

レビュー中のテクスチャとそのカテゴリを利用したレシピの評価と その可視化

執行 健人[†] 清光 英成[†] 大月 一弘[†]

[†] 神戸大学国際文化科学研究科 〒657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: [†]185c124c@stu.kobe-u.ac.jp, ^{††}{kiyomitu,ohtsuki}@kobe-u.ac.jp

あらまし レシピのレビューには、そのレシピを使用したレビュー投稿者の調理体験が書かれている。レビューコンテンツは、レビュー投稿者の調理体験を知りたいレシピ作成者・レシピ閲覧者にとって、有用な情報源となりうる。しかし人気レシピには数千件のレビューがつくことも少なくなく、レビューを一件ずつ読んでいくことは現実的ではない。本研究では、レシピ作成者・レシピ閲覧者が知りたい視点からレシピのレビューを集約・可視化することで、大量のレビューコンテンツを活用した情報を提供することを目的とする。レビュー中のテクスチャを集約することで、レビュー投稿者の調理結果を容易に把握できるようにした。レビュー中のテクスチャはレビュー投稿者の調理結果を表現する語彙であり、料理の出来上りを直感的に伝えることができる。さらにテクスチャの分類をレビュー集約・可視化に適用する方法を述べる。またレビュー投稿者の謝意を集約することで、レシピ評価を行う可能性について示す。キーワード 食メディア, レシピ, 味覚情報, 食感, テクスチャ, ソーシャルコンピューティング

1 はじめに

食品名や料理レシピのタイトルにテクスチャ (food texture) が使用されることは多い。ISO [1] の定義によれば、テクスチャとは「力学的、触覚的、および適切であれば視覚的、聴覚的な方法で感知できる食物の力学的、幾何学的、表面的属性の総体」である。テクスチャは食べ物の状態や食感といった複合的な情報を、一言で分かりやすく伝えることができる。例えば、「濃厚」という言葉からは色・味・香りなどが濃く、あっさりした味とは対照的食感がイメージできる。焼き芋の食感を「しっとり」と表現した場合、程よい潤いがあり舌触りも良く、パサパサしていないことが想像できるが、「ほくほく」を求める利用者には湿潤すぎることが想像できる。このように、テクスチャを用いることで、出来上りの歯ごたえや舌触りがとてもイメージしやすくなる。早川らにより、食感を表現する語としてよく使われるテクスチャは、日本語 445 語、英語 77 語、フランス語 227 語、中国語 144 語あることが調査された [2]。

料理レシピは雑誌や書籍などに記載され、読まれてきた。近年では Web 上でレシピを提供するレシピサイトが多くの人々に利用されている。中でも利用者が投稿したレシピを提供するレシピサイトがあり、代表的なものにクックパッド [3] や楽天レシピ [4] が挙げられる。本稿ではこのようなレシピサイトを投稿型レシピサイトと呼ぶことにする。投稿型レシピの特徴として、レシピのレビューの投稿や利用ができる点が挙げられる。レシピのレビューには、そのレシピを使用したレビュー投稿者の調理体験が記述されている。例えば、クックパッドでは「つくれば」、楽天レシピでは「つくったよレポート」がこれに当たる。レビューの記述内容としては調理の感想や料理の出来上がり、レシピ投稿者への謝意など様々なものがある。レ

ビューはレシピのメタデータとしても利用されている。例としてクックパッドでは 1000 件以上のレビューがついたレシピを「殿堂入りレシピ」という名称で分類し、利用者による評価が高いレシピとしている。またレビュー投稿者数が 10 人、100 人、1000 人以上という単位でレシピの絞り込みができる。一方で、レビューコンテンツ (レビューに書かれた内容) をレシピのメタデータとして捉え、活用する工夫はされてこなかった。レビューは、レシピ投稿者・レシピ閲覧者にとって有用な情報源となり得る。レビューを読むことで、レシピ投稿者はレシピ使用者 (レビュー投稿者) がどのような調理結果を得たのか、また、どのような感想を持ったのかというフィードバックを得ることができる。これにより、レシピの改善やレシピ作成のモチベーションの向上が期待できると考えられる。レシピ閲覧者は、レビュー投稿者の調理前と概ね同じ状況にあると考えられるため、レビュー投稿者の調理結果を参考にしてレシピを選ぶことができる。しかし人気レシピには数千件のレビューがつくことも少なくなく、レビューを一件ずつ読んでいくことは現実的ではない。レビューコンテンツを有効に活用するためには多数のレビューを何らかの方法で集約し、内容を容易に把握できるようにする必要がある。

本論文では、レビューコンテンツを活用し、レシピ投稿者・レシピ閲覧者に有用な情報を提供することを目的とする。そのために多数のレビューが投稿されたレシピについて、レビュー中のテクスチャを用いたレビューの集約方法とその可視化を議論する。これにより、大量のレビューを読まなくてもレビュー投稿者の調理結果を一目で把握できるようにする。さらにレシピ作成者に対するレビュー投稿者の謝意を表現する語彙を集約することで、レシピの評価を行う可能性についても述べる。本論文の構成は以下の通りである。第 2 節で食メディアにおける関連研究に触れ、本研究の位置付けを明らかにする。第 3 節で

レビューの集約について、その必要性や集約する情報、利用する語彙、集約方法と集約結果を示す。第4節ではレビューの集約結果に基づいたレシピ評価について述べる。第5節でレシピ評価の可視化法を示す。第6節はまとめである。

2 関連研究

食メディア研究はレシピ・献立検索、味覚情報提示など情報処理技術を駆使した研究が行われている。これらのうち、代表的なものを関連研究として紹介し、本研究の位置付けを明らかにする。

