

調理レシピテキストからの代替素材の発見

Discovery of replaceable materials from cooking recipe texts

志土地 由香*1

Yuka Shidochi

高橋 友和*2

Tomokazu Takahashi

井手 一郎*1*3

Ichiro Ide

村瀬 洋*1

Hiroshi Murase

*1名古屋大学大学院情報科学研究科

Graduate School of Information Science, Nagoya University

*2岐阜聖徳学園大学経済情報学部

School of Economics and Information, Gifu Shotoku Gakuen University

*3国立情報学研究所

National Institute of Informatics

The number of cooking recipe texts published on The Web is increasing in recent years. However, in general, cooking recipe texts have little flexibility. So, it is not always easy to retrieve cooking recipe texts that satisfy users' various demands. Therefore, it is necessary to create and offer recipes that suit the users' requirements, and to offer it. In this research, it propose a method for discovering an replaceable materials from a large amount of cooking recipe texts, so that we can use the knowledge to modify cooking recipes so that it should conform with the users' requirements.

1. はじめに

近年、料理本をはじめ、さまざまな形で、調理レシピテキストが公開されている。特に Web 上では、個人が調理レシピを気軽に公開できるようになり、Web 上の調理レシピテキストが急増している。長年、主婦にとって大変なことは献立決めであったため、そのレシピを活用しさまざまな献立を考えることができる。しかし、調理レシピテキストは、一般に素材や手順が決められており、柔軟性に欠ける。実際に料理を作る際は、既存の調理レシピを状況や気分に応じて変えるものである。例えば、以下のようなさまざまな条件を満たす料理を作りたいという要求がある。

- 冷蔵庫の中にある素材を用いた料理
- 安価で手に入りやすい素材を用いた料理
- 健康状態や栄養状態に適した料理
- 好き嫌いを考慮した料理

しかし、上述のように一般の調理レシピテキストは柔軟性に欠けるため、既存の料理レシピ検索システム*1、ゲーム機*2などを用いて要求を満たす調理レシピを見つけ出すことは困難である。そこで、我々は既存の調理レシピテキストに対して素材を適切に入れ替えることにより、ユーザの要求に即した調理レシピを創作することを考えた。

本発表では、そのレシピを創作するために必要な知識として、Web 上に存在する大量の調理レシピテキストから代替素材を発見する手法を提案する。

2. 関連研究

調理は多くの家庭で日常的に行われる高度な知識に基づく複雑な創作活動である。そのため、さまざまな方法で調理を支

連絡先: 志土地由香, 名古屋大学大学院 情報科学研究科 メディア科学専攻 村瀬研究室, 〒464-8601 愛知県名古屋市千種区不老町, (052)789-3310, shidochi@murase.m.is.nagoya-u.ac.jp

*1 “レシピ検索 No.1 / 調理レシピ載せるならクックパッド”, <http://cookpad.com/>

*2 Nintendo, “しゃべる! DS お料理ナビ”

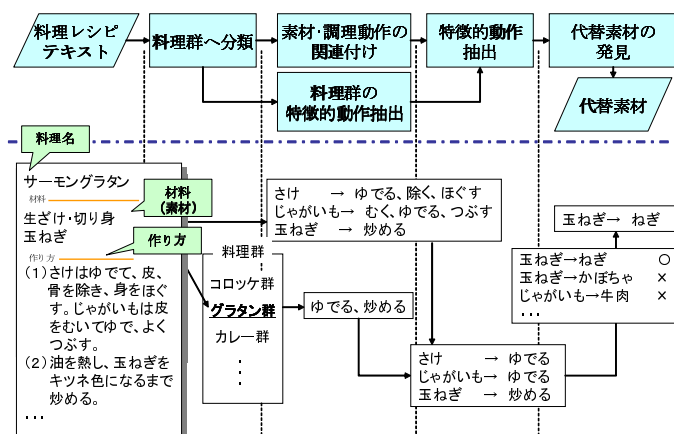


図 1: 代替素材発見までの処理の流れ

援しようとする研究がなされている。例えば、カメラなどを用いて調理動作を認識し、適切なタイミングで次の手順を教示する研究 [森 08] などがある。

本研究と共通の目的を持つ先行研究として、データマイニングの手法を用いて代替素材を発見する研究 [唐沢 06][大平 00]がある。しかし、これらの先行研究の手法では、得られる代替素材対が少なかったり、素材と調理動作の関連を考慮していないため、実用には不向きであったり、代替素材発見の精度が低下する。

