

H-013

ユーザニーズに応じた距離空間によるハイパースペクトル画像の分類

Classification of hyper-spectral images based on the metric space according to user needs

稲垣 佑哉† 石村 康生†* 和田 充雄† 佐鳥 新† 中澤 正博†*

Yuya Inagaki Kousei Ishimura Mitsuo Wada Shin Satori Masahiro Nakazawa

1. はじめに

リモートセンシングの技術は、鉱物資源の探索、土地の被服分類、農作物の生育状況の観察、海面温度の分布の観察などのさまざまな分野で用いられている[1]。この技術において近年注目を集めているのが、ハイパースペクトルセンサである。従来リモートセンシングで用いられてきたマルチスペクトルセンサは高々10バンドで対象物を観測するのに対し、ハイパースペクトルセンサは数十～数百ものバンドで対象物を観測する。つまり、マルチスペクトルセンサは対象物の離散的なスペクトル情報しか取得できないのに対し、ハイパースペクトルセンサは対象物の連続的なスペクトル情報を取得することができる[2]。よって、ハイパースペクトルセンサで取得されたデータに対しては、より高精度な画像分類が可能となる。

一方、画像分類においては、すべてのユーザが同じ分類結果を求めるとは限らない。また、最短距離法においては使用する距離定義によって分類結果は大きく異なってくる。例えば、コサイン距離は対象物の明度の変化に対応できる距離であり、マハラノビス距離は対象物の明度・彩度の変化に対応できる距離である。よって、本研究では各種距離の組み合わせによって、ニーズに適した距離を生成し、ユーザニーズに応じた分類結果を得ることを目的とする。

2. 提案手法

本研究では、まずテストデータを用いてユーザのニーズに適した距離定義を生成する。その後、その距離定義を使って実データを分類していく。また、距離生成には、各距離定義による分類精度の差異に注目して開発した組み合わせる距離定義の決定に優れたアプリケーションを用いる。提案手法のアルゴリズムを以下に説明する。

- a. テストデータをコサイン距離、マハラノビス距離、標準化ユークリッド距離、ミンコフスキー距離の4種類の距離により分類。
- b. 精度良く分類したいクラスを指定。
- c. bで指定したクラスの各距離による分類結果(誤棄却率・誤受入率)を提示。
- d. 独自の距離定義生成
 - d-1. 指定した各クラスの誤棄却率・誤受入率の閾値を設定する。
 - d-2. d-1で設定した閾値をすべてのクラスにおいて満たしていない距離定義を組み合わせの候補から

消去。

- d-3. d-2で残った距離定義の中から2つの距離を組み合わせ新たな距離定義とし、再度分類。
- d-4. d-3の分類結果がd-2で設定した閾値を満たすまでd-1～d-3を繰り返す。閾値を満たせば独自の距離定義決定。
- e. 生成された距離定義を用いて実データを分類。

各クラスの分類精度の閾値設定は、今回用いたアプリケーションによって縦軸誤棄却率、横軸誤受入率として二次元平面上で設定できる。そのためユーザにとって視覚的に分かりやすくニーズを設定することができ、かつ、ドラッグで操作可能であるため容易に組み合わせの距離候補を絞ることができる。よって、ユーザにとって独自の距離生成を容易に行うことができる。

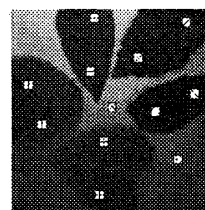
また、生成する距離定義は以下のように定義する。ここで、 d を生成した距離定義の値、 d_1 、 d_2 をそれぞれの組み合わせる距離定義の値、 α 、 β を d_1 、 d_2 をそれぞれ足し合わせる際の重み、ただし $\alpha + \beta = 1$ とする。

$$d = \alpha d_1 + \beta d_2$$

3. 評価実験

3.1 実験条件

画像サイズは 350×350 、観測波長帯は $0.4 \mu\text{m} \sim 0.8 \mu\text{m}$ 、観測バンド数は81バンドである。クラス数6のテストデータを使用する。このテストデータは、葉によって彩度の変化、照明の当たり具合による明度の変化が存在し、距離を組み合わせることによってそれぞれの変化に対応できるかを検証するために用いた。トレーニングデータは図1のように与える。また、提案手法におけるそれぞれの条件設定を表1のように設定する。このテストデータに対し、独自の距離定義を生成し実データの分類へと適用する。その分類結果を、単なる最短距離法の分類結果と比較し、提案手法の有効性を検証する。



class1 □ ピレア
class2 □ ブルーキャットアイ
class3 □ ハイビスカス
class4 □ スイセンジナ
class5 □ コリウス
class6 □ 背景