2.1 レシピ・献立検索と推薦

渡辺ら [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。さらに、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。

2.2 味覚情報提示

シズルワードは飲食行為以前に、飲食をしようとしている人々に対して伝達したいおいしさを表すことばである。大橋らは、おいしさを伝える単語や熟語をシズルワードとして調査・整理している。300を超える語を味覚系（味覚・嗅覚）90語、食感系（触覚・聴覚）102語、情報系（知識）113語の3分野に分けておいしさを感じるランキングなどの調査を行った [6] [8] [7]。飲食店や食品に関する訴求に優れる語と捉えることができるシズルワードは、その性質上 SNS などでのよい飲食店や秀逸なメニューを伝えるために用いられることが多いため、シズルワードを手掛かりにした SNS 研究も行われている [9] [10]。これらのレシピ検索に関する研究は利用者が表明できる検索意図を増やし、検索結果に反映させることを目的としている。味覚情報提示研究は、利用者にどのようにして味や食感を知覚・伝達させるかあるいは、知覚がどのように変化するかを解明することを目的としている。

本研究は、レシピのレビュー中のレビュー投稿者の調理結果・謝意を表現する語彙を集約し、レビュー投稿者による調理結果と謝意の情報表現とした可視化を提示する。これにより、大量のレビューを読まなくてもレビュー投稿者の調理結果を容易に把握できる。本研究では、レシピのレビューに含まれるテキストを、そのレビュー投稿者が自身の調理結果を認識して表現したものと捉える。これは、レビュー投稿者の口腔内の感覚器官で知覚され意味付けされることで認識が成立するが、テキストとして表現されたことでその種の認識があったことを客観的に確認できるためである。

3 レビューの集約

本節ではまずレビュー集約の必要性を示し、集約する情報や集約に用いる語彙を検討する。その後レビューの集約方法と集約結果について述べる。

3.1 レビュー集約の必要性

人気レシピには数千件のレビューが付くこともあり、一件ずつ読んでいくことは現実的ではないと上述した。ここではレビューを集約する必要性について言及する。まず一般の利用者が閲覧しようと思うレビュー数を調べるため、大学生 270 名を対象としてアンケート調査を行った。「レシピを参考にして料理をした感想（1件あたり 50 文字程度）がレビューとして投稿されています。あるハンバーグのレシピに 1,200 件のレビューがありました。何件までならレビューを読みますか？（ある程度時間に余裕があるという前提でお答えください）」という質問文に対して、何件までならレビューを読もうと思うかを答えてもらった。回答結果の平均は 35 件（小数点第一位で四捨五入）、最頻値は 10 件（66 名）であった。

次に実際のレシピサイトに掲載されたレシピについてのレビュー数を調べた。レシピデータとして、クックパッド株式会社が国立情報学研究所と協力して研究者を対象に提供しているクックパッドデータセット [11] を利用した。このデータセットには、2014 年 9 月 30 日までにクックパッドに掲載されたレシピに関するデータが収録されている。レシピ数は 1,715,595 件、レビュー数は 9,449,479 件である。全レシピのうちレビューが付いているレシピ数は 46.9% (805,018 件) であり、レビューが 1000 件以上付いているレシピは 601 件である。なおレビューが 36 件以上付いているレシピは 4,4682 件であった。上記のアンケート調査の結果から、これらのレシピには読みきれないほど大量のレビューがついており、その内容が一件ずつ読まれることは少ないと考えられる。大量のレビューの内容を有効利用するためには、何らかの視点からレビューを集約する必要がある。

3.2 集約する情報

本稿ではレビュー中の語彙を集約して得られる情報として、レビュー投稿者の調理結果とレビュー投稿者の謝意に着目する。各情報が利用者にもたらすメリットについて検討する。

3.2.1 レビュー投稿者の調理結果

レシピのレビューとは、レシピ作成者にとって自身が作成したレシピに対する評価やフィードバックである。レシピに少数のレビューがついている場合は、一件ずつ読むことでレビュー投稿者の調理結果を知ることができる。しかし大量のレビューがついている場合は、レビューテキストから調理結果を拾っていくことは容易ではない。また料理は作る人によって出来上がりが異なり、作った人が多ければそれだけ調理結果も多様になる場合もある。そのためレビューを全て読んだとしても、多様な調理結果を把握することは困難であると考えられる。レビュー投稿者の調理結果を集約すれば、レシピ作成者は大量の

レビューを読まなくてもレビュー投稿者がどのように調理したのかを俯瞰的に把握することができる。その結果、成功した調理結果があれば、レシピが役に立ったことを確認できる。また失敗した調理結果があれば、レシピの改善に役立てることができる。

レシピ閲覧者はレビュー投稿者と概ね同じ立場のため、レビュー投稿者の調理結果をレシピ選びの手掛かりにできる。多くのレビュー投稿者の調理結果が概ね同じレシピであれば、自分もそのように調理できると考えられる。逆に多くのレビュー投稿者の調理結果が異なるレシピであれば、調理者が持つ調理技術や調理道具に依存するレシピであると考えられる。

3.2.2 レビュー投稿者の謝意

食品関係の企業のレシピサイトでは、料理の専門家や企業関係者がレシピを作成する。レシピ作成のモチベーションの一つとして、金銭的な報酬が考えられる。一方で投稿型レシピサイトでは、金銭的な報酬は良質なコンテンツ作成の動機付けになるとは限らない。金銭的な報酬を目当てに質の低いレシピが大量に投稿されることも考えられるからである。レシピ投稿の動機として、レシピを他者と共有することで役に立ちたいという貢献欲が考えられる。そのような動機を持つレシピ作成者は、レビュー投稿者から感謝の言葉を得ることで、自身の貢献を確認できる。これにより、レシピ投稿のモチベーションがさらに高まると考えられる。この結果より多くのレシピが投稿され、それらが利用できるようになれば、レシピ閲覧者にとっても有用となる。

