

ユーザ調理環境の自動登録に向けた調理道具の認識と登録

吉藤 大貴[†] 宮森 恒[†]

[†] 京都産業大学先端情報学研究科

〒 603-8555 京都府京都市北区上賀茂本山

E-mail: †{i1358130,miya}@cse.kyoto-su.ac.jp

あらまし 毎日の食事における料理を考える際に、Web上のレシピ検索サイトが広く利用されている。ただし、現状の検索結果として得られるレシピには、ユーザの調理環境で利用できる調理道具が考慮されている訳ではない。調理道具を考慮したレシピ検索は、高いユーザ満足が得られることが期待されるものの、調理道具の登録に手間がかかる点が課題となり、現状では普及に至っていない。本稿では、カメラ付き携帯端末を持つ機会が増えていることを踏まえ、調理環境の自動登録化を目的として、ユーザの所持している調理道具を認識・登録する手法を提案する。98種類の調理道具に対し、局所的なエッジ特徴量を用いた識別器を構築し、識別精度を比較する。また、識別がうまくいかない場合の効率的な登録方法を調べるため、いくつかの手法で登録時間を比較する。

キーワード 料理レシピ, 調理道具, 画像認識, 調理環境, 登録, 携帯情報端末
登録時間を比較する。

1. はじめに

近年、インターネット環境の普及とともにウェブサイト上でCookpad(クックパッド)[1]のような数多くの料理レシピ検索サイトが存在し、日常的に利用されている。それに伴い、料理レシピの検索においても様々な手法が提案され、ユーザが目的に応じた料理レシピを効率よく検索することが出来るようになり、大変便利なものとなってきている。

現在の料理レシピ検索の多くは、使いたい食材名などをキーワードとして入力することで、その食材を含むレシピを検索結果として返すが、ユーザ個別のニーズに合わせた、より高度な検索を実現しようとすると、ユーザ属性をシステムに何らかの手段で入力する必要がある。ユーザにとって大きな手間がかかる点が問題となる。例えば、各家庭の冷蔵庫にある食材を考慮した検索[2]を実現するには、買い物の直後や料理をした後に、その時点での冷蔵庫内の食材情報を登録・更新する必要がある。また、各ユーザの調理環境を考慮した検索[3]を実現するには、各ユーザが台所で使用しているフライパンや鍋などの調理道具を少なくとも一度はシステムに登録する必要がある。これらの登録作業を全て手入力で行うのは、ユーザに過剰なシステム操作や時間的な負担を要求するものであり、自動化をはじめとする何らかの効率化を図る手段が必須となる。

一方、近年では個人でのスマートフォンやタブレットをはじめとするカメラ付き携帯情報通信端末の保有率が増加しており、身近な風景を手軽に撮影し、ネットを介したシステムへの登録作業が容易に行える環境となってきている。

そこで、本稿では、カメラ付き携帯端末を持つ機会が増えていることを踏まえ、調理環境の自動登録化を目的として、ユーザの所持している調理道具を認識・登録する手法を提案する。98種類の調理道具に対し、局所的なエッジ特徴量を用いた識別器を構築し、識別精度を比較する。また、識別がうまくいかない場合の効率的な登録方法を調べるため、いくつかの手法で登

本稿の構成は以下の通りである。2章では関連研究について、3章では提案システムについて説明する。4章では実験により、調理道具を対象とした画像についての識別精度、および、識別がうまくいかない場合の効率的な登録方法について調査し、それらに伴う結果と考察を述べる。5章でまとめる。

2. 関連研究

料理レシピ検索の従来研究として、冷蔵庫食材を考慮した料理レシピ検索の研究がある[2]。冷蔵庫食材の分量や賞味期限を考慮したレシピのランキング手法が検討されているが、ユーザによる冷蔵庫食材の登録や更新が必要であり、その負担を軽減する効率的な登録・更新方法が課題となっている。

