

# ショートレシピ動画を用いたユーザの料理レベルに基づく 補助レシピ情報の提示

米澤 拓也<sup>†</sup> 王 元元<sup>††</sup> 河合由起子<sup>†††,††††</sup> 角谷 和俊<sup>†</sup>

<sup>†</sup> 関西学院大学総合政策学部 〒 669-1337 兵庫県三田市学園 2 丁目 1

<sup>††</sup> 山口大学大学院創成科学研究科 〒 775-8611 山口県宇部市常盤台 2-16-1

<sup>†††</sup> 京都産業大学情報理工学部 〒 603-8555 京都府京都市北区上賀茂本山

<sup>††††</sup> 大阪大学サイバーメディアセンター 〒 567-0047 大阪府茨木市美穂ヶ丘 5 番 1 号

E-mail: <sup>†</sup>{fdu49780,sumiya}@kwansei.ac.jp, <sup>††</sup>y.wang@yamaguchi-u.ac.jp, <sup>†††</sup>kawai@cc.kyoto-su.ac.jp

あらまし 近年, Kurashiru や DELISH KITCHEN などのショートレシピ動画が普及している. これらのレシピ動画は, 短い時間で料理のポイントとなる場所を学ぶことができるため, 多くの人々に利用されている. しかし, これらの動画は, ユーザが一度の閲覧のみで, 全ての調理動作を理解することは困難である. また, これらのショートレシピ動画は, ユーザの料理の腕 (料理レベル) が考慮されていないため, 誰が観ても同じ動画しか閲覧することができない. そこで, 本研究では, 既存のショートレシピ動画に含まれているテキストレシピから調理動作を抽出し, それらの出現頻度から料理のジャンルごとに重み付けを行うことで, ユーザの料理レベルに基づいた多様な補助レシピ情報を抽出する. また, これらの手法によって抽出された補助レシピ情報をビデオタグクラウドというユーザインタフェースによって提示し, ユーザに補足を行う新たなシステムを提案する.

キーワード ショートレシピ動画, テキストレシピ, 補助レシピ情報

## 1. はじめに

近年, クックパッド<sup>(注1)</sup>や楽天レシピ<sup>(注2)</sup>などのユーザ投稿型レシピサイトが普及している. これらのサイトは, 無料で大量のレシピが閲覧可能である他に, ささまざまなユーザが作成したオリジナルのレシピなどを閲覧することが可能なサイトである. また, 携帯端末などの普及により, テキストレシピだけではなく, 料理番組やショートレシピ動画などの映像を用いたレシピが普及しつつある. 料理番組では, プロの料理人による解説をオープンキャプションなどで表示し, コメントーターによるアドバイスなどを表示することで, 視聴者の調理に対する理解を容易にする工夫が見られる. これらの料理番組は, 主に調理工程の複雑な料理を取り扱うものが多いため, 家庭で手軽に作られるような料理を取り上げることは稀である.

一方, Kurashiru や DELISH KITCHEN などのショートレシピ動画では, 料理番組で取り上げられるような料理と比較すると, 調理が容易な料理が中心に取り上げられている. これらの動画は, 料理番組に付与されるような音声による解説は付与されておらず, 調理動作・使用する食材・解説などがオープンキャプションで表示された, 定点カメラ撮影による動画を 1 分程度に編集した映像となっている. そのため, ショートレシピ動画の視聴者は短時間で料理の要点を概観することが可能である. それに加えて, これらのレシピ動画は, ノーカットの映像が早送り再生されているため, コロケやハンバーグの具材

などの途中生成物の様子を確認することができ, 視聴者は直感的に料理のコツを掴むことが可能である. しかし, このような動画の視聴者は, 一度動画を閲覧しただけで全ての調理動作を理解することは困難である. また, これらのショートレシピ動画は, 料理が得意な視聴者と不得意な視聴者のどちらが閲覧しても, 表示される料理動画は同一のものである. また, 料理の経験者の中にも, 「炒めもの」が得意な視聴者や, 「煮込みもの」が得意な視聴者など, 視聴者の嗜好も考慮されていない.

