

# 類似度分析に基づく未知レシピへの メタデータ自動付与方式の実験的評価

高田 夏彦<sup>1,a)</sup> 上田 真由美<sup>2,b)</sup> 森下 幸俊<sup>3,c)</sup> 中島 伸介<sup>1,d)</sup>

概要：近年，料理レシピの検索・推薦に関する取り組みが盛んに行われている．レシピ推薦サービスにおいて，各レシピに対してメタデータを付与することで推薦精度を向上させることが考えられる．しかしながら，全てのレシピに手動でメタデータを付与することは容易ではない．そこで我々は，既にメタデータが付与されている既知のレシピとの類似度分析に基づいた，未知レシピへのメタデータ自動付与方式を実装し，評価実験を行ったので本稿にて報告する．

## 1. はじめに

近年，料理レシピの検索・推薦に関する取り組みが盛んに行われている．クックパッドでは，登録レシピ約186万件以上，月間利用者数2000万人以上(2014年10月)，また，味の素レシピ大百科では，約11,000件以上のレシピが登録されており(2014年10月現在)，料理レシピの検索・推薦サービスの需要が伺える．

レシピ推薦サービスにおいて，高精度に推薦アイテムの選定を行うためには，各レシピデータを分析し，適切なメタデータを付与するというアプローチが考えられる．例えば，“パーティメニュー”や“夜食に最適”，“子供が喜ぶ弁当メニュー”等のメタデータを付与することで，利用者の目的に合致するアイテムの推薦が可能となる．また，森下らはその日の気分あった献立を検索するシステムを開発している[1]．これは利用者が「元気 or お疲れ」，「あっさり or こってり」，「お手軽 or 本格」等の軸に基づいて，その日の自分の気分を入力すると，これに合致した献立を提示するシステムである．

このように推薦対象アイテムに適切なメタデータが付与されていれば，多様な推薦もしくは推薦精度の向上が期待できる．しかしながら，多くの場合，人手によってメタデータが付与されており，大量の推薦対象データに対して適切なメタデータを付与することは容易ではない．そこで

我々は，未知のレシピデータに対して適切なメタデータを自動的に付与することを目指すことにした．既にこれまでの研究において，人手によりメタデータが付与されたマスターレシピを用いて，このマスターレシピとの類似度分析に基づいた，未知レシピへのメタデータ自動付与方式の提案を行っている[2]．そこで，上述のマスターレシピとの類似度分析に基づいた未知レシピへのメタデータ自動付与方式に関して検討し，プロトタイプシステムを構築した．また，実装したプロトタイプシステムを用いた評価実験により，提案手法の有効性を確認すると共に，解決すべき課題について検討を行ったので，本稿にて報告する．

## 2. 関連研究

レシピの類似度分析に関する研究として，苅米ら[3]は料理レシピの柔軟な検索を可能にするための関連検索を提案している．この研究では，レシピに含まれる各材料の分量から算出する「材料の類似」，調理手順から抽出した動詞の並びから算出する「調理手順の類似」，食品群別摂取量に基づいた栄養バランスから算出した「組み合わせの良さ」を関連度として扱っている．しかしながら，料理レシピに記載されているレシピは素材や手順が決まっており柔軟性が十分とはいえない．

白川ら[4]は料理レシピからアニメーションを作成し，ユーザに提示することを提案している．この取り組みでは，料理の基本的な動作に関して，レシピ中の表層表現とアニメーションを対応付ける情報を網羅的に格納した料理動作辞書を作成している．

志土地ら[5]は，同一料理カテゴリ中の料理群における特徴的な調理手順の類似度に基づく代替可能食材の発見手

<sup>1</sup> 京都産業大学コンピュータ理工学部

<sup>2</sup> 流通科学大学総合政策学部

<sup>3</sup> 大日本印刷株式会社 C & I 事業部

<sup>a)</sup> g1144696@cse.kyoto-su.ac.jp

<sup>b)</sup> Mayumi\_Ueda@red.umds.ac.jp

<sup>c)</sup> morishita-y4@mail.dnp.co.jp

<sup>d)</sup> nakajima@cse.kyoto-su.ac.jp

法を提案している。

岩上ら [6] は、ユーザの嗜好を考慮した料理推薦システムの試作を行っている。この取組では、料理間の類似度を算出し、料理間の関連ネットワークを生成する。料理間の類似度の算出には、化合物の類似度算出に用いられる tanimoto 係数を用いている。また、調理手順の類似度に関しては、編集距離を用いて算出している。

