。従来の教師つき分類は、教師データが十分に得られることを前提としている。ところが実際の画像中で、カテゴリーが既知の画素は大きく限られている。とりわけ教師データが空間的に偏った局所的な領域から与えられた場合、そのデータはしばしば特徴空間上でも偏っており、誤った分類結果をもたらす。これは、衛星画像の教師つき分類において、誤分類の避け難い要因となっている。

### 2. 研究の目的

本研究は衛星画像を対象として、限られた教師データから高い精度で画素のカテゴリー分類を実現する方法を開発することを目的とする。リモートセンシングによる観測画像の特徴をふまえ、この問題を解決する“半教師つき”アルゴリズムの確立をめざす。半教師つき(Semi-Supervised)手法は、教師データを用いる教師つき手法と、教師データを用いない教師なし手法の中間に位置するものであるが、ラベルづけされたデータとラベルのないデータが与えられた場合に、それらの両方を用いて学習することにより、ラベルづけされたデータだけで学習した場合よりも、精度の高い分類結果を実現しようとするものである。

### 3. 研究の方法

#### (1) マルチスペクトル画像のクラスタリングに基づく教師データの拡張

限られた教師データをもとに、未分類の画像全体から、新たな教師データとして利用できる画素を抽出する方法を開発する。新たな教師データは、画像中のトレーニングエリア(教師データの初期値)には含まれていないが、信頼性高くカテゴリーが推定できる画素を採用する。このとき、画素のスペクトルの類似性だけでなく、画像内の空間的な配置や、時間的な変化や類似性も考慮してクラスタリングを行うことで、信頼性高くカテゴリーが推定できる画素を抽出する。

#### (2) 拡張した教師データを用いる半教師つきカテゴリー分類

拡張された教師データが対象カテゴリーの分布をより適切に表現していれば、拡張教師データを用いることでカテゴリー分類の精度が向上すると期待できる。これは、教師なし分類(クラスタリング)と教師つき分類を組み合わせた半教師つき分類アルゴリズムである。ラベルなしの画素に対してもマージンを最大化するなどの方法で分類精度の向上をめざす。

### 4. 研究成果

#### (1) マルチスペクトル画像のクラスタリングに基づく教師データの拡張

まず画素のクラスタリングを行い、分光および時空間の特性が近い画素を抽出する。各クラス内のデータが同一のカテゴリーに由来するものと推定し、教師データが含まれるクラス内の画素を拡張教師データとして採用する。しかし、得られたクラスの中に異なる複数のクラスの教師データが含まれている場合には、クラスを単一のカテゴリーとして利用できないという問題がある。k-means 法によるクラスタリング結果は、最初に与えた初期クラスに大きく影響を受けることを考慮し、与えられた教師データに関して空間的な連続領域ごとの平均値を初期クラスタとして採用する方法を開発した。この方法をさらに改良し、与えられた教師データを用いてまず最尤法による予備的な分類(予備分類)を行い、次に分類の際の尤度が高い画素のみを選択的に抽出し、空間的に連続な領域を面積が大きい順に選択して初期クラスタとして設定し、クラスタリング結果に基づいて教師データを拡張する方法を開発した。

AVIRIS による観測画像を対象とし、教師データを大幅に限定して9クラスの分類処理の実験を行った結果を示す。

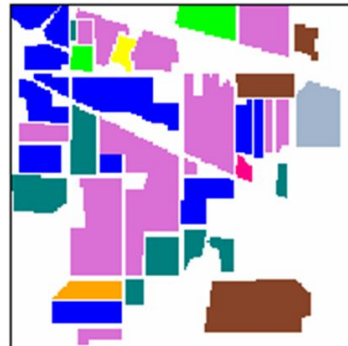


図1 AVIRIS 画像の Ground Truth

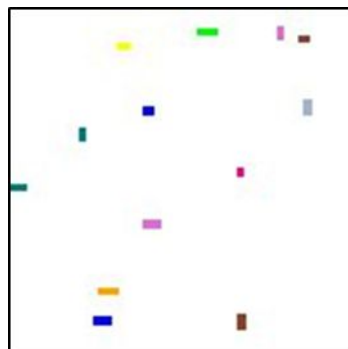


図2 設定したトレーニングエリア

図1の Ground Truth の中から図2に示す少数のトレーニングエリアを選択し、これらの少数のトレーニングデータを用いて分類が可能であるかを実験した。トレーニングデ

