

# 画像付き料理レシピの地域属性推定の検討

伊藤耀一朗<sup>†</sup> 道満 恵介<sup>††,†</sup> 川西 康友<sup>†</sup> 平山 高嗣<sup>†††</sup> 井手 一郎<sup>†</sup>

出口 大輔<sup>††††,†</sup> 村瀬 洋<sup>†</sup>

<sup>†</sup> 名古屋大学 大学院情報学研究科 〒464-8601 愛知県名古屋市千種区不老町

<sup>††</sup> 中京大学 工学部 〒470-0393 愛知県豊田市貝津町床立 101

<sup>†††</sup> 名古屋大学 未来社会創造機構 〒464-8601 愛知県名古屋市千種区不老町

<sup>††††</sup> 名古屋大学 情報戦略室 〒464-8601 愛知県名古屋市千種区不老町

E-mail: <sup>†</sup>itoy@murase.is.i.nagoya-u.ac.jp, kawanishi@i.nagoya-u.ac.jp, ide@i.nagoya-u.ac.jp,  
murase@i.nagoya-u.ac.jp, <sup>††</sup>kdoman@sist.chukyo-u.ac.jp, <sup>†††</sup>takatsugu.hirayama@nagoya-u.jp,  
<sup>††††</sup>ddeguchi@nagoya-u.jp

あらまし 近年、投稿型料理レシピポータルサイトが普及している。ウェブ上には膨大な数の料理レシピが存在し、掲載されている料理の種類は多岐にわたる。しかし、それに伴い、サイト利用者が検索の際に目的とする料理レシピ件数を絞りこむことが難しくなりつつある。そこで本研究では、地域属性（スペイン料理、中華料理など）ごとに料理の味や見た目が大きく異なることに着目し、料理レシピがもつ多様な情報から、機械学習を用いて作成される料理の地域属性を推定する手法を提案する。学習には、料理の完成画像から得られた画像特徴、素材一覧から得られた素材特徴、調理手順文の動詞から得られた手順特徴の3つを組み合わせる。結果として、これら3つの全特徴を用いた手法で8割近い推定精度が得られることを確認した。

キーワード 料理, 食メディア, 推定, 地域属性

## 1 はじめに

料理は我々の生活を豊かにする重要な要素の1つである。料理初心者や不慣れた料理を作る場合、料理レシピを参考にすることがある。近年では楽天レシピ<sup>1</sup> や COOKPAD<sup>2</sup> のような投稿型料理レシピポータルサイトが普及している。投稿型料理レシピは年々増加傾向にあり、投稿されている料理の属性も地域性や季節性をはじめ、多岐にわたっている。しかし、それに伴い、ユーザが検索を行う際に目的とする料理レシピを絞り込むことが困難になりつつある。料理レシピサイトでは、主に料理名や素材名など直接記載された情報に対して検索を行う。最近では、料理の付加的な情報を表す「タグ」を用いて料理レシピに直接記載されていない情報に対して検索を行うことができる料理サイトも存在する。しかし、ユーザ投稿型のサイトでは、タグの付与は主に料理レシピの投稿者に委ねられていることが多いため、適切なタグが付与されていない料理レシピも多く存在する。

関連研究として、様々なクエリにより料理レシピを検索するための支援をする研究が行われている。浅沼ら [1] は、食材に対する嗜好性と料理レシピの食材分量の情報をもとに検索を支援する研究を行っている。坪村ら [4] はユーザの運動量と料理レシピの栄養情報を基に料理レシピを推薦する研究を行っている。また、執行ら [3] は食感の可視化を通して料理レシピの検

索性を向上させる研究を行っている。これらの研究は、学習に素材情報やレビュー文などを各々単体で用いている。一方、料理レシピには画像や調理手順文などの情報も含まれている。それらの情報を用いれば更に高精度な推薦、検索が行えると考えられる。そこで、先行研究 [5] として、料理レシピ上の複数の種類の情報を基にその料理レシピの完成品の味を推定する研究が行われている。この研究では、本研究と同様に画像特徴と素材特徴と手順特徴を用いて、「甘味」、「酸味」、「辛味」、「塩味」、「苦味」の5つの味を推定している。そして、これらの特徴からある程度の味の推定が可能であることと、すべての特徴を結合させることで、各特徴を単体で使用するよりも推定精度が向上することを確認している。