橋ら [7] は、料理名に付与されている「簡単」や「ヘルシー」といった修飾表現の根拠をネーミングコンセプトと定義し、その抽出に取り組んでいる。材料と調理器具に着目し、同一料理で用いられる典型的な要素との差異を抽出し、その差異の代替関係を判定することによって実現している。

我々の手法は、食材、調理手順、レシピ名の 3 つの特徴ベクトルを抽出することで、より精密なレシピ間類似度を算出しようとするものであり、その新規性は高いと考えている。

### 3. レシピデータの特徴ベクトルの抽出とレシピ間類似度の算出

本研究では、メタデータ付きレシピデータを用いて、未知のレシピデータへのメタデータ自動付与を可能とする手法の提案を目指している。既にメタデータが付与されたレシピデータ（マスターレシピデータと呼ぶ）に基づいて、これらマスターレシピとの類似性を考慮して、未知レシピデータへのメタデータ付与を試みる。なお、本研究で扱ったマスターレシピ数は 493 であり、既に以下の 6 項目に関するメタデータ（-5 ~ +5 の値で表現）が付与されている。

- ・からだ（お疲れ 元気）
- ・こころ（シクシク ウキウキ）
- ・味（あっさり こってり）
- ・時間（お手軽 本格）
- ・お金（安く 豪華に）
- ・アレンジ（定番 アレンジ）

以下、本節ではレシピデータの特徴ベクトルの抽出とレシピ間類似度の算出方法について説明する。

#### 3.1 レシピデータの特徴ベクトル抽出

本節では、まずレシピデータの特徴ベクトル抽出手法について述べる。レシピデータの特徴量としては、レシピ名、食材、調理手順、栄養素、写真データなどが挙げられるが、全ての特徴量を考慮することは容易でない。したがって、本研究では、レシピ名、食材、調理手順を考慮した特徴ベクトル抽出に取り組む。なお、特徴ベクトルは各々に対して抽出する。すなわち、レシピ名特徴ベクトル、食材特徴ベクトル、調理手順特徴ベクトルである。また、マスターレシピとの類似度算出を目的としているので、特徴ベクトル

ル次元はマスターレシピのベクトル次元を採用する。各特徴ベクトルの抽出方法順を以下に示す。

##### 3.1.1 レシピ名特徴ベクトル

マスターレシピ群のレシピ名に対して形態素解析（MeCab[8] を使用）を行い、レシピ推薦に不必要な単語をストップワード（以下、ストップワードと呼ぶ）として取り除いた後の全ての語をレシピ名特徴ベクトルの次元として採用する。新規レシピデータに関しては、同等にレシピ名に対して形態素解析を行い、出現回数に基づいて各次元の値を決定する。

##### 3.1.2 食材特徴ベクトル

マスターレシピデータに対して形態素解析を行い、ストップワードを取り除いた後、少なくとも複数回出現するものを食材特徴ベクトルの次元として採用する。まだベクトルが作成されていない未知レシピのデータに関しては形態素解析を行い、同等の次元に関してベクトルを作成する。

##### 3.1.3 調理手順特徴ベクトル

マスターレシピデータに対して形態素解析を行い、ストップワードを取り除いた後の動詞及び名詞を調理手順特徴ベクトルの次元として採用する。新規レシピに関しては形態素解析を行い、同等の次元に関してベクトルを作成する。この際、「水をきる」などのようなレシピ特有の表現については形態素辞書をカスタマイズすることで効率的に抽出できるように工夫する。未知レシピに関しては、同様に手順に対して形態素解析を行い、対象単語の出現回数に基づいて各次元の値を決定する。以下に実際の例に基づいて、説明する。

対象となるレシピデータに対して MeCab を用いて形態素解析し、名詞、動詞、形容詞を抽出する。なお、事前にレシピの特徴を表現するのに不要な語句をストップワードとしてリストアップしておき、これらを取り除くことで、できる限りレシピの特徴を表現するのに適切な語句に限定したベクトル次元を決定する。以下に、レシピ『若竹汁』を例に説明する。

