

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

執行 健人[†] 清光 英成[‡] 大月 一弘[‡]

[†] 神戸大学国際文化学部 〒657-0011 兵庫県神戸市灘区鶴甲 1-2-1

[‡] 神戸大学国際文化学研究科 〒657-0011 兵庫県神戸市灘区鶴甲 1-2-1

E-mail: [†] 1416543c@stu.kobe-u.ac.jp, [‡] {kiyomitu, ohtsuki}@kobe-u.ac.jp

あらまし レシピの検索結果からレシピを選ぶ際の意味決定を支援するシステムの開発を目的とし、大量のレビューからテクスチャー表現を抽出・集計・可視化する。レビュー数が 1000 件を超えるレシピは少なくないが、本研究の成果により利用者は多数のレビューを読まなくても、他の利用者が調理した出来上がりを一目でイメージすることができる。レシピを参考に利用者自身が調理してどうなりそうか想像でき、レシピ選びの一助となる。

キーワード 食感, テクスチャ, レシピ検索, 集合知, データベース

1. はじめに

本研究では、Web 上のレシピを参考に「自分が調理すればどうなりそうか」知りたいという利用者の要求に応え、レシピ選択という行動の意味決定を支援することを目的とする。そのために各レシピの大量のレビューから食感語を抽出・集約し、他の利用者が調理した結果を提示・比較するシステムの開発を行った。これにより利用者のレシピ選択という行動の意味決定を支援することができる。利用者は自分で調理するレシピを選ぶため、検索結果からレシピを決定する際に「自分が作ったらどうなりそうか」知りたいと考えることは自然である。レビュワーが調理してどうなったかが分かるようにレビューを集約する。集約した結果を可視化すれば、利用者は多数のレビューを読まなくとも自分で調理すればどのようになりそうか一目で想像することができる。可視化を用いてレシピを比較することが容易になる。検索時には思いつかなかった食感語が可視化に現れていれば偶然の発見となり、利用者の潜在的な検索要求に気づかせる一助となる。

本文の構成は以下のとおりである。2 節で関連研究を紹介し、本研究の位置づけを明らかにする。3 節ではレシピの選定、4 節でレシピ可視化システムについて論じる。5 節はまとめである。

2. 関連研究

大橋ら[1]は「おいしいを感じる言葉」としてシズルワードを定義し、305 のシズルワードからどの語に「おいしい」を感じるか、また 92 のシズルワードからどんな食べ物や飲み物を想起するか、一般の消費者に対してインターネット調査を行った。調査により、シズルワードを「甘い」「辛い」などの味覚系、「もちもち」「サクサク」などの食感系、「季節限定」「朝採り」などの情報系に分類した。また性別や年齢、年代などによって、シズルワードの利用に関する特徴的な

違いを分析した。これらのシズルワードは「おいしい」を感じ、飲食の体験を共有する言葉である。

笹田ら[2]は、自然言語処理において、単語認識（形態素解析や品詞推定など）の次に解決すべき課題は、用語の抽出であると主張し、レシピ中に出現する重要な固有表現を「レシピ用語」と定義し、実際にコーパスに対してアノテーションすることで、実用的な精度の自動認識器の構築を行っている。レシピを利用するユーザーの嗜好や気分、調理履歴をもとに、レシピの推薦・評価を試みた研究として、森下ら [3] は、生活者の気分に合わせて献立を提案し食材決定を支援するシステムを開発し、「時間」「味」といった 6 つの気分検索軸の重要度を評価した。献立決定時の気分を評価軸として定めることで、レシピを活用した食品販売機能の市場ニーズの評価を可能とした。上田ら[4] は、レシピ利用者の過去の閲覧履歴や調理履歴をもとに、対象者の食材に対する好みを推定し、そこにレシピに使われる食材の量も加味した献立推薦手法を提案している。