また、ユーザの調理環境を考慮した料理レシピを検索する研究がある[3]。調理道具と調理動作のペアに難易度を付与し、ユーザの調理技能や調理環境に、よりマッチした料理レシピ検索について検討されている。ここでも、ユーザの所持する調理道具を登録する手段が、現状では、チェックボックスによる手動入力となっており、ユーザの負担を軽減する何らかの自動登録手段の実現が有用であると期待される。

一方、画像中に映っている各物体を認識し、その一般的な名称を認識結果として出力する一般物体認識は、コンピュータビジョンの重要な研究課題の一つとしてこれまでに数多くの研究がおこなわれている[4][5]。現在、「顔」「自動車」「ライオン」などの様々な対象に対して、大量の学習データを用意し、領域ごとのラベリング結果や局所的な特徴の組合せに対して、統計的機械学習を適用する手法が数多く提案されている。特に、料理に関連する画像を対象とした認識についても盛んに研究が行われている。食堂の定食のように、複数品目の料理が複数の皿に盛り付けられた食事画像を対象とした認識手法が提案されている[6]。料理部分を候補領域として検出し、複数の視覚的特徴を統合して利用したSVMによる分類を行うことで、85種類の

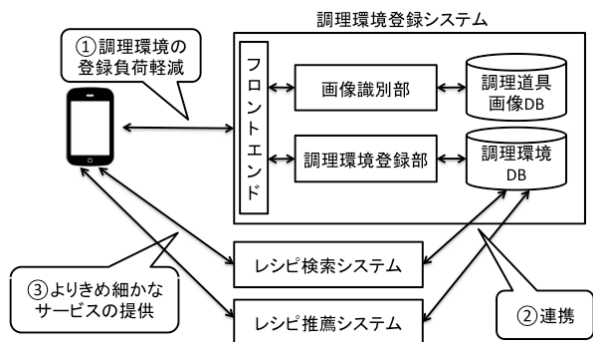


図1 システム構成

料理について 60%程度の分類率を達成している。

また、スマートフォンを用いて撮影された食材画像を対象とした認識を行うことで、その食材を使ったレシピを推薦するシステムが提案されている [7]。30 種類の食材に対して 45%程度の認識率を達成している。さらに、食器洗浄作業の自動化を想定し、上部から撮影した単一の皿やカップ等が映った食器画像を対象として、輪郭特徴量の累積ヒストグラムによりマッチングする手法が提案されている [8]。9 種類の食器について 90%以上の認識率が得られたとされている。

さらに、ロボットによる一般的な家事支援として調理行動をとりあげ、ロボットによる調理道具・野菜の発見と状態認識を試みる研究が報告されている [9]。「ボウル」や「包丁」といった調理道具やまな板上の食材を、画像中の色や特定の幾何形状といった基礎的な画像特徴の組み合わせで認識する手法が報告されている。

本稿では、スマートフォンで撮影された調理道具画像を対象として、該当する調理道具名を識別・登録することを目的とする。

3. 提案手法

本稿で検討するユーザ調理環境の登録システムの概要を図1に示す。ユーザが所持している調理道具を、カメラ付き携帯情報通信端末で撮影することで、調理道具が識別され、その結果が、ユーザ毎の調理道具環境として登録される。登録された調理道具環境は、他のレシピ検索や推薦と適宜連携、より高度な検索や推薦を実現するために活用させることができる。

本稿では、上記登録システムのうち、調理道具を撮影した画像を入力として、調理道具名を識別する手法について検討する。

3.1 調理道具の識別手順

調理道具画像の識別手順を図2に示す。

● 前処理部

まず、Web より、画像検索 API を利用して、調理道具名が既知の調理道具画像を収集する。次に、各調理道具画像に対して、画像特徴量を計算し、調理道具名とともに調理道具画像 DB に登録する。

