

同一料理に対する多様なレシピ集合からの 効率的な選別を目的とした可視化手法

村瀬 秀[†] 牛尼 剛聡[‡]

[†]九州大学芸術工学部 〒815-8540 福岡市南区塩原 4-9-1

[‡]九州大学大学院芸術工学研究院 〒815-8540 福岡市南区塩原 4-9-1

E-mail: [†] fonewmjp@gmail.com, [‡] ushiama@design.kyushu-u.ac.jp

あらまし 近年、投稿型レシピサイトが一般的に利用されるようになった。投稿型レシピサイトでは、利用者がレシピを投稿できるため、ひとつの料理に対して多数のレシピが存在する場合が多い。特定の料理を作ろうとしてレシピサイトでレシピを見ている利用者は、同一の料理に対する様々なレシピから自分が作るレシピを選別する必要がある。しかし、候補となる全てのレシピを閲覧することは現実的ではない。そこで、他のユーザが行った評価を利用してレシピを選別することが多い。しかしユーザの評価に基づいた選別では、新しいレシピは選別の候補にならない可能性が高く、その料理のバリエーション（料理の全体像）を把握して、自分の嗜好や目的に合致したレシピを選別したいという要求に対応できない。この問題を解決するために、本論文では、ユーザが指定した料理に対して、多数のレシピが存在する状況で、レシピの選別を支援する可視化手法を提案する。本研究では、レシピの構成要素を分析することでレシピのバラエティを発見し、ユーザが料理の全体像が理解できるような可視化を行う。

キーワード レシピ選別, 可視化手法, 階層的クラスタリング

1. まえがき

近年、投稿型レシピサイトには大量のレシピが投稿され、ユーザは様々なレシピから自分の目的にあったレシピを選択できるようになった。レシピサイトの代表的な利用方法の一つは、ユーザが調理したい料理を指定して、そのレシピを検索するというものである。しかし、投稿型サイトには膨大な数のレシピが登録されているため、料理を指定した場合には、大量のレシピが検索結果として返されることが多い。例えば、代表的なレシピサイトの一つである「楽天レシピ」[1]では、2014年1月の段階で「オムライス」のレシピとして1,755件が登録されており、今後もさらに増加していくと予想される。

レシピサイトの検索結果として返される大量のレシピを、ユーザがすべて確認することは現実的でない。そこで、多くのレシピサイトでは、人気順で検索結果をランキングしてユーザのレシピ選別を支援している。しかし、ユーザは必ずしも人気があるレシピを求めている場合がある。料理名をクエリとして返される結果のレシピには、利用する食材や、調理手順に様々なバリエーションがあり、それらの特徴を把握してからレシピを選びたい場合も多い。

上記の問題を解決するために、本研究では検索結果として返されるレシピ集合に対して、その全体的なカテゴ

リ構造を可視化する手法を提案する。具体的には、レシピに含まれる材料の出現頻度を調べ、料理をいくつかの部分要素（料理パーツ）に分割する。例えば、オムライスは、「ライス」、「オムレツ」、「ソース」という料理パーツに分解できる。次に材料の出現数から、料理パーツの代表の材料を決め、手順内で代表の材料と他の材料の共起数を調べ、共起の多い材料をパーツに分類させる。分類された材料の共起数から次の代表となる材料を決め、これを繰り返す。最終的に共起した順に材料を図示することで材料のカテゴリ構造の可視化を行う。