Lertsumruaypun ら[5]は、日本では料理や味覚を擬音語、擬態語を表すオノマトペを用いて表現することが多いとし、オノマトペを利用した料理レシピ推薦システムを開発した。web 上に掲載されているレシピ文章を収集し、レシピ内の文章を解析することで、オノマトペと料理名・食材などの固有名詞、形容詞、一般名詞、動詞の関連性を数値化した。Lertsumruaypun らは、この数値をレシピに含まれる語とオノマトペとの関連度として算出し、キーワードサーチと比べて精度の高いレシピ検索を実現した。

本研究の目的に適合する語彙は、シズルワードのようなおいしさを表現する語だけでは不十分である。おいしくないことを表現する語も必要である。そこで農業・食品産業技術総合研究機構 食品研究部門の調査で公開されている食感を表す語彙を利用する。消費者のテクスチャー表現における語彙を日本語テクスチ

ャー445語としてまとめている [7].

本稿で用いるテクスチャー語彙は、日本語テクスチャー語を、結果に大きな影響を与えると予想される表記揺れを考慮して調整した 526 語を使用する。

3. レシピの選定

現在、レシピを検索するツールとしてレシピサイトが多くの人に利用されている。レシピサイトによって形態やジャンル、テーマは多様である。料理の専門家や一般の利用者がレシピを提供するもの、和食や中華のレシピなど料理のジャンルに特化したもの、健康に配慮したレシピを提供するものなどがある。

レシピ検索において、利用者はキーワードやレシピのカテゴリを指定することで検索意図を表明することができる。キーワードとして料理名や食材名、あるいは運動会やお弁当など、食事の目的やシーンなどを指定することができる。カテゴリとしては食材やご飯もの、汁物といった料理のジャンルなどが指定可能だ。指定したキーワード・カテゴリで検索すると、それにマッチするレシピが返される。キーワードが 2,3 個であれば、返される検索結果候補は数千から数万件である。得られた大量の検索結果候補はアクセス数やレビュー数など妥当な評価軸で評価され、整列される。評価の高い順に 1 ページ当たり 10 件程度のレシピを提示することで、利用者はどれを選ぶか吟味するのに現実的な数のレシピを閲覧することができる。これにより、利用者は数万から数百万件のレシピの中から、表明した検索意図をよく満たすレシピを現実的な数まで絞り込むことができる。このようなレシピ検索は非常に有用なものであると考えられる。

検索結果候補から実際に調理するレシピを選ぶための手掛かりは、レシピの画像、調理の目安時間、レシピのタイトルや説明文中の単語やフレーズなどである。ただ、画像はレシピ投稿者が調理したものを撮影し、投稿したものである。料理の腕前は人によって異なるので、料理を作る人によって調理結果は当然異なる。そのため同じレシピを参考にしても、利用者は投稿者と同じように、つまり画像のように調理できるとは限らない。調理にかかる時間もまた個人差があるので、投稿者が提示する目安時間はあくまでも目安である。説明文に「失敗しない」などと書かれていても、それは投稿者が調理に失敗しなかったことを示すだけなので、利用者が失敗せずに調理できるとは限らない。利用者は「自分が作ったらどうなりそうか」を知りたいのである。これを知るためにはレビューを閲覧すればよい。レシピのレビューとはそのレシピを参考に調理した体験や結果を投稿者にフィードバック、あるい

は他の利用者と共有するものである。レビューの調理前と利用者の現在が、レシピ選びに関して同じ立場といえる。レビューを一件ずつ読むことは、そのレビューの調理体験を知ることである。利用者自身が調理するとどの様になるかは、レビューの調理結果を集約することで模倣する。つまり、「自分が作ったらどうなりそうか」イメージできるようにする。実際には、1000 件以上のレビューがつくレシピも少なくない。また、大量のレビューを一件ずつ読んでいくのは現実的ではない。つまり、レビューを集約する必要がある。