3.3 レビュー集約の語彙

レビュー投稿者の調理結果・謝意という視点からレビューを集約するために、レビューテキストから集約する語彙を検討する。

3.3.1 レビュー投稿者の調理結果

レビュー中でレビュー投稿者の調理結果が分かる語彙はいくつかある。「甘い」「辛い」など味覚を表現する語彙は、食べ物を食べたときに舌で感じる味覚である。調理の結果、食材や料理の味がそのようになったことが分かる。シズルワードは、食べ物の美味しさを表現する語彙である。「ジューシー」「ほくほく」など、食品や料理の美味しさをアピールする際によく用いられ、料理のレシピやレビューにも散見される。レビューにシズルワードが書かれていれば、食材や料理の出来上がりと共に美味しい料理を作れたということが分かる。また「ふわふわ」や「サクサク」など、食品の食感覚を表現する語彙としてテクスチャがある。テクスチャは食べ物の状態や食感などの複合的な情報を、一言で分かりやすく伝えることができる。調理結果を直感的に把握することができ、味覚と比べて認知の個人差が少ない。テクスチャには、大橋らによって整理された味覚系・食感系のシズルワードの一部が含まれる。シズルワードは食べ物の美味しさを表現する語彙だが、人によって食材の美味しい美味しさは異なる。また人によって料理の技術は異なり、同じレシピを参考に料理をしても、全ての人があまく調理できるわけではない。そのような調理結果がレビュー中で「美味しい」を

表現するシズルワードとして書かれることは考え難い。テクスチャには「ぐずぐず」や「もそもそ」など、調理に失敗した結果得られるような食べ物の食感を表現する語彙もある。そのためテクスチャであれば、失敗したような調理結果も表現できる。

以上の点から、好ましい意味を持つシズルワードだけでは、個人差のある好みの食感覚やうまくできなかった調理結果をも捉えることはできないと考えられる。そのため本研究ではレビューを集約する語彙としてテクスチャを利用する。早川 [12] らはアンケート調査や文献検索によって、日本語のテクスチャ 445 語を収集した。この語彙リストでは、カタカナとひらがなを統一し、音を表現することが明らかな場合および外来語はカタカナで表記されている。レシピのレビュー中のテクスチャは、ひらがな・カタカナどちらの表記でも書かれることを考慮する必要がある。そのためレビュー集約の際には、どちらでも表記が可能なテクスチャは、ひらがな・カタカナのテクスチャを両方とも抽出する。集約の際には語彙リストでの表記に統一し、ひらがな・カタカナどちらかのテクスチャとして頻度を集計する。またテクスチャの語彙リストでは「口あたりがよい」「ゼリー状の」など、テクスチャ以外の助詞やテクスチャを表現するための言葉が含まれている。レビュー中のテクスチャをできるだけ漏らさず抽出するために、テクスチャであると明らかに分かる場合は助詞等を除いて抽出対象のテクスチャとした。

また早川らはテクスチャの専門家 8 人を評価者とし、収集した 445 語のテクスチャを、ISO のテクスチャープロファイルの 3 特性である力学的特性、幾何学的特性、その他の特性 (水分・油脂に関する特性) に大別した [13]。日本語テクスチャには複数の意味を有していて場面によって使い分けられるものや、複合的な意味を表現するものがあるため、大別において重複を許している。3 特性の用語数は、力学的特性が 322 語、幾何学的特性が 135 語、その他の特性 (水分・油脂に関する特性) が 73 語である。445 語のうち、85 語は複数の特性に属している。本研究ではこの分類を、テクスチャを用いたレビューの集約と可視化に利用する。この分類によりテクスチャを構造化することで、集約結果からどのレシピのレビューにどの種類のテクスチャが多いかなどが分かる。また人気レシピについての大量のレビューには、様々な種類のテクスチャが書かれることがある。このような場合のレビュー集約結果の可視化において、テクスチャの分類情報を何らかの形で提示することもできる。

レビューテキストがテクスチャ文字列を含むならば、常にそのような調理結果が生じたとは限らない。例えば「サクサク」というテクスチャを含むレビューがあったとしても、「パリパリじゃない」「サクサクにはならなかったけど...」などテクスチャを否定するように書かれていれば、レビュー投稿者はそのテクスチャのように調理できなかったことになる。このようなテクスチャを集約する際には助動詞「ない」や形容詞「ない」、補助形容詞「ない」などがテクスチャに係っているかどうかを調べ、区別する必要がある。簡易的な係り受け解析により、このような事象の生起はテクスチャの出現百万回に対して一万五千回であった。重篤とは言えないが、精度の点から解決すべき問題であると認識している。こうした場合のテクスチャ集約方法

の検討は、今後の課題とする。

3.3.2 レビュー投稿者の謝意

一般的な場面で使われる感謝の言葉として「ありがとう」「感謝」「おかげさま」がある。また料理の場面で使われる感謝の言葉としては「ごちそうさま」「おいしかった」が挙げられる。本稿ではこれらの語彙をレビューから抽出し、集約することにする。