本発表では、先行研究の手法をもとに、料理レシピがもつ画像情報、素材情報、手順情報から、その完成品料理の属性を推定する手法を提案する。具体的には、中華料理の油通しなどの独特な調理方法の存在、スペイン風オムレツと卵焼きの見た目の違いなどに着目し、味や見た目に影響を与えると考えられる料理の地域属性を推定する。この推定結果に基づいてタグの自動付与を行えば、ユーザがより正確な精度で地域属性をクエリとした検索が可能になると考えられる。

## 2 提案手法

提案手法は、学習段階と識別段階の2段階の処理により、与えられた画像付き料理レシピの完成品の地域属性を推定する。図1に提案手法の処理手順を示す。以降、各段階の処理につい

1: 楽天株式会社, “楽天レシピ”, <http://recipe.rakuten.co.jp/>

2: クックパッド株式会社, “COOKPAD”, <http://cookpad.com/>

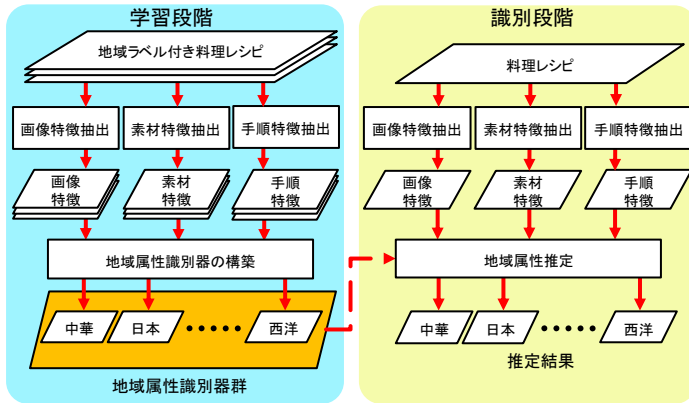


図 1 提案手法の処理手順

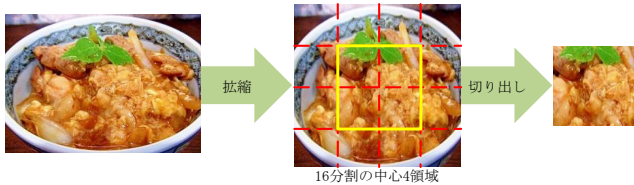


図 2 事前処理の手順

て詳述する。

## 2.1 学習段階：識別器の構築

学習段階では識別対象の 8 つの地域属性が料理レシピの完成品に含まれているか否かのラベルを教師データとし、SVM (Support Vector Machine) [6] によりそれぞれの地域属性に対する 2 クラス識別器を構築する。先行研究同様、料理の完成画像から得られる画像特徴、素材一覧から得られる素材特徴、調理手順文の動詞から得られる手順特徴を識別器の学習に用いる。以降、2.1.1 ~ 2.1.3 で各特徴の具体的な抽出方法についてそれぞれ述べる。

### 2.1.1 画像特徴の抽出

料理レシピ中の料理画像は、料理のみが大きく写っている画像だけでなく、器や背景も含まれている画像、画像全体の形が縦長・横長である画像など様々なものがある。そこで、本研究では、この構図の違いを吸収しつつ、より正確に特徴抽出を行うために、事前処理として料理画像の拡大と中心矩形の切り出しを行う。具体的には、図 2 のようにまず画像の縦横幅を 256 画素四方の正方形に拡大する。投稿型レシピサイトではユーザーによって料理写真を撮る角度や料理の置き方が様々なため、拡大による縦横比の変化は問題ないと考えられる。その後、この画像を 16 領域に等分割し、中心の 4 領域を特徴抽出領域として切り出す。この事前処理により、器や背景など料理そのものとの関係がない情報がほとんど除去され、画像の構図によらない特徴抽出が行えるようになると考えられる。