本研究では、レシピのレビューに書かれた食感語を抽出・集約し、その結果を可視化することで、レビューが調理した結果を一目で直感的に把握できるシステムを開発した。これにより、利用者はレシピ選びの際に、自分が調理すればどうなりそうかイメージすることが可能となる。これにより、本システムは利用者のレシピ選択を支援する一助となる。

現在の主なレシピサイトにおいて、レビューの調理体験はテキスト形式のレビューに書かれている。調理してどうなったかという観点でこれらを集約するためには、それが分かるような言葉で集約する必要がある。一般的に、レシピのレビューはレシピを参考にして調理がうまくできた時や、レシピ投稿者に感謝を伝えるために投稿される。そのためレビューにはプラス極性の言葉が書かれることが多い。調理結果を表すプラス極性の言葉として、まず「美味しくできた」や「うまかったです」などが考えられる。しかしそれらは抽象的な表現であり、具体的に調理してどうなったかが分からない。そのためそれらの語でレビューを集約した結果を見ても、自分が調理してどうなりそうかイメージできない。「ありがとう」といった感謝の言葉はどのレシピのレビューにも多く見られる。そのため感謝の言葉はレシピの特徴とならず、その集約結果はレシピ選びの参考にはならない。そもそも調理結果を表す言葉ではない。

レビューに書かれる語として、食感語がある。食感語は感性に基づく言葉である。そのため食感語であれば調理してどうなったかを一言で直感的に伝えられる。また食感の程度を評価する際に個人差はあるが、共通した認知を同じ言葉で表現できる。その上レビューが調理してどうなったかを表現する語であると考えられる。これらを踏まえ、本研究では食感語を利用し、レビューを集約した。食感語として農業・食品産業技術総合研究機構 食品研究部門の調査で収集された 445 語の食感語をもとに、表記ゆれに対応した 526 語の食感語を利用した。レビュー中の食感語は形態素解析器を用いて抽出した。あるレシピのレビューに食感語が書かれていれば、食感語を書いたレビュー者に同

様の認知がされたと考えられる。そのようなレビュー数が多ければ、利用者の調理においてもそのような出来上がる可能性が高いと考えられる。そこでレビュー中の食感語が何人によって書かれたかを調べ、レビューを集約した。

集約結果を Word Cloud として可視化した。Word Cloud とは文章中の単語の出現頻度に応じた大きさで単語を図示する可視化手法である。そのため単語の出現頻度が高いほど文字サイズが大きく、低いほど小さくなる。本研究ではレビュー中の食感語を、それを書いたレビュー数に応じて大きさをキャンパスに図示した。そのように Word Cloud として可視化すると、多くの人によって書かれた食感語ほど文字サイズが大きく表示される。他の食感語の大きさと比べることで、食感語を書いたレビュー数にどの程度の違いがあるか、実際の数値を見比べるよりもすぐに分かる。「自分が作ったらどうなりそうか」が一目で直観的に把握できる可視化手法であり、本研究のレビュー集約・可視化の目的に適っているとと言える。またレシピに対して順位や格付けがされず、参加の動機であるレシピ投稿者の貢献欲を損なうことがない。次元数の多少も問題にならない。可視化に検索意図として表明した以外の食感語が現れていることを確認でき、利用者に潜在的な検索意図を気づかせる一助にもなる。

4. レシピ可視化システム

本研究ではレシピのレビューに書かれた食感語を集約し、その結果を Word Cloud として可視化するシステムを開発した。図 1 はシステムの概要を示している。本システムはレシピ・レビューデータから可視化に適した形式でデータを抽出するユニットと、抽出したデータを Word Cloud として可視化するユニットに分けることができる。本研究の Word Cloud 可視化には、データとしてレビュー中の各食感語とそれを書いた人数を利用する。一般的なレシピデータセットに含まれるレシピ数は 100 万件以上、レビュー数は数百万件以上に上る。レシピデータセットにおいて、それらがすぐに取り出せるようになっていない。そのためデータベースで集約エンジンを構築し、可視化対象レシピの情報を入力すると、そのレシピに対する全レビューに書かれた食感語と、各語を書いたレビュー数を調べて出力するよう実装した。得られたテキスト表現をもとにし、Word Cloud を提示するプログラムを作成した。以下で各ユニットについて述べる。