=====

レシピ名：若竹汁

調理手順：

筍は薄切りにし、わかめは戻しておく。鍋に出し汁と筍を入れ、一煮立ちさせる。わかめを加え（A）で調味する。汁椀に注ぎ、吸い口に木の芽を浮かべる。

=====

これに対し形態素解析を行い、ストップワードを取り除くと表 1 の結果を得ることが出来る。

この処理をマスターレシピデータ全てに実施した上でとりまとめて、重複を整理することで、調理手順特徴ベクトルの次元を決定することができる（以下、マスターベクトル次元と呼ぶ）。各マスターレシピの特徴ベクトルは、マスターベクトル次元のうち、当該レシピに存在する語句に

表 1 若竹汁のレシピから形態素を抽出したもの

口	薄切りにする	入れる	汁
浮かべる	木の芽	筍	椀
加える	鍋	調味する	出し汁
吸う	注ぐ	おく	させる
わかめ	煮立つ		

関しては数値 1, 存在しない語句に関しては数値 0 とすることで得られる。なお, この中に「戻す」というワードがないのは以下で示す<ストップワードリスト>にて, 戻すというワードを弾いているからである。

次に未知レシピの特徴ベクトル抽出方法について述べる。最終的にはマスターレシピのメタデータを考慮した, 未知レシピへのメタデータ付与を目指しているので, 未知レシピの特徴ベクトル次元は, マスターレシピのベクトル次元をそのまま継承する。手順としては, 未知レシピについても形態素解析を行い, マスターベクトル次元のうち, 当該レシピに存在する語句に関しては数値 1, 存在しない語句に関しては数値 0 とすることで得られる。すなわち, マスターレシピに存在しない語句に関しては, 未知レシピに存在したとしても特徴ベクトルには採用されない。以下に, 『豚汁』を例に説明する。

=====

レシピ名: 豚汁

調理手順:

大根, 人参はいちょう切りにし, ごぼうはささがきにする。冷凍里芋は流水で半解凍し, 半分に切る。こんにゃくは下茹でし, 短冊切りにする。ねぎは小口切りにする。鍋に水とねぎ以外の材料をすべて入れ, やわらかくなるまで煮る。具材がやわらかくなったら, 味噌を加えてさっと煮る。器に盛り, ねぎを散らす。

=====

『豚汁』の未知レシピに対し形態素解析を行い, ストップワードを取り除いたものを表 2 に示す。

表 2 未知レシピ『豚汁』から形態素を抽出したもの

半分	解凍	小口切りにする
以外	里芋	ごぼう
やわらかい	水	ねぎ
短冊切り	冷凍	流水
散らす	入れる	煮る
すべて	材	味噌
加える	ささがきにする	豚汁
こんにゃく	大根	する
いちょう切りにする	具	茹でる
人参	器	切る
鍋	材料	盛る