2. 関連研究

石橋ら[2]は、食材の利用頻度と典型度から、ユーザの嗜好に合ったレシピ推薦を目的とし、食材の得点を以下の式から求めてレシピを推薦する手法を提案している。

食材 n の得点 =

食材 n の利用頻度 $\times$ 食材 n の典型度

ここで、食材の利用頻度は、ユーザの調理履歴から食材の出現頻度を求め、調理した日が近いほど重みを低くすることで同じ食材ばかりが出てこない工夫をしている。

望月ら[3]は、レシピサイトにおけるユーザ個人の嗜好に合わせた偏りのないレシピ推薦を目的とし、調理履歴から見つけ出したユーザの嗜好食材を用いた推薦、冷蔵庫の余剰食材を考慮した推薦、調理履歴のレシピの調理法を考慮した推薦の3つを提案している。ユーザの調理履歴から嗜好食材を見つけ出す方法については、食材 k の典型度を p^k 、利用頻度を u^k とし、レシピ r_j^k に含まれる食材を I_j^k とするとき、レシピの得点 $S_I(r_j^k)$ を以下の式から求めてレシピの推薦に用いる手法を提案している。

$$S_I(r_j^k) = \frac{\sum_{i \in I_j^k} u^i \cdot p^i}{\sum_{i \in I_j^k} p^i}$$

冷蔵庫の余剰食材を用いた推薦するために、まず、ユーザが冷蔵庫の余っている食材をデータベースに登録する。そして登録されている食材を多く含むレシピの得点を大きくする。また、余剰食材は古いものを優先的に使うことが望ましいとして、登録日を d とし、登録日によって、以下の式で食材に重み付けを行う。

$$w_{(c)} = \frac{(d-1)}{d}$$

調理履歴から求めた調理法の履歴による推薦のためには、まず、レシピを「煮物」「焼物」「炒物」「揚げ物」「蒸物」「和え物」「生煮」「汁物」の8つの調理法に分類した。そして、調理法 c のレシピの得点を、全調理履歴中の c が出現した回数によって求めている。そして、3つの推薦手法の推薦度を最適化することで最終的なレシピの推薦を行う。

福本ら[4]は、食材の重要度と食品群に注目し、レシピ間の類似度を算出する手法を提案した。例えば、「オムライス」という料理において「鶏肉」を材料に含むレシピと、それぞれ「豚肉」と「椎茸」を含む別のレシピの類似度を考える際、従来手法では、「豚肉」と「椎茸」は「鶏肉」と独立した別の食材であるため、類似度は高くない。しかし、この手法では食品群を考慮するため、「鶏肉」と「豚肉」が同じ「肉類」に分類され、「椎茸」を含むレシピに比べて類似度が高くなる。また、「大」「中」「小」の3つの食品群を作り、食品にそれぞれの食品群を対応付け、食材の重要度は料理における出現度だけでなく、その食材が属している食品群についても出現度を算出し合算した。最終的なレシピ間の類似度はコサイン類似度と食品群の重み付け w_n を利用して以下の式によって求める。

レシピ間類似度 $= w_1 \times$ 食材単体のコサイン類似度

$+ w_2 \times$ 食品群 [小] のコサイン類似度

$+ w_3 \times$ 食品群 [中] のコサイン類似度

$+ w_4 \times$ 食品群 [大] のコサイン類似度

$$(w_1 + w_2 + w_3 + w_4 = 1 \quad 0 \leq w_1, w_2, w_3, w_4 \leq 1)$$

これらの研究では、本研究で対象とするような、同一料理の複数のレシピを選別の支援することは考慮されていない。

3. 提案手法

本章では、同一料理として登録された様々なレシピのバリエーションを提示するために、レシピ集号の全体像抽出し、可視化する手法を提案する。具体的には、提案手法は、以下のステップから構成される。

1. レシピの代表的な材料の抽出
2. 料理の構成要素（パーツ）への分割
3. 可視化図の作成

以下に、それぞれのステップについて詳細に説明していく。なお、対象とする料理のレシピ集合は与えられているものとする。レシピは、材料一覧、及び手順から構成されるものとする。

3.1 代表的な材料の抽出

レシピ集合 R における代表的な材料を抽出する。そのためにレシピ集合 R における材料 f の出現頻度を $Freq(f, R)$ と表記する。レシピサイトに含まれる材料の集合を F とするとき、 F に含まれる全ての要素 f に対して $Freq(f, R)$ を求める。 $Freq(f, R)$ が大きい食材は、レシピ集合 R における代表的な材料であると考えられる。

