

# Tweet データに基づく料理画像の魅力度定量化モデル

1X14C111-1 藤波 英輝  
指導教員 後藤 正幸

## 1 研究背景・目的

近年、インターネット上でレストラン情報を掲載するグルメサイトが広く利用されている。一般にグルメサイトにおいて、ユーザはレストランの検索に加えて、レストランに対する評価や推薦記事および料理画像の投稿をすることができる。さらに、他ユーザが投稿した推薦記事に対して「いいね」などのリアクションを付与することが可能である。従って、グルメサイト運営企業の持つデータベース上には、各レストランに対する評価記事・料理画像・リアクションなどの多様なデータが蓄積されている。これに伴い、グルメサイト運営企業に対し、蓄積された大規模データを活用し、質の高いサービスを提供することが期待されている。

グルメサイト上でユーザが行きたいレストランを探す際の重要な判断材料の1つとして、各レストランのページに付与されている料理画像の魅力度が挙げられる。一般に、料理画像の魅力度は閲覧するユーザの主観によって決まるものであるが、もしこれが何らかの方法で定量化できれば、様々な施策やサービスに活用することができる。例えば、料理画像の魅力度を定量化することで、サイト運営企業はレストランを探すユーザにとって魅力的な料理画像を、ユーザの目に留まりやすい位置に配置するといった施策に結び付けることが可能である。

そこで本研究では、料理画像の魅力度を定量化するための足がかりとして、リアクションを多く獲得している画像を魅力的な画像と定義し、各料理画像の魅力度を考慮した画像データ処理の方法、並びに魅力的な料理画像を抽出する手法の提案を目的とする。一般に画像解析を行う際には、画像をベクトル表現したデータが用いられることが多い。従来の画像のベクトル表現方法では、色と質感の局所特徴ベクトルを用いて画像をベクトル表現する [1] [2]。しかし、このベクトル表現は料理画像の識別 [1] などには用いられるが、料理画像の魅力度を表現するには不十分である。そこで、Social Networking Service(以下、SNS)上でユーザが画像を投稿する際に、画像をより魅力的に見えるように加工する画像加工技術に着目する。そして、その技術から着想を得て、画像に付与されたリアクション数の多寡に基づき学習した特徴量を従来の特徴量に追加した画像のベクトル表現を提案し、料理画像の魅力度を定量的に表現する。加えて、代表的な SNS として知られている Twitter に蓄積された料理画像とリアクション数 (以下、いいね数) が付与されたデータ (以下、Tweet データ) [3] を用いた魅力的な画像の分類実験を行い、提案手法の有用性を示す。

## 2 従来研究

画像を学習に用いる際、各画像はベクトルで表現される。画像をベクトル表現する手法の一つに Patch based Bag of Colors and Textures [1] (以下、PBCT) がある。一般に色や質感などの局所特徴量を用いた画像のベクトル表現は非常に高次元であるため、スパース性が高い。そこで PBCT では、全局所特徴ベクトルを  $k$ -means 法でクラスタリングし、得られる各クラスターの中心ベクトル (以下、Visual Words) の出現頻度のベクトルで画像を表現する。以下に PBCT による画像表現の流れを示す。

### STEP1 各画像から局所特徴ベクトルを抽出

#### i) 色の局所特徴ベクトル抽出

各局所特徴ベクトルは  $3 \times 3$  ピクセル (1 パッチ単位) で抽出する。ピクセル単位ではなく、1 パッチ単位で特徴ベクトルを抽出することで、局所的な変化を捉えることが可能となる。全画像から抽出される全局所特徴ベクトル数を  $I$  とし、各局所特徴ベクトルを  $\mathbf{x}_i: 1 \leq i \leq I$  とする。このときカラー画像において、あるピクセルを表現する値は 0 以上 255 以下の 3 つの整数値  $(R, G, B)$  であり、抽出される 1 パッチ単位の画像の