ータを用いて最尤法で画素を分類し、尤度の高い画素を抽出した結果を図3に示す。

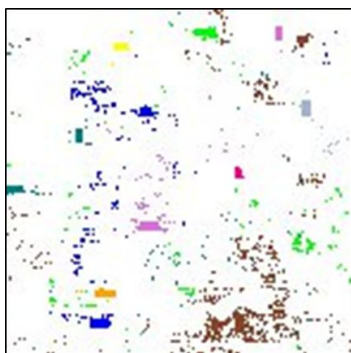


図3 予備分類における抽出画素

これらの中から画素が空間的に連続した領域を抽出し、面積が大きい順に初期クラスタとして設定してクラスタリングを行い、得られたクラスタについてカテゴリーの割り当てが可能と判定された画素を表示した結果を図4に示す。

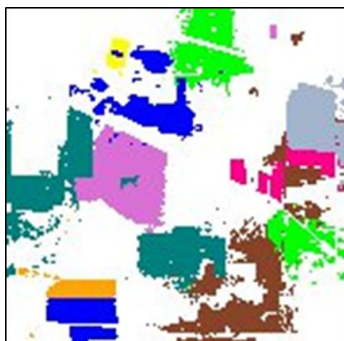


図4 クラスタリングによって抽出された画素

(2) 拡張トレーニングデータによる分類

図4の画素を拡張トレーニングデータとして、分類を行った結果を図5に示す。



図5 拡張トレーニングデータを用いた分類結果

初期クラスタ設定方法を上記のように改良したことによって、拡張教師データの適切性（正しい教師データが設定できた割合）が従来は約75%であったものを約80%へ改善することができた。この結果として得られた拡張教師データを分類処理に用いることにより、

分類正答率が約65%から約70%へと約5%の向上を達成することができた。

(3) 分光特性の空間的な変動と、カテゴリーの空間的な分布状態に着目した改良

分光特性の空間的な変動に対応するため、マルチスペクトル画像全体をメッシュ状に分割し、メッシュ領域内で最も多く分類されたカテゴリーの占める割合が一定値を超えていれば、その領域内の全画素を新たに該当するカテゴリーの拡張トレーニングデータとする方法を開発した。

人工衛星 Landsat による画像を利用した実験結果を示す。図6の画像に対して図7のような少数の初期トレーニングデータを選択して分類を試みた。抽出された拡張トレーニングデータを図8に示す。