我々は, 過去にショートレシピ動画に含まれる「食材」や「調理動作」の出現時間と出現頻度に基づいて分析を行った [1]. 動画内のコンテンツでは, 主にそれぞれの出現頻度や使用頻度から重み付けを行うことにより, ショートレシピ動画内で, 「より詳細に説明されるべき箇所」を判定する研究を行ってきた. また, ショートレシピ動画に含まれている, 「調理動作」の出現頻度や出現時間から, シーンごとの難易度を判定・抽出し, レシピ動画の難易度提供ショートレシピ動画に対する難易度抽出及び, 補助レシピ情報抽出のためのクエリ生成手法を提案した. 本研究では, 既存のショートレシピ動画に含まれるテキストレシピから調理動作を抽出し, それらの出現頻度から料理のジャンルごとに重み付けを行うことで, 視聴者の料理レベルに基づいた多様な補助レシピ情報を抽出する. また, これらの手法によって抽出された補助レシピ情報をビデオタグクラウドと呼ばれるユーザインタフェース上に提示し, 補足情報の推薦を行うシステムを提案する.

## 2. 関連研究

Wang ら [2] は, 視聴者が閲覧中の映像コンテンツに各視聴

(注1) : <https://cookpad.com/>

(注2) : <https://recipe.rakuten.co.jp/>



図1 ビデオタグクラウドの画面遷移

者の興味や関心を反映させることができていない問題を解決するために、視聴者がどのレベルで映像コンテンツが含んでいる興味・関心を反映させた映像をどの程度の詳細度で所望するかを、シーンの話題性を X 軸、編集した映像の長さを Y 軸とした 2 軸で分析した。その後、話題性と詳細度に基づく 4 種類の映像コンテンツの中から適切な映像を自動生成する手法の提案を行った。この研究は、視聴者の興味・関心に基づいて映像を生成しているが、本研究では、ショートレシピ動画に含まれる調理動作の重要度を分析することにより、視聴者の料理レベルに基づいた補助レシピ情報を抽出するという点で差異がある。

秋口ら [3] は、料理番組をレシピ動画と定義し、映像内の「作業種類の多さ」「調理動作の並列性」「編集意図 (カメラワーク)」の 3 点の要素を抽出することで、レシピ動画の時間的特性に基づいた難易度判定手法を提案した。料理動画の調理工程の複雑性に注目する点において本研究と類似するが、本研究では、料理番組ではなく、ショートレシピ動画をレシピ動画として用い、動画内に含まれる調理動作の分析に基づき補足情報を推薦するという点で異なる。牧野ら [4] は、学校教育において教材として採用されている家庭科の教科書に出現する調理動作を抽出し、客観的に調理動作の難易度を定義する研究を行った。また、定義した難易度をもとに料理レシピの難易度を算出する手法を提案している。この研究では、教科書に記載されている調理動作から難易度判定を行っているが、本研究では、ショートレシピ動画に出現する調理動作に重要度を付与し、補足情報を抽出するという点で異なる。

嗣ら [5] は、料理番組をレシピ動画として定義し、映像に含まれる調理動作のうち、繰り返し行われる動作を解析することで、調理動作の分類を行う手法を提案している。この研究は、料理動画中の調理動作の分類を行う研究であるが、本研究ではショートレシピ動画の調理動作から補足情報を抽出する点で異なる。大滝ら [6] は、料理レシピ動画に含まれる調理動作・器具・食材の 3 点に着目し、それぞれの出現回数や料理の移動回数などを考慮し、調理にかかるコストを算出する手法を提案している。本研究では、ショートレシピ動画を扱い、動画内に出現する調理動作の出現頻度から重要度を算出する。

祖父江ら [7] は、ユーザが撮影した動画から特徴ベクトルを抽出し、SNS に投稿される料理レシピ動画と形式の似た動画を作成する要約動画システムを提案した。この研究は、料理動画中の要素に注目し、特徴中抽出を行う点で本研究と類似するが、

本研究では、料理レシピ動画に対して補足動画を抽出し、ユーザインタフェース上に表示するシステムを提案するという点で異なる。牛久ら [8] は、数十分程度の調理動画を Faster R-CNN を用いて物体認識を行い、それらの認識結果からテンプレートモデルを用いて、テキストレシピを自動生成する手法を提案した。この研究は、ショートレシピ動画からテキストレシピを自動生成する研究である。本研究は、ショートレシピ動画から補助レシピ情報を抽出するという点で異なる。

