

# 料理画像認識の高精度化

## Accuracy Improvement for Recognition of Food Images

堂土 奨†      佐野睦夫†  
Sho Dodo      Mutsuo Sano

### 1. はじめに

毎日の食事の栄養推定は、健康サポートや食生活の管理に重要な情報源である。最近では、健康などのアプリケーションツールとして Food Log や食事管理のインターネットサービスが実装されている。そのツールでは、ユーザが何を食べたかをログとして残し食生活の改善が行える。食生活の改善には食事内容の記録が効果的であるということが実証されていることもあり、食事の記録は摂取した栄養バランスを知ることができ、摂取不足や過剰摂取を防ぐことができる。これを実装するために、食材認識が必要とされている。食材学習を用いない従来方式では、Bag of Features (BoF) を使用し、画像から局所特徴量を抽出しクラスタリングすることで、辞書を作成し、認識系を構成している [1]。このような食材学習を用いない方式では、認識精度に限界がある。一方で、認識精度を向上させるために食材学習を用いて食事画像認識している研究もある [2]。しかしながら、これらの研究ではクラスタ数が与える認識精度に与える影響はあまり考慮されていない。本研究では、今までに k-means 法の手法を取り入れ、最適なクラスタ数を推定し、認識精度の向上を試みてきた。しかしながら、クラスタ数を自分で決める必要がある。この問題を打開する為に自動でクラスタ数を推定する混合ディリクレ過程の手法を本研究では取り入れる [3]。また、レシピと食材学習に基づく料理画像認識において、混合ディリクレ過程を用いた認識率と k-means 法を用いた認識率の違いを検討し、混合ディリクレ過程の必要性を検討する。

### 2. k-means 法による食材学習に基づく食事画像クラスタ推定

最初に、M 個の認識する料理名を決め、食材リストを作成する。決めた料理名 1 つに対して P 個のレシピデータを Web サイト上から集め、統計を取り多かった食材の上位 N 個を食材リストとして選出する。合計 M×N 個の食材リストになる。しかし、同じ食材が選ばれた場合はその個数分、合計個数から引く。

次に、食材学習を行う。食材学習には R×(食材リストの数)枚の画像を使用する。X から Y 個のクラスタ数を決め k-means 法を用い、各食材のキーポイントをクラスタリングする。全体が料理の画像は、R×M 枚の画像を使用する。V から W 個のクラスタ数を決め k-means 法を用い、各料理画像のキーポイントをクラスタリングする。一方、料理全体画像からは、平均カラーヒストグラムを求める。平均カラーヒストグラムは、R 枚の画像に対して 180 個に分けた色相に画像全体から抽出した色情報を投票し、R で割ることで算出する。これを M 個作成する。下記に学習工程と認識工程の概略ステップを示す。

#### <学習の工程>

- STEP1 食材リストを作成
- STEP2 食材の SIFT 情報量を学習
- STEP3 料理全体画像の SIFT 情報量を学習
- STEP4 料理全体画像の平均カラーヒストグラム算出

#### <認識の工程>

- STEP1 未知画像を入力
- STEP2 未知画像から色情報を抽出
- STEP3 平均カラーヒストグラムのデータと比較し料理名の候補を M 個から α 個まで絞る
- STEP4 未知画像から SIFT 情報を抽出
- STEP5 料理画像全体の SIFT 学習データと比較し料理名の候補を α 個から β 個まで絞る
- STEP6 食材の SIFT 学習データと比較し料理名の候補を β 個から 1 個まで絞る
- STEP7 最も認識率の高いクラスタ数を求める
- STEP8 最適なクラスタ数に対する料理名を出力

食材は、加工により色が変化しやすいので本研究では、食材に対する平均カラーヒストグラムは使用しないものとする。

### 3. 混合ディリクレ過程によるクラスタ自動推定の導入

k-means 法の手法では、k の値によって結果が大きく変わってしまう。混合ディリクレ過程を用いたクラスタリングを行うことにより、統計的により尤もらしい k の値を自動で判断する。

混合ディリクレ過程のクラスタリングにおける入出力を以下に示す。

- ・入力データ：n 個のデータ ( $X_1 \dots X_n$ ) とする。
- ・出力データ：各データの所属クラスタを表す潜在変数の値は、 $s_1 \dots s_n$  とする。

混合ディリクレ過程によるクラスタリングのアルゴリズムを以下に示す。

- STEP1 初期設定
  - ・潜在変数  $s_1 \dots s_n$  を初期化する。このときのクラスタ数を  $c$  とし、クラスタ  $\omega_i$  に更新するパターン数を  $n_i$  とする ( $i = 1 \dots c$ )。
  - ・この初期結果と各クラスタに所属するパターンを用いて、クラスタのパラメータ  $\theta_1 \dots \theta_c$  を初期化する。
  - ・さらに、事後確率の最大値  $P_{max}$  を 0 として初期化する。

- STEP2 所属クラスタの更新
  - ・ここでは、 $k = 1 \dots n$  に対して実行する。パターン  $X_k$  の所

属クラスタ $s_k (= \omega_j)$ と仮定する)を更新するため、 $X_k$ を現在のクラスタから除外し、 $n_j$ を $n_j - 1$ とする。