これが未知レシピ『豚汁』の特徴を表現する語句集合であるが, 特徴ベクトルの次元としては前述のマスターベ

クトル次元を採用するため, 表 2 で示した語句のうち, マスターベクトル次元に存在する語句のみ, 未知レシピ『豚汁』の特徴ベクトルとして考慮することになる。

ここで, レシピ特徴ベクトル次元としては採用しない語句リストであるストップワードリストを以下に示す。

~~~~~<ストップワードリスト>~~~~~  
か所, 供する, 人, 針, れる, きする, 前, れん, がく, く, のく, せる, ねぐ, は種, じす, まあ, スー, cm, よい, よう, ほう, 部, うち, さとい, なる, こ, の, め, とも, とわ, もの, す, げ, さ,, ぼる, いぬ, そう, つき, つま, きん, えむ, つる, ちる, ら, い, ん, てる, つけ, とさつ, ,, \*, 普通, か所, やすい, パッド, 二, 人, 縦, 度, からず, スー, もともの, 休む, にくい, その他, 保存, つる, 翌日, せる, 刷毛, 空気, ない, できる, ところ, いる, 向う, 止める, おきる, プロセッサー, 左, 時, ひく, ふう, 他, れる, 画びょう, 使用, とめる, 位, 1 つ, 同様, 安定, く, 半量, 中央, 的, 直径, cm, よい, よう, ほう, 七, うち, 以上, これ, さ, ケ所, 下, 上, 三, ビニール, くる, 一, 適当, 台, 取りはずす, 座る, 市販, 立つ, 立て, てる, 前方, ねる, たんぱ, 終わる, いかめしい, 残り, 残る, 操作, ころがす, 見る, 分, 各種, 浮き上がる, の, 順番, 時間, 管理, 各々, 全て, スナップ, ごと, 手早い, 大, 入手, 角, えむ, 戻す, 7, 外, えらい, はう, かぶる, 次, 倍, 化粧, 日, それぞれ, 秒, 幅, ある, 長い, 0, 1, 2, 3, 4, 5, 8, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, サイズ, 放す, 四隅, とも, 耐熱, 一緒, 用

このストップワードリストを充実させることにより, データの特徴表現および類似度計算をより正確に行うことができる。

表 3 に, マスターレシピ次元および未知レシピ『豚汁』に含まれる語句から作成される『豚汁』の特徴ベクトル次元の一部を示す。レシピ間類似度を計算するため, ベクトル次元はマスターレシピ次元を使うことで統一されている。したがって, 表 3 に示すとおり, 「小口切りにする」や「こんにゃく」のように, マスターレシピ次元に存在し, レシピ内にも記述されている語句は, 数値 1 を与える。表 3 内で数値が 0 である語句は, マスターレシピ次元には含まれているが, 『豚汁』レシピには含まれない語句である。

以上の通り, マスターベクトル次元の抽出を行い, 未知レシピの特徴ベクトル抽出を行う。なお, 我々が用意したマスターレシピ 493 件に対して, マスターベクトル次元の抽出を行ったところ, 次元数は 1185 次元となった。

### 3.2 レシピ間類似度算出方式

本節では, 抽出された特徴ベクトルを用いた, レシピデータ間の類似度計算方法について述べる。図 1 に, レシピ間類似度計算のイメージを示す。

各レシピは, レシピ名特徴ベクトル, 食材特徴ベクト

表 3 マスターレシピ次元を用いた『豚汁』レシピ特徴ベクトルの一部

|         |   |           |   |
|---------|---|-----------|---|
| 小口切りにする | 1 | 材料        | 1 |
| ほぐす     | 0 | こんにゃく     | 1 |
| すべて     | 1 | 以外        | 1 |
| 大根      | 1 | 器         | 1 |
| お湯      | 0 | 切る        | 1 |
| 味噌      | 1 | いちよう切りにする | 1 |
| 茹でる     | 1 | おく        | 0 |
| 散らす     | 1 | 人参        | 1 |
| やわらかい   | 1 | 加える       | 1 |
| 石づき     | 0 | ポウル       | 0 |