局所特徴ベクトルは、 $\mathbf{x}_i = (R_{i1}, G_{i1}, B_{i1}, \dots, R_{i9}, G_{i9}, B_{i9})$  の 27 次元で与えられる。

#### ii) 質感の局所特徴ベクトル抽出

質感の局所特徴ベクトル抽出には、Local Binary Pattern [2] (以下、LBP) を用いる。LBP では、グレースケール画像の各ピクセルの輝度値を、周辺ピクセルとの大小関係から、周辺ピクセルとの相対的な明るさを表す LBP value に変換する処理を全ピクセルに施すことで、画像を変換する。そして、変換した画像から、1 パッチごとに特徴ベクトルとして抽出する。局所特徴ベクトル  $\mathbf{x}_i$  のあるピクセルの LBP value を  $LBP_{ik} (k \in \{1, 2, \dots, 8, 9\})$  とする。このとき質感の局所特徴ベクトルは  $\mathbf{x}_i = (LBP_{i1}, LBP_{i2}, \dots, LBP_{i9})$  の 9 次元で与えられる。

### STEP2 局所特徴ベクトルの $k$ -means 法による圧縮

STEP1 で抽出した局所特徴ベクトルを  $k$ -means 法を用いてクラスタリングし、各クラスターの中心ベクトルを Visual Words とする。

### STEP3 各画像を Visual Words の頻度ベクトルに変換

各局所特徴ベクトルをその所属クラスターに対応する Visual Words に変換し、Visual Words の頻度ベクトルにより画像を表現することで、学習画像の局所特徴ベクトルを反映する。

## 3 新たな特徴ベクトルの提案

従来研究における画像表現は主に物体認識を目的としており、料理画像の魅力度を定量的に表現するには不十分であると考えられる。そこで、従来の画像表現に用いられる特徴ベクトルに加え、料理画像の魅力度を表現することに適した新たな特徴ベクトルを提案する。

### 3.1 着想

近年、SNS 上でリアクション数などの指標に基づいた高い評価を得ることを目的として、多数の画像加工技術が開発されている [4]。故に、これらの加工技術により加工が施された画像は、SNS 上においてリアクション数を多く獲得することができると考えられる。そこで、画像加工技術によって、加工を行う前後の画像の変化を特徴ベクトルとして用いることで、魅力的な画像の特徴を捉えた画像のベクトル表現が可能となると考える。

本研究では、近年開発されている多くの画像加工技術の中でも、画像の中心を明るく、周辺に向かって暗くする処理に着目する。そして、中心から等間隔に明るさ情報を切り出す際に、各画像データのリアクション数の大小を考慮してベクトル化した、画像の特徴ベクトルを提案する。

### 3.2 特徴ベクトル抽出アルゴリズム

以下に提案する特徴ベクトル抽出のイメージ図を示す。下記の図 1 のように、画像の中心から周辺に向かって、各画像データのリアクション数の大小を考慮して、ピクセルの輝度値を階層的に抽出し、提案特徴ベクトルを求める。

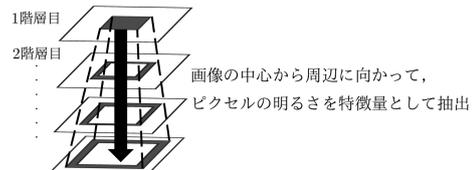


図 1. 提案する画像の特徴ベクトル抽出法のイメージ

以下に提案する特徴ベクトル抽出のアルゴリズムを示す。下記のアルゴリズムに基づいて、まず、Tweet データにおいて、いいね数が多いか否かの 2 つのクラスが付与された学習画像を用意する。そして、各画像において、輝度値の閾値  $\alpha$  を超えるピクセルの割合を階層ごとに求め、特徴ベクトルとする。ここで、輝度値は RGB 値の加重平均として、式 (1) で定義する。そして  $\alpha$  を変化させ、いいね数の多い画像群

