

## 食事ログシステムの解析の個人への適応と検討 Personalization of FoodLog System

丸山 悠斗<sup>†</sup> 山崎 俊彦<sup>‡</sup> 相澤 清晴<sup>†</sup>  
Yuto Maruyama Toshihiko Yamasaki Kiyoharu Aizawa

### 1. はじめに

インターネットの普及によって、我々の生活も変化しており、日常の出来事をいつでもどこでも簡単に管理することができるようになってきた。その結果、日々の様々な記録を活用する「ライフログ」という研究は、より重要度が増し、特に生活に最も密着している「食事」が重要な要素であることは疑いない。近年、食事の記録に関する研究・サービスも盛んに行われており、食事の前後の画像を撮影し、料理部分をセグメントして、料理の分類、量の推定を行うもの[1]や、機械学習の Support Vector Machine (SVM) を用いて、50種類内でのメニュー推定を行うもの[2]、さらには、ファーストフードに特化し、食事の様子を動画撮影して、得られる画像の SIFT 特徴量のマッチングを行うことで、食事を分類し、摂取カロリーを推定するもの[3]などがある。

本研究は、我々が研究・開発している食事記録サービス FoodLog [4][5]での、食事画像判定とバランス推定での精度向上を目指すものである。インクリメンタルに精度を高められる Naïve Bayes と、既存の SVM との精度を比較し、前者は、食事判定が 43%、非食事判定は 95.8%、後者は交差検定で 86%であった。また、ユーザの食傾向の違いから、実際に Naïve Bayes によってインクリメンタルな学習を疑似的に行ってみたが、優位な差は得られなかった。

### 2. FoodLog

#### 2.1 システム概要

FoodLog は、最も簡単に食事を記録できるサービスの一つとして、2010年4月にプレスリリースした[6](図1)。今まで、食事の記録としては手入力一般的なものであり、持続することが困難であったが、FoodLog ではその負担を改善すべく、食事の写真を撮るだけで記録するシステムを採用している。ユーザは写真を撮って画像共有サイトの Flickr[7]か直接 FoodLog に送るだけで、自動的に食事画像だけを抽出し、更に、SV を単位とする栄養バランス(主食・副菜・主菜)の推定を行う。また、ユーザは Web 上で自分の食事記録を修正できる機能があり、FoodLog が行った食事画像抽出や栄養バランス推定において、ユーザからのフィードバックを得ることができる。

#### 2.2 既存手法

現在のシステムは、画像から特徴量を抽出し、そのデータを機械学習にかけることで実現させている。食事画像抽出においては SVM を、栄養バランス推定においては決定木を弱分類器とした Adaboost を用いている。また、特徴量は、色・円情報の大域特徴量と、SIFT 特徴量を k-means に

<sup>†</sup> 東京大学大学院 情報学環・学際情報学府  
Interfaculty Initiative in Information Studies, Interdisciplinary  
Information Studies, The University of Tokyo

<sup>‡</sup> 東京大学大学院 情報理工学系研究科  
Department of Information and Communication Engineering,  
The University of Tokyo



図1: FoodLog (食事時間分布表示)

よって 500 個にクラスタリングした局所特徴量の大きく 2 種類を用いている。食事画像抽出においては、現在 90%以上の精度であるが、バランス推定においては 40%前後である。

#### 2.3 提案手法

既存手法では、栄養バランス推定において向上の余地が多く、更にユーザからのフィードバックを学習に生かしていない。図2は、3人のユーザのそれぞれ連続した12枚の料理画像であるが、これを見ても分かる通り、ユーザの食事画像は各々傾向があることが伺える。これは、ユーザのカメラの特性であったり、普段使用する食器であったり、ある食事の時の写真が多く撮られたり、もちろん食生活などに依存していると推測できる。よって、ユーザごとに特化した処理を行うことで、推定精度を向上させることが期待できる。そこで、ユーザからの修正情報を学習器に反映させ、徐々に精度を高める手法を提案する。今回は、最もシンプルな Naïve Bayes について検討してみる。そもそも、ベイズ推定とは、最も確率の高い事後確率を用いるものである。ある画像から、特徴量  $F$  (次元数  $j$ ) が抽出された時に、その画像から推測されるべき値  $\theta$  (クラス数  $i$ ) の確率は、

$$P(\theta_i | F_j) = \frac{P(F_j | \theta_i) P(\theta_i)}{P(F_j)}$$

と与えられる。この時の右辺の  $P(F_j | \theta_i)$  は尤度、 $P(\theta_i)$  は事前確率と呼ばれる。さらに、右辺分母  $P(F_j)$  は、 $\theta$  のクラス  $i$  によらず一定なので、解くべき式は下記のように書き換えられる。

$$\arg \max P(\theta_i | F_j) \propto P(F_j | \theta_i) P(\theta_i)$$

よって、ユーザのフィードバックによって精度を向上させていくには、尤度と事前確率を更新させていくことになる。事前確率は統計的に求める(更新する)ことができ、尤度についても、 $N$  を  $j$  の総数、 $f_j$  を  $F_j$  のベクトル平均として