| 水       | 1 | 一口        | 0 |
| 短冊切り    | 1 | 酒粕        | 0 |
| 入れる     | 1 | ねぎ        | 1 |
| 鮭       | 0 | する        | 1 |
| 小房に分ける  | 0 | 盛る        | 1 |
| 注ぐ      | 0 | 鍋         | 0 |
| 溶く      | 0 | しめじ       | 0 |
| 取る      | 0 | /         | / |

て特徴ベクトルを抽出し、個々の類似度を算出可能にすることで、種々の目的に応じた類似度計算が可能になる。例えば、「食材は類似しているが、調理手順は似ていないレシピの推薦」などが実現可能となる。

ここで、実際に未知のレシピと 493 件のマスターレシピとの類似度計算を行い、最も類似度の高いマスターレシピ上位 10 件を示す。表 4 には、未知レシピ「肉じゃが」との類似レシピ、表 5 には、未知レシピ「カレーライス」との類似レシピ、表 6 には、未知レシピ「粥」との類似レシピ、を示す。

表 4 「肉じゃが」と似ているレシピ上位 10 件

| レシピ名           | 類似度   |
|----------------|-------|
| おからの煮物         | 0.420 |
| かぼちゃの含め煮       | 0.389 |
| ビーフカレー         | 0.369 |
| 里芋の煮っころがし      | 0.363 |
| クリームシチュー       | 0.356 |
| チンジャオロース       | 0.348 |
| さやいんげんと豚バラ肉の煮物 | 0.348 |
| 味噌煮込みうどん       | 0.342 |
| がんもどきと小松菜の含め煮  | 0.338 |
| 牛肉の柳川風         | 0.335 |

表 5 「カレーライス」と似ているレシピ上位 10 件

| レシピ名         | 類似度   |
|--------------|-------|
| 肉じゃが         | 0.445 |
| ビーフカレー       | 0.400 |
| おからの煮物       | 0.388 |
| ポトフ          | 0.380 |
| クラムチャウダー     | 0.375 |
| 金時豆煮         | 0.350 |
| 鶏肉と蕪のクリーム煮   | 0.349 |
| 味噌煮込みうどん     | 0.342 |
| 白菜と豚バラ肉の重ね蒸し | 0.342 |
| オニオンスープ      | 0.327 |

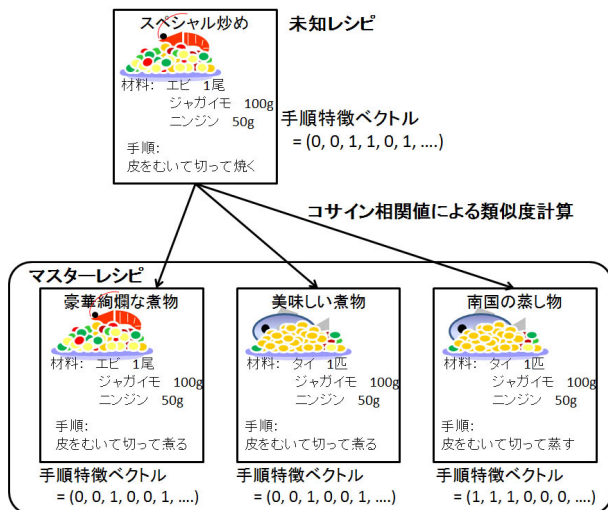


図 1 レシピ間類似度計算

ル、調理手順特徴ベクトルを有する。すなわち、レシピ間の類似度判定としては、レシピ名類似度、食材類似度、調理手順類似度を個別に計算し、以下の式により統合することになっている。

$$Rsim = \alpha \cdot Tsim + \beta \cdot Isim + \gamma \cdot Psim \quad (1)$$

ただし、 $Rsim$  はレシピ間類似度、 $Tsim$  はレシピ名類似度、 $Isim$  は食材類似度、 $Psim$  は調理手順類似度である。それぞれ、3.1.1 節のレシピ名特徴ベクトル、3.1.2 節の食材特徴ベクトル、3.1.3 節の調理手順特徴ベクトルを用いて算出されるものである。また、 $\alpha, \beta, \gamma$  は重みであり、 $\alpha + \beta + \gamma = 1$  を満たす。なお、特徴ベクトル間の類似度計算には、コサイン類似度を用いる。

このように、1つのレシピを1つの特徴ベクトルで表すのではなく、レシピ名、食材、調理手順のそれぞれに対し

表 6 「粥」と似ているレシピ上位 10 件

| レシピ名    | 類似度   |
|---------|-------|
| 中華粥     | 0.764 |
| 牛肉の柳川風  | 0.420 |
| うなぎの柳川風 | 0.420 |
| いか飯     | 0.414 |
| 赤飯      | 0.395 |
| 黒めめ     | 0.393 |