図 1：システム概要図

4.1 データ抽出

データ抽出ユニットでは、可視化するレシピ情報を入力すると、そのレシピのレビューに書かれた食感語と各語を書いたレビュー者数を出力する。そのためにレシピ ID を入力すると、ID とそのレシピのレビューに書かれた人数分の食感語を集計テーブルに出力する集約エンジンをデータベースで構築した。図 2 は集計テーブル生成の流れを示している。準備としてデータベース内にレシピ情報を格納したレシピテーブル、レビュー情報を格納したレビューテーブル、抽出する食感語を格納した食感語テーブルを作成する。レシピテーブルには各レシピに固有のレシピ ID、レビューテーブルには各レビューに固有のレビュー ID、レビュー先のレシピ ID、レビューテキストを属性値として持たせる。食感語として農業・食品産業技術総合研究機構 食品研究部門の調査で収集された 445 語の食感語を参考に、526 語を用意した。

食感語をキーとしたデータの抽出を行う際には、何度も同じテーブル群を利用するため、SQL だけでは処理に膨大な時間がかかる。高速化・効率化を図るために VIEW を積極的に活用する。また通常の VIEW と別にマテリアライズドビューを使用して実行に掛かる時間の短縮を試みた。マテリアライズドビューとは VIEW と同様に特定のテーブル群からデータを抜き出したものである。VIEW にある程度の永続性を持たせ

ることで、VIEW を複数回参照する際の検索処理コストを削減することができる。そのため通常の VIEW よりも検索スピードの向上が期待できる。実装実験では Intel(R)Core(TM)i7-4470 プロセッサ、8GB メモリ、Windows10(32bit)、DBMS は PostgreSQL を利用した。

実装実験では集計関数を用いてレビューの多い順にレシピ ID を並び替え、レビューが一つも付いていないレシピをデータ群から落とした。その上で食感語テーブルとレビューテーブルを利用し、食感語ごとにマテリアライズドビューを作成した。各ビューは属性としてレビュー者のユーザー ID、レシピ ID、レビューテキスト、レビュー中の食感語名を持つ。全レビュー文に対し、各食感語が含まれているか Like 文で検索する。しかし例えば「チョコクリームがまるで市販品みたい!!大満足です♪」というレビューでは「コク」という食感語が「チョコクリーム」から抽出されてしまう。正確に食感語の抽出するため形態素解析器を用いた。デフォルトで形態素解析器を使用すると食感語群の中には想定通りに抽出できないものもある。例えば「クリーム状」という食感語を解析すると「クリーム」と「状」の二語に分かれてしまう。そのため予め 526 語の食感語を辞書に登録しておき、形態素解析を行う。その後正しく抽出された食感語と該当レシピ ID を属性に持つ集計テーブルを作成した。食感語ごとに頻度を数えると、その語を書いたレビュー者数となる。