● システム利用時

ユーザによる調理道具識別を行う際には、名前が未知の調理道具を撮影した画像が入力され、上記と同様の手順で画像特徴

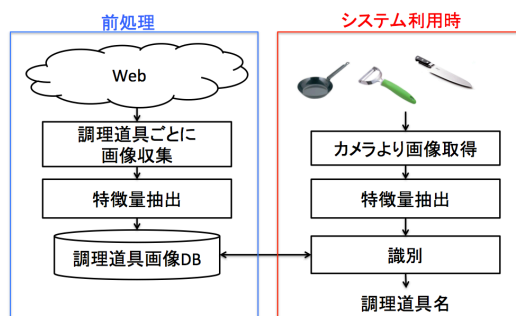


図2 識別手順

量が計算される。得られた画像特徴量を用いて、調理道具画像 DB 内の画像群を k 近傍法で探索し、該当する調理道具名を識別結果として出力する。

3.2 識別カテゴリの整理

本稿で対象とする調理道具は、野田ら [2] の研究を参考にした。野田らは、クックパッドのレシピ約 43 万件から調理道具名 966 種類を抽出し、表記ゆれや打ち間違いを修正した 202 種類の調理道具名を使用している。202 種類の調理道具の中には、同一の道具で呼称が異なるもの (例えば、「お玉」と「レードル」) や、材質・大きさのみが異なり、調理における役割としては同一の道具 (例えば、「へら」と「ゴムベラ」、「小鍋」と「大鍋」) が存在した。これらについては、調理時に差し替えが効くものと考え、それぞれ識別カテゴリとしては同一のものとしてまとめることとした。また、特定のメーカー名がある調理道具を指す呼称として一般化しているもの (例えば、「ル・クルーゼ」) については、調理道具の基本機能が同一であっても、特定の付加価値機能をもつ (例えば、「熱が逃げにくい」) ことが多いと考えられるが、同カテゴリとして分類することとした。以上より、識別対象とする調理道具カテゴリを 98 種類に整理した。

本研究で検討するユーザ調理環境の登録システムにおいて、この 98 種類の調理道具名を識別し登録できるようにすることを目標とする。

3.3 調理道具画像の収集

次に、98 種類の調理道具名に該当する画像を収集する。収集時には、該当する調理道具画像をより網羅的に集めるため、野田らが使用した 202 種類の調理道具名を用いることとした。画像収集には、検索エンジン Bing の検索 API を利用し、各調理道具名をクエリとして与え、それぞれ約 1000 件ずつ画像を収集した。収集する画像中の調理道具の向きや角度、色や大きさについて特に制限は設けていないが、調理道具を学習させる上で必ずしも適当とはいえない画像が含まれており、それらを人手で判断し除外した。具体的には、複数の調理道具が写り込んでいる画像、文字やイラストが画像中で大きな割合を占め混



図3 除外した画像の例

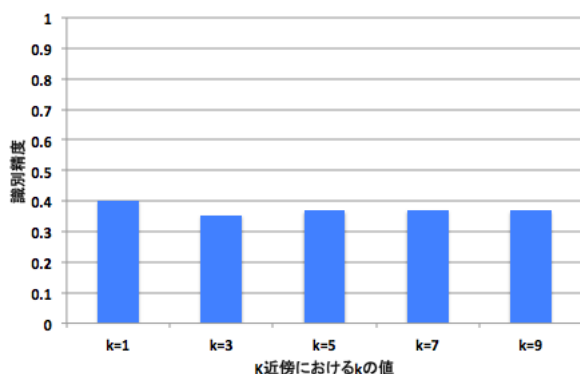


図4 k近傍法での識別精度

する画像、写真ではなくイラストである画像、調理道具の一部が他の物体に隠され、全体の形を捉えられない画像、調理道具の一部が画像から見切れている画像を除外した。図3に除外した画像の例を示す。

各クエリに基づき収集した画像を、3.2節の分類に従って各カテゴリに対応づけた。以上より、各識別カテゴリの調理道具画像が、14~1714件取得された。

3.4 特徴量抽出

収集した調理道具画像から識別に有用と思われる特徴量を抽出する。筆者らは、従来研究[10]において、類似画像検索に用いられる大局特徴量および局所特徴量を使用し、23種類の調理道具カテゴリについて、それぞれどの程度の識別性能が得られるかを調べた。その結果、局所的なエッジ特徴のみを用いた特徴量の場合、最も高い識別精度が得られることが明らかとなった。本稿では、そこで用いられた以下のEdge histogramを採用することとした。