### 3. システム概要

本研究では、ユーザの料理レベルに基づく多様な補助レシピ情報を提示するビデオタグクラウドという補足情報推薦システムの提案を行う。ビデオタグクラウドによる画面の遷移は、図 1 のように表示する。まず、ユーザの調理レベルを判定するために、料理の経験者か否かを判定する。次に、ユーザの得意分野を判定するため、得意な料理のジャンルを選択させる。そして、ユーザの料理レベルと得意分野の判定を行い、その結果に基づく多様な補助レシピ情報を提示するビデオタグクラウドを図 2 に示す。インタフェースの左端に、既存のショートレシピ動画を表示し、提案手法により抽出された補助レシピ動画をビデオタグクラウドとして表示する。このビデオタグクラウドでは、各調理動作のサムネイル画像が難易度に沿って大きさが変化しつつ、画面の中心に位置する調理動作から外に向かって難易度が下がるように表示される。たとえば、カレーのレシピに含まれる「みじん切り」が最も難しいと判定された場合、「みじん切り」が画面の中心に配置され、その他の調理動作は、難易度順に円を描くように配置され、サムネイルが縮小する。このインタフェース上の表示により、ユーザは調理動作の難しさを直感的に理解することができる。

- A: メイン画面 ユーザに料理の経験の有無や得意なジャンルを選択
- B: メニュー選択 メニューを選択
- C: ビデオタグクラウド 画面の左にレシピ動画、右にビデオタグクラウドによって補足する調理動作を表示
- D: 補足レシピ動画 選択された調理動作を表示

### 4. 提案手法

#### 4.1 調理動作の抽出

はじめに、ショートレシピ動画に含まれる調理動作の抽出を

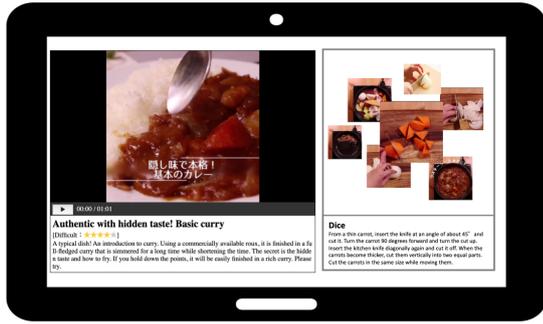


図2 ビデオタグクラウド

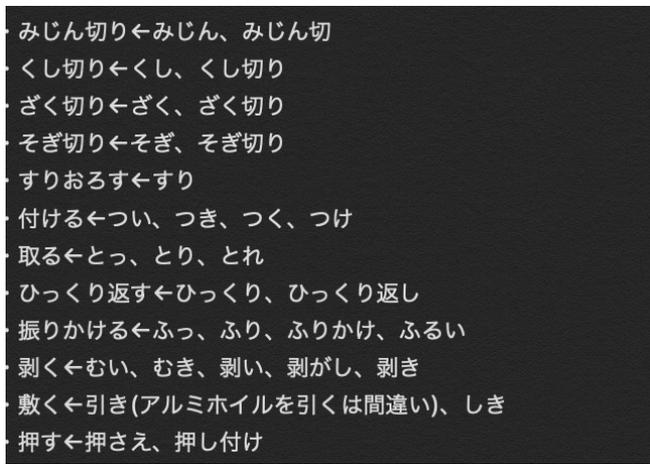


図3 調理動作の辞書作成

行う。Kurashiru から選出した 100 件のショートレシピ動画に含まれるテキストレシピから、MeCab を用いて形態素解析を行い、120 種類の調理動作を抽出した。

#### 4.2 調理動作のルール作成

また、今回形態素解析によって調理動作には多数の表記揺れが確認された。例をあげると、「みじん切り」という調理動作では、「みじん」や「みじん切」など複数の表現で抽出される。この抽出された表現を直接流用してしまうと、重要度判定を行う際に、出現回数の精度に差異が生じてしまうため、これを解決する必要がある。そこで、調理動作のデータを正規化するために、調理動作のルールを作成した。