- これにより、 $n_j = 0$ となり空きクラスタが発生した場合は、この空きクラスタを除去すべく、 $c$ を $c - 1$ とするとともに、 $j$ 以降のクラスタインデックスを全て1つ減らす。また、それに対応するパラメータのインデックスも更新する。
- 次に、既存クラスタに所属するパターン数 $n_i'$ 、新規クラスタ数を $\alpha$ 、 $G_0(\theta_{new})$ を事前確率分布とした場合、下式(3-1)を用いて $s_k$ の値を確率的に決定する。  
 $s_k = \omega_i$  ( $i = 1 \dots c$ ) に対して、下式(3-1)の式上段で計算する。また、 $s_k = \omega_{new}$  に対しては、下式(3-1)の式下段で計算する。また、 $c + 1$ 個の値の比の確率で $s_k$ の値を決定する。もし  $s_k = \omega_i$  と更新されたとき、 $n_i$ を $n_i + 1$ とする。それ以外で  $s_k = \omega_{new}$  と更新されたとき、 $n_{c+1}$ を1とするとともに  $c$ を $c + 1$ として総クラスタ数を更新する。

$$P(s_k | X_k \cdot s_{-k} \cdot X_{-k} \cdot \theta) \quad (3-1)$$

$$\propto \begin{cases} \frac{n_i'}{n-1+\alpha} \cdot p(X_k | \theta_i) & (s_k = \omega_i \text{ のとき}) \\ \frac{n}{n-1+\alpha} \cdot \int p(X_k | \theta_{new}) G_0(\theta_{new}) d\theta_{new} & (s_k = \omega_{new} \text{ のとき}) \end{cases}$$

#### STEP3 各クラスタのパラメータ更新

- 上記STEP2で得た  $s = \{s_1 \dots s_n\}$  の値に基づいて、事前確率及びクラスタに所属するパターンに対する尤度を  $p(\{k: X_k \in \omega_i\} | \theta_i)$  として、下式(3-2)より、各クラスタのパラメータ $\theta_i$ を確率的に決定し、更新する。

$$p(\theta_i | \{X_k: X_k \in \omega_i\}) \quad (3-2)$$

$$\begin{aligned} &= \frac{p(\{k: X_k \in \omega_i\} | \theta_i) \cdot G_0(\theta_i)}{p(\{k: X_k \in \omega_i\})} \\ &= \frac{G_0(\theta_i) \cdot \prod_{X_k \in \omega_i} p(X_k | \theta_i)}{\int G_0(\theta_i) \cdot \prod_{X_k \in \omega_i} p(X_k | \theta_i) d\theta_i} \end{aligned}$$

#### STEP4 事後確率最大化

- 現時点での  $s = \{s_1 \dots s_n\}$  及び  $\theta = \{\theta_1 \dots \theta_c\}$  の値を用い、 $P(S)$  をイェンセンの公式を用いて求め、ベイズの定理を用いて下式(3-3)により、事後確率 $\gamma$ を計算する。

$$p(s, \theta | X) \propto P(S) p(\theta) p(X | \theta, s) \quad (3-3)$$

$$\begin{aligned} &= P(S) \prod_{i=1}^c \left( G_0(\theta_i) \prod_{X_k \in \omega_i} p(X_k | \theta_i) \right) \\ &= \gamma \end{aligned}$$

- 事後確率最大化において以下の処理を行う。

(1)  $\gamma > P_{max}$  ならば以下の更新を行う。

- $P_{max}$  を  $\gamma$  とする
- $s = \{s_1 \dots s_n\}$  とする

(2) それ以外ならば、これまでの  $P_{max}$  と  $s$  を保存する

#### STEP5 終了判定

- 以上のことを繰り返すことにより、 $P_{max}$  の更新されない状態が継続された場合、実行を終了し、現時点でのクラスタ数やクラスタの割り振られた情報を出力する。そうでない場合は、STEP2の工程まで戻り再び実行を繰り返す[4]。

本研究では、収束条件を  $P_{max}$  の更新されない状態が継続された場合まで続けるには、時間が多く掛かってしまう。ここで、状態変化が経験的に決定されたしきい値以内であれば収束したとみなし、更新を終了したとする。

## 4. 実験方法

入力画像 100 枚に対して何枚料理名を認識できたかで認識率を計算する。この実験では、10 個の料理・20 のレシピ・食材上位 5 個・R=50 枚として行う。k-means 法を用いたクラスタ数に関しては、食材画像に対しては 50 から 500 個・料理画像に対しては 500 から 5000 個で行う。料理名の候補は、始め 10 個から 5 個まで絞り、次に 5 個から 2 個まで絞り最後は 1 個まで絞る。カラーのみ使用した場合・SIFT 情報量のみ使用した場合・カラーと SIFT の 2 つを組み合わせた場合と比較する。混合ディリクレ過程を用いた認識については、自動でクラスタ数が推定されるので出力されたクラスタ数に対して認識を行う。この 2 つの手法を比較し、混合ディリクレ過程の必要性を見つけ出す。実験画像の一例を以下に示す。



図 4.1 料理画像と食材画像

## 5. まとめ

現段階では、混合ディリクレ過程の自動クラスタ数の推定に成功している。ここから、認識率を算出し、k-means 法との違いを図る。

## 文 献

- [1] 田中雄翔, 滝口哲也, 有木康雄” ウェブ画像を用いたカテゴリ別 Visual Words による未知物体判別,” 画像の認識・理解シンポジウム (MIRU2012), 2012
- [2] 数藤恭子, 村崎和彦, 神谷叔季, 谷口行信,” 料理画像の素材セグメンテーションに基づく成分推定,” 電子情報通信学会誌, No. 468 (IMQ), No. 469 (IE), No. 470 (MVE), pp. 73-76, 2014.
- [3] 上田修功, 山田武士,” ノンパラメトリックベイズモデル,” 電子情報通信学会技術研究報, pp. 81-86, 2007.
- [4] 上田修功, 石田健一郎,” 続・わかりやすいパターン認識—教師なし学習入門,” pp. 260-261, 2014, オーム社.