| キムチ鍋    | 0.378 |
| ひじきの煮物  | 0.364 |
| 肉ご飯     | 0.364 |
| 卵雑炊     | 0.364 |

類似度算出結果より、肉じゃがと似ているレシピについては調理方法が類似しているものが多く、カレーライスと似ているものは食材が類似しており、粥はレシピ名や調理方法が類似しているものが上位となった。それぞれの結果について、ある程度レシピの類似性が考慮できていると

思えるが、ランキング最上位のレシピが最も類似しているとは言い難い状況である。したがって、まだリストアップできていないストップワードがノイズになっていることを含めて、特徴ベクトルの作成方法に改良の余地があると考えている。今後はストップワードリストの充実を含め、類似度算出方式の改良に取り組む。

## 4. 未知レシピへのメタデータ自動付与方式

### 4.1 類似度を考慮したメタデータ自動付与方式の検討

本節では、類似度を考慮した未知レシピへのメタデータ自動付与方式について説明する。本研究では、メタデータ付与済みの493件のマスターレシピを用いて、未知レシピへのメタデータ自動付与を目指す。類似するマスターレシピの気分データに基づいた、未知レシピへの気分データ付与方式について、現状では以下の3つの方法を検討している。

(1) 類似度が最も高いマスターレシピのメタデータを、該当する未知レシピへそのままコピーする。

$$M_u(i) = M_1(i)$$

ただし、

$M_u(i)$  : 未知レシピの各気分データ

$M_1(i)$  : 類似度ランク1位のレシピの各気分データ

$i = \{ \text{からだ, ところ, 味, 時間, お金, アレンジ} \}$

(2) 類似度が上位数件(例えば5件)のレシピ集合のメタデータの平均を、該当する未知レシピへコピーする。

$$M_u(i) = \frac{M_1(i) + M_2(i) + M_3(i) + M_4(i) + M_5(i)}{5}$$

ただし、

$M_n(i)$  : 対象未知レシピとの類似度順位が  $n$  番のマスターレシピの各気分データ

(3) 類似度が上位数件(例えば5件)のレシピ集合のメタデータに対して、類似度順位の重みを考慮したメタデータ平均値を、該当する未知レシピへコピーする。

$$M_u(i) = \frac{S_1 \cdot M_1(i) + S_2 \cdot M_2(i) + S_3 \cdot M_3(i) + S_4 \cdot M_4(i) + S_5 \cdot M_5(i)}{S_1 + S_2 + S_3 + S_4 + S_5}$$

ただし、

$S_m$  : 対象未知レシピと類似度順位が  $m$  番のマスターレシピとの類似度

上記を踏まえて、メタデータの自動付与に関する実験を行い、その妥当性について検証すると共に、類似度算出方法も含めて、未知レシピへのメタデータ自動付与を精度良く行う手法の開発に取り組む。

## 4.2 未知レシピへのメタデータ自動付与方式の実装

本節では、類似度を考慮した未知レシピへのメタデータ自動付与方式を実装したソフトウェアに関して説明する。実装したソフトウェアは、形態素解析手法として MeCab[8]を採用し、Pythonにより実装した。図2に、実装したソフトウェアのユーザインタフェースを示す。

本ソフトウェアは、事前にメタデータが付与されたマスターレシピ、ストップワードリスト、および未知レシピを入力すると、入力された未知レシピに対して自動でメタデータを付与するものである。メタデータの自動付与方式としては、4.2節にて説明した(1)~(3)の方法を同時に実行することが可能である。

## 5. メタデータ自動付与方式の実験的評価

本研究で提案し実装した、未知レシピに対するメタデータ自動付与方式に対する評価実験を行ったので説明する。

### 5.1 レシピ間類似度算出方法の妥当性評価

提案手法では、未知のレシピに対して類似するマスターレシピを正しく判定する必要がある。したがって、ここでは提案手法によるレシピ間類似度の算出方式が適切に行われているかを評価する。