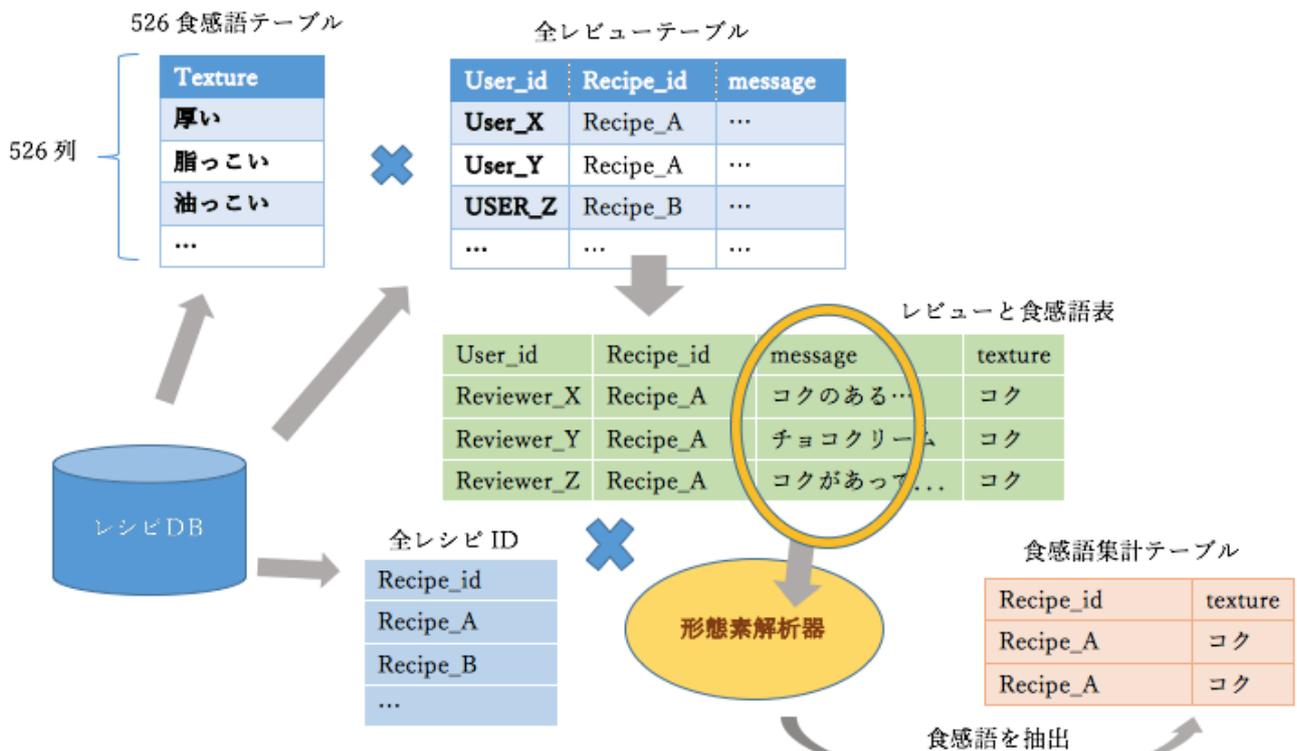


図 2：食感語集計テーブルの生成

4.2 データ可視化

データ可視化ユニットでは、データ抽出ユニットから得られたテキストチャ表現を Word Cloud として可視化する。Word Cloud として各単語をキャンパスに配置する際に最低限必要となるデータは、配置する単語、単語の文字サイズ、単語の座標である。このうち単語と各単語の文字サイズは、データ抽出ユニットから得られる食感語と各食感語を書いたレビュー数をもとにした値とする。食感語を書いたレビュー数を文字サイズとし、可視化するとキャンパスの大きさによっては文字サイズが小さすぎたり大きすぎたりする場合がある。そのためキャンパスに各食感語を適当な大きさで表示できるよう、各食感語を書いたレビュー数を一定の範囲に収まるよう調整し、その値を文字サイズとした。

調整にはレビュー数範囲の縮尺を線形的に一定割合で、また平方根をとって変えるなどの方法がある。縮尺を線形的に変換すると、食感語の頻度の最大値と最小値の差が大きい場合に、頻度が最大の食感語だけが突出して大きくなり、他の食感語が一様に小さく表示されてしまう。これでは他の食感語間の相関関係が一目で分からない。平方根をとって変換すると頻度の最大値と最小値の差が大きい場合でも差が少ない値にでき、各語を適当な大きさで表示することができた。

実装実験ではキャンパスの大きさを 600 ピクセル四方とした。その上で 0 から食感語の最頻値の範囲をとり、各語の頻度が 10 から 110 までの値に収まるよう縮尺を変換した。変換後の値を文字サイズとし、文字を描画する。図 3 は「簡単*濃厚チョコレート」というレシピを各調整方法で可視化したものである。レ