切り出した矩形から色特徴として HS 色ヒストグラム (90 次元) と HS 色コリログラム (116 次元)、テクスチャ特徴として SIFT 特徴 [7] の Bag-of-Features 表現 (90 次元)、形状特徴として HOG 特徴 [8] (90 次元) を抽出し画像特徴とする。

表 1 素材特徴の例

素材名	しお	たまご	...	だいこん	とりにく
素材特徴	1	1	...	0	1

表 2 手順特徴の例

手順名	にる	まぜる	...	やく	ひきのばす
手順特徴	0	1	...	1	0

### 2.1.2 素材特徴の抽出

素材特徴はその料理に各素材が含まれているか否かを 2 値表現したものを用いる。抽出手順としてまず、全料理レシピから素材一覧を抽出し、記号や数字を除去し、すべて平仮名表記に統一した素材名辞書を作成する。表 1 上部は素材辞書中の素材名の一覧を表す。その後、各料理レシピの素材一覧に基づいて、表 1 に示すように辞書中の各素材がその料理レシピで使用されているか否かを 2 値表現し、その料理レシピの素材特徴とする。

### 2.1.3 手順特徴の抽出

手順特徴も素材特徴と同様、料理レシピの料理手順文に各調理手順動詞が含まれているか否かを 2 値表現したものを用いる。抽出手順として、まず、全料理レシピから調理手順を抽出し、森らによって提案された料理レシピ専用の形態素解析ツール [9] を適用することで、調理手順文を品詞分解し、調理手順とタグ付けされた動詞 (以下、手順動詞) のみを抽出する。次に、抽出した手順動詞をすべて平仮名表記に統一し、手順辞書を作成する。表 2 上部は手順辞書中の手順動詞の一覧を表す。その後、各料理レシピ中の調理手順より、辞書作成時と同様の方法を用いて手順動詞を抽出し、表 2 に示すように辞書内の各調理手順がその料理レシピで行われるか否かを 2 値表現し、その料理レシピの手順特徴とする。

## 2.2 識別段階

入力された料理レシピから学習段階と同様に画像特徴、素材特徴、手順特徴をそれぞれ抽出する。学習段階で構築した識別器を用いて、入力した料理レシピから作成される料理が各地域属性をもつか否かをそれぞれ推定する。その結果、各地域属性をもつか否かが 2 値表現された推定結果が得られる。

## 3 評価実験

提案手法の有効性を確認するための評価実験を行った。

### 3.1 データセット

楽天レシピ<sup>1</sup> にタグが存在する中華、エスニック、フレンチ、イタリアン、日本、韓国、沖縄、西洋の 8 種類の地域属性を識別対象とした。識別対象の中には、沖縄と日本、西洋とイタリアンやフレンチなど、地理的な包含関係があるものが存在するが、本実験では、上記のラベルの中から複数のラベルが付与さ

表 3 地域属性別料理レシピの件数

地域属性	件数
中華	351
エスニック	1,175
フレンチ	871
イタリアン	379
日本	987
韓国	251
沖縄	620
西洋	746
合計	5,380

表 4 特徴量毎の SVM のパラメータ

使用特徴量	Cost	Kernel	$\gamma$
画像のみ	1,000	RBF	0.0010
素材のみ	1	RBF	0.0001
手順のみ	1	RBF	0.0001
画像+素材	1	Linear	—
画像+手順	1,000	RBF	0.0001
素材+手順	1	Linear	—
画像+素材+手順	1	Linear	—

れている料理レシピを除去し、ラベルが重複していないもののみを使用した。データセットとしては、楽天レシピのウェブサイトに掲載されている料理レシピの中から、上記の 8 種類いずれかのタグが付与されているものを抽出した。抽出したレシピのうち、楽天データセット [10] に画像、素材、手順に関するすべての情報をもつレシピを選択した。最後に、複数の地域属性をもつ料理レシピを除外したところ、表 3 に示すように、5,380 件の料理レシピから構成されるデータセットになった。

