

IV-268

衛星リモートセンシングデータを用いた土地被覆分類における
トレーニングデータの選定方法について

東京理科大学 正会員 大林 成行

東京理科大学 正会員 藤田 圭一

東京理科大学 正会員 小島 尚人

○東京理科大学 学生員 中島 剛彦

1. はじめに 衛星リモートセンシングデータから土地被覆分類図を作成する場合、トレーニングデータの良否が分類精度に大きく影響をおよぼす。広く一般に用いられている最尤法分類を用いる場合、トレーニングデータの均一性や代表性、そして正規性が満たされていなければならないことは言うまでもない。トレーニングデータを選定する場合、現地調査の結果や航空写真等の情報を基にして各土地被覆項目に対応する領域を取捨選択する。しかし、そのトレーニングデータが必ずしも現地の土地被覆状況を代表しているとは限らず、トレーニングデータの選定方法に関する研究は今なお重要な研究課題となっている¹⁾。本来理想的なトレーニングデータとは、人間が教師として選定するクラスがスペクトル特性によって忠実に現地の土地被覆項目を代表できるものでなければならない。そこで、本研究では予備的に選定したトレーニング領域に対して「教師つき分類」と「教師なし分類」をそれぞれ施し、両者の分類結果が一致した画素を新たなトレーニングデータとして再抽出する方法を提案し、最尤法による分類精度が向上するか否かについて検討した。

2. 使用データおよび評価領域 本研究で提案したトレーニングデータの選定方法の効果を評価するために使用した衛星データは1988年10月14日に観測されたSPOT・HRVデータ(PATH332-ROW278)である。また、評価領域は代表的な土地被覆項目が混在し、詳細に現地調査を実施しやすくグランドトゥルスデータが収集しやすいことを念頭に置き、千葉県印旛沼周辺を選定した。

3. 本研究で提案するトレーニングデータの再抽出方法

本研究で提案するトレーニングデータの再抽出方法は次に述べる5つのステップから構成される。

STEP 1 (トレーニングデータの予備選定) : 詳細な現地調査の結果をもとに候補となるトレーニング領域を予備的に選定する。

STEP 2 (トレーニング領域の教師つき分類) : 設定したトレーニングクラスを教師として、トレーニング領域の全画素を最尤法によって分類する。トレーニングデータにおいて自分自身のクラスに判別された画素を正当な画素とする。このステップでは、あらかじめ統計的に不適切な画素(混在ピクセル)をトレーニングデータから除去し、洗練化したトレーニングデータを抽出したこととなる。なお、本研究の提案手法の効果を検証する上で、分類に際してはHRVの3つのバンドすべてを用いた。

STEP 3 (トレーニング領域の教師なし分類) : STEP 2とは別の処理として、トレーニング領域の全画素を対象に、ISODATA法を用いてクラスターリングする。あらかじめ与える分類項目数(クラス数)は、STEP 2で設定したトレーニングデータのクラス数と同一とする。さらに、クラスターリング後の各クラスの統計量(バンド別平均値、分散共分散行列)を算出しておき、STEP 4の処理に備える。

STEP 4 (STEP 2とSTEP 3の分類後のクラスが一致した画素の抽出) : STEP 3の処理によって決定されたクラスは、スペクトル的にクラス間の分離性は保証されるものの、現地の土地被覆項目との対応付けができていない。一方、STEP 2で決定された各クラスは現地の土地被覆項目との対応付けができており、最尤法によって混在ピクセルが除去され洗練化された教師データではあるもののスペクトル的にクラス間の分離性が保証されるとは限らない。そこで、これら両者の欠点を相互に補うために、STEP 2とSTEP 3の処理後のクラスが一致した画素を抽出する。この時、一致する画素を抽出する際の各クラスの対応付けが問題となる。本研究で提案する手法では、STEP 1で予備的に選定したトレーニングクラスとSTEP 3で決定されたそれぞれのクラス間の「マハラノビスの分布間距離」を指標として、クラスの対応付けを統計的かつ自動的に行う。すなわち、クラス間のマハラノビスの距離が最も短いものを同一クラス

に対応するものと見なす。本研究で予備的に選定したトレーニングデータに対して、以上の処理を行った結果を表-1に示す。表中、「A」欄は、予備的に選定したトレーニングクラスを示し、「B」欄はこのトレーニングデータの全ての画素を対象に ISODATA法によって分類されたクラスである。マハラノビスの距離が最も小さい値