ビュー数は 1412 件であり、そのうち食感語を書いたレビュー数は 512 名である。「濃厚」の頻度が 353 で突出して高く、「しっとり」の頻度が 105、残りの食感語の頻度は 10 未満である。左図では線形的に、右図では平方根で文字サイズを調整し、可視化した。右図の方が「濃厚」以外の語も適度な大きさで表示され、食感語間の相対関係が分かりやすくなっている。

Word Cloud ではキャンパス内に出現頻度に応じた大きさで単語を並べて提示する。キャンパス内では単語が重ならないように単語の座標を調整する。実装実験では d3-cloud ライブラリ [8] を使用し、各単語の座標を決定した。座標の候補は複数個生まれる場合があるが、画面が更新される度にそのうちの 하나가 選ばれる。食感語を何らかの基準で分類できれば座標に何らかの意味を付与でき、より情報量の多い可視化となることが考えられる。単語の座標を決定する際には、表示するキャンパスサイズや単語の密度によって単語間に適度な間隔を空けるようにする。

食感語を分類すると、分類に応じて単語に配色することができ、複数の Word Cloud 可視化結果を見る際、似た食感語ごとに色が統一されていれば色を見比べることでどの様な食感語が提示されているかある程度把握でき、より容易な比較が可能となる。また各食感語にその語のイメージに合った色やテキストマップピングを施すことで、より直観的な可視化になると考えられる。実装実験では各語の色は 20 色から文字サイズの大きい順に決定されている。



図 3：食感語頻度の調整方法による可視化の違い

Word Cloud では単語ごとにキャンパスにおける配置を決めていく。一般的には文字サイズに応じて各単語が収まる仮想的な長方形をとり、それらがキャンパス内でお互いに重ならないよう配置し、その座標に単語を描画する。重ならないければ長方形をキャンパスの下辺に対して斜め、あるいは直角に配置することもできる。そうすると単語は斜めに、あるいは90度傾いて見える。こうした場合、仮想的な長方形を真横にのみ配置する場合よりも単語の配置パターンが増え、より視覚に訴える可視化を提示できる可能性が高まると考えられる。人や動物の輪郭を持つ Word Cloud も見られるが、これらの中には単語の傾きを利用してのものもある。本研究の可視化に対しても応用の余地は考えられるが、実装実験では単語の読みやすさを考慮し、単語が真横に描画されるようにした。

レシピによってはレビューに食感語が書かれていないものもある。そうしたレビューを本研究の手法で可視化してもキャンパスには何も描画されず、利用者はレシピ選びに役立つ情報を何も得られない。レビューに食感語が一つも書かれていない場合でも、本研究の可視化手法によって利用者が何らかの情報を得られるようにする。そのためにレビューテキスト全体を形態素解析し、得られた語群のうち、記号や句読点、助詞を除く各語をキャンパスの背景に Word Cloud として可視化した。レビューに食感語が含まれる場合でも前景の食感語を見やすくするため、背景の文字サイズは頻度に関わらず小さくし、各語を半透明で描画した。

レビューには突出して頻度の高い食感語以外は頻度が同程度低い食感語が多い。そのため文字サイズを調整していても、そうした食感語は同じような大きさで提示される。その場合、一度に食感語を Word Cloud として提示すると、文字サイズの大小関係、つまり頻度の高低関係を把握することが困難である。そこで文字の大きい順、つまり頻度の高い順に食感語を表示するアニメーション処理を可視化に加えた。実装実験では表示間隔を0.5秒としている。これにより、利用者は各食感語頻度の大小関係をより細かく把握することができる。

レシピ検索では利用者が表明した検索意図を満たすレシピを返す。そのため検索結果候補は類似したレシピになる。特にレシピのタイトルに食感語が含まれていると、レビューにその食感語を書くレビューが多くなる。こうしたレビューを食感後で集約し、Word Cloud として可視化すると、その食感語が最も大きく、その他の食感語が比較的小さく提示され、同じ様な見た目となる。絞り込まれた候補からレシピを選ぶとい