一方、ユーザが必要とする調理レシピを提案する関連研究がある。例えば、ユーザの食材利用頻度、食材の特異度から食材の特徴を尺度化し、個人の嗜好を考慮するもの [石原 08]、基本的なユーザの要求事項をデータベース化することで、栄養や価格などの面からユーザの要求を満たしたものを素材の代替を含めて提案する研究 [大平 01] がある。しかし、これらの研究は既存のレシピを提案するため、ユーザが求める調理レシピを見つけ出すことは困難であると思われる。

表 1: 料理群の生成結果 (抜粋)

料理群名	レシピ数	レシピ例
サラダ	496	かつおの刺身サラダ、マカロニサラダ
煮	447	大根のべっこう煮、豚すき煮
汁	272	たけのこ汁、芋煮汁
スープ	259	キャベツスープ、セロリのスープ
焼き	244	ラム肉のみそくわ焼き、マヨ卵焼き
炒め	239	うずら卵の甘酢炒め、かきの辛味炒め
かけ	132	もずく山かけ、油菜のカキ油かけ
揚げ	131	たらの白扇揚げ、豚肉の南蛮揚げ
添え	130	焼き卵の野菜添え、春巻のミント添え
漬け	128	大根の黒酢漬け、かぶの即席漬け

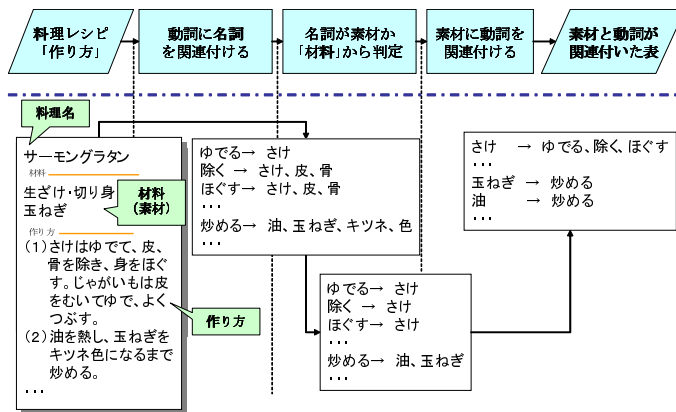


図 2: 素材と調理動作の関連付け処理の流れ

3. 調理レシピテキストからの代替素材の発見

本章では、調理レシピテキスト中の素材・動作を抽出し、代替素材を発見するまでの手順について述べる。3.1 節で全体的な処理の流れを示し、3.2 節以降でそれぞれの処理の詳細を述べる。

3.1 処理の流れ

本研究では素材と調理動作の関連付けを行い、素材間の類似性を考えることで、新たな代替素材を得ることを目指す。本研究で提案した処理の流れを図 1 に示す。

調理レシピテキストは、特定のレシピサイトの HTML テキストの規則性を利用して「料理名」「材料」「作り方」を構造化したものである。

3.2 料理群の作成:料理名によるレシピの分類

料理は多様であるため、収集した全ての調理レシピテキストから料理の特徴を求めることは困難である。そのため、ある程度似た種類の料理に分類をして、その特徴を求める必要がある。そこで、「料理名の語尾が同一であることが、ある程度共通の手順により作られる料理である」と仮定して、調理レシピテキストを分類した。ここで、分類された調理レシピテキストのグループを料理群と名付けた。分類された料理群の上位 10 位を表 1 に示す。

3.3 素材・調理動作の関連付け

調理動作が類似しているという観点から代替素材を見つけるためには素材、調理動作を関連付ける必要がある。そのためにまず調理レシピテキスト中から素材、調理動作を抽出する。抽出するために、調理レシピテキストを形態素解析する^{*3}。解析の結果より品詞の種類に注目して、関連付けを行う。ここで、調理レシピテキスト内の「材料」に出現し、かつ「作り方」にも出現する名詞は素材、「作り方」に出現する動詞は調理動作であると考えた。

以下のようなものは、処理に悪影響を与えるため、必要に応じて辞書を作成して表現を統一して抽出を行う。

- 形態素解析により、本来 1 つの単語であるものが 2 つの単語に分かれるもの
- 有対動詞 (例「揚げる」、「揚がる」)
- 表記の揺らぎ (例「ぎょうざ」、「餃子」、「ギョーザ」)