図6 LANDSAT 画像

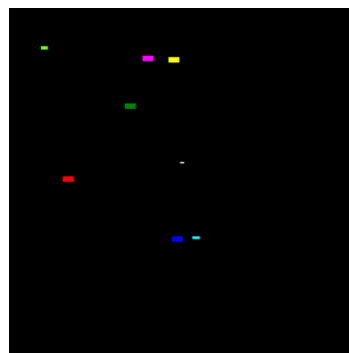


図7 初期トレーニングデータ

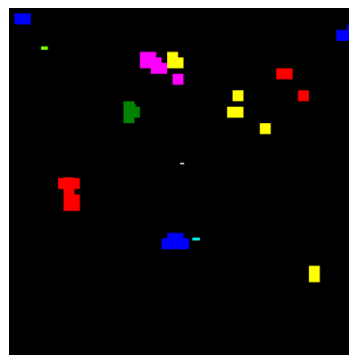


図8 拡張トレーニングデータ

これらの拡張トレーニングデータを用いて分類を行った結果を図9に示す。この例では、分類精度は64.5%から69.4%へ約5%

向上した。

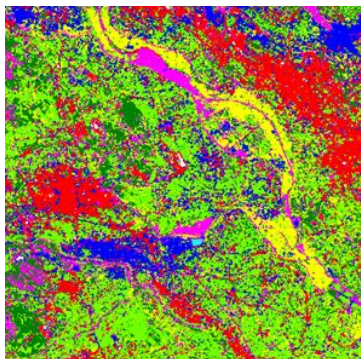


図9 拡張トレーニングデータを用いた分類結果

このとき、対象物の分光特性が画像内で変動する場合には、拡張トレーニングデータの偏りにより十分な効果が得られない場合がある。そこで画像全体をあらかじめ分光特性の変動を考慮して分割し、領域ごとにその領域内から得られた拡張トレーニングデータを用いて分類を行う方法も開発し、改良方法を用いることによって、分類正答率が向上を確認した。

#### (4) 今後の課題・展望

開発してきたアルゴリズムを改良しつつ実用性の検証を行ったが、対象データによってアルゴリズムが有効に機能する場合とそうでない場合があることが確かめられた。また、パラメータの選択が重要であることがわかった。半教師付きの手法は、データの特性によって有効な場合が異なることに注意しなければならないが、半教師付き手法の有効性をデータ自身から判定しつつアルゴリズムを自動的に切り替える方法などが必要になることがわかった。今後は、どのようなデータに対しても適切な結果を与えるロバストな手法に発展させていく必要がある。

## 5. 主な発表論文等

〔雑誌論文〕(計4件)

Senya Kiyasu, Daichi Maruyama, Kotaro Sonoda, Semi-Supervised Land Cover Classification of Remotely Sensed Image with k-menas++ Clustering, Proc. of the 37th RESES Symposium, 査読無, pp.52-57, 2015

Senya Kiyasu, Daichi Maruyama, Kotaro Sonoda, Semi-Supervised Land Cover Classification of Remotely Sensed Image with Improved Clustering Seeds Initialization, Proc. of SICE Annual Conference 2014, 査読有, pp.101-104, 2014

堀多恵美, 馬場涼介, 園田光太郎, 酒井智弥, 喜安千弥, マルチテンポラ画像を教師データとする衛星画像のカテゴリー分類, 電子情報通信学会九州支部 第22回学生会講演会講演論文集, 査読無, D-28, 2014

馬場涼介, 園田光太郎, 酒井智弥, 喜安千弥, 土地被覆の連続性を考慮した衛星画像の半教師付き分類, 電子情報通信学会九州支部 第21回学生会講演会講演論文集, 査読無, D-36, 2013

〔学会発表〕(計4件)

Senya Kiyasu, Semi-Supervised Land Cover Classification of Remotely Sensed Image with k-menas++ Clustering, the 37th RESES Symposium, 2015年9月6日, やすらぎ伊王島(長崎県・長崎市)

Senya Kiyasu, Daichi Maruyama, Kotaro Sonoda, Semi-Supervised Land Cover Classification of Remotely Sensed Image with Improved Clustering Seeds Initialization, SICE Annual Conference 2014, 2014年09月10日, 北海道大学(北海道・札幌市)

堀多恵美, 馬場涼介, 園田光太郎, 酒井智弥, 喜安千弥, マルチテンポラ画像を教師データとする衛星画像のカテゴリー分類, 電子情報通信学会九州支部 第22回学生会講演会, 2014年09月20日, 鹿児島大学(鹿児島県・鹿児島市)

馬場涼介, 園田光太郎, 酒井智弥, 喜安千弥, 土地被覆の連続性を考慮した衛星画像の半教師付き分類, 電子情報通信学会九州支部 第21回学生会講演会, 2013年09月23日, 熊本大学(熊本県・熊本市)

## 6. 研究組織

### (1) 研究代表者

喜安 千弥 (KIYASU, Senya)  
長崎大学・工学研究科・教授  
研究者番号: 20234388

### (2) 研究分担者

酒井 智弥 (SAKAI, Tomoya)  
長崎大学・工学研究科・准教授  
研究者番号: 30345003

園田 光太郎 (SONODA, Kotaro)  
長崎大学・工学研究科・助教  
研究者番号: 90415852