3.2 料理の構成要素（パーツ）への分割

料理の全体像をユーザに提示する際、その料理のレシピ集合における代表的な材料を提示するだけでは、その料理の手順を知らない利用者にとっては、全体像を構築することが困難である。

一般的に、料理は要素から構成される。例えば、オムライスに対しては、「オムレツ」、「ライス」、「ソース」という3種類の要素に分割できる。このとき、料理のバリエーションは、それぞれの構成要素のバリエーションの組み合わせであると考えられる。したがって、レシピを構成要素に分割し、それぞれの構成要素を作成する際に利用する材料を提示することにより、ユーザは、レシピ集合のバリエーションの全体像を効率的に把握できると考えられる。オムライスにおける材料の分割イメージを図2に示す。

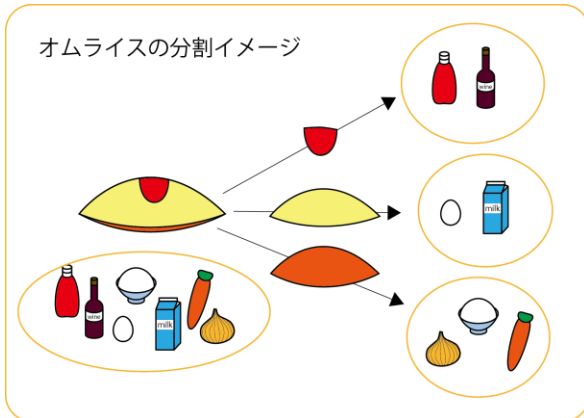


図2：オムライスの材料の分割

次に、構成要素を抽出する手法について述べる。我々は、構成要素は、代表的な材料とそれに付随する材料によって構成されると考えた。レシピは、複数のステップから構成されるが、一つのステップには、一つの構成要素に関する調理方法が記述されている場合が多い。すなわち、一つのステップにおいて、複数の構成要素に関する調理を行うことは稀であると考えられる。本手法では、上記の考えに基づいて、「同じステップに出現する材料は同じ構成要素のために使用される材料である可能性が高い」という仮説を立てた。

上記の考えに基づいて、材料を構成要素に基づいてグループ化するために、代表的な材料と他の材料との共起回数を利用する。材料の共起例を図3に示す。図3より同じ手順にある「玉ねぎ」「ハム」「コーン」は共起していると判断する。また、ステップ2においては、「ご飯」「塩」「コショウ」が共起している。

レシピ集合に含まれる全ての材料の共起を集計し、最も多く共起している代表的な材料が表す構成要素に対して、代表的な材料以外の材料を分類する。例として、オムライスのレシピ集合における、代表的な食材「ご飯」との共起回数の上位を図4に示す。その結果、玉ねぎ、ニンジン、鳥肉の共起回数が多かったため、これらは、「ご飯」の構成要素に所属する材料であるとする。

共起回数が最も多い代表的な材料のパーツにすべての材料を分割した後に、分割された材料の中で共起回数の上位3件を新しい代表的な食材として、それ以外の材料を分割する。これを繰り返すことで関係が強い材料から順に木構造を作ることが可能である。図4にパーツへの分割工程の例を示す。

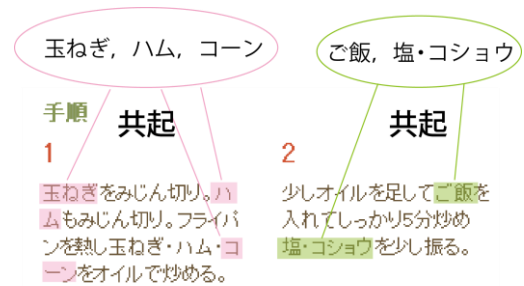


図3：材料の共起

ご飯パーツ		ご飯との共起数	
1回目	玉ねぎ	200	玉ねぎパーツ
	ニンジン	180	
	鳥肉	150	
	ウインナー	120	ニンジンパーツ
	ピーマン	110	
	コーン	80	鳥肉パーツ
		.	.
		.	.
玉ねぎパーツ		玉ねぎとの共起数	
2回目	ウインナー	40	ウインナーパーツ
	ピーマン	25	
	ハム	20	ピーマンパーツ
	コーン	10	
	ほうれん草	8	ハムパーツ
	ネギ	5	
		.	.
		.	.