図1: テストデータ

† 北海道大学大学院情報科学研究科
†* 北海道大学 / 現在 宇宙航空研究開発機構
‡ 北海道工業大学
‡* 北海道総合研究調査会

表 1:条件設定

	誤受入率(%)	誤棄却率(%)
class1	15.0	15.0
class2	10.0	10.0
class4	10.0	10.0

3.2 実験結果

表 1 の条件設定において生成された距離は、コサイン距離とマハラノビス距離を 3:7 で組み合わせた距離であった。テストデータの分類結果を図 2 に示す。

テストデータの分類結果(図 2)から、コサイン距離は class4, マハラノビス距離は class1, class2 を精度良く分類できている。一方、生成した距離は両者の特徴を受け、class1, class2, class4 を精度良く分類できていることがわかる。よって、ユーザのニーズに対して適切な距離定義を生成できたといえる。

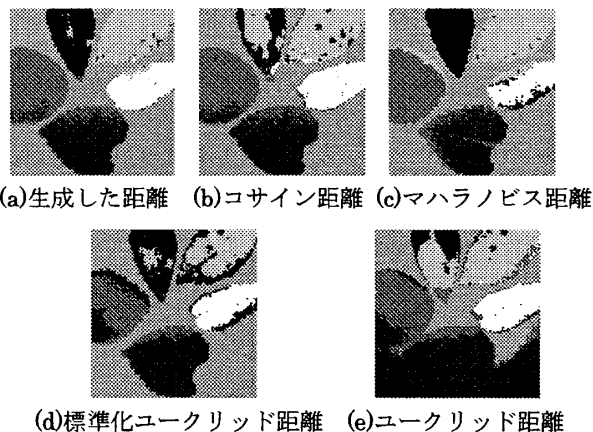


図 2:テストデータ(6 クラス)の分類結果

次に、生成した距離を使い、テストデータにおいて葉の位置や向きがそれぞれ異なる 5 種類の実データを分類する。実データ 5 種類における表 1 で示した 3 クラスの分類精度(誤受入率, 誤棄却率)の平均値を、生成された距離と他の距離と比較し、生成された距離の汎用性を検証する。

実データの分類結果を(図 3, 図 4, 図 5)に示す。ただし、マハはマハラノビス距離, 標準は標準化ユークリッド距離, ユーはユークリッド距離を表すものとする。誤受入率に関してはマハラノビス距離が他の距離と比べて一番小さい値

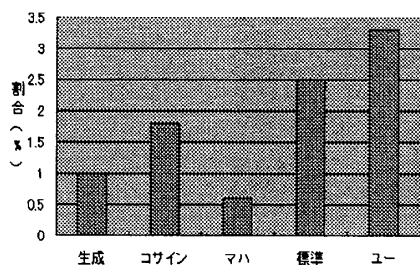


図 3:実データ 5 種類(3 クラス)の誤受入率の平均値

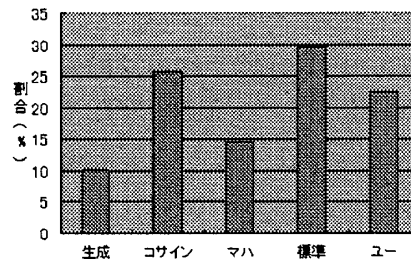


図 4:実データ 5 種類(3 クラス)の誤棄却率の平均値

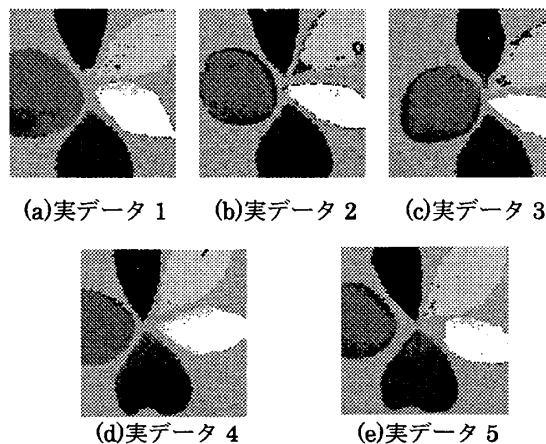


図 5:実データ(6 クラス)5 種類の分類結果

となっている。一方、誤棄却率に関しては生成した距離が他の距離と比べて一番小さい値となっている。生成した距離とマハラノビス距離では、誤受入率の差より誤棄却率の差の方が大きいため、生成した距離の方が良い結果を出しているといえる。このことから、提案手法によって生成した距離が汎用性を持つことができたといえる。

4. おわりに

本研究では、最短距離法においてユーザのニーズに応じた距離を生成し、ハイパースペクトル画像の分類に適用することでユーザのニーズに応じた分類結果を得るための手法を提案した。ユークリッド距離やコサイン距離を単体で用いる単なる最短距離法では、必ずしもユーザが指定した個々のクラス分類精度が良いわけではない。それに対し提案手法は、ユーザの注目しているクラスを全体的に精度良く分類することができた。

よって、提案手法は従来のユークリッド距離やコサイン距離などを単体で用いる最短距離法に比べ、ユーザのニーズに応じた分類結果を得ることができたといえる。

参考文献

[1] 古濱 洋治, 岡本 謙一, 増子 治信:人工衛星によるマイクロ波によるリモートセンシング, 社団法人 電子通信学会出版, pp.3-10, March 1986
 [2] 資源・環境観測解析センター:地球観測データからの情報抽出, 東京:資源・環境観測解析センター, pp.98-103, March 2001