$$P(F_j | \theta_i) \approx \prod_j P(f_j | \theta_i)$$

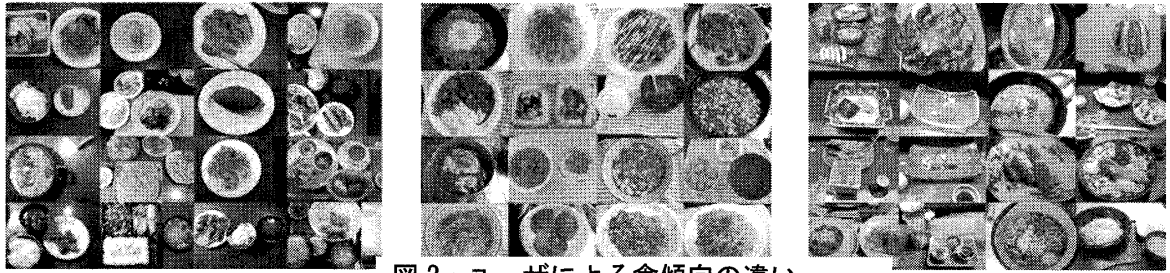


図2: ユーザによる食傾向の違い

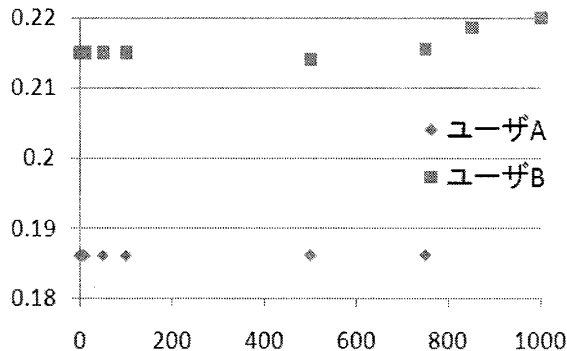


図3: フィードバック数と食事判定率の関係

と表せる。よって、新たな画像が増え、 $F_i$ が増えれば増えるほど、ベクトル平均の  $f_i$  も更新され、こちらもより精度を向上させることができる。以上より、よりパーソナライズされた推定を行うためには、インクリメンタルな精度向上が見込めるベイズ推定を用いることが有効であると考えられる。

### 3. 実験結果

#### 3.1 SVM と Naive Bayes との比較

今回用いた特徴量は、画像の大域特徴量である色情報 (RGB, HSV の平均, 分散) と、円情報 (数, 面積) の計 46 次元だけである。今回は単純に SVM と Naive Bayes の比較を目的としたため、局所特徴量は用いなかった。また、推定に関しても、食事画像判定について行った。テスト用画像には、食事であるものとそうでないものをそれぞれ 1000 枚ずつ、計 2000 枚を用意し、今回の研究は全てこの 2000 枚を基準に行った。SVM に関しては交差検定を行い、Naive Bayes については全ての画像を用いて判定を行った。結果は、SVM に関しては、線形カーネルを用いた時が 51.9%、多項式カーネルを用いた時が 86.7%であった。一方、ベイズ推定に関しては、食事画像を食事画像であると判別できたものが 43.0%、食事画像でないものを食事画像でないとして判別できたものが 95.8%であった。

#### 3.2 ユーザの食傾向の影響

ユーザの食傾向の影響を調べるために、上記 2000 枚の写真以外に、特定の 2 ユーザの食事画像をそれぞれ 500 枚、1000 枚用意した。今回はユーザの写真を、1, 5, 10, 50...1000 と増やして行った。Naive Bayes でのモデルに関

して、その増分も合わせて毎回生成し直し、疑似的なインクリメンタルな学習結果とした。図 3 は、その時の食事判定の正答率についてである。ユーザ B に関しては、元から 2.3%の精度向上が見られたが、優位な差とは言えない。また、モデルに使わないユーザの食事画像も全て入れて、非食事判定については、フィードバック数に関係なく、95.8%であった。

### 4. まとめ

本研究では、ユーザの食事画像抽出と、バランス推定において、ユーザの食傾向の違いに着目した。そこでまず、インクリメンタルな精度向上が見込まれるベイズ推定と既存手法の SVM についての比較を行った。SVM については、カーネルを調節することで、大域特徴量だけで 86.7%の精度が出ることを確認した。Naive Bayes については、食事画像抽出においてはランダムよりも低い値となってしまうが、非食事抽出に関しては有効な手法だと言える。

そして、ユーザがデータの修正データを学習器にフィードバックさせた時の精度は、少なくとも今回用いた特徴量では優位な差は生まれなかった。また、3.1 と 3.2 での Naive Bayes の食事判定率の違いについては、後者はフィードバック数に関わらず全ての画像でテストしているのに対し、前者では最初からテストされる全ての画像はモデル生成に利用されている点が考えられる。しかしながら、フィードバック数最大の時と、3.1 での判定率に差がある点については、検討すべきことである。その他、インクリメンタルな学習は、SVM においても Incremental SVM[8][9]があり、特徴量と合わせて、用いるアルゴリズムについても今後の課題としたい。

#### 参考文献

- [1] F. Zhu, A. Mariappan, C. J. Boushey, D. Kerr, K. D. Lutes, D. S. Ebert, and E. J. Delp "Technology-assisted dietary assessment", Proc.SPIE, Vol.6814, No.681411 (2008).
- [2] T. Joutou and K. Yanai "A food image recognition system with multiple kernel learning", ICIP, pp.285-288 (2009).
- [3] W. Wu and J. Yang "Fast food recognition from videos of eating for calorie estimation", ICME, pp. 1210-1213 (2009).
- [4] K. Kitamura, T. Yamasaki, K. Aiazawa "Food Log by Analyzing Food Images", Proc. ACM Multimedia, pp. 999-1000, (2008).
- [5] 北村 圭吾, 山崎 俊彦, 相澤 清晴 "食事ログの取得と処理—画像処理による食事記録—", Proc. MIRU, pp.676-681, (2008)
- [6] "FoodLog", <http://www.foodlog.jp/>
- [7] "Flickr", <http://www.flickr.com/>
- [8] G. Fung and O. L. Mangasarian "Incremental Support Vector Machine Classification", ACM SIGKDD, 2002.
- [9] G. Cauwenberghs and T. Poggio "Incremental and Decremental Support Vector Machine Learning", NIPS, Vol.13, 2000.