と少ない画像群で特徴ベクトルの距離が最大のとき、料理画像の魅力度を最も説明できていると考えられる。そこで、閾値を5~255の範囲で5ずつ増やし、それぞれの閾値において、特徴ベクトルで表現されたいね数の多い画像と少ない画像の全組み合わせにおいて、それぞれユークリッド距離を求める。その全組み合わせにおけるユークリッド距離の和を距離和と定義し、距離和が最大となる閾値を探索的に求める。以下に輝度値の定義式(1)と提案特徴ベクトル抽出のアルゴリズムの詳細を示す。

$$\text{輝度値} = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (1)$$

**STEP0**  $\alpha = 5$  とする。

**STEP1** 各画像の中心の  $n \times n$  ピクセルのうち、ピクセルの輝度値が閾値  $\alpha$  を超えるピクセルの割合を求める。これを特徴ベクトルの1つ目の要素とし、 $i = 1$  とする。

**STEP2** 画像の中心の  $(2i + 1)n \times (2i + 1)n$  ピクセルのうち、中心の  $(2i - 1)n \times (2i - 1)n$  ピクセル以外の部分において、輝度値が閾値  $\alpha$  を超えるピクセルの割合を求め、1つの特徴量とし、 $i = i + 1$  とする。

**STEP3** “画像の幅(横幅のピクセル数)”,  $g_i$  が  $(2i + 1)n$  以上であればSTEP2へ戻る。さもなくばSTEP4へ進む。

**STEP4** STEP1~3によって得られた、いいね数の多い画像の特徴ベクトルと、少ない画像の特徴ベクトルの全組み合わせにおいて、ユークリッド距離を求め、その和  $dis_\alpha$  を求める。

**STEP5**  $\alpha = \alpha + 5$  とし、 $\alpha \leq 255$  ならば、STEP1へ戻る。それ以外はSTEP6へ進む。

**STEP6** 最適な閾値  $T$  は距離和  $dis_\alpha$  が最大になる  $\alpha$  とし、式(2)から求める。この閾値  $T$  において、STEP1~3から得られる特徴ベクトルがいいね数の多い画像の魅力度を最もよく定量的に表現できるものとし、提案特徴ベクトルとする。

$$T = \arg \max_{\alpha \in \{5, 10, \dots, 250, 255\}} dis_\alpha \quad (2)$$

## 4 実験

いいね数が相対的に多い画像を魅力的な画像(正例)とし、いいね数が相対的に少ない画像を魅力的ではない画像(負例)とラベル付けする。そして、提案した特徴ベクトルによって優れた二値分類器を学習することができるかを検証する。

この際、従来手法として、“色の特徴ベクトル”、“質感の特徴ベクトル”、“色+質感の特徴ベクトル”を用いて学習を行う。提案手法として、“色+提案特徴ベクトル”、“色+質感+提案特徴ベクトル”を用いて分類器の学習を行う。そして全ての画像表現における分類結果を比較することで、提案手法の有効性を示す。

### 4.1 実験条件

提案手法の有効性を確認するために、グルメサイトAが運営するグルメマガジンBのTwitterアカウントから取得したTweetデータに含まれる画像データといいね数を用いて実験を行う。

Tweetデータの観察期間は2014年8月21日から2017年11月12日、総ツイート数14,428件である。全ての投稿には画像があるが、複数の料理画像が付与されている場合がある。そこで分析には最も目に入りやすい1枚目の画像のみを対象とする。

実験には、2017年11月12日以前の直近のツイートに付与されている料理画像のうちいいね数が上位1/4以上の画像1,000枚を正例、いいね数が下位1/4以下の画像1,000枚を負例とした計2,000枚を用いた。また、画像表現手法の過程で用いるk-means法におけるクラス数  $k=256$  と設定する。提案特徴ベクトルにおける閾値は、 $T = 135$  となった。

### 4.2 実験方法

4.1章で述べた料理画像に対しランダムフォレストによる分類を行い料理の魅力度を表現することができているか検証する。ランダムフォレストにおけるパラメータはGrid Searchにより決定した。2,000枚の画像に対し、10分割交差検証を用いた二値分類を行い、再現率と適合率、F値を式(3)~

