

# シーンの曖昧性を考慮した画像認識のためのファジィクラスタリング Fuzzy Clustering for Image Recognition Considering The Ambiguity of Scene

島崎 謙†  
Ken Shimazaki

長尾 智晴†  
Tomoharu Nagao

## 1. まえがき

近年、自律移動ロボットやデジタルカメラ、そして画像検索などへの利用を目的とした、計算機に画像中の環境を認識させる「シーン認識」に関する研究が盛んに行われている。シーンを認識することができれば、環境中の物体を認識することにも繋がるため、非常に重要な研究であるといえる。しかし、「シーン」という枠組は曖昧であり、具体的な定義は存在しない。シーンの曖昧性の例を図 1 に示す。画像の見た目として非常に類似した画像であっても、異なるカテゴリとして扱われている画像が多く見受けられる。これまでに提案されてきた多くの手法は単一のカテゴリへの分類、認識を行っているものが多い。しかし、明確な定義ができない枠組みを単一のカテゴリに割り当てることはあまり有用ではないと考えられる。

そこで本報告では、画像を単一のシーンには分類せず、複数のシーンに属することを許容するファジィクラスタリングに着目し、シーンの見た目や構造の曖昧性を重視した画像分類や解析を行う。

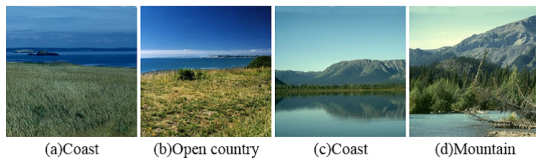


図 1: シーンの曖昧性の例

## 2. シーンを考慮した特徴算出

本報告では、シーン画像を考慮した特徴量の算出方法についても検討を行う。シーンは曖昧性を含むカテゴリであると述べたが、やはり各シーンには共通した特徴があるのではないかと考えられる。その例を今回実験に使用したデータセット [1] の画像を用いて図 2 に示す。図 2 は各シーンカテゴリの平均画像であり、特に Open country と Highway では、画像の上部と下部で色成分に大きな違いが見受けられる。また、Forest は画像の位置によらず色成分が一樣であることがわかる。そこで、原画像  $I$  から特徴量を算出するのではなく、以下の式を用いて原画像の座標  $(x, y)$  ごとに重み付けを行った画像  $W$  から特徴量の算出を行う。

$$W(x, y) = N(\boldsymbol{\mu}_{(x,y)}, \boldsymbol{\sigma}_{(x,y)}^2) * I(x, y) \quad (1)$$

ここで、 $N(\boldsymbol{\mu}_{(x,y)}, \boldsymbol{\sigma}_{(x,y)}^2)$  は座標ごとの平均、分散の正規分布を表している。なお、今回は特徴量として色成分に関する統計量とテクスチャ特徴 [2] を用いた。

†横浜国立大学大学院環境情報学部  
Graduate School of Environment and Information  
Sciences, Yokohama National University

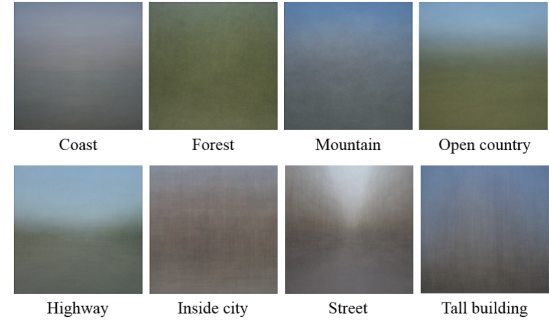


図 2: データセット [1] 中の各カテゴリの平均画像

## 3. クラスタリング手法

本報告ではシーン画像のクラスタリングを行うために、代表的なクラスタリング手法である K-means 法 (KM)[3] とファジィの概念を導入した Fuzzy C-means 法 (FCM)[4] を用いた。それぞれのアルゴリズムについて以下に示す。

### 3.1. K-means 法 (KM)

非階層型のクラスタリング手法の一つで、クラスタの平均を用いて、あらかじめ与えられたクラスタの個数に分類する手法であり、以下の目的関数を最小化するクラスタ集合  $V$  を求めることに等しい。

$$I(V) = \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} \|\mathbf{x}_j - \mathbf{v}_i\| \quad (2)$$

ここで、 $K$  はクラスタの総数、 $N$  はデータの総数、 $\|\cdot\|$  はユークリッド距離を、 $\mathbf{v}_i$  はクラスタの重心を表している。

### 3.2. ファジィ C-means 法 (FCM)

主なアルゴリズムは KM と同様である。しかし、KM がデータを単一のクラスタのみに割り当てるのに対し、FCM はデータに各クラスタに所属する割合（以下、帰属度と呼ぶ）を持たせることで複数のクラスタに属することを許容しており、以下の目的関数を最小化するクラスタ集合  $V$  と帰属度集合  $U$  を求めることに等しい。

$$J(U, V) = \sum_{i=0}^{K-1} \sum_{j=0}^{N-1} (u_{ij})^m \|\mathbf{x}_j - \mathbf{v}_i\| \quad (3)$$

ここで、 $u_{ij}$  は  $j$  番目のデータの各クラスタへの帰属度、 $m$  は  $m > 1$  を満たす定数を表している。なお、本報告では  $m = 1.2$  とした。また、式 (3) は以下の制約条件の下で最小化しなければならない。

$$\sum_{i=0}^{K-1} u_{ij} = 1 \quad (0 \leq u_{ij} \leq 1) \quad (4)$$

#### 4. クラスタリング実験

Oliva らが作成したデータセット [1] から各カテゴリの画像 100 枚, 計 800 枚を用いたクラスタリング実験を行う. なお, クラスタリング結果の性能評価の指標にはエントロピー (Entropy) と純度 (Purity) を使用した. それぞれの評価指標の算出方法を以下に示す.