こちらの図3は、調理動作を正規化の際に使用したルールの一部である。これらのルールに従って、正規化を行い、調理動作の重要度判定に使用する。

#### 4.3 ジャンルに基づく重み付け

ショートレシピ動画に対する補助レシピ情報の抽出・表示を行うために、補足する対象の調理動作を選定し、クエリとして

表1 実験1の結果

野菜のかき揚げ			ちくわの磯辺焼き		
順位	調理動作	重要度	順位	調理動作	重要度
1	入れる	0.267	1	くぐらせる	0.192
2	角切り	0.200	2	入れる	0.170
3	落とし	0.200	3	斜め切り	0.167
4	揚げる	0.087	4	下ろす	0.152
5	敷く	0.087	5	油をひく	0.134
6	そぎ切り	0.087	6	敷く	0.132
7	絡める	0.079	7	添える	0.121
8	水気を切る	0.070	8	盛り付ける	0.099
9	添える	0.070	9	混ぜる	0.096
10	まぶし	0.063	10	焼く	0.094
11	盛り付ける	0.057	11	熱する	0.094
12	油を切る	0.055			
13	切る	0.055			
14	混ぜる	0.052			
15	熱し	0.050			
	閾値	0.331		閾値	0.151

生成する必要がある。そのため、動画に含まれる調理動作の中から、補足情報が必要であると予想される動作を選出する。この抽出では、ユーザの料理レベルや得意分野を考慮するために、40 件のレシピをジャンルごとに分類した。本研究では、「煮込みもの」「炒めもの」「揚げもの」「○○焼き」「丼もの」の計5種類のジャンルに分類し、それぞれのジャンルごとに  $TF-iDF$  法による重み付けを行なった。重み付けの算出には以下の式(1)を用いる。ここで、 $I$  は調理動作の重要度を、 $k$  はジャンルごとに分類されたレシピ動画を、 $s$  は調理動作を、 $N$  はジャンルごとのレシピの集合を表す。

$$I(k, s) = TF(k, s) * \log \frac{N}{df(k) + 1} \quad (1)$$

#### 4.4 補足を行う調理動作の抽出

次に、重み付けの結果から閾値を設定し、補足すべき調理動作の判定を行う。設定した閾値よりも低い数値の調理動作は、「補足する必要がない調理動作」として除去する。一方、閾値よりも高い数値の調理動作は「補足する必要のある調理動作」と判定される。また、本研究における閾値には、算出した  $TF-iDF$  値を料理ごとに平均した値を閾値として使用した。この重み付けと補足対象の判定に基づき、上位に出現する調理動作に対応する補助レシピ動画を抽出するためのクエリとして扱い、このクエリを用いて Web 上から補助レシピ動画をビデオタグクラウドによって表示する。

### 5. 実験と結果

#### 5.1 調理動作の重要度判定の検証

本研究では、3 種類の実験を行った。実験1では、提案した式による  $TF-iDF$  値の算出を100件のレシピに対して行い、重要だと考えられる調理動作抽出の検証を行った。また、実験2では、それらの結果をジャンルごとに比較することで、ジャンルごとのレシピの特徴抽出の検証を行った。また、実験3では、10人のユーザに評価実験を行い、提案手法の有効性を検証した。

#### 5.2 考察

重み付けと閾値による「補足する必要のある調理動作」の判定を行った実験1の結果を表1に示す。実験1の結果、「盛り付ける」「熱する」などの補足する必要がないと考えられる調理

表 2 実験 2 の結果

ミニマヨコーンピザ			ミネストローネ		
順位	調理動作	重要度	順位	調理動作	重要度
1	薄切り	0.222	1	入れる	0.236
2	かける	0.189	2	炒める	0.207
3	敷く	0.178	3	角切り	0.177
4	散らす	0.178	4	調える	0.140
5	切る	0.150	5	輪切り	0.140
6	乗せる	0.136	6	油を引く	0.123
7	盛り付ける	0.132	7	熱し	0.112
8	混ぜる	0.128	8	薄切り	0.108
9	入れる	0.114	9	盛り付ける	0.104
	閾値	0.331	10	煮る	0.091
				閾値	0.151

表 3 実験 3 の結果

白菜と鶏肉の塩あかけ丼			鶏団子の味噌煮込み		
調理動作	重要度	評価値 (順位)	調理動作	重要度	評価値 (順位)
混ぜる	0.254	1(9位)	入れる	0.236	1(5位)
角切り	0.152	3(2位)	斜め切り	0.177	1(5位)
細切り	0.152	1.5(5位)	丸める	0.177	2.5(2位)
止める	0.142	4(1位)	おろす	0.154	1.5(4位)
閾値	0.132		閾値	0.144	