実験手順としては、マスターレシピには含まれない未知のレシピを10件用意し、各レシピに対して類似度の高いレシピ上位10件を提案手法により判定した。該当する未知レシピと共に類似度上位のレシピ10件を被験者に提示し、提示された10件のレシピそれぞれが、実際に類似しているといえるかどうかを判定してもらった。判定は{似ている(3点)、やや似ている(2点)、やや似ていない(1点)、似ていない(0点)}の4択とした。図3に、未知レシピとの類似度上位10件のレシピに対する被験者8名の評価結果を示す。

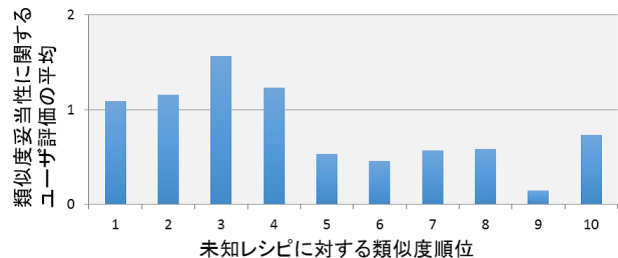


図3 類似度上位10件のレシピに対するユーザ評価結果

すなわち、提示された上位10件のレシピに対するユーザ評価(似ている or 似ていない)のスコア(0~3点)を、類似度順位毎に平均したものである。提案手法の類似度判定が適切に行われていれば、類似度判定上位のレシピの方が下位のレシピに比べて評価値が高いはずである。図3より、上位のレシピ(特に1~4位)の評価が相対的に高く、

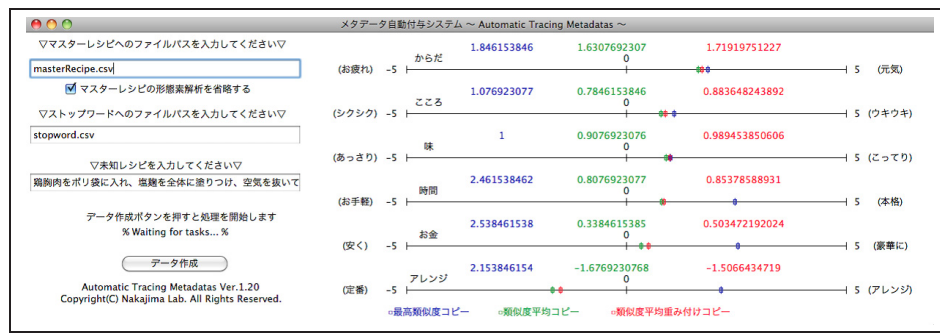


図 2 実装したメタデータ自動付与ソフトウェアの UI

5 位以下のレシピの評価は相対的に低いことが確認できる。評価値の値が必ずしも類似度順位通りになっていないことに関しては、現状の被験者数が 8 名と十分ではないことが大きな原因ではないかと考えている。現在も実験を継続しており、データ数を増やしながら引き続き評価を行うつもりである。また、評価値そのものが全体的に低めであることについては、そもそも 500 件弱のマスターレシピの中から類似レシピを判別しているため、未知レシピに対して類似したレシピがそれほど多く存在しないということも原因の一つである。未知データに対するメタデータの自動付与を行うにあたって、マスターデータ量がどの程度必要なのかということについても検討したいと考えている。

### 5.2 未知レシピに対するメタデータ付与の妥当性評価

提案手法により、未知レシピに対して付与されるメタデータは、当然ながら適切である必要がある。したがって、未知レシピに対して付与されたメタデータの妥当性評価を行う。

実験手順としては、5.1 節と同様にマスターレシピには含まれない未知のレシピを 10 件用意し、各メタデータ付与方式により付与されたメタデータを被験者に提示する。被験者には、付与されたメタデータが妥当かどうかを判定してもらう。判定は { 妥当である (3 点), やや妥当である (2 点), やや妥当でない (1 点), 妥当でない (0 点) } の 4 択とした。なお、メタデータの付与方式としては、

- 4.1 節 (1) の方式 (1 件コピー),
- 4.1 節 (2) で上位 3 件による方式 (1 ~ 3 件平均)
- 4.1 節 (3) で上位 3 件による方式 (1 ~ 3 件重み平均)
- 4.1 節 (2) で上位 5 件の方式 (1 ~ 5 件平均)
- 4.1 節 (3) で上位 5 件の方式 (1 ~ 5 件重み平均)

の 5 つの方式である。図 4 に、付与されたメタデータの妥当性に関するユーザ評価結果 (6 軸) を、図 5 に、6 軸全てを平均した結果を示す。

図 4 より、各 6 軸の評価で若干のバラつきはあるものの、(1 ~ 5 件平均) もしくは (1 ~ 5 件重み平均) の評価が相対的に高いことが分かる。図 5 から明らかであるが、やはり未知レシピに類似したマスターレシピを判定する際

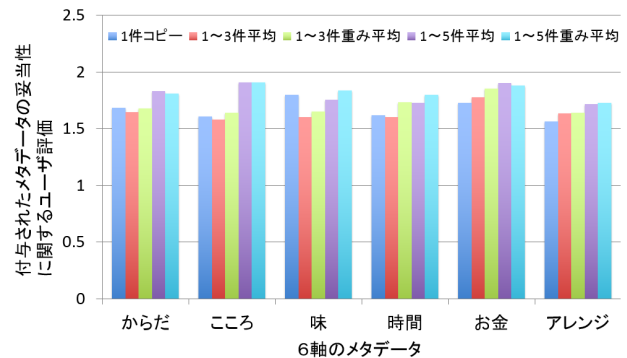


図 4 付与されたメタデータの妥当性に関するユーザ評価結果 (6 軸)

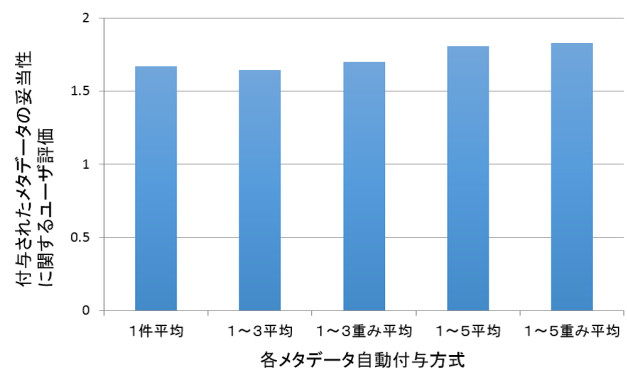


図 5 付与されたメタデータの妥当性に関するユーザ評価結果 (平均)

に、非常に高い類似度のマスターレシピが存在する訳ではないため、1 件もしくは極少数のレシピのメタデータから未知レシピに適切なメタデータを推定することは困難であるということがいえる。また、類似度上位のマスターレシピのメタデータを平均する場合には、その類似度に応じて平均する方が、若干ではあるが比較的高い評価となった。今回の評価実験では (1 ~ 5 件重み平均) の評価が最も高い結果となったが、5 件が最適なのかということに関しては今後さらに検討する必要がある。

### 5.3 各メタデータ自動付与方式の比較評価

各メタデータ自動付与方式のうち、どの方式が最もユーザの印象がよかったのかを調べるために、各方式に対する比較評価を行った。