図4：パーツへの分割の繰り返し

ご飯パーツに分割された材料の内、共起数の上位3件の「玉ねぎ」「ニンジン」「鳥肉」がそれぞれ代表の材料となりパーツ化される。次に玉ねぎパーツに注目すると、同様に「ウインナー」「ピーマン」「ハム」がパーツ化される。

3.3 可視化図の作成手法

上記の処理によって得られた構造に基づいて、可視化図を作成する。図5に可視化図の描き方を示す。具体的には、まず中央に料理名、次の階層に代表となる材料を配置する。その後共起数に基づいてそれぞれの構成要素を生成された順に配置する。この時、階層が一つ変わるに従って材料名の文字サイズを小さくする。こうすることで、その料理にとって重要な材料と、階層構造が認識しやすくなる。また各材料は単純に直線で繋ぐだけにし、無駄な情報を極力入れない構造とする。

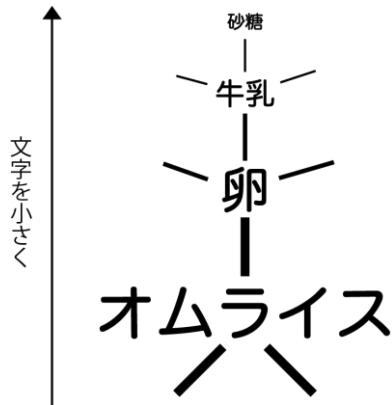


図 5：文字のサイズの変更

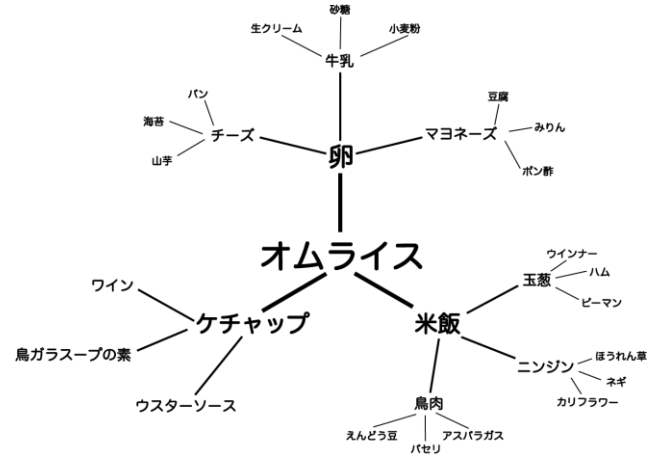


図 6：オムライスの可視化図

4. 実験・評価

提案手法の有効性を評価するために被験者実験を行った。

4.1 対象とするデータ

実験には、楽天レシピのレシピデータを用いた。データは2012年の総レシピ数442,504件までのものについて、以下の項目をデータベースに格納した。

- ・レシピの ID
- ・レシピの料理名
- ・レシピに含まれる材料名
- ・レシピの手順

また、レシピの手順については文章を形態素解析し、食材名のみを抽出した。なお、材料の表記揺れの問題を解決するため、5,513 件の辞書データを用意し、データベースに格納した。

4.2 提案図の作成

楽天レシピにおいて、料理名に「オムライス」を含むレシピ数 862 件に基づいて、提案手法によって作成した可視化図を図 6 に示す。なお、可視化図を作成する際、代表的な材料の抽出と構造の取得はプロトタイプシステムを利用して行い、得られた構造からの可視化図の作成は手動で行った。