(5)で算出する。そして、それぞれの平均値をモデルの評価指標とする。

$$\text{再現率} = \frac{\text{正例と正しく分類された画像枚数}}{\text{テストデータ中の正例画像枚数}} \quad (3)$$

$$\text{適合率} = \frac{\text{正例と正しく分類された画像枚数}}{\text{正例として抽出された画像枚数}} \quad (4)$$

$$F \text{ 値} = \frac{2 \times \text{再現率} \times \text{適合率}}{\text{再現率} + \text{適合率}} \quad (5)$$

## 4.3 実験結果と考察

実験結果を以下の表1に示す。

表1. 従来と提案の画像表現ごとの画像分類の結果

区分	学習に使用する特徴ベクトル	再現率	適合率	F 値
従来	色の特徴ベクトル	0.6830	0.7190	0.6990
従来	質感の特徴ベクトル	0.5250	0.6836	0.5926
従来	色+質感の特徴ベクトル	0.6710	0.7216	0.6934
提案	色+提案特徴ベクトル	<b>0.6870</b>	0.7364	0.7088
提案	色+質感+提案特徴ベクトル	0.6720	<b>0.7568</b>	<b>0.7094</b>

表1より、再現率は“色の特徴ベクトル+提案特徴ベクトル”が最も優れていることがわかる。また適合率とF値は“色+質感+提案特徴ベクトル”が最良となっている。この結果より、従来で用いられていた特徴ベクトルに提案した特徴ベクトルを加えることで、より精度の高い二値分類器を学習できていることがわかる。

この結果は、画像表現空間上において、正例と負例のクラスの重なりが従来法よりも小さいことに起因していると考えられる。これは、表2に示すように、分類器の性能を示すROC曲線の曲線下面積(AUC: Area under the curve)の値を比較したとき、“従来の特徴ベクトル+提案特徴ベクトル”による画像表現において、よりAUCが高くなっていることから裏付けられる。

表2. 従来と提案の局所特徴ベクトルにおけるAUCの値

区分	学習に使用する特徴ベクトル	AUC
従来	色の特徴ベクトル	0.7808
従来	質感の特徴ベクトル	0.6934
従来	色+質感の特徴ベクトル	0.7757
提案	色+提案特徴ベクトル	<b>0.7897</b>
提案	色+質感+提案特徴ベクトル	<b>0.8049</b>

また提案特徴ベクトルを加えることで、従来の特徴ベクトルによる画像表現に比べ、適合率は向上したが、再現率の上昇率は小さかった。このことから、提案特徴ベクトルは、負例の誤判別を防ぐ効果があると考えられる。

## 5 まとめと今後の課題

色と質感の特徴ベクトルに加えて、リアクション数の獲得に起因する階層ごとの明るさに着目した特徴ベクトルを用いることで、リアクション数が付与されやすい魅力的な料理画像の魅力度をより高い精度で表現する手法を提案し、Tweetデータを用いた実験により、その有効性を示した。

今後の課題として、リアクション数の高い画像を正しく分類する精度(再現率)を高める画像の特徴ベクトルを提案すること、リアクション数の特に高い画像を正しく分類できているか検証することなどが挙げられる。

## 参考文献

- [1] Shota Sasao, and Xian-Hua Han, and Yen-Wei Chen, “Food Recognition by Combined Bags of color Features and Texture Features,” *Proc. of the International Conference on International Congress on Image and Signal Processing, BioMedical Engineering and Informatics '9*, pp.815-819, 2016.
- [2] T. Maenpaa, et al., “Robust texture classification by subset of local binary pattern,” *Proc. of the International Conference on Pattern Recognition '15*, pp.947-950, 2000.
- [3] “食通の厳選グルメマガジン「メシコレ」,” <https://twitter.com/mecicolle/>
- [4] “【2017年最新】インスタグラマーが使うおしゃれ加工アプリ・サービスまとめ,” <http://social-buzz.site/instagram-edit-photo-apps-2017/>, 2017年1月6日