*3 形態素解析には、和布蕪 (<http://mecab.sourceforge.net/>) を用いた。

以降の処理においても、素材、調理動作は、全て上記のように表現を統一する。

料理群中の素材と調理動作を関連付けてセットとして考えると、料理群中の素材に対する調理動作に共通性が見られる。よって調理動作の共通性を考えるには、まず素材と調理動作の関連付けが必要である。素材と調理動作を関連付けるための詳しい手順を図 2 に示す。

関連付けを行う際に、調理レシピテキスト中の「作り方」に注目する。この部分は以下の特徴を持つ。

- 手順番号 ((1), (2), ...) を持つ。
- ある手順に従って生成されたものを次の手順で用いる際、簡略化のためにその手順番号で表す。
- 複数の素材を同時に用いる際、簡略化のためにまとめて「A」や「B」などの記号で表す。

このような特徴を持つ調理レシピテキストから素材と調理動作を関連付けるための条件を次に述べる。

1. ある 1 文中に出現する動詞の前にある全ての「素材名」は、その動詞に係る。
2. 条件 1 で素材名がない場合、その前文に出現する全ての素材名がその動詞に係る。
3. 手順中に手順番号が使われる場合、その 1 文に出現する調理動作に係る素材名とともに、手順番号の手順に出現する素材名も対応付ける。

日本語は、主に「主述構造」をとる。また、一般に主語には名詞を用い、述語に動詞を用いる。このため、ある動詞に係る名詞は、一般にその動詞より前にあると考え、条件 1 を決めた。主語が省略された場合、その前文あるいはそれ以前の手順で出現する素材が主語である可能性が高いと考えたため、条件 2 を決めた。

3.4 料理群の特徴的動作抽出

素材と関連付いた調理動作が全て一致するものは代替可能であると考えられるが、レシピの書き方や細かい調理方法の違いにより関連付く調理動作は変化する。そのため、素材間の調理動作が全て一致するものを発見するのは困難である。そこで、共通性を見つけるために、最低限必要な料理群の特徴を表す調理動作と素材の関連に注目する。調理レシピテキスト中の各調理動詞につき、経験的に TFIDF 値を求める。4 章で示す実験では、TFIDF 値が 0.8 以上である場合に料理群における特徴的な動作とした。

3.5 レシピにおける特徴的動作の抽出

3.3節で求めた素材に関連付いた調理動作から、3.2節で求めた料理群に対して特徴的な調理動作のみを抽出し、素材に関連付く調理動作を限定した。しかし、料理群内の調理レシピテキストには、同一素材も数多く使われている。よって、代替する素材を発見するためには、料理群の特徴的な調理動作を素材に関連付けるだけでなく、「料理群中に出現する同一素材」に共通する特徴的な調理動作も考慮しなければならない。料理群内でその素材が出現する調理レシピテキスト数の75%を上回るものを、ある素材に対する特徴的な調理動作として抽出する*4。素材に関連付けられる調理動作のうち、代替素材を発見するために必要な素材間の調理動作の類似度を求めるために、上記2種類の特徴的な調理動作を組み合わせて素材ごとの共通調理動作ベクトルを作成する。

素材ごとの共通調理動作ベクトルの類似度を求める際にはコサイン距離を用いる。コサイン距離を用いる際、素材が同種類か否かで代替可能性は変化すると考えられるため、重み付けを行う。素材が同種類か否かは食品成分表 [食品成分表] を用いて判断する。食品成分表に掲載されていない素材に関しては経験的に判断をした。

3.6 代替素材の発見

2つの素材の共通調理動作ベクトル内の類似度が小さいほど、その2つの素材における特徴的動作が類似し、料理群中で同じように調理されると考えられる。よって、共通調理動作ベクトルの類似度が小さいほど2つの素材の代替可能性が高いと考える。

素材代替をするために作成した料理群は、料理の性質に深く関係している。本研究では、料理の性質を変えることを目的としていないため、料理の性質に関わりがある調味料類は素材の代替対象外とした。

3.5節の処理で、代替可能と判断された素材対の中には、料理群中の同じ調理レシピテキスト中で何度も同時に使われることが多いものがある。

この場合は代替素材対がその料理群における重要な素材であると考えられるため、以下の実験では、料理群中の全レシピのうち、3%以上で同時に使われる素材対は代替できないとした。