Edge histogram 画像の局所的なエッジ特徴の分布を表す[11]。画像を4×4ブロックに分割し、各ブロックでエッジを検出し、垂直、水平、斜め45度、斜め135度、エッジなしの5タイプのいずれかに分類する。2つのEdge histogramの類似度計算には、L1-ノルムを用いる。

4. 実験

4.1 実験1: 調理道具画像に対する識別精度

4.1.1 目的および方法

調理道具画像に対する識別精度を明らかにする。

被検索対象となる98カテゴリの調理道具画像に対し、識別精度を調べる。3.3節で説明した局所的なエッジ特徴量を用いたk近傍法で識別する。98カテゴリの調理道具画像計22682枚に対し、5-fold交差確認で評価する。k近傍法での識別におけるパラメータkについては、1,3,5,7,9と変化させる。また、

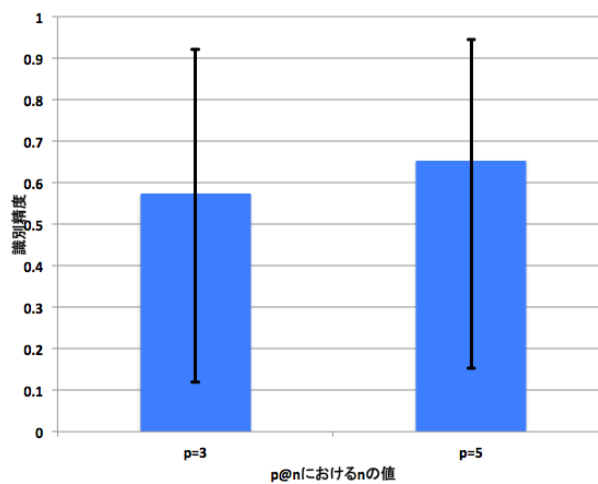


図5 p@nでの識別精度



図6 同一調理道具でも形状が大きく異なる多様な見え方の例（ピーラーの場合）

異なる調理道具カテゴリの上位n件に正しいカテゴリが含まれていれば正解と数える方法(本稿では、カテゴリ上位n件の精度、p@nと呼ぶこととする)でも精度を調べる。ここでは、n=1, 3, 5とした。

4.1.2 結果

実験結果を図4, 図5に示す。図中の識別精度は、各テストカテゴリによる識別精度のマクロ平均を示している。

4.1.3 考察と課題

実験結果より、p@5においては、マクロ平均で0.6を超える識別結果が得られた。ただし、識別精度のばらつきが大きく、カテゴリによっては、精度が0.9を超えるものや0.1程度にしかないものも存在した。これは、主に、カテゴリごとの学習データ数にばらつきが大きいこと、および、図6で示すように、同一の調理道具の場合でも多様な見え方が存在し、それらを学習データが十分にカバーできていないことが挙げられる。より多様な見え方を含む、多くの正例画像を収集し、新たに実験をする必要がある。

4.2 実験2: 代替手段による登録時間の比較

4.2.1 目的と方法

調理道具画像の識別がうまくいかない場合の効率的な代替手段を明らかにすることを目的とする。調理道具名のテキスト入力(サジェスト機能を含む)による登録方法と、調理道具名が50音順に表示されている一覧でチェックボックス選択する登録方法を代替手段の候補とし、一般的な家庭に存在すると考えられる33件の調理道具を登録し終えるまでの時間を計測した。普段から携帯端末やPCをよく利用している情報分野の大学生3名を被験者とし、登録作業は、携帯端末とPC上のそれぞれで行ってもらうこととした。

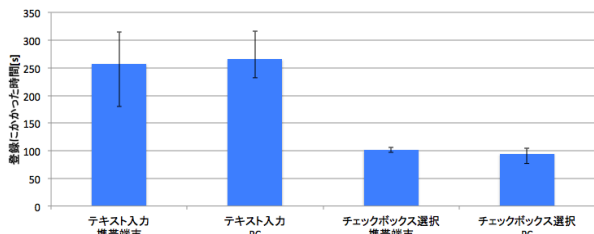


図 7 代替となる登録方法による登録時間の比較

4.2.2 結 果

実験結果を図 7 に示す。

テキスト入力による登録より、チェックボックス選択による登録の方がより短い時間で登録できることが確認された。また、携帯端末、PC のいずれの端末を用いる場合も、登録時間に大きな差は見られなかった。