3.4 レビューの集約方法

3.4.1 テクスチャの集約

レビュー中のテクスチャを集約する方法について述べる。レシピのレビューデータの属性としては対応するレシピ ID、レビュー投稿者のユーザー ID・レビューテキストがある。まずテクスチャごとに全レビューテキストから当該テクスチャの部分一致検索を行い、テクスチャを含むレビューを抽出する。その結果をレシピ ID、レビュー投稿者のユーザー ID、レビューテキスト、テクスチャという組を得る。その後レビューテキストを形態素解析器 MeCab を用いて形態素に分解し、テクスチャと等しい形態素があるかどうかを調べる。形態素解析の際にはテクスチャのユーザー辞書を作り、コスト値を全て 1 に設定した。これによりテクスチャの優先度を高め、レビューからテクスチャの形態素を正しく抽出できるようにした。テクスチャと等しい形態素がある場合は、そのレビューテキストは当該テクスチャを含むとし、レシピ ID、レビュー投稿者のユーザー ID、テクスチャという組を作る。レビューに対する部分一致検索だけでは、テクスチャではない単語に適合する場合がある。例えば「コク」というテクスチャを抽出したい場合に、レビューテキストに「チョコクリーム」という単語が含まれていれば、その単語中の「コク」が抽出されてしまう。このような場合を考慮し、部分一致検索の後に形態素解析によるテクスチャの抽出を行った。テクスチャを抽出して得られた組みから、当該レシピのレビューにテクスチャを用いた人数をユーザー ID により算出する。これにより、レシピ ID、テクスチャ、レビュー投稿者数という組を得てレビューの集約データとする。こうすることで同一のレシピに対して一人で複数回のテクスチャを用いたレビューの重複集計を避けることができる。あるレシピのレビューにテクスチャが書かれていれば、そのレビュー投稿者にテクスチャと同様の認知がされたと考えられる。そのようなレビュー投稿者数が多い場合、利用者はそのレシピを使用して調理すると、そのテクスチャで表現されるような調理結果になる可能性が高いと考えられる。そのためレビューにテクスチャを書いた人数を集計し、テクスチャの頻度とした。

3.4.2 テクスチャの集約結果：テクスチャ表現

レビュー集約結果のデータ構造は、レシピ ID・ユーザー ID と 445 種類のテクスチャの有無を 0/1 表現した組を情報単位とする。レシピごとにこの組数の集計結果から得られた情報をテクスチャ表現とする。これは 445 種類の各テクスチャを何人のレビュー投稿者が用いたかの総計である。テクスチャ表現は以下の t^r で表すことができる。

$$t^r = (t_1^r, \dots, t_{445}^r)$$

次に、レビュー中でレビューアーが用いたテクスチャは何であったのかを以下のように表現する。 t_u^r をレビューアー u がレシピ r に対するレビュー中で用いたテクスチャ組とし、

$$t_u^r = (t_{u1}^r, \dots, t_{ui}^r, \dots, t_{u445}^r)$$

で表す。各 t_{ui}^r ($1 \leq i \leq 445$)、はレビューアー u がテクスチャ t_i をレシピ r のレビュー中で用いた場合 1、用いなかった場合 0 である。レシピのテクスチャ表現 t_i^r を、レシピ r に投稿されたレビュー中でテクスチャ t_i 用いたレビューアー数とする。レシピ r にレビューを投稿したレビューアー集合を U_r とした時、

$$t_i^r = \sum_{u \in U_r} t_{ui}^r \quad (1)$$

である。

ここで、レビュー中に出現したテクスチャと、そのテクスチャを用いたレビューアー数の分布について、「最も多くのレビューアーが用いたテクスチャと 2 番目以降のテクスチャのレビューアー数の差が著しい」という、概ねの傾向が見られたので言及する。図 1 に 1000 件以上のレビューがついた全 545 件のレシピに関し、各レシピのレビューで各テクスチャを用いたレビュー投稿者数の分布を示す。レシピごとに散布図を作成し、それらを重ね合わせた。横軸は各レシピのレビューで用いられたテクスチャを、レビュー投稿者数の多い順に 1 位から並べている。縦軸は各レシピのレビューでテクスチャを用いた総レビュー投稿者数に対し、当該順位のテクスチャを用いたレビュー投稿者数の割合である。対象とするレシピのうち、そのレシピのレビューに書かれたテクスチャの種類数は累計で最大値 66、中央値 24、最小値 12 であった。テクスチャを用いたレビュー投稿者数の割合上位 1 位は最大で 92.0%、2 位で最大 41.6%、3 位で 23.6%となっている。各レシピの分布は概ね上位 1 位と 2 位以下のレビュー投稿者数の割合の差が著しく、それ以降は減少している。このようにレビュー数が 1000 件以上のレシピでは、レビューでテクスチャを用いたレビュー投稿者数の分布は概ねべき乗側に従うことが言えそうである。このため、1 番目と 2 番目のレビューアー数の比が 2 倍程度であっても、1 番目と 3 番目との比は 10 倍以上となるレシピが散見される。この場合、レビュー投稿者