4. 代替素材発見実験

本章では、第3章で述べた手法を用い、代替素材の発見実験を行った結果と考察をまとめる。

4.1 実験条件・手順

第3章で述べた手順に従って、料理群の中から9群391件の調理レシピテキスト表2に関して代替素材の抽出実験を行った。これらの料理群の含まれる素材数は168種類であった。

4.2 実験結果

料理群(1)~(9)の中で92組の代替素材案が得られた。結果を見ると、「玉ねぎ」「ねぎ」「鶏ひき肉」「牛ひき肉」など、日常的に代替することがある同種類の素材間の代替結果が数多く見られた。また、「じゃがいも」(いも及びでんぷん類)「かぼちゃ」(野菜類)など異種類間での代替可能性に関しても極端に違和感を覚える結果は少なかった。

*4 例えば、コロケ群の特徴的な動作が「揚げる」であったとする。そのとき、コロケ群中の調理レシピテキストに出現するじゃがいもにおける特徴的な動作は「揚げる」ではない。なぜなら、じゃがいもはまず、「つぶす」のが一般的だからである。よって、じゃがいもに対する特徴的な動作は「つぶす、揚げる」である。

表 2: 実験に用いた料理群

番号	料理群	レシピ数
(1)	コロケ	26
(2)	マリネ	67
(3)	オムレツ	32
(4)	トースト	32
(5)	カレー	51
(6)	グラタン	70
(7)	巻き	72
(8)	きんぴら	27
(9)	肉じゃが	14

表 3: 評価実験

指標	1	2	3	4
レシピ数	15	15	1	9

調理レシピテキスト中には、料理本体に使われる素材と飾りに使われる素材など異なった性質をもつ素材が混合して記載されることがある。本実験では、素材と調理動作を関連付けたことで、これらが混同されることなく飾りに使われる素材と料理本体に使われる素材の代替が区別できた。

4.3 評価実験

この代替素材発見結果をもとに、日常的に料理に従事し、栄養士の資格を持つ被験者に評価をしてもらった。評価にあたり、Web上の調理レシピから分量を消去した調理レシピテキスト13種類をテンプレートとし、得られた代替素材を1つずつ入れ替えた調理レシピ40種類を生成した。これにより、13+40種類の創作調理レシピが用意された。実験にあたっては公平を期すために、既存レシピであるか、創作した調理レシピであるかを伏せて、53種類のレシピを被験者に提示した。また、料理名に代替する素材名が入っている場合、その素材名も伏せた。

評価として、その調理レシピが以下のいずれであるか選んでもらった。

1. 素材・調理法ともに妥当である。
2. 素材は妥当であるが、作り方における調理法(動作)の表現方法を変える必要がある。
3. 素材の入れ替えは無理である。
4. 判別不可能である。

表3に創作レシピ40種類に関する評価結果を示す。なお、テンプレートとした調理レシピ13件は全て指標1であった。この結果、代替素材が妥当である割合は75%であり、比較的良好に素材の代替に成功していることがわかった。

指標1または指標2と評価された素材の代替を、表4に示す。

4.4 考察

評価実験では指標1と指標2が同数であったので、素材名を単純に置き換えるだけでは、整合性をもつレシピを創作するのに不十分であることが分かる。指標3の代替は、マカロニサラダのマカロニとご飯の代替であった。ご飯ではなく、米であれば代替は可能である。「材料」中で、素材に何らかの調理がなされていることで代替できる可能性が低下するものがあると考えられる。

表 4: 被験者実験により妥当と評価された素材対

料理群	代替する素材	代替される素材
グラタン	ベーコン	カリフラワー、ハム
	マッシュルーム	しめじ
	玉ねぎ	キャベツ、ねぎ カリフラワー
	溶き卵	溶き卵黄
	じゃがいも	豆腐、カリフラワー
	ほうれん草	ベーコン、ブロッコリー
カレー	じゃがいも	かぼちゃ
	ピーマン	マッシュルーム
マリネ	白魚	さけ
	パプリカ(黄)	トマト
コロッケ	ねぎ	玉ねぎ
	鶏ひき肉	牛ひき肉
	牛ひき肉	鶏ひき肉、ねぎ
	エンダイブ	ラディッシュ、青じそ チャービル
	玉ねぎ	ねぎ
トースト	サラミ	にんじん
オムレツ	ベーコン	にんじん
きんぴら	ごぼう	うど
巻き	スモークサーモン	サラミ、まぐろ
	グリーンアスパラ	トマト、まぐろ