$$Entropy = -\frac{1}{\log N_k} \sum_{i=0}^{N_c-1} \frac{n_i}{n} \left( \sum_{j=0}^{N_k-1} \frac{n_{ij}}{n_i} \log \frac{n_{ij}}{n_i} \right) \quad (5)$$

$$Purity = \frac{1}{n} \sum_{j=0}^{N_k-1} \max(n_{ij}, j = 0, 1, \dots, N_c - 1) \quad (6)$$

ここで,  $N_c$  はクラスタの総数,  $N_k$  はカテゴリの総数,  $n_i$  は  $i$  番目クラスタの総データ数,  $n_{ij}$  は  $i$  番目のクラスタ内のカテゴリ  $j$  のデータ数を表わしている. エントロピーはクラスタ内でのカテゴリの混ざり具合を表す指標であり, 結果が良いほど低い値を示す. 純度はクラスタ内の最も多くのデータ数を占めるカテゴリの割合を表わす指標であり, 結果が良いほど高い値を示す.

今回, FCM を用いたクラスタリングでは帰属度が最大のクラスタへの割り当てを行う. また, KM, FCM ともアルゴリズムにランダム要素を含むため, 結果の評価には 10 試行の平均を用いる. 各クラスタリング手法, 各特徴算出方法を用いた結果を, クラスタ数を横軸としたエントロピーと純度の推移グラフ図 3, 図 4 にそれぞれ示す.

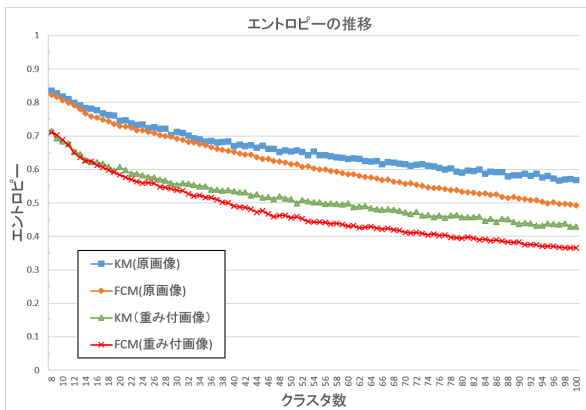


図 3: エントロピーの推移のグラフ

クラスタの総数が増えると各クラスタに含まれるデータの数が減少するため, エントロピーは減少し, 純度は上昇している. クラスタリング手法の違いによる結果では, 特徴量の算出方法に関係なく KM よりも FCM を用いたクラスタリング結果がわずかに良好である. これは, FCM が KM に比べ単純にデータ間の距離を用いるのではなく, 各クラスタへの帰属度に応じて距離を算出できるため, 妥当なクラスタ重心へ収束できたためと考えられる. また, 特徴量の算出方法の違いによる結果の違いでは, 原画像よりも重み付画像から算出

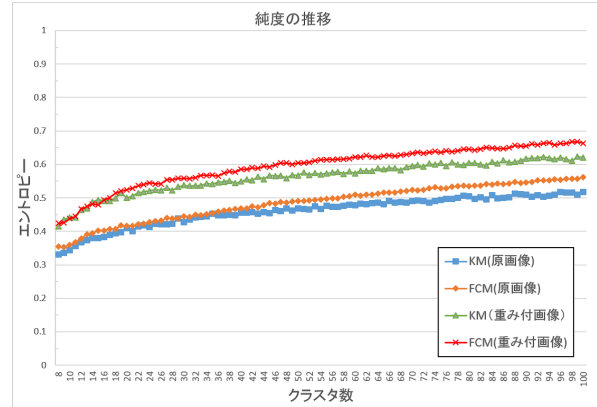


図 4: 純度の推移のグラフ

した特徴量を用いたクラスタリングの精度が高く, 少ないクラスタ数であってもエントロピー, 純度ともに原画像を用いた結果に比べ精度に大きな差が生じている. これは, 各シーンに顕著な色成分のみが強調されたため, 特徴空間における各カテゴリ間のデータのばらつきが原画像を用いた特徴空間よりも大きくなったためだと考えられる. 以上の結果を踏まえ図 3, 図 4 のグラフをみると, 重み付画像を用いた FCM によるクラスタリングの結果が最も良好であることが分かる.

#### 5. まとめと今後の課題

本報告ではシーンカテゴリごとの平均画像を用いて重み付を行った画像から特徴を算出し, それらを用いた KM と FCM によるクラスタリングの性能を評価した. 結果として, 重み付画像を用いたファジィクラスタリングにおけるエントロピー, 純度, ともに最も結果が良かった.

今後の課題として, クラスタリングの精度向上だけでなく, クラスタ内のシーン画像の類似度についても検討を行う必要がある. また, 今回はファジィクラスタリングに従来手法である FCM を用いたが, クラスタリング手法についても提案していきたいと考えている.

#### 参考文献

- [1] Oliva. A and Torralba. A, "Modeling the shape of the scene, A holistic representation of the spatial envelope," International Journal of Computer Vision, Vol.42, no.3, pp.145-175, 2001.
- [2] Jin, Hongliang and Liu, Qingshan and Lu, Hanqing and Tong, Xiaofeng, "Face detection using improved LBP under bayesian framework," Image and Graphics, 2004. Proceedings. Third International Conference on, pp.306-309, 2004.
- [3] M.R.Anderberg, "Cluster Analysis for Applications," Academic Press, New York, 1973.
- [4] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithm," Plenum Press, 1981.