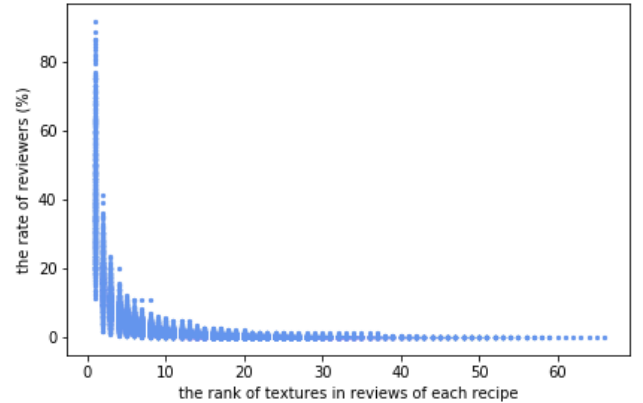


図 1 1000 件以上のレビューがついた全レシピのレビューでテクスチャを用いたレビュー投稿者数の分布

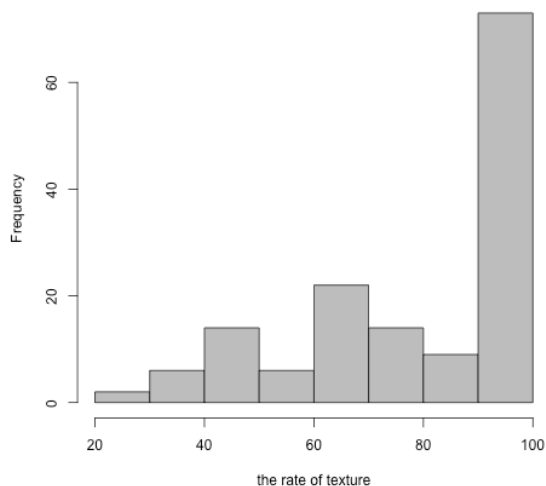


図2 レシピタイトルに「ふわふわ」を含むレシピ

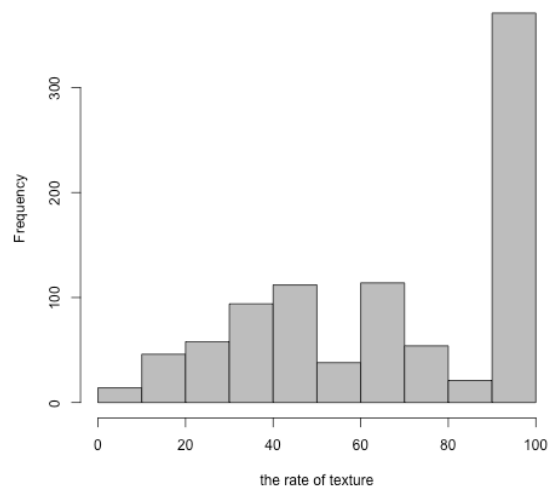


図3 レシピタイトルに「ふわふわ」を含まないレシピ

数の総和をそのまま利用することは現実的でない．テクスチャによるレシピの評価は，名義尺度に基づくため名義数のある程度保持してレシピの評価を行いたい．そこで，差異減衰に対数を用いてレシピのテクスチャ表現とすることにした．

$$t_i^r = \log \sum_{u \in U_r} t_{ui}^r \quad (2)$$

なおレシピタイトルにテクスチャが含まれていれば，そのテクスチャがレビューに書かれやすくなると考えられる．例としてまづクックパッドデータセットでレビュー中にテクスチャ「ふわふわ」を含み，レシピタイトルに「ハンバーグ」且つ「ふわふわ」が含まれているレシピ (146 件)，「ハンバーグ」を含み「ふわふわ」が含まれないレシピ (922 件) を抽出した．図2と図3では，各レシピのレビューに書かれたテクスチャのうち，「ふわふわ」の割合を横軸に，レシピ数を縦軸にそれぞれに示している．レシピタイトルに「ふわふわ」を含む・含まないどちらのレシピでも，レビューに書かれたテクスチャのうち，「ふわふわ」の割合が60%以上を占めるレシピが多い．しかしレシピタイトルに「ふわふわ」を含むレシピの方が，レビューに書かれたテクスチャのうち「ふわふわ」の割合が60%未満であるレシピの分布が少ないことが分かる．これは，レシピタイトルにテクスチャが含まれるレシピには，レビューにそのテクスチャが比較的多く書かれるという一例であると考えられる．

3.4.3 レビュー投稿者の謝意の集約

テクスチャと同様に，レビュー投稿者の謝意と捉えられる「ありがとう」「感謝」「ごちそうさま」「おかげさま」「美味しかった」という語彙をレビューから抽出し，集約する．「感謝」を除く各語彙に関して，一部または全て漢字表記されている語彙はひらがなの語彙にまとめた．

まず語彙ごとに全レビューテキストから当該語彙の部分一致検索を行い，その語彙を含むレビューを抽出する．その結果をレシピ ID，レビュー投稿者のユーザー ID，レビューテキスト，謝意の語彙という組みを得る．謝意の語彙の場合はテクスチャ

とは異なり，その一部が別の単語に含まれることはない．従って形態素解析による抽出は行わない．

3.4.4 レビュー投稿者の謝意の集約結果

前節で上述した組みからレビューテキストに謝意の語彙を用いた人数をユーザー ID により算出し，レシピ ID，謝意の語彙，レビュー投稿者数という集計結果を得る．表1では，各語彙の集約結果を当該語彙，その語彙をレビューに含むレシピ件数，データセットの全レシピ数に対する当該レシピ件数の割合として示している．レシピ件数が最多の語彙は「おいしかった」で517,759件 (64.3%)，次いで「ありがとう」が165,884件 (20.6%) であった．一方で「おかげさま」を含むレビューがついたレシピ件数は3045件 (0.38%) であった．このように「おかげさま」がレビューに書かれたレシピは全レシピの1%に満たなかったため，この語はレビューの集約に利用しないことにする．「おいしかった」「ありがとう」「ごちそうさま」「感謝」がレビューに書かれたレシピは少なくないことが分かった．レビュー投稿者の謝意という視点からレビューを集約する語彙として，これらの語彙を利用することにする．今後，この集約結果に基づいたレシピの評価法を検討したい．その際にはレビュー投稿者の謝意の量化や分類などが考えられる．またレビュー投稿者に感謝されるレシピの特徴や，感謝を表現する語彙をよく用いるレビュー投稿者の特徴などを分析し，レシピの評価に応用したい．