### 3.2 実験方法

構築したデータセットを用いて、提案手法の有効性を確認する実験を行った。評価には 8 分割交差検定法を用いた。提案手法の有効性を比較評価するために、画像特徴、素材特徴、手順特徴単体もしくはその組み合わせをそれぞれ特徴量とした 7 通りの手法を比較した。各使用特徴量にスケーリングを施した。また、SVM のパラメータについて使用特徴毎に事前に Grid Search により、表 4 に示すような最適な組み合わせを求め、以下の実験で用いた。各特徴量の次元数は画像特徴 386 次元、素材特徴 7,941 次元、手順特徴 2,836 次元であった。本実験の推定結果は検索の支援に用いられるため、評価指標には、識別結果中における正しい識別結果の割合を示す適合率を用いた。

### 3.3 実験結果

各地域属性に対する評価結果を表 5 に示す。全地域平均では素材特徴のみを使用する手法が最も高精度、次に全特徴を使用する手法が高精度であった。地域別では、中華、イタリアン、韓国、沖縄料理において全特徴を使用した手法、エスニック、



(a) キッシュの画像例<sup>3</sup>

(b) チーズフォンデュの画像例<sup>4</sup>

図 3 フレンチの地域属性を持つ画像例

日本料理においては素材特徴のみを使用した手法、フレンチ、西洋料理では画像特徴のみを使用した手法がそれぞれ最も高精度であった。また手順特徴のみを使用した手法がどの地域においても最も低精度であった。地域別では西洋の推定精度が他の地域に比べて低精度であった。

### 3.4 考察

フレンチの属性推定について、画像特徴のみを使用した手法が最も高精度となった。フレンチのタグが付与された料理はキッシュ、ムニエル、チーズフォンデュ、マリネが多くを占めていた。図 3 の例のように、これらの料理は黄色や橙色等の暖色系の色特徴をもつものが多いため、画像特徴を使った手法が高精度になったと考えられる。また、チーズフォンデュのチーズ以外の素材やマリネで使用する素材は料理レシピによって異なる。このように、料理名では素材の特定が困難であるため、他の地域と異なり素材特徴を追加した手法の精度が低下したと考えられる。

次に、手順特徴を用いた手法が低精度であった点について考察する。これは手順動詞の表記ゆれが影響していると考えられる。例えば、中華の属性推定について、中華料理の調理手法には、卵を絡めて揚げる「炸」や仕上げに油をかける「油淋」など独特な調理手法が多かった。しかし、データセット中の料理レシピは日本人が記述した日本人向けのものが多いため、独特な調理動作は簡略化して記述されている可能性がある。このことから、手順特徴を用いた時の推定精度の向上が期待よりも現れなかったと考えられる。

次に、沖縄の属性の推定精度が最も高精度であった点について考察する。特に、素材特徴を用いた推定では、9 割の推定精度が得られた。これは、図 4 の例のように、沖縄料理のほとんどがゴーヤチャンプルのゴーヤ、海ぶどう、ミミガーの豚耳など、日本料理を含め他の地域属性をもつ料理ではあまり使われない素材の素材特徴をもっているためと考えられる。

最後に、西洋の属性の推定精度が最も低精度であった点について考察する。これは、西洋料理の定義が広く、含まれる料理の種類が多岐に渡り、各特徴が現れにくかったためと考えられる。

3 : <https://recipe.rakuten.co.jp/recipe/1700011297/>

4 : <https://recipe.rakuten.co.jp/recipe/1660006926/>

5 : <https://recipe.rakuten.co.jp/recipe/1310001611/>

6 : <https://recipe.rakuten.co.jp/recipe/1220001283/>

表 5 評価結果 (適合率)