5. おわりに

本発表では、さまざまな調理レシピテキストから代替素材を発見する手法を提案した。提案手法では、「玉ねぎ」「ねぎ」「じゃがいも」「かぼちゃ」など比較的良好な結果が得られた。また、栄養士の資格を持つ被験者による評価実験からは、本研究で得られた代替可能素材の75%が適切な結果であることが確認された。

今後の課題として以下のようなことが挙げられる。

料理群に分類する際の問題点

語尾が素材であるものや語尾に調理法が表われないものが数多く存在する。そのため、料理名に注目した分類だけでなく、作り方や、素材を用いた分類などと組み合わせることで料理群の分類精度を高める必要がある。

代替素材を発見する際の問題点

発見した代替素材を用いて調理レシピを創作する際、素材と調理動作が対応しないため、素材を代替するだけでは不十分なときがある。よって、調理動作の表現も素材に対応したものに置き換える必要がある。今後、素材に対する調理動作の順序を考慮した関連付けを行うなど、工夫しなければならない。また、調理レシピを創作する際には、素材ごとに対応する調理動作を連動して代替することで矛盾点を解消することが必要である。今後は、どのように素材に対応付いた調理動作を抽出し、素材と調理動作をセットとして置き換えられるか考えていく。

全体を通じた問題点

提案手法は、料理群の作成や素材、調理動作の抽出などテキスト処理部分の精度に大きく影響される。料理名は

表記のゆれや連濁があり、さらに、調理レシピテキストは一般的ではない語句が出現したり、素材、調理動作の表記方法が多岐にわたることが多い。本研究では、必要に応じて辞書を作成したが、将来は調理レシピテキストに網羅的に対応した辞書や形態素解析が必要である。

また、この手法を広く多様な料理群に適用しても、素材が多様なため、対応する共通した調理動作が少なく、代替可能と判断される素材が多岐にわたってしまう。このことから、どこまで詳細な料理群を作成すべきか検討する。

一方、本研究の実験結果より、調理動作のみではなく色や形状で代替可能素材を発見できる可能性があることがわかった。食卓の彩りを良くするシステム[森 08]などがあるように、料理は味はもちろんのこと彩りも大事である。よって素材の色や形状の代替も考えられる。これをふまえて料理群内の代替素材発見ではなく、料理群という分類の枠を超えた代替可能素材を発見することもできると考えられる。さまざまな代替素材を発見するには、料理群内だけではなく料理という分野全体に共通する知識を抽出することも視野に入れる必要がある。

今後、さらにデータを増やして料理群と知識を抽出したいと考えている。

このような課題があるが、最終的にはユーザが作りたと思うレシピのほか、ユーザが希望する素材や彩りを入力することで新たな創作レシピを提供できるようなシステムの実現を考えていく。

参考文献

- [森 08] 森直幸, 船富卓哉, 山肩洋子, 角所孝, 美濃導彦, “調理者の手の動きを時空間制約とした調理中の食材追跡”, 信学技報, MVE2007-76, Jan. 2008.
- [唐沢 06] 唐沢隆, 浜田玲子, 井手一郎, 坂井修一, 田中英彦, “料理教材テキストからの素材と調理法に関する知識の抽出”, 第 66 回情報処理学会全国大会講演論文集, vol.2, pp.119-120, Mar. 2006.
- [大平 00] 大平峰子, 大園忠親, 新谷虎松, “料理における類似性評価のための知識の発見手法について”, 第 14 回人工知能学会全国大会論文集, pp.396-399, July 2000.
- [石原 08] 石原和幸, 上田真由美, 平野靖, 梶田将司, 間瀬健二, “FF-IRF を用いた個人嗜好レシピ推薦手法の有効性検証”, 信学技報, MVE2007-77, Jan. 2008.
- [大平 01] 大平峰子, 大園忠親, 新谷虎松, “類似性評価知識を用いたレシピ検索システム mineRecipe の実装について”, 第 62 回情報処理学会全国大会講演論文集, vol.3, pp.129-130, Mar. 2001.
- [食品成分表] 食品成分研究調査会, “五訂日本食品成分表”, 医歯薬出版, 2001.
- [森 08] 森麻紀, 栗原一貴, 塚田浩二, 椎尾一郎, “いろどりん: 食卓の彩り支援システム”, 信学技報, MVE2007-80, Jan. 2008.