表1 レビュー投稿者の謝意の集約結果

謝意の語彙	レシピ数 (件)	レシピの割合 (%)
おいしかった	517,759	64.3
ありがとう	165,884	20.6
ごちそうさま	120,879	15.0
感謝	114,114	14.1
おかげさま	3046	0.38

4 レビュー集約結果の提示

本節ではまず、レビューの集約結果を表として提示する例を挙げる。その上で、調理結果をさらに直感的に想像できるような提示手法をいくつか検討し、テキストチャ表現をワードクラウドとして可視化する方法を示す。

4.1 表による提示

レビューの集計結果を表で提示する例を示す。表 2 に「秘密の豆腐ハンバーグ」というレシピのレビューに書かれたテキストチャ表現値が大きい上位 10 件のテキストチャと、各テキストチャを書いたレビュー投稿者数、テキストチャを書いた全レビュー投稿者数のうち、各テキストチャを書いたレビュー投稿者数の割合を示す。このレシピに対するレビュー数は 7819 件であり、レビューにテキストチャを書いた人数は延べ 3559 人であった。レビューに書かれたテキストチャの種類は 53 種類であった。表 2 は、テキストチャを用いた延べ数のレビュー投稿者に対して 63.4% のレビュー投稿者が「ふわふわ」というテキストチャを用いたことを示している。ふわっ、ふんわり、ふわふわは促音や撥音、長音、り語尾を接辞としている、或いは反復型のオノマトペであり、擬態語「ふわ」の変形と考えてよい それらを合わせるとテキストチャを用いたレビュー投稿者の 8 割以上に「ふわ」という認知が生じていると考えられる。これにより、利用者が調理すると「ふわふわ」なハンバーグになると想像できる。このように、レビュー中のテキストチャの集約結果を表にすることで、大量のレビューを読まなくてもレビュー投稿者の調理結果を容易に把握できるようになる。ただし表による提示では、レビューに書かれたテキストチャの種類が多くなれば、表の行数も多くなる。大量の行を一行ずつ読んでいくことは現実的ではない。表による提示に加え、さらに容易に調理結果を想像できる可視化ができればよい。

表 2 レビュー中のテキストチャ例

テキストチャ	レビュー投稿者数 (人)	レビュー投稿者数率 (%)
ふわふわ	2257	63.41
ジューシー	348	9.77
ふわっ	317	8.90
ふんわり	211	5.92
ふっくら	197	5.53
柔らかい	40	1.12
パリッ	32	0.89
しっかり	27	0.75
なめらか	20	0.56
しっとり	16	0.44

4.2 可視化手法の検討

テキストチャ表現の可視化手法について議論する。レーダーチャートは複数の次元を各軸にとり、原点から放射状に軸を描くことで対象全体の傾向 (バランス) を把握できるように可視化する提示方法である。対象がどのような傾向を持っているか把握するのに便利であり、複数の対象を比較する場合、グラフ

の形を照らし合わせることで違いや類似点を容易に把握できる。評価の次元数は現実的には多くはできない。複数の対象を比較するためには、評価項目とその数をそろえる必要がある。レシピによってレビューに用いられるテキストチャの種類と数は異なる。そのためテキストチャ表現をもとにしたレーダーチャートで複数のレシピを比較するためには、当該レシピのレビューで用いられたテキストチャを、いくつで評価するか決める必要がある。レシピを比較する際には、レビューで用いられたテキストチャ全てを軸にする必要はない。利用者が評価したいテキストチャを軸として選べれば良い。評価の軸を適切に設定することでレシピの比較に有用な可視化になると考えられる。

ワードクラウドは文章中の単語の出現頻度に応じた大きさで単語を画面に敷き詰めて提示する手法である。そのため単語の出現度が高いほど大きく、低いほど小さく文字が表示される。本研究の目的を考慮し、テキストチャ表現をワードクラウドとして可視化することにした。提示する次元数の多少が問題にならず、テキストチャ表現を構成する 445 の軸を表現できる。また順位や格付けがされず、参加の動機である貢献欲を損なうことが少ないと考えられる。レシピサイトの目的は、利用者の要求に合うレシピを提供することである。利用者の検索要求が明らかでない場合、探したいレシピに対する明確な要求を数個のキーワードで表明することは困難である。ワードクラウドであれば検索意図として表明した以外のテキストチャ表現が評価に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なると考えられる。

以上の検討により、本研究ではテキストチャ表現をレーダーチャートとワードクラウドとして可視化する。利用者には両方の手法でテキストチャ表現を可視化した図を提供できるようにし、要求に応じてどちらかを選べるようにする。本稿では以降の節で、レビューの集約結果をワードクラウドとして可視化する方法について述べる。

4.3 ワードクラウド可視化

本節ではレビューの集約結果をワードクラウドとして可視化する方法について述べる。例として図 4 に 4.1 節で表による提示を行った「秘密の豆腐ハンバーグ」のテキストチャ表現をワードクラウドとして可視化した。以降ではこのような可視化を行うために、まずテキストチャ表現に基づいた文字サイズの決定方法に触れ、その後に文字色・文字の位置について述べる。

4.3.1 文字サイズ

本研究のワードクラウドでは、5 ページの式 2 によって定義されるテキストチャ表現の値に基づいて単語の文字サイズを決定する。つまり、レビュー中でその単語を用いたレビュー投稿者数が多いテキストチャほど文字サイズを大きく、少ないテキストチャほど小さくする。レビューにあるテキストチャが書かれている場合、そのレシピを参考に調理した多くのレビュー投稿者にそのような認知が生まれたと考えられる。利用者は各テキストチャの大きさを見比べることによって、各テキストチャが相対的にどれくらい多くのレビュー投稿者に認知されたかを知ることができる。レシピ閲覧者はそのようなテキストチャを目にすることで、