使用特徴量	中華	エスニック	フレンチ	イタリアン	日本	韓国	沖縄	西洋	全地域
画像のみ	0.69	0.81	0.75	0.56	0.67	0.65	0.78	0.68	0.72
素材のみ	0.79	0.92	0.65	0.61	0.88	0.69	0.90	0.61	0.78
手順のみ	0.26	0.50	0.41	0.29	0.46	0.12	0.42	0.35	0.40
画像+素材	0.79	0.79	0.64	0.62	0.78	0.73	0.85	0.55	0.72
画像+手順	0.49	0.70	0.66	0.46	0.59	0.32	0.60	0.55	0.59
素材+手順	0.77	0.78	0.58	0.63	0.77	0.80	0.83	0.47	0.70
画像+素材+手順	0.81	0.83	0.71	0.69	0.81	0.86	0.91	0.59	0.77

(a) ゴーヤチャンプルの画像例<sup>5</sup>(b) ミミガーの画像例<sup>6</sup>

図 4 沖縄の地域属性を持つ画像例

## 4 ま と め

本発表では、料理レシピがもつ画像特徴、素材特徴、手順特徴を使用して料理の地域属性を推定する手法を提案した。地域属性(スペイン料理、中華料理など)ごとに料理の味や見た目が大きく異なることに着目し、先行研究の手法[5]をもとに、料理レシピがもつ画像情報、素材情報、手順情報からその完成品料理の属性を推定する手法を提案した。また各地域属性に対して、最も高精度となる特徴量の組み合わせを分析した。実験の結果、全体において、使用特徴量を画像特徴、素材特徴、手順特徴すべての組み合わせとした識別器で8割の推定精度を確認し、手法の有効性を確認した。また、地域別では、イタリアン、韓国、沖縄、西洋料理において全特徴を使用した手法、中華、エスニック、日本料理においては素材特徴のみを使用した手法、フレンチでは画像特徴のみを使用した手法がそれぞれ最も高精度となった。

今後は、素材と手順を組み合わせさせてそれらの関係を学習したり、画像特徴の学習に深層学習を使用することで推定精度の向上を目指す。

謝辞 本研究では、楽天株式会社が国立情報学研究所の協力により研究目的で提供している「楽天公開データ[10]」を利用した。

## 文 献

- [1] 浅沼 駿佑, 中川 明莉沙, 宮脇 佑介, 上田 真由美, 中島 伸介, “食材の嗜好と使用分量を考慮したレシピ推薦システム,” 第5回データ工学と情報マネジメントに関するフォーラム, P3-4, 2013.
- [2] 渡辺 知恵美, 中村 聡史, “オノマトペロリ: 味覚や食感を表すオノマトペによる料理レシピのランキング,” 人工知能学会論文

- 誌, Vol. 30, No. 1, pp. 340–352, 2015.
- [3] 執行 健人, 清光 英成, 大月 一弘, “レシピ選択を支援する食感テクスチャ可視化システムの開発,” 第10回データ工学と情報マネジメントに関するフォーラム, C4-2, 2018.
- [4] 坪村 佳奈, 関 亜紀子, “活動量と栄養摂取量を考慮した献立推薦システム,” 第10回データ工学と情報マネジメントに関するフォーラム, P-84, 2018.
- [5] 伊藤 耀一郎, 道満 恵介, 川西 康友, 平山 高嗣, 井手 一郎, 出口 大輔, 村瀬 洋, “画像付き料理レシピからの味推定の改良 調理手順の利用に関する検討,” 信学技報, MVE2017-97, 2018.
- [6] V.N. Vapnik, “The nature of statistical learning theory,” Springer, 1998.
- [7] G. Csurka, C. Bray, C. Dance, and L. Fan, “Visual categorization with bags of keypoints,” Proc. ECCV2004 Workshop on Statistical Learning in Computer Vision, pp. 59–74, 2004.
- [8] N. Dalal and W. Triggs, “Histograms of oriented gradients for human detection,” Proc. 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 886–893, 2005.
- [9] 森 信介, 笹田 鉄郎, 前田 浩邦, “Flow graph corpus from recipe texts,” Proc. 9th Int. Conf. on Language Resources and Evaluation, pp.2370–2377, 2014.
- [10] 国立情報学研究所, “楽天データセット,” <http://www.nii.ac.jp/cscenter/idr/rakuten/rakuten.html>. (2019/02/04 参照).