4.3 評価実験の内容

可視化図を見る場合と見ない場合でレシピ検索に差が出るかを被験者実験により調査した。被験者は、大学生 5 人である。

被験者には、制限時間 5 分の間に楽天レシピでオムライスのレシピをなるべく多く見て、見た中で一番作りたいと思ったレシピを一つ選択してもらい、次の 3 つの設問に 5 段階評価で回答してもらった。

- ①選んだレシピに対しての確信度
- ②よく使われる材料、あまり使われない材料の把握
- ③バリエーションの把握

4.4 実験結果及び考察

被験者の評価点を平均した結果を図 7 に示す。

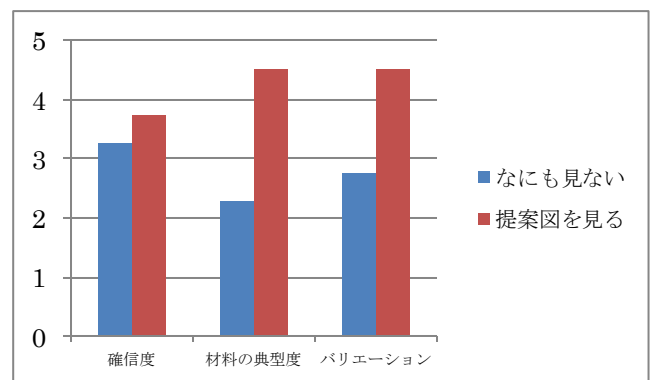


図 7：実験結果

実験結果より、よく使われる材料やあまり使われない材料の把握と料理のバリエーションを把握することに関して提案図が特に有効であることが示された。

確信度については、それほど差がでなかったが、これは、制限時間に設定した 5 分だと、人気順を閲覧しても確信をもてるレシピを容易に見つけられることが原因であると思われる。材料の典型度に関しては典型

度の高い食材は大部分提案図に記載しているので、あまり使われない材料についても把握でき、見ない場合に比べ大きな差が出たと考えられる。バリエーションについては料理名を提案図に記載されている気になる材料によって絞り込み検索を行うことで、新しいバリエーションを発見できたと考えられる。

5. まとめ

レシピに含まれる食材の出現頻度と手順を解析し食材ごとの共起数から、料理を構成要素（パーツ）に分割し、パーツ内でさらにパーツに分割するという作業を繰り返すことで、料理における材料のカテゴリ構造を算出した。またその結果からユーザのレシピ選別を支援する可視化手法を提案した。実験ではレシピを検索する際に、よく使われる材料やあまり使われない材料の把握、レシピのバリエーションの把握という2点について提案図が特に有効であるという結果が得られた。

本研究では、レシピの可視化を行う際に、食材の手順に注目し共起数から食材をパーツに分割した。しかし、現状ではパーツに分割される食材数に偏りができるという問題がある。これは代表的な材料の出現数に差があるためである。この問題を解決するために、食材の分量を考慮するということが考えられる。だが、レシピサイトの分量の表記は、「少々」、「好きなだけ」など、ユーザのさじ加減で決まってしまうため、研究に反映させることが出来なかった。今後は、分量の表記ゆれも計算し、可視化に反映出来るようなシステムの開発なども検討したい。

また今回は提案図の最終的な配置を手動で行ったが自動で行う手法を検討する必要がある。

参考文献

- [1]“楽天レシピ:レシピを見つける, 料理する, 投稿する, みんなで作る”, <http://recipe.rakuten.co.jp/>
- [2] 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. MVE,マルチメディア・仮想環境基礎, vol.107, no.454, pp.51-56(2008)
- [3]望月美里, 高橋裕樹:“余剰食材の有効活用と嗜好に基づく多彩なレシピ推薦”, 映像情報メディア学会技術報告. vol.36(8), pp.127-130, (2012)
- [4] 福本亜紀, 井上悦子, 中川優: “食材の重要度と食品群を考慮したレシピ感類似度の算出手法”, DEIM Forum 2012 D9-2