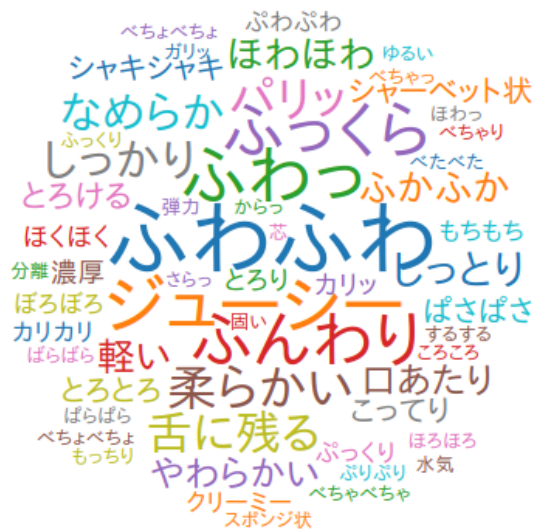


図 4 「秘密の豆腐ハンバーグ」テキスト表現のワードクラウド可視化

そのレシピを使用した場合の調理結果を容易に想像できる。

テキスト表現の値をそのまま文字サイズとすると、文字が小さすぎてテキストを判読することは難しい。またどのテキストも同様に小さくなるため、どのテキストがレビューに相対的に多く書かれたのかも分からない。そこでテキスト表現の値を、ワードクラウドで利用する文字サイズに適した大きさに変換する。レシピごとにテキスト表現の値の範囲を一定の範囲に収まるよう変換し、各テキストの縦幅の文字サイズとする。テキストの横幅は縦幅とテキスト構成する各文字幅によって決まる。まず可視化するレシピごとに式 2 におけるテキスト表現の値の範囲を求める。これを定義域 (domain) とする。定義域の最小値は、あるレシピのレビューに書かれたテキストの内、最も用いられなかったテキストを使ったレビュー投稿者数に対数を取った値となる。

最下位のテキストが複数ある場合、そのレビュー投稿者数に対数をとった値を定義域の最小値とする。逆に、テキスト頻度の最大値は、あるレシピのレビューに書かれたテキストの内、最もよく用いられたテキストを書いたレビュー投稿者数に対数を取った値となる。最も多くのレビュー投稿者に用いられたテキストが複数ある場合は、そのレビュー投稿者数に対数をとった値を定義域の最大値とする。

ワードクラウドで提示するテキストの文字サイズの範囲は、ワードクラウドを描画する画面やキャンバスサイズに応じて適宜決定する。この範囲を値域 (range) とする。実装実験ではスマートフォンの画面幅を考慮し、キャンバスサイズを 310 ピクセル四方とした。レビューから抽出する日本語テキストの最小文字列は 1 文字 (芯)，最大文字列は 7 文字 (シャーベット状，サンドイッチ状) であった。最大文字列のテキストを最大文字とした際に、キャンバス内に収まるようにする必要がある。そのためテキストの最大文字サイズは、キャンバスサイズをテキストの最大文字列数で割った商の小数点以下を切り

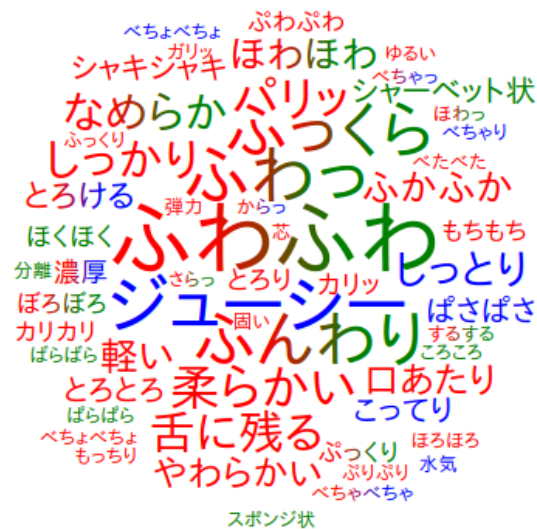


図 5 「秘密の豆腐ハンバーグ」テキストの分類による配色

捨てた値とした。実装実験での最大文字サイズは 44 ピクセルとなる。テキストの最小文字サイズは 10 ピクセルとした。これにより、値域の最大値が 44，最小値が 10 となる。

まず、定義域の最小値を値域の最小値，定義域の最大値を値域の最大値に置き換え文字サイズとする。それ以外のテキストの文字サイズは以下の式によって決定される。

$$Size(t_i^r, R) = \frac{t_i^r}{Max(T^r)} \times (R_{max} - R_{min}) + R_{min}$$

この式では、あるレシピのレビューに書かれた一つ，または複数個のテキストが、それぞれ一人のレビュー投稿者によって書かれた場合、つまり定義域が (1,1) の時、式 の分母と分子が 0 となり値は不定となる。そのためこのような場合には、各テキストの文字サイズを縦幅 10 ピクセルとする。

4.3.2 文字色

各テキストには利用者の要望に応じて単色，または多色を配色することができる。大学生 24 人にアンケートをとり、ワードクラウド中の各テキストに単色を配色した場合と多色を配色した場合のどちらが見やすいかを回答してもらった。単色で配色した図，多色で配色した図を選んだ回答者数はそれぞれ 7 名，17 名であった。各テキストに多色を配色した方が見やすいと回答した人の方が多かったため，本論文のワードクラウドの図では各テキストに多色を配色している。同じテキストであれば同じ色を用い，頻度が高いテキスト順に 20 色を循環利用した。また 3.3.1 節で述べたテキストの分類を文字色に適用する試みも行った。図 5 では「秘密の豆腐ハンバーグ」のテキスト表現をワードクラウドとして可視化している。各テキストの文字色は，テキストの 3 特性によって塗り分けられている。ここでは簡易的に力学的特性は赤色，幾何学的特性は緑色，その他の特性 (水分・油脂に関する特性) は青色を用いた。テキストには複数の分類に属するものもある。3つの分類に属するテキストは「クリーミー」「ぼそぼそ」など 6 語，2つの分類に属するテキストは「ふわふわ」「濃厚」な

