

軌道衛星画像による教師なし分類法の適用とその処理時間*

1P-3

小原 裕史, 星 仰†

茨城大学‡

1 はじめに

リモートセンシングの衛星画像の中でもここでは軌道衛星の画像を取り上げ（例えばJERS-1）、この画像に対して教師なし分類法のアルゴリズムを適用し分類を行う。これらのアルゴリズムは、階層的手法と非階層的手法に分かれる。階層的手法では、最短、最長距離法、メディアン法、重心法、群平均法、Ward法を用い、非階層的手法ではK-means法を用いる。これらの手法に対するアルゴリズムに対してJERS-1等の画像データを適用し、その処理時間を比較検討する。

2 教師なし分類法

2.1 K-means法

本研究で適用したK-meansクラスタリングをアルゴリズム化したときの概略手順を要約する。

1. 目的画像からランダムまたは一定間隔に画素をN個抽出する。これ以外にクラスタ間の距離Tを設定する。
2. 抽出したデータに対して、まず既存クラスタの中で最も近いクラスタ K_i を求める。その距離DがT以下ならば未知画素データをクラスタ K_i に融合し、クラスタセンタを再度計算し直す。Tより大きければその未知画素データを新しいクラスタ K_n とする。
3. 全画素に対してクラスタ K_i との距離 D_i を計算し、一番近いクラスタをその未知画素データのクラスタ K_0 とする。

2.2 階層的手法

教師なし分類の階層的手法のアルゴリズムは式(1),(2)のクラスタ間の距離に基づいて作成する。

$$D_{fg} = \alpha_h D_{fh} + \alpha_l D_{fl} + \beta D_{hl} + \gamma |D_{fh} - D_{fl}| \quad (1)$$

$$D_{fg}^2 = \alpha_h D_{fh}^2 + \alpha_l D_{fl}^2 + \beta D_{hl}^2 \quad (2)$$

このときの距離空間は図1のようなユークリッド距離である。ここで α, β, γ は定数であり、各手法によって表1のようになる。

また処理手順は下記のようになる。

1. あらかじめ求めたいクラスタ数Nを決定する。
2. 原画像データからランダムまたは一定間隔（メッシュ状）に画素データをn個抽出する。

3. データの総数nをそれぞれクラスタとし、クラスタ間の距離を計算する。
4. 抽出したデータに対し、クラスタ数がNになるまで以下の処理を続ける。
 - (a) aとbのクラスタの距離を D_{ab} とするとき、 D_{ab} の中で最小な距離 D_{hl} を求める。
 - (b) クラスタhとlを融合する。融合されたクラスタをgとする。
 - (c) gクラスタに属していないクラスタfとの距離 D_{fg} を式(1)または式(2)によって求める。
5. 全画素に対してクラスタ K_i との距離 D_i を計算し、一番近いクラスタをその未知画素データのクラスタ K_0 とする。

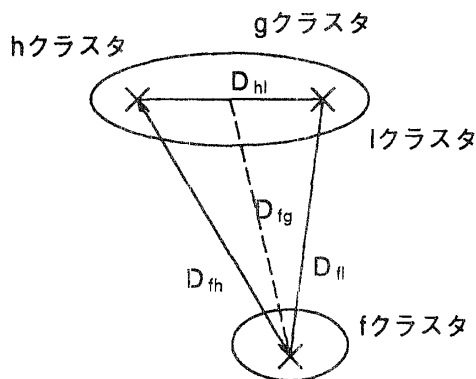


図1 クラスタの距離

表1 階層的クラスタリング手法

手法	α_h	α_l	β	γ	距離の定義
最短距離法	0.5	0.5	0	-0.5	D
最長距離法	0.5	0.5	0	0.5	D
メディアン法	0.5	0.5	-0.25	0	D
重心法	n_h/n_g	n_l/n_g	$-n_h n_l / n_g^2$	0	D^2
群平均法	n_h/n_g	n_l/n_g	0	0	D^2
Ward法	$\frac{n_l + n_h}{n_l + n_g}$	$\frac{n_l + n_h}{n_l + n_g}$	$-\frac{n_l}{n_l + n_g}$	0	D^2

3 適用結果および考察

すでに述べてきた階層的手法と非階層的手法の教師なし分類法をJERS-1.OPSの画像に適用した。このJERS-1.OPSは1993年3月に抽出された画像データである。

*Application of unsupervised classification for orbital satellite images and its computer processing time

†Yuji Obara, Takashi Hoshi

‡Ibaraki University

4-12-1 Naka-narasawa, Hitachi, Ibaraki 316, Japan

分析地区は茨城県の日立市周辺の久慈川・日立港を含む。分析地区は512x512画素と1024x1024画素サイズを用い、各々の分析地区の教師なし分類をかけたときのその処理時間を求めた。これらの結果を図2、3に示す。また分類結果画像の一部を図4、5(共にクラス数は15)に示す。

上記の分類法のアルゴリズムから推察されるように、本研究で適用したK-means法は、1つのデータがいずれかのクラスに融合されたら、そのデータは他のクラスに融合されることはない。また、最終的なクラス数Nは距離Tに大きく依存しており、1度分類結果を算出しなければ最終的なクラス数が分からない。

しかし、階層的手法は目的クラス数Nだけを与えるだけで分類可能である。だが、最小距離のクラスを何度も求め、データの所属クラスは変化しつづけるため処理に長時間を要する。