4.2.3 考察と課題

実験結果より、調理道具画像の識別がうまくいかない場合の代替登録機能としては、チェックボックス選択による登録方法を実装することが効果的だと考えられる。

画像識別による登録方法では、各ユーザの撮影条件の違いにより、調理道具の見え方や背景等に多様性が出てくるため、識別精度に大きな違いが生じ、登録に要する時間が大きく変動する可能性がある。登録部を実装した上で、改めて登録に要する時間を評価する必要がある。

5. ま と め

本稿では、レシピ検索やレシピ推薦における、ユーザの調理環境の自動登録の課題を取り上げ、カメラ付き携帯情報通信端末等で撮影されたユーザの調理道具画像から、その調理道具名を識別・登録する手法を提案した。98 種類の調理道具に対し、局所的なエッジに基づく特徴量を用いた識別器を構築し、識別精度を比較した。また、調理道具の識別が誤った場合を想定し、代替登録方法のユーザビリティについて評価した。

調理道具の識別については、カテゴリ上位 5 件の精度で評価した結果、マクロ平均で 0.6 を超える識別結果が得られた。ただし、カテゴリ毎で精度の差が大きいため、より多様な見え方を含む、より多くの正例画像を収集し、実験することが課題である。また、他の識別手法についても検討を進めていきたい。

今後、提案手法のスマートフォンへの実装を進め、既存レシピ検索システムと連動させていく予定である。

文 献

- [1] レシピ検索 No.1/料理レシピ載せるならクックパッド
<http://cookpad.com>
- [2] 赤澤康幸, 宮森恒. 冷蔵庫食材を考慮した料理レシピ検索システムの提案, DEIMForum2012, E1-2, 2012.

- [3] 野田真, 宮森恒. 料理レシピにおける調理動作の道具別難易度付与の一検討, データ工学研究会, 2012, 信学技報, vol. 112, no.5, DE2012-3, pp. 13-18, 2012.
- [4] 柳井啓司. 一般物体認識の現状と今後, 情報処理学会論文誌:コンピュータビジョンとイメージメディア, Vol.48, No.SIG16(CVIM19), 2007.
- [5] 秋山瑞樹, 柳井啓司. 特定物体認識手法による大量画像を用いた一般物体認識, 画像の認識・理解シンポジウム (MIRU2010), 2010.
- [6] 甫足創, 松田裕司, 柳井啓司. 候補領域推定による複数品目に対応した食事画像認識, 画像の認識・理解シンポジウム (MIRU2011), 2011.
- [7] 丸山拓馬, 秋山瑞樹, 柳井啓司. 食材画像認識を用いたレシピ推薦システム, 電子情報通信学会技術研究報告, 信学技報 111(478), 43-48, 2012.
- [8] 野村英祐, 林悠, 鏡慎吾, 橋本浩一. 画像上の輪郭特徴量を用いた食器認識システム, 計測自動制御学会東北支部第 236 回研究集, 資料番号 236-12, 2007.
- [9] 渡辺義明, 長濱虎太郎, 山崎公俊, 岡田慧正, 稲葉雅幸. 一般調理道具を扱う等身大ヒューマノイドの認識行動システム統合とサラダ調理の実現, ロボティクス・メカトロニクス講演会講演概要集 2010, “2A2-A17(1)” - “2A2-A17(4)”, 2010.
- [10] 吉藤大貴, 宮森恒. ユーザ調理環境の自動登録に向けた調理道具の識別手法の検討, データ工学研究会, 2014, 信学技報, vol. 114, no.204, DE2014-23, pp. 1-5, 2014.
- [11] P. Salembier and T. Sikora. Introduction to MPEG-7: Multimedia Content Description Interface. John Wiley & Sons, Inc., New York, NY, USA, 2002.