ど 78 語ある。これら複数の分類に属するテクスチャの配色には、当該の色を用いたグラデーションを適用した。例えば図 5 の中央に見られる「ふわふわ」は力学的特性と幾何学的特性に属するテクスチャである。そのためこのテクスチャには赤色と緑色を用いて左から右へ色が遷移するグラデーションが配色されている。複数の分類に属するテクスチャを赤・青・緑以外の別の色を用いて配色する実装も行ったが、これでは色数が多くなり、色と分類の対応が分かりにくくなると考えられた。今後の評価により、グラデーションを利用したテクスチャの配色方法の有用性を検証したい。

4.3.3 文字の位置

各テクスチャの描画位置は頻度の高い順に中心から外側に、重ならないように配置される。ただしテクスチャが文字の隙間に収まる大きさであれば、その隙間に挿入する。

5 ま と め

本論文では、Web 上に投稿された料理レシピに対する大量のレビューの内容を活用することを目的とし、レビューから日本語テクスチャと謝意の語彙を集約して利用者に提供できるようにした。大量のレビューを読まなくても、利用者はレビュー投稿者の調理結果や謝意を容易に把握できるようになる。テクスチャであれば料理の出来上がり直感的に伝えることができ、シズルワードでは表現できない失敗した調理結果をも捉えることができる。これにより他の利用者の調理結果を知りたいレシピ投稿者やレシピ閲覧者にとって有用な情報を提供できる。レビューの集約結果であるテクスチャ表現の提示方法を検討し、ワードクラウドによる可視化法を示した。またレビュー投稿者の謝意という視点からのレビュー集約方法を述べた。集約結果に謝意を表現する語彙があれば、レシピ作成者は自身の貢献を確認することができる。これにより、レシピ作成のモチベーションを高めることができると考えられる。レビュー集約に用いる謝意を表現する語彙を特定し、集約結果に基づいたレシピ評価の方向性を示した。今後の展望として、レビュー中の食材やテクスチャとの関連性やテクスチャの共出現を分析して、レシピの評価やレビュー可視化への応用の可能性を探りたい。また本研究のレシピ調理結果の情報提示法が、レシピに対する要求を顕在化させることで「偶然の発見」を促す可能性についても検証したい。

文 献

- [1] ISO 11036. Sensory analysis -methodology- texture profile. Technical report, International Organization for Standardization, 1994.
- [2] Katsuyoshi Nishinari, Fumiyo Hayakawa, Chong- FeiXia, Long Huang, Jean- FrancoisMeulenet, Jean- MarcSieffermann. Comparative study of texture terms: English, french, japanese and chinese. *Journal of Texture Studies*, Vol. 39, No. 5, pp. 530–568, 2008.
- [3] クックパッド株式会社. クックパッド.
- [4] 楽天株式会社. 楽天レシピ.
- [5] 渡辺知恵美, 中村聡史. オノマトペロリ: 味覚や食感を表すオノマトペによる料理レシピのランキング. 人工知能学会論文誌, Vol. 30, No. 1, pp. 340–352, 2015.
- [6] 大橋正房, シズル研究会. 「おいしい」感覚と言葉: 食感の世代. B・M・FT 出版部, 2010.
- [7] B・M・FT. シズルワードの現在: 「おいしいを感じる言葉」調査報告. B・M・FT 出版部, 2015.
- [8] B・M・FT ことばラボ. ふわとろ: Sizzle word: 「おいしい」言葉の使い方. B・M・FT 出版部, 2016.
- [9] 森田真季, 荒牧英治, 灘本明代, 宮部真衣. マイクロブログにおける“おいしさ”情報の自動抽出: 嗜好とリアルタイム性を考慮した飲食店検索システムの提案. 電子情報通信学会技術研究報告 = IEICE technical report: 信学技報, Vol. 115, No. 486, pp. 63–68, mar 2016.
- [10] 森田真季, 荒牧英治, 灘本明代, 宮部真衣. マイクロブログにおける“おいしさ”情報分類器の適用可能性の検証. 電子情報通信学会技術研究報告 = IEICE technical report: 信学技報, Vol. 116, No. 488, pp. 229–234, mar 2017.
- [11] クックパッド株式会社. クックパッドデータ.
- [12] 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, pp. 337–346, 2005.
- [13] 早川文代. 日本語テクスチャー用語体系 (ver. 1.1).

- [1] ISO 11036. Sensory analysis -methodology- texture profile. Technical report, International Organization for Standardization, 1994.
- [2] Katsuyoshi Nishinari, Fumiyo Hayakawa, Chong- FeiXia, Long Huang, Jean- FrancoisMeulenet, Jean- MarcSieffermann. Comparative study of texture terms: English, french, japanese and chinese. *Journal of Texture Studies*, Vol. 39, No. 5, pp. 530–568, 2008.
- [3] クックパッド株式会社. クックパッド.
- [4] 楽天株式会社. 楽天レシピ.
- [5] 渡辺知恵美, 中村聡史. オノマトペロリ: 味覚や食感を表すオノマトペによる料理レシピのランキング. 人工知能学会論文誌,