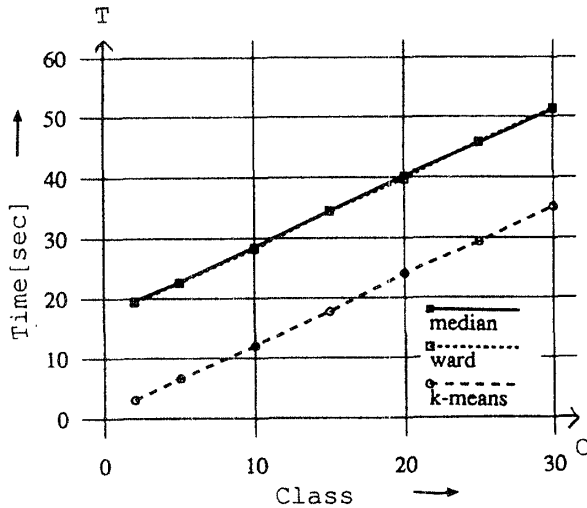


図2 処理時間の比較 (512x512)

図2、図3から分類処理時間はクラス数に比例していることが理解できる。アルゴリズムからメディアン法よりWard法の方が分類処理に長時間かかるが、距離の定義の違いのため、メディアン法の方が初期のデータ間の距離を求めるのに長時間かかり、最終的にはほぼ処理時間が等しくなっている。分類画像を比較してもK-means法とWard法ではかなり類似している。

4 おわりに

以上から教師なし分類においてK-means法は高速で処理できました、分類画像においても主なクラスタについては階層的手法と大差ないため高速性を要求されたとき有効であろう。

参考文献

- [1] 星 仰, 小原 裕史, 鳥井 清司, 佐藤 一紘: 画像データのK-meansクラスタリングの高速処理、写真測量学会平成8年度秋季学術講演会発表論文集、pp9-12
- [2] 星 仰: 地形情報処理学、森北出版、1991、pp165-168

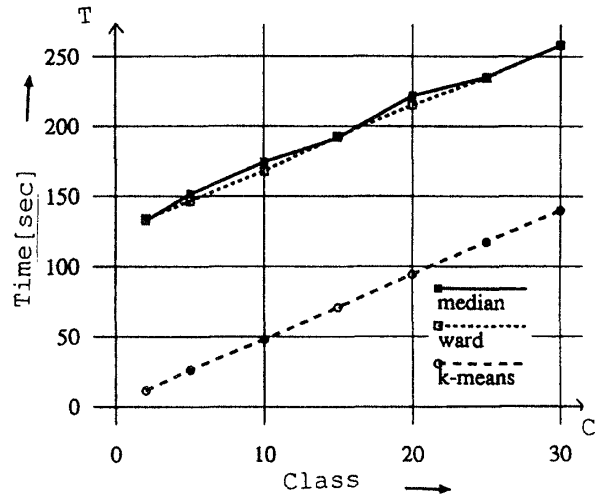


図3 処理時間の比較 (1024x1024)

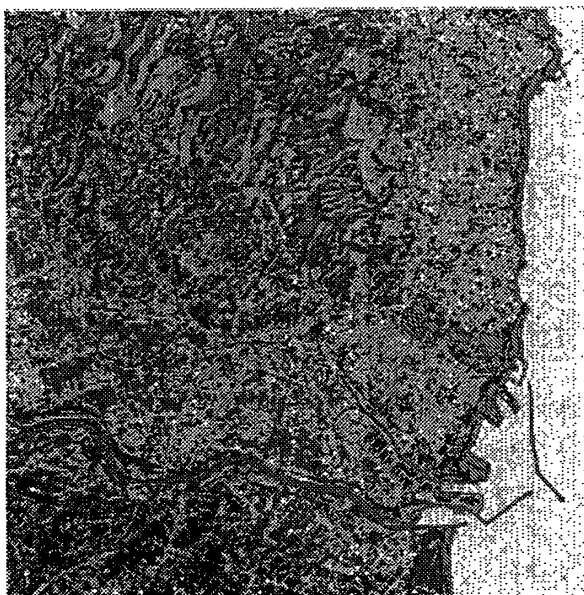


図4 Ward法による分類結果 (512x512)

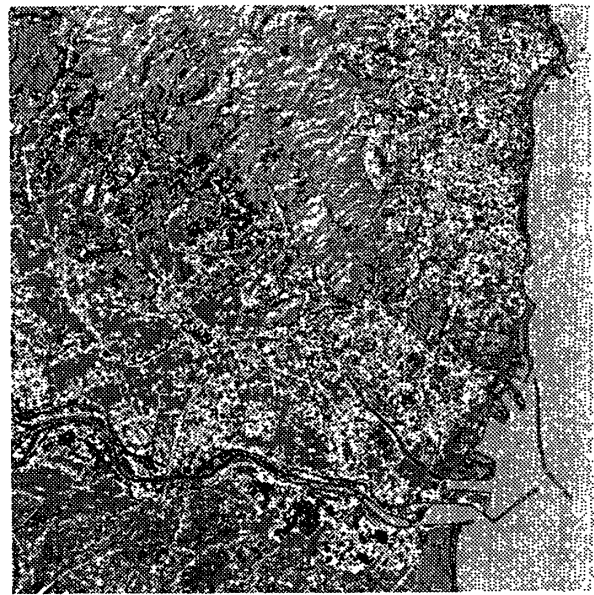


図5 K-means法による分類結果 (512x512)