う段階ではレシピを比較し、どのような違いがあるかを吟味する。その際に、同様の見た目の可視化を見比べてもレシピの違いが分かりにくい。複数のレシピを比較する際、可視化においてレシピの違いを増幅できれば、容易な比較が可能となり、利用者のレシピ選びを支援することができる。そのための方法として、TF-IDF を利用し、レシピごとに特徴となる食感語を見つけることも考えられる。本稿では可視化データから最頻出した食感語を除く前処理をした後、Word Cloud として可視化した。

例として図4に「さくさく☆うの花（おから）クッキー♪」と「☆クリームチーズのさくさくクッキー☆」というレシピのレビューを可視化した。それぞれレビュー数は768件、1038件であり、レビューに食感語を書いたレビューラー数は216名、211名であった。両レシピとも「さくさく」を書いたレビューラー数が最も多く、それぞれ159名、134名であった。図4の上段では前処理をせず可視化した。両方のレシピとも「さくさく」の頻度が他の食感後より突出して高いので、その文字サイズが格段に大きくなっている。単体で可視化結果を見た場合、利用者は自分で調理すると「さくさく」なクッキーを作れるとイメージできる。ただレシピを比較するとなると「さくさく」以外の文字サイズが一樣に小さいため、レシピの違いが分かりにくい。レシピの差を増幅し、レシピを容易に比較可能にするため、図4の下段では「さくさく」を除いて可視化した。すると残りの食感語が頻度に応じてより大きく提示され、片方にしかない食感語を見つけやすくなっている。これにより、レシピの違いが分かりやすくなり、利用者のレシピ選びを支援できる可視化となった。

各食感語の文字サイズは、頻度に応じて一定の範囲内の値になるよう調整されている。そのため複数のキャンパスにおいて、実際の出現頻度が異なる場合でも、食感語が同じ文字サイズで提示されることがある。複数の可視化結果を用いてレシピを比較する際には、比較対象レシピの中でレビュー数を考慮し、各食感語の頻度を相対化して文字サイズの調整に反映させることが考えられる。そうすれば可視化結果を見ることで、レシピをより正確に比較できるようになる。現段階では食感語にマウスポインタをあてると、その語の右上付近に実際の頻度が表示されるようにした。

参考文献

- [1] 大橋正房・光岡祐子・須藤正彦・北村奈津子・米山真知子・猿渡一兵「sizzle word シズルワードの現在 「おいしいを感じる言葉」調査報告」株式会社 B・M・FT 出版部 2015 年 12 月
- [2] 笹田 鉄郎・森 信介・山肩 洋子・前田 浩邦・河原 達也「レシピ用語の定義とその自動認識のためのタグ付与コーパスの構築」自然言語処理 22(2), 107-131, 2015 一般社団法人 言語処理学会
- [3] 森下幸俊, 中村富予「気分による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,
- [4] 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(DEIM Forum 2013) P3-4, 2013 年 3 月
- [5] Lertsumruaypun Kanwipa・渡辺知恵美・中村聡史「オノマトペロリ: オノマトペを利用した料理推薦システム」研究報告デジタルドキュメント(DD) 2009-DD-73(6), 1-7, 2009-09-18 情報処理学会
- [6] 早川文代・井奥加奈・阿久澤さゆり・米田千恵・風見由香利・西成勝好・馬場康維・神山かおる「質問紙法による消費者のテクスチャー語彙調査」日本食品化学工学会誌 Vol.53,No.6 (2006)
- [7] 早川文代・井奥加奈・阿久澤さゆり・齋藤昌義・西成勝好・山野善正・神山かおる「日本語テクスチャー用語の収集」日本食品科学工学会誌 Vol.52,No.8 (2005)
- [8] <https://github.com/jasondavies/d3-cloud>