を枠で囲って示してある。この結果を見ると、例えば畑として選定したClass-3は、ISODATA法によって分類された Class-5に対応することが判る。このようにクラスに対応付けが完了した後に、STEP 2とSTEP 3のそれぞれの処理後のクラスで一致する画素を抽出する。以上のような統計処理に基づいて自動的に、しかも簡便な手順で精度の高いトレーニングデータを再抽出できる点が本提案手法の特徴である。

STEP 5 (最尤法による画像全体の分類) : STEP 4の処理によって再抽出されたトレーニングデータを教師とし、最尤法を用いて画素単位で画像全体を分類する。

3. 堤案手法の適用効果 再抽出前後のトレーニングクラス間の分離性の向上を確認するために「マハラノビスの距離=1」の等確率楕円を図-1に示した。「畑-樹林-芝地」間、「裸地-市街地」間の分離性の向上が顕著であり、全体的に見てもクラス間の分離性は向上している。また、本研究では χ^2 検定によってバンド別にトレーニングクラスの正規性が良くなることも確認している。さらに、再抽出前後のそれぞれのトレーニングデータを用いて最尤法分類を行い、分類精度を比較した結果、評価用データを基に算出されるPCC (Probability of Correct Classification) の値は84.9%から92.6%まで向上した。その他、トレーニングデータに対する区分精度や誤分類率、クラス間分散度の向上についても確認しているが、これらの結果の詳細については紙面の都合上、口頭発表時に説明する。

4. まとめ 本研究で提案したトレーニングデータの再抽出方法を用いれば、各クラスの正規性も改善するとともに分類精度が向上することが判った。さらに、この方法は複雑な処理手順を踏むことなく、統計処理に基づいて精度の高いトレーニングデータの再抽出と混在ピクセルの除去が自動的に行え、極めて実用性の高いことが判った。今後は空間情報を利用して混在ピクセルを除去する方法等を併用し、さらに分類精度が向上するか否かを検討することが課題としてあげられる。

【参考文献】 1) 飯倉善和、安岡善文：EMアルゴリズムを用いたトレーニングデータの代表性の改善、日本リモートセンシング学会誌、Vol.9、No.4、PP.5~PP.13、1989年

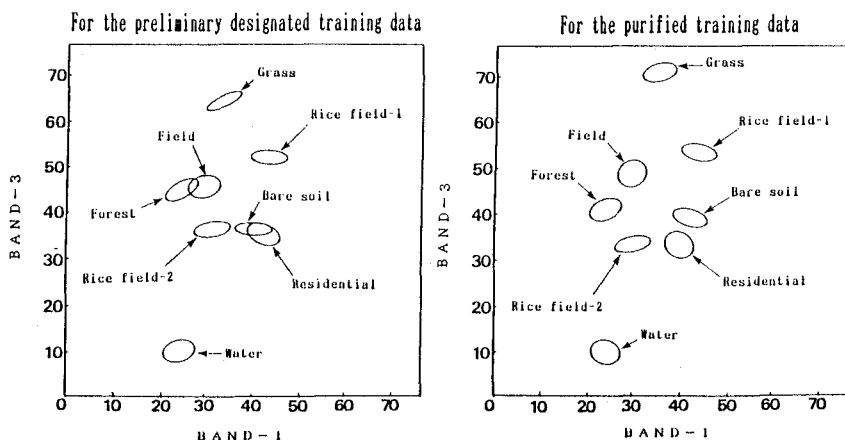


図-1 再抽出前後のトレーニングクラス間の分離性の比較

(マハラノビスの距離 $D^2=1$ の等確率楕円: HRVデータ)

表-1 トレーニングクラス間のマハラノビスの距離

	A							
	Class-1 Rice field-1	Class-2 Rice field-2	Class-3 Field	Class-4 Forest	Class-5 Residen- tial	Class-6 Grass	Class-7 Bare Soil	Class-8 Water
B	Class-1	31.6	14.5	14.8	13.2	24.1	22.1	26.0
	Class-2	6.4	0.7	2.1	6.6	5.1	5.5	5.4
	Class-3	11.0	5.0	3.4	0.9	9.7	5.7	11.0
	Class-4	4.9	3.4	5.4	13.0	1.8	7.8	1.9
	Class-5	6.2	4.0	1.2	3.4	6.3	2.4	8.5
	Class-6	4.6	5.4	8.0	17.0	2.2	10.4	1.8
	Class-7	0.4	4.8	7.7	17.3	5.0	8.0	4.4
	Class-8	8.5	6.7	4.2	7.3	10.0	0.9	12.2

Notes) A : Classes of the preliminary designated training data

B : Classes classified with unsupervised classification (ISODATA)
for the preliminary designated training data