

NOAA-AVHRR画像の雲域の分類

5Q-7

松本 裕司

高木 幹雄

東京大学工学部

東京大学生産技術研究所

1 はじめに

近年、地球的規模の環境問題に対する関心が高まってきており、このような広範囲を同時観測できる衛星データはとても有効であると考えられる。東大生産技術研究所では、アメリカの地球観測衛星NOAAと日本の気象衛星「ひまわり」からのデータを受信し、一括保管する画像データベースの構築を推し進めている¹⁾。

より利用しやすいデータベースを構築するにあたり、衛星画像中の雲領域は大きな問題の一つである。画像中の雲領域の有無は、陸・海に関心のある利用者にとっても、雲そのものに関心のある利用者にとっても重要な問題である。

そこで利用者の検索効率の向上をはかることを研究の目的とし、各々のデータに対し雲領域の割合、あるいは雲のおおまかな種別を付加価値として与えることを検討する。

2 気象衛星画像の特徴

NOAAのAVHRRセンサでは、可視2channel赤外3channel、GMSのVISSRでは可視1channel、赤外3channelから構成される。それぞれのバンドは、異なる放射スペクトル特性を持っているため、地表面の状況によって異なる輝度値を示す。また、可視光帯の画像は、夜間は使用できない。

3 海・雲領域の分類

本研究においては経験的な知識をなるべく利用しない方法をとっていくことにした。この方法は、特に地域、季節に限定することはないので、汎用性があるという長所が考えられる。

また、当面海と雲のみを含む画像を処理の入力画像とした。処理画像中に陸地を含むとカテゴリの数が増えるため、より処理が難しくなることが懸念されたためである。陸を含む画像に関しては、地球に関するグローバルデータベースに含まれる海岸線データ等を用いることも検討している。

本研究は、次のような段階から構成される。

- 特徴空間の設定、特徴抽出
- 特徴量からクラスタへの分類
- クラスタからカテゴリへのラベリング

3.1 特徴空間の設定、特徴の抽出

特徴量の選択として、各バンドのデータをそのまま用いたり、あるいは各バンドのデータの差(たとえば、ch.1 - ch.2)、画像を小領域に分けた時の分散の値、アルベド値、NDVI等様々なパラメータを用いることができる。本研究では、分散が大きいデータを選択して特徴量として用いることにした。しかしながら、データの意味を考える必要もあり、やみくもに使用することはできない。

3.2 特徴データの分類

特徴空間において画像の領域分割を行なう方法としては様々な方法を挙げることができるが、「各クラスタの標準パターンが不定」「クラスタの個数が未知」である本研究の場合の特徴である。

また、膨大なデータを少時間で扱うことが必要であるため、処理時間の短縮も要求される。

そこで本研究では、クラスタの分離にヒストグラムを用いることにした。特徴空間を一つの特徴軸によって定義される1次元特徴空間に写像した時の分布をヒストグラムで表す。この時、ヒストグラムの横軸に、特徴量をK-L変換し得られた高位の主成分軸を特徴軸とする。このようにして、一度に複数のクラスタに分類が可能にする。そして、さらに分類されたクラスタのうち絶対数の少ないものに対しては、重心の最も近いクラスタに結合させるものとする。

一般にヒストグラムによる分類は、ヒストグラムの横軸の間隔のとり方、閾値の決め方が大きな問題となってくる。ここでは、横軸の間隔は経験的なものにたよるとして、閾値の決定は、得られたヒストグラムを適当なフィルタをかけることによって小規模の振動を除去し、大局的な極小値において閾値を設ける。

3.3 カテゴリの割り当て

分類された各々のクラスタがいったい何を意味しているクラスタなのか個々のクラスタについて意

味との対応づけをする必要がある。「海」、「雲」を表すクラスタ、雲のカテゴリはさらに「上層」、「中層」、「下層」程度の分類を考えている。上層、中層、下層という分け方は、赤外面像の温度によってある程度求めることができる²⁾。そこで本研究においては、クラスタの重心の赤外領域における温度によってカテゴリの判断を行うことを検討している。