動作を除去することができた。それに対して、「角切り」「斜め切り」などの補足する必要があると考えられる調理動作を上位に抽出することができた。これらの結果から、簡単な調理動作ほどレシピの中での出現頻度が多く、難しい調理動作ほど出現頻度が少ないと考えられる。しかし、上位として選出された項目の中に「入れる」「油を引く」などの補足する必要がないと考えられる調理動作も抽出されたため、精度向上のために閾値の設定や前処理の手法を今後の課題とする。

また、実験 2 の結果を表 2、実験 3 の結果を表 3 に示した。それぞれの結果をレシピごとに比較することで、料理のジャンルにおける特徴の抽出を行った。比較した結果、ミニマヨコーンピザなどの焼きもの系とミネストローネなどの煮込みもの系を比較すると、揚げもの系では「薄切り」が上位に選出され、それに対して、煮込みもの系であるミネストローネでは下位に選出されていることが分かる。これらの差異から、揚げ物が得意なユーザと比べて、炒め物が得意なユーザには「薄切り」に関する情報を補足するべきであると考えられる。最後に、評価実験では、提案手法による重要度判定結果とユーザ評価値を比較した結果、角切りや丸めるなどは、一致するものの全体を見ると良い結果は得ることができなかった。ユーザ評価では、主に「丸める」「熱する」「揚げる」などの感覚的な表現の多い調理動作がユーザにとって補足の必要があることが判明した。

## 6. おわりに

本論文では、ショートレシピ動画のテキストレシピを用いて重要度を算出し、閾値を設定することで補足するべき調理動作を抽出する手法の提案を行った。また、抽出された情報から補助レシピ動画をユーザに提供するユーザインタフェースを提案した。

今後の課題は、抽出する調理動作の精度を上げるために詳細な閾値を設定することである。これにより、視聴者の調理レベルに対応できる幅を増やすことが可能になると考える。また、提案したユーザインタフェースの有効性を示すために評価実験を行う予定である。実験結果をもとに、より視聴者が調理動作

の理解を容易にさせるユーザインタフェースへの改良を目指す。

## 謝 辞

本研究の一部は、この研究は 2019 年度国立情報学研究所公募型共同研究 (19FC03) の助成を受けたものである。ここに記して謝意を表す。

## 文 献

- [1] Takuya Yonezawa, Yuanyuan Wang, Yukiko Kawai, Kazutoshi Sumiya: A Cooking Support System by Extracting Difficult Scenes for Cooking Operations from Recipe Short Videos, Proc. of the 27th ACM International Conference on Multimedia (ACM Multimedia 2019), pp. 2225-2227, October 2019.
- [2] Yuanyuan Wang, Yukiko Kawai, Kazutoshi Sumiya, Yoshiharu Ishikawa: An Automatic Video Reinforcing System based on Popularity Rating of Scenes and Level of Detail Controlling, Proc. of the 2015 IEEE International Symposium on Multimedia (ISM 2015), pp. 529-534, December 2015.
- [3] 秋口いくみ, 王元元, 河合由起子, 角谷和俊: 料理レシピ動画の時間権限特性抽出による難易度判定, 第 10 回データ工学と情報マネジメントに関するフォーラム, C4-4, 2018.
- [4] 牧野望, 塩井隆円, 楠和馬, 波多野賢治: 調理動作に基づく料理レシピ検索のための難易度算出法の提案, 第 9 回データ工学と情報マネジメントに関するフォーラム, C4-5, 2017.
- [5] 蒯承穎, 志土地由香, 高橋友和, 井手一郎, 村瀬洋: 料理動画における調理動作の解析, 第 4 回デジタルコンテンツシンポジウム講演予稿集, 2014.
- [6] 大滝健太郎, 鷹野孝典: 調理動作・器具・食材に注目した調理動画からの調理コスト判定手法の検討, 情報処理学会第 79 回全国大会, Vol. 2017.
- [7] 祖父江亮, 中澤満, 益子宗, 山下隆義, 藤吉弘巨: ソーシャルネットワークサービスに適した料理レシピ動画の生成, 第 26 回インタラクティブシステムとソフトウェアに関するワークショップ予稿集, 3-A17, 2018.
- [8] 牛久敦, 橋本敦史, 森信介: 調理動画からのレシピの自動生成, 言語処理学会第 23 回年次大会発表論文集, 2017.