実験手順としては、5.2節と同様にマスターレシピには含まれない未知のレシピを10件用意し、各メタデータ付与方式により付与されたメタデータを被験者に提示する。被験者には、5.2節でも説明した5つのメタデータ付与方法に対して、適切と思われる順序を付けてもらった。

表7に、被験者による比較評価結果（順位の平均）を示す。

表7 各メタデータ自動付与方式の比較評価結果（順位の平均）

| 1件<br>コピー | 1～3件<br>平均 | 1～3件<br>重み平均 | 1～5件<br>平均 | 1～5件<br>重み平均 |
|-----------|------------|--------------|------------|--------------|
| 3.16      | 3.68       | 3.35         | 2.51       | 2.30         |

当然ともいえるが、被験者による評価順位についても（1～5件平均）および（1～5件重み平均）の評価が高く、特に（1～5件重み平均）の評価が最も良いという結果となった。

以上より、提案手法である類似度分析に基づく未知レシピへのメタデータ自動付与方式の特性について、幾つかの知見を得ることができた。

なお、今後解決すべき課題としては、以下のようなものが挙げられる。

- レシピ特徴ベクトル抽出方式の改良。
- ユーザの評価に合致する類似度判定方法の検討。
- マスターレシピデータの量および質の向上。
- メタデータの各軸に特化した自動付与方式の検討。

継続して実施している評価実験結果を踏まえて、上記課題に取り組みつつ、最終的には実用化に耐えうる高精度なシステムの構築を目指す。

## 6. おわりに

本稿では、既にメタデータが付与されている既知のレシピとの類似度分析に基づいた、未知レシピへのメタデータ自動付与方式を実装し、評価実験を行った結果について報告した。これにより、提案手法である類似度分析に基づく未知レシピへのメタデータ自動付与方式の特性について、幾つかの知見を得ることができた。

今後実用化に耐えうるシステムを構築するためにも、評価実験にて明らかになった検討課題を克服しつつ、提案手法の改良に取り組むつもりである。

## 謝辞

本研究の一部は、文部科学省科学研究費助成事業（学術研究助成基金助成金）基盤研究（C）（課題番号：#26330351）による。ここに記して謝意を表します。

## 参考文献

- [1] 森下幸俊, 中村富予, 気分により献立検索システムの検索軸の評価とレシピを活用した食品販売機能の市場ニーズの評価, 電子情報通信学会技術研究報告 Vol.112, No.75, DE2012-14, pp.79-84, 2012.
- [2] 高田夏彦, 佐々江駿, 上田真由美, 中島伸介, 森下幸俊, 類似度分析に基づく未知レシピへのメタデータ自動付与方式の提案, 第6回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2014) P2-3, 2014年3月.
- [3] 苅米志帆乃, 藤井敦, 料理どうしの類似と組合せに基づく関連レシピ検索システム, 言語処理学会第14回年次大会発表論文集, pp.959-962, 2008.
- [4] 白井清昭, 大川寛志, アニメーション生成のための料理動作辞書の構築, 情処研報 自然言語処理研究会報告, Vol.2004 No.108, pp.123-128, 2004.
- [5] 志土地由香, 井手一郎, 高橋友也, 村瀬洋, 料理レシピマイニングにいる代替可能食材の発見, 電子情報通信学会論文誌 (料理を取り巻く情報メディア技術論文特集 Vol. J94-A No.7 pp.532-535, 2011)
- [6] 岩上将史, 伊藤孝行, ユーザの嗜好順序を利用した料理推薦システムの試作, 第24回人工知能学会全国大会. 1D2-4, 2010.
- [7] 橋 明穂, 若宮翔子, 難波英嗣, 角谷和俊, 料理名の修飾表現の関係性に基づくレシピのネーミングコンセプト抽出, 信学技報, vol. 113, no. 214, DE2013-36, pp. 19-24, 2013年9月.
- [8] MeCab (形態素解析エンジン):  
<http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>