4 実験

以上、述べてきたことをもとに NOAA の AVHRR 画像を用いて実験を行った。

4.1 実験の方法

1. 1993/06/01 17 時, NOAA11 号 (東経 135.5-142.5, 北緯 29.0-34.0) の画像データを用いた
2. 特徴量として「ch.1」、「ch.4」、「ch.5」を用いた
3. 特徴量を K-L 変換し、高位の主成分軸上へ投影する (第 1、第 2)
4. 得られた各主成分軸上でのヒストグラムを Gaussian filter($g(x, \sigma)$) をかけてぼかした ($\sigma = 10$)

$$F(x, \sigma) = f(x) * g(x, \sigma) = \int_{-\infty}^{\infty} f(u) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-u)^2}{2\sigma^2}} du \quad (1)$$

5. ぼかしたヒストグラムの極小値をクラスタの閾値とし、クラスタを決定する

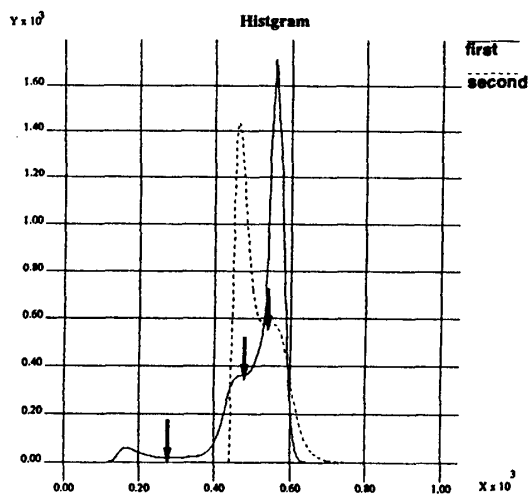


図 1: 第 1、2 主成分軸上でのヒストグラム (フィルタ ($\sigma = 10$) を通した結果)

4.2 結果、考察

この 3 次元データを K-L 変換し第 1,2 主成分軸上において投影したのが図 1 である。

また、この 3 つのクラスタを図に示したものが図 2 である。ch.4 のデータに基づき、クラスタの重心の最も温度の高いクラスタ (黒色) を海と判断する。雲領域は 2 つに大きく分かれたが、重心の温度値の低い方が高度の雲 (白色) であることが分かる。また、海 76.477%、下層の雲 19.420%、上層の雲 4.102% だった。

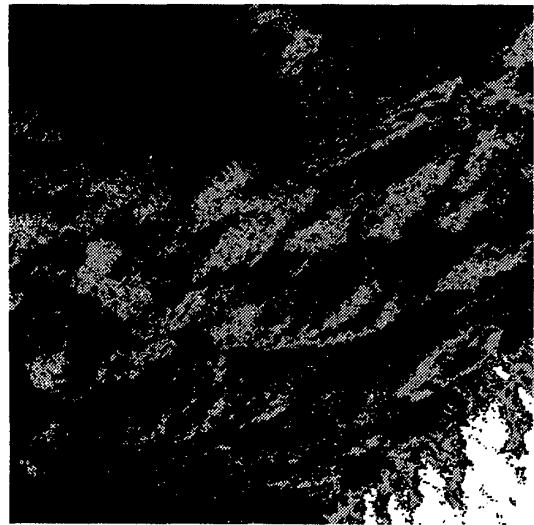


図 2: クラスタリングの結果

5 おわりに

衛星画像データベースの付加価値として、マルチチャンネル画像の主成分軸上におけるヒストグラムを利用して雲領域を分類する方法を提案した。

今後は、いろいろな画像で試行をしていくとともに、カテゴリへの対応付けも検討していく。また時間、地域との関連付け、クラスタリングする際のエリアの大きさの検討もしていかなければならない。

参考文献

- [1] RGOE Home Page. UTL: <http://www.tkl.iis.u-tokyo.ac.jp/RGOE/>.
- [2] Elik Liljas. USE OF THE AVHRR 3.7 MICROMETER CHANNEL IN MULTISPECTRAL CLOUD CLASSIFICATION. Number 2. SMHI promis-rapporter, 1986.